

STATISTICAL METHODS FOR THE ANALYSIS OF CASE
SERIES DATA

by

Shen Ling

Submitted in partial fulfillment of the
requirements for the degree of
Master of Science

at

Dalhousie University
Halifax, Nova Scotia
August 2014

© Copyright by Shen Ling, 2014

Table of Contents

List of Tables	iv
List of Figures	v
Abstract	vi
Acknowledgements	vii
Chapter 1 Introduction	1
1.1 Background1
1.2 Development1
1.3 Advantages and Limitations3
1.4 Application5
1.5 Our problem5
Chapter 2 Likelihood and Theoretical Inference	7
2.1 Conditional Poisson Likelihood8
2.2 Unconditional Poisson Likelihood11
2.3 SCCS (Likelihood Conditioning on at least one event)14
2.4 Power17
2.5 Efficiency18
Chapter 3 Simulation	21
3.1 Estimation22
3.2 Power24
3.3 Mean Square Error30
Chapter 4 Random Effect Model	38
4.1 General Unconstrained Random Effect Model38
4.2 Random Effect for Conditional Poisson Model41

4.3 Random Effect Hierarchical Model	41
Chapter 5 Discussion	44
Bibliography	46
Appendix Appendix A R Code	48

List of Tables

3.1	Estimated different power of three methods30
3.2	MSE and Sample Size for $\beta = 0.5$ and $\lambda = 0.5$32
3.3	Bias and MSE comparison for $\beta = 0.5$ and $\lambda = 0.5$35

List of Figures

3.1	Parameters estimation of method 2, unconditional cohort method, under full cohort sampling (Figure A) and case series sampling (Figure B).23
3.2	Parameters estimation of method 3, self-controlled case series method, under case series sampling.23
3.3	Histogram of MLE of conditional Poisson for β under $\lambda = 0.2$ and $\lambda = 1$, here $\beta = 0.02, 0.2, 0.5$ in three rows.25
3.4	Histogram of MLE of unconditional Poisson for β under $\lambda = 0.2$ and $\lambda = 1$, here $\beta = 0.02, 0.2, 0.5$ in three rows.26
3.5	Histogram of MLE of SCCS method for β under $\lambda = 0.2$ and $\lambda = 1$, here $\beta = 0.02, 0.2, 0.5$ in three rows.27
3.6	Power of the conditional Poisson method (one-side, 5% nominal type I error risk) with sample size 100.28
3.7	Power of the conditional Poisson method (one-side, 5% nominal type I error risk) with sample size 1000.29
3.8	(Figure A) Compare different $\hat{\beta}$ under conditional Poisson and unconditional Poisson with fixed λ . (Figure B) compare $\hat{\beta}_u$ and $\hat{\beta}_c$ using a single data set.32
3.9	Compare different λ values with fixed β under three methods. Figures A, B, C represent conditional method, unconditional method and SCCS or method of conditional on $n_i \geq 1$33
3.10	Comparison of sampling designs with $\lambda = 0.2$. Particularly, SCCS (method 3) is not applicable under full cohort sampling.	.34

Abstract

This paper compares the three methods -conditional Poisson method, unconditional Poisson method and self-controlled case series (SCCS) method, based on the retrospective cohort study with full cohort and case series sampling designs, with particular emphasis on their assumptions, power, MSE, relative efficiency, and handling of confounding. The performance of the three methods is contrasted in a study investigating the causality between vaccination and rare adverse events seizures. And we extend these methods to random effect model.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to my supervisor Prof. Bruce Smith for the continuous support during my Master study and research, for his patience, motivation, enthusiasm, and immense knowledge. His generous guidance helped me in all the time of research and writing of this thesis. He is truly the best advisor and mentor beyond my imagination not only for my Master study but also for my life.

I would like to thank Prof. Hong Gu who has given me a lot of advice, encouragement and kindly support in my study during this year.

I also would like to thank the Department of Statistics and all my friends for their help and valuable opinion in the last year.

A special thank to my boyfriend and my mother and father. Words cannot express how grateful I am for all the sacrifices that they have made.

Chapter 1

Introduction

1.1 Background

Immunization is the most effective intervention in public health nowadays, especially in child health. Reductions over recent decades in the morbidity and mortality attributable to smallpox, measles, polio, diphtheria, whooping cough, and tetanus are eloquent reminders of this fact. However, no intervention is entirely without risk, even though an adverse event is very rare. Evaluation of vaccine safety with respect to rare adverse events is an important issue in epidemiology and public health [3]. In other words, we want to investigate an association between transient exposures and acute outcome events. In this thesis, we investigate the possible causality between vaccines and seizures (acute outcome events).

1.2 Development

Such studies have mostly used two main approaches, case-control study and cohort study. The cohort study starts with the putative cause of disease, and observes the occurrence of disease relative to the hypothesized causal agent, while the case-control study proceeds from documented disease and investigates possible causes of the disease [7]. In other words, the cohort study compares the rate of events (incidence rate) between vaccinated and unvaccinated individuals, while the case-control study compares the rate of vaccination between individuals with and without events. However, both of their original versions contain control selection bias, because comparison individuals cannot be chosen to be completely the same.

According to Fine and Chen [7], many potential sources of bias have been identified. One source of bias comes from the problem of ensuring that adverse events

are ascertained independent of vaccination history. Failure to control for this factor may lead to creation, or overestimation, of an association between a vaccine and an adverse event. Another source of bias comes from control selection, the confounding between the risk factor (vaccination) and outcome measure (adverse event) of interest.

Therefore, the modified methods for both cohort and case-control studies are developed which eliminate the control selection bias. As mentioned by Fine and Chen [7], the cohort study comparing incidence rates between vaccinated and unvaccinated individuals is equivalent to comparing between (age-specific) incidence rates in successive intervals before and after vaccination, in population with high vaccine coverage. Later, Farrington [15] developed self-controlled case series (SCCS) method using only data on cases. No separate controls are required as the cases act as their own controls. Each case's given observation time is divided into control and risk periods. The method found a relative incidence, that is, the incidence in risk periods relative to the incidence in control periods. Time-varying confounding factors such as age can be allowed for by dividing up the observation period further into age categories. An advantage of the method is that confounding factors that do not vary with time, such as genetics, location, socio-economic status, are controlled implicitly, see [5]. This kind of design combines the simplicity of cohort method and economy of case series, while reducing confounding factors [3]. This method also yields a direct estimate of the relative incidence of events.

The case-crossover method is the modified version of case-control study. Maclure [11] suggested the Mantel-Haenszel estimator and gives the definition. The term "crossover" is mainly used to describe experiments in which all subjects pass through both the exposure and unexposure phases [11]. In these studies, each subject serves as his or her own control— "the ultimate form of statistical adjustment" for confounding by constant subject characteristics. However, "crossover" needs not to imply randomization. It is an apt term for intermittent exposure to factors with transient effects. Maclure [12] gave an example about car collisions and cell phone calls to describe this method. In traditional case-control study, the control is different person at similar time, while in case-crossover design, the control is the same person at a

different time [3]. This method yields an estimate of the odds ratio of an event after vaccination. For rare events and short risk intervals, this is equivalent to the relative incidence. Both methods are data-intensive, involving large cohorts or careful selection and matching of controls. Moreover, Vines and Farrington [14] pointed out if the exposure distribution is stationary over time, the Mantel-Haenszel estimator for the odds ratio would be approximately asymptotically unbiased.

Depending on different sampling designs, the studies can be categorized as full cohort and case series or case only [6]. Full cohort includes all individuals with events or without events while case series only includes individuals with events. The case series sampling would be less expensive but with more restrictions and assumptions when constructing models or likelihoods. Hence, the methodologies are categorized by sampling and approach as cohort studies, case-cohort studies, case-control studies and case-crossover studies, etc.

From the methods of data collection, studies can be categorized as retrospective and prospective. In the case of a retrospective cohort study, the investigator collects data from past records and does not follow patients unlike a prospective study. In a nutshell, in retrospective cohort study, all the events - exposure, latent period, and subsequent outcome like development of disease have already occurred in the past [12]. If the data collection is prospective, one can avoid recall bias and so it is suitable for chronic disease infection. A prospective cohort study follows over time a group of similar individuals (cohorts) who differ with respect to certain factors under study, to determine how these factors affect rates of a certain outcome. It needs more time for longitudinal observation experiment and hence is suitable for short time disease or new disease infection.

1.3 Advantages and Limitations

Cohort studies involve following a sample from a population over time, either prospectively or retrospectively [1]. The analysis may be carried out using log-linear modeling after stratifying by other potential confounding factors, such as other vaccinations,

age, sex and region. In practice, most cohort studies are retrospective rather than prospective and they are often used with mass vaccination campaigns in order to compare incidence rate pre- and post-vaccination. The cohort studies have the advantages of enabling direct estimation of absolute and attributable risk as well as generally having the highest power for a given number of cases and allowing comparisons of background incidence between groups, such as regions, which cannot be carried out with the SCCS method or for matched co-variate in a case-control study.

The self-controlled case series (SCCS) method was developed in early 1990s to handle the type of data obtained from administrative databases and has since been widely applied to vaccine safety evaluation in a variety of settings. The SCCS method has power very similar to that of a cohort study, particularly, if the proportion of individuals vaccinated is high or if the risk period is short in comparison to the observation period. For rare events, the SCCS method is much quicker and cheaper than cohort method because it only requires information on cases. A further advantage of the SCCS method is the fact that it does not require a population to be defined from which the case arose.

In case-control studies, each case is matched with one or more controls according to potential confounding variables, such as age and region. The number of cases and controls with vaccination within specified intervals prior to onset are then compared using conditional logistic regression. For rare events, the odds ratio obtained from logistic regression approximates very closely to the relative incidence rate. Case-control studies are more proper for long-term effects, where effects of vaccination are usually more difficult to assess. This is because long term effects can only be studied by direct comparison of vaccinated and unvaccinated individuals. The problem with such studies is that it is likely that the unvaccinated group will be very different from the vaccinated group in many aspects, particularly, for high vaccine coverage population. Therefore, we only choose cohort studies and SCCS method as candidate approaches for our problem.

Methods involving only cases are attractive for three reasons, as mentioned by

Farrington [4]. First, they can usually be implemented using data extracted from readily available databases such as hospital admission data or other case reporting mechanisms. Second, they can produce results quickly, for example, in response to public concerns or media attention about vaccine methods. Third, they are usually cheaper to carry out than methods requiring explicit denominators or separate controls.

1.4 Application

As mentioned in Farrington [15], most applications of cohort studies and case series method so far are to vaccine safety and other areas of epidemiology. Navidi [12] applied case series method in studies of air pollution which is described as a bi-directional case-crossover method. Case-crossover analysis has been used to study air pollution in relation to daily mortality and injuries in racehorses. A pilot study indicated the feasibility of using the case-crossover method to investigate nonmedical triggers of visits to general practitioners in Denmark, which suggests further applications in health services research. In theory, the method can be used to study triggers of engineering failures and other acute events in inanimate systems.

1.5 Our problem

This thesis compares the three methods based on the retrospective cohort study with full cohort and case series sampling designs, with particular emphasis on their assumptions, relative efficiency, and handling of confounding. The performance of the three methods is contrasted in a study of seizures after vaccination on children with epilepsy.

We first review methods for investigating causality between vaccination and the rare adverse event, seizures, with data collection of everybody and only of cases, and then compare their underlying assumptions and performances. The three methodologies include a Poisson cohort model by conditioning on total occurrence of events,

unconditional Poisson cohort method and self-controlled case series method (a Poisson cohort model by conditioning on at least one occurrence of events). Since three methods have different likelihood and construction, the estimator of relative incidence rate would be slightly different. Then we compare three estimators from various perspective including power of test, asymptotic efficiency and mean square error to investigate their advantages and limitations.

The thesis is organized as follows. In Chapter 2, we give theoretical inference for the construction of these methodologies for fixed baseline risk, including likelihood, power of the test, bias of the relative incidence [13] and variance efficiency, to make comparisons. In Chapter 3, we carry out some simulation for these approaches where we present their asymptotic inference and make comparison with theoretical inference. In Chapter 4, we apply hierarchical model and Bayesian method to solve random effect (random baseline risk) mixed Poisson processes. In Chapter 5, we discuss at some level the assumptions, advantages and limitations of the methods, as well as areas for further research.

Chapter 2

Likelihood and Theoretical Inference

We assume that events, in our example seizures, arise according to a Poisson process, that there are two observation intervals, a low risk interval before vaccination, and a high risk interval after vaccination, and that the rates of the process are constant within each of the interval. That is, within each interval the events occur according to a homogeneous Poisson process, but the rates may differ across intervals.

To keep matters simple throughout this and the next chapter, we will be concerned only with situations where the baseline incidence rate is constant for all individuals, and does not vary with time or age. In this case the baseline event process is a homogeneous Poisson process, and in the absence of exposure to vaccination, is parameterized by the rate e^ϕ of the Poisson process. In the period after vaccination, events are assumed to occur according to a homogenous Poisson process with rate $e^{\phi+\beta}$. ϕ and β are real valued parameters to be estimated, and e^β is the relative event rate, so that β is the parameter of interest.

For a individual i , let N_{i0} denote the number of events occurring in the unexposed interval of length t_0 having rate e^ϕ , and N_{i1} denotes the number of events in the exposed interval of length t_1 , with incidence rate $e^\phi e^\beta$. Assuming that the event process is Poisson, it follows that N_{i0} and N_{i1} are independent, with

$$\begin{aligned} N_{i0} &\sim \text{Poisson}(e^\phi t_0) \\ N_{i1} &\sim \text{Poisson}(e^\phi e^\beta t_1) \end{aligned}$$

Furthermore, N_{i0} and N_{i1} are jointly sufficient statistics for ϕ and β , so we can focus our attention on these total numbers of events.

We consider three methods for estimating the rate parameters – a conditional method which conditions on a subject’s total number of events in the two intervals, the SCCS method which conditions on a subject having at least one event over both intervals, and an unconditional method which does not constrain the underlying Poisson processes. The likelihood for the conditional and unconditional methods are described, for example, in guttorp [8], while the SCCS likelihood is set down in Farrington [15].

We discuss the likelihoods, estimation of relative incidence estimation, LR test, power and efficiency of these likelihoods in the following section.

2.1 Conditional Poisson Likelihood

In the conditional Poisson model we condition on a subject’s total number of events over the two intervals. From the Poisson assumption, it follows that the total number of events $N_i = N_{i0} + N_{i1}$ has a Poisson distribution with mean $e^\phi e^\beta t_1 + e^\phi t_0$, and that the conditional distribution of the number of events N_{i1} in the exposed interval given the total N_i is binomial with parameters N_i and $\pi = \frac{t_1 e^\beta}{t_0 + t_1 e^\beta}$. Note that the baseline hazard rate parameter ϕ disappears from this conditional probability, so that after conditioning, we are left with a one parameter model.

It follows that if M individuals are observed over two periods of durations t_1 and t_0 , then the conditional likelihood, given the observed totals $n_i, i = 1, 2, \dots, M$, is given by

$$\begin{aligned} L(\beta) &\propto \prod_{i=1}^M \left(\frac{t_0}{t_0 + t_1 e^\beta} \right)^{n_{i0}} \left(\frac{t_1 e^\beta}{t_0 + t_1 e^\beta} \right)^{n_{i1}} \\ &= \left(\frac{t_0}{t_0 + t_1 e^\beta} \right)^{n_0} \left(\frac{t_1 e^\beta}{t_0 + t_1 e^\beta} \right)^{n_1}, \end{aligned}$$

which is a function of the scalar parameter β . Here $n_0 = \sum_{i=1}^M n_{i0}$, $n_1 = \sum_{i=1}^M n_{i1}$, and $n = n_0 + n_1$.

Note that if the baseline parameter ϕ is replaced by subject specific parameters (ϕ_i for subject i), this method allows for variation between individuals in their baseline susceptibility to the events under consideration. The multiplicative effect assumption allows the "between individual" variation to be eliminated by conditioning, and the vaccine effect to be estimated "within individuals".

The log-likelihood is:

$$\begin{aligned} l(\beta) &\propto n_0 \log \left(\frac{t_0}{t_0 + t_1 e^\beta} \right) + n_1 \log \left(\frac{t_1 e^\beta}{t_0 + t_1 e^\beta} \right) \\ &\propto n_1 \beta - n \log (t_0 + e^\beta t_1). \end{aligned}$$

During the risk period, the baseline incidence rate of seizure is increased by a factor e^β , which represents the relative incidence rate. When testing, the null hypothesis of interest is $H_0 : \beta = 0$, with the one sided alternative $\beta > 0$ representing an increased incidence rate, or $\beta < 0$ for an decreased incidence rate, after vaccination.

Parameterizing in terms of the relative seizure rate, $\lambda = e^\beta$, the conditional log likelihood is

$$\begin{aligned} l(\lambda) &\propto n_0 \log \left(\frac{t_0}{t_0 + \lambda t_1} \right) + n_1 \log \left(\frac{\lambda t_1}{t_0 + \lambda t_1} \right) \\ &\propto n_1 \log (\lambda t_1) - n \log (t_0 + \lambda t_1). \end{aligned}$$

The maximum likelihood estimator is $\hat{\lambda} = \frac{t_0 n_1}{n_0 t_1}$, which has the natural interpretation of the number of events per unit time in the risk interval divided by the number of events per unit time in the baseline interval. By the invariance property of the maximum likelihood estimator, the MLE of β is $\log(\hat{\lambda})$.

In the limiting cases where there are no events in one of the two intervals ($n_1 = 0$ exclusive or $n_0 = 0$) the relative incidence is estimated as 0 or ∞ . The case where no events are observed ($n_1 = n_0 = 0$) is non-informative. In this case, interpreting $\frac{0}{0} = 1$, the estimated relative rate is 1.

If we have full cohort sampling, we can generalize the model by adding an indicator variable, setting $z_i = 1$ if individual i was vaccinated prior to the risk period and 0 otherwise. In this case the conditional Poisson log likelihood has the form of a binomial log likelihood, as:

$$\begin{aligned} l(\beta) &\propto \sum_{i=1}^M \left[n_{i0} \log \left(\frac{t_0}{t_0 + t_1 e^{\beta z_i}} \right) + n_{i1} \log \left(\frac{t_1 e^{\beta z_i}}{t_0 + t_1 e^{\beta z_i}} \right) \right] \\ &\propto \sum [n_{i1} \beta z_i - n_i \log(t_0 + t_1 e^{\beta z_i})] \\ &= -(n_v \log(t_0 + t_1 e^{\beta}) + n_u \log(t_0 + t_1)) + \sum n_{i1} z_i \beta, \end{aligned}$$

where n_v denotes number of events in vaccinated individuals and n_u denotes number of events in unvaccinated individuals.

The first derivative of the log likelihood is:

$$\frac{\partial l}{\partial \beta} = -\frac{n_v t_1 e^{\beta}}{t_0 + t_1 e^{\beta}} + \sum n_{i1} z_i$$

and the second derivative is:

$$\frac{\partial^2 l}{\partial \beta^2} = -\frac{n_v t_1 t_0 e^{\beta}}{(t_0 + t_1 e^{\beta})^2}$$

Let $\widehat{\beta}_c$ denote the conditional maximum likelihood estimator of β with full cohort sampling, The asymptotic variance is

$$Var(\widehat{\beta}_c) = -E \left[\frac{\partial^2 l}{\partial \beta^2} \right] = \frac{(t_0 + t_1 e^{\beta})^2}{n_v t_1 t_0 e^{\beta}}. \quad (2.1)$$

We will use the variance to compute relative efficiency later in this chapter.

2.2 Unconditional Poisson Likelihood

Unconditional Poisson likelihood would be slightly more complicated than conditional one mentioned above, because it cannot cancel out the baseline hazard e^ϕ . That means we have two parameters β and ϕ .

A sample of M individuals are observed over two periods of durations t_0 and t_1 . We consider the more general scenario which allows both vaccinated and unvaccinated individuals, again by setting $z_i = 1$ if individual i was vaccinated prior to the risk period, and 0 otherwise. That is, each individual is observed over a baseline period of length t_0 with hazard rate e^ϕ , and then over an exposed period of length t_1 with hazard rate e^ϕ if $z_i = 0$ or hazard rate $e^{\phi+\beta}$ if $z_i = 1$.

Suppose that V individuals are vaccinated at the start of exposed period, and are consequently at increased risk during that period. Once again, N_{i0} and N_{i1} denote the number of events in the two periods.

Now consider estimation in this unconditional cohort model, with vector parameter $\boldsymbol{\theta} = [\phi, \beta]^T$. The parameter of primary interest is β , while ϕ is just a nuisance parameter. The parameterisation is different from the conditional model, where the nuisance parameter ϕ was eliminated by conditioning. For the i th individual, the distributions are:

$$\begin{aligned} N_{i0} &\sim \text{Poisson}(e^\phi t_0) \\ N_{i1} &\sim \text{Poisson}(e^\phi e^{\beta z_i} t_1). \end{aligned}$$

Let $\alpha_{1i} = e^\phi e^{\beta z_i} t_1$ and $\alpha_0 = e^\phi t_0$. Then the unconditional Poisson likelihood is:

$$\begin{aligned}
L(\phi, \beta) &\propto \prod_{i=1}^M \left(\frac{e^{-\alpha_0} \alpha_0^{n_{i0}}}{n_{i0}!} \right) \left(\frac{e^{-\alpha_{1i}} \alpha_{1i}^{n_{i1}}}{n_{i1}!} \right) \\
&= \left(\frac{e^{-\sum(\alpha_0 + \alpha_{1i})} \alpha_0^{\sum n_{i0}} \prod_1^M \alpha_{1i}^{n_{i1}}}{\prod_1^M n_{i0}! n_{i1}!} \right).
\end{aligned}$$

Then log-likelihood is:

$$l(\phi, \beta) \propto \sum_{i=1}^M [-(\alpha_0 + \alpha_{1i}) + n_{i0} \log(\alpha_0) + n_{i1} \log(\alpha_{1i})],$$

so that

$$\begin{aligned}
l(\beta, \phi) &\propto \sum [-e^\phi (t_0 + t_1 e^{\beta z_i}) + n_{i0}(\phi + \log(t_0)) + n_{i1}(\phi + \log(t_1) + \beta z_i)] \\
&\propto \sum [-e^\phi (t_0 + t_1 e^{\beta z_i}) + n_{i1} \beta z_i + n_i \phi].
\end{aligned}$$

If all individuals are vaccinated prior to the exposed period ($z_i = 1, i = 1, \dots, M$), then the log likelihood would be:

$$l(\beta, \phi) = -M e^\phi (t_0 + t_1 e^\beta) + n_1 \beta + n \phi,$$

where n_1 is the total number of events in the risk period and n is the total number of events over both periods, summed over all subjects. In this case, evaluating first derivatives with respect to β and λ , the likelihood equations are

$$\begin{aligned}
\frac{\partial l}{\partial \beta} &= -M e^\phi t_1 e^\beta + n_1 = 0 \\
\frac{\partial l}{\partial \phi} &= -M e^\phi (t_0 + t_1 e^\beta) + n = 0
\end{aligned}$$

$$\begin{aligned} \implies \frac{t_1 e^\beta}{t_0 + t_1 e^\beta} &= \frac{n_1}{n} \\ \implies e^{\hat{\beta}} &= \frac{t_0 n_1}{n_0 t_1} \end{aligned}$$

giving $e^{\hat{\phi}} = n_0/Mt_0$

The estimators again have a natural interpretation. The estimate of the baseline rate is $e^{\hat{\phi}}$, which is the average number of events per unit time in the baseline interval, and the estimated relative incidence rate is $e^{\hat{\beta}}$, which is the average number of events per unit time in the exposed interval divided by the average number of events per unit time in the unexposed interval.

We see that if all individuals are vaccinated, the unconditional Poisson maximum likelihood estimator of the relative rate is the same as the binomial (or conditional Poisson) maximum likelihood estimator.

For the generalized unconditional model with full cohort sampling, the information matrix for the parameter $\boldsymbol{\theta} = [\phi, \beta]^T$ is:

$$-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = \begin{bmatrix} \sum e^\phi (t_0 + t_1 e^{\beta z_i}) & \sum e^\phi t_1 e^{\beta z_i} z_i \\ \sum e^\phi t_1 e^{\beta z_i} z_i & \sum e^\phi t_1 e^{\beta z_i} z_i^2 \end{bmatrix}$$

Since z_i is an indicator, $z_i^2 = z_i$. Suppose the sample size is M and V is the number of vaccinated individuals. Then M-V is the number of unvaccinated individuals, and

$$-\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} = \begin{bmatrix} e^\phi (t_0 M + t_1 (e^\beta V + M - V)) & e^\phi t_1 e^\beta V \\ e^\phi t_1 e^\beta V & e^\phi t_1 e^\beta V \end{bmatrix}$$

After simplification:

$$I(\theta) = e^{\phi} t_1 e^{\beta} V \begin{bmatrix} k & 1 \\ 1 & 1 \end{bmatrix}$$

$$\text{where } k = \frac{t_0 M + t_1 (M - V)}{t_1 e^{\beta} V} + 1.$$

Let $\hat{\beta}_u$ denote the unconstrained maximum likelihood estimator of β . Its large sample variance is given by

$$\text{Var}(\hat{\beta}_u) = \frac{1}{e^{\phi} t_1 e^{\beta} V} \left(1 + \frac{P t_1 e^{\beta}}{t_0 + (1 - P) t_1} \right), \quad (2.2)$$

where $P = \frac{V}{M}$ is the proportion of individuals vaccinated.

For the unconditional model with all the individuals vaccinated. $P = 1$ and $V = M$. The large sample variance of β is given by

$$\text{Var}(\hat{\beta}_u) = \frac{1}{e^{\phi} t_1 e^{\beta} M} \left(1 + \frac{t_1}{t_0} e^{\beta} \right),$$

2.3 SCCS (Likelihood Conditioning on at least one event)

Self-controlled case series method has two key words. One is self-controlled which means the vaccine effect is estimated “within individuals”. Another one is case series which means a subject having at least one event over both intervals.

This approach described in this section is derived from a cohort study by conditioning on occurrence of at least one event and on the individual’s vaccination history

during a specified observation period. Conditioning on occurrence of at least one event means the approach is only suitable for case series sampling. This approach will have the most complicated construction of likelihood and non-explicit solution, because any constant covariates cannot be eliminated and there exists nuisance parameter ϕ as well.

The construction of likelihood combines both truncated Poisson likelihood and binomial likelihood. Here let $\mathbf{n}_i = [n_{i1}, n_{i0}]^T$ denote the number of events in each of the two intervals, and $n_{i.} = n_{i1} + n_{i0}$ denotes the total number of events for the i th individual in the two intervals. Then

$$P(\mathbf{n}_i | n_{i.} \geq 1) = \frac{P(\mathbf{n}_i, n_{i.} \geq 1)}{P(n_{i.} \geq 1)} = \frac{P(\mathbf{n}_i)}{P(n_{i.} \geq 1)}. \quad (2.3)$$

If we know $\mathbf{n}_i = [n_{i1}, n_{i0}]^T$, that is, the number of events happened in each interval, we will know $n_{i.}$ and whether $n_{i.} \geq 1$ or not.

$$P(n_{i.} | n_{i.} \geq 1) P(\mathbf{n}_i | n_{i.}) = \frac{P(n_{i.}, n_{i.} \geq 1)}{P(n_{i.} \geq 1)} \times \frac{P(\mathbf{n}_i, n_{i.})}{P(n_{i.})} \quad (2.4)$$

$$= \frac{P(n_{i.})}{P(n_{i.} \geq 1)} \times \frac{P(\mathbf{n}_i)}{P(n_{i.})} \quad (2.5)$$

$$= \frac{P(\mathbf{n}_i)}{P(n_{i.} \geq 1)}. \quad (2.6)$$

Combining (2.3) and (2.6), it follows that

$$P(\mathbf{n}_i | n_{i.} \geq 1) = P(n_{i.} | n_{i.} \geq 1) P(\mathbf{n}_i | n_{i.}),$$

Here $P(n_{i.} | n_{i.} \geq 1)$ is a truncated Poisson distribution with incidence rate parameter $\lambda_1 + \lambda_0$ and $P(\mathbf{n}_i | n_{i.})$ is the binomial distribution with index $n_{i.} (\geq 1)$ and parameter $\left(\frac{\lambda_1}{\lambda_1 + \lambda_0} \right)$.

Therefore, the construction of the likelihood conditional on a subject having at least one event is the product of a truncated Poisson probability and a binomial probability, as follows.

$$\begin{aligned}
N_{i1} &\sim \text{Poisson}(\lambda_1), \\
N_{i0} &\sim \text{Poisson}(\lambda_0), \\
N_i &= N_{i1} + N_{i0} \sim \text{Poisson}(\lambda_1 + \lambda_0).
\end{aligned}$$

The truncated Poisson distribution is:

$$P(n_i | n_i. \geq 1) = \frac{(\lambda_1 + \lambda_0)^{n_i} e^{-(\lambda_1 + \lambda_0)}}{n_i!} / (1 - e^{-(\lambda_1 + \lambda_0)}), n_i. = 1, 2, \dots$$

and the binomial distribution is:

$$P(\mathbf{n}_i | n_i.) = \binom{n_i.}{n_{i1}} \left(\frac{\lambda_1}{\lambda_1 + \lambda_0} \right)^{n_{i1}} \left(\frac{\lambda_0}{\lambda_1 + \lambda_0} \right)^{n_{i0}}, n_{i1} = 0, 1, \dots, n_i.$$

so that the likelihood, conditional on at least one event, is:

$$\begin{aligned}
P(\mathbf{n}_i | n_i. \geq 1) &= P(n_i. | n_i. \geq 1) P(\mathbf{n}_i | n_i.) \\
&= \left(\frac{\lambda_1^{n_{i1}} e^{-\lambda_1}}{n_{i1}!} \right) \left(\frac{\lambda_0^{n_{i0}} e^{-\lambda_0}}{n_{i0}!} \right) / (1 - e^{-(\lambda_1 + \lambda_0)}) \\
&= \frac{P(n_{i1}) P(n_{i0})}{1 - P(n_i. = 0)},
\end{aligned}$$

where $\lambda_1 = e^\phi e^\beta t_1$, $\lambda_0 = e^\phi t_0$ and $\lambda_1 + \lambda_0 = e^\phi e^\beta t_1 + e^\phi t_0$, $\left(\frac{\lambda_1}{\lambda_1 + \lambda_0} \right) = \frac{t_1 e^\beta}{t_0 + t_1 e^\beta}$.

It is seen that the likelihood, conditional on $n_i. \geq 1$, is the product of two Poisson probabilities, divided by one minus another Poisson probability. There is no explicit solution for the maximum likelihood estimator, and so we use the general purpose optimizer “optim” in R to approximate the MLE and the information matrix, as described in the next chapter.

2.4 Power

Power is the probability of correctly rejecting the null hypothesis when the alternative hypothesis is true, that is, the ability of a test to detect an effect, if the effect actually exists. Equivalently, the power is one minus the probability of a type II error. For a fixed sample size, it is usually impossible to make both types of error probabilities arbitrarily small. For instance, if we minimize the probability of type I error by restricting $\alpha = 0$, which means that we never reject H_0 a.s.. This implies that $\text{prob}(\text{Type II error})=1$. Similarly, if we let $\text{prob}(\text{Type II error})=0$, we get $\text{prob}(\text{Type I error})=1$. Therefore, in searching for a good test, it is common to control the Type I error probability at a specified level. Within this class of tests, we then search for tests with minimum Type II error probability, or maximum power.

Among the three models considered, it is only possible to set down an explicit formula for the power for the conditional Poisson model, in which case, the variance was given by (2.1). Here we use generalized model (there is indicator z_i).

The regularity conditions underlying the asymptotic normality of the maximum likelihood estimator are satisfied in this case, so that

$$\hat{\beta} \sim AN \left(\beta, \frac{(t_0 + e^\beta t_1)^2}{n_v t_0 t_1 e^\beta} \right)$$

where the notation “AN” denotes asymptotic normality.

Under the null and alternative hypotheses:

$$\begin{aligned} \text{Under } H_0, \hat{\beta} &\sim N(\beta_0, \sigma_0^2), & \rightarrow \frac{\beta - \beta_0}{\sigma_0} &\sim N(0, 1) \\ \text{Under } H_1 : \hat{\beta} &\sim N(\beta_1, \sigma_1^2), & \rightarrow \frac{\beta - \beta_1}{\sigma_1} &\sim N(0, 1) \end{aligned}$$

$$\text{where } \beta_0 = 0, \sigma_0^2 = \frac{(t_0 + t_1)^2}{t_0 t_1 n_v}; \beta_1 = \beta, \sigma_1^2 = \frac{(t_0 + e^\beta t_1)^2}{n_v t_0 t_1 e^\beta}$$

In the seizure example, the alternative hypothesis of interest is that vaccination increases the seizure rate, or

$$H_0 : \beta = 0$$

$$H_1 : \beta > 0$$

The test statistic is

$$Z = \frac{\hat{\beta} - \beta_0}{\sigma_0}$$

and the usual test rejects H_0 if $Z \geq Z_{1-\alpha} \iff \hat{\beta} \geq \beta_0 + Z_{1-\alpha}\sigma_0$, where $Z_{1-\alpha}$ is the $(1 - \alpha)$ -quantile of standard normal distribution, and α is the asymptotic significance level of the test.

The power of this test is given by

$$\begin{aligned} \text{Power}(\beta_1) &= P_{H_1}(\text{Reject } H_0) && (2.7) \\ &= P_{\beta_1}(T \geq Z_{1-\alpha}) = P_{\beta_1}(\hat{\beta} \geq \beta_0 + Z_{1-\alpha}\sigma_0) \\ &= P_{\beta_1}\left(\frac{\hat{\beta} - \beta_1}{\sigma_1} \geq \frac{\beta_0 + Z_{1-\alpha}\sigma_0 - \beta_1}{\sigma_1}\right) \\ &= P\left(Z \leq \frac{(\beta_1 - \beta_0) - Z_{1-\alpha}\sigma_0}{\sigma_1}\right) \\ &= P(Z \leq Z_{\beta_1}) \end{aligned}$$

where Z has a standard normal distribution, and $Z_{\beta_1} = \frac{(\beta_1 - \beta_0) - Z_{1-\alpha}\sigma_0}{\sigma_1}$.

For the unconditional Poisson and Self-Controlled Case Series methods, we cannot write down an explicit form for the power calculation. In chapter 3 we use simulation to approximate the power in those cases.

2.5 Efficiency

The variance of conditional Poisson estimate is close to that of the unconditional Poisson estimate when the risk period is short in comparison with the observation

period, or when the population vaccine coverage in the observation period is high. In particular, if vaccine coverage is 100%, the full cohort sampling and case series sampling have the same efficiency [5].

We are interested in the relative efficiency of estimate of β using the conditional method compared with the unconditional method, when using generalized forms, and we now derive the asymptotic relative efficiency in a simple case.

The conditional likelihood, which had a binomial form, can be thought of as a case series (case only) Poisson likelihood, since individuals without any events contribute $\binom{0}{0}\pi^0(1-\pi)^0 = 1$, and so provide no information ($\log(1) = 0$) to the log-likelihood. In other words, the comparison of the conditional method and unconditional method is just the comparison of case series and cohort sampling. Therefore, as compared with full cohort sampling, case series sampling might lose some power by conditioning (noncases are ignored), and its asymptotic efficiency might be reduced.

For the generalized model with indicator of vaccination z_i , the unvaccinated individuals are valuable because they contribute information about the baseline rate, although they contribute nothing on the relative incidence associated with vaccination.

The variances of the conditional and unconditional MLE's of β are given by (2.1) and (2.2). Replacing n_v by its expected value $E[n_v] = e^\phi(t_0 + t_1 e^\beta)V$ in the conditional variance formula (2.1), the asymptotic (as M increases) relative efficiency [10] of the conditional $\hat{\beta}_c$ relative to the unconditional $\hat{\beta}_u$, when using generalized models, is given by

$$\widetilde{eff} = \frac{Var(\hat{\beta}_u)}{\widetilde{Var}(\hat{\beta}_c)} = \left(1 + \frac{Pt_1 e^\beta}{t_0 + (1-P)t_1}\right) / \left(1 + \frac{t_1 e^\beta}{t_0}\right).$$

here the hat of $\widetilde{Var}(\hat{\beta}_c)$ means approximation, because n_v was replaced by its expected value $E[n_v] = e^\phi(t_0 + t_1 e^\beta)V$ in the conditional variance formula (2.1). Because $0 < P \leq 1$, it's straight forward to see $\widetilde{eff} \leq 1$.

When the proportion of vaccinated individuals P increases, the number of vaccinated individuals V becomes larger, the difference between conditional Poisson and unconditional Poisson becomes smaller and smaller, that is, the decline in asymptotic relative efficiency for conditional Poisson is small, as little information is lost if individuals without any events are ignored.

Therefore, this suggests that the conditional Poisson method retains high efficiency relative to the unconditional Poisson analysis as well as economy in sampling, at least in large samples, under conditions which commonly apply to studies of vaccine safety.

Then we consider the situation that all the individuals take vaccination, that is $P = 1$. So the asymptotic relative efficiency of the conditional $\widehat{\beta}_c$ relative to the unconditional $\widehat{\beta}_u$

$$\begin{aligned} \widetilde{eff} &= \frac{Var(\widehat{\beta}_u)}{\widetilde{Var}(\widehat{\beta}_c)} = \left(1 + \frac{t_1}{t_0} e^\beta\right) / \left(1 + \frac{t_1}{t_0} e^\beta\right) \\ &= 1 \end{aligned}$$

We see that if all individuals are vaccinated, the unconditional Poisson variance of the relative rate is the same as the binomial (or conditional Poisson) variance.

Since the complicated likelihood of self-controlled case series, we cannot get explicit information matrix theoretically. However, we apply numerical algorithm, R optimization, which returns the maximum likelihood estimators and hessian matrix to approximate the information matrix.

Chapter 3

Simulation

A simulation study was conducted to examine the performance of the three estimation procedures under both full cohort sampling and case series sampling, with the focus being on the relative incidence rate parameter β .

We derived the explicit form of the MLE $\hat{\beta}$ for conditional Poisson in Chapter 2. No explicit solution exists for the other two methods, unconditional method and SCCS method, and so we maximized the likelihood using numerical optimization, using the R function “optim”, which returns not only an estimate of the MLE, but also, the sample Fisher information matrix, as the Hessian. The code is given in the appendix.

We also obtained the power function, the bias, MSE, and relative efficiency of different estimators, under the two different sampling scenarios. These criteria provide different perspectives, and give insight into the performance of the three methods.

When it was not possible to calculate the quantities of interest explicitly, they were simulated using multiple simulation batches. In all cases presented we used 1000 batches.

Prior to the simulation, our expectation is that the SCCS method should have the best behavior with case sampling, provided that the probability of a non-case is moderately large, while the conditional and unconditional methods should perform reasonably well even with case sampling, provided that the probability of a non-case is close to 0. Conditioning removes the baseline parameter ϕ , and the simulation results will also provide information on the advantage or disadvantage of conditioning, if any.

In all simulations the number of batches was fixed at 1000, and we used a moderate sample size of $M = 100$ subjects. We looked at the scenario where the probability of a non-case was non-negligible ($\phi = 0$, $\beta = .2$, $t_0 = 1$, $t_1 = .5$, giving $P(\text{non-case}) = 20\%$, relative incidence $\rho = 1.22$).

3.1 Estimation

To begin, we consider the scenario where non-cases are very rare, in which the data sets from case and cohort sampling designs are very similar, often the same.

We used parameters $\phi = 0$, $\beta = .5$, $t_0 = 4$, and $t_1 = 2$, $P(\text{non-case}) = 0.068\%$, relative incidence $\rho = 1.65$.

Figure 3.1 shows contours of the joint likelihood for β and ϕ for a single simulation batch, using the unconditional Poisson method under full cohort and case series sampling designs. The contours appear elliptically symmetric, as would be expected from the asymptotic normal distribution of the MLE.

The contours under full cohort and case series sampling designs appear the same, which is to be expected, as all observations are expected to be cases. The contours are centred near $\hat{\beta} = 0.5$ and $\hat{\phi} = 0$, and demonstrate that the unconditional method is able to provide relatively accurate estimation.

Similarly, Figure 3.2 shows contours of the joint likelihood for β and ϕ for a single simulation batch, using the SCCS method under case series sampling designs. The contours appear elliptically symmetric, as would be expected from the asymptotic normal distribution of the MLE, and are centered near the true parameter $(0, .5)$.

Figures 3.3, 3.4, 3.5 show histograms of the maximum likelihood estimators for the three methods with case series sampling. Here in simulation, we choose $\beta = 0.02, 0.2, 0.5$ and $\lambda = e^\phi = 0.2, 1$ respectively, and then plot six histograms to see the variance and accuracy of estimator $\hat{\beta}$. When $\lambda = 1$, frequent baseline incidence, the variance of estimation becomes small. When $\lambda = 0.2$ the baseline incidence rate is

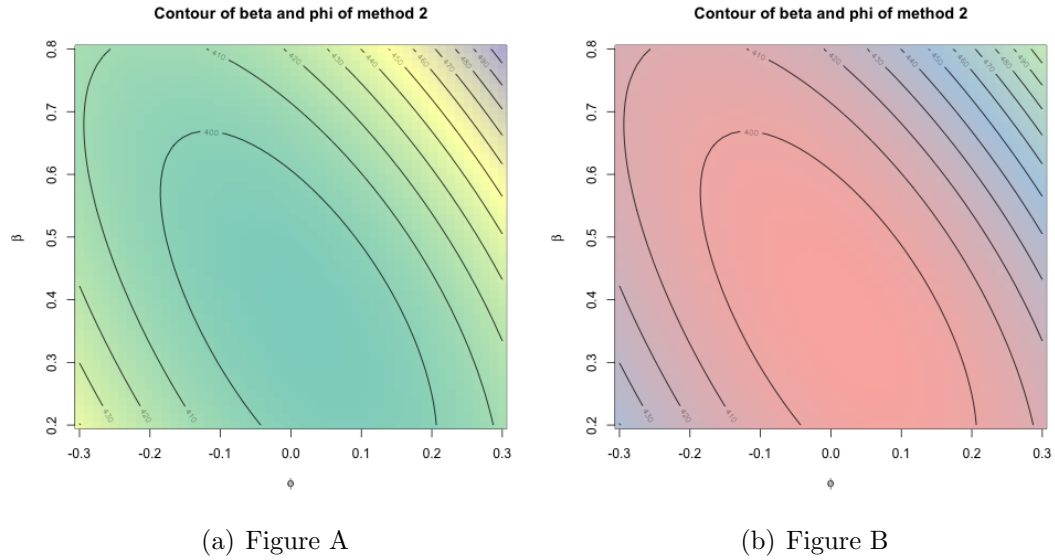


Figure 3.1: Parameters estimation of method 2, unconditional cohort method, under full cohort sampling (Figure A) and case series sampling (Figure B).

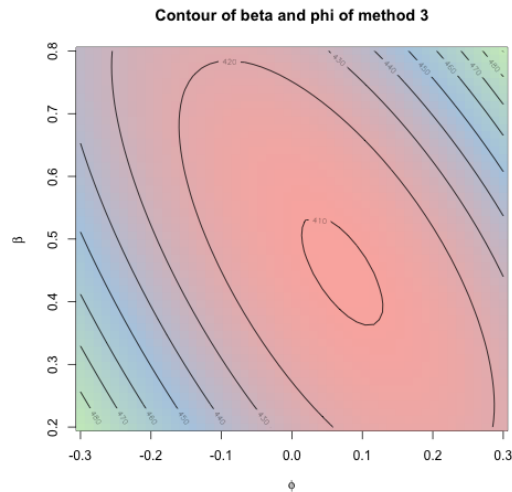


Figure 3.2: Parameters estimation of method 3, self-controlled case series method, under case series sampling.

low, resulting in a large variance of the estimators. Based on these plots we see no obvious difference of estimation among the three methods.

3.2 Power

For the unconditional and SCCS methods, we used the likelihood ratio test to test for $\beta \leq 0$ against the alternative $\beta > 0$. In this case, the critical region is the upper 5% point of the χ^2 distribution with one degree of freedom, and the empirical power is the proportion of simulation batches rejecting the null hypothesis (when twice the log likelihood ratio exceeds the critical value).

Figure 3.6 presents the power of the conditional Poisson method with sample size 100. Since this method has an explicit solution, power can be written down as an explicit expression as proved in Section 2.4. We calculate the power function from equation (2.7). We use $\beta_0 = 0$ as null hypothesis and choose a collection of β_1 's as alternative hypothesis to see the change of power.

We observed that the power of the unconditional method under full cohort sampling was very similar to the conditional Poisson method. So here we just put conditional power figure in the thesis. We did not compare conditional Poisson and unconditional Poisson under case series sampling, because they have the same power, as noted in Section 2.2 (the same estimator) and Section 2.5 (the same variance).

Figure 3.7 presents the power of conditional Poisson under full cohort sampling with sample size 1000. From two different sample size 100 and 1000, it demonstrates that variance increases when sample size becomes smaller, as expected. Moreover, when β_1 increases and becomes large enough, the power will increase. Additionally, the variance on both two extreme sides 0 and 1 are very small, because when β_1 is too small or too large, we get no power or full power, while in the middle, the variance becomes larger.

Tables 3.1 shows the mean power from three methods with sample size 100, estimated using 1000 repetitions. Since the fixed parameter in the Table is $\lambda = 1$ with

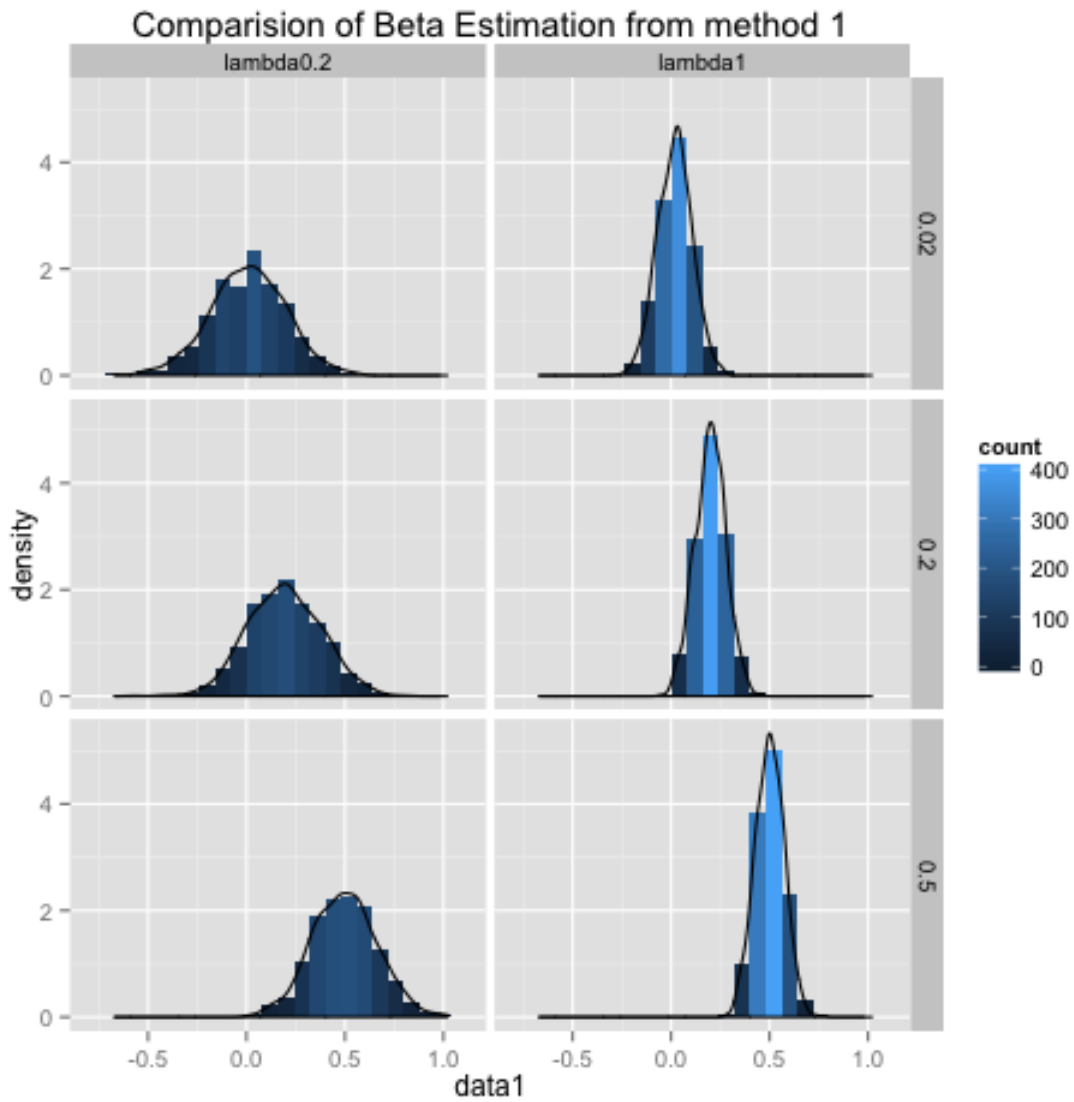


Figure 3.3: Histogram of MLE of conditional Poisson for β under $\lambda = 0.2$ and $\lambda = 1$, here $\beta = 0.02, 0.2, 0.5$ in three rows.

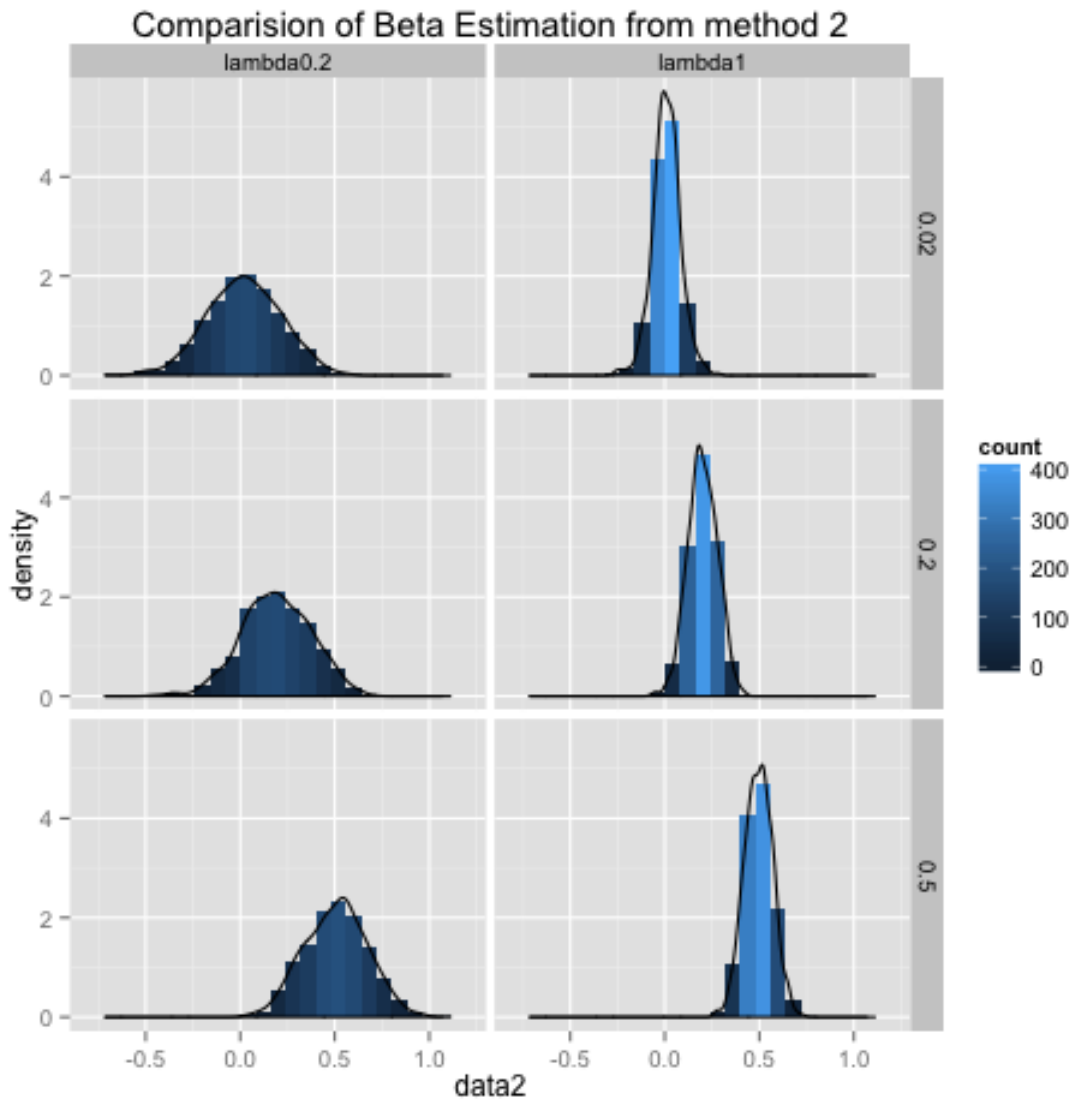


Figure 3.4: Histogram of MLE of unconditional Poisson for β under $\lambda = 0.2$ and $\lambda = 1$, here $\beta = 0.02, 0.2, 0.5$ in three rows.

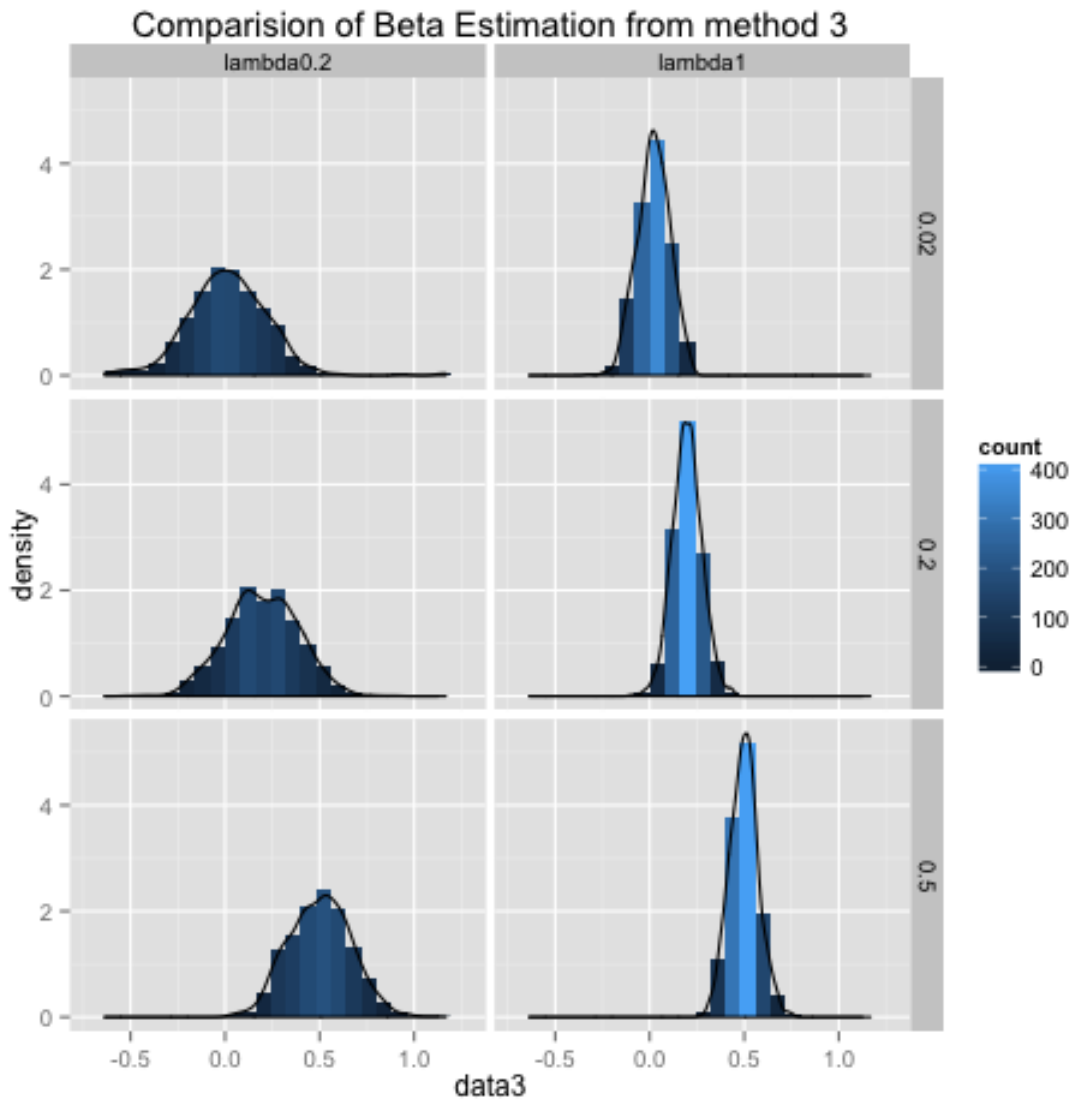


Figure 3.5: Histogram of MLE of SCCS method for β under $\lambda = 0.2$ and $\lambda = 1$, here $\beta = 0.02, 0.2, 0.5$ in three rows.

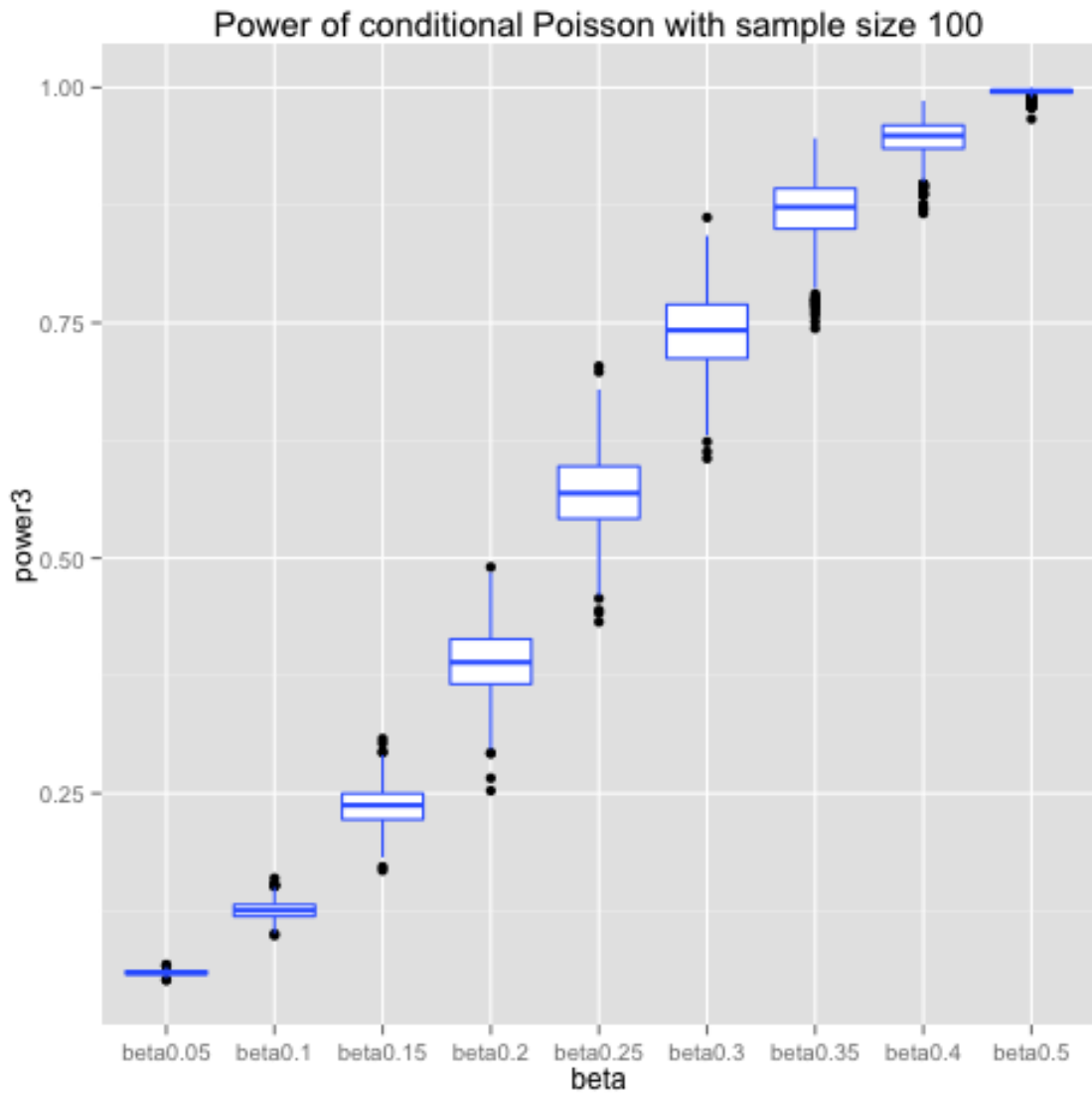


Figure 3.6: Power of the conditional Poisson method (one-side, 5% nominal type I error risk) with sample size 100.

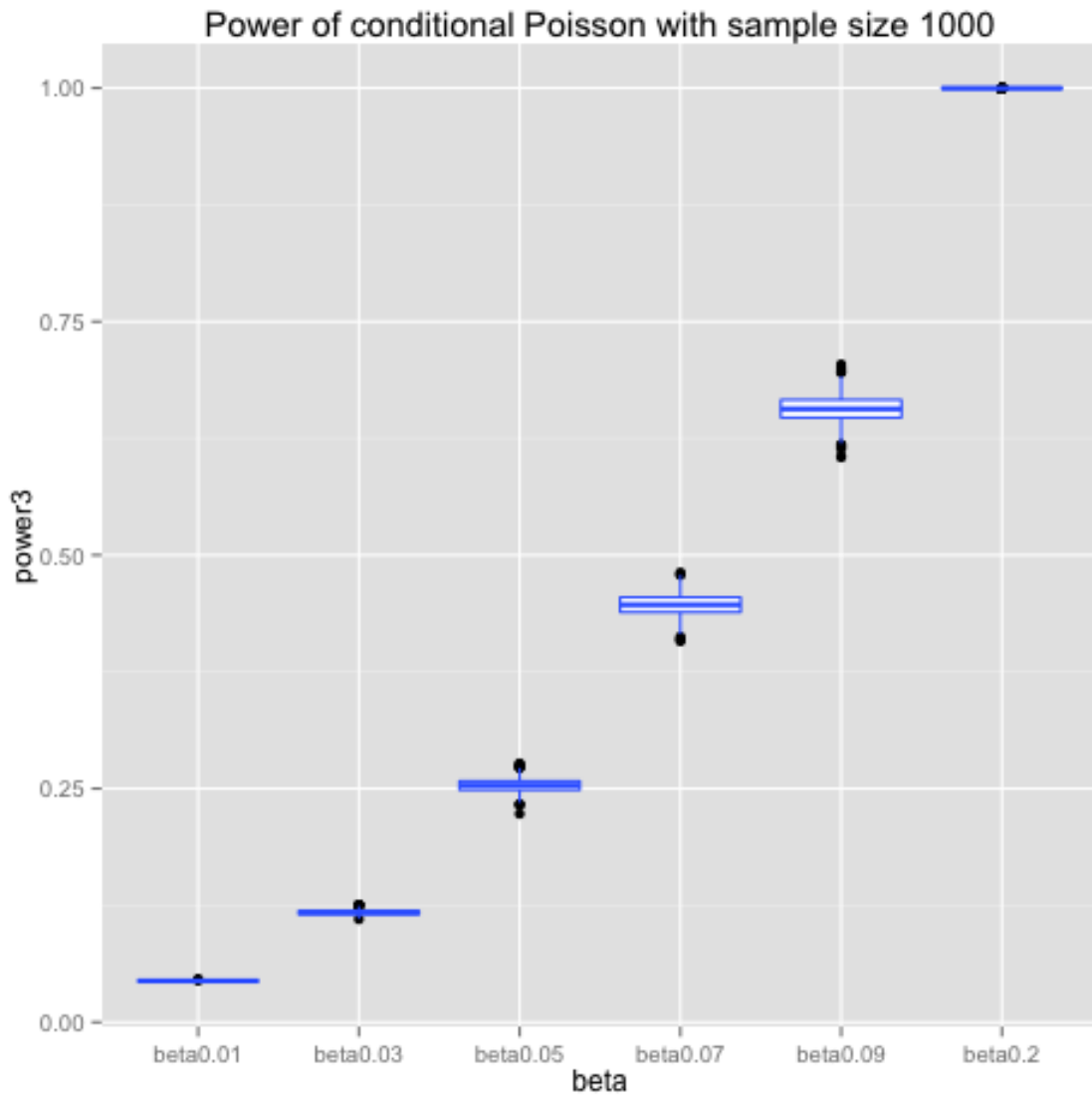


Figure 3.7: Power of the conditional Poisson method (one-side, 5% nominal type I error risk) with sample size 1000.

$P(\text{non-case})=0.22\%$, we can look at them as the case series sampling. There is no obvious difference between three methods, and the estimation of β doesn't seem to depend too much on the methods used, due to the high proportion of cases.

Theoretically, when β_1 is close to 0, power would not be less than type I error α , that means $\text{power} \geq \alpha = 0.05$. However, here the corresponding power of $\beta_1 = 0.01$ is 0.044, which is less than 0.05. That is because we use estimated power, the proportion $\hat{\pi}$ of rejecting H_0 of $M=1000$ samples. Then the confidence interval of true power π is

$$\hat{\pi} \pm Z_{\frac{\alpha}{2}} \sqrt{\frac{\pi(1-\pi)}{M}},$$

Table 3.1: Estimated different power of three methods

β	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.5	Methods
power	0.06	0.13	0.27	0.39	0.57	0.74	0.87	0.95	0.995	conditional Poisson
power	0.08	0.14	0.26	0.44	0.59	0.75	0.83	0.97	0.99	unconditional Poisson
power	0.08	0.12	0.26	0.42	0.57	0.81	0.84	0.96	0.99	SCCS method

3.3 Mean Square Error

Mean Square Error is another criteria to compare the performance of three methods. The MSE of an estimator measures the average of the squares of the “errors”, that is, the difference between the estimator and what is estimated. It is the sum of the variance and the squared bias of the estimator or of the predictions.

$$MSE(\hat{\theta}) = Var(\hat{\theta}) + E[(\hat{\theta} - \theta)^2]$$

Let λ denote e^ϕ . We want to know how incidence rate ($\lambda e^{\beta t}$) affects the estimation $\hat{\beta}$. Here, the parameter is $\theta = [\beta, \lambda]^T$. We chose two different values of λ , $\lambda = 0.2$, $\lambda = 1$, which represent low frequency of seizures in baseline and high frequency of seizures in baseline (seizure happened 0.2 times per week and once per week). We also choose three different values of β , $\beta = 0.02$, $\beta = 0.2$, $\beta = 0.5$. Since we again used a sample size 100, and 1000 replications in order to estimate the bias and MSE using simulation.

In Figure 3.8, we compare different values of β , $\beta = 0.02$, $\beta = 0.2$, $\beta = 0.5$ and different λ 's under conditional Poisson and unconditional Poisson with full cohort sampling. The variance is much larger when $\lambda = 0.2$ than when $\lambda = 1$ which means that the baseline frequency of events affects the estimation. Low frequency or rare events leads to high variance. From Figure B, we use a single data set to compare $\hat{\beta}_u$ and $\hat{\beta}_c$. Here we choose $\lambda = 1$ and this is for the case where probability of a non-case is close to 0. It shows that the conditional and unconditional estimates of β are virtually identical under full cohort sampling.

Figure 3.9 compares the estimation of β using three methods with fixed λ under case series sampling. We see that as β increases, the variance $\hat{\beta}$ decreases. If there is high baseline risk ($\lambda = 1$), then probability of case will be close to 1 and the probability of non-case will close to 0 ($\phi = 0$, $\beta = .2$, $t_0 = 4$, $t_1 = 2$, giving $P(\text{non-case}) = 0.16\%$). If λ is low ($\lambda = 0.2$), then the probability of case will be relatively small and the probability of non-case will be relatively high ($\phi = \log(.2)$, $\beta = .2$, $t_0 = 4$, $t_1 = 2$, giving $P(\text{non-case}) = 27.6\%$).

Since high baseline frequency of events like $\lambda = 1$ cannot generate case only data with moderate sample sizes, we use baseline frequency $\lambda = 0.2$. Figure 3.10 compares

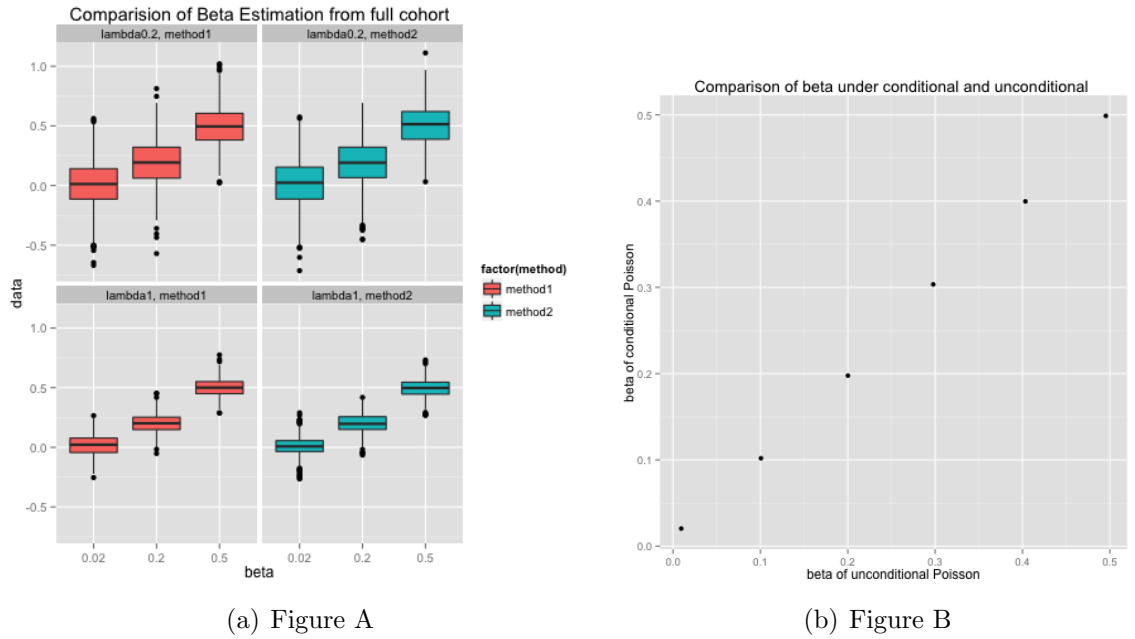


Figure 3.8: (Figure A) Compare different $\hat{\beta}$ under conditional Poisson and unconditional Poisson with fixed λ . (Figure B) compare $\hat{\beta}_u$ and $\hat{\beta}_c$ using a single data set.

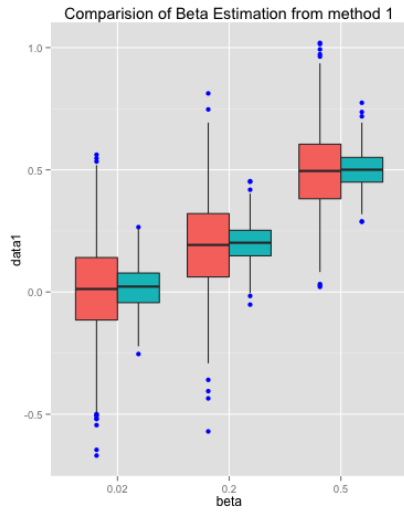
full cohort and case series sampling designs with fixed β and λ ($\lambda = 0.2$).

The parameters in the Table 3.2 are for fixed $\beta = 0.5$ and $\lambda = .5$ with $P(\text{non-case})=2.6\%$, and the estimation of β doesn't seem to depend too much on the methods used when there is a high proportion of cases.

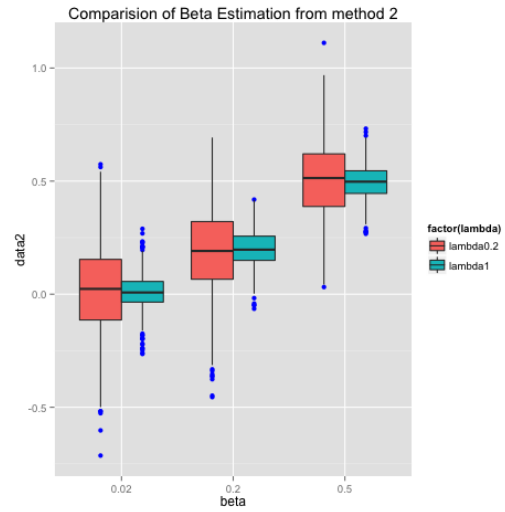
Table 3.2: MSE and Sample Size for $\beta = 0.5$ and $\lambda = 0.5$

sample size	20	50	100	200	method
MSE	3.21×10^{-2}	1.11×10^{-2}	5.60×10^{-3}	2.62×10^{-3}	Conditional on n_i .
	2.90×10^{-2}	9.93×10^{-3}	5.51×10^{-3}	2.70×10^{-3}	Unconditional
	2.88×10^{-2}	1.11×10^{-2}	5.70×10^{-3}	2.81×10^{-3}	Conditional on $n_i \geq 1$

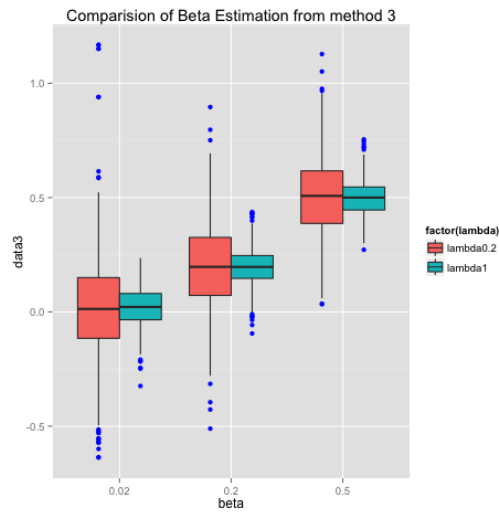
We tried four different sample sizes 20, 50, 100 and 200 to compare MSE of



(a) Figure A



(b) Figure B



(c) Figure C

Figure 3.9: Compare different λ values with fixed β under three methods. Figures A, B, C represent conditional method, unconditional method and SCCS or method of conditional on $n_i \geq 1$.

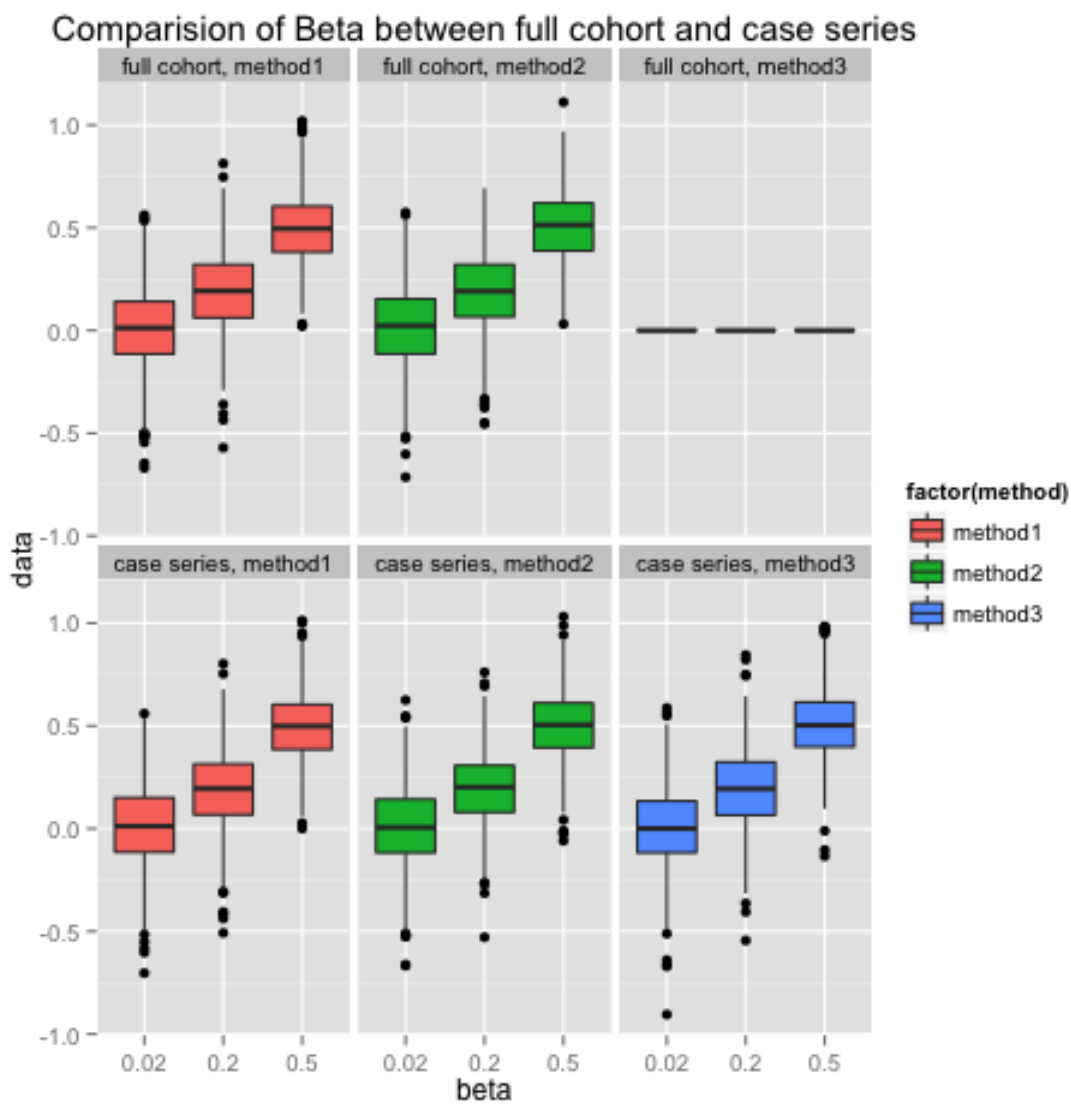


Figure 3.10: Comparison of sampling designs with $\lambda = 0.2$. Particularly, SCCS (method 3) is not applicable under full cohort sampling.

Table 3.3: Bias and MSE comparison for $\beta = 0.5$ and $\lambda = 0.5$

sampling design	full cohort		case series		
	Bias	MSE	Bias	MSE	
β	5.22×10^{-3}	5.33×10^{-2}	5.22×10^{-3}	5.33×10^{-2}	Conditional on n_i .
ϕ	1.15×10^{-2}	2.185×10^{-2}	5.88×10^{-1}	3.580×10^{-1}	Unconditional
β	1.581×10^{-2}	5.906×10^{-2}	1.582×10^{-2}	5.907×10^{-2}	
ϕ			1.87×10^{-2}	4.52×10^{-2}	Conditional on $n_i \geq 1$
β			1.12×10^{-2}	5.21×10^{-2}	

estimator under fixed $\beta = 0.5$ and $\lambda = 0.5$. We did 1000 times repetition and used full cohort sampling design for our simulation. We can get one $\hat{\beta}$ from each sample and $\hat{\beta} \sim AN\left(\beta_0, I_n^{-1}(\hat{\beta})\right)$, here “AN” represents asymptotic normality. For 1000 repetition, we got 1000 $\hat{\beta}$. Then

$$\frac{1}{1000} \sum_{i=1}^{1000} \hat{\beta}^{(i)} \implies E(\hat{\beta})$$

$$\sqrt{\frac{1}{1000} \sum_{i=1}^{1000} (\hat{\beta}^{(i)} - \bar{\hat{\beta}})^2} \implies S.E.(\hat{\beta})$$

can approximate expectation and variance of $\hat{\beta}$. Therefore, we can use them to approximate β_0 and $I_n^{-1}(\hat{\beta})$

$$\implies \hat{\beta} \sim AN\left(\beta_0, I_n^{-1}(\hat{\beta})\right).$$

And MSE and bias are:

$$\begin{aligned} \implies MSE &= Var(\hat{\beta}) + bias^2(\hat{\beta}) \\ bias(\hat{\beta}) &= E(\hat{\beta}) - \beta_0. \end{aligned}$$

From Table 3.2, for each method, when sample size increases, the MSE decreases, as we might expect, because it would be more accurate with large sample size. However, for the same sample size, there is no obvious difference amongst these three methods. For sample size from 20 to 100, MSE improved a lot by adding just a few observations.

Table 3.3 compares the bias and MSE of parameters ϕ and β for the three estimation methods. The parameters underlying this simulation were $\phi = 0(\lambda = 1)$, $t_0 = 1$ for baseline interval, and $\beta = 0.2$, $t_1 = 0.5$ for the exposed interval. The sample size was $M = 100$, and the results are based on 1000 simulation batches. With these parameters the probability of a non-case is 20%, and for case series sampling, all cases with 0 events in both the exposed and unexposed intervals were dropped.

First, ϕ has larger bias (5.88×10^{-1}) and MSE (3.580×10^{-1}) for case series sampling under unconditional Poisson method, whereas with full cohort sampling, the estimate $\hat{\phi}$ has Bias= 1.15×10^{-2} , and MSE= 2.185×10^{-2} . Interestingly, the estimate β does not change much as we move from full cohort to case series sampling. The estimation of β in either full cohort or case series sampling are almost unbiased, and the estimator would be slightly better in full cohort for unconditional Poisson. In a nutshell, if we choose unconditional Poisson and are only interested in relative incidence β , the sampling design does not affect estimation much, while if we are also interested in baseline risk ϕ , we have to use full cohort sampling to obtain correct estimation. Second, there is no difference for conditional Poisson, because non-case individuals contribute nothing to the likelihood.

Last but not least, SCCS is only suitable for case series sampling, but it can provide good estimators of both ϕ and β . In terms of estimating ϕ , SCCS seems not to

do a better job than unconditional method under full cohort sampling. However, it's still a good estimation, because when the sample size is reduced by 20%, the MSE of SCCS only twice those of the MSE of unconditional methods, and the MSE of SCCS is not too far from unconditional methods under case series sampling. In terms of estimating β , it does not change too much when sample size drops. In other words, estimation of β does not depend too much on sample design.

Consequently, if we apply wrong model in case series sampling, it does not matter in terms of estimating β . However, it does make difference in ϕ by choosing different methods under case series sampling.

We compare different sampling design, full cohort and case series, for each method in Figure 3.10. For Method 1, conditional Poisson method, there would be no difference between full cohort and case series theoretically, because non-case individuals contribute nothing to the likelihood. Method 2, the unconditional Poisson method, is nearly unbiased under full cohort sampling and shows a little bias under case series sampling. The unconditional method uses all the information from the data, and it is more suitable and performs well for full cohort sampling. The last method, the self-controlled case series or conditional on $n_i \geq 1$ method, is only suitable for case series sampling and performs best under this situation.

Chapter 4

Random Effect Model

We have already mentioned fixed effect model in Chapters 2 and 3, which means each individual has the same baseline risk e^ϕ . However, individuals are very different from each other and it may be unreasonable to assume each individual has the same fixed baseline risk. Then how about random effect? Cook [2] developed guidelines of random effect models based on mixed Poisson processes and Bayesian methods.

4.1 General Unconstrained Random Effect Model

Similar to the unconditional Poisson Model in Chapter 2, Cook gave a general unconstrained likelihood and Bayesian model. In this model, individuals are vaccinated at random prior to the second period with z_i being the indicator of vaccination for the i th individual. Suppose the study consists of M subjects. Let t_0 denote the duration of the period prior to randomization for all subjects and n_{i0} the number of events experienced for subject $i, i = 1, 2, \dots, M$. Let t_1 denote the duration of the observation period after randomization with the corresponding event count $n_{i1}, i = 1, 2, \dots, M$. Let $z_i = 1$ if subject i is in the vaccination group and $z_i = 0$ otherwise. To induce an association between N_{i0} and N_{i1} we introduce a subject-specific random effect U_i which will vary for different individuals. Then N_{i0} and N_{i1} are assumed conditionally independent Poisson distributed with

$$\begin{aligned} N_{i0} &\sim \text{Poisson}(u_i \lambda t_0), \\ N_{i1} &\sim \text{Poisson}(u_i \lambda e^{\beta z_i} t_1). \end{aligned}$$

Here we add a random effect u_i in the relative incidence rate $u_i \lambda t_0$ instead of λt_0 or $e^\phi t_0$ ($\lambda = e^\phi$) in the previous fixed effect model. Cook assumed that the random effects $U_i, i = 1, 2, \dots, M$, are independently gamma distributed with

$$u_i \sim \text{Gamma}(\phi^{-1}, \phi^{-1}),$$

$$P(u_i) = \frac{(\phi^{-1})^{\phi^{-1}}}{\Gamma(\phi^{-1})} u_i^{\phi^{-1}-1} e^{-u_i \phi^{-1}}.$$

and where $\Lambda_{i1} = \lambda t_1 e^{\beta z_i}$, the conditional p.m.f. of N_{i1} is

$$P(n_{i1}|u_i) = \frac{(\Lambda_{i1} u_i)^{n_{i1}} e^{-\Lambda_{i1} u_i}}{n_{i1}!}, n_{i1} = 0, 1, 2, \dots$$

After updating, we get the joint probability and posterior probability

$$P(n_{i1}, u_i) = P(n_{i1}|u_i)P(u_i)$$

$$u_i|n_{i1} \sim \text{Gamma}(\phi^{-1} + n_{i1}, \phi^{-1} + \Lambda_{i1}).$$

Then the marginal probability of N_{i1} would be a negative binomial distribution with p.m.f.

$$P(n_{i1}) = \int_{u_i=0}^{\infty} P(n_{i1}, u_i) du_i$$

$$= \frac{\Gamma(\phi^{-1} + n_{i1})}{\Gamma(\phi^{-1})n_{i1}!} \left(\frac{\phi^{-1}}{\phi^{-1} + \Lambda_{i1}} \right)^{\phi^{-1}} \left(\frac{\Lambda_{i1}}{\phi^{-1} + \Lambda_{i1}} \right)^{n_{i1}}$$

$$= \frac{\Gamma(\phi^{-1} + n_{i1})}{\Gamma(\phi^{-1})n_{i1}!} \left(\frac{1}{1 + \Lambda_{i1}\phi} \right)^{\phi^{-1}} \left(\frac{\Lambda_{i1}\phi}{1 + \Lambda_{i1}\phi} \right)^{n_{i1}}.$$

Similarly, we obtain the marginal p.m.f. of N_{i0}

$$P(n_{i0}) = \int_{u_i=0}^{\infty} P(n_{i0}, u_i) du_i$$

$$= \frac{\Gamma(\phi^{-1} + n_{i0})}{\Gamma(\phi^{-1})n_{i0}!} \left(\frac{1}{1 + \Lambda_0\phi} \right)^{\phi^{-1}} \left(\frac{\Lambda_0\phi}{1 + \Lambda_0\phi} \right)^{n_{i0}}, n_{i0} = 0, 1, 2, \dots$$

N_{i0} and N_{i1} are conditionally independently Poisson distributed with

$$\begin{aligned} P(n_{i0}, n_{i1}|u_i) &= P(n_{i0}|u_i)P(n_{i1}|u_i) \\ &= \frac{e^{-u_i(\Lambda_0+\Lambda_{i1})} u_i^{n_i} \Lambda_0^{n_{i0}} \Lambda_{i1}^{n_{i1}}}{n_{i0}!n_{i1}!}, \end{aligned}$$

where $\Lambda_0 = \lambda t_0$.

After updating, we get joint probability and posterior probability

$$\begin{aligned} P(n_{i0}, n_{i1}, u_i) &= P(n_{i0}, n_{i1}|u_i)P(u_i), \\ u_i|n_{i0}, n_{i1} &\sim \Gamma(\phi^{-1} + n_i, \phi^{-1} + \Lambda_0 + \Lambda_{i1}). \end{aligned}$$

Therefore, the marginal joint p.m.f. of n_{i0} and n_{i1} is

$$\begin{aligned} P(n_{i0}, n_{i1}) &= \int_{u_i=0}^{\infty} P(n_{i0}, n_{i1}|u_i) du_i \\ &= \frac{\Gamma(\phi^{-1} + n_{i1})}{\Gamma(\phi^{-1})n_{i0}!n_{i1}!} \left(\frac{1}{(\Lambda_0 + \Lambda_{i1})\phi + 1} \right)^{\phi^{-1}} \left(\frac{\Lambda_0\phi}{(\Lambda_0 + \Lambda_{i1})\phi + 1} \right)^{n_{i0}} \\ &\quad \left(\frac{\Lambda_{i1}\phi}{(\Lambda_0 + \Lambda_{i1})\phi + 1} \right)^{n_{i1}} \\ &= \frac{\Gamma(\phi^{-1} + n_{i1})}{\Gamma(\phi^{-1})n_{i0}!n_{i1}!} \frac{(\Lambda_0\phi)^{n_{i0}} (\Lambda_{i1}\phi)^{n_{i1}}}{((\Lambda_0 + \Lambda_{i1})\phi + 1)^{\phi^{-1}+n_{i0}+n_{i1}}}. \end{aligned}$$

Then we can obtain a conditional negative binomial likelihood, to which the i th subject contributes

$$\begin{aligned} P(n_{i1}|n_{i0}) &= \frac{P(n_{i0}, n_{i1})}{P(n_{i0})} \\ &= \frac{\Gamma(\phi^{-1} + n_{i1})}{\Gamma(\phi^{-1})n_{i0}!n_{i1}!} \frac{\Gamma(\phi^{-1})n_{i0}!}{\Gamma(\phi^{-1} + n_{i0})} \\ &\quad \frac{(\Lambda_0\phi)^{n_{i0}} (\Lambda_{i1}\phi)^{n_{i1}}}{((\Lambda_0 + \Lambda_{i1})\phi + 1)^{\phi^{-1}+n_{i0}+n_{i1}}} \frac{(1 + \Lambda_0\phi)^{\phi^{-1}+n_{i0}}}{(\Lambda_0\phi)^{n_{i0}}} \\ &= \frac{\Gamma(\phi^{-1} + n_{i1})}{\Gamma(\phi^{-1})n_{i0}!n_{i1}!} \frac{(\Lambda_{i1}\phi)^{n_{i1}} (1 + \Lambda_0\phi)^{\phi^{-1}+n_{i0}}}{((\Lambda_0 + \Lambda_{i1})\phi + 1)^{\phi^{-1}+n_{i0}+n_{i1}}}. \end{aligned}$$

Note that the individual random effects do not appear in this conditional probability. We estimate $\boldsymbol{\theta} = (\beta, \psi, \phi)$ where $\psi = \Lambda_1\phi/(1 + \Lambda_0\phi)$, and $\Lambda_1 = \lambda t_1$. As Cook mentioned, the conditional negative binomial model is adequate, if interest only lies in the relative incidence rate β . Otherwise, a joint model for (N_{i0}, N_{i1}) is required for other parameters. The log-likelihood also arises from the conditional negative binomial distribution and the observed information matrix is obtained after twice differentiating with $\boldsymbol{\theta}$. The inverse of the information matrix gives the covariance of $\hat{\boldsymbol{\theta}}$. After that we can obtain power and variance efficiency, and make comparisons.

However, It is unreasonable to assume both parameters of the gamma distribution are the same in Cook's assumptions. We will discuss a full hierarchical model more reasonable below.

4.2 Random Effect for Conditional Poisson Model

For a given subject, the conditional likelihood given the total number of events can be thought as a binomial probability that the event occurred in the exposure interval, and the subject specific baseline risk e^{ϕ_i} is cancelled out in the probability $\pi = \frac{t_1 e^\beta}{t_0 + t_1 e^\beta}$ of the binomial distribution. This means that the baseline risks do not affect the likelihood, and the random effect is eliminated in the same way as a fixed effect for the conditional Poisson model.

4.3 Random Effect Hierarchical Model

First we consider the unconditional Poisson model with no constraints on the number of events n_{i0}, n_{i1} . We can develop a fully Bayesian hierarchical model which incorporates random subject effects, as follows. The likelihood is

$$\begin{aligned} n_{i0}, n_{i1} | u_i &\sim P(n_{i0}, n_{i1} | u_i) \\ &= P(n_{i0} | u_i) P(n_{i1} | u_i) \end{aligned}$$

If the prior distribution for u_i is

$$u_i | \boldsymbol{\alpha} \sim P(u_i | \boldsymbol{\alpha}),$$

and the hyperprior is

$$\boldsymbol{\alpha} \sim P(\boldsymbol{\alpha}),$$

then the joint posterior distribution of all parameters is

$$P(u_i, \boldsymbol{\alpha}, \beta | n_{i0}, n_{i1}) \propto P(n_{i0}, n_{i1} | u_i, \beta) P(u_i | \boldsymbol{\alpha}) P(\boldsymbol{\alpha}) P(\beta).$$

Here the likelihood is assumed to follow a joint Poisson distribution. We might assume, for example, a Gamma distribution as a prior for the random effects, but in general, the prior and hyperprior can be any distributions, provided they lead to a proper posterior distribution.

Now that we have established a full probability model for the data and the parameters, and we could compute the marginal posterior density of hyperparameters through the Metropolis–Hastings algorithm.

The Metropolis–Hastings algorithm [9] can draw samples from any probability distribution $P(\mathbf{u}_i, \boldsymbol{\alpha}, \beta | \mathbf{n}_0, \mathbf{n}_1)$, provided we can compute the value of a function $f(u_i, \alpha, \beta | n_{i0}, n_{i1})$ which is proportional to the joint density $P(\mathbf{u}_i, \boldsymbol{\alpha}, \beta, \mathbf{n}_0, \mathbf{n}_1)$. It works by generating a sequence of sample values in such a way that, as more and more sample values are produced, the empirical distribution of values approximates more closely the desired distribution, in this case $P(\mathbf{u}_i, \boldsymbol{\alpha}, \beta | \mathbf{n}_0, \mathbf{n}_1)$.

If interest focuses on the relative incidence parameter β , the joint samples can be marginalized to get a sample from the marginal posterior of β .

Next we consider self-controlled case series method, which has the constraint of at least one event in observation period, $n_{i0} + n_{i1} \geq 1$. The only difference to the hierarchical model is in the likelihood. As in chapter 3, under the assumption of the conditional joint Poisson model, the contribution of the i 'th subject to the conditional likelihood is

$$P(n_{i0}, n_{i1} | u_i, n_{i0} + n_{i1} \geq 1) = \frac{P(n_{i0} | u_i) P(n_{i1} | u_i, \beta)}{1 - P(n_{i0} + n_{i1} = 0 | u_i, \beta)}.$$

In principle, we can again use the Metropolis–Hastings algorithm to draw sample points from the full conditional distribution $P(u_i, \alpha, \beta | n_{i0}, n_{i1}, n_{i0} + n_{i1} \geq 1)$.

Chapter 5

Discussion

We cannot apply Poisson model or self-controlled case series method to random effect problem directly. However, Bayesian and hierarchical models combined with mixed a Poisson model provide another way to solve random effect problem.

First, ϕ has larger bias and MSE of unconditional Poisson for case series sampling than for full cohort sampling. Interestingly, the estimate β does not change much as we move from full cohort to case series sampling. The estimation of β in either full cohort or case series sampling are almost unbiased, and the estimator would be slightly better in full cohort for unconditional Poisson. In a nutshell, if we choose unconditional Poisson and are only interested in relative incidence β , the sampling design does not affect estimation much, while if we are also interested in baseline risk ϕ , we have to use full cohort sampling to obtain correct estimation. What's more, there is no difference for conditional Poisson, because non-case individuals contribute nothing to the likelihood. Last but not least, SCCS is only suitable for case series sampling, but it can obtain good estimators of both ϕ and β .

Consequently, if we apply wrong model in case series sampling, it does not matter in terms of estimating β . However, it does make a difference in ϕ by choosing different methods under case series sampling.

The random effect models would be more complicated and realistic than fixed effect models. Since random effect model has more complicated posterior and likelihood. it needs more work on importance sampling to make accurate approximation in the future.

A major advantage of the SCCS method is that the analysis adjusts for fixed

covariates. Thus, individual specific characteristics need not be included as main effects in analysis. However, some time varying covariates like age may act as effect modifiers. One example is in Farrington's [15] ITP and MMR vaccination problem.

Therefore, it would be interesting to develop models with time varying covariates. If we add more covariates into three models, unconditional Poisson will work worse because of more parameters while conditional Poisson and SCCS will still works well because they eliminate the constant covariates.

Another interesting result is that Cox proportional hazard model and homogeneous Poisson process have a lot of similarity. First, their expressions look the same as follows:

$$\lambda(t, \boldsymbol{\beta}) = \lambda_0(t)e^{\boldsymbol{\beta}^T z_i},$$

$$\lambda(t, \boldsymbol{\beta}) = \begin{cases} \lambda_0(t), & \text{baseline risk} \\ \lambda_0(t)e^{\boldsymbol{\beta}^T z_i}, & \text{high risk} \end{cases}$$

Second, we can estimate β without estimating $\lambda(or\phi)$. In other words, we can estimate β correctly even we get a wrong λ estimation in Poisson process. For Cox PH model, relative incidence β can be estimated even we do not know the baseline rate λ . Third, We may try to rearrange inter-event time for single individual in a Poisson process into several pseudo individuals in a Cox model.

Bibliography

- [1]N.J. Andrews. Statistical assessment of the association between vaccination and rare adverse events post-licensure. *Vaccine*, 20:S49–S53, 2001.
- [2]R.J. Cook and W. Wei. Conditional analysis of mixed poisson processes with baseline counts: implications for trial design and analysis. *Biostatistics*, 4(3):479–494, 2003.
- [3]C.P. Farrington. Relative incidence estimation from case series for vaccine safety evaluation. *Biometrics*, 51(1):228–235, 1995.
- [4]C.P. Farrington. Control without separate controls: evaluation of vaccine safety using case-only methods. *Vaccine*, 22(15):2064–2070, 2004.
- [5]C.P. Farrington, J. Nash, and E. Miller. Case series analysis of adverse reactions to vaccines: a comparative evaluation. *American journal of epidemiology*, 143(11):1165–1173, 1996.
- [6]U. Feldmann. Epidemiologic assessment of risks of adverse reactions associated with intermittent exposure. *Biometrics*, 49(2):419, 1993.
- [7]P.E.M. Fine and R.T. Chen. Confounding in studies of adverse reactions to vaccines. *American journal of epidemiology*, 136(2):121–135, 1992.
- [8]Peter Guttorp and Vladimir N Minin. *Stochastic modeling of scientific data*. CRC Press, 1995.
- [9]W.K. Hastings. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109, 1970.
- [10]Erich Leo Lehmann and George Casella. *Theory of point estimation*, volume 31. Springer, 1998.
- [11]M. Maclure. The case-crossover design: a method for studying transient effects on the risk of acute events. *American journal of epidemiology*, 133(2):144–153, 1991.
- [12]M. Maclure and M.A. Mittleman. Should we use a case-crossover design? *Annual review of public health*, 21(1):193–221, 2000.
- [13]C. Quantin, E. Benzenine, M. Velten, F. Huet, C. P. Farrington, and P. Tubert-Bitter. Self-controlled case series and misclassification bias induced by case selection from administrative hospital databases: application to febrile convulsions in pediatric vaccine pharmacoepidemiology. *American journal of epidemiology*, 178(12):1731–1739, 2013.

- [14]S.K. Vines and C.P. Farrington. Within-subject exposure dependency in case-crossover studies. *Statistics in medicine*, 20(20):3039–3049, 2001.
- [15]J.H. Whitaker, C.P. Farrington, B. Spiessens, and P. Musonda. Tutorial in biostatistics: the self-controlled case series method. *Statistics in medicine*, 25(10):1768–1797, 2006.

Appendix A

R Code

Empirical Power – The numerical algorithm based on LR test calculate empirical power of unconditional Poisson and self-controlled case series method.

```
datasim<-function(beta,lambda=0.5,M=100){
  mu1=lambda*exp(beta); mu0=lambda
  #rate for n1 is mu0*e1*exp(beta), for n0 is mu0*e0;
  e1=2; e0=4
  n1=rpois(M,mu1*e1); n0=rpois(M,mu0*e0)
  #returns (n1,n0)
  return(cbind(n1,n0))}

power2<-function(beta,n=100,thetastart=c(0,0)){
  count <- NULL
  c=qchisq(.95,1)
  for (i in 1:n) {
    data=datasim(beta)
    #maximize unconstrained likelihood
    optim.out=optim(par=thetastart,loglikn,data=data,
                   method="BFGS",hessian=F)
    betahat=optim.out$par[2];
    lambdahat=optim.out$par[1];
    L1=loglikn(theta=c(lambdahat,betahat),data) #negative log L unconstrained
    #maximize constrained likelihood
    optim.out=optim(par=thetastart[1],loglik0n,data=data,method="BFGS",hessian=F)
    lambdahat=optim.out$par[1];
    L0=loglik0n(lambdahat,data) #constrained negative log L
    fun=2*(L0-L1) #should be > 0
    if(fun>c){
      count=c(count,1)
    }else{
      count=c(count,0)
    }
    #print(c(sum(count)/i))
  }
  sum(count)/n #this is the estimated power
}
```

Optimization – The numerical algorithm, R optimization, which return the maximum likelihood estimators and hessian matrix to approximate variance.

```

#simulation function -----
datasim<-function(beta=0.5,lambda=1,M){
  mul=lambda*exp(beta); mu0=lambda
  e1=2; e0=4
  n1=rpois(M,mul*e1); n0=rpois(M,mu0*e0)
  return(cbind(n1,n0))}

#conditional on n (binomial)-----
loglik1=function(phi,e1,e0,beta,n1,n0){
  t2=dbinom(n0,n1+n0,2/(exp(beta)+2))
  t1=dbinom(n1,n1+n0,exp(beta)/(exp(beta)+2))
  return(log(t1)+log(t2))}

#uncondition -----
loglik2=function(phi,e1,e0,beta,n1,n0){
  t2=dpois(n0,exp(phi)*e0)
  t1=dpois(n1,exp(phi+beta)*e1)
  return(log(t1)+log(t2))}

#conditional on n>1-----
loglik3=function(phi,e1,e0,beta,n1,n0){
  t2=dpois(n0,exp(phi)*e0)
  t1=dpois(n1,exp(phi+beta)*e1)
  t3=1-dpois(0,exp(phi)*(e0+e1*exp(beta)))
  return(log(t1)+log(t2)-log(t3))}

#likelihood -----
loglikn=function(theta,data){
  temp=0
  phi=theta[1]; beta=theta[2] #parameters phi and beta
  for(i in 1:(dim(data)[1])){
    n1=data[i,1]; n0=data[i,2]
    e1=2; e0=4
    temp=temp-loglik1(phi,e1,e0,beta,n1,n0)}
  return(temp)}

thetastart=c(0,0)

#optimization -----
data=datasim(M=20)
optim.out=optim(par=thetastart,fn=loglikn,gr=NULL,data=data,hessian=T) #ML Estimation

```



```
hess=optim.out$hessian #information matrix
sds=sqrt(diag(solve(hess)))

sds=sqrt(solve(hess[2,2]))
```