

Exact and Asymptotic Inference in Clinical  
Trials with Small Event Rates under Inverse

*Original*

Exact and Asymptotic Inference in Clinical  
Trials with Small Event Rates under Inverse

Sampling / Heimann, G.; Von Tress, M.; Gasparini, Mauro. - ELETTRONICO. - (2014), pp. 1-42.

*Availability:*

This version is available at: 11583/2563549 since:

*Publisher:*

*Published*

DOI:

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*

(Article begins on next page)

Exact and Asymptotic Inference in Clinical Trials with Small  
Event Rates under Inverse Sampling

**Günter Heimann, Mark Von Tress, Mauro Gasparini**

**Preprint**  
Version 1.0

July 31, 2014

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Motivating Example</b>	<b>5</b>
<b>3</b>	<b>Statistical Model and Notation</b>	<b>6</b>
<b>4</b>	<b>Exact Inference</b>	<b>6</b>
4.1	An Alternative to Fisher's Exact Test . . . . .	8
<b>5</b>	<b>Interim Analyses Using Exact Distributions</b>	<b>9</b>
5.1	Motivating Example Continued . . . . .	11
<b>6</b>	<b>Asymptotic Inference</b>	<b>13</b>
<b>7</b>	<b>Inverse Versus Conventional Sampling</b>	<b>18</b>
<b>8</b>	<b>Appendix</b>	<b>23</b>
8.1	Proofs and Derivations for Section 4 . . . . .	23
8.2	Proofs and Derivations for Section 6 . . . . .	27
8.3	Sample Size and Asymptotic Relative Efficiency . . . . .	36

# 1 Introduction

Inverse sampling schemes are frequently applied in retrospective epidemiology studies when the event of interest is rare. Inverse sampling (also called negative binomial sampling) refers to situations where one samples binary data until one has observed a fixed number of the event of interest. The sample size is a random variable, which impacts the (conditional) distributions of some of the test statistics and estimates.

Research on inverse sampling schemes has a long history. Haldane (Haldane (1945b) and Haldane (1945a)) describes a haematological experiment which investigates the occurrence of a rare type of red blood cells, and derives the expected value and variance of the corresponding estimates under inverse sampling. Steyn (see Steyn (1959)) describes an experiment where two binary variables are measured jointly and sampling is complete when a fixed number of joint failures is observed. He derives the correct distribution of the  $\chi^2$ -test for association under such a sampling scheme. This work is extended by Rudolph to a multinomial experiment (see Rudolph (1967)).

In the 1990's, Lui developed methods to compare two groups based on binary endpoints and inverse sampling (Lui (1995), Lui (1996), Lui (1997), and Lui (2000)). More recently, different exact and asymptotic methods to compare risk ratios or risk differences across two treatment groups with inverse sampling have been discussed (Tian et al. (2008), Tang & Tian (2009), and Tang & Tian (2010)). This work is further improved by Röhmel (see Röhmel (2010)).

An example of a recent epidemiological study which compares conventional and independent inverse sampling is published by Aggarwal & Pandey (2010). This study explores the prevalence of leprosy in two different areas of India. The authors conclude that "the cost of data collection was more in conventional sampling as compared to inverse sampling", acknowledging that the precision of the estimator obtained in the conventional sampling area was better. Another example is the design of a cytogenetic study discussed in Moreno et al. (2002). This paper also compares independent inverse with conventional sampling to conclude that inverse sampling "could save 24% of the fixed sample size".

All these authors consider the case of two groups and *independent* inverse sampling. With such a scheme, the number of cases is fixed separately for each treatment group. Sampling ends in any given treatment group if the target number of cases has been observed for that group. Independent inverse sampling is applied in retrospective experiments, where the sample sizes of both groups can vary considerably, if the underlying event rates differ. Hence, such a scheme is not feasible in a prospective randomized clinical trial under double blind conditions.

In such a prospective randomized trial one needs to assign patients to both groups throughout the entire study period. This can be achieved by fixing the total number of cases across both treatment groups when designing the trial. Recruitment of patients ends when the target number of overall cases has been observed, regardless of the treatment group. The random sample sizes in both groups will be of comparable size, even if the underlying event rates differ across the groups. We refer to such a scheme as *inverse sampling* to distinguish it from *independent inverse sampling*.

To illustrate the difference, consider the  $2 \times 2$ -table

	event	no event	total
placebo	$R_0$	$Q_0$	$N_0$
active	$R_1$	$Q_1$	$N_1$
total	$R$	$Q$	$N$

summarizing the results of a corresponding trial. With conventional sampling,  $N_0$  and  $N_1$  are fixed by design, and  $R_0$  and  $R_1$  are the random variables which contain the relevant information about the treatment difference. With independent inverse sampling,  $R_0$  and  $R_1$  are fixed by design, and  $N_0$  and  $N_1$  are random variables which contain the relevant information about the treatment difference. In this paper, we consider the situation where only  $R$  is fixed by design. We show that  $R_1$  and  $(Q_1, N)$  are the (independent) random variables of interest under inverse sampling, and derive the corresponding exact and asymptotic distributions.

Inverse sampling is not very common in the context of clinical trials. Chan & Bohidar (1998) mention an example of a pivotal trial to develop a new vaccine. In their example, incidence rates are low and large sample sizes would be required with conventional sampling. Alternatively, sampling is done until a specified number of cases is observed. Chan & Bohidar (1998) conclude that such a design "avoids the situation where the anticipated power is not achieved because the number of cases that occurs by the study completion is fewer than expected". In Bolland & Whitehead (2000), an example from safety monitoring is discussed, where only the relevant adverse events (cardiac arrests and all cause deaths) are reported to a drug safety monitoring committee. The committee only has access to these cases, but not to the other patient data for reasons of blinding. The corresponding safety interim analyses are then purely based on the cases, and are analogous to the analyses discussed in this paper.

Our motivating example is similar to the example in Chan & Bohidar (1998): a clinical trial with small incidence rates, designed to run until a fixed number of cases is observed rather than fixing the sample size. In this paper we discuss the consequences of such a sampling scheme on statistical inference, including interim analyses, and we explain the differences to and advantages over conventional sampling designs.

Our paper is organized as follows. In the next section we present the motivating example. Then we introduce the model and the necessary notation. Section 4 is devoted to exact inference under inverse sampling: the unconditional distributions of the relevant test statistics differ from those under conventional sampling, which allows us to present a simple exact alternative to Fisher's exact test. Exact interim analyses are discussed in section 5. In section 6 we show that the asymptotic distributions of many relevant test statistics (like the log odds ratio) and estimators are identical under inverse and conventional sampling. We also provide an asymptotic argument to show that our alternative exact test has similar power as compared to the usual tests for small incidence rates. In the last section we discuss the advantages of inverse sampling for particular situations with very small incidence rates. Proofs are deferred to the appendix.

## 2 Motivating Example

In cataract surgery one often observes an undesired side effect, endophthalmitis. This serious ocular infection occurs in approximately 0.1% of the patients. There is reason to believe that intracameral injection of an antibiotic may reduce this rate to 0.05%. In order to test this hypothesis, a trial with active treatment (the antibiotic) and a placebo control is planned, with the primary endpoint being presence or absence of the ocular infection.

Since the anticipated event rate is very low, inverse sampling is considered. The target number of cases  $r$  is to be selected to achieve a power of  $1 - \beta = 0.9$  when comparing the treatment arms, assuming event rates of 0.1% and 0.05%, a type I error of  $\alpha = 0.025$ , and a one-sided test. Unequal allocation of patients to the treatment groups is also considered.

Even with such a set-up, the expected sample size is in the order of 100,000 to 120,000 patients. This is a huge investment, and hence the sponsor is interested in performing a number of interim analyses to stop the trial early, either for success, for futility, or for both. Also, the observed placebo rates may be lower in a controlled clinical trial as compared to usual clinical practice, again posing a challenge for the overall sample size, and the interim analyses are designed to provide an early indication of this problem.

In this paper we describe the statistical aspects of designing such a trial. We provide an alternative and simpler test than Fisher's exact test, and we describe simple approaches to perform the interim analyses based on the ideas of stochastic curtailment, as described in Lan et al. (1982).

### 3 Statistical Model and Notation

Let's begin with the underlying sampling model. For each subject  $k$  one observes a pair  $(Z_k, Y_k)$  of binary random variables. The first component describes the randomization to placebo ( $Z_k = 0$ ) or treatment ( $Z_k = 1$ ). The second component describes absence ( $Y_k = 0$ ) or presence ( $Y_k = 1$ ) of the event of interest. The trial ends once  $r$  cases (= events) have been observed, so that the corresponding overall sample size  $N$  is a random variable.

The variables  $Z_k$  are binary random variables with  $\mathbb{P}rob\{Z_k = 1\} = q$ , where  $q \in ]0, 1[$  is the probability to be randomized to the treatment group 1. The variables  $Y_k$  are also binary with  $\mathbb{P}rob\{Y_k = 1\} = qp_1 + (1 - q)p_0 =: \pi$ , where  $p_l = \mathbb{P}rob\{Y_k = 1 | Z_k = l\}$  denotes the probability of experiencing endophthalmitis under treatment ( $l = 1$ ) or placebo ( $l = 0$ ).

The primary objective of the trial is to demonstrate that active treatment prevents the adverse event of interest as compared to placebo. The corresponding testing problem is

$$H_0 : p_1 \geq p_0 \quad \textit{versus} \quad H_1 : p_1 < p_0 . \quad (1)$$

With  $\lambda := \frac{p_1}{p_0}$  one can also state the hypotheses in terms of the parameter  $\lambda$ :

$$H_0 : \lambda \geq 1 \quad \textit{versus} \quad H_1 : \lambda < 1 \quad . \quad (2)$$

Note that the parameter  $\pi = \pi(\lambda, p_0)$  defined in the previous section depends both on  $\lambda$  and  $p_0$ .

Throughout this paper we use bold letters  $\mathbf{x}$  to indicate column vectors. The corresponding row vector is denoted as  $\mathbf{x}^T$ .  $\Phi$  denotes the distribution function of a standard normal distribution  $\mathcal{N}(0, 1)$ , and  $u_\alpha = \Phi^{-1}(\alpha)$  denotes the corresponding critical value. The notation  $\xrightarrow{\mathbf{P}}$  and  $\xrightarrow{\mathcal{D}}$  stand for convergence in probability and convergence in distribution, respectively.

### 4 Exact Inference

The results of a two-group comparison with binary endpoints are usually summarized in a  $2 \times 2$ -table like the one given in the introduction. It is important to note that inverse sampling alters the distributions for the variables  $R_0, R_1, Q_0,$  and  $Q_1$ , as compared to conventional sampling. This is shown in this section under general alternatives, not just under  $p_0 = p_1$ . Moreover,  $N, N_0,$  and  $N_1$  are random variables under inverse sampling, and we derive their distributions.

**Theorem 4.1** *Under inverse sampling, with  $R = r$  fixed by design, the density of the joint distribution of  $(R_1, Q_1, N)$  is*

$$\begin{aligned} & \mathbb{P}rob\{R_1 = r_1, Q_1 = q_1, N = n\} \\ &= \binom{n-1}{n-r} \binom{r}{r_1} \binom{n-r}{q_1} \rho_1^{r_1} (1-\rho_1)^{r-r_1} \rho_2^{q_1} (1-\rho_2)^{n-r-q_1} \pi^r (1-\pi)^{n-r}, \end{aligned}$$

where

$$\rho_1 = \frac{\lambda q}{\lambda q + 1 - q} \quad \text{and} \quad \rho_2 = q \frac{1 - \lambda p_0}{1 - (\lambda q + 1 - q)p_0} \quad (3)$$

are functions of  $q$  and  $\lambda$ , or  $q$ ,  $\lambda$ , and  $p_0$  respectively.

The proof of this theorem can be found in the appendix. The next corollary follows directly from Theorem 4.1.

**Corollary 4.2** *Under inverse sampling, the marginal distribution of  $R_1$  is binomial with parameters  $r$  and  $\rho_1$ , denoted by  $R_1 \sim \mathcal{B}(r, \rho_1)$ . Moreover,  $R_1$  and  $(Q_1, N)$  are independent, and the conditional distribution of  $Q_1$  given  $N$  is a binomial distribution with parameters  $N - r$  and  $\rho_2$ . The marginal distribution of  $N$  is negative binomial with parameters  $r$  and  $\pi$ , e.g.  $N \sim \mathcal{NB}(r, \pi)$ . Such a distribution has a density*

$$\mathbb{P}rob\{N = n\} = \binom{n-1}{r-1} \pi^r (1-\pi)^{n-r} \quad \forall n \geq r, \quad (4)$$

expected value  $r/\pi$ , and variance equal to  $r(1-\pi)/\pi^2$ . Finally, the marginal distribution of  $Q_1$  is also negative binomial with parameters  $r$  and  $\rho_3 = 1 - \frac{q(1-\lambda p_0)}{q+p_0-qp_0}$ . In this case we use the alternative representation of a negative binomial density, namely

$$\mathbb{P}rob\{Q_1 = q_1\} = \binom{q_1 + r - 1}{q_1} \rho_3^r (1-\rho_3)^{n-r} \quad \forall q_1 \geq 0. \quad (5)$$

Note that the situation under inverse sampling is very different from conventional sampling, where  $R_1$  and  $R_0$  would be independent binomial  $\mathcal{B}(N_1, p_1)$  and  $\mathcal{B}(N_0, p_0)$  distributions. Under inverse sampling,  $R_0 = r - R_1$ , whereas  $R_1$  and  $Q_1$  are independent.

An immediate consequence of Theorem 4.1 is that one can simplify simulations. Since the (conditional) distributions of the sufficient statistics  $N$ ,  $Q_1$ , and  $R_1$  are known under both under  $H_0$  and under  $H_1$ , one can simulate them directly, rather than having to simulate the patient level data  $(Z_1, Y_1), \dots, (Z_N, Y_N)$ . To do this, fix the parameter values  $q$ ,  $\lambda$ , and  $p_0$ . Obtain  $\pi$ ,  $\rho_1$ , and  $\rho_2$ . Then generate  $N$  from a negative binomial  $\mathcal{NB}(r, \pi)$  distribution and  $R_1$  and  $Q_1$  from the corresponding binomial distributions  $\mathcal{B}(r, \rho_1)$  and  $\mathcal{B}(N - r, \rho_2)$ . All other relevant statistics such as  $N_1 = R_1 + Q_1$ ,  $N_0$ , etc. can be obtained from the triplets  $(N, R_1, Q_1)$ .



Under conventional sampling (with  $n_0$ ,  $n_1$ , and  $n = n_0 + n_1$  fixed by design) and very small event rates, one would use Fisher's exact test to compare the two treatment groups. This test would use the fact that the conditional distribution of the test statistic  $R_1$ , given  $R = R_0 + R_1$  and  $Q$ , is hypergeometric under  $H_0$ . The next corollary (again an immediate consequence of Theorem 4.1) shows that this is also true under inverse sampling.

**Corollary 4.3** *Under inverse sampling, the conditional distribution of  $R_1$  given  $N = Q + r$  and  $N_1$  is an extended hypergeometric distribution with parameter  $\kappa = \lambda \frac{1-p_0}{1-\lambda p_0}$  and density*

$$\mathbb{P}rob\{R_1 = r_1 | N_1 = n_1, N = n\} = \binom{n_1}{r_1} \binom{n - n_1}{r - r_1} \kappa^{r_1} C_r, \quad (6)$$

where

$$C_r^{-1} = \sum_{\rho=\max(0, n_1-n+r)}^{\min(r, n_1)} \binom{n_1}{\rho} \binom{n - n_1}{r - \rho} \kappa^\rho. \quad (7)$$

Under  $H_0$  :  $\lambda = 1$  this extended hypergeometric distribution simplifies to a hypergeometric distribution.

Corollary 4.3 shows that Fisher's exact test is a valid test under inverse sampling as well. It rejects  $H_0$  if  $R_1 \leq c_{N, N_1, r, \alpha}$ , with the critical value  $c_{N, N_1, r, \alpha}$  being determined from the hypergeometric distribution.

#### 4.1 An Alternative to Fisher's Exact Test

Under inverse sampling there is a simple alternative to Fisher's exact test: one can use the test statistic  $R_1$  and the binomial distribution derived in Theorem 4.1. This test rejects  $H_0$  if  $R_1 \leq c_{r, \alpha}$ , with the critical value  $c_{r, \alpha}$  being determined from the  $\mathcal{B}(r, q)$  distribution. This is possible since  $\lambda < 1 (= 1, > 1)$  is equivalent to  $\rho_1 < q (= q, > q)$ , so that one can replace the testing problems (1) or (2) by

$$H_0 : \rho_1 \geq q \quad \text{versus} \quad H_1 : \rho_1 < q \quad . \quad (8)$$

This resulting test is effectively an unconditional test. Its distribution depends solely on  $r$  and  $q$ , which are known at the design stage of the trial, and not upon  $N$  and  $N_1$ . Hence, from a planning perspective, this test is more convenient than Fisher's exact test. We will discuss this further in Section 5.

Note that the conditional and the unconditional distributions for  $R_1$  are very close for very small proportions  $p_0$  and  $p_1$  since we then expect rather large values of  $N$

and  $N_1$ . Hence  $\frac{N_1}{N}$  will be very close to  $q$ , implying that the probabilities from the binomial  $\mathcal{B}(r, \rho_1)$  distribution and those from an extended hypergeometric distribution are very close. This in turn implies  $c_{N, N_1, r, \alpha} \approx c_{r, \alpha}$ , and that the power functions are approximately the same. More details can be found in the appendix.

Of course, one could also use  $Q_1$  as a test statistic for  $H_0$ , because  $Q_1$  (which conditionally on  $N$  follows a  $\mathcal{B}(N - r, \rho_2)$  distribution) also carries information about the original parameters  $p_0$  and  $\lambda$ : here  $\lambda < 1 (= 1, > 1)$  is equivalent  $\rho_2 > q (= q, < q)$ .

Given that  $R_1$  and  $Q_1$  are independent, it is also possible to combine  $R_1$  and  $Q_1$  into a test which is more powerful than a test based on  $R_1$  (or  $Q_1$ ) alone. How much power is lost when using the unconditional test based on  $R_1$ ? In Section 6 we show that the loss of information associated with neglecting the information in  $Q_1$  is small in our case, when the proportions  $p_0$  and  $p_1$  are very small.

## 5 Interim Analyses Using Exact Distributions

In this section we show how to obtain simple formulae for interim stopping rules or the probability for early stopping based on the exact unconditional  $\mathcal{B}(r, \rho_1)$  distribution of the test statistic  $R_1$ . The simplicity of the unconditional test based on  $R_1$  proves to be very useful in deriving these stopping rules. In contrast, the exact conditional hypergeometric distribution of the test statistic  $R_1$  would not lead to such simple rules. We use stochastic curtailment here (see Lan et al. (1982)), but other approaches (see for example Jennison & Turnbull (1999), section 12.1.2) could also be used.

With inverse sampling, an interim analysis can be conducted after the first  $s < r$  cases have been observed. Using stochastic curtailment, one can stop early to reject  $H_0$  if the conditional probability to reject  $H_0$  after  $r$  cases, given the observed data after  $s$  cases, is large enough (larger than some  $\gamma \geq \frac{1}{2}$ ). More formally, one stops early and declares success if

$$\text{Prob}_{\rho_1(1)=q}\{R_1 \leq c_{r, \alpha\gamma} | R_{1[s]}\} \geq \gamma. \quad (9)$$

Here,  $R_{1[s]}$  denotes the number of cases in the active treatment group amongst the first  $s$  cases in the study, and  $R_{1[r-s]}$  denotes the number of cases in the active treatment amongst the second  $r - s$  cases.  $R_{1[r-s]}$  follows a  $\mathcal{B}(r - s, \rho_1)$  distribution and is independent of  $R_{1[s]}$ . Note that  $R_1 = R_{1[s]} + R_{1[r-s]}$ .

The significance level at the final analysis has to be set to  $\alpha\gamma$  in order to control the overall type I error by  $\alpha$ , see Lan et al. (1982). The probability in (9) is evaluated under  $H_0$ , i.e. under  $\rho_1(1) = q$  or equivalently  $\lambda = 1$ .

In our specific case, the early stopping rule (9) can be simplified to

$$\sum_{\nu=0}^{c_{r,\alpha\gamma}-R_{1[s]}} \binom{r-s}{\nu} q^\nu (1-q)^{r-s-\nu} \geq \gamma, \quad (10)$$

because of Theorem 4.1 and the independence of  $R_{1[s]}$  and  $R_{1[r-s]}$ .

The probability to stop early for success (which is a function of  $\rho_1$ ) can be obtained as

$$\sum_{\nu=0}^{c_{r,\alpha\gamma}-b_{r-s,\gamma}} \binom{s}{\nu} \rho_1^\nu (1-\rho_1)^{s-\nu}, \quad (11)$$

where  $b_{r-s,\gamma}$  is the smallest value such that  $\mathbb{P}rob_{\rho_1=q}\{R_{1[r-s]} \leq b_{r-s,\gamma}\} \geq \gamma$ . Since  $R_{1[r-s]}$  has a  $\mathcal{B}(r-s, \rho_1)$  distribution, one can obtain  $b_{r-s,\gamma}$  as

$$b_{r-s,\gamma} = \operatorname{argmin}_b \left\{ \sum_{\nu=0}^b \binom{r-s}{\nu} q^\nu (1-q)^{r-s-\nu} \geq \gamma \right\}. \quad (12)$$

Simple expressions to define a stopping rule for futility can also be obtained when using the binomial distribution of the test statistic  $R_1$ . Following the ideas in Lan et al. (1982), one may stop early for futility if

$$\mathbb{P}rob_{\rho_1(\lambda)}\{R_1 \leq c_{r,\alpha\gamma} | R_{1[s]}\} \leq \kappa, \quad (13)$$

with  $0 \leq \kappa \leq \frac{1}{2}$ . Here the probability is evaluated under the alternative hypothesis, for example under the parameter  $\lambda = 0.5 \Leftrightarrow \rho_1(0.5)$  used for sample size calculations. Note that (13) assumes that one may stop early for futility and efficacy (with stopping boundary  $\gamma$ ). If one only wants to stop early for futility, but not for efficacy, then choose  $\gamma = 1$  in (13).

The stopping rule (13) can be evaluated using the binomial distribution of  $R_{1[r-s]}$  and the independence of  $R_{1[s]}$  and  $R_{1[r-s]}$ , just as shown in (10). Moreover, if  $d_{r-s,\kappa}$  is defined to be the largest value such that  $\mathbb{P}rob_{\rho_1(0.5)}\{R_{1[r-s]} \leq d_{r-s,\kappa}\} \leq \kappa$ , one can calculate the probability to stop early for futility as

$$\sum_{\nu=c_{r,\alpha\gamma}-d_{r-s,\kappa}}^s \binom{s}{\nu} \rho_1^\nu (1-\rho_1)^{s-\nu}. \quad (14)$$

Note that a value  $d_{r-s,\kappa}$  with  $\mathbb{P}rob_{\rho_1(0.5)}\{R_{1[r-s]} \leq d_{r-s,\kappa}\} \leq \kappa$  may not exist (i.e. if  $\mathbb{P}rob_{\rho_1(0.5)}\{R_{1[r-s]} = 0\} > \kappa$ ). In that case  $d_{r-s,\kappa}$  doesn't exist for any  $s' > s$ , and the probability to stop early for futility is 1 for all  $s' \geq s$ .

It is well known that interim analyses for futility decrease the power of the test, and that one has to adjust the required number of cases accordingly. This can be done by

using  $\frac{\beta}{1-\kappa}$  instead of  $\beta$  when calculating the required number of cases (see Lan et al. (1982)).

In practice one often wants to conduct several interim analyses. For example, one may want to conduct a sequence of  $J$  interim analyses, each of these after a predefined number of  $s_j$  cases ( $s_1 < s_2 < \dots < s_J < r$ ) have been observed in the trial. In our setting, one may even want to conduct an interim analysis after each observed case. The formulae derived above are valid in this case as well, regardless of the number of interim analyses. The quantities  $c_{r,\alpha\gamma} - b_{r-s,\gamma}$ ,  $s = 1, \dots, r-1$ , serve as stopping boundaries for efficacy, and correspondingly  $c_{r,\alpha\gamma} - d_{r-s,\kappa}$  serve as stopping boundaries for futility. An interim outcome where  $R_{1[s]}$  lies between these two boundaries would lead to continuation of the trial.

## 5.1 Motivating Example Continued

Let us return to our motivating example. In order to achieve a power of  $1 - \beta = 0.9$  with a type I error of  $\alpha = 0.025$  and a one-sided test,  $r = 94$  cases are required. This assumes that there are event rates of 0.1% for placebo and 0.05% under active treatment (i.e.  $\lambda = 0.5$ ), and a randomization proportion of  $q = 0.6$  in favor of the active treatment. The required number of cases  $r$  when no interim analysis is planned is displayed as a function of  $\lambda$  in the upper left graph of Figure 1 (solid red points).

Now assume that we want to conduct an interim analysis after every case, using the interim stopping rule (9) with  $\gamma = 0.8$ , say. The corresponding required number of cases is also shown in the upper left graph of Figure 1 (blue squares). For the case of a clinically relevant effect of  $\lambda = 0.5$ , the graph shows that the required number of cases is also  $r = 94$  when interim analyses (with  $\gamma = 0.8$ ) are planned. This is due to the discreteness of the underlying binomial distribution.

This effect becomes even clearer when one displays the required number of cases  $r$  as a function of  $\gamma$ , as is done in the upper right graph of Figure 1. This graph is calculated assuming a relevant effect of  $\lambda = 0.5$ . It presents the number of cases both when no interim analysis for futility is to be conducted (blue diamonds), as well if an interim analysis for futility ( $\kappa = 0.2$ , green points) is planned. Again these numbers are exact calculations using the binomial distribution.

In the case when only interim analyses for efficacy, but no interim analysis for futility are planned, one can see from this graph that the required number of cases is  $r = 94$ , as long as  $\gamma \geq 0.78$ . For these values of  $\gamma$ , there are no extra costs in terms of power or number of cases required when including an interim analysis for efficacy. This effect also comes into play if one includes an interim analysis for futility ( $\kappa = 0.2$ ), but now

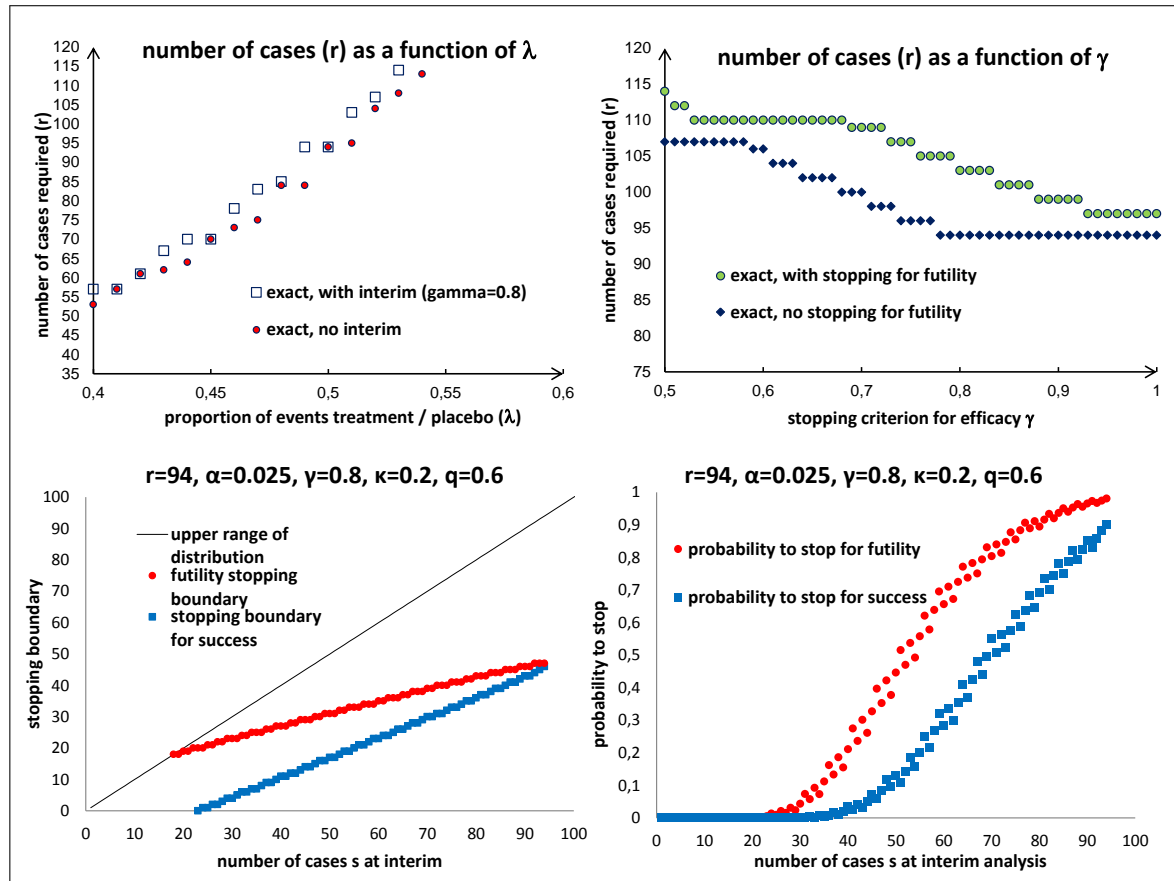


Figure 1: Upper Left: Required overall number of cases  $r$  as a function of  $\lambda$  for  $\alpha = 0.025$ ,  $1 - \beta = 0.9$ ,  $q = 0.6$ , with ( $\gamma = 0.8$ ) and without an interim analysis for efficacy. Upper right: Required overall number of cases  $r$  as a function of  $\gamma$  for  $\alpha = 0.025$ ,  $1 - \beta = 0.9$ ,  $q = 0.6$ , and  $\lambda = 0.5$ , with ( $\kappa = 0.2$ ) and without an interim analysis for futility. Lower left: Stopping boundaries for efficacy and futility ( $r = 94$ ,  $\alpha = 0.025$ ,  $q = 0.6$ ,  $\gamma = 0.8$ , and  $\kappa = 0.2$ ). Lower right: Probability to stop early for either futility (assuming  $\lambda = 1$ ) or for early success (assuming  $\lambda = 0.5$ ), as a function of the number of cases  $s$  at the interim analysis; exact calculations with  $r = 94$ ,  $\alpha = 0.025$ ,  $q = 0.6$ ,  $\gamma = 0.8$  and  $\kappa = 0.2$ .

only as long as  $\gamma \geq 0.93$ .

One can quite generally exploit the discrete nature of the test statistic to optimize the choice of the parameter  $\gamma$ . The exact boundary can be determined from the nominal level of the test ( $\alpha = 0.025$  in our case) and the actual level of the test ( $\approx 0.0193$  with  $r = 94$  in our example) as  $0.025\gamma = 0.0193$ , which is more or less equal to  $\gamma = 0.77372$ .

In the lower right graph of Figure 1 we present the operating characteristics of the procedure with  $r = 94$  when adding interim analysis for both efficacy and futility. We have chosen  $\gamma = 0.8$  as the criterion to stop for early success, and  $\kappa = 0.2$  as the criterion to stop early for futility. Both probabilities are obtained as functions of the number of cases  $s$  required to do the interim analysis. The probability to stop early for futility (red circles) is evaluated assuming  $\lambda = 1$ . The probability to stop early for success (blue squares) is evaluated assuming  $\lambda = 0.5$ . One can see that it makes little sense to conduct an interim analysis with less than 30 cases (i.e.  $s < 30$ ), because the probabilities to stop early are close to zero. A 20% chance to stop early for futility can be achieved with  $s = 40$  cases at the interim analysis. One needs to wait for 57 cases in order to have a 20% chance to stop early for success.

The lower left graph of Figure 1 displays the corresponding stopping boundaries, under the same scenario ( $r = 94$ ,  $\lambda = 0.8$ ,  $\kappa = 0.2$ ,  $\alpha = 0.025$ ) as before. The stopping boundaries  $c_{r,\alpha\gamma} - d_{r-s,\kappa}$  for futility (red circles) and  $c_{r,\alpha\gamma} - b_{r-s,\gamma}$  for efficacy (blue squares) are shown.

## 6 Asymptotic Inference

In this section we explore the asymptotic distributions of estimators and test statistics under inverse sampling. In the first part we show that the standard estimators known from conventional sampling are valid estimators under inverse sampling as well. For example, under conventional sampling  $p_1$  would be estimated by  $\hat{p}_1 = \frac{R_1}{n_1}$ . Under inverse sampling,  $N_1 = R_1 + Q_1$  is a random number, but it turns out that  $\hat{p}_1 = \frac{R_1}{N_1}$  is still a consistent estimator of  $p_1$ , and this estimator is asymptotically normally distributed with mean 0 and variance  $p_1(1 - p_1)$  as in the conventional case. In the second part of this section we demonstrate that the potential power loss associated with the binomial estimator based on  $R_1$ , disregarding the information on  $\lambda$  contained in  $(Q_1, N)$ , is minimal for small values of  $p_0$ .

Under inverse sampling,  $r$  is fixed by design rather than the sample size. Asymptotic results are obtained for increasing values of  $r$ .

**Lemma 6.1** *Under inverse sampling with  $r \rightarrow \infty$  we obtain*

$$\left(\frac{1}{r}R_1, \frac{1}{r}Q_1, \frac{1}{r}N\right) \xrightarrow[r \rightarrow \infty]{\mathbf{P}} \left(\rho_1, \frac{1-\pi}{\pi}\rho_2, \frac{1}{\pi}\right). \quad (15)$$

The notation  $\xrightarrow[r \rightarrow \infty]{\mathbf{P}}$  stands for convergence in probability.

This lemma follows from the weak law of large numbers. For details we refer to the appendix.

**Theorem 6.2** *Under the conditions of Lemma 6.1 we have*

$$\sqrt{r} \begin{pmatrix} \frac{1}{r}R_1 - \rho_1 \\ \frac{1}{r}Q_1 - \frac{1-\pi}{\pi}\rho_2 \\ \frac{1}{r}N - \frac{1}{\pi} \end{pmatrix} \xrightarrow[r \rightarrow \infty]{\mathcal{D}} \mathcal{N}(\mathbf{0}_3, \Sigma), \quad (16)$$

where  $\mathbf{0}_3 = (0, 0, 0)^T$  is a three-dimensional column vector with all entries equal to zero, and where

$$\Sigma = \begin{pmatrix} \rho_1(1-\rho_1) & 0 & 0 \\ 0 & \frac{1-\pi}{\pi^2}[\pi\rho_2(1-\rho_2) + \rho_2^2] & \frac{1-\pi}{\pi^2}\rho_2 \\ 0 & \frac{1-\pi}{\pi^2}\rho_2 & \frac{1-\pi}{\pi^2} \end{pmatrix}. \quad (17)$$

The notation  $\xrightarrow[r \rightarrow \infty]{\mathcal{D}}$  stands for convergence in distribution.

The proof of the theorem is provided in the appendix.

As an immediate consequence of lemma 6.1 we obtain  $\hat{p}_1 \xrightarrow[r \rightarrow \infty]{\mathbf{P}} p_1$ , i.e. the consistency of the conventional maximum likelihood estimator. Moreover, with a little bit of algebra one can directly derive

$$\sqrt{N_1}(\hat{p}_1 - p_1) \xrightarrow[r \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, p_1(1-p_1)) \quad (18)$$

from Theorem 6.2. Hence, the asymptotic tests and confidence intervals for  $p_1$  which one would use under conventional sampling are valid under inverse sampling as well. Also, the usual asymptotic confidence interval for the risk difference  $p_1 - p_0$  known from conventional sampling can be used under inverse sampling.

The odds ratio  $\theta = \frac{p_1(1-p_0)}{(1-p_1)p_0} = \frac{\rho_1(1-\rho_2)}{(1-\rho_1)\rho_2}$  can be estimated as

$$\hat{\theta} = \frac{R_1(N-r-Q_1)}{(r-R_1)Q_1}. \quad (19)$$

**Corollary 6.3** *Under inverse sampling,  $\hat{\theta}$  and  $\log(\hat{\theta})$  are consistent estimators for  $\theta$  and  $\log \theta$  (a direct consequence of Lemma 6.1). Theorem 6.2 and the delta method imply*

$$\sqrt{r} \left( \log(\hat{\theta}) - \log(\theta) \right) \xrightarrow[r \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \sigma_{IS}^2). \quad (20)$$

The asymptotic variance  $\sigma_{IS}^2$  equals

$$\sigma_{IS}^2 = \frac{1}{\rho_1(1-\rho_1)} + \frac{\pi}{1-\pi} \frac{1}{\rho_2(1-\rho_2)} \quad (21)$$

(see Lemma 6.1, (20), and (21)), and can be estimated consistently by  $\hat{\sigma}_{IS}^2 = \frac{r}{R_1} + \frac{r}{Q_1} + \frac{r}{r-R_1} + \frac{r}{N-r-Q_1}$ . This implies that the usual confidence interval

$$\log(\hat{\theta}) \pm u_{1-\alpha/2} \sqrt{\frac{1}{R_1} + \frac{1}{Q_1} + \frac{1}{r-R_1} + \frac{1}{N-r-Q_1}} \quad (22)$$

is valid under inverse sampling.

The corollary follows directly from Theorem 6.2, using a Taylor expansion for  $\sqrt{r}(\log(\hat{\theta}) - \log(\theta))$ . Note that under conventional sampling, the asymptotic limiting distribution is a normal distribution with mean 0 and variance  $\sigma_{CS}^2 = \frac{1}{qp_1(1-p_1)} + \frac{1}{(1-q)p_0(1-p_0)}$ , which can be estimated by  $\hat{\sigma}_{CS}^2 = \frac{N}{r} \hat{\sigma}_{IS}^2$ , with  $\hat{\sigma}^2$  as defined in the corollary.

We now return to a question raised at the end of section 4, where we propose to use the binomially distributed test statistic  $R_1$  for inference about  $H_0 : \lambda = 1$ . As a matter of fact, the independent test statistics  $R_1$  and  $(Q_1, N)$  both carry information about the parameter  $\lambda$ . Hence, in order to optimally exploit the data, one would need to use a test statistics which uses both  $R_1$  and  $(Q_1, N)$ . How much "information" do we lose if we base inference just on  $R_1$ ? In order to answer this question, we now look at the asymptotic distribution of  $R_1$  and  $Q_1$  under contiguous alternatives.

**Theorem 6.4** *Let  $\bar{R}_1 = \frac{1}{r}R_1$  and  $\bar{Q}_1 = \frac{1}{N-r}Q_1$ . Define a sequence  $\lambda_r = 1 - \frac{t}{\sqrt{r}}$  of contiguous alternatives. Then*

$$\begin{pmatrix} \frac{\sqrt{r}(\bar{R}_1 - q)}{\sqrt{\bar{R}_1(1-\bar{R}_1)}} \\ \frac{\sqrt{N-r}(\bar{Q}_1 - q)}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}} \end{pmatrix} \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r)} \mathcal{N} \left( \begin{pmatrix} -\sqrt{1-p_0} t \Delta \\ \sqrt{p_0} t \Delta \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \quad (23)$$

holds, where  $\Delta = \sqrt{\frac{q(1-q)}{1-p_0}}$ , and where  $\xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r)}$  indicates that the result holds under the assumption that the sequence  $(p_0, \lambda_r)$  are the true parameters. Moreover, for the log odds ratio we can show

$$\sqrt{r} \frac{\log(\hat{\theta})}{\hat{\sigma}_{IS}} \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r)} \mathcal{N}(-t\Delta, 1). \quad (24)$$



The proof is sketched in the appendix.

One can directly obtain the asymptotic power functions from this result. For example,  $\lambda < 1$  implies  $\rho_1 < 1$ , so that the asymptotic test based on  $R_1$  (referred to as the "R-based" test) rejects  $H_0 : \lambda \geq 1$  if  $\frac{\sqrt{r}(\bar{R}_1 - q)}{\sqrt{\bar{R}_1(1 - \bar{R}_1)}} < u_\alpha$ . From (23) we derive the corresponding asymptotic power function to be

$$\mathbb{P}rob_{\lambda_r} \left\{ \frac{\sqrt{r}(\bar{R}_1 - q)}{\sqrt{\bar{R}_1(1 - \bar{R}_1)}} < u_\alpha \right\} \xrightarrow{r \rightarrow \infty} \Phi \left( u_\alpha + \sqrt{1 - p_0} t \Delta \right). \quad (25)$$

Correspondingly, the asymptotic test based on  $Q_1$  (the "Q-based" test) rejects  $H_0$  if  $\frac{\sqrt{N-r}(\bar{Q}_1 - q)}{\sqrt{\bar{Q}_1(1 - \bar{Q}_1)}} > u_{1-\alpha}$ . Its asymptotic power function equals  $\Phi(u_\alpha + \sqrt{p_0} t \Delta)$ . The log odds ratio test (the "LOR-based" test) rejects  $H_0$  if  $\sqrt{r} \frac{\log(\hat{\theta})}{\hat{\sigma}_{IS}} < u_\alpha$ , and its power function equals  $\Phi(u_\alpha + t \Delta)$ .

These asymptotic power functions demonstrate that the loss in power between the LOR-based test and the R-based test is negligible for small rates  $p_0$ : the log odds ratio test (which uses the entire information contained in  $R_1$  and  $(Q_1, N)$ ) provides very little gain in power as compared to the the R-based test (based on  $R_1$  alone). Similarly, these power functions show that a test based on  $(Q_1, N)$  alone (the "Q-based" test) contains very little information about the parameter of interest  $\lambda$ .

Our theoretical results are confirmed in Figure 2, where we simulated the power of the three asymptotic tests (R-based, Q-based, and LOR-based), and compared these with the simulated power of the binomial test, which is based on  $R_1$  and its exact binomial distribution. For small values of  $r$  (such as  $r = 10$  in the top row) the asymptotic tests do not necessarily control the type I error. Hence, the power of the R-based and the LOR-based tests appears to be better than that of the binomial test, which however fully controls the type I error and is even conservative. However, this is due to the asymptotic tests not controlling the type I error. Also, the R-based test may have better power than the LOR-based test in some scenarios, but largely because the type I error control is more liberal for the R-based test in these scenarios. For larger values of  $r$  (such as  $r = 30$  in the bottom row) the asymptotic tests control the type I error well, and hence the simulated power of the asymptotic (R-based) test and the binomial test are indistinguishable. For the situation considered in this paper (i.e. with small values of  $p_0 = 0.01$  as in the left column), the asymptotic Q-based test has no power, and the simulated power of the R-based and the LOR-based tests are indistinguishable. For large values of  $p_0$  (such as  $p_0 = 0.5$  as in the right column) the power of the Q-based test is getting closer to that of the R-based test, and one can see that the LOR-based test has more power than the R-based test.

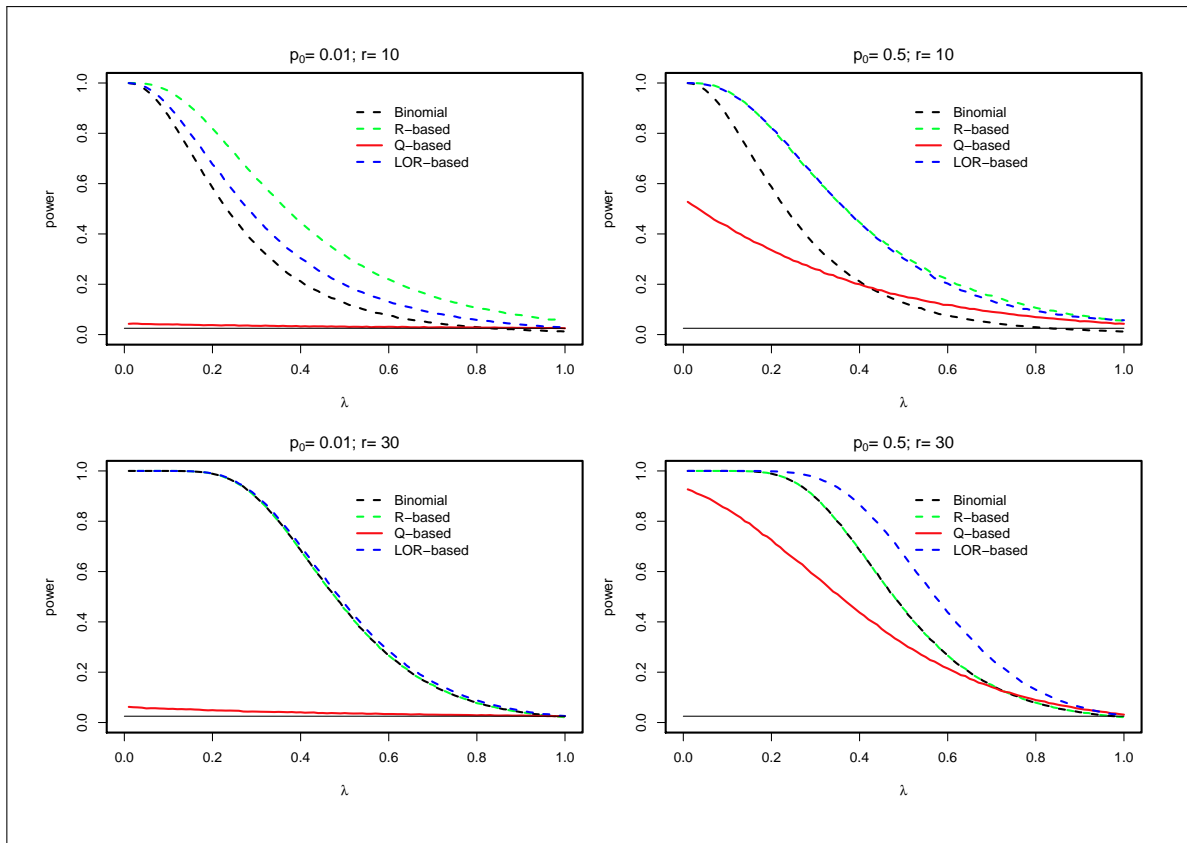


Figure 2: Simulated power functions based on 100,000 simulations for three asymptotic tests (based on  $R_1$  (R-based),  $Q_1$  (Q-based), and  $\log(\hat{\theta})$  (LOR-based) as described in Theorem 6.4) and for the exact test based on  $R_1$  using the binomial distribution; top row  $r = 10$ , bottom row  $r = 30$ , left column  $p_0 = 0.01$ , right column  $p_0 = 0.5$ ; all with  $\alpha = 0.025$ .

## 7 Inverse Versus Conventional Sampling

In this paper, we derived the exact unconditional distributions for a two-by-two table under inverse sampling, both under the null-hypothesis as well as under alternatives. These exact unconditional distributions differ from those obtained under conventional sampling. The exact unconditional distribution of the test statistic  $R_1$  under inverse sampling allowed us to obtain stopping boundaries for a sequential trial in a particularly simple manner. Moreover, we can calculate the corresponding probabilities to stop early for futility or for efficacy exactly. In contrast, under conventional sampling the interim stopping boundaries are conditional, and stopping probabilities would have to be evaluated via simulations.

We have also shown that the loss of power when basing inference solely on the exact unconditional binomial distribution of  $R_1$  is negligible in situations with very small event rates. The corresponding theoretical result is confirmed via simulations, as shown in Figure 2.

From a sample size perspective, there is not much difference between the two sampling schemes. Under conventional sampling, the sample size for the log odds ratio test is  $n_* = \sigma_{CS}^2 \frac{[u_\alpha + u_\beta]^2}{\log(\theta_*)^2}$ . The index “\*” emphasizes that the corresponding quantities depend upon the relevant treatment effect  $\lambda_*$  and the nuisance parameter  $p_{0*}$ . For the same design scenario (i.e.  $\lambda_*$  and  $p_{0*}$ ) one needs  $r_* = \sigma_{IS}^2 \frac{[u_\alpha + u_\beta]^2}{\log(\theta_*)^2}$  cases under conventional sampling. The expected sample size for this design scenario under inverse sampling is  $\mathbb{E}_{\lambda_*, p_{0*}}[N|r_*] = \frac{r_*}{\pi_*}$ , with  $\pi_* = \pi(\lambda_*, p_{0*})$ . With a little bit of algebra one can show that the expected sample size is equal to  $n_*$  (see appendix).

The above comparison is valid under the assumption that the design scenario is the true scenario. In order to fully understand the performance of the two designs, one needs to compare the conventional sample size  $n_*$  to the expected sample size  $\mathbb{E}_{\lambda, p_0}[N|r_*] = \frac{r_*}{\pi}$  from inverse sampling obtained under parameter values  $(\lambda, p_0)$  different from the design parameters  $(\lambda_*, p_{0*})$ . It is easy to see that  $\frac{\mathbb{E}_{\lambda, p_0}[N|r_*]}{n_*} = \frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_0)}$ . The average sample size required under inverse sampling is smaller than the sample size  $n_*$  under conventional sampling for all values of  $(\lambda, p_0)$  where  $\frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_0)} < 1$ .

However, the smaller average sample size comes with a decrease in power as compared to inverse sampling when  $\frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_0)} < 1$ . This is confirmed from the simulated power functions of the log odds ratio test under the two sampling schemes presented in Figure 3. The power functions correspond 90 cases (inverse sampling) and 127482 patients (conventional sampling). These numbers were derived from the design scenario with  $\lambda_* = 0.5$  and  $p_{0*} = 0.001$ . The simulations were done with  $p_0 = 0.001$  and for different

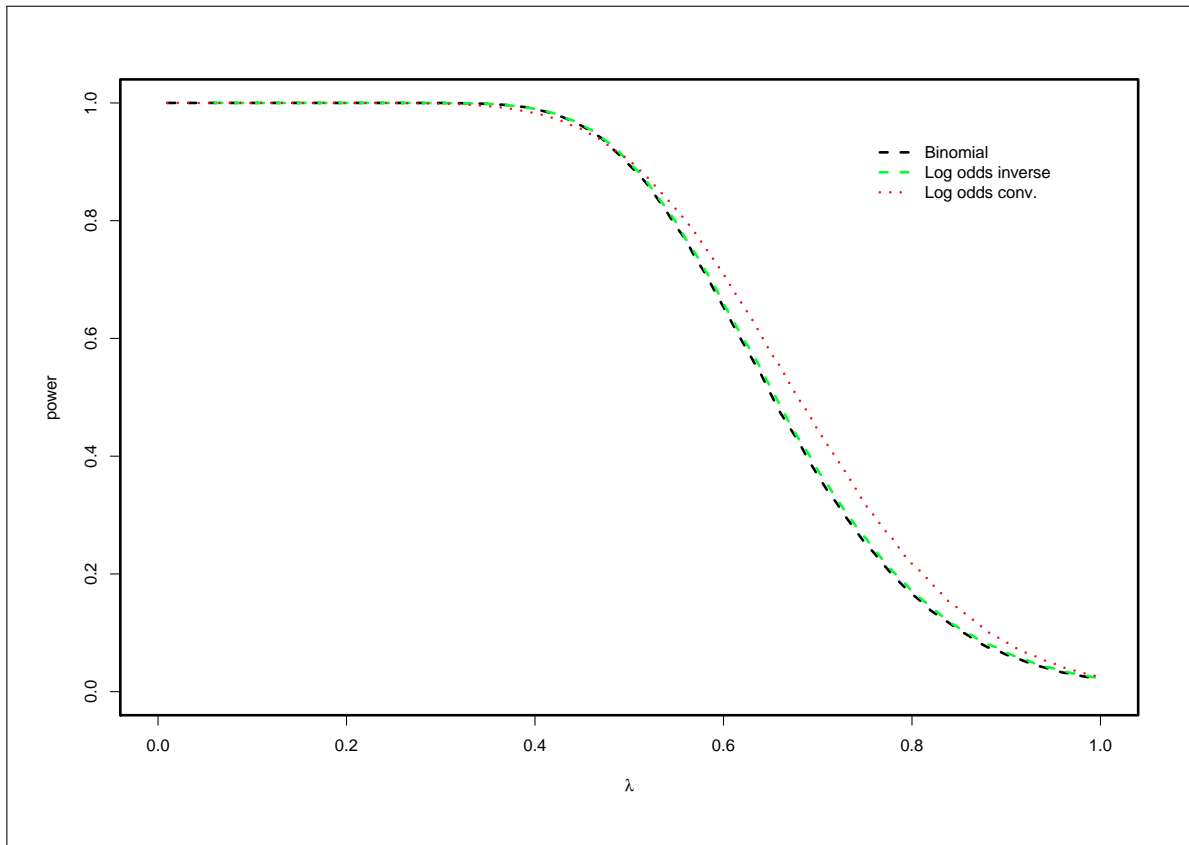


Figure 3: Simulated power function for the log-odds ratio test under inverse sampling (dashed green line,  $r = 90$  cases) and conventional sampling (dotted red line,  $n = 127482$  patients), and for the binomial test under inverse sampling (solid black line); 100,000 simulations.

values of  $\lambda$ . The power of the log odds ratio test under inverse sampling is smaller when  $\lambda > \lambda_* = 0.5$ , which is exactly the region where  $\frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_{0*})} < 1$ . However, the average sample size under inverse sampling is smaller than  $n_* = 127482$  in this region, as discussed before.

From these asymptotic considerations one cannot see too much of a difference between the two sampling schemes. The difference becomes more apparent if one looks at the (simulated) sampling distributions of the log odds ratio  $\ln(\hat{\theta})$  under inverse and conventional sampling, the corresponding variance estimators  $\hat{\sigma}_{IS}^2$  and  $\hat{\sigma}_{CS}^2$ , and the test statistics  $\sqrt{r} \frac{\ln(\hat{\theta})}{\hat{\sigma}_{IS}}$  and  $\sqrt{n} \frac{\ln(\hat{\theta})}{\hat{\sigma}_{CS}}$ . One can see that the sampling distributions for the test statistics are normally distributed. However, the sampling distributions of the log-odds ratio  $\ln(\hat{\theta})$  and the variances differ between the two sampling schemes. They vary around a common mean, but the sampling distributions under inverse sampling are much more concentrated than their counterparts under inverse sampling. This is demonstrated in Figure 4. There is an apparent difference in the sampling distributions

of the standard errors  $\sqrt{r}\hat{\sigma}_{IS}$  and  $\sqrt{N}\hat{\sigma}_{CS}$ , but no difference in the sampling distribution of the estimators and the test statistics. This is due to the higher correlation between estimator and standard error under inverse sampling, as demonstrated in Figure 5.

This stabilizing effect of inverse sampling is also obvious when performing Fisher's exact test. Under inverse sampling, the resulting tables will always include  $r$  cases. Under conventional sampling, the observed number  $R$  may be small. This does not play a role for asymptotic considerations or power calculations at the design stage, because the chance to observe a small  $R$  is balanced by the possibility to observe a very large number of cases  $R$ . But nevertheless, a study with a very small number of actual cases is a very unfortunate outcome of such a large investment. We refer to Chan & Bohidar (1998), page 1312, for a similar discussion.

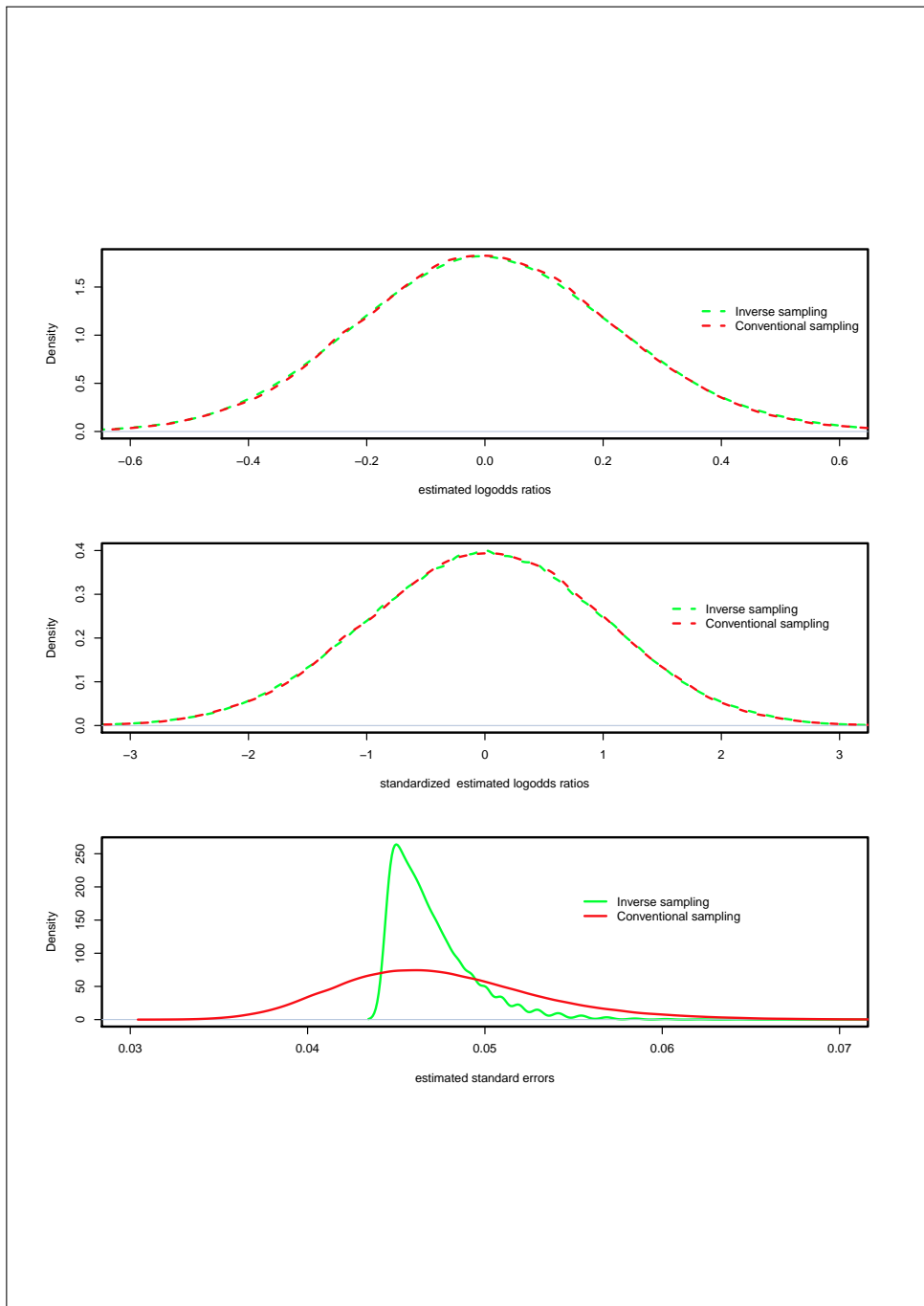


Figure 4: Sampling distributions based on 100,000 simulations for the log-odds ratio, the standard error, and the LOR based test statistic under inverse (solid green line, 90cases) and conventional (dotted red line, 127482 patients). The underlying parameters are  $p_0 = 0.001$ ,  $\lambda = 0.5$ , and  $q = 0.6$

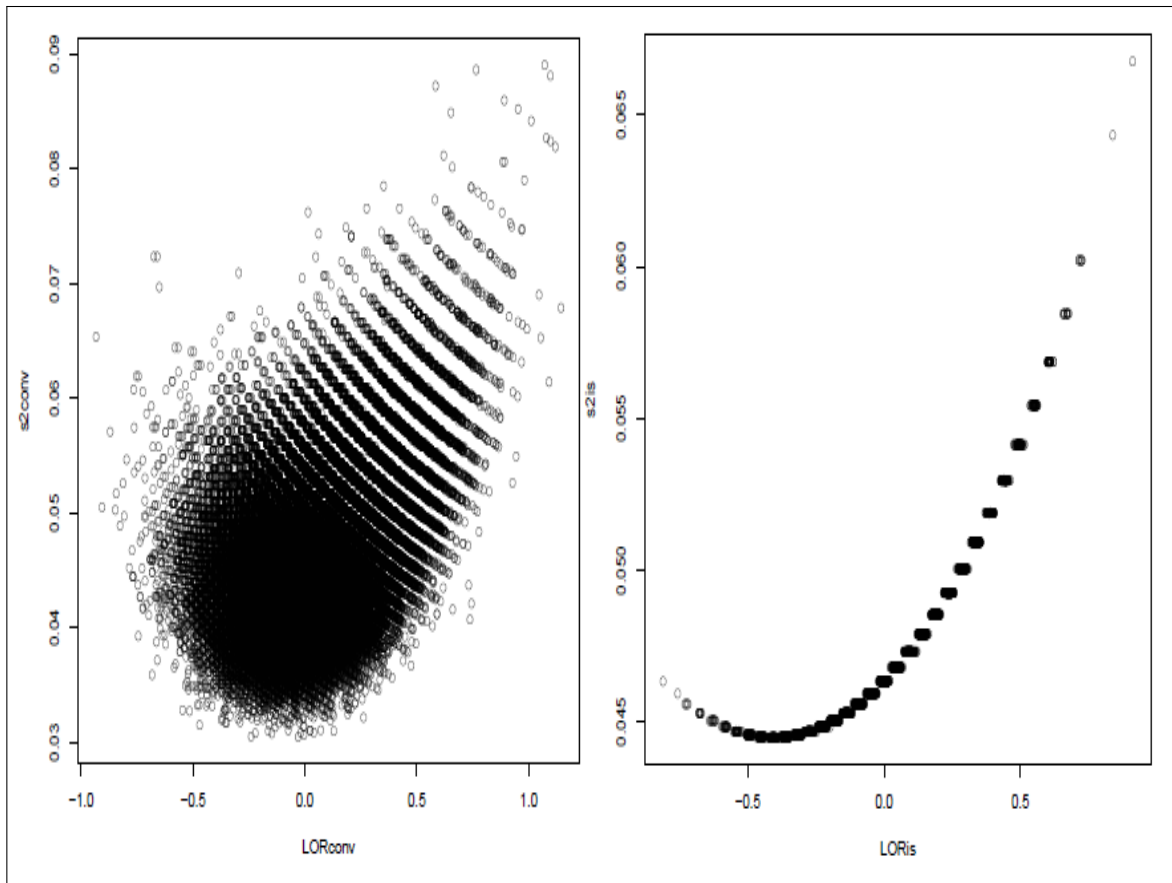


Figure 5: Sampling distributions based on 100,000 simulations the log-odds ratio (x-axis) versus the standard error (y-axis) for conventional (left, 127482 patients) and inverse sampling (right, 90 cases) sampling. The underlying parameters are  $p_0 = 0.001$ ,  $\lambda = 0.5$ , and  $q = 0.6$

## 8 Appendix

### 8.1 Proofs and Derivations for Section 4

**Proof of Theorem 4.1:** Note that the  $k$ -th observation  $(Z_k, Y_k)$  can be regarded as a multinomial random variable with four possible outcomes  $(0, 0)$ ,  $(0, 1)$ ,  $(1, 0)$ , and  $(1, 1)$ . For any given  $\nu$  define  $R_{(\nu)} = \sum_{k=1}^{\nu} Y_k$ ,  $R_{(1\nu)} = \sum_{k=1}^{\nu} Z_k Y_k$ ,  $Q_{(\nu)} = \sum_{k=1}^{\nu} (1 - Y_k)$ ,  $Q_{(1\nu)} = \sum_{k=1}^{\nu} Z_k (1 - Y_k)$ , and  $N = \min_{\nu} \{R_{(\nu)} = r\}$ . Moreover, use  $R_1 = R_{(N)}$  and  $Q_1 = Q_{(N)}$ .

Since the  $N$ -th observation is an event, we either have  $(Z_N, Y_N) = (1, 1)$  with probability  $qp_1$  (case I), or  $(Z_N, Y_N) = (0, 1)$  with probability  $(1 - q)p_0$  (case II). In both cases there are  $\binom{N-1}{N-r}$  ways to position the  $N - r$  patients without the event amongst the first  $N - 1$  observations. Furthermore, in both cases there are then  $\binom{N-r}{Q_1}$  ways to position the  $Q_1$  patients from the active treatment group without the event amongst all  $N - r$  patients without the event. Each such positioning has probability  $(q(1 - p_1))^{Q_1}((1 - q)(1 - p_0))^{N-r-Q_1}$ .

In case I, there are  $\binom{r-1}{R_1-1}$  ways to position the remaining  $R_1 - 1$  actively treated patients with events amongst the first  $r - 1$  events, and such positioning has probability  $(qp_1)^{R_1-1}((1 - q)p_0)^{r-R_1}$ . In case II, there are  $\binom{r-1}{R_1}$  ways to position the  $R_1$  actively treated patients with event between the first  $r - 1$  positive responses, and each such positioning has probability  $(qp_1)^{R_1}((1 - q)p_0)^{r-1-R_1}$ .

Overall, the above implies

$$\begin{aligned} & \mathbb{P}rob\{R_1 = r_1, Q_1 = q_1, N = n\} \\ &= qp_1 \binom{n-1}{n-r} \binom{n-r}{q_1} \binom{r-1}{r_1-1} (q(1-p_1))^{q_1} ((1-q)(1-p_0))^{n-r-q_1} (qp_1)^{r_1-1} ((1-q)p_0)^{r-r_1} \\ &+ (1-q)p_0 \binom{n-1}{n-r} \binom{n-r}{q_1} \binom{r-1}{r_1} (q(1-p_1))^{q_1} ((1-q)(1-p_0))^{n-r-q_1} (qp_1)^{r_1} ((1-q)p_0)^{r-1-r_1} \\ &= \binom{n-1}{n-r} \binom{n-r}{q_1} \binom{r}{r_1} (q(1-p_1))^{q_1} ((1-q)(1-p_0))^{n-r-q_1} (qp_1)^{r_1} ((1-q)p_0)^{r-r_1} . \end{aligned}$$

Use  $\rho_1 = \frac{qp_1}{\pi}$ ,  $\rho_2 = \frac{q(1-p_1)}{1-\pi}$ , and some additional algebra to finalize the proof.

**Proof of Corollary 4.2:** From Theorem 4.1 we immediately conclude the indepen-



dence of  $R_1$  and  $(Q_1, N)$ . Hence

$$\mathbb{P}rob\{Q_1 = q_1, N = n\} = \binom{n-1}{n-r} \binom{n-r}{q_1} \rho_2^{q_1} (1-\rho_2)^{n-r-q_1} \pi^r (1-\pi)^{n-r}. \quad (26)$$

With this we get

$$\begin{aligned} \mathbb{P}rob\{N = n\} &= \sum_{q_1=0}^{n-r} \mathbb{P}rob\{Q_1 = q_1, N = n\} \\ &= \sum_{q_1=0}^{n-r} \binom{n-1}{n-r} \binom{n-r}{q_1} \rho_2^{q_1} (1-\rho_2)^{n-r-q_1} \pi^r (1-\pi)^{n-r} \\ &= \binom{n-1}{n-r} \pi^r (1-\pi)^{n-r}, \end{aligned}$$

which proves the well known fact that  $N$  has a negative binomial distribution. Next,

$$\begin{aligned} \mathbb{P}rob\{Q_1 = q_1\} &= \sum_{n=r}^{\infty} \mathbb{P}rob\{Q_1 = q_1, N = n\} = \sum_{n=r+q_1}^{\infty} \mathbb{P}rob\{Q_1 = q_1, N = n\} \\ &= \sum_{n=r+q_1}^{\infty} \binom{n-1}{n-r} \binom{n-r}{q_1} \rho_2^{q_1} (1-\rho_2)^{n-r-q_1} \pi^r (1-\pi)^{n-r} \\ &= \sum_{n=r+q_1}^{\infty} \binom{n-1}{n-r} \binom{n-r}{q_1} \left( \frac{q(1-\lambda p_0)}{1-\pi} \right)^{q_1} \left( \frac{(1-q)(1-p_0)}{1-\pi} \right)^{n-r-q_1} \pi^r (1-\pi)^{n-r} \\ &= \binom{q_1+r-1}{q_1} (q(1-\lambda p_0))^{q_1} \sum_{n=r+q_1}^{\infty} \binom{n-1}{q_1+r-1} \pi^r ((1-q)(1-p_0))^{n-r-q_1} \\ &= \binom{q_1+r-1}{q_1} \left( \frac{q(1-\lambda p_0)}{q+p_0-qp_0} \right)^{q_1} \left( \frac{\pi}{q+p_0-qp_0} \right)^r \\ &\quad \sum_{n=r+q_1}^{\infty} \binom{n-1}{q_1+r-1} (q+p_0-qp_0)^{q_1+r} ((1-q-p_0+qp_0))^{n-r-q_1} \\ &= \binom{q_1+r-1}{q_1} \left( \frac{q(1-\lambda p_0)}{q+p_0-qp_0} \right)^{q_1} \left( 1 - \frac{q(1-\lambda p_0)}{q+p_0-qp_0} \right)^r. \end{aligned}$$

**Poof of Corollary 4.3:** First we show

$$\begin{aligned}
& \rho_1^{r_1} (1 - \rho_1)^{r-r_1} \rho_2^{q_1} (1 - \rho_2)^{n-r-q_1} \pi^r (1 - \pi)^{n-r} \\
&= (qp_1)^{r_1} ((1-q)p_0)^{r-r_1} (q(1-p_1))^{q_1} ((1-q)(1-p_0))^{n-r-q_1} \\
&= q^{r_1+q_1} p_1^{r_1} (1-p_1)^{q_1} (1-q)^{n-r_1-q_1} p_0^{r-r_1} (1-p_0)^{n-r-q_1} \\
&= q^{n_1} (1-q)^{n-n_1} p_0^r (1-p_0)^{n-r} \left( \frac{1-p_1}{1-p_0} \right)^{n_1} \left( \frac{p_1}{p_0} \cdot \frac{1-p_0}{1-p_1} \right)^{r_1} \\
&= q^{n_1} (1-q)^{n-n_1} p_0^r (1-p_0)^{n-r} \left( \frac{1-p_1}{1-p_0} \right)^{n_1} [\kappa(\lambda)]^{r_1}
\end{aligned}$$

using (3) and  $\kappa(\lambda) = \lambda \frac{1-p_0}{1-\lambda p_0}$ . Next, remember that  $R = r$  was fixed by design, and use the above chain of equalities to obtain

$$\begin{aligned}
& \mathbb{P}rob\{R_1 = r_1 | N_1 = n_1, N = n\} \\
&= \frac{\mathbb{P}rob\{R_1 = r_1, N_1 = n_1, N = n\}}{\mathbb{P}rob\{N_1 = n_1, N = n\}} \\
&= \frac{\mathbb{P}rob\{R_1 = r_1, Q_1 = n_1 - r_1, N = n\}}{\sum_{\rho=\max(0, n_1-n+r)}^{\min(r, n_1)} \mathbb{P}rob\{R_1 = \rho, Q_1 = n_1 - \rho, N = n\}} \\
&= \frac{\binom{r}{r_1} \binom{n-r}{n_1-r_1} [\kappa(\lambda)]^{r_1}}{\sum_{\rho=\max(0, n_1-n+r)}^{\min(r, n_1)} \binom{r}{\rho} \binom{n-r}{n_1-\rho} [\kappa(\lambda)]^\rho} \\
&= \frac{\binom{n_1}{r_1} \binom{n-n_1}{r-r_1} [\kappa(\lambda)]^{r_1}}{\sum_{\rho=\max(0, n_1-n+r)}^{\min(r, n_1)} \binom{n_1}{\rho} \binom{n-n_1}{r-\rho} [\kappa(\lambda)]^\rho} \\
&= \binom{n_1}{r_1} \binom{n-n_1}{r-r_1} [\kappa(\lambda)]^{r_1} C_r(\lambda),
\end{aligned}$$

which completes the proof.

**Derivation of  $c_{N, N_1, r, \alpha} \approx c_{r, \alpha}$ :**

First, note that

$$\frac{\binom{n_1}{r_1} \binom{n-n_1}{r-r_1} (\kappa(\lambda))^{r_1}}{\sum_{\rho=0}^r \binom{n_1}{\rho} \binom{n-n_1}{r-\rho} (\kappa(\lambda))^\rho} = \frac{\frac{\binom{n_1}{r_1} \binom{n-n_1}{r-r_1}}{\binom{n}{r}} (\kappa(\lambda))^{r_1}}{\sum_{\rho=0}^r \frac{\binom{n_1}{\rho} \binom{n-n_1}{r-\rho}}{\binom{n}{r}} (\kappa(\lambda))^\rho}$$

This implies

$$\frac{\binom{n_1}{r_1} \binom{n-n_1}{r-r_1} (\kappa(\lambda))^{r_1}}{\sum_{\rho=0}^r \binom{n_1}{\rho} \binom{n-n_1}{r-\rho} (\kappa(\lambda))^\rho} = \frac{\frac{\binom{r}{r_1} \binom{n-r}{n_1-r_1}}{\binom{n}{n_1}} (\kappa(\lambda))^{r_1}}{\sum_{\rho=0}^r \frac{\binom{r}{\rho} \binom{n-r}{n_1-\rho}}{\binom{n}{n_1}} (\kappa(\lambda))^\rho}.$$

Now, for very small  $p_0$  and  $p_1$  we expect very large  $N$  and  $N_1$ , so that  $\frac{N_1}{N} \approx \frac{N_1-1}{N-2} \approx \dots \approx \frac{N_1-r_1}{N-r} \approx q$ . This implies

$$\binom{n-r}{n_1-\rho} / \binom{n}{n_1} \approx q^\rho (1-q)^{r-\rho} \quad (27)$$

for all  $\rho \leq r_1$ , and hence

$$\frac{\binom{n_1}{r_1} \binom{n-n_1}{r-r_1} (\kappa(\lambda))^{r_1}}{\sum_{\rho=0}^r \binom{n_1}{\rho} \binom{n-n_1}{r-\rho} (\kappa(\lambda))^\rho} \approx \frac{\binom{r}{r_1} q^{r_1} (1-q)^{r-r_1} (\kappa(\lambda))^{r_1}}{\sum_{\rho=0}^r \binom{r}{\rho} q^\rho (1-q)^{r-\rho} (\kappa(\lambda))^\rho}. \quad (28)$$

Next, observe that

$$(\kappa(\lambda))^\rho = \left[ \frac{p_1}{p_0} \frac{1-p_0}{1-p_1} \right]^\rho \approx \left[ \frac{p_1}{p_0} \right]^\rho = \left[ \frac{p_1}{\pi} \right]^\rho \left[ \frac{p_0}{\pi} \right]^{r-\rho} \left[ \frac{\pi}{p_0} \right]^r. \quad (29)$$

In the above equation we have again made use of the fact that we have very small  $p_0$  and  $p_1$ , which implies that  $1-p_0 \approx 1$  and  $1-p_1 \approx 1$ . With this in mind, and remembering

$\rho_1(\lambda) = \frac{qp_1}{\pi}$  and  $1 - \rho_1(\lambda) = \frac{(1-q)p_0}{\pi}$ , we can now conclude

$$\frac{\binom{n_1}{r_1} \binom{n-n_1}{r-r_1} (\kappa(\lambda))^{r_1}}{\sum_{\rho=0}^r \binom{n_1}{\rho} \binom{n-n_1}{r-\rho} (\kappa(\lambda))^\rho} \approx \frac{\binom{r}{r_1} \rho_1(\lambda)^{r_1} (1 - \rho_1(\lambda))^{r-r_1}}{\sum_{\rho=0}^r \binom{r}{\rho} \rho_1(\lambda)^\rho (1 - \rho_1(\lambda))^{r-\rho}}. \quad (30)$$

Since

$$\sum_{\rho=0}^r \binom{r}{\rho} \rho_1(\lambda)^\rho (1 - \rho_1(\lambda))^{r-\rho} = 1, \quad (31)$$

we have shown that for very small  $p_0$  and  $p_1$  we get

$$\frac{\binom{n_1}{r_1} \binom{n-n_1}{r-r_1} (\kappa(\lambda))^{r_1}}{\sum_{\rho=0}^r \binom{n_1}{\rho} \binom{n-n_1}{r-\rho} (\kappa(\lambda))^\rho} \approx \binom{r}{r_1} \rho_1(\lambda)^{r_1} (1 - \rho_1(\lambda))^{r-r_1}. \quad (32)$$

Note that one can always derive (32) as an exact equality via integration for this type of sampling scheme with  $R = r$  fixed, not just with very small  $p_0$  and  $p_1$ . The approximate calculation presented in this section is only valid for very small  $p_0$  and  $p_1$ .

## 8.2 Proofs and Derivations for Section 6

In this section we will derive basic asymptotic results under inverse sampling. The proofs are based on a simple fact: the basic  $2 \times 2$  table

	event	no event	total
placebo	$R_0$	$Q_0$	$N_0$
active	$R_1$	$Q_1$	$N_1$
total	$r$	$Q$	$N$

which summarizes the results of our study with a fixed number of  $r$  cases can be obtained as the sum of  $\nu = 1, \dots, r$  independent tables

	event	no event	total
placebo	$R_{0[1],\nu}$	$Q_{0[1],\nu}$	$N_{0[1],\nu}$
active	$R_{1[1],\nu}$	$Q_{1[1],\nu}$	$N_{1[1],\nu}$
total	1	$Q_{[1],\nu}$	$N_{[1],\nu}$

The table corresponding to  $\nu = 1$  represents the data which have accumulated until the first case was observed. The table corresponding to  $\nu = 2$  corresponds to the data accumulated between the first and the second case, and so on.

The relevant vector  $(R_1, Q_1, N)$  of observations from the overall table can be expressed as a sum of independent vectors

$$(R_1, Q_1, N) = \sum_{\nu=1}^r (R_{1[1],\nu}, Q_{1[1],\nu}, N_{[1],\nu}) . \quad (33)$$

From theorem 4.1 we can immediately derive that the random variables  $R_{1[1],\nu}$  are  $\mathcal{B}(1, \rho_1)$  and independent of  $(Q_{1[1],\nu}, N_{[1],\nu})$ . The random variables  $N_{[1],\nu}$  are  $\mathcal{NB}(1, \pi)$ , and the  $Q_{1[1],\nu}$  are distributed as a  $\mathcal{B}(N_{[1],\nu} - 1, \rho_2)$  distribution (conditionally given  $N_{[1],\nu}$ ). For a definition of  $\pi = qp_1 + (1 - q)p_0$ ,  $\rho_1$ , and  $\rho_2$ , see Section 3, and (3).

**Proof of Lemma 6.1:** First use the conditional distribution of  $Q_{1[1],\nu}$  to establish

$$\mathbb{E}[Q_{1[1],\nu}] = \mathbb{E}[\mathbb{E}[Q_{1[1],\nu} | N_{[1],\nu}]] = \mathbb{E}[(N_{[1],\nu} - 1)\rho_2] = \frac{1 - \pi}{\pi} \rho_2 , \quad (34)$$

$$\text{Var}[Q_{1[1],\nu}] = \frac{1 - \pi}{\pi} \rho_2 (1 - \rho_2) + \rho_2^2 \frac{1 - \pi}{\pi^2} . \quad (35)$$

Note that (35) follows from

$$\begin{aligned} \mathbb{E} \left[ \left( Q_{1[1],\nu} - \frac{1 - \pi}{\pi} \rho_2 \right)^2 \right] &= \mathbb{E} \left[ \left( Q_{1[1],\nu} - (N_{[1],\nu} - 1)\rho_2 \right)^2 \right] + \rho_2^2 \mathbb{E} \left[ \left( N_{[1],\nu} - \frac{1}{\pi} \right)^2 \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ \left( Q_{1[1],\nu} - (N_{[1],\nu} - 1)\rho_2 \right)^2 | N_{[1],\nu} \right] \right] + \rho_2^2 \frac{1 - \pi}{\pi^2} \\ &= \mathbb{E} \left[ (N_{[1],\nu} - 1)\rho_2 (1 - \rho_2) \right] + \rho_2^2 \frac{1 - \pi}{\pi^2} \\ &= \frac{1 - \pi}{\pi} \rho_2 (1 - \rho_2) + \rho_2^2 \frac{1 - \pi}{\pi^2} , \end{aligned}$$

i.e. the variance of  $Q_1$  is the expectation of the conditional variance plus the variance of the conditional expectation. In the above chain of equations we have used

$$\mathbb{E} \left[ \left( Q_{1[1],\nu} - (N_{[1],\nu} - 1)\rho_2 \right) \rho_2 \left( N_{[1],\nu} - \frac{1}{\pi} \right) \right] = 0 . \quad (36)$$

Now, (35) and Tchebichev's inequality (see Roussas (1997)) imply  $\frac{1}{r}Q_1 - \frac{1 - \pi}{\pi} \rho_2 \xrightarrow[r \rightarrow \infty]{\mathbf{P}} 0$ .

The other two statements  $\frac{1}{r}R_1 - \rho_1 \xrightarrow[r \rightarrow \infty]{\mathbf{P}} 0$  and  $\frac{1}{r}N - \frac{1}{\pi} \xrightarrow[r \rightarrow \infty]{\mathbf{P}} 0$  can be shown analogously.

As an immediate consequence of Lemma (6.1) we get

$$\frac{R_1}{R_1 + Q_1} \xrightarrow[r \rightarrow \infty]{\mathbf{P}} p_1 , \quad (37)$$

which proves that the maximum likelihood estimator  $\hat{p}_1 = \frac{R_1}{R_1 + Q_1}$  is consistent, and

$$\frac{1}{r}N_1 \xrightarrow[r \rightarrow \infty]{\mathbf{P}} \frac{1 - q}{\pi} . \quad (38)$$

**Poof of Theorem 6.2:** The asymptotic limits of the marginal distributions

$$\frac{1}{\sqrt{r}}(N - \frac{1}{\pi}) \xrightarrow[r \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \frac{1 - \pi}{\pi^2}), \quad (39)$$

$$\frac{1}{\sqrt{r}}(R_1 - \rho_1) \xrightarrow[r \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \rho_1(1 - \rho_1)), \quad (40)$$

and

$$\frac{1}{\sqrt{r}}(Q_1 - \frac{1 - \pi}{\pi}\rho_2) \xrightarrow[r \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \frac{1 - \pi}{\pi}\rho_2(1 - \rho_2) + \rho_2^2 \frac{1 - \pi}{\pi^2}) \quad (41)$$

are a direct consequence of the central limit theorem. The proof of (41) uses (35). With a similar argument one can show

$$\frac{1}{\sqrt{r}} \sum_{\nu=1}^r (Q_{1[1],\nu} - (N_{[1],\nu} - 1)\rho_2) \xrightarrow[r \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \frac{1 - \pi}{\pi}\rho_2(1 - \rho_2)). \quad (42)$$

To establish the joint distribution of the three quantities, we apply the Crámer-Wold device (see Lehmann (1999)), using

$$\begin{aligned} & \mathbb{E} \left[ (Q_{1[1],\nu} - \frac{1 - \pi}{\pi}\rho_2)(N_{[1],\nu} - \frac{1}{\pi}) \right] \\ &= \mathbb{E} \left[ (Q_{1[1],\nu} - (N_{[1],\nu} - 1)\rho_2)(N_{[1],\nu} - \frac{1}{\pi}) \right] + \rho_2 \mathbb{E} \left[ (N_{[1],\nu} - \frac{1}{\pi})^2 \right] \\ &= \rho_2 \mathbb{E} \left[ (N_{[1],\nu} - \frac{1}{\pi})^2 \right] = \frac{1 - \pi}{\pi^2} \rho_2, \end{aligned}$$

and

$$\begin{aligned} & \mathbb{E} \left[ a(Q_{1[1],\nu} - \frac{1 - \pi}{\pi}\rho_2) + b(N_{[1],\nu} - \frac{1}{\pi}) \right]^2 \\ &= a^2 \mathbb{E} \left[ Q_{1[1],\nu} - \frac{1 - \pi}{\pi}\rho_2 \right]^2 + b^2 \mathbb{E} \left[ N_{[1],\nu} - \frac{1}{\pi} \right]^2 + ab \mathbb{E} \left[ (Q_{1[1],\nu} - \frac{1 - \pi}{\pi}\rho_2)(N_{[1],\nu} - \frac{1}{\pi}) \right] \\ &= a^2 \left[ \frac{1 - \pi}{\pi}\rho_2(1 - \rho_2) + \rho_2^2 \frac{1 - \pi}{\pi^2} \right] + b^2 \frac{1 - \pi}{\pi^2} + ab \frac{1 - \pi}{\pi^2} \rho_2. \end{aligned}$$

**Derivation of (18):** First, simple algebra shows that  $\sqrt{N_1}(\hat{p}_1 - p_1)$  equals

$$\sqrt{\frac{r}{N_1}} \left[ (1 - p_1)\sqrt{r} \left( \frac{1}{r}(R_1 - \rho_1) - p_1 \right) - p_1 \sqrt{r} \left( \frac{1}{r}Q_1 - \frac{1 - \pi}{\pi}\rho_2 \right) \right], \quad (43)$$

because

$$\begin{aligned}
\sqrt{N_1}(\hat{p}_1 - p_1) &= \sqrt{\frac{N_1}{r}}\sqrt{r} \left[ \frac{R_1}{r} - \rho_1 + p_1 \left( \frac{q}{\frac{N_1}{r}} - 1 \right) \right] \\
&= \sqrt{\frac{r}{N_1}}\sqrt{r} \left[ \frac{R_1}{r} - \rho_1 - p_1 \left( \frac{N_1}{r} - \frac{q}{\pi} \right) \right] \\
&= \sqrt{\frac{r}{N_1}}\sqrt{r} \left[ \frac{R_1}{r} - \rho_1 - p_1 \left( \frac{R_1}{r} - \rho_1 + \frac{Q_1}{r} - \frac{q(1-p_1)}{\pi} \right) \right] \\
&= \sqrt{\frac{r}{N_1}} \left[ (1-p_1)\sqrt{r} \left( \frac{R_1}{r} - \rho_1 \right) - p_1\sqrt{r} \left( \frac{Q_1}{r} - \rho_2 \frac{1-\pi}{\pi} \right) \right].
\end{aligned}$$

Now, Lemma 6.1, (38), (40), and (41) imply that the asymptotic limit distribution of  $\sqrt{N_1}(\hat{p}_1 - p_1)$  is normal with mean 0 and variance

$$\frac{\pi}{q} \left[ (1-p_1)^2 \rho_1(1-\rho_1) + p_1^2 \frac{1-\pi}{\pi} \rho_2(1-\rho_2) + p_1^2 \frac{1-\pi}{\pi^2} \rho_2^2 \right] \quad (44)$$

Finally, the following chain of equalities then completes the proof:

$$\begin{aligned}
&\frac{\pi}{q} \left[ (1-p_1)^2 \rho_1(1-\rho_1) + p_1^2 \frac{1-\pi}{\pi} \rho_2(1-\rho_2) + p_1^2 \frac{1-\pi}{\pi^2} \rho_2^2 \right] \\
&= \frac{\pi}{q} \left[ \frac{(1-p_1)^2 p_1 p_0 q (1-q)}{\pi^2} + \frac{p_1^2 (1-p_1)(1-p_0)q(1-q)}{\pi(1-\pi)} + \frac{1-\pi}{\pi^2} \frac{p_1^2 (1-p_1)^2 q^2}{(1-\pi)^2} \right] \\
&= p_1(1-p_1) \left[ \frac{(1-p_1)p_0(1-q)}{\pi} + \frac{p_1(1-p_0)(1-q)}{1-\pi} + \frac{p_1(1-p_1)q}{\pi(1-\pi)} \right] \\
&= p_1(1-p_1) \left[ (1-p_1)(1-\rho_1) + p_1(1-\rho_2) + p_1\rho_2 \frac{1}{\pi} \right] \\
&= p_1(1-p_1) \left[ 1 + p_1(\rho_1 - \rho_2) - \rho_1 + p_1\rho_2 \frac{1}{\pi} \right] \\
&= p_1(1-p_1) \left[ 1 + p_1 \frac{(1-\pi)p_1q - \pi(1-p_1)q + q(1-p_1)}{\pi(1-\pi)} - \rho_1 \right] \\
&= p_1(1-p_1) \left[ 1 + p_1 \frac{q(1-\pi)}{\pi(1-\pi)} - \rho_1 \right] = p_1(1-p_1)
\end{aligned}$$

**Proof of Corollary 6.3:** To start, consider a first order Taylor expansion

$$\begin{aligned}
& \sqrt{r} \left( \log(\hat{\theta}) - \log(\theta) \right) \\
&= \sqrt{r} \left( \log\left(\frac{1}{r}R_1\right) - \log(\rho_1) \right) - \sqrt{r} \left( \log\left(1 - \frac{1}{r}R_1\right) - \log(1 - \rho_1) \right) \\
&\quad - \sqrt{r} \left( \log\left(\frac{1}{r}Q_1\right) - \log(\rho_2) \right) + \sqrt{r} \left( \log\left(\frac{1}{r}N - 1 - \frac{1}{r}Q_1\right) - \log(1 - \rho_2) \right) \\
&= \sqrt{r} \left( \log\left(\frac{1}{r}R_1\right) - \log(\rho_1) \right) - \sqrt{r} \left( \log\left(1 - \frac{1}{r}R_1\right) - \log(1 - \rho_1) \right) \\
&\quad - \sqrt{r} \left( \log\left(\frac{1}{r}Q_1\right) - \log\left(\frac{1-\pi}{\pi}\rho_2\right) \right) + \sqrt{r} \left( \log\left(\frac{1}{r}N - 1 - \frac{1}{r}Q_1\right) - \log\left(\frac{1-\pi}{\pi}(1 - \rho_2)\right) \right) \\
&\approx \frac{1}{\rho_1} \sqrt{r} \left( \frac{1}{r}R_1 - \rho_1 \right) + \frac{1}{1 - \rho_1} \sqrt{r} \left( \frac{1}{r}R_1 - \rho_1 \right) - \frac{\pi}{(1 - \pi)\rho_2} \sqrt{r} \left( \frac{1}{r}Q_1 - \frac{1 - \pi}{\pi}\rho_2 \right) \\
&\quad + \frac{\pi}{(1 - \pi)(1 - \rho_2)} \sqrt{r} \left( \frac{1}{r}N - 1 - \frac{1 - \pi}{\pi} \right) - \frac{\pi}{(1 - \pi)(1 - \rho_2)} \sqrt{r} \left( \frac{1}{r}Q_1 - \frac{1 - \pi}{\pi}\rho_2 \right) \\
&= \left( \frac{1}{\rho_1(1 - \rho_1)}, -\frac{\pi}{1 - \pi} \frac{1}{\rho_2(1 - \rho_2)}, \frac{\pi}{(1 - \pi)(1 - \rho_2)} \right)^T \sqrt{r} \begin{pmatrix} \frac{1}{r}R_1 - \rho_1 \\ \frac{1}{r}Q_1 - \frac{1 - \pi}{\pi}\rho_2 \\ \frac{1}{r}N - \frac{1}{\pi} \end{pmatrix} \\
&=: \boldsymbol{\xi}^T \sqrt{r} \begin{pmatrix} \frac{1}{r}R_1 - \rho_1 \\ \frac{1}{r}Q_1 - \frac{1 - \pi}{\pi}\rho_2 \\ \frac{1}{r}N - \frac{1}{\pi} \end{pmatrix}
\end{aligned}$$

Now Theorem 6.2 implies

$$\sqrt{r} \left( \log(\hat{\theta}) - \log(\theta) \right) \xrightarrow[r \rightarrow \infty]{\mathcal{D}} \mathcal{N}(0, \boldsymbol{\xi}^T \boldsymbol{\Sigma} \boldsymbol{\xi}) \quad (45)$$

with  $\boldsymbol{\Sigma}$  defined as in (17). Simple algebra provides

$$\boldsymbol{\xi}^T \boldsymbol{\Sigma} \boldsymbol{\xi} = \boldsymbol{\xi}^T \begin{pmatrix} \rho_1(1 - \rho_1) & 0 & 0 \\ 0 & \frac{1 - \pi}{\pi} \rho_2(1 - \rho_2) & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{\xi} = \frac{1}{\rho_1(1 - \rho_1)} + \frac{\pi}{1 - \pi} \frac{1}{\rho_2(1 - \rho_2)},$$

i.e. (21). The proof of (22) follows directly from Lemma 6.1.

**Proof of Theorem 6.4:** First note that one can show

$$\left( \frac{1}{r}R_1, \frac{1}{r}Q_1, \frac{1}{r}N \right) \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r, \mathbf{p}_0)} \left( q, \frac{1 - p_0}{p_0} q, \frac{1}{p_0} \right). \quad (46)$$

in exactly the same way as shown in the proof of Lemma 6.1. The notation  $\xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r, \mathbf{p}_0)}$  stands for convergence in probability when  $(\lambda_r, p_0)$  are the true underlying parameters.



Similarly, with obvious notation,

$$\sqrt{r} \begin{pmatrix} \frac{1}{r} R_1 - \rho_1(\lambda_r) \\ \frac{1}{r} Q_1 - \frac{1 - \pi(\lambda_r, p_0)}{\pi(\lambda_r, p_0)} \rho_2(\lambda_r, p_0) \\ \frac{1}{r} N - \frac{1}{\pi(\lambda_r, p_0)} \end{pmatrix} \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r, p_0)} \mathcal{N}(\mathbf{0}_3, \boldsymbol{\Sigma}_0), \quad (47)$$

is a direct consequence of the central limit theorem, just like demonstrated in the proof of Lemma 6.2. The matrix  $\boldsymbol{\Sigma}_0$  is defined as

$$\boldsymbol{\Sigma}_0 = \begin{pmatrix} q(1 - q) & 0 & 0 \\ 0 & \frac{1 - p_0}{p_0^2} [p_0 q(1 - q) + q^2] & \frac{1 - p_0}{p_0^2} q \\ 0 & \frac{1 - p_0}{p_0^2} q & \frac{1 - p_0}{p_0^2} \end{pmatrix}. \quad (48)$$

Define  $\boldsymbol{\xi}^T = (0, 1, -q)$  to conclude

$$\boldsymbol{\xi}^T \sqrt{r} \begin{pmatrix} \frac{1}{r} R_1 - \rho_1(\lambda_r) \\ \frac{1}{r} Q_1 - \frac{1 - \pi(\lambda_r, p_0)}{\pi(\lambda_r, p_0)} \rho_2(\lambda_r, p_0) \\ \frac{1}{r} N - \frac{1}{\pi(\lambda_r, p_0)} \end{pmatrix} \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r, p_0)} \mathcal{N}(\mathbf{0}_3, \boldsymbol{\xi}^T \boldsymbol{\Sigma}_0 \boldsymbol{\xi}) = \mathcal{N}\left(0, \frac{1 - p_0}{p_0^2} [p_0 q(1 - q)]\right), \quad (49)$$

and hence

$$\sqrt{r} \left[ \frac{1}{r} Q_1 - \frac{1 - \pi}{\pi} \rho_2 - \left( \frac{1}{r} N - 1 - \frac{1 - \pi}{\pi} \right) \rho_2 \right] \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r, p_0)} \mathcal{N}\left(0, \frac{1 - p_0}{p_0} [q(1 - q)]\right), \quad (50)$$

because of  $\rho_2(\lambda_r, p_0) \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r, p_0)} q$ .

Let's move to the first marginal distribution of (23), namely

$$\sqrt{r} \frac{\bar{R}_1 - q}{\sqrt{\bar{R}_1(1 - \bar{R}_1)}} = \sqrt{r} \frac{\bar{R}_1 - \rho_1(\lambda_r)}{\sqrt{ar \bar{R}_1(1 - \bar{R}_1)}} - \sqrt{r} \frac{\rho_1(1) - \rho_1(\lambda_r)}{\sqrt{\bar{R}_1(1 - \bar{R}_1)}}. \quad (51)$$

Formulae (46) and (47) imply that the first summand of the right hand side of (51) converges in distribution to a standard normal distribution, and the second summand converges in probability to  $\sqrt{1 - p_0} t \Delta$ , both under the assumption that  $\lambda_r$  is the true parameter sequence. We use

$$\sqrt{r} \frac{\rho_1(1) - \rho_1(\lambda_r)}{\sqrt{\bar{R}_1(1 - \bar{R}_1)}} = \sqrt{r} \frac{q - \rho_1(\lambda_r)}{\sqrt{\bar{R}_1(1 - \bar{R}_1)}} = \frac{q(1 - q)t}{\left(1 - \frac{tq}{\sqrt{r}}\right) \sqrt{\bar{R}_1(1 - \bar{R}_1)}} \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r)} t \sqrt{q(1 - q)}, \quad (52)$$

and  $\Delta = \sqrt{\frac{q(1 - q)}{1 - p_0}}$  to obtain the limit of the second summand, and to conclude

$$\frac{\sqrt{r}(\bar{R}_1 - q)}{\sqrt{\bar{R}_1(1 - \bar{R}_1)}} \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r)} \mathcal{N}(-\sqrt{1 - p_0} t \Delta, 1). \quad (53)$$

Let's turn to the second marginal distribution in (23), i.e.

$$\sqrt{N-r} \frac{\bar{Q}_1 - q}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}}, \quad (54)$$

which is equivalent to

$$\sqrt{N-r} \frac{\bar{Q}_1 - q}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}} = \sqrt{N-r} \frac{\bar{Q}_1 - \rho_2(\lambda, p_0)}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}} - \sqrt{N-r} \frac{q - \rho_2(\lambda, p_0)}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}}. \quad (55)$$

We have

$$\begin{aligned} \sqrt{N-r} [\bar{Q}_1 - \rho_2] &= \sqrt{\frac{r}{N-r}} \sqrt{r} \left[ \frac{1}{r} Q_1 - \left( \frac{1}{r} N - 1 \right) \rho_2 \right] \\ &= \sqrt{\frac{r}{N-r}} \sqrt{r} \left[ \frac{1}{r} Q_1 - \frac{1-\pi}{\pi} \rho_2 - \left[ \frac{1}{r} N - 1 - \frac{1-\pi}{\pi} \right] \rho_2 \right]. \end{aligned}$$

This,

$$\bar{Q}_1 = \frac{r}{N-r} \frac{Q_1}{r} \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r, p_0)} = q, \quad (56)$$

and (46) together with (50), imply

$$\sqrt{N-r} \frac{\bar{Q}_1 - \rho_2(\lambda, p_0)}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}} \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r, p_0)} \mathcal{N}(0, 1). \quad (57)$$

Note that (56) follows from (46). Statement (56) also implies

$$\frac{\frac{1}{r} N - 1}{\bar{Q}_1(1-\bar{Q}_1)} \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r, p_0)} = \frac{1-p_0}{p_0 q(1-q)}. \quad (58)$$

Next,

$$\begin{aligned} \sqrt{N-r} \frac{q - \rho_2(\lambda, p_0)}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}} &= \sqrt{\frac{N-r}{r}} \sqrt{r} \frac{q - \rho_2(\lambda, p_0)}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}} \\ &= \sqrt{\frac{\frac{1}{r} N - 1}{\bar{Q}_1(1-\bar{Q}_1)} \frac{-q(1-q)tp_0}{1-p_0 + \frac{qp_0 t}{\sqrt{r}}}} \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r)} -\sqrt{p_0} t \Delta \end{aligned}$$

to conclude the proof of the second marginal distribution in (23), i.e.

$$\frac{\sqrt{r}(\bar{Q}_1 - q)}{\sqrt{\bar{Q}_1(1-\bar{Q}_1)}} \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r)} \mathcal{N}(\sqrt{p_0} t \Delta, 1) \quad (59)$$

Asymptotic independence of the two marginal distributions follows from the Crámer-Wold device (see Lehmann (1999)).

Let's now turn to the proof of (24). A similar Taylor expansion as in the proof of Corollary 6.3 yields

$$\sqrt{r} \left( \log(\hat{\theta}) - \log(\theta(\lambda_r, p_0)) \right) \approx \boldsymbol{\xi}_r^T \sqrt{r} \begin{pmatrix} \frac{1}{r} R_1 - \rho_1(\lambda_r) \\ \frac{1}{r} Q_1 - \frac{1 - \pi(\lambda_r, p_0)}{\pi(\lambda_r, p_0)} \rho_2(\lambda_r, p_0) \\ \frac{1}{r} N - \frac{1}{\pi(\lambda_r, p_0)} \end{pmatrix} \quad (60)$$

with

$$\boldsymbol{\xi}_r = \begin{pmatrix} \frac{1}{\rho_1(\lambda_r)(1 - \rho_1(\lambda_r))} \\ -\frac{\pi(\lambda_r, p_0)}{1 - \pi(\lambda_r, p_0)} \frac{1}{\rho_2(\lambda_r, p_0)(1 - \rho_2(\lambda_r, p_0))} \\ \frac{\pi(\lambda_r, p_0)}{(1 - \pi(\lambda_r, p_0))(1 - \rho_2(\lambda_r, p_0))} \end{pmatrix} \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r)} \begin{pmatrix} \frac{1}{q(1 - q)} \\ -\frac{1}{1 - p_0} \frac{1}{q(1 - q)} \\ \frac{p_0}{1 - p_0} \frac{1}{1 - q} \end{pmatrix}. \quad (61)$$

Hence (60), (61), and (47) imply

$$\sqrt{r} \left( \log(\hat{\theta}) - \log(\theta(\lambda_r, p_0)) \right) \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r)} \mathcal{N}\left(0, \frac{1}{1 - p_0} \cdot \frac{1}{q(1 - q)}\right). \quad (62)$$

From (46) we conclude that

$$\hat{\sigma}^2 = \frac{r}{R_1} + \frac{r}{Q_1} + \frac{r}{r - R_1} + \frac{r}{N - r - Q_1} \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r)} \frac{1}{q(1 - q)} + \frac{p_0}{1 - p_0} \frac{1}{q(1 - q)}, \quad (63)$$

so that

$$\sqrt{r} \frac{\log(\hat{\theta}) - \log(\theta(\lambda_r, p_0))}{\hat{\sigma}} \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r)} \mathcal{N}(0, 1). \quad (64)$$

Next, a further Taylor expansion provides

$$\sqrt{r} (\log(\theta(\lambda_r, p_0)) - \log(\theta)) = \frac{1}{q(1 - q)} (\rho_1(\lambda_r, p_0) - q) - \frac{1}{q(1 - q)} (\rho_1(\lambda_r, p_0) - q) + o(1) \quad (65)$$

This uses  $\rho_1(1) = \rho_2(1, p_0) = q$ . By definition

$$\sqrt{r} (\rho_1(\lambda_r) - q) = -t \frac{q(1 - q)}{1 - (1 - \frac{qt}{\sqrt{r}}) p_0} \xrightarrow[r \rightarrow \infty]{} -tq(1 - q) \quad (66)$$

and

$$\sqrt{r} (\rho_2(\lambda_r, p_0) - q) = tp_0 \frac{q(1 - q)}{1 - \frac{qt}{\sqrt{r}}} \xrightarrow[r \rightarrow \infty]{} t \frac{p_0}{1 - p_0} q(1 - q), \quad (67)$$

so that

$$\sqrt{r} \frac{\log(\theta(\lambda_r, p_0)) - \log(\theta)}{\hat{\sigma}} \xrightarrow[r \rightarrow \infty]{\mathbf{P}(\lambda_r)} -t \sqrt{\frac{q(1 - q)}{1 - p_0}} = -t\Delta \quad (68)$$

Now (65) and (68) imply (24), which concludes the proof of Theorem 6.4.

In Theorem 6.4 we used the conditional distribution of  $Q_1$  given  $N$  to motivate the test statistic 54. If we were to use the marginal distribution of  $Q_1$  as given in (5), then one would need to use a test statistic like

$$\sqrt{r} \left( \frac{1}{r} Q_1 - \frac{1 - \rho_3(1, p_0)}{\rho_3(1, p_0)} \right) \quad (69)$$

This is a theoretical test statistic, since  $p_0$  is an unknown parameter and would need to be replaced by an estimate. Now, since  $\frac{1}{r}N - 1 - \frac{1-\pi}{\pi} \xrightarrow[r \rightarrow \infty]{P(\lambda_r)} 0$ , and since

$$\frac{1 - \rho_3}{\rho_3} = \frac{1 - \pi}{\pi} \rho_2 \quad (70)$$

a practical estimate would be

$$\sqrt{r} \left( \frac{1}{r} Q_1 - \left( \frac{1}{r} N - 1 \right) \rho_2(1, p_0) \right) = \sqrt{r} \left( \frac{1}{r} Q_1 - \left( \frac{1}{r} N - 1 \right) q \right). \quad (71)$$

The asymptotic distribution of this estimate is

$$\sqrt{r} \left( \frac{1}{r} Q_1 - \left( \frac{1}{r} N - 1 \right) \rho_2(1, p_0) \right) \xrightarrow[r \rightarrow \infty]{\mathcal{D}(\lambda_r)} \mathcal{N}(tq(1 - q), \frac{1 - p_0}{p_0} [q(1 - q)]). \quad (72)$$

This follows from

$$\begin{aligned} & \sqrt{r} \left( \frac{1}{r} Q_1 - \left( \frac{1}{r} N - 1 \right) \rho_2(1, p_0) \right) \\ &= \sqrt{r} \left( \frac{1}{r} Q_1 - \frac{1 - \pi(\lambda_r, p_0)}{\pi(\lambda_r, p_0)} \rho_2(\lambda_r, p_0) \right) - \sqrt{r} \left( \left( \frac{1}{r} N - 1 - \frac{1 - \pi(\lambda_r, p_0)}{\pi(\lambda_r, p_0)} \right) \rho_2(\lambda_r, p_0) \right) \\ & \quad + \sqrt{r} \left( \left( \frac{1}{r} N - 1 \right) (\rho_2(\lambda_r, p_0) - q) \right) \end{aligned}$$

From (46) we get

$$\sqrt{r} \left( \left( \frac{1}{r} N - 1 \right) (\rho_2(\lambda_r, p_0) - q) \right) \xrightarrow[r \rightarrow \infty]{P(\lambda_r)} \left( \frac{1 - p_0}{p_0} \right) \left( \frac{tq(1 - q)p_0}{1 - p_0} \right), \quad (73)$$

and from (50) we conclude that the first summand converges to a normal distribution with mean 0 and variance  $\frac{1-p_0}{p_0} [q(1-q)]$ . Overall, an appropriately standardized version  $\sqrt{\frac{r}{\hat{\sigma}^2}} \left( \frac{1}{r} Q_1 - \left( \frac{1}{r} N - 1 \right) q \right)$  of our estimate is then asymptotically normally distributed with variance 1 and mean  $\sqrt{p_0} t \Delta$ .

If one would just calculate the asymptotic distribution of the ideal test  $\sqrt{r} \left( \frac{1}{r} Q_1 - \frac{1-p_0}{p_0} q \right)$  under  $\lambda_r$  and  $p_0$ , pretending one knew  $p_0$ , then the asymptotic distribution would be a normal distribution with variance 1 and mean  $\sqrt{p_0} t \sqrt{\frac{q(1-q)}{1-p_0} + \frac{q^2}{p_0(1-p_0)}}$ . This of course dominates the practical test.

### 8.3 Sample Size and Asymptotic Relative Efficiency

Under inverse sampling, the log odds ratio test for  $H_0 : \lambda \geq 1$  rejects if  $\sqrt{r} \frac{\log(\hat{\theta})}{\hat{\sigma}} < u_\alpha$ , with  $\hat{\sigma}^2 = \frac{r}{R_1} + \frac{r}{Q_1} + \frac{r}{r-R_1} + \frac{r}{N-r-Q_1}$  (see Lemma 6.3). The corresponding asymptotic power function is equal to

$$\Phi \left( u_\alpha - \sqrt{r} \frac{\log(\theta)}{\sigma_{IS}} \right) \quad (74)$$

where  $\theta = \theta(\lambda, p_0)$  is a function of  $\lambda$  and  $p_0$ , and  $\sigma_{IS}^2$  is defined in (21). From this one can immediately derive a formula for the required number of cases under inverse sampling:

$$r = \sigma_{IS}^2 \frac{[u_\alpha + u_\beta]^2}{[\log(\theta)]^2}. \quad (75)$$

Note that  $\sigma_{IS}^2 = \sigma_{IS}^2(\lambda, p_0)$  is also a function of  $\lambda$  and  $p_0$ .

Next, under conventional sampling we have

$$\sqrt{n} \left( \log(\hat{\theta}) - \log(\theta) \right) \xrightarrow[n \rightarrow \infty]{\mathcal{D}(\theta)} \mathcal{N}(\mathbf{0}, \sigma_{CS}^2), \quad (76)$$

where

$$\sigma_{CS}^2 = \frac{1}{qp_1(1-p_1)} + \frac{1}{(1-q)p_0(1-p_0)} \quad (77)$$

is also a function of  $\lambda$  and  $p_0$ . The above two statements are true because

$$\log(\hat{\theta}) = \log\left(\frac{R_1}{R-R_1} \frac{n-R-Q_1}{Q_1}\right) = \log\left(\frac{\frac{1}{n_1}R_1(1-\frac{1}{n_0}R_0)}{\frac{1}{n_0}R_0(1-\frac{1}{n_1}R_1)}\right), \quad (78)$$

so that a Taylor expansion along the same lines as before implies

$$\sqrt{n}(\log(\hat{\theta}) - \log(\theta)) = \frac{1}{p_1(1-p_1)} \sqrt{n}(\hat{p}_1 - p_1) + \frac{1}{p_0(1-p_0)} \sqrt{n}(\hat{p}_0 - p_0). \quad (79)$$

From the standard central limit theorem we obtain (76) and (77).

Hence, under conventional sampling the log odds ratio test for  $H_0 : \lambda \geq 1$  rejects if  $\sqrt{n} \frac{\log(\hat{\theta})}{\hat{\sigma}_{CS}} < u_\alpha$ , with  $\hat{\sigma}_{CS}^2 = \frac{n}{R_1} + \frac{n}{n_1-R_1} + \frac{n}{R_0} + \frac{n}{n_0-R_0}$  being the consistent estimator for  $\sigma_{CS}^2$  under conventional sampling. The corresponding asymptotic power function is equal to

$$\Phi \left( u_\alpha - \sqrt{n} \frac{\log(\theta)}{\sigma_{CS}} \right). \quad (80)$$

From this one can immediately derive a formula for the required sample size under conventional sampling:

$$n = \sigma_{CS}^2 \frac{[u_\alpha + u_\beta]^2}{[\log(\theta)]^2}. \quad (81)$$

Now assume that one wants to detect a relevant difference  $\lambda_* = 0.5$ , say. We assume a placebo incidence rate of  $p_{0*} = 0.001$ ,  $\alpha = 0.025$ , and  $\beta = 0.1$ . Under inverse sampling, applying formula (75), one gets

$$\sigma_{IS}^2(\lambda_*, p_{0*}) \frac{[u_\alpha + u_\beta]^2}{[\log(\theta(\lambda_*, p_{0*}))]^2} \approx 89.23674, \quad (82)$$

so that the required number of cases is  $r_* = 90$ . The corresponding expected sample size (assuming  $\lambda_*$  and  $p_{0*}$  are the true underlying parameters) is

$$N_* = \mathbb{E}_{\lambda_*, p_{0*}}[N|r_*] = \frac{90}{\pi(\lambda_*, p_{0*})} \approx 128571.4. \quad (83)$$

Under conventional sampling one gets

$$\sigma_{CS}^2(\lambda_*, p_{0*}) \frac{[u_\alpha + u_\beta]^2}{[\log(\theta(\lambda_*, p_{0*}))]^2} \approx 127481.1, \quad (84)$$

so that the required sample size is  $n_* = 127482$ . At a first glance, this number differs from the expected sample size under inverse sampling, namely 128571.4, but this is due to rounding errors. More precisely,

$$\mathbb{E}_{\lambda_*, p_{0*}}[N|\sigma_{IS}^2(\lambda_*, p_{0*}) \frac{[u_\alpha + u_\beta]^2}{[\log(\theta(\lambda_*, p_{0*}))]^2}] \approx \frac{89.23674}{\pi(\lambda_*, p_{0*})} = 127481.1. \quad (85)$$

Note that  $\log(\theta) = \log\left(\frac{p_1(1-p_0)}{p_0(1-p_1)}\right) = \log\left(\frac{\lambda p_0(1-p_0)}{p_0(1-\lambda p_0)}\right) \log\left(\frac{\rho_1(\lambda, p_0)(1-\rho_2(\lambda, p_0))}{\rho_2(\lambda, p_0)(1-\rho_1(\lambda, p_0))}\right)$  for all values of  $(\lambda, p_0)$ . Moreover,  $\frac{\sigma_{IS}^2}{\pi\sigma_{CS}^2} = 1$  holds for all values of  $\lambda$  and  $p_0$ , which is why (84) and (85) provide identical numbers.

The equality  $\frac{\sigma_{IS}^2}{\pi\sigma_{CS}^2} = 1$  holds because of

$$\sigma_{CS}^2 = \frac{(1-q)p_0(1-p_0) + qp_1(1-p_1)}{q(1-q)p_0(1-p_0)p_1(1-p_1)}, \quad (86)$$

$$\frac{\sigma_{IS}^2}{\pi} = \frac{\pi(1-p_1)(1-p_0) + (1-\pi)p_1p_0}{q(1-q)p_0(1-p_0)p_1(1-p_1)}, \quad (87)$$

and because of

$$\pi(1-p_1)(1-p_0) + (1-\pi)p_1p_0 = (1-q)p_0(1-p_0) + qp_1(1-p_1). \quad (88)$$

We can define the **relative efficiency** of conventional to inverse sampling as the ratio of the expected sample size under inverse sampling and the required sample size under conventional sampling, i.e.

$$\frac{\mathbb{E}_{\lambda_*, p_{0*}}[N|r_*]}{n_*} = \frac{\sigma_{IS}^2(\lambda_*, p_{0*})}{\pi(\lambda_*, p_{0*})\sigma_{CS}^2(\lambda_*, p_{0*})}. \quad (89)$$

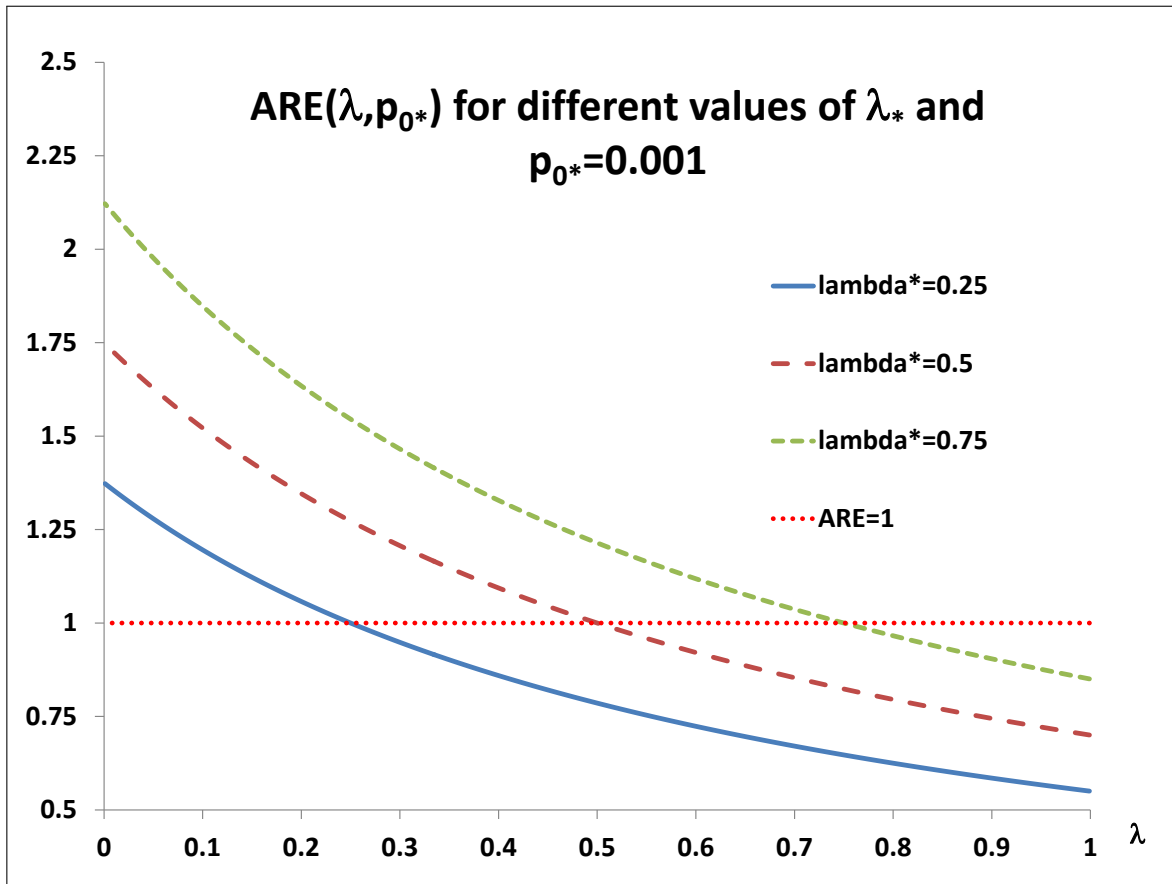


Figure 6: Relative efficiency as a function of  $\lambda$  for three different values of  $\lambda_*$ . Areas where the ARE is below 1 indicate where inverse sampling is more efficient than conventional sampling.

As just shown, this equals 1. However, if the real value of  $\lambda$  and  $p_0$  deviates from the values  $\lambda_* = 0.5$  and  $p_{0*} = 0.001$  which were used to obtain  $r_*$  or  $n_*$ , then the relative efficiency is

$$\frac{\mathbb{E}_{\lambda, p_0}[N|r_*]}{n_*} = \frac{\sigma_{IS}^2(\lambda_*, p_{0*})}{\pi(\lambda, p_0)\sigma_{CS}^2(\lambda_*, p_{0*})} = \frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_0)}. \quad (90)$$

This calculation is approximate and does not take the rounding of  $r_*$  and  $n_*$  to integer (or even integer) values into account!

In Figure 6 we present the relative efficiency for  $p_{0*} = 0.001$  and three different values of  $\lambda_*$ . One can see that the function  $\lambda \rightarrow \frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_0)}$  is always below 1 for  $\lambda > \lambda_*$ , and above one for  $\lambda < \lambda_*$ . This indicates that the expected sample size under inverse sampling is smaller (larger) than the sample size under conventional sampling if the true parameter  $\lambda$  is larger (smaller) than the value  $\lambda_*$  used for sample size calculations.

In principle, it is good if the expected sample size under inverse sampling is less than the

sample size under conventional sampling. Unfortunately, the smaller expected sample size in certain regions of the parameter space goes along with a loss of power in the same regions of the parameter space.

This point can be seen from the following argument. The power function for the log odds ratio test under conventional sampling for a given  $n_*$  can be approximated by

$$\Phi \left( u_\alpha - [u_\alpha + u_\beta] \frac{\sigma_{CS}(\lambda_*, p_{0*})}{\sigma_{CS}(\lambda, p_0)} \frac{\log(\theta(\lambda, p_0))}{\log(\theta(\lambda_*, p_{0*}))} \right). \quad (91)$$

Correspondingly, the power function for inverse sampling with  $r_*$  fixed is approximately

$$\Phi \left( u_\alpha - [u_\alpha + u_\beta] \frac{\sigma_{IS}(\lambda_*, p_{0*})}{\sigma_{IS}(\lambda, p_0)} \frac{\log(\theta(\lambda, p_0))}{\log(\theta(\lambda_*, p_{0*}))} \right). \quad (92)$$

Given that

$$\frac{\sigma_{IS}^2(\lambda, p_0)}{\pi(\lambda, p_0)\sigma_{CS}^2(\lambda, p_0)} = 1 \quad (93)$$

for all  $\lambda, p_0$ , one can easily see that the power function of the log odds ratio test for inverse sampling is approximately

$$\Phi \left( u_\alpha - [u_\alpha + u_\beta] \frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_0)} \frac{\sigma_{CS}(\lambda_*, p_{0*})}{\sigma_{CS}(\lambda, p_0)} \frac{\log(\theta(\lambda, p_0))}{\log(\theta(\lambda_*, p_{0*}))} \right). \quad (94)$$

It differs from that under conventional sampling only by the factor

$$\frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_0)}, \quad (95)$$

This relates the improvement in relative efficiency to the corresponding loss in power.

Note that these are asymptotic calculations, and that we again did not take the rounding into account (i.e. we treated  $r_*$  and  $n_*$  as real numbers, not as integers or even integers).

From these figures one can see the following: After we fix all the parameters for sample size calculations, including the relevant effect  $\lambda_*$  and the nuisance parameter  $p_{0*}$ , one can calculate a sample size  $n_*$  for conventional sampling, or the required number of cases  $r_*$  for inverse sampling. The expected sample size for inverse sampling  $N_* = \mathbb{E}_{\lambda_*, p_{0*}}[N|r_*]$  and the sample size  $n_*$  are the same, as long as one calculates the expected sample size under inverse sampling using the parameters  $\lambda_*$  and  $p_{0*}$  used for sample size calculations, and as long as one ignores impression due to rounding. From this perspective there is no difference between inverse and conventional sampling. However, the parameters  $\lambda_*$  and  $p_{0*}$  are design parameters, and do not necessarily reflect reality. If one calculates the relative efficiency of inverse versus conventional sampling with fixed  $n_*$  and  $r_*$  under general parameter values  $\lambda$  and  $p_0$ , then we start to see differences, because the sample



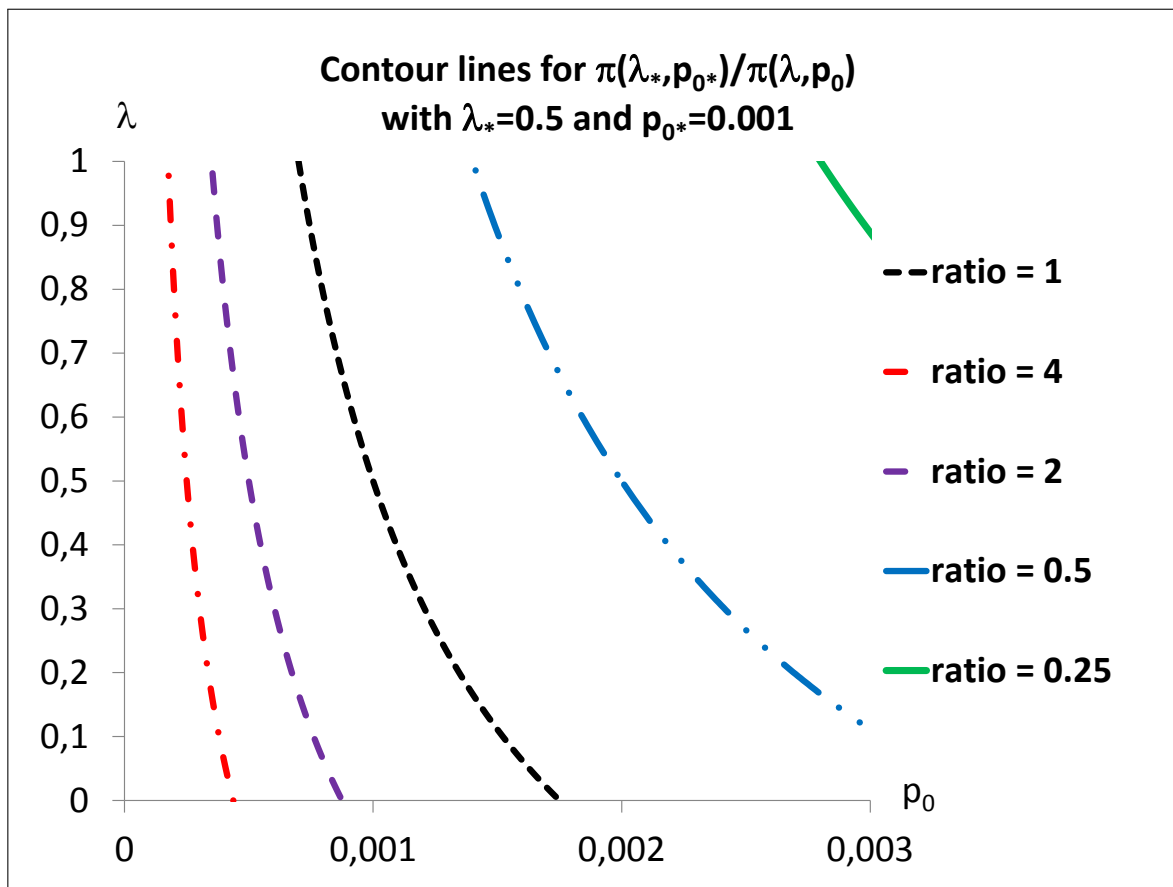


Figure 7: Relative efficiency as a function of  $p_0$  and  $\lambda$  for three values of  $p_{0*}$  and  $\lambda_* = 0.5$ . Areas where the ARE is below 1 indicate where inverse sampling is more efficient than conventional sampling.

size  $n_*$  is fixed, whilst the expected sample size  $\mathbb{E}_{\lambda, p_0}[N|r_*]$  is dependent on the true parameter values. If the true  $\lambda$  is smaller than the relevant effect  $\lambda_*$ , then the relative efficiency of inverse versus conventional sampling is smaller than 1, indicating less cost for inverse sampling. However, this comes with a decrease in power.

Contour lines for the ratio  $\frac{\pi(\lambda_*, p_{0*})}{\pi(\lambda, p_0)}$  are presented in Figure 7.

## References

- AGGARWAL, A. & PANDEY, A. (2010). Inverse sampling to study disease burden of leprocy. *Indian Journal of Medical Research* **132**, 438–441.
- BOLLAND, K. & WHITEHEAD, J. (2000). Formal approaches to safety monitoring of clinical trials in life-threatening conditions. *Statistics in Medicine* **19**, 2899–2917.
- CHAN, I. S. & BOHIDAR, N. R. (1998). Exact power and sample size for vaccine efficacy studies. *Communications in Statistics-Theory and Methods* **27**, 1305–1322.
- HALDANE, J. B. (1945a). A labour-saving method of sampling. *Nature* **155**, 49–50.
- HALDANE, J. B. (1945b). On a method of estimating frequencies. *Biometrika* **33**, 222–225.
- JENNISON, C. & TURNBULL, B. W. (1999). *Group Sequential Methods with Applications to Clinical Trials*. Chapman and Hall, 1st ed.
- LAN, K.-K. G., SIMON, R. & HALPERIN, M. (1982). Stochastically curtailed tests in long-term clinical trials. *Communications in Statistics, Part C* **1**, 207–219.
- LEHMANN, E. L. (1999). *Elements of Large Sample Theory*. Springer Verlag, 1st ed.
- LUI, K.-J. (1995). Notes on conditional confidence limits under inverse sampling. *Statistics in Medicine* **14**, 2051–2056.
- LUI, K.-J. (1996). Sample size for the exact conditional test under inverse sampling. *Statistics in Medicine* **15**, 671–678.
- LUI, K.-J. (1997). Conditional maximum likelihood estimate and exact test of the common relative difference in combination of 2x2 tables under inverse sampling. *Biometrical Journal* **39**, 215–225.
- LUI, K.-J. (2000). Asymptotic conditional test procedures for relative difference under inverse sampling. *Computational Statistics & Data Analysis* **34**, 335–343.
- MORENO, V., MARTÍN, I., TORRES, F., HORAS, M., RIOS, J. & GÓNZALES, J. R. (2002). Inverse sampling and triangular sequential designs to compare a small proportion with a reference value. *Qüestió* **26**, 259–271.
- RÖHMEL, J. (2010). Comparing two independent samples using the risk difference under inverse sampling. *Computational Statistics and Data Analysis* **54**, 804–805.
- ROUSSAS, G. G. (1997). *A Course in Mathematical Statistics*. Academic Press, 2nd ed.
- RUDOLPH, G. (1967). A quasi-multinomial type of contingency table. *South African Statistical Journal* **1**, 59–65.

- STEYN, H. (1959). On  $\chi^2$ -tests for contingency tables of negative binomial types. *Statistica Neerlandica* **13**, 433–444.
- TANG, M.-L. & TIAN, M. (2009). Asymptotic confidence interval construction for risk difference under inverse sampling. *Computational Statistics and Data Analysis* **53**, 621–631.
- TANG, M.-L. & TIAN, M. (2010). Asymptotic confidence interval construction for risk difference under inverse sampling. *Statistical Computation and Simulation* **26**, 87–98.
- TIAN, M., TANG, M.-L., NG, H. T. & CHAN, P.-S. (2008). Confidence intervals for the risk ratio under inverse sampling. *Statistics in Medicine* **27**, 3301–3324.