Summer 2015

Small sample saddlepoint confidence intervals in epidemiology

Pasan Manuranga Edirisinghe

Department: Mathematics and Statistics

SMALL SAMPLE SADDLEPOINT CONFIDENCE INTERVALS IN EPIDEMIOLOGY

by

PASAN MANURANGA EDIRISINGHE

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MATHEMATICS

2015

Approved by:

Dr. Robert Paige, Advisor
Dr. V.A. Samaranayake
Dr. Xuerong Wen
Dr. Gayla Olbricht
Dr. Xiaoping Du

ABSTRACT

In section 1, we develop a novel method of confidence interval construction for directly standardized rates. These intervals involve saddlepoint approximations to the intractable distribution of a weighted sum of Poisson random variables and the determination of hypothetical Poisson mean values for each of the age groups. Simulation studies show that, in terms of coverage probability and length, the saddlepoint confidence interval (SP) outperforms four competing confidence intervals obtained from the moment matching (M8), gamma-based (G1,G4) and ABC bootstrap (ABC) methods.

In section 2, we first consider Brillinger's classical model for a vital rate estimate with a random denominator. We derive statistical properties for this rate estimate and investigate difficulties encountered while trying to perform statistical inference about its expected value. Since inference about this expected value is not possible, we consider instead confidence intervals for covariance of the bivariate Poisson distribution which underlies Brillinger's model, on the way to proposing a new model which is a modification of Brillinger's model and which has numerous theoretical and computational advantages over the latter. A simulation study for our new model shows that in terms of coverage probability, our novel two-dimensional mid-P "Clopper-Pearson" type confidence interval (CP2) outperforms the "Clopper-Pearson" type interval with no mid-P correction (CP0) and the "Clopper-Pearson" type interval with a classical one-dimensional mid-P correction (CP1). In addition, CP2 was found to be more or less equivalent, in terms of coverage probabilities, to CDF0, CDF1 and CDF2 which are the CDF pivot methods with no mid-P correction, a one-dimensional mid-P correction and a two-dimensional mid-P correction, respectively. Furthermore, method CP2 performed essentially as well as the saddlepoint approximation (SP0) to the CDF0 method. Finally, all of the above-mentioned methods (CDF0, CDF1, CDF2, CP0, CP1, CP2 and SP0) substantially outperform the large sample (LS) method of confidence interval construction, in terms of coverage probability.

ACKNOWLEDGMENTS

I would never have been able to finish my dissertation without the guidance of my advisor, committee members, help from friends, and support from my family and wife.

I would like to express my deepest gratitude to my advisor, Dr. Robert Paige, for his excellent guidance, care, patience, and for providing me with an excellent atmosphere for doing research. I would also like to thank my committee members Dr. V.A. Samaranayake, Dr. Xuerong Wen, Dr. Gayla Olbricht and Dr. Xiaoping Du for guiding my research for the past several years and helping me to develop my background in mathematics and statistics.

I would also like to thank my parents, elder brother and his family and all my friends. They were always supporting and encouraging me with their best wishes.

Finally, I would like to thank my wife, Madhuka Weerasinghe and son, Thinuga Daham Edirisinghe. They are always there to cheer me up and stand by me through the good times and bad.

DEDICATION

I would like to dedicate this Doctoral dissertation to my parents, Siripala Edirisinghe, Hemalatha Jayasinghe and my wife, Madhuka Weerasinghe. Without their continued support and counsel, I could not have completed this process.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. DIRECTLY STANDARDIZED RATES

1.1. INTRODUCTION

In this section we consider standardization which is a common strategy of controlling for confounding in the analysis of epidemiologic data. The age of a subject will act as an important confounder in the comparison of incidence or mortality rates between two or more groups when the age distributions of these groups differ significantly from one another; see Rothman *et al.* (2008) and Woodard (2004). Epidemiologists often use the Poisson distribution to model the occurrence of a specific disease. With this assumption the sample directly (age) standardized (incidence) rate (DSR) of a disease is a weighted sum of independent Poisson random variables. Here the weight for the ith age group, w_i, is determined as the ratio of the proportion of the standard population in this group to the number of person-years observed for the group. As a result, the associated population DSR is a weighted sum of Poisson means and this fact complicates the construction of good confidence intervals for this parameter Dobson *et al.* (1991), Fay and Feuer (1997).

The classical asymptotic method based upon the large-sample normal approximation for the Poisson distribution may not work well when the number disease occurrences is small Dobson *et al.* (1991), Fay and Feuer (1997). Ng *et al.* (2008) performed extensive simulation studies comparing twenty different DSR confidence intervals methods. The methods they considered fell into four basic classes: (i) the asymptotic normal confidence interval and three modifications; (ii) the moment matching method of Dobson et al. (1991) and eight variants thereof; (iii) the gamma-based confidence interval proposed by Fay and Feuer (1997) with three modifications and (iv) two methods based on beta distribution both of which were proposed in Tiwari *et al.* (2006).

Ng *et al.* (2008) recommend three of these twenty methods for general use. These methods are (in the original notation of Ng *et al.* (2008)): M8, a variant of the moment matching method of Dobson *et al.* (1991) which is an approximation to the mid-P

confidence interval proposed in Cohen and Yang (1994); G1, the gamma-based confidence interval proposed by Fay and Feuer (1997) and G4, a modification of this method proposed by Tiwari *et al.* (2006). Swift (2010) compared these three methods with the approximate bootstrap method (ABC) method proposed in an earlier work of Swift (1995) and found the ABC method to be competitive with the M8, G1 and G4 methods.

All of the aforementioned methods have coverage probabilities which tend to vary significantly from the nominal level, usually taken to be 95%; Ng *et al.* (2008) and Swift (2010). Trends in their coverage probabilities are roughly linear in $Var\,(w_i)$, the variance of the standardization weights. The M8 method is slightly liberal for small values of $Var\,(w_i)$ and becomes increasing more liberal as $Var\,(w_i)$ increases. The G1 method is conservative for small values of $Var\,(w_i)$ and becomes increasing more conservative for larger values of $Var\,(w_i)$. In contrast, the G4 method is conservative for small values of $Var\,(w_i)$ and becomes increasing less conservative as $Var\,(w_i)$ as increases. Finally, the ABC bootstrap method tends to be slightly conservative for all values of $Var\,(w_i)$.

In the section 1.2, we develop a saddlepoint-based method with average coverage probabilities that are close to the nominal 0.95 value for all values of $Var\,(w_i)$. In section 1.3, we describe the four competing methods from Ng *et al.* (2008) and Swift (2010). In section 1.4, we present an application of all methods to the Ausburg myocardial infarction data from Dobson *et al.* (1991) and summarize the results of Monte Carlo studies which compare our saddlepoint method with the M8, G1, G4 and ABC methods. We conclude with a discussion in section 1.5.

1.2. SADDLEPOINT CONFIDENCE INTERVAL

The population DSR is defined as

$$\mu = \sum\nolimits_{i=1}^{n} w_i \mu_i \tag{1.1}$$

where $\mu_i = n_i\theta_i$, n_i is the number of person-years observed for the ith age group, θ_i is the associated incidence rate, and w_i is the associated standardization weight for this group. Here w_i is given as

$$w_i = c_i \left(n_i \sum\nolimits_{j=1}^{n} c_j\right)^{-1}$$

where c_i is person-years in the standard population for the ith age group. We let Poisson random variable X_i, with mean μ_i, represent the number of disease occurrences in the ith group for $i = 1, \ldots, n$. Furthermore, we assume that the X_i's are independent of one another. The maximum likelihood estimator (MLE) for DSR μ is given as

$$\hat{\mu} = \sum\nolimits_{i=1}^{n} w_i X_i. \tag{1.2}$$

Note that while the distribution of $\hat{\mu}$ is intractable its cumulant generating function (CGF) is easily obtained in closed-form as

$$K_{\hat{\mu}}(s) = \sum\nolimits_{i=1}^{n} \mu_i \left[\exp(s w_i) - 1\right] \tag{1.3}$$

and from this one can easily verify that $E(\hat{\mu}) = \mu$ and $Var(\hat{\mu}) = \sum\nolimits_{i=1}^{n} w_i^2 \mu_i$. Furthermore, in what follows we shall let $\hat{\mu}_{obs}$ denote the observed value $\hat{\mu}$ from a random sample of the n age groups, let μ_0 denote the true value of the population DSR value, and let μ denote the assumed value of μ_0 which is used in the construction of the saddlepoint confidence interval.

The motivation for the saddlepoint-based method lies in the observation that probability integral transform $P(\hat{\mu} \leq \hat{\mu}_{\mathrm{obs}}; \mu = \mu_0)$ has a standard uniform distribution, therefore is a pivotal quantity and, as such, can be used to construct a confidence interval for μ_0; see Casella and Berger (2002) for further discussion of confidence intervals of this type. Here, one would determine $(\hat{\mu}_L, \hat{\mu}_U)$, a 95% confidence interval for μ_0, as the solution of the following equations:

$$P(\hat{\mu} \leq \hat{\mu}_{\mathrm{obs}}; \mu = \hat{\mu}_L) = 0.975 \quad \text{and} \quad P(\hat{\mu} \leq \hat{\mu}_{\mathrm{obs}}; \mu = \hat{\mu}_U) = 0.025. \tag{1.4}$$

This confidence interval is guaranteed to have exact coverage under the assumption that the family of cumulative distribution functions (CDFs) $\{P(\hat{\mu} \leq \hat{\mu}_{\text{obs}}; \mu)\}$ is stochastically decreasing in μ, the assumed μ_0 value.

The implementation of this method in practice is hindered by (i) the lack of a tractable expression for the CDF of $\hat{\mu}$ and (ii) uncertainty about how to relate an assumed value for μ to assumed values for the μ_i's; from equation (1.1) we see that there are an infinite number of ways of doing this.

The first issue is easily solved with a saddlepoint approximation to the CDF of $\hat{\mu}$. Saddlepoint approximations have been found to be remarkably accurate in approximating nonnormal distributions in a wide variety of situations; see Butler (2007). We use the Luganani and Rice (1980) saddlepoint approximation to the CDF of $\hat{\mu}$ which is given as

$$\hat{P}(\hat{\mu} \leq \hat{\mu}_{\text{obs}}; \mu) = \begin{cases} \Phi(\hat{t}) + \phi(\hat{t})\left[\hat{t}^{-1} - \hat{u}^{-1}\right], & \text{if } \hat{\mu}_{\text{obs}} \neq \mu \\ \frac{1}{2} + K_{\hat{\mu}}^{(3)}(0)\left[72\pi K_{\hat{\mu}}^{(2)}(0)^3\right]^{-1/2}, & \text{if } \hat{\mu}_{\text{obs}} = \mu \end{cases} \tag{1.5}$$

where $\Phi(\cdot)$ and $\phi(\cdot)$ are the standard normal CDF and normal (probability density function) PDF respectively, $K_{\hat{\mu}}^{(i)}(s)$ is the ith derivative of CGF $K_{\hat{\mu}}(s)$ for $i = 1, 2, 3$,

$$\hat{t} = sgn(\hat{s})\sqrt{2\left[\hat{s}\hat{\mu}_{\text{obs}} - K_{\hat{\mu}}(\hat{s})\right]}, \;\; \hat{u} = \hat{s}\sqrt{K_{\hat{\mu}}^{(2)}(\hat{s})}$$

and saddlepoint $\hat{s}$ is the solution to saddlepoint equation

$$K_{\hat{\mu}}^{(1)}(\hat{s}) = \hat{\mu}_{\text{obs}}.$$

With regards to the second issue of distributing a single assumed value for μ to the μ_i parameters we write μ as

$$\mu = \hat{\mu} + \Delta\mu$$

where $\hat{\mu}$ is the MLE for μ_0 and $\Delta\mu$ is the deviation of assumed value μ from $\hat{\mu}$. From equations (1.1, 1.2) we have that

$$\mu = \hat{\mu} + \Delta\mu = \sum\nolimits_{i=1}^{n} w_i \left(X_i + d_i \Delta\mu \right) = \sum\nolimits_{i=1}^{n} w_i \mu_i$$

where $d_i\Delta\mu$ represents the deviation of assumed value μ_i from MLE $\hat{\mu}_i = X_i$ for $i = 1, \ldots, n$. Clearly, $\sum\nolimits_{i=1}^{n} w_i d_i = 1$. Note however that the assumed value of μ_i cannot be negative and as such we take it to be

$$\mu_i = \max\left\{ X_i + d_i\Delta\mu, 0 \right\}. \tag{1.6}$$

Furthermore, setting $\mu_i = 0$ in CGF (1.3) is equivalent to assuming that X_i is a degenerate random variable with unit point mass at 0. Finally, we define

$$d_i = \left[\sum_{i:X_i \neq 0} w_i \right]^{-1} \tag{1.7}$$

and, as a result, assumed value (1.6) is a the conditional maximum likelihood estimate for that parameter where deviation $\Delta\mu$ is distributed equally among the non-zero X_i values.

Our proposed saddlepoint method is a CDF pivot method (1.4) where (i) the intractable CDF functions are replaced by their saddlepoint approximations given in (1.5) and (ii) method (1.6) is used to determine assumed values for the μ_i Poisson mean parameters from a single assumed value for population DSR μ. Pivotal CDF confidence intervals which make use of saddlepoint CDF approximations often yield lengths and coverage probabilities that compare favorably with those from basically any competing method; see for instance Paige and Trindade (2008) and Paige *et al.* (2009).

1.3. COMPETING METHODS

In this section we briefly describe the competing methods that we will compare to our saddlepoint confidence interval. They are the M8 method from Ng *et al.* (2008);

method G1, the gamma-based confidence interval proposed by Fay and Feuer (1997); method G4, a modification of Fay and Feuer's method considered in Tiwari *et al.* (2006) and the ABC bootstrap method proposed in Swift (2010).

1.3.1. M8 Confidence Interval. The upper and lower bounds of the $100\,(1-\alpha)\,\%$ M8 confidence interval method are given as

$$\hat{\mu}_{_L}^{M8} = \tilde{X}.\sqrt[3]{1 - \frac{1}{9}\tilde{X}.^{-1} - \frac{Z\,(1-\alpha/2)}{3}\tilde{X}.^{-1/2}}$$
$$\hat{\mu}_{_U}^{M8} = \tilde{X}.\sqrt[3]{1 - \frac{1}{9}\tilde{X}.^{-1} + \frac{z_{1-\alpha/2}}{3}\tilde{X}.^{-1/2}}$$

where $\tilde{X}. = \sum_{i=1}^{n} X_i + 1/2$ and $Z\,(1-\alpha/2)$ is the $(1-\alpha/2)$th quantile of the standard normal distribution.

1.3.2. G1 and G4 Confidence Intervals. The upper and lower bounds for the $100\,(1-\alpha)\,\%$ G1 interval are

$$\hat{\mu}_{_L}^{G1} = \frac{\widehat{Var}\,(\hat{\mu})}{2\hat{\mu}}\chi^2\,(\alpha/2)_{\left(2\hat{\mu}^2\left[\widehat{Var}(\hat{\mu})\right]^{-1}\right)}$$
$$\hat{\mu}_{_U}^{G1} = \frac{\widehat{Var}\,(\hat{\mu}) + (w_{\max})^2}{2\,(\hat{\mu} + w_{\max})}\chi^2\,(1-\alpha/2)_{\left(2(\hat{\mu}+w_{\max})^2\left[\widehat{Var}(\hat{\mu})+(w_{\max})^2\right]^{-1}\right)}$$

where $\hat{\mu} = \sum_{i=1}^{n} w_i X_i$, $\widehat{Var}\,(\hat{\mu}) = \sum_{i=1}^{n} w_i^2 X_i$, $w_{\max} = \max\{w_1, \ldots, w_n\}$ and $\chi^2\,(\gamma)_d$ is the γth quantile of the chi-square distribution with d degrees of freedom. Method G4 is a continuity-corrected version of method G1 developed in Tiwari *et al.* (2006). Here the upper and lower $100\,(1-\alpha)\,\%$ confidence bounds are given as

$$\hat{\mu}_{_L}^{G4} = \hat{\mu}_{_L}^{G1}$$
$$\hat{\mu}_{_U}^{G4} = \frac{\widehat{Var}^*\,(\hat{\mu})}{2\hat{\mu}^*}\chi^2\,(1-\alpha/2)_{\left(2(\hat{\mu}^*)^2\left[\widehat{Var}^*(\hat{\mu})\right]^{-1}\right)}$$

where $\hat{\mu}^* = \hat{\mu} + \frac{1}{n}\sum_{i=1}^{n} w_i$ and $\widehat{Var}^*\,(\hat{\mu}) = \widehat{Var}\,(\hat{\mu}) + \frac{1}{n}\sum_{i=1}^{n} w_i^2$.

1.3.3. ABC Confidence Interval. The $100\,(1-\alpha)\,\%$ ABC bootstrap confidence interval is given as

$$\hat{\mu}_L^{ABC} = \hat{\mu} + \frac{a - Z\,(1-\alpha/2)}{(\{1 - a[a - Z\,(1-\alpha/2)]\}^2}\sqrt{\widehat{Var}\,(\hat{\mu})}$$
$$\hat{\mu}_U^{ABC} = \hat{\mu} + \frac{a + Z\,(1-\alpha/2)}{(\{1 - a[a + Z\,(1-\alpha/2)]\}^2}\sqrt{\widehat{Var}\,(\hat{\mu})}$$

where bootstrap acceleration constant a is given as $\sum_{i=1}^{n} w_i^2 X_i \left[6\widehat{Var}\,(\hat{\mu})\right]^{-1}$. When the crude incidence rate is zero, meaning that $\sum_{i=1}^{n} X_i = 0$, this interval defaults to an interval of the form $\left(0, \hat{\mu}_U^P \sum_{i=1}^{n} w_i\right)$ where $\hat{\mu}_U^P$ the exact upper confidence limit for the mean of a Poisson random variable.

1.4. AN EXAMPLE AND MONTE CARLO STUDY

For our example we consider data from the WHO MONICA Project which is shown in Table 1.1. This data has previously been considered in Dobson *et al.* (1991) and Ng *et al.* (2008). Here incidence rates of myocardial infarction in 1986 for women age 35-64 are recorded from an urban reporting unit of the study area in Ausburg, Germany. Table 1.2 provides the 95% confidence intervals for the age-standardized incidence rate per 10,000 obtained from the five different methods we consider. Here $Var\,(w_i) = 1.843 \times 10^{-3}$ and we see that in terms of length, the shortest to longest intervals are M8, ABC, SP G4 and G1. These lengths are consistent with the liberalness of the M8 method and the conservativeness of the G4 and G1 methods which we observe in our simulation studies for small $Var\,(w_i)$ values.

For these Monte Carlo studies we adopt a design motivated by those used in Fay and Feuer (1997), Ng *et al.* (2008) and Swift (2010) and randomly generate the standardization weights and age group means. In our study, we worked with $n = 6$ age groups, generated independent standard uniform weights, $\{w_i\}$, and Poisson mean parameters $\{\mu_i\}$ and then standardized their values so that $\sum_{i=1}^{n} w_i = 6$ and $\mu. = \sum_{i=1}^{n} \mu_i = (10, 20)$. This process was repeated 500 times where $\mu. = 10$ and another 500 times with $\mu. = 20$. For each of these 1,000 weight-mean configurations we generated

10,000 Poisson counts according to the true μ_i values and then computed confidence intervals for population DSR μ using the saddlepoint (SP) method, the (M8) method from Ng *et al.* (2008), method (G1) in Fay and Feuer (1997), the (G4) method from Tiwari *et al.* (2006) and the (ABC) bootstrap method in Swift (2010). For each of these sets of 10,000 confidence intervals the coverage probability, for a particular method, was estimated as the proportion of times the interval contained μ_0, the true value of μ. We use the terms "underage" and "overage" to denote the probability the interval lies completely below μ_0 and completely above μ_0, respectively. Furthermore, we estimated underage by the proportion of times the upper confidence bound fell below μ_0 and overage by the proportion of times the lower confidence bound was above μ_0. Minitab 16 © was used to generate plots of the estimated coverage probabilities versus $Var\,(w_i)$ and the ratio of underage and overage versus $Var\,(w_i)$. Figure 1.1 shows the estimated coverage probabilities versus the $Var\,(w_i)$ values, as well as 95% reference lines and LOWESS smooths. Figure 1.2 shows the plot of the estimated underage/overage values versus $Var\,(w_i)$, along reference line of 1 which corresponds to a symmetric or equal-tailed confidence interval method.

Figure 1.1 shows that, for $\mu. = 10$ as well as $\mu. = 20$, the SP method exhibits a coverage probability trend which is quite close to the nominal 95% level, the ABC method is quite similar in this regard, albeit while being slightly conservative, the M8 method is liberal and the G1 and G4 methods are generally more conservative. Figure 1.2 shows, for both $\mu. = 10$ and $\mu. = 20$, that the ABC method is the most symmetric of all the methods, followed by the SP and G1 methods and then G4 and M8.

Table 1.3 provides the mean (CP Mean) , median (CP Median) and standard deviation (CP SD) of the coverage probabilities for each of the five methods as well as the average confidence interval lengths (CI Length). We see that the SP method always has an average coverage probability which closest to the nominal 95% value and exhibits the smallest variance. For $\mu. = 10$ and $\mu. = 20$, all differences in average coverage probabilities for all the methods are statistically significant and the SP and M8 methods always yield the shortest average lengths. For the $\mu. = 10$ simulations the lengths of the SP and M8 methods are not significantly different from one another, but are significantly less than the lengths from the G1, G4 and ABC methods. In contrast, for $\mu_. = 20$ the lengths of the SP and M8 methods are significantly different from one another. Furthermore, the means for the SP and ABC methods are not significantly different from one another, but are significantly different from the lengths for the other three methods.

1.5. CONCLUSIONS

We have developed a novel method of confidence interval construction for directly standardized rates using saddlepoint approximations. These intervals involve working with assumed values for μ, the population DSR. However this parameter is a weighted sum of Poisson means $\{\mu_i\}$ and as such we had to infer reasonable values for these means. We developed a maximum likelihood solution to this problem and the resulting SP procedure performed well in practice. In particular, simulation results showed that, in terms of coverage probability and length our intervals outperformed competing methods obtained from the moment matching, gamma-based and ABC bootstrap methods. However, in terms of symmetry the ABC bootstrap method outperformed our method and the G1 method slightly but the M8 and G4 methods by a wide margin.

Table 1.1. Incidence rates of myocardial infarction for women in Augsburg in 1986 by age group.

Age	c_i	n_i	w_i	X_i
$35-39$	$6/31$	$7,971$	0.243	0
$40-44$	$6/31$	$7,084$	0.273	0
$45-49$	$6/31$	$9,291$	0.208	1
$50-54$	$5/31$	$7,743$	0.208	2
$55-59$	$4/31$	$7,798$	0.165	4
$60-64$	$4/31$	$8,809$	0.146	10

Table 1.2. 95% confidence intervals for the age-standardized incidence rate of myocardial infarction per 10,000 people.

Method	LCB	UCB	Length
SP	1.685	4.428	2.744
M8	1.653	4.321	2.668
G1	1.593	4.608	3.015
G4	1.593	4.430	2.836
ABC	1.643	4.358	2.715

Table 1.3. The mean (CP Mean), median (CP Median) and standard deviation (CP SD) of the coverage probabilities as well as the average confidence interval lengths (CI Length) for SP, M8, G1, G4 and ABC methods.

$\mu_{\cdot} = 10$

Method	CP Mean	CP Median	CP SD	CI Length
SP	94.854	94.890	0.486	14.177
M8	93.702	93.935	1.122	14.068
G1	97.569	97.520	0.494	16.374
G4	96.044	96.100	0.649	15.055
ABC	95.255	95.555	1.134	14.885

$\mu_{\cdot} = 20$

Method	CP Mean	CP Median	CP SD	CI Length
SP	94.941	94.930	0.293	19.873
M8	94.420	94.490	0.499	19.808
G1	96.841	96.780	0.423	21.906
G4	95.819	95.840	0.379	20.790
ABC	95.248	95.280	0.383	20.360

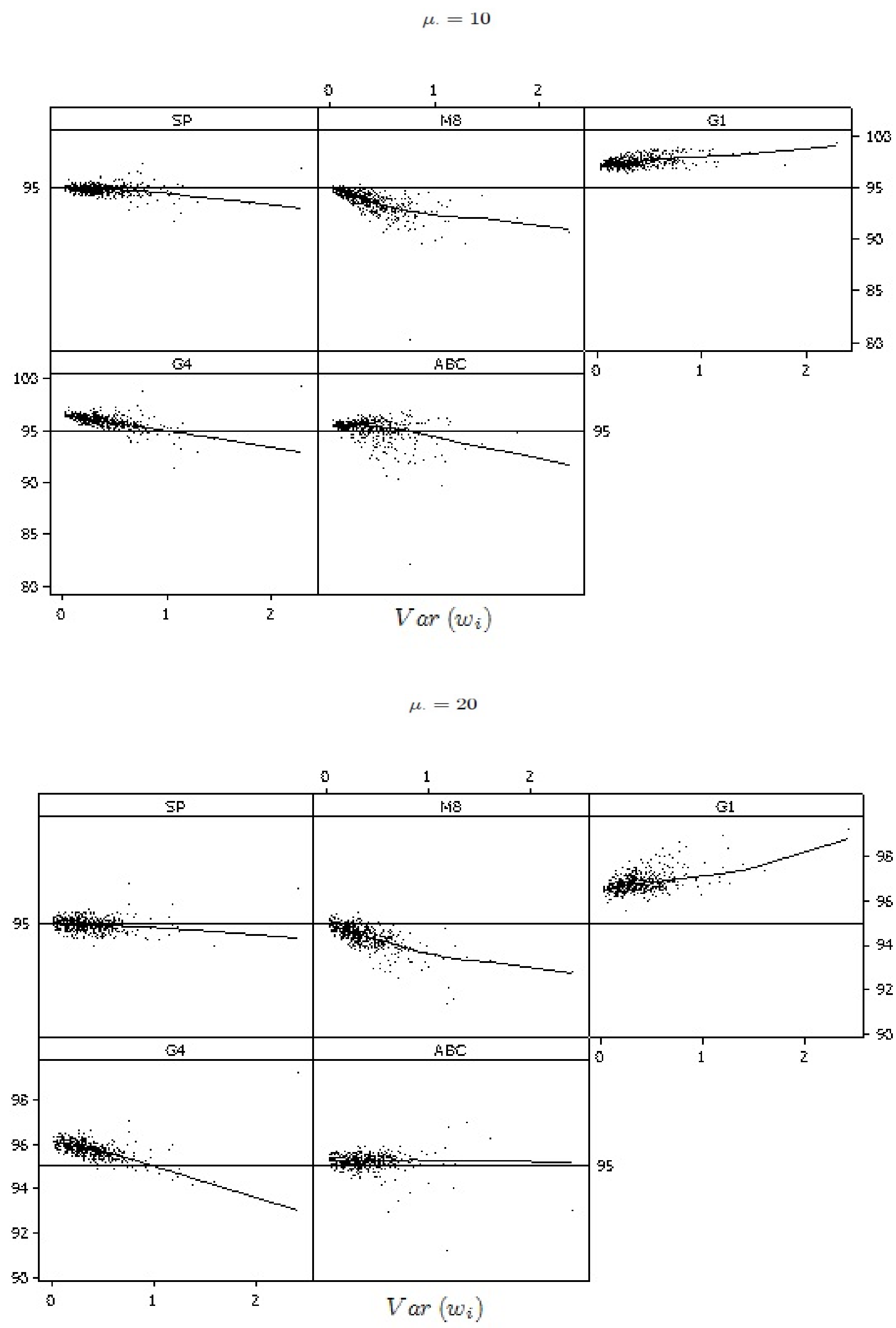
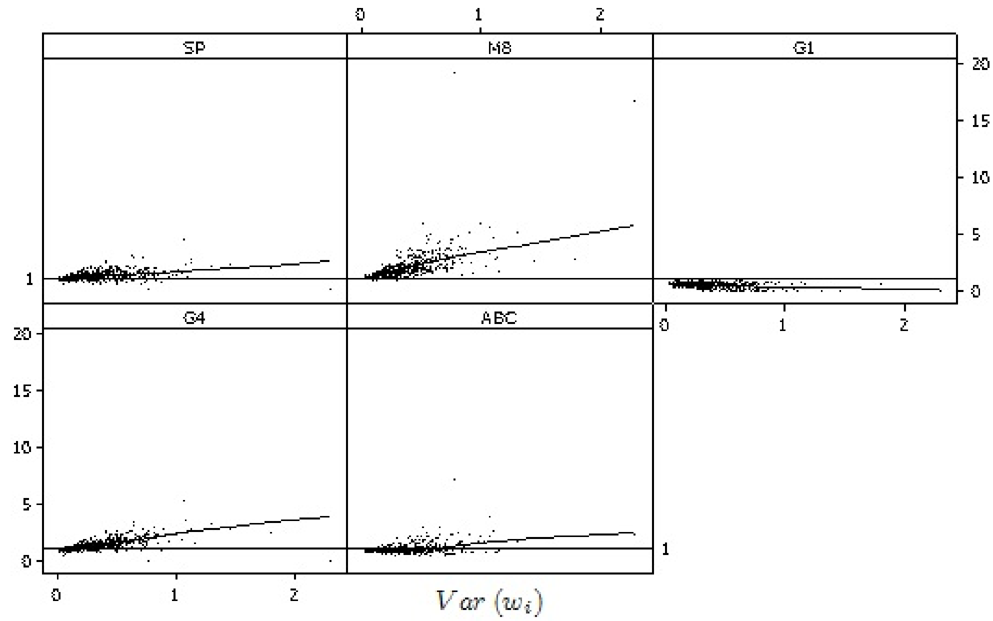


Figure 1.1. Plots of estimated coverage probabilities (in percents) for each of the five confidence interval methods (SP, M8, G1, G4 and ABC) versus $Var\,(w_i)$ with horizontal 95% reference lines and LOWESS smooths.

$\mu_{\cdot} = 10$



$\mu_{\cdot} = 20$

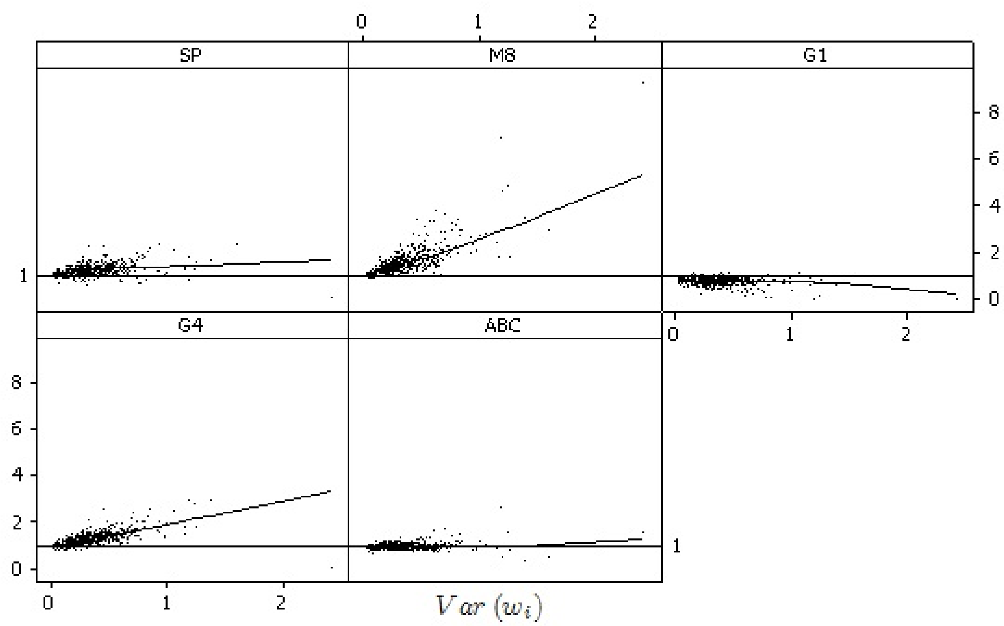


Figure 1.2. Plots of the estimated underage/overage values versus $Var\,(w_i)$ for each of the five confidence interval methods (SP, M8, G1, G4 and ABC) with horizontal reference lines at 1 and LOWESS smooths.

2. RATIO OF TWO POISSON RANDOM VARIABLES

2.1. INTRODUCTION

In this section, we consider confidence intervals for the ratio of two Poisson random variables. As such, we consider inference about a ratio with a random denominator. One can think about this denominator as being a random sample size. When the denominator is random standard distribution theory becomes more difficult to apply; see Molenberghs *et al.* (2014) and Bain *et al.* (1990). For example, the mean of a ratio of two Poisson random variables is not equal to the ratio of means, even when the Poisson random variables are independent. Our methodologies are motivated by Brillinger's work on vital rates (Brillinger, 1986) which involve the ratio of two random variables. Here it was shown why it is reasonable to have random denominator for vital rates. The primary reason is that the denominator is an approximation to the population size. Brillinger's approach has a natural motivation involving planar Poisson processes and a bivariate Poisson distribution on the number of deaths and mid-year population in connection with Lexis diagrams (see section 2.2). One problem, however, with Brillinger's approach is that mid-year population size can be zero which is a hindrance to the development of inferential techniques for his vital rates.

To remedy this, we consider restrictions of two types for Brillinger's model; (i) we constrain the mid-year population size to be strictly greater than zero and (ii) we require the number of deaths to be less than or equal to mid-year population size. With the latter restriction probability calculations become prohibitively complicated, so that we develop methodologies only for the former restriction. Even with this tractable restriction we find that Brillinger's model lacks interpretability and it is difficult to make inference about the model parameters. As a result, we propose a new model which is a limiting case of Brillinger's model and which has greater interpretability and better inferential properties. In particular, the expectation of vital rate has a very simple form and is interpretable unlike the expectation of vital rate in Brillinger's model wherein we require the mid-year population size to exceed zero.

The CDF pivot method of confidence interval construction for vital rates is based on probability integral transform where we use the cumulative distribution function (CDF) as a pivotal quantity, much like in section 1.2. In the proposed model we also consider confidence interval construction technique motivated by Clopper and Pearson (1934) to overcome difficulties encountered when the number of deaths and population size are equal. Also we consider confidence intervals based on the large sample theory for the our proposed model. Since Brillinger's model does not lend itself to large sample theory we consider maximum likelihood estimates and their exact small sample properties. For both models (Brillinger's and proposed) we consider highly accurate saddlepoint approximations and obtain confidence intervals based on the Luganani and Rice (1980) saddlepoint approximation to the CDF of a random variable. In the two types of models we estimate the nuisance parameters with constrained maximum likelihood estimates. For Brillinger's model we also introduce another constrained estimate based on the method of moments principle. In the CDF pivot and "Clopper-Pearson" type methods we introduce two mid-P correction methods which are novel in part because they are developed for a bivariate discrete distribution.

The rest of this section is organized as follows. In section 2.2, we discuss Lexis diagrams which are used to motivate and develop Brillinger's model as well as the proposed model. In section 2.3, we discuss probabilistic properties, methods of confidence interval construction and parameter estimation for Brillinger's model. Here we also prove the corollary from Brilliger (1986) which is not at all obvious but was stated without proof in that paper. In section 2.4, we discuss probabilistic properties, methods of confidence interval calculation and parameter estimation for the model we propose. Note that we do not consider as wide a range of inferential techniques for Brillinger's model as we consider for our proposed model since for the former we cannot make inference about the expected value of the vital rate estimate. In section 2.5, we discuss the saddlepoint approximations to the CDF pivot method considered for Brillinger's model and the proposed model. In section 2.6, we present simulation studies involving all confidence interval methods considered for both models. Finally, in section 2.7, we present our conclusions.

2.2. LEXIS DIAGRAMS

The Lexis diagrams were first introduced by Lexis (1875). Here the concept of representing the three demographic coordinates on one plane is considered, where (i) X_1 = the moment of death; (ii) X_2 = the age of the deceased at the moment of death and (iii) X_3 = the moment of birth of the deceased. Vandeschrick (2001) explains how the Lexis diagram represents a projection of a three-dimensional demographic point onto the two-dimensional X_2-X_3 plane. This projection results in a ray of unit slope representing the lifetime of a person. In effect this ray in the X_2-X_3 plane represents X_1. This is the basic setting in which Brillinger (1986) introduces a planar point Poisson process to model counts within the various planar regions of a Lexis diagram. A more detailed explanation and history of Lexis diagrams is provided by Vandeschrick (2001). In addition, the various uses for Lexis diagrams and statistical techniques for Lexis diagrams are reviewed in Keiding (1990).

2.2.1. Lexis Diagram for Brillinger's Model. We will first introduce the Lexis diagram for Brillinger's model. A reproduction of this diagram from Brillinger (1986) is shown in the Figure 2.1 below. We present this reproduction to improve legibility. This diagram describes how the mortality rate is defined in terms of the counts for regions B and C which are represented by $N(B)$ and $N(C)$, respectively. The associated mortality rate is

$$\frac{N(B)}{N(C)}.$$

Here, $N(B)$ represents the number of deaths in that region and $N(C)$ is the mid-year population for the corresponding year. Note that region C extends indefinitely in the direction of a 45^0 ray emanating from the origin. A rectangular region in a Lexis diagram represents deaths from a certain age group during a particular time period. In Figure 2.1, area B represents the number of people died in the age group 40 - 44 years during the 1980 calendar year. Parallelograms like C represent the total population of 40 - 44 years old who were alive in mid-1980. One can extend the sides of this parallelogram so that they intersect the horizontal axis and in the process determine a range of birthdays for those persons in the study. Therefore, the region C in Figure 2.1 represents people who were born after July 1st of 1935 and before July 1st of 1940. Recall that in Brillinger's model, denominator count $N(C)$ can be zero. This will happen if all study subjects die in the first half of the time period.

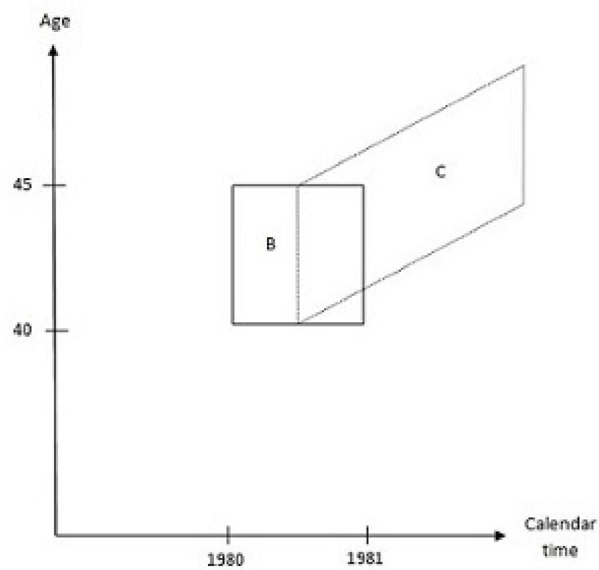


Figure 2.1. Lexis diagram for Brillinger's model.

2.2.2. Lexis Diagram for Proposed Model. Our proposed model is like that of Brillinger in that we define a mortality rate as

$$\frac{N(BC)}{N(C)}$$

where, much like before, $N(BC)$ and $N(C)$ represent the number of counts (death) in regions BC and C, respectively. Note however that region BC represents the deaths for a particular age-calendar year combination and has a different shape than Brillinger's region B.

Figure 2.2 shows the shapes for our BC and C regions, where region C contains region BC as a subset. Tetrahedron-shaped region BC is discussed in Keiding (1990) but is not used to define mortality rate models. Region BC can be seen to be a collection of right triangles as discussed in Keiding (1990). Its tetrahedron shape comes from the fact that vertical range of the age-calendar year combination is five units (years), from 40-44 years, but the horizontal range is one-half unit (year), from mid 1980 to beginning of 1981. In the example discussed below and illustrated in Figure 2.4, we can see that ten-by-five region BC becomes collection of three triangles since we consider age values from 50 to 59 years and calender times from 1968 to 1972, in the latter part of the time period. Furthermore, region BC represent the number of people who died in a given time period and age limit, and are born in a given time period. For example if we consider Brillinger's set-up given in Figure 2.2, region BC represents the number of people born from mid 1935 to mid 1939 and who died in the six month time period mid 1980 to beginning of 1981, and so were aged 40 - 44 years. Here, the population at risk is represented by region C like that in Figure 2.1. One important difference between the proposed model and Brillinger's model is that in the proposed model we do not count deaths which occur before the mid-year or include deaths of subjects who enter the age group after mid year. Therefore, deaths which occur before mid-year for people aged 40-44 years and deaths for people enter the age 40-44 group after mid-year are not counted as shown in Figure 2.2.

The "Epi" package in statistical software R provides code for generating a Lexis diagram for the Danish male lung cancer data (Carstensen *et al.* 2014). We reproduce diagrams given in Figures 2.3 and 2.4, of the "Epi" manual, to illustrate regions used in the development of models (Brillinger's model and proposed model). Note that in these graphs, numbers in each triangle represent number of rays terminating in that region since graphing a large number of rays will result in a graph which cannot be read.

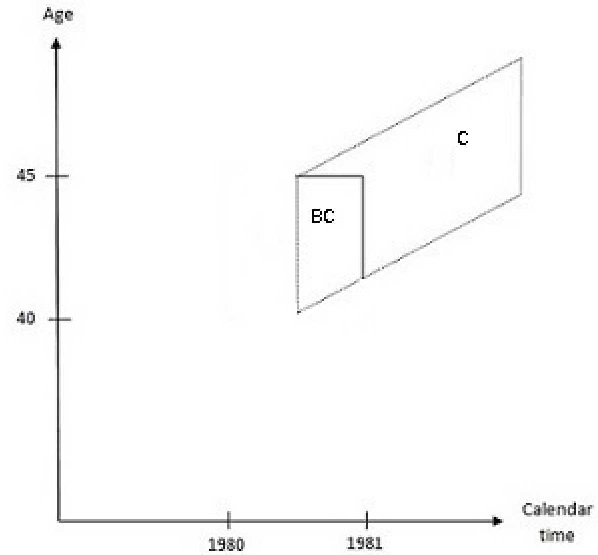


Figure 2.2. Lexis diagram for proposed model.

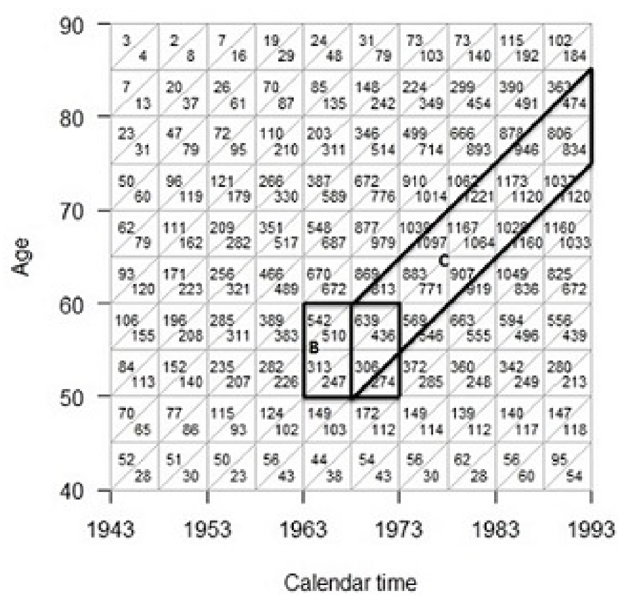


Figure 2.3. Lexis diagram for Danish male lung cancer data. This graph represents number of males died in years 1963 - 1972 due to lung cancer in Denmark who were in the age group 50 - 59 (Area B) and mid-population for that time period (Area C).

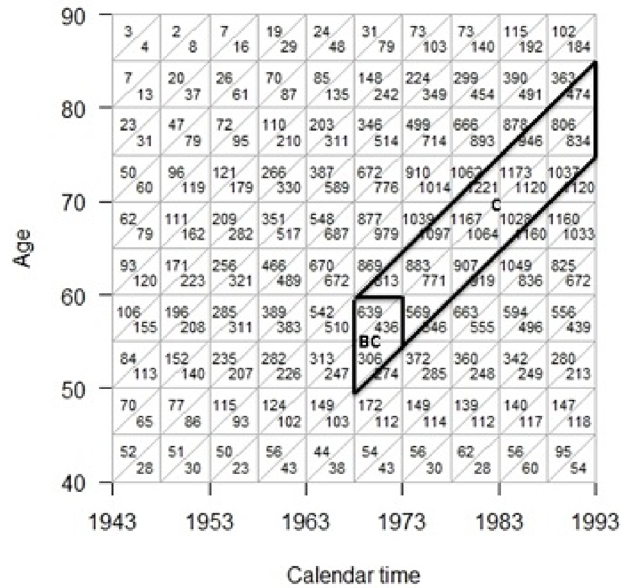


Figure 2.4. Lexis diagram for Danish male lung cancer data. This graph represents number of males died in years 1968 - 1972 before reaching the age 60 due to lung cancer in Denmark who were aged 50 - 59 years at the beginning of 1968 (Area BC) and mid-population for the time period 1963 - 1972(Area C).

2.3. BRILLINGER'S MODEL

Brillinger (1986) defines a model for vital rates in which the estimator has a random denominator. The denominator, which represents sample size, is often fixed in statistical applications. There are however situations in which it makes sense to treat the sample size as random, as described in the introduction. Brillinger (1986) takes the estimator of the population size to be the size of the mid-year sample population. In principle, the midyear population could consist of zero individuals and in such a setting Brillinger's vital rate estimator is undefined. We rectify this situation by restricting the vital rate estimator to non-zero sample populations sizes by conditioning upon a denominator which is non-zero.

The main idea for the Brillinger's model is explained in the singular corollary of that paper via a Lexis diagram much like that which was shown in Figure 2.1. Here a planar Poisson process with arbitrary intensity functions is defined in each of the three disjoint regions of this graph. This implicitly defines three independent Poisson random variables;

$$X_{B-BC},\ X_{C-BC} \text{ and } X_{BC},$$

where $BC \equiv B \cap C$ and $A - B$ denotes the set difference of sets A and B, in regions $[B - BC]$, $[C - BC]$ and $[BC]$, respectively, and with means λ_{B-BC}, λ_{C-BC} and λ_{BC}. From this development the corollary states, without proof, that $N(B)$ and $N(C)$ have a bivariate Poisson distribution, where $N(B)$ is given as $X_{B-BC}+X_{BC}$ and $N(C)$ is given as $X_{C-BC}+X_{BC}$. Note that the mean of X_{BC} represents the covariance of the bivariate Poisson distribution for $N(B)$ and $N(C)$ (see Johnson *et al.* 1997 and Kawamura 1984). Following Brillinger's notation we represent $N(B)$ as D and $N(C)$ as P. It turns out that D and P have a bivariate Poisson distribution, i.e.

$$(D, P) \sim \text{bivariate Poisson}(\lambda_{B-BC} + \lambda_{BC}, \lambda_{C-BC} + \lambda_{BC})$$

with associated probability mass function (PMF);

$$\begin{aligned}
&\Pr(D = d, P = p) \qquad (2.1)\\
&= \Pr(X_{B-BC} + X_{BC} = d, X_{C-BC} + X_{BC} = p)\\
&= \sum_{k=0}^{\min(d,p)} \Pr(X_{B-BC} + X_{BC} = d, X_{C-BC} + X_{BC} = p | X_{BC} = k) \Pr(X_{BC} = k)\\
&= \sum_{k=0}^{\min(d,p)} \Pr(X_{B-BC} = d - k, X_{C-BC} = p - k) \Pr(X_{BC} = k)\\
&= \sum_{k=0}^{\min(d,p)} \Pr(X_{B-BC} = d - k) \Pr(X_{C-BC} = p - k) \Pr(X_{BC} = k)\\
&= e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}} \sum_{k=0}^{\min(d,p)} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!\,(p-k)!k!}.
\end{aligned}$$

where $d = 0, 1, 2, ..., \infty$ and $p = 0, 1, 2, ..., \infty$.

Brillinger (1986) estimates the vital rate, for a given age group and a time period, as D/P where D and P distributed as bivariate Poisson. We consider the distribution of D/P and associated probability calculation

$$\Pr\left(D/P \leq (D/P)_{\text{obs}}\right), \tag{2.2}$$

where $(D/P)_{\text{obs}} = D_{\text{obs}}/P_{\text{obs}}$, D_{obs} is the observed value of D and P_{obs} is the observed value of P, in hopes of developing a confidence interval for $E\left(D/P\right)$. The idea behind this CDF pivot method of confidence interval construction involves the use of the CDF value (equation 2.2), which is in fact the probability integral transform, as a pivotal quantity and is described in greater detail in section 2.3.2.

Note however that this expected value is undefined since denominator P can be zero. Furthermore in practice a sample proportion based upon a sample of size zero is not of interest. To overcome these computational and practical difficulties when the denominator is zero, we consider the distribution of D/P given that $P > 0$. The joint distribution of random variable D and truncated random variable P is given in equation (2.3),

$$\Pr(D = d, P = p|P > 0) = \frac{\Pr(D = d, P = p)}{\Pr(P > 0)} \tag{2.3}$$
$$= \frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{1 - e^{-\lambda_{C-BC}-\lambda_{BC}}} \sum_{k=0}^{\min(d,p)} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}.$$

where $d = 0, 1, 2, ..., \infty$ and $p = 1, 2, ..., \infty$

We can easily see that this sums to one as follows,

$$\sum_{p=1}^{\infty}\sum_{d=0}^{\infty}\Pr(D = d, P = p|P > 0)$$
$$= \sum_{p=1}^{\infty}\sum_{d=0}^{\infty} \frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{1 - e^{-\lambda_{C-BC}-\lambda_{BC}}} \sum_{k=0}^{\min(d,p)} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}$$
$$= \frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{1 - e^{-\lambda_{C-BC}-\lambda_{BC}}} \sum_{p=1}^{\infty}\sum_{d=0}^{\infty}\sum_{k=0}^{\min(d,p)} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}$$

$$= \frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}}\left(\sum_{p=0}^{\infty}\sum_{d=0}^{\infty}\sum_{k=0}^{\min(d,p)}\frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}-\sum_{d=0}^{\infty}\frac{\lambda_{B-BC}^{d}}{d!}\right)$$

$$= \sum_{p=0}^{\infty}\sum_{d=0}^{\infty}\sum_{k=0}^{\min(d,p)}\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{\left(1-e^{-\lambda_{C-BC}-\lambda_{BC}}\right)(d-k)!(p-k)!k!}-\frac{e^{-\lambda_{C-BC}-\lambda_{BC}}}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}}$$

$$= \sum_{p=0}^{\infty}\sum_{d=0}^{\infty}\frac{\Pr\left(D=d,P=p\right)}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}}-\frac{e^{-\lambda_{C-BC}-\lambda_{BC}}}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}}$$

$$= \frac{1}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}}-\frac{e^{-\lambda_{C-BC}-\lambda_{BC}}}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}}$$

$$= 1$$

We would like to make inference about D/P given that $P > 0$ and are able to derive its CDF in closed-from as follows,

$$\Pr\left(D/P \leq (D/P)_{\text{obs}}\,|P>0\right) \tag{2.4}$$

$$= \Pr(D \leq (D/P)_{\text{obs}}\,P|P>0)$$

$$= \sum_{p=1}^{\infty}\sum_{d=0}^{\lfloor p(D/P)_{\text{obs}}\rfloor}\Pr(D=d|P=p,P>0)\Pr(P=p|P>0)$$

$$= \sum_{p=1}^{\infty}\sum_{d=0}^{\lfloor p(D/P)_{\text{obs}}\rfloor}\frac{\Pr(D=d,P=p|P>0)}{\Pr(P=p|P>0)}\Pr(P=p|P>0)$$

$$= \sum_{p=1}^{\infty}\sum_{d=0}^{\lfloor p(D/P)_{\text{obs}}\rfloor}\Pr(D=d,P=p|P>0)$$

$$= \frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}}\sum_{p=1}^{\infty}\sum_{d=0}^{\lfloor p(D/P)_{\text{obs}}\rfloor}\sum_{k=0}^{\min(d,p)}\frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}$$

where $\lfloor x \rfloor$ is the floor function which gives largest integer less than or equal to argument x. Ideally we would like to construct confidence intervals for parameter $E\left(D/P|P>0\right)$ but note however that

$$E\left(D/P|P>0\right) \tag{2.5}$$

$$= E_P E\left(D/P|P,P>0\right)$$

$$= E_P\left(\sum_{d=0}^{\infty}\frac{d}{p}\Pr(D=d|P=p,P>0)\right)$$

$$= \sum_{p=1}^{\infty}\sum_{d=0}^{\infty}\frac{d}{p}\Pr(D=d,P=p|P>0)$$
$$= \sum_{p=1}^{\infty}\sum_{d=0}^{\infty}\frac{d}{p}\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{(1-e^{-\lambda_{C-BC}-\lambda_{BC}})}\sum_{k=0}^{\min(d,p)}\frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}.$$

This doubly infinite sum cannot be simplified to closed form expression. Due to the complicated and non-interpretable nature of $E\left(D/P|P>0\right)$ we consider instead confidence intervals for λ_{BC} to illustrate the proposed methodologies in sections (2.3.2) and (2.5.1). Note that we have chosen λ_{BC} as our parameter of interest since it represents the covariance of the bivariate Poisson distribution for the unconditional distribution of D and P and has previously been consider by Johnson *et al.* (1997) and Kawamura (1984).Ultimately however, we consider λ_{BC} for illustrative purposes only and we could in fact just as easily consider confidence intervals for λ_{B-BC} and λ_{C-BC}, as well. Nonetheless, we forgo these confidence interval computations, in detail, since they are as complicated as those for covariance parameter λ_{BC}.

2.3.1. Maximum Likelihood Estimates for Brillinger's Model. First, let's consider maximum likelihood estimates for Brillinger's model. Teicher (1954) and Holgate (1964) considered inference for parameters of the bivariate Poisson distribution. Holgate (1964) also explains how to obtain maximum likelihood estimates for bivariate Poisson distribution for arbitrary n. Our focus is on a sample of size of $n=1$ and the results shown in Holgate (1964) simplify considerably in this case with key quantity

$$R(D,P)=\frac{\Pr(D-1,P-1)}{\Pr(D,P)}=\frac{\sum_{k=0}^{\min(d-1,p-1)}\frac{\lambda_{B-BC}^{d-1-k}\lambda_{C-BC}^{p-1-k}\lambda_{BC}^{k}}{(d-1-k)!(p-1-k)!k!}}{\sum_{k=0}^{\min(d,p)}\frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}}$$

where $\Pr(D,P)$ is given in equation (2.1). We note, however, that

$$R(D,P)=\frac{\sum_{k=0}^{\min(d,p)}\frac{k\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}}{\lambda_{BC}\sum_{k=0}^{\min(d,p)}\frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}}$$

where we can consider it to be an expected value in the following form,

$$\lambda_{BC}R(D,P) = \frac{\sum_{k=0}^{\min(d,p)} \frac{k\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}}{\sum_{k=0}^{\min(d,p)} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}} = E\left(T\right)$$

where T has PMF

$$\begin{aligned}
&\Pr\left(T=k\right) \\
&= \frac{\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}}{\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}} \sum_{k=0}^{\min(d,p)} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}} \\
&= \frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}}{\sum_{k=0}^{\min(d,p)} e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}} \frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}} \\
&= \frac{\Pr\left(X_{B-BC}=d-k, X_{C-BC}=p-k, X_{BC}=k\right)}{\sum_{k=0}^{\min(d,p)} \Pr\left(X_{B-BC}=d-k, X_{C-BC}=p-k, X_{BC}=k\right)} \\
&= \frac{\Pr\left(X_{B-BC}=d-k, X_{C-BC}=p-k, X_{BC}=k\right)}{\Pr\left(D=d, P=p\right)} \\
&= \frac{\Pr\left(D=d, P=p | X_{BC}=k\right)\Pr\left(X_{BC}=k\right)}{\Pr\left(D=d, P=p\right)}.
\end{aligned}$$

Using Bayes' theorem we can express $P\left(T=k\right)$ as follows,

$$\Pr\left(T=k\right) = \Pr\left(X_{BC}=k | D=d, P=p\right)$$

with expectation

$$\lambda_{BC}R(D,P) = E(T) = E_{X_{BC}|D=d,P=p}\left(X_{BC}\right).$$

First partial derivatives of the log-likelihood for Brillinger's model with no positivity condition on P from equation (2.1), with respect to λ_{B-BC} , λ_{C-BC} and λ_{BC} respectively, yields equations which can be expressed in terms of $E_{X_{BC}|D,P}\left(X_{BC}\right)$ as follows,

$$-1 + \frac{d}{\lambda_{B-BC}} - \frac{E_{X_{BC}|D=d,P=p}\left(X_{BC}\right)}{\lambda_{B-BC}} = 0 \tag{2.6}$$

$$-1 + \frac{p}{\lambda_{C-BC}} - \frac{E_{X_{BC}|D=d,P=p}\left(X_{BC}\right)}{\lambda_{C-BC}} = 0$$
$$-1 + \frac{E_{X_{BC}|D=d,P=p}\left(X_{BC}\right)}{\lambda_{BC}} = 0.$$

These equations simplify to

$$E_{X_{BC}|D=d,P=p}\left(X_{BC}\right) = \lambda_{BC}$$
$$\lambda_{B-BC} = d - \lambda_{BC}$$
$$\lambda_{C-BC} = p - \lambda_{BC}.$$

Here

$$E_{X_{BC}|D=d,P=p}\left(X_{BC}\right) = \lambda_{BC}$$

is similar to the expression

$$R(D,P) = 1$$

in Holgate (1964) when we consider a sample of size $n = 1$. Holgate (1964) suggests that to solve $E_{X_{BC}|D,P}\left(X_{BC}\right) = \lambda_{BC}$ for λ_{BC} one should profile out parameters λ_{B-BC} and λ_{C-BC} via a substitution of the other two equations to obtain the maximum likelihood estimate (MLE) of λ_{BC}. From this MLE one can obtain the MLE's for λ_{B-BC} and λ_{C-BC} as well. The equation $E_{X_{BC}|D,P}\left(X_{BC}\right) = \lambda_{BC}$ then simplifies to following univariate equation in λ_{BC},

$$\sum_{k=0}^{\min(d,p)} \frac{\left(d-\lambda_{BC}\right)^{d-k}\left(p-\lambda_{BC}\right)^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!}\left(\lambda_{BC}-k\right) = 0 \tag{2.7}$$

whose solution corresponds to the root of a polynomial of degree $(d+p)$, with a parameter space of $0 < \lambda_{BC} \leq \min(d,p)$. For arbitrary n, Holgate (1964) also obtains a univariate polynomial equation in λ_{BC} whose order depends upon the observed D and P values and notes that in general its root needs to be determined numerically. We note that for our $n = 1$ case the parameter space admits only one solution to equation (2.7) which is $\lambda_{BC} = \min(d, p)$. This in turn yields the following MLE's for λ_{B-BC} and λ_{C-BC};

$$\widehat{\lambda}_{B-BC} = d - \min(d, p)$$

and

$$\widehat{\lambda}_{C-BC} = p - \min(d, p).$$

Note that in this situation either $\widehat{\lambda}_{B-BC}$ or $\widehat{\lambda}_{C-BC}$ must be zero depending on the value of $\min(d, p)$. We derived the solution of the equation (2.7) while adopting the convention that $0^0 := 1$. This convention leads to equating a Poisson random variable with zero mean to the degenerate random variable with unit point mass at zero. We also see that convention of $0^0 := 1$ makes sense in the model we propose, in section (2.4), where we modify Brillinger's model by taking $\lambda_{B-BC} = 0$.

The exact distributions of MLE's $\widehat{\lambda}_{B-BC}$, $\widehat{\lambda}_{C-BC}$ and $\widehat{\lambda}_{BC}$ are as follows,

$$\begin{aligned}
\Pr\left(\widehat{\lambda}_{B-BC} \leq k\right) &= \Pr\left(D - \min(D, P) \leq k\right) \\
&= \Pr\left(\min(D, P) \geq D - k\right) \\
&= \Pr\left(D \geq D - k, P \geq D - k\right) \\
&= \Pr\left(P \geq D - k\right) \\
&= \Pr\left(X_{C-BC} + X_{BC} \geq X_{B-BC} + X_{BC} - k\right) \\
&= \Pr\left(X_{C-BC} \geq X_{B-BC} - k\right) \\
&= \sum_{i=0}^{\infty} \Pr\left(X_{C-BC} \geq X_{B-BC} - k | X_{B-BC} = i\right) \Pr\left(X_{B-BC} = i\right) \\
&= \sum_{i=0}^{\infty} \Pr\left(X_{C-BC} \geq i - k\right) \Pr\left(X_{B-BC} = i\right) \\
&= \sum_{i=0}^{\infty} \sum_{j=i-k}^{\infty} \Pr\left(X_{C-BC} = j\right) \Pr\left(X_{B-BC} = i\right)
\end{aligned}$$

$$= \sum_{i=0}^{\infty} \sum_{j=i-k}^{\infty} \frac{e^{-\lambda_{C-BC}} \lambda_{C-BC}^{j}}{j!} \frac{e^{-\lambda_{B-BC}} \lambda_{B-BC}^{i}}{i!}$$

and

$$\begin{aligned}
\Pr\left(\widehat{\lambda}_{C-BC} \leq k\right) &= \Pr\left(P - \min(D, P) \leq k\right) \\
&= \Pr\left(\min(D, P) \geq P - k\right) \\
&= \Pr\left(D \geq P - k, P \geq P - k\right) \\
&= \Pr\left(D \geq P - k\right) \\
&= \Pr\left(X_{B-BC} + X_{BC} \geq X_{C-BC} + X_{BC} - k\right) \\
&= \Pr\left(X_{B-BC} \geq X_{C-BC} - k\right) \\
&= \sum_{i=0}^{\infty} \sum_{j=i-k}^{\infty} \Pr\left(X_{B-BC} = j\right) \Pr\left(X_{C-BC} = i\right) \\
&= \sum_{i=0}^{\infty} \sum_{j=i-k}^{\infty} \frac{e^{-\lambda_{B-BC}} \lambda_{B-BC}^{j}}{j!} \frac{e^{-\lambda_{C-BC}} \lambda_{C-BC}^{i}}{i!}.
\end{aligned}$$

Note that above probability calculations for $\widehat{\lambda}_{B-BC}$ and $\widehat{\lambda}_{C-BC}$ involve infinite summations whereas the following probability calculation for $\widehat{\lambda}_{BC}$ involves only finite summations.

$$\Pr\left(\widehat{\lambda}_{BC} \leq k\right) = \Pr\left(\min(D, P) \leq k\right)$$

$$= 1 - \Pr\left(\min(D, P) > k\right) = 1 - \Pr\left(D > k, P > k\right)$$

$$= 1 - \Pr\left(X_{B-BC} + X_{BC} > k, X_{C-BC} + X_{BC} > k\right)$$

$$= 1 - \sum_{i=0}^{\infty} \Pr\left(X_{B-BC} + X_{BC} > k, X_{C-BC} + X_{BC} > k | X_{BC} = i\right) \Pr\left(X_{BC} = i\right)$$

$$= 1 - \sum_{i=0}^{\infty} \Pr\left(X_{B-BC} > k - i, X_{C-BC} > k - i\right) \Pr\left(X_{BC} = i\right)$$

$$= 1 - \sum_{i=0}^{\infty} \Pr\left(X_{B-BC} > k - i\right) \Pr\left(X_{C-BC} > k - i\right) \Pr\left(X_{BC} = i\right)$$

$$= 1 - \sum_{i=0}^{\infty} \left(\sum_{j=k-i+1}^{\infty} \Pr\left(X_{B-BC} = j\right) \sum_{j=k-i+1}^{\infty} \Pr\left(X_{C-BC} = j\right) \right) \Pr\left(X_{BC} = i\right)$$

$$= 1 - \sum_{i=0}^{\infty} \left(1 - F_{\lambda_{B-BC}}(k-i)\right)\left(1 - F_{\lambda_{C-BC}}(k-i)\right) \frac{e^{-\lambda_{BC}}\lambda_{BC}^i}{i!}$$
$$= \sum_{i=0}^{k} \left(F_{\lambda_{B-BC}}(k-i) + F_{\lambda_{C-BC}}(k-i) - F_{\lambda_{B-BC}}(k-i)F_{\lambda_{C-BC}}(k-i)\right) \frac{e^{-\lambda_{BC}}\lambda_{BC}^i}{i!}$$

where the above CDFs are

$$F_{\lambda_{B-BC}}(k-i) = \sum_{j=0}^{k-i} \frac{e^{-\lambda_{B-BC}}\lambda_{B-BC}^j}{j!} \text{ and } F_{\lambda_{C-BC}}(k-i) = \sum_{j=0}^{k-i} \frac{e^{-\lambda_{C-BC}}\lambda_{C-BC}^j}{j!}.$$

Note that one can obtain a confidence interval for covariance parameter λ_{BC} with the CDF pivot method. In that case, one needs to solve following equations after replacing λ_{B-BC} and λ_{C-BC} with constrained MLE's $\widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right)$ and $\widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right)$.

$$\Pr\left(\min(D,P) \le \min(d,p) | \lambda_{BCL}, \widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right), \widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right)\right) = 0.975$$
$$\Pr\left(\min(D,P) \le \min(d,p) | \lambda_{BCU}, \widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right), \widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right)\right) = 0.025.$$

This is the CDF pivot method which will be discussed later in the section 2.3.2.

In a similar fashion one can compute the first partial derivatives of the natural logarithm of equation (2.3) to obtain MLE's for Brillinger's model under the restriction that $P > 0$. The derivative with respect to λ_{B-BC} is exactly equal to the left side of the first equation in (2.6) and derivatives with respect to λ_{C-BC} and λ_{BC} yields the following score equations,

$$-1 + \frac{p}{\lambda_{C-BC}} - \frac{E_{X_{BC}|D=d,P=p}\left(X_{BC}\right)}{\lambda_{C-BC}} - \frac{e^{-\lambda_{C-BC}-\lambda_{BC}}}{1 - e^{-\lambda_{C-BC}-\lambda_{BC}}} = 0 \tag{2.8}$$
$$-1 + \frac{E_{X_{BC}|D=d,P=p}\left(X_{BC}\right)}{\lambda_{BC}} - \frac{e^{-\lambda_{C-BC}-\lambda_{BC}}}{1 - e^{-\lambda_{C-BC}-\lambda_{BC}}} = 0. \tag{2.9}$$

These equations will not simplify further in closed-form like those for the unconditional distribution. But equation (2.6) can be written as

$$\lambda_{B-BC} = d - \frac{p\lambda_{BC}}{\lambda_{C-BC} + \lambda_{BC}} \tag{2.10}$$

equations (2.8) and (2.9) simplify to

$$
\begin{aligned}
p\left(1-e^{-\lambda_{C-BC}-\lambda_{BC}}\right)-\left(\lambda_{C-BC}+\lambda_{BC}\right)&=0\\
E_{X_{BC}|D=d,P=p}\left(X_{BC}\right)-\frac{p\lambda_{BC}}{\lambda_{C-BC}+\lambda_{BC}}&=0.
\end{aligned}
$$

By substituting λ_{B-BC} in to $E_{X_{BC}|D=d,P=p}\left(X_{BC}\right)$ from equation (2.10) one can solve above two equations numerically to obtain MLE's of λ_{C-BC} and λ_{BC}. Note however that given the fact that these equations cannot be solved in closed-form and the above-mentioned problems with the MLE's for Brillinger's unrestricted model we do not pursue inference based upon the MLE's but instead consider a CDF pivot method of confidence interval construction next.

2.3.2. CDF Pivot Method. We construct a confidence interval for covariance of bivariate Poisson distribution (λ_{BC}) using the probability integral transform, which corresponds to the CDF given in equation (2.4) as a function of λ_{BC}, provided that one knew the true values for λ_{B-BC}, λ_{C-BC}. In setting $n=1$ case, one can obtain regular MLE's for λ_{B-BC}, λ_{C-BC} and λ_{BC}, as discussed in the previous section. For a sample of size $n>1$, numerical determination of these MLE's is discussed in Kawamura (1984).

Since we always consider the $n=1$ case, we do not have access to large sample Wald confidence intervals like those considered in section 2.4. Nonetheless, one can determine constrained estimates for two of the parameters given a third. Once we have some form of constrained estimates, $\widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right)$ and $\widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right)$, for λ_{B-BC} and λ_{C-BC} respectively, we may obtain a confidence interval $\left(\hat{\lambda}_{BCL},\, \hat{\lambda}_{BCU}\right)$ for λ_{BC} by solving following equations:for $(D/P)_{\text{obs}}<1$

$$
\begin{aligned}
\Pr\left(D/P \le (D/P)_{\text{obs}}\,|\hat{\lambda}_{BCL}, \widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right), \widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right), P>0\right) &= 0.975\\
\Pr\left(D/P \le (D/P)_{\text{obs}}\,|\hat{\lambda}_{BCU}, \widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right), \widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right), P>0\right) &= 0.025
\end{aligned}
\tag{2.11}
$$

for $(D/P)_{\text{obs}} > 1$

$$\begin{aligned}\Pr\left(D/P \le (D/P)_{\text{obs}} \,|\hat{\lambda}_{BCL}, \widehat{\lambda}_{B-BC}(\lambda_{BC}), \widehat{\lambda}_{C-BC}(\lambda_{BC}), P > 0\right) &= 0.025 \\ \Pr\left(D/P \le (D/P)_{\text{obs}} \,|\hat{\lambda}_{BCU}, \widehat{\lambda}_{B-BC}(\lambda_{BC}), \widehat{\lambda}_{C-BC}(\lambda_{BC}), P > 0\right) &= 0.975.\end{aligned} \tag{2.12}$$

These two different sets of equations result from the fact that the above CDF for D/P is a stochastically decreasing function in λ_{BC} when $(D/P)_{\text{obs}} < 1$ and is stochastically increasing when $(D/P)_{\text{obs}} > 1$.

Note that one can obtain confidence intervals for λ_{B-BC} and λ_{C-BC} by replacing constrained estimates $\left(\widehat{\lambda}_{B-BC}(\lambda_{BC}), \widehat{\lambda}_{C-BC}(\lambda_{BC})\right)$ in equations (2.11, 2.12) by

$$\left(\widehat{\lambda}_{C-BC}(\lambda_{B-BC}), \widehat{\lambda}_{BC}(\lambda_{B-BC})\right) \text{ and } \left(\widehat{\lambda}_{B-BC}(\lambda_{C-BC}), \widehat{\lambda}_{BC}(\lambda_{C-BC})\right)$$

respectively. These constrained estimates are described next in section (2.3.3).

2.3.3. Constrained Estimates.

2.3.3.1. Constrained MLE's. Lee and Young (2005), DiCiccio *et al.* (2001) and Diciccio and Romano (1995) discuss the optimal properties associated with the substitution of unknown nuisance parameters with constrained maximum likelihood estimates (constrained MLE's). They show that a lower level of error is achieved by replacing nuisance parameters with these constrained MLE's instead of regular MLE's. As such, we replace λ_{B-BC} and λ_{C-BC} with their constrained MLE's. We used multivariate form of Newton's method to solve the equations given in (2.6) and (2.8) for a fixed value of λ_{BC}.

Note that, to consider inference about λ_{B-BC} one can compute the first partial derivatives of joint probability density function given in equation (2.3) with respect to λ_{C-BC} and λ_{BC} which is given in (2.8) and (2.9). In a similar fashion the equations for inference about λ_{C-BC} are given in (2.6) and (2.9).

We also considered an alternative way to obtain constrained estimates and in particular developed method of moments constrained estimates, which will be discussed in the next section.

2.3.3.2. Method of moments constrained estimates. Constrained parameter estimates are obtained via the method of moments principle by solving equations (2.13). Which are based on the univariate distributions of D and P and the sample moments.

$$E(D|P>0) = d \text{ and } E(P|P>0) = p. \tag{2.13}$$

These equations simplify to

$$\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}}}{1-e^{-\lambda_{C-BC}-\lambda_{BC}}}\sum_{d=0}^{\infty}\sum_{p=1}^{\infty}\sum_{k=0}^{\min(d,p)} d\frac{\lambda_{B-BC}^{d-k}\lambda_{C-BC}^{p-k}\lambda_{BC}^{k}}{(d-k)!(p-k)!k!} = d, \tag{2.14}$$

$$\left(\lambda_{C-BC}+\lambda_{BC}\right)/(1-\exp(-\lambda_{C-BC}-\lambda_{BC})) = p. \tag{2.15}$$

Here, equation (2.14) can be simplified to

$$\lambda_{B-BC} + \frac{\lambda_{BC}}{1-\exp(-\lambda_{C-BC}-\lambda_{BC})} = d. \tag{2.16}$$

The equations given in (2.15, 2.16) can be solved separately. First, by solving equation (2.15) we can obtain $\widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right)$, then by substituting $\widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right)$ in equation (2.16) can obtain $\widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right)$ for given values of λ_{BC}. Note that the derivative of the equation (2.16) with respect to λ_{B-BC} is 1 for all the values of λ_{B-BC} and has an intercept of

$$\lambda_{BC}/(1-\exp(-\lambda_{C-BC}-\lambda_{BC})).$$

This yields $\widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right)$ in closed-form as

$$\widehat{\lambda}_{B-BC}\left(\lambda_{BC}\right) = d - \lambda_{BC}/(1-\exp(-\widehat{\lambda}_{C-BC}\left(\lambda_{BC}\right)-\lambda_{BC})).$$

We can use equations $(2.15, 2.16)$ to make inference about λ_{B-BC} and λ_{C-BC} using method of moment constrained estimates. Here we need to treat λ_{B-BC} as a constant and solve for λ_{C-BC} and λ_{BC} to obtain confidence interval for λ_{B-BC}. Similarly we can obtain confidence interval for λ_{C-BC} by solving for λ_{B-BC} and λ_{BC} by treating λ_{B-BC} as a constant.

With the method of moments estimates the lower bound of the CDF pivot confidence interval cannot be obtained. Numerical investigations indicate that the CDF given in equation (2.4) approaches a number close to 0.5 as $\lambda_{BC} \to 0$. Here, equation (2.4) reduces to following equation as $\lambda_{BC} \to 0$,

$$\Pr\left(D/P \leq (D/P)_{\text{obs}} \,|P > 0\right) = \frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}}}{1 - e^{-\lambda_{C-BC}}} \sum_{p=1}^{\infty} \sum_{d=0}^{\lfloor p(D/P)_{\text{obs}} \rfloor} \frac{\lambda_{B-BC}^{d} \lambda_{C-BC}^{p}}{d!p!}.$$

Note that the terms in the summation of variable k in equation (2.4) will be defined only for $k = 0$ and for all the other values of k those terms are zero. This probability fluctuates around 0.5 and it approaches the target percentile values of 0.025 and 0.975 as λ_{B-BC} goes to zero. But in this setting λ_{B-BC} will not be close to zero as $\lambda_{BC} \to 0$. Therefore we cannot obtain lower bound of the confidence interval using method of moment constrained estimates, and as a result we default to finding a 95% upper bound. Also, the regular MLE of λ_{B-BC} is zero when $(D/P)_{\text{obs}} < 1$ and the MLE of λ_{C-BC} is zero when $(D/P)_{\text{obs}} > 1$. Therefore, if one were to replace the nuisance parameters with regular MLE's they would not have a problem finding the lower bound. Numerical investigations of this procedure indicate that the resulting confidence intervals typically have very low coverage close to 80%.

2.3.4. Proof of Brillinger's Corollary. Corollary: Under the conditions of the theorem 1 in Brillinger (1986), for any regions in the Lexis diagram: (a) $\{D, P\}$ is distributed as $\{X_{B-BC} + X_{BC}, X_{C-BC} + X_{BC}\}$, where X_{B-BC}, X_{C-BC} and X_{BC} are independent Poissons with means λ_{B-BC}, λ_{C-BC} and λ_{BC}; (b) D/P is distributed as $(X_{B-BC} + X_{BC})/(X_{C-BC} + X_{BC})$. Further, (c) D given P is distributed as $U + S$ where U is Poisson with mean λ_{B-BC} and S is independently binomial with $n = P$ and proportion $\lambda_{BC}/\left(\lambda_{C-BC} + \lambda_{BC}\right)$.

Proof: The three regions $[B-BC]$, $[C-BC]$, $[BC]$ in Lexis diagram from Figure 2.1 are disjoint. Therefore counts in each region X_{B-BC}, X_{C-BC} and X_{BC} is distributed as Poisson with means respectively λ_{B-BC}, λ_{C-BC} and λ_{BC}. Then it follows

$$D = X_{B-BC} + X_{BC} \sim \text{Poisson}\left(\lambda_{B-BC} + \lambda_{BC}\right)$$

and

$$P = X_{C-BC} + X_{BC} \sim \text{Poisson}\left(\lambda_{C-BC} + \lambda_{BC}\right).$$

Therefore D and P are two correlated Poisson random variables and $\{D, P\}$ has a bivariate Poisson distribution determined from the relation

$$\{D, P\} = \{X_{B-BC} + X_{BC}, X_{C-BC} + X_{BC}\}.$$

Now consider the distribution of D given P.

$$\begin{aligned}\Pr\left(D = a|P = b\right) &= \Pr\left(X_{B-BC} + X_{BC} = a|X_{C-BC} + X_{BC} = b\right) \\ &= \frac{\Pr(X_{B-BC} + X_{BC} = a, X_{C-BC} + X_{BC} = b)}{P(X_{C-BC} + X_{BC} = b)}.\end{aligned}$$

The numerator of the above expression can be expressed as

$$\begin{aligned}&\Pr\left(X_{B-BC} + X_{BC} = a, X_{C-BC} + X_{BC} = b\right) \\ &= \sum_{k=0}^{\infty} \Pr(X_{B-BC} + X_{BC} = a, X_{C-BC} + X_{BC} = b|X_{BC} = k) \\ &\times \Pr\left(X_{BC} = k\right) \\ &= \sum_{k=0}^{\infty} \Pr(X_{B-BC} = a - k, X_{C-BC} = b - k) \Pr(X_{BC} = k) \\ &= \sum_{k=0}^{\min(a,b)} \Pr(X_{B-BC} = a - k) \Pr(X_{C-BC} = b - k) \Pr(X_{BC} = k) \\ &= \sum_{k=0}^{\min(a,b)} \Pr(X_{B-BC} = a - k) \Pr(X_{C-BC} = b - k) \Pr(X_{BC} = k).\end{aligned}$$

This yields

$$\begin{aligned}
&\Pr(X_{B-BC} + X_{BC} = a | X_{C-BC} + X_{BC} = b) \\
&= \sum_{k=0}^{\min(a,b)} \frac{\Pr(X_{B-BC} = a-k)\Pr(X_{C-BC} = b-k)\Pr(X_{BC} = k)}{\Pr(X_{C-BC} + X_{BC} = b)}.
\end{aligned}$$

Consider the summand of the finite summation above

$$\begin{aligned}
&\frac{\Pr(X_{B-BC} = a-k)\Pr(X_{C-BC} = b-k)\Pr(X_{BC} = k)}{\Pr(X_{C-BC} + X_{BC} = b)} \\
&= \frac{\frac{\lambda_{B-BC}^{a-k} e^{-\lambda_{B-BC}} \lambda_{C-BC}^{b-k} e^{-\lambda_{C-BC}} \lambda_{BC}^{k} e^{-\lambda_{BC}}}{(a-k)!(b-k)!k!}}{\frac{(\Lambda(C-BC)+\Lambda(BC))^b e^{-\Lambda(C-BC)-\Lambda(BC)}}{b!}} \\
&= \frac{b!\lambda_{B-BC}^{a-k} e^{-\lambda_{B-BC}} \lambda_{C-BC}^{b-k} \lambda_{BC}^{k}}{(b-k)!k!(a-k)!(\lambda_{C-BC}+\lambda_{BC})^b}.
\end{aligned}$$

These terms can be rearranged in to a product of following terms,

$$\binom{b}{k}, \left(\frac{\lambda_{C-BC}}{\lambda_{C-BC}+\lambda_{BC}}\right)^{b-k}, \left(\frac{\lambda_{BC}}{\lambda_{C-BC}+\lambda_{BC}}\right)^{k} \text{ and } \frac{\lambda_{B-BC}^{a-k} e^{-\lambda_{B-BC}}}{(a-k)!}.$$

From this we obtain

$$\begin{aligned}
&\Pr\left(X_{B-BC} + X_{BC} = a | X_{C-BC} + X_{BC} = b\right) \\
&= \sum_{k=0}^{\min(a,b)} \binom{b}{k} \left(\frac{\lambda_{C-BC}}{\lambda_{C-BC}+\lambda_{BC}}\right)^{b-k} \left(\frac{\lambda_{BC}}{\lambda_{C-BC}+\lambda_{BC}}\right)^{k} \\
&\times \frac{\lambda_{B-BC}^{a-k} e^{-\lambda_{B-BC}}}{(a-k)!}.
\end{aligned}$$

This is equivalent to the convolution of two Poisson and binomial random variables. Consider for instance

$$X \sim \text{Poisson}(\lambda) \quad \text{and} \quad Y \sim \text{Bin}(n, \theta)$$

then we have

$$\begin{aligned}
\Pr(X+Y=i) &= \sum_{q=0}^{\min(i,n)} \Pr(X+Y=i|Y=q)\Pr(Y=q) \\
&= \sum_{q=0}^{\min(i,n)} \Pr(X=i-q)\Pr(Y=q) \\
&= \sum_{q=0}^{\min(i,n)} \frac{\lambda^{k-q}e^{-\lambda}}{(i-q)!}\binom{n}{q}\theta^q(1-\theta)^{n-q}.
\end{aligned}$$

2.3.5. Issues with Brillinger's Model. There are several problems we faced in making inference about λ_{BC} in Brillinger's model. The biggest problem was computational errors in the $D_{\text{obs}} = 0$ and $D_{\text{obs}} = P_{\text{obs}}$ cases, as such our simulations results in section 2.6 for Brillinger's model are exclude the $D_{\text{obs}} = 0$ and $D_{\text{obs}} = P_{\text{obs}}$ cases. A secondary problem was the inability to obtain a lower bound when using the method of moments constrained estimate for the nuisance parameters in the Brillinger's model. This phenomenon occurs in CDF pivot method as well as its saddlepoint approximation, which is discussed in section 2.5.

We tried imposing an additional restriction that $D \le P$, which was discussed in the introduction, to get a confidence interval for covariance λ_{BC} and obtain the probability calculation given in equation (2.17). Unfortunately, this results in a very complicated CDF expression involving five summations in which two are infinite. The CDF is as follows;

$$\begin{aligned}
&\Pr(D/P \le (D/P)_{\text{obs}} | P > 0, D \le P) \qquad (2.17)\\
&= \Pr(D \le aP | P > 0, D \le P) \\
&= \sum_{p=0}^{\infty} \sum_{d=0}^{\lfloor p(D/P)_{\text{obs}} \rfloor} \Pr(D=d|P=p, D\le P, P>0)\Pr(P=p|D\le P, P>0) \\
&= \sum_{p=1}^{\infty} \sum_{d=0}^{\lfloor p(D/P)_{\text{obs}} \rfloor} \frac{\Pr(D=d|P=p, D\le P)\Pr(P=p|D\le P)}{\Pr(P>0)} \\
&= \sum_{p=1}^{\infty} \sum_{d=0}^{\lfloor p(D/P)_{\text{obs}} \rfloor} \frac{\Pr(D=d, P=p, D\le P)}{\Pr(P>0)\Pr(P=p, D\le P)}\Pr(P=p|D\le P)
\end{aligned}$$

$$= \sum_{p=1}^{\infty} \sum_{d=0}^{\lfloor p(D/P)_{\text{obs}} \rfloor} \frac{\Pr(D=d, P=p, D \leq P)}{\Pr(P>0)\Pr(P=p, D \leq P)} \frac{\Pr(P=p, D \leq P)}{\Pr(D \leq P)}$$

$$= \sum_{p=1}^{\infty} \sum_{d=0}^{\lfloor p(D/P)_{\text{obs}} \rfloor} \frac{\Pr(D=d, P=p, D \leq P)}{\Pr(P>0)\Pr(D \leq P)}.$$

If we consider the summand of the above infinite double summations

$$\begin{aligned}
&\frac{\Pr(D=d, P=p, D \leq P)}{\Pr(P>0)\Pr(D \leq P)} \\
&= \frac{\Pr(X_{B-BC} + X_{BC} = d, X_{C-BC} + X_{BC} = p, X_{B-BC} \leq X_{C-BC})}{\Pr(P>0)\Pr(X_{B-BC} \leq X_{C-BC})} \\
&= \frac{\sum_{k=p-d}^{p} \Pr(X_{B-BC} + X_{BC} = d, X_{C-BC} + X_{BC} = p, X_{B-BC} \leq X_{C-BC} | X_{C-BC} = k)}{\Pr(P>0)\sum_{w=0}^{\infty} \Pr(X_{B-BC} \leq w)\Pr(X_{C-BC} = w)} \\
&\times \Pr(X_{C-BC} = k) \\
&= \frac{\sum_{k=p-d}^{p} \Pr(X_{B-BC} + X_{BC} = d, X_{BC} = p-k, X_{B-BC} \leq k)\Pr(X_{C-BC} = k)}{\Pr(P>0)\sum_{w=0}^{\infty}\sum_{r=0}^{w} \Pr(X_{B-BC} = r)\Pr(X_{C-BC} = w)} \\
&= \frac{\sum_{k=p-d}^{p} \Pr(X_{B-BC} = k-(p-d))\Pr\left(X_{BC} = p-k\right)\Pr(X_{C-BC} = k)}{\Pr(P>0)\sum_{w=0}^{\infty}\sum_{r=0}^{w} \Pr(X_{B-BC} = r)\Pr(X_{C-BC} = w)} \\
&= \frac{\sum_{k=p-d}^{p} \left(\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}} \lambda_{B-BC}^{k-(p-d)} \lambda_{C-BC}^{k} \lambda_{BC}^{p-k}}{(k-(p-d))!(p-k)!k!}\right)}{\sum_{w=0}^{\infty}\sum_{r=0}^{w} \left(\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}} \lambda_{B-BC}^{r} \lambda_{C-BC}^{w}}{r!w!}\right)(1-e^{-\lambda_{BC}-\lambda_{C-BC}})}.
\end{aligned}$$

So that;

$$\begin{aligned}
&\Pr\left(D/P \leq (D/P)_{\text{obs}} | P>0, D \leq P\right) \\
&= \sum_{p=1}^{\infty} \sum_{d=0}^{\lfloor p(D/P)_{\text{obs}} \rfloor} \frac{\sum_{k=p-d}^{p} \left(\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}-\lambda_{BC}} \lambda_{B-BC}^{k-(p-d)} \lambda_{C-BC}^{k} \lambda_{BC}^{p-k}}{(k-(p-d))!(p-k)!k!}\right)}{\sum_{w=0}^{\infty}\sum_{r=0}^{w} \left(\frac{e^{-\lambda_{B-BC}-\lambda_{C-BC}} \lambda_{B-BC}^{r} \lambda_{C-BC}^{w}}{r!w!}\right)(1-e^{-\lambda_{BC}-\lambda_{C-BC}})}.
\end{aligned}$$

In the next section we introduce a much simpler model that enforces the $D \leq P$ restriction and has number of theoretical and inferential advantages over Brillinger's model.

2.4. PROPOSED MODEL

This model is developed from Brillinger's model by setting $X_{B-BC} = 0$ in equation 2.3. As discussed in section 2.3.1 under the convention that $0^0 := 1$ this results in $\lambda_{B-BC} = 0$. Note that with this restriction we preserve the condition that $P > 0$ and in section 2.3.1 we point out that the regular MLE of λ_{B-BC} for Brillinger's unrestricted model is equal to zero when $D_{\text{obs}} \leq P_{\text{obs}}$. Hence, the setting of $X_{B-BC} = 0$ in the proposed model is consistent with the fitted Brillinger's model in the $D_{\text{obs}} \leq P_{\text{obs}}$ and $P_{\text{obs}} > 0$ setting.

Here, we estimate the mortality rate using the estimator D/P where we define

$$D = X_{BC} \sim \text{Poisson}(\lambda_{BC}) \quad \text{and} \quad P = X_{BC} + X_{C-BC} \sim \text{Poisson}(\lambda_{BC} + \lambda_{C-BC})$$

where, as before, X_{BC} and X_{C-BC} are two independent Poisson random variables with means λ_{BC} and λ_{C-BC}, respectively. The corresponding Lexis diagram for this model is given in Figure 2.2. The new D and P random variables are correlated, as were the old ones in Brillinger's model. We obtain the distribution of D/P given that $P > 0$ by first considering the joint distribution of (D, P) given that $P > 0$ that is obtained by taking $\lambda_{B-BC} = 0$ in equation 2.3 for Brillinger's model. More formally, we note the continuity in λ_{B-BC} of the expressions for joint PMF and joint CDF of D and P in Brillinger's model and take the limit as $\lambda_{B-BC} \to 0$ of these expressions. As such, the joint distribution of D and P given that $P > 0$ is

$$\begin{aligned}
&\Pr\left(D = d, P = p | P > 0\right) \\
&= \frac{e^{-\lambda_{BC} - \lambda_{C-BC}} \lambda_{BC}^{d} \lambda_{C-BC}^{p-d}}{d!\,(p-d)!\,\left(1 - e^{-\lambda_{BC} - \lambda_{C-BC}}\right)}.
\end{aligned}$$

where $D = 0, 1, ..., p$ and $P = 1, 2, ..., p, ..., \infty$.

Note that in principle one can obtain expected value of D/P given that $P > 0$ from Brillinger's model by setting $\lambda_{B-BC} = 0$ in the equation 2.5 as follows,

$$E\left(D/P|P>0\right)=\sum_{p=1}^{\infty}\sum_{d=0}^{\infty}\frac{d}{p}\frac{e^{-\lambda_{C-BC}-\lambda_{BC}}}{\left(1-e^{-\lambda_{C-BC}-\lambda_{BC}}\right)}\frac{\lambda_{BC}^{d}\lambda_{C-BC}^{p-d}}{d!(p-d)!}$$
$$=\frac{e^{-\lambda_{BC}-\lambda_{C-BC}}}{\left(1-e^{-\lambda_{BC}-\lambda_{C-BC}}\right)}\sum_{p=1}^{\infty}\frac{1}{p\times p!}\left(\lambda_{BC}+\lambda_{C-BC}\right)^{p}$$
$$\times\sum_{d=0}^{p}d\binom{p}{d}\left(\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\right)^{d}\left(\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\right)^{p-d}$$

where the second summation is the expectation of a binomial random variable with sample size p and proportion of success $\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}$, which can be denoted as $\mathrm{Bin}\left(p,\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\right)$. Now this will simplify to

$$E\left(D/P|P>0\right)$$
$$=\frac{e^{-\lambda_{BC}-\lambda_{C-BC}}}{1-e^{-\lambda_{BC}-\lambda_{C-BC}}}\sum_{p=1}^{\infty}\frac{1}{p\times p!}\left(\lambda_{BC}+\lambda_{C-BC}\right)^{p}\sum_{d=0}^{p}d\mathrm{Bin}\left(p,\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\right)$$
$$=\frac{e^{-\lambda_{BC}-\lambda_{C-BC}}}{1-e^{-\lambda_{BC}-\lambda_{C-BC}}}\sum_{p=1}^{\infty}\frac{1}{p\times p!}\left(\lambda_{BC}+\lambda_{C-BC}\right)^{p}p\left(\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\right)$$
$$=\frac{e^{-\lambda_{BC}-\lambda_{C-BC}}}{1-e^{-\lambda_{BC}-\lambda_{C-BC}}}\sum_{p=1}^{\infty}\frac{1}{p!}\left(\lambda_{BC}+\lambda_{C-BC}\right)^{p}\left(\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\right)$$
$$=\frac{1}{1-e^{-\lambda_{BC}-\lambda_{C-BC}}}\left(\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\right)\sum_{p=1}^{\infty}\frac{e^{-\lambda_{BC}-\lambda_{C-BC}}\left(\lambda_{BC}+\lambda_{C-BC}\right)^{p}}{p!}$$
$$=\frac{1}{1-e^{-\lambda_{BC}-\lambda_{C-BC}}}\left(\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\right)\left(1-e^{-\lambda_{BC}-\lambda_{C-BC}}\right)$$
$$=\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}.$$

This simple and easily interpretable closed-form expression of the expectation of D/P allows us to obtain confidence interval for the mean of the estimator D/P. We use the reparameterization

$$\theta_1=\frac{\lambda_{BC}}{\lambda_{BC}+\lambda_{C-BC}}\text{ and }\theta_2=\lambda_{C-BC}$$

in the forthcoming development of the confidence intervals for population vital rate θ_1.

2.4.1. Large Sample Method. The large sample 95% confidence interval for θ_1 is of the form

$$\widehat{\theta}_1 \pm 1.96\,(SE)\,.$$

where

$$\widehat{\theta}_1 = D/P$$

and standard error (SE) given as the observed Fisher information for θ_1;

$$J\left(\theta_1\right) = \frac{-C}{AC - B^2} = SE^2.$$

The log-likelihood function for θ_1 and θ_2 after reparameterization is

$$\begin{aligned} l(\theta_1, \theta_2 | D = d, P = p) = &-\theta_2/(1-\theta_1) + d\log(\theta_1/(1-\theta_1)) + p\log(\theta_2) \\ &- \log\left(1 - e^{-\theta_2/(1-\theta_1)}\right) - \log(d!(p-d)!). \end{aligned}$$

Taking second derivatives of this loglikelihood function, yields the following expressions for the components of $J\left(\theta_1\right)$,

$$\begin{aligned} A &= -\frac{D}{\widehat{\theta}_1^2} + \frac{D}{\left(1-\widehat{\theta}_1\right)^2} - \frac{\widehat{\theta}_2\left(2\left(1-\widehat{\theta}_1\right)\left(1-\psi\left(\widehat{\theta}_1,\widehat{\theta}_2\right)\right) + \widehat{\theta}_2\psi\left(\widehat{\theta}_1,\widehat{\theta}_2\right)\right)}{\left(1-\psi\left(\widehat{\theta}_1,\widehat{\theta}_2\right)\right)^2\left(1-\widehat{\theta}_1\right)^4}, \\ B &= \frac{-\left(1-\psi\left(\widehat{\theta}_1,\widehat{\theta}_2\right) - \frac{\widehat{\theta}_2}{1-\widehat{\theta}_1}\psi\left(\widehat{\theta}_1,\widehat{\theta}_2\right)\right)}{\left(1-\psi\left(\widehat{\theta}_1,\widehat{\theta}_2\right)\right)^2\left(1-\widehat{\theta}_1\right)^2}, \\ C &= -\frac{P}{\widehat{\theta}_2^2} - \frac{\psi\left(\widehat{\theta}_1,\widehat{\theta}_2\right)}{\left(1-\psi\left(\widehat{\theta}_1,\widehat{\theta}_2\right)\right)^2\left(1-\widehat{\theta}_1\right)^2}, \end{aligned}$$

where

$$\psi\left(\widehat{\theta}_1, \widehat{\theta}_2\right) = \exp\left(\frac{-\widehat{\theta}_2}{1-\widehat{\theta}_1}\right),$$

and $\widehat{\theta}_1$ and $\widehat{\theta}_2$ are maximum likelihood estimations of θ_1 and θ_2, respectively.

2.4.2. CDF Pivot Method. Using the CDF in equation 2.19 as a pivotal quantity, as was done for Brillinger's model in section 2.3, we obtain 95% confidence intervals for θ_1 which we denote as $\left(\hat{\theta}_{1L}^{CDF}, \hat{\theta}_{1U}^{CDF}\right)$ and which satisfy the following equations;

$$\Pr\left(D/P \leq (D/P)_{\text{obs}} \,|\hat{\theta}_{1L}^{CDF}, \widehat{\theta}_{2|\hat{\theta}_{1L}^{CDF}}, P>0\right) = 0.975 \tag{2.18}$$
$$\Pr\left(D/P \leq (D/P)_{\text{obs}} \,|\theta_{1U}^{CDF}, \widehat{\theta}_{2|\theta_{1U}^{CDF}}, P>0\right) = 0.025.$$

Note that we used the constrained MLE of θ_2 ($\widehat{\theta}_{2|\theta_1}$), as discussed in section 2.3.3 to solve the equations given in (2.18). Also, $(D/P)_{\text{obs}}$ is the observed value of estimator D/P. The constrained MLE of θ_2 can be obtained by solving following equation for θ_2 with a fixed value of θ_1;

$$\frac{P_{\text{obs}}}{\theta_2} - \frac{1}{(1-\theta_1)\left(1-\exp\left(\frac{-\theta_2}{1-\theta_1}\right)\right)} = 0.$$

The resulting CDF of estimator D/P, as a function of θ_1, is given as follows;

$$\Pr\left(D/P \leq (D/P)_{\text{obs}} \,|\theta_1, \widehat{\theta}_{2|\theta_1}, P>0\right) \tag{2.19}$$
$$= \sum_{p=1}^{\infty} \sum_{d=0}^{\lfloor p(D/P)_{\text{obs}} \rfloor} \frac{e^{-\widehat{\theta}_{2|\theta_1}/(1-\theta_1)} \widehat{\theta}_{2|\theta_1} \widehat{\theta}_{2|\theta_1}^{p} \left(\theta_1/\left(1-\theta_1\right)\right)^d}{d!(p-d)!\left(1-e^{-\widehat{\theta}_{2|\theta_1}/(1-\theta_1)}\right)}.$$

For the CDF pivot method we encountered numerical issues when $D_{\text{obs}} = P_{\text{obs}}$ ($X_{C-BC} = 0$) since according to our model the above CDF given in equation (2.19) is equal to one for any value of λ_{BC} and λ_{C-BC}. To overcome this problem, we use confidence intervals generated from a "Clopper-Pearson" type method which we discuss next.

2.4.3. "Clopper-Pearson" Type Method. The method we discuss here is motivated by the classical Clopper and Pearson (1934) confidence interval for a proportion. With this in mind we determine upper and lower 95% confidence bounds $\left(\hat{\theta}_{1L}^{CP}, \hat{\theta}_{1U}^{CP}\right)$ as a solution to equations given below:

$$\begin{aligned} \hat{\theta}_{1L}^{CP} &= \inf\left\{\theta_1 : \Pr\left(D/P \geq (D/P)_{\text{obs}} | \theta_1, \hat{\theta}_{2|\theta_1}, P > 0\right) > 0.025\right\} \\ \hat{\theta}_{1U}^{CP} &= \sup\left\{\theta_1 : \Pr\left(D/P \leq (D/P)_{\text{obs}} | \theta_1, \widehat{\theta}_{2|\theta_1}, P > 0\right) > 0.025\right\} \end{aligned} \tag{2.20}$$

where again $\widehat{\theta}_{2|\theta_1}$ is the constrained MLE of θ_2. Note that the "Clopper-Pearson" type upper confidence bound is equal to upper confidence bound of CDF pivot method in section 2.4.2. To obtain the lower confidence bound one needs to solve following equation for $\hat{\theta}_{1L}^{CP}$;

$$\Pr\left(D/P \geq (D/P)_{\text{obs}} | \hat{\theta}_{1L}^{CP}, \widehat{\theta}_{2|\theta_{1\hat{\theta}_{1L}^{CP}}}, P > 0\right) = 0.025.$$

The following calculation shows how to obtain lower bound for the $D_{\text{obs}} = P_{\text{obs}}$ case, which is a problem for the CDF pivot method as discussed above in section 2.4.2. Here we have suppressed much of the dependence on parameters and their estimates for the sake of clarity.

$$\begin{aligned} &\inf\{\theta_1 : \Pr(D/P \geq 1|, P > 0) > 0.025\} \\ &= \inf\{\theta_1 : \Pr(D \geq P|P > 0) > 0.025\} \\ &= \inf\{\theta_1 : \Pr(D = P|P > 0) > 0.025\} \\ &= \inf(\{\theta_1 : \Pr(X_{BC} = X_{C-BC} + X_{BC}|P > 0) > 0.025\} \\ &= \inf\{\theta_1 : \Pr(X_{C-BC} = 0|X_{C-BC} + X_{BC} > 0) > 0.025\} \\ &= \inf\left\{\theta_1 : \frac{\Pr(X_{C-BC} = 0, X_{C-BC} + X_{BC} > 0)}{\Pr(X_{C-BC} + X_{BC} > 0)} > 0.025\right\} \\ &= \inf\left\{\theta_1 : \frac{\Pr(X_{C-BC} = 0, X_{BC} > 0)}{\Pr(X_{C-BC} + X_{BC} > 0)} > 0.025\right\} \\ &= \inf\left\{\theta_1 : \frac{e^{-\theta_2}\left(1 - e^{-\theta_1\theta_2/(1-\theta_1)}\right)}{1 - e^{-\theta_1\theta_2/(1-\theta_1)}} > 0.025\right\}. \end{aligned}$$

The obtainment of the infimum yields the following equation:

$$\frac{e^{-\theta_2}\left(1-e^{-\theta_1\theta_2/(1-\theta_1)}\right)}{1-e^{-\theta_1\theta_2/(1-\theta_1)}}=0.025.$$

Here again, we solve this equation in θ_1 after having replaced θ_2 with its constrained MLE $\widehat{\theta}_{2|\theta_1}$.

The resulting "Clopper-Pearson" type confidence interval can be denote as $(\hat{\theta}_{1L}^{CP}, 1)$ since the upper bound is one. The calculation for the upper bound is as follows;

$$\begin{aligned}
&\sup\left\{\theta_1 : \Pr\left(D/P \leq 1|P>0\right)>0.025\right\}\\
&=\sup\left\{\theta_1 : \Pr\left(D \leq P|P>0\right)>0.025\right\}\\
&=\sup\left\{\theta_1 : \Pr\left(X_{BC} \leq X_{C-BC}+X_{BC}|P>0\right)>0.025\right\}\\
&=\sup\left\{\theta_1 : \Pr\left(X_{C-BC} \geq 0|X_{C-BC}+X_{BC}>0\right)>0.025\right\}\\
&=\sup\left\{\theta_1 : \frac{\Pr\left(X_{C-BC}\geq 0, X_{C-BC}+X_{BC}>0\right)}{\Pr\left(X_{C-BC}+X_{BC}>0\right)}>0.025\right\}\\
&=\sup\left\{\theta_1 : \frac{\Pr\left(X_{C-BC}+X_{BC}>0\right)}{\Pr\left(X_{C-BC}+X_{BC}>0\right)}>0.025\right\}\\
&=\sup\left\{\theta_1 : 1>0.025\right\}
\end{aligned}$$

and the supremum of the above set is

$$\theta_1 = 1.$$

Finally, note that for Brillinger's model, in section 2.3, we discuss two types of constrained estimates (constrained MLE's and method of moment constrained estimates) for the nuisance parameters, but here we consider only the constrained MLE. This is because simulation results for Brillinger's model show that the method of moment constrained estimate yields clearly inferior confidence intervals as the theory of Lee and Young (2005), DiCiccio *et al.* (2001) and Diciccio and Romano (1995) would suggest. Therefore we did not consider method of moment constrained estimates for our proposed model.

2.4.3.1. Comparison of lower bounds. In this subsection we investigate how the lower bounds of CDF pivot method and the "Clopper-Pearson" type method are related. Consider

$$
\begin{aligned}
\hat{\theta}_{1L}^{CP} &= \inf\left\{\theta_1 : \Pr\left(D/P \geq (D/P)_{\text{obs}} \,|P > 0\right) > 0.025\right\} \\
&= \inf\left\{\theta_1 : 1 - \Pr\left(D/P < (D/P)_{\text{obs}} \,|P > 0\right) > 0.025\right\} \\
&= \inf\left\{\theta_1 : 1 - \Pr\left(D < (D/P)_{\text{obs}} P|P > 0\right) > 0.025\right\} \\
&= \inf\{\theta_1 : 1 - \Pr\left(D \leq (D/P)_{\text{obs}} P|P > 0\right) + \\
&\Pr\left(D = (D/P)_{\text{obs}} P|P > 0\right) \; > \; 0.025\} \\
&= \sup\{\theta_1 : \Pr\left(D \leq (D/P)_{\text{obs}} P|P > 0\right) - \\
&\Pr\left(D = (D/P)_{\text{obs}} P|P > 0\right) \; < \; 0.975\} \\
&= \Pr\left(D \leq (D/P)_{\text{obs}} P|P > 0\right) - \\
&\Pr\left(D = (D/P)_{\text{obs}} P|P > 0\right) \; = \; 0.975
\end{aligned}
$$

and note that

$$
\hat{\theta}_{1L}^{CDF} = \left\{\theta_1 | \Pr\left(D \leq (D/P)_{\text{obs}} P|P > 0\right) = 0.975\right\}.
$$

2.4.4. Mid-P Correction. The random variable D/P is a discrete random variable taking on values over $\mathbb{Q}^+$, the set of non-negative rational numbers. As such it is different from a typical discrete random variable which takes on only integer values. The regular one-dimensional mid-P correction is given in Berry and Armitage (1995) and Agresti and Gottard (2005). Therefore to apply this correction one would have to model D/P as having a lattice distribution with span $1/P_{\text{obs}}$. With this in mind, we introduce two methods for mid-P correction.

First, by way of review, we consider the classical setting from Berry and Armitage (1995) and Agresti and Gottard (2005) and how the mid-P correction would be used to generate a 95% confidence interval for mean of a Poisson random variable with mean λ

where

$$\Pr(X = x; \lambda) = \frac{e^{-\lambda}\lambda^{x}}{x!}.$$

If the observed value of X is 3, then confidence interval based upon a mid-P correction is given as the solution to following equations

$$\begin{aligned}\Pr(X = 0; \lambda) + \Pr(X = 1; \lambda) + \Pr(X = 2; \lambda) + \frac{1}{2}\Pr(X = 3; \lambda) = 0.975\\ \Pr(X = 0; \lambda) + \Pr(X = 1; \lambda) + \Pr(X = 2; \lambda) + \frac{1}{2}\Pr(X = 3; \lambda) = 0.025.\end{aligned}$$

The resulting 95% confidence interval is $(0.763, 8.164)$.

2.4.4.1. One-dimensional mid-P correction. For the usual discrete random variable defined over an integer lattice the mid-P correction is defined as the subtraction one half of the probability of the boundary point from the P-value of the observed data, as shown above in section 2.4.4. For random variable D/P we introduce first method of mid-P correction involving the subtraction of $1/(2P_{\text{obs}})^{th}$ portion of the probability from the P-value of the observed data. The CDF of D/P given in equation (2.18); and used in the solution of

$$\begin{aligned}\Pr\left(D/P \leq (D/P)_{\text{obs}} \,|\hat{\theta}_{1L}^{CDF}, \widehat{\theta}_{2|\hat{\theta}_{1L}^{CDF}}, P > 0\right) = 0.975\\ \Pr\left(D/P \leq (D/P)_{\text{obs}} \,|\theta_{1U}^{CDF}, \widehat{\theta}_{2|\theta_{1U}^{CDF}}, P > 0\right) = 0.025\end{aligned}$$

would therefore be modified as

$$\begin{aligned}&\Pr\left(D/P \leq (D/P)_{\text{obs}} \,|\theta_{1L}^{CDF}, \widehat{\theta}_{2|\theta_{1L}^{CDF}}, P > 0\right) -\\ &\frac{1}{2P_{\text{obs}}}\Pr\left(D/P = (D/P)_{\text{obs}} \,|\theta_{1L}^{CDF}, \widehat{\theta}_{2|\theta_{1L}^{CDF}}, P > 0\right) \;=\; 0.975\\ &\Pr\left(D/P \leq (D/P)_{\text{obs}} \,|\theta_{1U}^{CDF}, \widehat{\theta}_{2\theta_{1U}^{CDF}}, P > 0\right) -\\ &\frac{1}{2P_{\text{obs}}}\Pr\left(D/P = (D/P)_{\text{obs}} \,|\theta_{1U}^{CDF}, \widehat{\theta}_{2|\theta_{1U}^{CDF}}, P > 0\right) \;=\; 0.025.\end{aligned}$$

For example, consider data in which $D_{\text{obs}} = 4$ and $P_{\text{obs}} = 10$. The 95% confidence interval for the mean of D/P using CDF pivot method with one-dimensional mid-P correction given above is $(0.1448, 0.7291)$

2.4.4.2. Two-dimensional mid-P correction. In this section we develop a novel two-dimensional mid-P correction. Here we consider the joint distribution of the random variables D and P as shown in Figure 2.5. To motivate this method consider the following modification to equation (2.18). Notice that the boundary for the region being computed by this equation is the

$$D = (D/P)_{\text{obs}} P$$

line. For the mid-P calculation we consider mass points which are first-nearest neighbors of the boundary line which are shown in Figure 2.5 as solid dots. For each pair of vertically aligned neighboring points, we determine a mass on the border line point in between them via linear interpolation. In effect we are smoothing the underlying discrete joint distribution of D and P via linear interpolation. Hence the required correction is $1/(2P_{\text{obs}})^{th}$ of the sum of the linearly interpolated probabilities on boundary line

$$D = (D/P)_{\text{obs}} P.$$

The formula for $(IP(p))$, the interpolated probability at $P = p$ is given as

$$\begin{aligned} IP(p) = {} & (1 - (p\,(D/P)_{\text{obs}} - \lfloor p\,(D/P)_{\text{obs}} \rfloor)) \Pr\left(D/P = \lfloor p\,(D/P)_{\text{obs}} \rfloor | P > 0\right) + \\ & (p\,(D/P)_{\text{obs}} - \lfloor p\,(D/P)_{\text{obs}} \rfloor) \Pr\left(D/P = \lfloor p\,(D/P)_{\text{obs}} \rfloor + 1 | P > 0\right). \end{aligned} \tag{2.21}$$

The resulting modified version of equation(2.18) , is as follows,

$$\Pr\left(D/P \leq (D/P)_{\text{obs}} | \theta_{1L}^{CDF}, \widehat{\theta}_{2|\theta_{1L}^{CDF}}, P > 0\right) - \frac{1}{2P_{\text{obs}}} \sum_{p=1}^{\infty} IP(p) = 0.975$$

$$\Pr\left(D/P \le (D/P)_{\text{obs}} \left| \theta_{1U}^{CDF}, \widehat{\theta}_{2|\theta_{1U}^{CDF}}, P > 0\right.\right) - \frac{1}{2P_{\text{obs}}} \sum_{p=1}^{\infty} IP(p) \;=\; 0.025.$$

If we consider the hypothetical example from section 2.4.4.1, wherein $D_{\text{obs}} = 4$ and $P_{\text{obs}} = 10$, then the 95% confidence interval for the mean of D/P using CDF pivot method with two dimensional mid-P correction is $(0.1407, 0.7247)$.

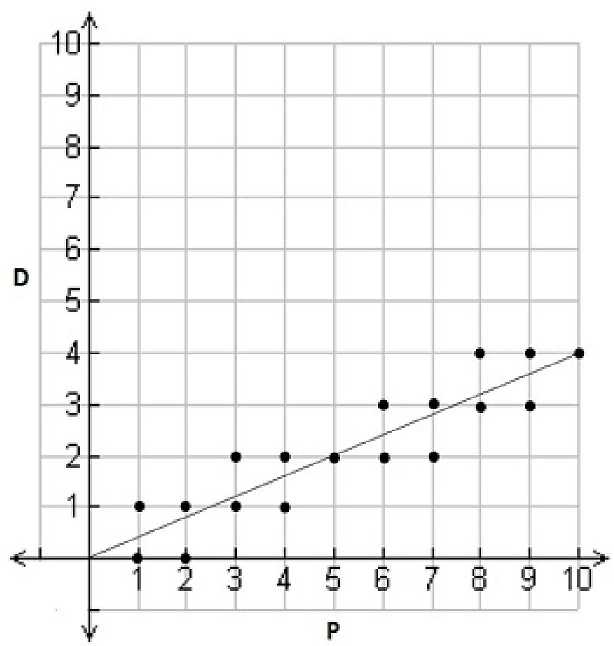


Figure 2.5. Illustration of points considered in mid-P correction method 2 for the case $(D/P)_{\text{obs}} = 0.4$.

2.5. SADDLEPOINT CONFIDENCE INTERVALS

2.5.1. Brillinger's Model. Here we present a method to approximate the CDF pivot confidence interval for the covariance (λ_{BC}) of the bivariate Poisson distribution which underlies Brillinger's model, that involves saddlepoint approximations. Note that the CDF for D/P given in equation (2.4) which can be rewritten as follows:

$$\Pr\left(D/P \le (D/P)_{\text{obs}} \left| P > 0\right.\right) \tag{2.22}$$

$$
\begin{aligned}
&= \Pr\left(D \leq P\left(D/P\right)_{\text{obs}} | P > 0\right) \\
&= \Pr\left(D - P\left(D/P\right)_{\text{obs}} \leq 0 | P > 0\right) \\
&= \Pr\left(X_{B-BC} + X_{BC} \leq \left(X_{C-BC} + X_{BC}\right)\left(D/P\right)_{\text{obs}} | P > 0\right) \\
&= \Pr\left(-\left(D/P\right)_{\text{obs}} X_{C-BC} + \left(1 - \left(D/P\right)_{\text{obs}}\right) X_{BC} + X_{B-BC} \leq 0 | P > 0\right).
\end{aligned}
$$

One can easily obtain the moment generating function (MGF) of a linear combination of three independent random variables and from that obtain the cumulant generating function (CGF) as follows,

$$
\begin{aligned}
&K_{X_{B-BC},X_{C-BC},X_{BC}|X_{C-BC}+X_{BC}>0}\left(s, -\left(D/P\right)_{\text{obs}} s, (1 - \left(D/P\right)_{\text{obs}})s\right) \\
&= \log\left(M_{X_{C-BC}}(-\left(D/P\right)_{\text{obs}} s) M_{X_{BC}}((1 - \left(D/P\right)_{\text{obs}})s) - e^{(-\lambda_{C-BC} - \lambda_{BC})}\right) \\
&- \log\left(1 - e^{-\lambda_{C-BC} - \lambda_{BC}}\right) + \log\left(M_{X_{B-BC}}(s)\right) \\
&= \log\left(\exp(\lambda_{C-BC}(e^{-(D/P)_{\text{obs}} s} - 1) + \lambda_{BC}(e^{(1-(D/P)_{\text{obs}})s} - 1) - e^{-\lambda_{C-BC} - \lambda_{BC}}\right) \\
&- \log\left(1 - e^{-\lambda_{C-BC} - \lambda_{BC}}\right) + \lambda_{B-BC}(e^{s} - 1).
\end{aligned}
$$

The first and second derivatives of this CGF and the Luganani and Rice (LR) approximation for a CDF provides access to the saddlepoint approximation to the CDF given in equation (2.22) .The LR saddlepoint approximation to the CDF of

$$
\begin{aligned}
Y &= \left[D - P\left(D/P\right)_{\text{obs}}\right] | \left\{X_{C-BC} + X_{BC} > 0\right\} \\
&= \left[-\left(D/P\right)_{\text{obs}} X_{C-BC} + \left(1 - \left(D/P\right)_{\text{obs}}\right) X_{BC} + X_{B-BC}\right] | \left\{X_{C-BC} + X_{BC} > 0\right\}
\end{aligned}
$$

can be obtained by modifying the equation given in equation (1.5) as follows:

$$
\widehat{\Pr}\left(Y \leq 0; \lambda_{BC}\right) = \begin{cases} \Phi(\hat{t}) + \phi(\hat{t})\left[\hat{t}^{-1} - \hat{u}^{-1}\right], & \text{if } \; E(Y) \neq 0 \\ \frac{1}{2} + K_Y^{(3)}(0)\left[72\pi K_Y^{(2)}(0)^3\right]^{-1/2}, & \text{if } \; E(Y) = 0 \end{cases} \tag{2.23}
$$

where

$$
\hat{t} = sgn\left(\hat{s}\right)\sqrt{2\left[-K_Y\left(\hat{s}\right)\right]}
$$

$$\hat{u} = \hat{s}\sqrt{K_Y^{(2)}(\hat{s})}$$

and where saddlepoint $\hat{s}$ is the solution to saddlepoint equation

$$K_Y^{(1)}(\hat{s}) = 0.$$

From here we simply replace the exact CDF in (2.4) with its LR saddlepoint approximation and proceed as we did in section 2.3.2.

Note that for these saddlepoint confidence intervals we use a saddlepoint approximation to the root of random estimating equation

$$Y = \left[-\left(D/P\right)_{\text{obs}} X_{C-BC} + \left(1 - \left(D/P\right)_{\text{obs}}\right) X_{BC} + X_{B-BC}\right] \left| \left\{X_{C-BC} + X_{BC} > 0\right\} = 0\right.$$

whereas in section 1.2 our saddlepoint confidence intervals are based upon the saddlepoint approximation of the CDF for the directly standardized rate; see Butler (2007, chapter 12) for more details on saddlepoint approximation of roots of estimation equations. The Luganani and Rice saddlepoint CDF approximation is obtained by approximating the Riemann-Lebesque integral of the saddlepoint density approximation, which is a continuous approximation to the discrete mass function, by Temme's method (Butler,2007, chapter 2.3). As such, the saddlepoint approximation automatically provides an one-dimensional mid-P type correction for discreteness.

2.5.2. Proposed Model. Here we obtain a confidence interval for parameter θ_1 in our proposed model (see section (2.4)) wherein the exact CDF of D/P given in equation (2.19) is replaced by its LR saddlepoint CDF approximation. The exact CDF can be written as

$$\begin{aligned}
&\Pr\left(D/P \leq \left(D/P\right)_{\text{obs}} | P > 0\right) \\
&= \Pr\left(D \leq P\left(D/P\right)_{\text{obs}} | P > 0\right) \\
&= \Pr\left(D - P\left(D/P\right)_{\text{obs}} \leq 0 | P > 0\right) \\
&= \Pr\left(X_{BC} - \left(X_{BC} + X_{C-BC}\right)\left(D/P\right)_{\text{obs}} \leq 0 | P > 0\right)
\end{aligned}$$

$$= \Pr\left(\left(1-(D/P)_{\text{obs}}\right)X_{BC} - (D/P)_{\text{obs}}\,X_{C-BC} \le 0 | P > 0\right).$$

The CGF for

$$Y = \left[\left(1-(D/P)_{\text{obs}}\right)X_{BC} - (D/P)_{\text{obs}}\,X_{C-BC}\right] | \{X_{BC} + X_{C-BC} > 0\}$$

can be found as,

$$\begin{aligned}
&K_{X_{BC},X_{C-BC}|X_{BC}+X_{C-BC}>0}\left(\left(1-(D/P)_{\text{obs}}\right)s, -(D/P)_{\text{obs}}\,s\right) \\
&= \log\left(M_{X_{C-BC}}\left(-(D/P)_{\text{obs}}\,s\right)M_{X_{BC}}\left(\left(1-(D/P)_{\text{obs}}\right)s\right) - e^{-\lambda_{BC}-\lambda_{C-BC}}\right) - \\
&\log\left(1-e^{-\lambda_{BC}-\lambda_{C-BC}}\right) \\
&= \log\left(\exp\left(\lambda_{C-BC}(e^{-(D/P)_{\text{obs}}s}-1\right) + \lambda_{BC}\left(e^{(1-(D/P)_{\text{obs}})s}\right) - e^{-\lambda_{BC}-\lambda_{C-BC}}\right) - \\
&\log\left(1-e^{-\lambda_{BC}-\lambda_{C-BC}}\right) \\
&= \log\left(\exp\left(\theta_2(e^{-(D/P)_{\text{obs}}s}-1\right) + \left(\theta_1\theta_2/\left(1-\theta_1\right)\right)\left(e^{(1-(D/P)_{\text{obs}})s}\right) - e^{-\theta_2/(1-\theta_1)}\right) - \\
&\log\left(1-e^{-\theta_2/(1-\theta_1)}\right).
\end{aligned}$$

Using the same technique described in the section (2.5.1) we obtained an approximate confidence interval for the θ_1 by applying the LR approximation to the CDF given in equation (2.23). Here, again θ_2 is estimated using constrained MLE $\widehat{\theta}_{2|\theta_1}$ as described in section 2.4.2. In the simulation studies which follow we observe that the saddlepoint confidence interval for the two cases $D_{\text{obs}} = 0$ and $D_{\text{obs}} = P_{\text{obs}}$ did not exist since these cases correspond to boundaries of the support and so saddlepoint $\hat{s}$ is infinite in absolute value. Therefore, for these cases, we use the "Clopper-Pearson"type method discussed in section 2.4.3. Nonetheless, the SP approximation provides an automatic one-dimensional mid-P type correction for discreteness in this setting as well.

2.6. SIMULATION STUDIES

In this section we simulate coverage probabilities for the confidence interval methods applied to Brillinger's model and to the proposed model, as described in the sections 2.3 and 2.4.

2.6.1. Brillinger's Model. Here we generate $10,000$ (D, P) values for each combination of the λ_{B-BC} , λ_{C-BC} and λ_{BC} parameters and for each we construct 95% confidence intervals for λ_{BC}, the correlation parameter in Brillinger's model. From these confidence intervals we compute empirical coverage probabilities. We vary each of λ_{B-BC} , λ_{C-BC} and λ_{BC} from 1 to 9 in steps of 2. The confidence intervals we consider are the CDF pivot method with constrained MLE's (CDFML), the CDF pivot method with method of moment constrained estimates (CDFMM), the saddlepoint approximation to CDFML (SPML), and the saddlepoint approximation to CDFMM (SPMM). Table 2.1 presents the coverage probabilities for methods CDFML, CDFMM, SPML and SPMM under the various combinations of the λ_{B-BC} , λ_{C-BC} and λ_{BC} parameters. For methods CDFMM and SPMM we were unable to obtain lower confidence bounds. Therefore we obtained 95% upper confidence bounds and report the resulting coverage probabilities for them. Note that for the coverage probabilities, of all the methods considered, we remove the $D_{\text{obs}} = 0$ and $D_{\text{obs}} = P_{\text{obs}}$ cases due to the aforementioned computational problems.

The performance of the CDF pivot method and its saddlepoint approximations is comparable for the constrained MLE's as well as method of moment constrained estimates. However, the constrained MLE's result in more conservative coverage probabilities in comparison to method of moment constrained estimates for small sample size estimates and small numbers of deaths. Future work might consider confidence interval constructions for the $D_{\text{obs}} = 0$ and $D_{\text{obs}} = P_{\text{obs}}$ cases to see whether this results in improved coverage probabilities. Note also that with the method of moment constrained estimates we construct a 95% upper bound while taking the lower bound to be zero. Therefore we observe much conservative results when true value of λ_{BC} is small. The conservative results, for all of the confidence intervals considered, are mainly due to the need to estimate two nuisance parameters. In the next section, the model we proposed entails the estimation of only one nuisance parameter and as we will see, this results in much better coverage probabilities overall.

2.6.2. Proposed Model. Here again we generate $10,000$ (D, P) values for each combination of the λ_{C-BC} and λ_{BC} parameters and for each we construct 95%

Table 2.1. Coverage Probabilities for Brillinger's Model

λ_{B-BC}	λ_{C-BC}	λ_{BC}	CDFML	CDFMM	SPML	SPMM	λ_{B-BC}	λ_{C-BC}	λ_{BC}	CDFML	CDFMM	SPML	SPMM
1	1	1	100.00	85.25	100.00	100.00	5	1	1	99.91	98.17	99.86	100.00
1	1	3	100.00	73.32	100.00	76.27	5	1	3	99.97	88.97	99.97	90.70
1	1	5	98.61	69.48	99.46	68.95	5	1	5	98.47	98.43	98.51	89.42
1	1	7	99.60	66.82	99.66	66.82	5	1	7	98.94	85.18	99.26	85.18
1	1	9	99.61	65.00	99.70	65.00	5	1	9	99.26	82.96	99.41	82.96
1	3	1	99.80	97.83	99.95	100.00	5	3	1	99.59	99.37	99.63	100.00
1	3	3	100.00	89.50	100.00	89.90	5	3	3	99.97	96.71	99.98	97.36
1	3	5	99.84	83.46	99.64	81.46	5	3	5	99.64	95.72	99.67	95.62
1	3	7	99.71	77.97	99.62	77.95	5	3	7	99.82	92.93	99.88	92.93
1	3	9	99.78	75.52	99.78	75.11	5	3	9	99.85	90.11	99.88	90.09
1	5	1	99.20	99.78	99.59	100.00	5	5	1	97.46	99.93	98.60	100.00
1	5	3	99.98	96.55	99.99	96.60	5	5	3	99.91	99.31	99.95	99.51
1	5	5	99.97	91.37	99.19	89.86	5	5	5	99.94	98.41	99.92	98.24
1	5	7	99.61	85.60	99.18	85.14	5	5	7	99.94	96.25	99.94	96.25
1	5	9	99.53	83.65	99.45	82.33	5	5	9	99.98	94.80	99.95	94.65
1	7	1	97.80	99.96	98.69	100.00	5	7	1	89.32	99.98	93.01	100.00
1	7	3	99.88	98.97	99.90	98.99	5	7	3	99.47	99.82	99.62	99.84
1	7	5	99.92	93.41	98.77	92.67	5	7	5	99.94	99.18	99.93	99.06
1	7	7	99.47	90.13	99.01	89.36	5	7	7	99.98	98.15	99.97	98.07
1	7	9	99.29	88.72	99.19	86.83	5	7	9	99.98	96.98	99.98	96.73
1	9	1	96.45	99.99	97.52	100.00	5	9	1	75.12	100.00	80.53	100.00
1	9	3	99.62	99.68	99.69	99.68	5	9	3	97.89	99.99	98.41	99.99
1	9	5	99.75	95.17	98.51	94.40	5	9	5	99.65	99.56	99.63	99.53
1	9	7	99.13	92.85	98.76	92.06	5	9	7	99.82	99.18	99.82	99.12
1	9	9	99.22	91.81	99.16	90.36	5	9	9	99.94	98.50	99.95	98.20
3	1	1	100.00	92.84	100.00	100.00	7	1	1	99.71	99.49	99.56	100.00
3	1	3	100.00	83.33	100.00	87.14	7	1	3	99.94	91.82	99.93	92.48
3	1	5	98.24	82.24	98.59	82.19	7	1	5	98.51	92.20	98.36	92.20
3	1	7	99.27	78.24	99.59	78.24	7	1	7	98.75	89.23	98.91	89.23
3	1	9	99.62	74.65	99.75	74.64	7	1	9	99.19	87.01	99.27	87.01
3	3	1	99.80	98.09	99.93	100.00	7	3	1	98.34	99.85	97.72	100.00
3	3	3	100.00	93.85	100.00	95.07	7	3	3	99.75	98.02	99.76	98.33
3	3	5	99.75	91.16	99.77	90.69	7	3	5	99.66	97.27	99.65	97.26
3	3	7	99.81	87.42	99.87	87.41	7	3	7	99.74	95.51	99.77	95.51
3	3	9	99.95	84.48	99.95	84.42	7	3	9	99.70	93.96	99.78	93.96
3	5	1	97.88	99.67	99.07	100.00	7	5	1	95.04	99.99	95.31	100.00
3	5	3	99.98	98.25	99.99	98.43	7	5	3	99.60	99.59	99.62	99.65
3	5	5	99.93	96.23	99.84	95.72	7	5	5	99.83	99.12	99.81	99.10
3	5	7	99.92	92.56	99.88	92.51	7	5	7	99.93	98.01	99.98	98.01
3	5	9	99.91	91.11	99.89	90.57	7	5	9	99.96	97.05	99.96	97.05
3	7	1	92.55	99.93	95.41	100.00	7	7	1	88.49	99.98	90.93	100.00
3	7	3	99.72	99.50	99.78	99.54	7	7	3	99.16	99.93	99.36	99.97
3	7	5	99.92	97.97	99.72	97.72	7	7	5	99.83	99.78	99.88	99.73
3	7	7	99.86	95.78	99.74	95.56	7	7	7	100.00	99.24	99.99	99.23
3	7	9	99.85	94.14	99.79	93.32	7	7	9	99.97	98.52	99.97	98.50
3	9	1	84.20	99.99	88.05	100.00	7	9	1	74.51	100.00	78.4	100.00
3	9	3	98.70	99.85	99.00	99.87	7	9	3	96.95	100.00	97.57	100.00
3	9	5	99.76	98.36	99.56	98.04	7	9	5	99.43	99.89	99.55	99.88
3	9	7	99.70	97.19	99.62	97.03	7	9	7	99.91	99.80	99.91	99.78
3	9	9	99.81	96.09	99.80	95.43	7	9	9	99.98	99.27	99.98	99.19

Table 2.1. Coverage Probabilities for Brillinger's Model (cont.)

λ_{B-BC}	λ_{C-BC}	λ_{BC}	CDFML	CDFMM	SPML	SPMM
9	1	1	99.02	99.93	98.72	100.00
9	1	3	99.81	92.47	99.73	92.67
9	1	5	98.83	93.65	98.47	93.68
9	1	7	98.87	92.56	98.93	92.56
9	1	9	99.01	89.80	99.06	89.80
9	3	1	94.94	99.95	93.97	100.00
9	3	3	99.31	98.48	99.25	98.53
9	3	5	99.68	98.02	99.53	98.02
9	3	7	99.79	97.27	99.82	97.23
9	3	9	99.78	96.25	99.80	96.25
9	5	1	89.49	99.99	88.55	100.00
9	5	3	98.27	99.55	98.28	99.58
9	5	5	99.78	99.55	99.72	99.54
9	5	7	99.85	98.96	99.84	98.96
9	5	9	99.88	98.18	99.88	98.17
9	7	1	83.43	100.00	83.76	100.00
9	7	3	97.65	99.95	97.91	99.95
9	7	5	99.71	99.85	99.65	99.85
9	7	7	99.96	99.66	99.92	99.66
9	7	9	99.98	99.21	99.98	99.21
9	9	1	73.69	100.00	76.04	100.00
9	9	3	95.79	99.98	96.55	99.98
9	9	5	99.32	99.99	99.40	99.98
9	9	7	99.87	99.87	99.85	99.86
9	9	9	99.97	99.69	99.97	99.67

confidence intervals for θ_1 the expected value of the vital rate estimator D/P. From these confidence intervals we compute empirical coverage probabilities. We vary each of value of λ_{C-BC} and λ_{BC} from 1 to 9 in steps of 2. The confidence intervals we consider are the CDF pivot method with no mid-P correction (CDF0), one-dimensional mid-P correction (CDF1) and two-dimensional mid-P correction (CDF2), respectively, the "Clopper-Pearson" type method with no mid-P correction (CP0), one-dimensional mid-P correction (CP1) and two-dimensional mid-P correction (CP2), respectively, the large sample method (LS) and saddlepoint approximation to CDF0 (SP0). Table 2.2 presents the coverage probabilities for methods CDF0, CDF1, CDF2, CP0, CP1, CP2, LS and SP0 under the various settings for the λ_{C-BC} and λ_{BC} parameters. For the CDF pivot method it turns out that a confidence interval cannot be obtained when we observe $D_{\text{obs}} = P_{\text{obs}}$. In the "Clopper-Pearson" type method we do not have this problem, and therefore for the $D_{\text{obs}} = P_{\text{obs}}$ cases we replaced in the unobtainable CDF pivot confidence intervals with "Clopper-Pearson" type confidence intervals and the proportion of the time we made this replacement for CDF pivot method is given in the "Patch 1" column of Table 2.2. For the large sample method and the saddlepoint approximation to the CDF0 method we cannot find confidence intervals at the two boundaries where $D_{\text{obs}} = 0$ and $D_{\text{obs}} = P_{\text{obs}}$. For the large sample method, this situation occurs due to an infinite standard error and in the saddlepoint method the saddlepoint $\hat{s}$, is infinite at the boundary of the support. Therefore for these cases we also replaced the confidence intervals which cannot be computed with "Clopper-Pearson" type confidence intervals and the proportion of the time we made this substitution is given in the "Patch 2" column of Table 2.2.

To further investigate the performance of the various confidence interval methods for the rate parameter θ_1, we consider the simulation over a finer grid of λ_{C-BC} and λ_{BC} values and plot the resulting coverage probabilities below in Figure 2.6. Here we vary λ_{C-BC} and λ_{BC} independently from 1 to 10 with increments of 0.25 and simulated 10,000 data sets for each of the 1,600 parameter settings. In Figure 2.6 we also include horizontal 95% reference lines and LOWESS smooths fits to our coverage probabilities. Notice that the coverage probabilities for the LS method have significantly higher variability than the other methods.

From Table 2.2 and Figure 2.6 we see that for the proposed model the coverage probabilities are generally conservative for small sample size estimates and small numbers of deaths. But these coverages are substantially better than those of Brillinger's model. The LS method performs poorly compared to all other methods (CDF0, CDF1, CDF2, CP0, CP1, CP2 and SP0) and is often liberal. The "Clopper-Pearson" type method with two-dimensional mid-P correction (CP2) performs better than the CP0

Table 2.2. Coverage Probabilities for Proposed Model

λ_{C-BC}	λ_{BC}	CDF0	CDF1	CDF2	CP0	CP1	CP2	LS	SP0	Patch1	Patch2
1	1	99.90	99.90	99.89	99.90	99.90	99.89	99.01	99.90	27.19	53.87
1	3	99.40	99.40	99.40	99.40	99.40	99.40	98.14	99.38	35.06	38.43
1	5	98.95	98.95	98.40	98.95	98.95	98.40	97.84	98.59	36.34	36.72
1	7	98.67	98.67	98.60	98.67	98.67	98.60	99.21	98.21	36.09	36.14
1	9	98.18	98.18	97.78	98.18	98.18	97.78	99.39	97.79	36.97	36.99
3	1	99.13	99.13	99.45	99.45	99.45	99.45	95.33	99.45	2.93	38.24
3	3	98.55	98.55	98.56	99.01	99.01	98.18	89.53	98.89	4.76	9.23
3	5	97.78	97.78	97.47	98.07	98.07	97.47	91.38	98.07	5.03	5.68
3	7	97.60	97.42	97.12	97.72	97.54	97.12	92.46	97.30	5.00	5.13
3	9	96.97	96.97	96.97	97.22	97.22	97.22	92.41	96.78	5.22	5.22
5	1	97.31	98.11	98.87	99.00	99.00	98.41	97.83	98.68	0.40	37.42
5	3	97.13	97.13	97.16	98.00	98.00	97.40	91.02	98.00	0.77	5.78
5	5	97.08	97.08	96.83	97.85	97.85	96.52	88.98	97.25	0.64	1.24
5	7	96.44	96.26	96.03	96.45	96.27	96.04	90.76	96.44	0.60	0.70
5	9	95.97	95.97	96.39	96.46	96.46	96.40	91.32	96.14	0.64	0.64
7	1	98.03	97.98	98.33	98.76	98.71	98.58	99.12	98.33	0.07	37.11
7	3	97.55	97.55	97.19	97.99	97.76	96.86	92.59	97.18	0.07	5.18
7	5	96.27	96.61	96.52	96.90	96.87	96.52	90.69	96.98	0.09	0.72
7	7	96.27	96.27	96.15	96.85	96.85	96.33	90.94	96.53	0.04	0.08
7	9	96.13	96.24	95.99	96.89	96.89	96.09	91.53	95.76	0.06	0.06
9	1	97.15	97.15	98.23	98.23	98.23	97.83	99.25	97.88	0.00	36.58
9	3	97.21	97.21	97.22	97.93	97.93	97.53	92.86	97.11	0.02	4.48
9	5	96.17	96.17	96.26	96.56	96.56	96.46	91.56	96.17	0.02	0.70
9	7	96.32	96.32	95.87	96.72	96.72	96.05	92.01	95.80	0.00	0.14
9	9	95.58	95.58	95.52	95.91	95.91	95.75	92.10	95.76	0.00	0.00

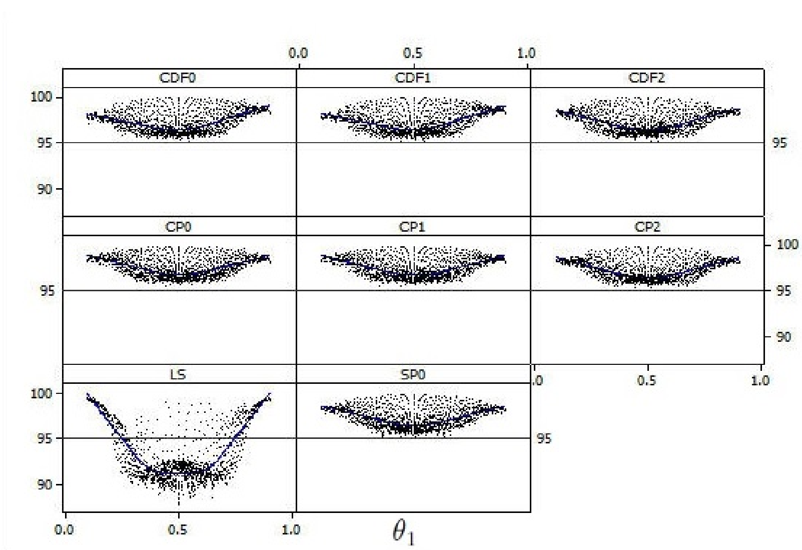


Figure 2.6. Plot of the coverage probabilities for methods CDF0, CDF1, CDF2, CP0, CP1, CP2, LS and SP0 verses θ_1.

and CP1 methods. The CDF pivot methods (CDF0, CDF1 and CDF2) and SP0 method perform more or less as well as the CP2 method in terms of coverage probability.

2.7. CONCLUSIONS

We developed various methods of confidence interval construction for vital rates. Our work was is motivated by Brillinger's seminal work on vital rates. Here the vital rates are modeled as the ratio of two Poisson random variables and therefore we have a random denominator. We discussed the Lexis diagrams for the Brillinger's model and introduced a novel Lexis diagram for our proposed model. Our investigation of the statistical properties of and inference for Brillinger's classical model shows that its maximum likelihood estimates do not have standard large sample properties and as such do not support the development of a large sample confidence interval for any parameter including the covariance λ_{BC}. Nonetheless, we developed a CDF pivot method with its saddlepoint approximated confidence interval, by applying a positivity condition to the Brillinger's classical model. For those methods we estimate our nuisance parameters using constrained MLE's and method of moment constrained estimates to obtain the confidence intervals. Simulation studies show that inference about correlation parameter in Brillinger's model appears to be difficult in small samples as evidenced by its very conservative confidence intervals.

We next proposed a new model which is also a limiting result of the Brillinger's model to construct confidence intervals for the expected value of the vital rate estimator D/P and investigated statistical properties and inference for this expected rate parameter. We developed a CDF pivot method with its saddlepoint approximation, a "Clopper-Pearson" type method, and a large sample method of confidence interval construction. Furthermore, we introduced a novel two-dimensional mid-P correction for the CDF pivot method and "Clopper-Pearson" type method. Simulation studies show that "Clopper-Pearson" type method with two-dimensional mid-P correction performed better than the "Clopper-Pearson" type method with no mid-P correction and "Clopper-Pearson" type method with one-dimensional mid-P correction. The CDF pivot method with no mid-P correction, one-dimensional mid-P correction and two-dimensional mid-P correction, "Clopper-Pearson" type method with two-dimensional mid-P correction and saddlepoint approximation to the CDF pivot method are not significantly different in terms of coverage probabilities. Their coverage probabilities are generally closer to the nominal 95% value than the coverage probabilities for the confidence interval methods used with Brillinger's model since we need to estimate more nuisance parameters for that model.

BIBLIOGRAPHY

[1] Agresti A, Gottard A. Comment: Randomized Confidence Intervals and the mid-P Approach. *Statistical Science* 2005, **20**(4), 367–371.

[2] Berry G, Armitage P. Mid-P confidence intervals: a brief review, *The Statistician* 1995; **44**(4), 417-423.

[3] Brillinger D. The Natural Variability of Vital Rates and Associated Statistics *Biometrics* December 1986; **42**, 693-734.

[4] Butler RW. *Saddlepoint Approximations with Applications.* Cambridge University Press: New York, 2007.

[5] Carstensen B. Age-Period-Cohort models for the Lexis diagram. *Statistics in Medicine* 2007; **26**, 3018-3045.

[6] Carstensen B, Plummer M, Laara E, Hills M. Epi: A Package for Statistical Analysis in Epidemiology. R package 2014 version 1.1.67.

[7] Casella G, Berger RL. *Statistical Inference*, 2nd ed., Duxbury Press: Pacific Grove CA, 2002.

[8] Clopper C, Pearson E. The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika* 1934; **26**, 404–413.

[9] Cohen GR, Yang S-Y. Mid-p confidence intervals for the Poisson expectation. *Statistics in Medicine* 1994; **13**, 2189-2203.

[10] DiCiccio T, Martin M, Stern S. Simple and accurate one-sided inference from signed roots of likelihood ratios. *Canad. J. Statist.* 2001; **29**, 67–76.

[11] DiCiccio T, Romano J. On bootstrap procedures for second-order accurate confidence limits in parametric models. *Statistica Sinica.* 1995; **5**, 141-160.

[12] Dobson AJ, Kuulasmaa K, Eberle E, Scherer J. Confidence intervals for weighted sums of Poisson parameters. *Statistics in Medicine* 1991; **10**, 457-462.

[13] Fay MP, Feuer EJ. Confidence intervals for directly standardized rates: A method based on the gamma distribution. *Statistics in Medicine* 1997; **16**, 791-801.

[14] Holgate P. Estimation for the bivariate Poisson distribution *Biometrika* 1964; **51(**1/2**)**, 241-245.

[15] Johnson N, Kotz S, Balakrishnan N. *Discrete Multivariate Distributions* New York: Wiley, 1997.

[16] Kawamura K. Direct calculation of maximum likelihood estimator for the bivariate Poisson distribution. *Kodai math J.* 1984; **7**, 211-221.

[17] Keiding N. Statistical inference in the Lexis diagram *Philosophical Transactions: Physical Sciences and Engineering* September 1990; **332**(1627), 487-509.

[18] Lee SMS, Young GA. Parametric bootstrapping with nuisance parameters. *Statistics & Probability Letters* 2005; **71**, 143–153.

[19] Lexis W. Einleitung in die Theorie der Bevölkerungsstatistik. Strassburg: Karl J. Trübner. 1875.

[20] Lugannani R, Rice SO. Saddlepoint approximations for the distribution of sums of independent random variables. *Advances in Applied Probability* 1980; **12**, 475-490.

[21] Molenberghs G, Kenward M, Aerts M, Verbeke G, Tsiatis A, Davidian M, Rizopoulos D. On Random Sample Size, Ignorability, Ancillarity, Completeness, Separability, and Degeneracy: Sequential Trials, Random Sample Sizes, and Missing Data *Stat Methods Med Res.* February 2014; **23**(1), 11-41.

[22] Ng HKT, Filardo G, Zheng G. Confidence interval estimating procedures for standardized incidence rates. *Computational Statistics and Data Analysis* 2008; **52**, 3501-3516.

[23] Paige RL, Trindade AA, Fernando PH. Saddlepoint-based bootstrap inference for quadratic estimating equations. *Scandinavian Journal of Statistics* 2009; **36**, 98-111.

[24] Paige RL, Trindade AA. Practical small sample inference for single lag subset autoregressive models. *Journal of Statistical Planning and Inference* 2008; **138**, 1934-1949.

[25] Rothman KJ, Greenland S, Lash TL. *Modern Epidemiology*, 3rd ed. Lippincott Williams & Wilkins: Philadelphia, 2008.

[26] Swift, MB. A simulation study comparing methods for calculating confidence intervals for directly standardized rates. *Computational Statistics and Data Analysis* 2010; **54**, 1103-1108.

[27] Swift, MB. Simple confidence intervals for standardized rates based on the approximate bootstrap method. *Statistics in Medicine* 1995; **14**, 1875-1888.

[28] Tiwari RC, Clegg LX, Zou Z. Efficient interval estimation for age-adjusted cancer rates. *Statistical Methods in Medical Research* 2006; **15**, 547-569.

[29] Vandeschrick C. The Lexis diagram, a misnomer. *Demographic Research* 2001; **4**(3), 97-124.

[30] Woodard M. *Epidemiology: Study Design and Data Analysis*, 2rd ed. Chapman & Hall: Boca Raton, 2004.

VITA

Pasan Manuranga Edirisinghe was born in 1980 in Colombo, Sri Lanka. After completing his schoolwork at Asoka Vidyalaya and Ananda College in Colombo, Sri Lanka in 1999, Pasan entered University of Colombo, Sri Lanka for his Bachelor of Science degree in 2001. He received a Bachelor of Science with a major in mathematics from University of Colombo, Sri Lanka in October 2005. During the following two years, he was employed as an instructor at University of Colombo, Sri Lanka. Subsequently, in 2007, he joined the Department of Statistics, University of Colombo, to pursue a post-graduate diploma in statistics which was obtained in 2009. In 2008, he joined Sri Lanka Institute of Information Technology as an instructor of mathematics and statistics. In August 2009, he entered the Graduate School of the Missouri University of Science and Technology and obtained his Masters degree in applied mathematics, with emphasis in statistics in 2012. He received his Ph.D. in Mathematics from the Missouri University of Science and Technology in December 2015.