ISSN 1995-0802, Lobachevskii Journal of Mathematics, 2016, Vol. 37, No. 4, pp. 466–496. © Pleiades Publishing, Ltd., 2016.

Confidence Intervals for a Ratio of Binomial Proportions Based on Direct and Inverse Sampling Schemes

T. Ngamkham*, A. Volodin, and I. Volodin*****

(Submitted by Alexandr Elizarov)

Thammasat University, Department of Mathematics and Statistics, Rangsit Center, Pathumthani 12121, Thailand,
University of Regina, Department of Mathematics and Statistics, Regina, Saskatchewan S4S0A2, Canada,
Kazan (Volga Region) Federal University, Institute of Computational Mathematics and Information Technologies, Department of Mathematical Statistics, Kremlevskaya st. 35, Kazan, 420008, Russia

Received November 27, 2015

Abstract—A general problem of the interval estimation for a ratio of two proportions p_1/p_2 according to data from two independent samples is considered. Each sample may be obtained in the framework of direct or inverse binomial sampling. Asymptotic confidence intervals are constructed in accordance with different types of sampling schemes with an application, where it is possible, of unbiased estimations of success probabilities and also their logarithms. Since methods of constructing confidence intervals in the situations when values for the both samples are obtained for identical sample schemes (for only direct or only inverse binomial sampling) are already developed and well known, the main purpose of this paper is the investigation of constructing confidence intervals in two cases that correspond to different sampling schemes (one is direct, another is inverse). In this situation it is possible to plan the sample size for the second sample according to the number of successes in the first sample. This, as it is shown by the results of statistical modeling, provides the intervals with confidence level which closer to the nominal value. Our goal is to show that the normal approximations (which are relatively simple) for estimates of p_1/p_2 and their logarithms are reliable for a construction of confidence intervals. The main criterion of our judgment is the closeness of the confidence coefficient to the nominal confidence level. It is proved theoretically and shown by statistically modeled data that the scheme of inverse binomial sampling with planning of the size in the second sample is preferred. Main probability characteristics of intervals corresponding to all possible combinations of sampling schemes are investigated by the Monte-Carlo method. Estimations of coverage probability, expectation and standard deviation of interval widths are collected in tables and some recommendations for an application of each of the intervals obtained are presented. Finally, a sufficient and complete review of the literature for the problem is also presented.

DOI: `10.1134/S1995080216040132`

Keywords and phrases: *Confidence limits, ratio of binomial proportions, inverse binomial sampling, asymptotic confidence limits.*

Dedicated to the memory
of Daniar Hamidovich Mushtari

1. INTRODUCTION

The problem we are solving in this article can be formulated in the following general way. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be two independent Bernoulli sequences with probabilities of success p_1 and p_2, respectively. Observations are done in sequential schemes of samplings with Markov's stopping times ν_1 and ν_2. According to the results of the observations $X^{(\nu_1)} = (X_1, \ldots, X_{\nu_1})$ and $Y^{(\nu_2)} = (Y_1, \ldots, Y_{\nu_2})$ it is required to construct a confidence interval for the parametric function $\theta = p_1/p_2$.

*E-mail: `thuntida@mathstat.sci.tu.ac.th`
**E-mail: `andrei@uregina.ca`
***E-mail: `Igor.Volodin@kpfu.ru`

A complexity of the problem stated can be explained by two reasons. The first is the absence of uniformly most powerful test for hypothesis testing $\theta = \theta_0$ with one-sided or two-sided alternative hypotheses in the case of an arbitrary value θ_0. It is known (see, for example Lehmann [1], Section 4.5) that the uniformly most powerful unbiased test exists for the values of the cross-product ratio $\rho = p_1(1-p_2)/p_2(1-p_1)$, however this is not what we are interested in this paper. Hence, it seems to be impossible to use the standard method of uniformly most precise confidence boundaries construction based on an acceptance region of the corresponding test. Therefore other tests should be applied or the method of pivot functions with an additional estimation of the nuisance parameter should be used.

The second method of confidence intervals construction is based on an application of pivot functions with good precision properties. It has similar difficulties because of the absence of an unbiased estimation for the parametric function $1/p$ for Bernoulli trials with the fixed sample size n (see Lehmann [2], Chapter 2, Section 1, Example 1.2; general theory of unbiased estimation is presented in the monograph Voinov and Nikulin (1993) [3]). However if the inverse, not direct binomial sampling method is used, then such an unbiased estimation exists. This is the starting point for our investigation on the confidence limits construction for a ratio of probabilities of success.

Now we provide a brief discussion of the literature pertaining to this subject in order to compare our results what is already known.

As far as we know, the statistical problem of interval estimation for ratio of binomial proportions has been solved only for schemes of sampling with a fixed number of observations $\nu_i = n_i$, $i = 1, 2$, or if observations in both samples are done in schemes of inverse binomial sampling. We start with a review of the literature for the case of direct binomial sampling in both samples.

The first easily computed methods of confidence estimation of θ have been suggested by Noether (1957) [4] and Guttman (1958) [5]. A review of these early methods may be found in Sheps (1959) [6]. Methods of confidence estimation of the ratio of proportions as a diagnostic test that can detect a disease, are used in McNeil et al. (1975) [7].

Next, some methods based on the corresponding tests for significance have been developed. For example, Thomas et al. (1977) [8] suggests to apply the method based on fixed marginals in the two-by-two tables. Santher et al. [9] develops and generalizes this method and suggests three related exact methods for finding confidence intervals.

Katz et al. [10] suggests three methods of lower confidence limit for θ, and the limits are defined as solutions of some equations. Numerical comparison shows that the method in which the logarithmic transformation is applied to the ratio of estimates of probabilities is preferential. Some modifications of these methods, that take their origin in Fieller's method, are discussed in Bailey (1987) [11].

Santher et al. [9] derives exact intervals for the risk ratio from Cornfield's [12] confidence interval for the odds ratio.

Koopman (1984) [13], and Miettinen and Nurminen (1985) [14] propose methods based on asymptotic likelihood for hypothesis $\theta = \theta_0$ testing with the alternative hypotheses $\theta \neq \theta_0$. In Koopman (1984) [13] this method compared with the one recommended by Katz et all (1978) [10].

All the results until the end of 1980's were summarized in Gart and Nam (1988) [15]. In this paper they provide a comprehensive survey of various approximation methods of confidence interval constructions for the ratio of probabilities based on the properties of goodness of fit with Pearson's chi-square test, invariance, universality of an application for all observations and computational simplicity. Also, asymptotic methods were improved by taking into account the asymptotic asymmetry of statistics (see also Gart and Nam (1990) [16]. The results obtained are extended for the case of estimating the common ratio in a series of two-by-two tables, which was considered before in Gart (1985) [17]. Extensive numerical illustrations are provided, which allow the comparison of precision properties to the methods of interval estimation of probabilities ratio. Instead of iterative algorithms for calculating the approximate confidence intervals that have been provided by Koopman (1984) [13], Gart and Nam (1988) [15], Nam (1995) [18] gives the analytical solutions for upper and lower confidence limits in closed form.

For an interval estimator construction, Bedrick (1987) [19] uses the special power divergent family of statistics. Intervals based on inverting the Pearson, likelihood-ratio, and Freeman–Tukey statistics are included in this family. Asymptotic efficiency, coverage probability, and expected interval width are investigated. Comparisons of methods are provided by numerical examples.

The bootstrap method of a confidence interval construction for θ is suggested in Kinsella (1987) [20].

Coe and Tamhane (1993) [21] provide a method for small sample confidence intervals construction for the difference of probabilities, based on an extension of known Sterne's method for constructing small sample confidence intervals for a single success probability. Modifications of the algorithm for ratio probabilities are also indicated.

The paper by Nam and Blackwelder (2002) [22] develops a superior alternative to the Wald's interval and gives corresponding sample size formulas. Bonett and Price (2006) [23] propose alternatives to the Nam–Blackwelder confidence interval based on combining two Wilson score intervals. Two sample size formulas are derived to approximate the sample size required to achieve an interval estimate with desired confidence level and width.

Extensive numerical illustrations for comparison of exact and asymptotic methods for θ confidence intervals constructions are presented in the thesis by Mukhopadhyay (2003) [24].

The starting point of the research on confidence intervals in the situation when the data from both samples are taken by the inverse binomial sampling most probably was done in the paper by Bennet (1981) [25]. In this paper the preferable usage of inverse sampling scheme in epidemiology studies is mentioned, especially in the case of a small probability of a success. In the same paper, an asymptotic F-test for hypothesis testing of the value of the ratio of proportions is suggested, which has been used later by other statisticians for confidence intervals construction. Advantages of the choice of unbiased estimates with uniformly minimal risk in the case of the inverse binomial sampling (we are using exactly these estimates) among other estimates from the point of view of quadratic risk is shown in Roberts (1993) [26], Bennet (1981) [25], and especially Lui (1996) [27].

A development of basic methods of confidence intervals construction and investigation of their precision properties for the case of inverse sampling schemes belongs to K.J. Lui. Three methods of confidence estimation with a comparative analysis of their properties on statistically modeled data are developed in Lui (1995) [28]. These methods are based on the normal approximation of the estimate of probabilities ratio and also the logarithm of this estimate. The third method is based on F-test suggested in Bennet (1981) [25]. Confidence intervals based on the maximum likelihood estimate and unbiased estimate with uniformly minimal risk are presented in Lui (2001) [29]. A series of similar methods is suggested in Lui (1995b) [30] and (1996a) [31]. A comparative review of characteristics of methods suggested by Lui with data obtained by statistical experiments is presented in Lui (2006) [32]. Confidence intervals based on Wald's test statistics and its logarithm are suggested in Wu et al. (2009) [33]. We should note that in comparison with direct binomial sampling, the number of publications for confidence intervals for inverse binomial sampling scheme is not too large.

We construct asymptotic confidence intervals for a few schemes of direct and inverse sampling and illustrate their characteristics by the results of statistical modeling (see Tables). In each cell of tables the following characteristics are presented: actual confidence level (the nominal confidence level is chosen to be 0.95), expectation, and standard deviation of the width of a corresponding confidence interval. For each interval 10^6 random numbers with Bernoulli and/or Negative Binomial Distributions were generated with parameters (probabilities of success) p_1, $p_2 = 0.1$ (0.1) 0.9. The tables presented contain only a part of the results for the probability values 0.2, 0.5, 0.8, but the conclusions (see the last section) are made according to all obtained results of statistical modeling. Moreover for confidence intervals based on the logarithm of an estimate of the ratio p_1/p_2, an additional investigation of precision and reliability properties of these intervals for small values of probabilities $p_1,\ p_2 = 0.05, 0.1, 0.2$ is conducted. Analysis of the results from the tables shows the preference of the inverse binomial sampling scheme for the planning of experiments in the second sample. We provide a theoretical justification of this recommendation.

The initial objective of the investigation resulting in the present paper was only to justify an application of the delta-method and corresponding normal approximations of probability ratio estimates and logarithms of these estimates for a construction of confidence intervals. We were not planning to construct confidence intervals that have better properties than the intervals presented in the papers of our predecessors and discussed in the review above. But we were able to construct asymptotic intervals for different sampling schemes by considering the situation when the stopping moment for observations in the second sample depends on the result of observation in the first one. Moreover, it appears that confidence intervals formulae that we obtain in this paper are much simpler for practical applications than the formulae presented in some of our predecessor's papers. We can achieve the desired precision and reliability in confidence intervals without solving complicated system of nonlinear equations.

2. ESTIMATION OF PROPORTIONS AND THEIR RATIO

For a solution of the problems stated in Introduction, it is necessary to construct estimates, preferably unbiased, for a proportion p and parametric function p^{-1} for two schemes of Bernoulli trials.

From the point of view of unbiased estimation of the parametric function p^{-1}, the most simple is the case of inverse binomial sampling. In this scheme a Bernoulli sequence $Y^{(\nu)} = (Y_1, \ldots, Y_\nu)$ is observed with a stopping time

$$\nu = \min \left\{ n \colon \sum_{k=1}^{n} Y_k \geq m \right\}.$$

That is, the components of the sequence $Y_1, Y_2, \ldots$ is observed until the given number m of successes will appear. As it is known, ν has Pascal distribution $P(m, p)$, which is a particular case of negative binomial distribution with integer parameters. When $m \to \infty$, the random variable $\overline{Y}_m = \nu/m$ is asymptotically normal with mean $\mu_Y = 1/p$ and variance $\sigma_Y^2 = (1-p)/mp^2$. Statistics $\hat{p}_m^{-1} = \overline{Y}_m = \nu/m$ and (see Guttman, I. (1958)) $\hat{p}_m = (m-1)/(\nu-1)$ are unbiased estimates of p^{-1} and p, respectively.

In the case of direct binomial sampling a random vector $X^{(n)} = (X_1, \ldots, X_n)$ with Bernoulli components and fixed number of observations n is observed. For the proportion p there exists an unbiased estimate with uniformly minimal variance $\hat{p}_n = \overline{X}_n = T/n$, where the statistics

$$T = \sum_{k=1}^{n} X_k$$

has the binomial distribution $B(n, p)$. The estimate $\overline{X}_n$ is asymptotically normal with mean $\mu_X = p$ and variance $\sigma_X^2 = p(1-p)/n$. As we already mentioned in the Introduction, there is no unbiased estimate for the parametric function $\vartheta = p^{-1}$. Below we propose an estimate of ϑ with bias that decreases with an exponential rate as $n \to \infty$.

It is suggested to use the statistics ν which equals to the number of the last trial with $X_\nu = 1$. Then, by the analogy with the inverse binomial sampling, it is natural to suggest the statistic $\hat{\vartheta}_\nu = \nu/T_\nu$ where

$$T_\nu = \sum_{k=1}^{\nu} X_k$$

as the estimate of ϑ. If the value of ν is unknown, so it is better to use the projection $\hat{p}_n^{-1} = \mathbf{E}\{\hat{\vartheta}_\nu | T\}$ of this statistic on the sufficient statistic T. As it is known, (see Lehmann (1998), Chapter 2, Section 1), a projection does not cause an increase of the risk if the loss function is convex.

Proposition . *The projected estimator has the following representation* $\hat{p}_n^{-1} = (n+1)/(T+1)$ *and its mean value is*

$$\mathbf{E}\hat{p}_n^{-1} = \frac{1}{p}\left(1 - (1-p)^{n+1}\right).$$

Proof. The joint distribution of statistics ν and T is defined by the probabilities

$$\mathrm{P}(\nu = k, T = t) = \begin{cases} 0, & \text{if} \quad k = 0, \quad t \geq 1, \\ (1-p)^n, & \text{if} \quad k = 0, \quad t = 0, \\ \binom{k-1}{t-1} p^t (1-p)^{n-t}, & \text{if} \quad t = 1, \ldots, n, \quad k = t, \ldots, n. \end{cases}$$

The marginal distribution of statistic T is

$$\mathrm{P}(T = t) = \binom{n}{t} p^t (1-p)^{n-t}, \quad t = 0, 1, \ldots, n,$$

then the conditional distribution

$$P(\nu = k|T = t) = \begin{cases} 0, & \text{if} \quad k = 0, \quad t \geq 1, \\ 1, & \text{if} \quad k = 0, \quad t = 0, \\ \binom{k-1}{t-1}/\binom{n}{t}, & \text{if} \quad t = 1, \ldots, n, \quad k = t, \ldots, n. \end{cases}$$

All further calculations for mean values are trivial, if we use the well known combinatorial formula

$$\sum_{k=1}^{N} \binom{n+k}{n} = \binom{n+N+1}{n+1}.$$

□

Therefore, in the case of the direct binomial sampling, we will use the estimate $\hat{p}_n = \overline{X}_n$ for the proportion p and the estimate $\hat{p}_n^{-1} = (n+1)/(n\overline{X}_n + 1)$ for p^{-1}.

We now discuss estimates of the parametric function $\theta = p_1/p_2$ for all possible combinations of the direct and inverse sampling schemes for data.

Direct-direct. Let both samples are obtained in the scheme of direct sampling with probabilities p_1 and p_2 of successes and sample sizes n_1 and n_2, respectively. It follows from the Proposition 1 proved above that for an estimate of the parametric function $\theta = p_1/p_2$, it is appropriate to take the statistic

$$\hat{\theta}_{n_1,n_2} = \frac{\overline{X}_{n_1}(n_2+1)}{n_2\overline{X}_{n_2} + 1}$$

with the mean value

$$\mathbf{E}\hat{\theta}_{n_1,n_2} = \theta\left(1 - (1-p_2)^{n_2+1}\right).$$

Obviously, when $n_2 \to \infty$ this estimate is equivalent to $\tilde{\theta}_{n,m} = \overline{X}_{n_1}/\overline{X}_{n_2}$, but the application of $\hat{\theta}_{n_1,n_2}$ allows to avoid the division by zero when all the outcomes in the second sample are failures (compare with a solution of this problem in the paper Cho (2007) [34]).

Direct-inverse. The first sample is obtained by the scheme of direct binomial sampling with probability p_1 of a success and fixed sample size n, while the second sample is obtained by the scheme of inverse binomial sampling with the probability p_2 and stopping time which is defined by the fixed number m of successes in the sample. An unbiased estimate p_2^{-1} is given by statistics $\hat{p}_m^{-1} = \overline{Y}_m = \nu/m$. In the framework of the sampling schemes under consideration we suggest the unbiased estimate $\hat{\theta}_{n,m} = \overline{X}_n\overline{Y}_m$ for the parametric function p_1/p_2. This estimate minimizes any risk function with convex loss function uniformly by all values of p_1, p_2 (see Lehmann (1998), Chapter 2, Section 1).

Inverse-direct. For the first sample obtained by the scheme of inverse binomial sampling with parameters (p_1, m), an unbiased estimate of p_1 is the statistics $\hat{p}_m = (m-1)/(\nu - 1)$. For the second sample obtained by the scheme of direct sampling with parameters (p_2, n), an estimate of p_2^{-1} with exponentially decreasing rate of the bias is $\hat{p}_n^{-1} = (n+1)/(n\overline{X}_n + 1)$. Therefore, as asymptotically $(n \to \infty)$ unbiased estimate of θ we suggest to apply the statistics

$$\hat{\theta}_{m,n} = \frac{(m-1)(n+1)}{(\nu-1)(n\overline{X}_n + 1)}.$$

The same as in the "direct-direct" case, this estimate for large values of m and n is equivalent to the estimate $\tilde{\theta}_{m,n} = [\overline{Y}_m\overline{X}_n]^{-1}$, but an application of $\theta_{m,n}$ solves the problem of proportions ration estimation when the number of successes in the second sample is zero.

Inverse-inverse. In the case when both samples are obtained in the scheme of inverse binomial sampling with parameters (p_1, m_1) and (p_2, m_2), respectively, the estimate

$$\hat{\theta}_{m_1,m_2} = \frac{m_1 - 1}{\nu_1 - 1} \cdot \frac{\nu_2}{m_2}$$

is an unbiased estimation of the probabilities ratio θ which minimizes any risk function with convex loss function uniformly by all values of p_1, p_2. Nothing prohibits to apply instead of θ_{m_1,m_2} its stochastic approximation $\tilde{\theta}_{m_1,m_2} = \overline{Y}_{m_2}/\overline{Y}_{m_1}$ for large values of m_1 and m_2.

3. ASYMPTOTIC DISTRIBUTION OF ESTIMATES FOR PROBABILITY RATIO AND THEIR LOGARITHMS

For large values of n and m, all four estimates of probability ratio θ are continuous functions of statistics $\overline{X}_n$ and $\overline{Y}_m$ with finite second moments and therefore the estimates are asymptotically normal. Our immediate task is to find the asymptotic of the mean and variance of these estimates, for which we explore the standard delta-method. The method is based on Taylor series expansion in the neighborhoods of the mean values of the statistics $\overline{X}_n$ and $\overline{Y}_m$. We keep only linear terms because for our estimates the remainder term of the expansion converges in probability to zero with the rate $O\left([\min\{n,m\}]^{-1/2}\right)$.

It is natural to expect that the distribution of the statistic $\ln\hat{\theta}$ possesses a better approximation by the normal distribution in comparison with the estimate $\hat{\theta}$. At least this is confirmed by the results of statistical modeling that were obtained by us and our predecessors. Because of that below we will calculate the parameters for normal approximation not only for the estimates of θ, but also for their logarithms.

Direct-direct. With the notation

$$X_i = \overline{X}_{n_i}, \quad \mu_i = \mathbf{E}X_i = p_i, \quad \sigma_i^2 = \mathbf{Var}X_i = p_i(1-p_i)/n_i, \quad i = 1,2$$

for an estimate of θ we obtain the following asymptotic representation

$$\hat{\theta}_{n_1,n_2} \sim \frac{X_1}{X_2} \sim \theta + \frac{X_1-\mu_1}{\mu_2} - \frac{\mu_1}{\mu_2^2}(X_2-\mu_2).$$

Therefore the estimate $\hat{\theta}_{n_1,n_2}$ is asymptotically normal with mean θ and variance

$$\sigma_{n_1,n_2}^2 = \mathbf{Var}\hat{\theta}_{n_1,n_2} = \frac{\sigma_1^2}{\mu_2^2} + \frac{\mu_1^2}{\mu_2^4}\sigma_2^2 = \theta^2\left(\frac{1-p_1}{n_1p_1} + \frac{1-p_2}{n_2p_2}\right).$$

Similarly

$$\ln\hat{\theta}_{n_1,n_2} \sim \ln X_1 - \ln X_2 \sim \ln\theta + \frac{X_1-\mu_1}{\mu_1} - \frac{(X_2-\mu_2)}{\mu_2},$$

which implies the asymptotic normality of $\ln\hat{\theta}_{n_1,n_2}$ with mean $\ln\theta$ and variance

$$\mathbf{Var}\ln\hat{\theta}_{n_1,n_2} = \frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2} = \frac{1-p_1}{n_1p_1} + \frac{1-p_2}{n_2p_2}.$$

Direct-inverse. Let

$$X = \overline{X}_n, \quad \mu_1 = \mathbf{E}X = p_1, \quad \sigma_1^2 = \mathbf{Var}X = p_1(1-p_1)/n;$$

$$Y = \overline{Y}_m, \quad \mu_2 = \mathbf{E}Y = p_2^{-1}, \quad \sigma_2^2 = \mathbf{Var}Y = (1-p_2)/(mp_2^2).$$

For the estimate of θ the following asymptotic representation is true

$$\hat{\theta}_{n,m} = XY \sim \theta + (X-\mu_1)\mu_2 + (Y-\mu_2)\mu_1$$

which implies that $\hat{\theta}_{n,m}$ is asymptotically normal with mean θ and variance

$$\sigma_{n,m}^2 = \sigma_1^2\mu_2^2 + \sigma_2^2\mu_1^2 = \theta^2\left(\frac{1-p_1}{np_1} + \frac{1-p_2}{m}\right).$$

Since

$$\ln\hat{\theta}_{n,m} = \ln XY \sim \ln\theta + \frac{X-\mu_1}{\mu_1} + \frac{Y-\mu_2}{\mu_2},$$

the asymptotic mean of the logarithm of the estimate is $\ln\theta$ and asymptotic variance is

$$\mathbf{Var}\ln\hat{\theta}_{n,m} = \frac{\sigma_1}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2} = \frac{1-p_1}{np_1} + \frac{1-p_2}{m}.$$

Inverse-direct. By the analogy with "direct-inverse" case, denote

$$Y = \overline{Y}_m, \quad \mu_1 = \mathbf{E}Y = p_1^{-1}, \quad \sigma_1^2 = \mathbf{Var}Y = (1-p_1)/(mp_1^2);$$

$$X = \overline{X}_n, \quad \mu_2 = \mathbf{E}X = p_2, \quad \sigma_2^2 = \mathbf{Var}X = p_2(1-p_2)/n.$$

Asymptotic representation of the estimate of θ in this case can be written as

$$\hat{\theta}_{m,n} \sim \frac{1}{YX} \sim \theta - \frac{Y-\mu_1}{\mu_1^2\mu_2} - \frac{X-\mu_2}{\mu_1\mu_2^2}.$$

Therefore the estimate $\hat{\theta}_{m,n}$ is asymptotically normal with mean θ and variance

$$\sigma_{n,m}^2 = \frac{\sigma_1^2}{\mu_1^4\mu_2^2} + \frac{\sigma_2^2}{\mu_1^2\mu_2^4} = \theta^2\left(\frac{1-p_1}{m} + \frac{1-p_2}{np_2}\right).$$

For the logarithm of the estimate we have the representation

$$\ln\hat{\theta}_{m,n} = -\ln YX \sim \ln\theta - \frac{Y-\mu_1}{\mu_1} - \frac{X-\mu_2}{\mu_2}$$

which implies that the statistics $\ln\hat{\theta}_{m,n}$ is asymptotically normal with mean $\ln\theta$ and variance

$$\mathbf{Var}\ln\hat{\theta}_{m,n} = \frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2} = \frac{1-p_1}{m} + \frac{1-p_2}{np_2}.$$

Inverse-inverse. For these sampling schemes denote

$$Y_i = \overline{Y}_{m_i}, \quad \mu_i = \mathbf{E}Y_i = p_i^{-1}, \sigma_i^2 = \mathbf{Var}Y_i = (1-p_i)/(m_ip_i^2), i = 1,2$$

and obtain the asymptotic representation of the estimate

$$\hat{\theta}_{m_1,m_2} \sim \frac{Y_2}{Y_1} \sim \theta - \frac{\mu_2(Y_1-\mu_1)}{\mu_1^2} + \frac{Y_2-\mu_2}{\mu_1}.$$

Therefore, the estimate $\hat{\theta}_{m_1,m_2}$ is asymptotically normal with mean θ and variance

$$\sigma_{m_1,m_2}^2 = \frac{\mu_2^2}{\mu_1^4}\sigma_1^2 + \frac{\sigma_2^2}{\mu_1^2} = \theta^2\left(\frac{1-p_1}{m_1} + \frac{1-p_2}{m_2}\right).$$

Since

$$\ln\hat{\theta}_{m_1,m_2} = \ln\frac{Y_2}{Y_1} \sim \ln\theta - \frac{Y_1-\mu_1}{\mu_1} + \frac{Y_2-\mu_2}{\mu_2},$$

the logarithm of the estimate $\hat{\theta}_{m_1,m_2}$ is asymptotically normal with mean $\ln\theta$ and variance

$$\mathbf{Var}\ln\hat{\theta}_{m_1,m_2} = \frac{\sigma_1^2}{\mu_1^2} + \frac{\sigma_2^2}{\mu_2^2} = \frac{1-p_1}{m_1} + \frac{1-p_2}{m_2}.$$

The asymptotic formulae for variances of the estimates allow us to make two important conclusions that in large extend predetermine the precision properties of the asymptotic confidence intervals based on these estimates. In the following this conclusions will be supported by the results of statistical modeling.

(I) For the schemes of inverse binomial sampling with parameters (p, m) the mean sample size is $\mathbf{E}\nu = m/p$. If the observations are obtained in the scheme of direct sampling with the same probability p of the success and sample size $n = m/p$, then "on average" it is equivalent to the scheme of inverse sampling from the point of view of the cost for the experiment. Variance of the estimate $\hat{\theta}_{m_1,m_2}$ coincides with variance of the estimate $\hat{\theta}_{n_1,n_2}$, if $m_1 = n_1p_1$ and $m_2 = n_2p_2$. Therefore *schemes direct-direct and inverse-inverse are equivalent in the same sense from the same point of view of asymptotic precision of the estimates for probability ratio*. Of course, the same conclusion is true for all pairs of sampling schemes with the corresponding substitution of m by np.

(II) From the formula of asymptotic variance of the estimate $\hat{\theta}_{n_1,n_2}$ of θ for the direct-direct scheme it follows that the corresponding asymptotic confidence interval will obtain poor precision properties for small true values of the probability p_2 of the success in the second sample because for $p_2 \to 0$ the variance $\sigma^2_{n_1,n_2} = O(p_2^{-3})$. The application of $\ln \hat{\theta}_{n_1,n_2}$ for the interval estimation of θ significantly improves the precision properties of the estimate because in this case the dispersion of the supporting estimate is of the order $O(p_2^{-1})$ when $p_2 \to 0$. However for $p_1 \to 0$ the order of the growth for the variance is the same $O(p_1^{-1})$. If the second sample is obtained by the inverse scheme, then for $p_2 \to 0$ the variance $\sigma^2_{n,m}$ of the supporting estimate has even slower growth, namely as $O(p_1^{-2})$. An application of the logarithm of the estimate in this scheme solves the problem of poor precision properties of the asymptotic interval estimate for small values of p_2 completely because the variance of the supporting estimate when $p_2 \to 0$ is a bounded function of p_2. However for $p_1 \to 0$ the order of growth for the variance is the same as in the direct-direct scheme. Finally, *when both samples are obtained by the inverse schemes, all troubles with a poor behaviour of variance of the supporting estimate can be overcome by an application of* $\ln \hat{\theta}_{m_1,m_2}$, in this case the variance is the bounded function by both arguments p_1, p_2.

4. CONFIDENCE LIMITS

The estimates $\hat{\theta}$ for the probabilities ratio θ and asymptotics for their mean values and variances obtained in the previous section, show that for all sampling schemes means and variances have the same structure: $\mathbf{E}\hat{\theta} = \theta$, $\mathbf{Var}\hat{\theta} = \theta^2 s^2(p_1, p_2)$. If the sample sizes for both sampling schemes tend to infinity, then

$$\mathbf{P}\left(|\hat{\theta} - \theta| \le \lambda_\alpha \theta s(p_1, p_2)\right) \sim 1 - \alpha,$$

where λ_α is $(1 - \alpha/2)$-quantile of the standard normal distribution. Since $s(p_1, p_2)$ is a continuous function of its arguments, replacing $\theta s(p_1, p_2)$, or only $s(p_1, p_2)$ by their plug-in estimate, we obtain the same asymptotic equality.

Similarly, because the mean values and variances for logarithms of estimates of θ have the same form: $\mathbf{E} \ln \hat{\theta} = \ln \theta$, $\mathbf{Var} \ln \hat{\theta} = s_l^2(p_1, p_2)$, using the inequity

$$|\ln \theta - \ln \hat{\theta}| \le \lambda_\alpha s_l^2(p_1, p_2)$$

and replacing p_1 and p_2 by their estimates that correspond to sampling schemes, we obtain asymptotically $(1 - \alpha)$-confidence intervals. Therefore, the following theorem is true.

Theorem 1. *If the sample sizes in both sample schemes tend to infinity, then the intervals with the following end-points*

$$\hat{\theta}[1 \mp \lambda_\alpha s(\hat{p}_1, \hat{p}_2)], \tag{1}$$

$$\hat{\theta}[1 \pm \lambda_\alpha s(\hat{p}_1, \hat{p}_2)]^{-1}, \tag{2}$$

$$\hat{\theta} \exp\{\mp \lambda_\alpha s_l(\hat{p}_1, \hat{p}_2)\}, \tag{3}$$

are the asymptotic $(1 - \alpha)$*-confidence sets for the ratio of probabilities* $\theta = p_1/p_2$.

Remark. It is simple to observe that interval (2) has bigger width than interval (1) for all possible observations. However, it is possible, and this is confirmed by the results of statistical modeling, interval (2) has confidence coefficient which is more close to the nominal confidence level $1 - \alpha$. Nevertheless, if the value $1 - \lambda_\alpha s(\hat{p}_1, \hat{p}_2)$ appears to be close to zero, we obtain very wide confidence interval. The fact that this sometimes happens is possible to see from the results of statistical modeling presented in the tables below.

In the following we will call interval (1) as *linear*, interval (2) as *hyperbolic*, and interval (3) *as logarithmic*. Now we consider each of the intervals corresponding to four possible combinations of the sampling schemes separately and provide the results of their statistical modeling for precision and confidence properties.

Direct-direct. When both samples are obtained by direct sampling scheme, asymptotic ($n_1, n_2 \to \infty$) confidence intervals (1)–(3) are based on the relative frequencies X_{n_2} and X_{n_2} of successes (sample means) in each sample and can be written as

$$\frac{\overline{X}_{n_1}(n_2+1)}{n_2\overline{X}_{n_2}+1}\left(1 \mp \lambda_\alpha\sqrt{\frac{1-\overline{X}_{n_1}}{n_1\overline{X}_{n_1}+1}+\frac{1-\overline{X}_{n_2}}{n_2\overline{X}_{n_2}+1}}\right), \tag{dd 1}$$

$$\frac{\overline{X}_{n_1}(n_2+1)}{n_2\overline{X}_{n_2}+1}\left(1 \pm \lambda_\alpha\sqrt{\frac{1-\overline{X}_{n_1}}{n_1\overline{X}_{n_1}+1}+\frac{1-\overline{X}_{n_2}}{n_2\overline{X}_{n_2}+1}}\right)^{-1}, \tag{dd 2}$$

$$\frac{\overline{X}_{n_1}(n_2+1)}{n_2\overline{X}_{n_2}+1}\exp\left\{\mp\lambda_\alpha\sqrt{\frac{1-\overline{X}_{n_1}}{n_1\overline{X}_{n_1}+1}+\frac{1-\overline{X}_{n_2}}{n_2\overline{X}_{n_2}+1}}\right\} \tag{dd 3}$$

(expressions of the form $(n+1)/n$ are replaced by 1).

Below we provide the results of statistics modeling (Tables 1–4). For each pair (n_1, n_2) of observations and values (p_1, p_2) of success probabilities, we present Monte-Carlo estimations of the coverage probability, mean width, and standard deviation of the width for each of intervals (dd 1)–(dd 3). The nominal level is assumed to be 0.95.

The results of Table 1 show that the linear interval (dd 1) has the confidence level lower than nominal and an error not larger than 0.01 is possible only for $n_1, n_2 \geq 100$ and $p_1, p_2 \geq 0.5$. Hyperbolic interval (dd 2) (see Table 2) has coverage probability undoubtedly more in line with the nominal level 0.95 and an error more than 0.01 is observed only for some values of $p_1 \leq 0.2$ and $n_2 \leq 50$, for other values of probabilities of successes and sample sizes the coverage probability is practically the same as the nominal level. Nevertheless the hyperbolic intervals, as it is expected, has poorer precision properties, but from a practical point of view this is noticeable only if sample sizes $n_1, n_2 \leq 100$.

Logarithmic interval (dd 3) (Table 3) has high correspondence for coverage probability to the nominal and insignificantly worse precision properties in comparison with the linear interval. Table 4 shows that for $n_1, n_2 \geq 100$ the logarithmic interval still has high coverage probability even for small values of success probabilities such as 0.05–0.1, but its precision properties in this region are sufficiently poor.

Of course, these properties of asymptotic confidence intervals for direct scheme of binomial sampling are already well known and similar conclusions can be found in some review papers that we mentioned in Introduction. We provide here more extensive numerical illustrations in order to have certain standard for a comparison of precision and reliability properties for intervals based on different schemes of sampling in first and second samples.

Direct-inverse. If the first sample is obtained by the direct binomial sampling with sample size n and the second sample corresponds to the inverse binomial sampling scheme with fixed number of successes m, then asymptotic ($n, m \to \infty$) $(1-\alpha)$-confidence intervals (1)–(3) with

$$\hat{p}_1 = \overline{X}_n, \quad \hat{p}_1^{-1} = \frac{n+1}{n\overline{X}_n+1}, \quad \hat{p}_2 = \frac{m-1}{\nu-1}, \quad \overline{Y}_m = \frac{\nu}{m}$$

have the form

$$\overline{X}_n\overline{Y}_m\left(1 \mp \lambda_\alpha\sqrt{\frac{1-\hat{p}_1}{n\hat{p}_1}+\frac{1-\hat{p}_2}{m}}\right), \tag{di 1}$$

$$\overline{X}_n\overline{Y}_m\left(1 \pm \lambda_\alpha\sqrt{\frac{1-\hat{p}_1}{n\hat{p}_1}+\frac{1-\hat{p}_2}{m}}\right)^{-1}, \tag{di 2}$$

$$\overline{X}_n\overline{Y}_m\exp\left\{\mp\lambda_\alpha\sqrt{\frac{1-\hat{p}_1}{n\hat{p}_1}+\frac{1-\hat{p}_2}{m}}\right\}. \tag{di 3}$$

Characteristics of this interval are presented in Tables 5–8.

Table 1. Coverage probability, mean width, and standard deviation for the interval (dd 1)

n_1	n_2	50			100			200		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
50	0.2	0.888	0.918	0.915	0.893	0.913	0.917	0.894	0.911	0.919
		1.563	0.472	0.268	1.264	0.439	0.263	1.115	0.423	0.261
		0.811	0.115	0.043	0.401	0.081	0.037	0.252	0.066	0.035
	0.5	0.896	0.932	0.944	0.913	0.939	0.941	0.909	0.938	0.939
		3.189	0.779	0.379	2.319	0.662	0.357	1.834	0.597	0.346
		1.771	0.185	0.043	0.729	0.096	0.024	0.364	0.054	0.015
	0.8	0.899	0.934	0.948	0.906	0.939	0.949	0.908	0.943	0.943
		4.747	0.997	0.389	3.237	0.758	0.333	2.346	0.613	0.302
		2.687	0.268	0.064	1.071	0.128	0.035	0.498	0.065	0.027
100	0.2	0.913	0.936	0.936	0.924	0.936	0.935	0.924	0.934	0.932
		1.451	0.382	0.202	1.106	0.342	0.196	0.925	0.321	0.192
		0.824	0.091	0.027	0.358	0.056	0.020	0.198	0.039	0.018
	0.5	0.916	0.938	0.948	0.924	0.939	0.947	0.928	0.946	0.946
		3.209	0.691	0.299	2.222	0.553	0.271	1.667	0.474	0.256
		2.235	0.179	0.041	0.752	0.090	0.021	0.361	0.048	0.012
	0.8	0.916	0.937	0.944	0.924	0.944	0.949	0.925	0.945	0.949
		4.918	0.959	0.339	3.289	0.703	0.276	2.322	0.541	0.238
		3.148	0.273	0.066	1.158	0.129	0.031	0.529	0.064	0.017
200	0.2	0.918	0.939	0.942	0.932	0.941	0.942	0.935	0.942	0.944
		1.383	0.318	0.153	0.991	0.271	0.145	0.782	0.245	0.141
		0.987	0.080	0.020	0.335	0.044	0.013	0.171	0.027	0.010
	0.5	0.921	0.937	0.946	0.934	0.946	0.948	0.937	0.948	0.949
		3.207	0.635	0.245	2.159	0.483	0.211	1.558	0.391	0.192
		2.042	0.176	0.041	0.775	0.086	0.020	0.356	0.044	0.011
	0.8	0.921	0.936	0.938	0.933	0.942	0.945	0.937	0.947	0.950
		4.031	0.939	0.311	3.317	0.671	0.241	2.305	0.496	0.195
		3.266	0.276	0.067	1.198	0.131	0.032	0.543	0.064	0.015

All asymptotic confidence intervals for the direct-inverse sampling scheme have an interesting property (see Tables 5–8) that the coverage probability is almost free of the choice of m (number of successes in the second sample) in the range 20–100, while the precision properties are improved significantly when m increases (of course, the similar picture appears if the sample size n for the first same increases, too). Choosing again as the acceptable error of 0.01 in the correspondence of the coverage probability to the nominal 0.95, it is possible to state that the linear interval di 1 (Table 5) satisfies this requirement only in the region $p_1 \geq 0.5$ for all $n, m \geq 50$ and all $p_2 \geq 0.2$. Hyperbolic interval (di 2) (Table 6) has undoubtedly better properties even for $p_1 = 0.2$ and, what is surprising, its precision properties are only a little inferior to the linear interval. Logarithmic interval (di 3) (Table 7) is acceptable for almost all region with better precision properties than linear and hyperbolic intervals.

Table 2. Coverage probability, mean width, and standard deviation for the interval (dd 2)

n_1	n_2	50			100			200		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
50	0.2	0.928	0.938	0.929	0.954	0.935	0.926	0.958	0.934	0.926
		5.080	0.808	0.407	2.457	0.686	0.390	1.913	0.643	0.383
		6.353	2.220	0.113	1.219	0.218	0.083	0.578	0.182	0.088
	0.5	0.959	0.948	0.949	0.967	0.960	0.945	0.974	0.954	0.942
		5.844	0.928	0.421	3.074	0.749	0.391	2.173	0.662	0.377
		5.413	0.245	0.049	1.218	0.109	0.026	0.452	0.058	0.016
	0.8	0.954	0.952	0.954	0.964	0.963	0.950	0.965	0.961	0.946
		7.795	1.111	0.403	3.985	0.805	0.343	2.611	0.638	0.310
		6.742	0.336	0.070	1.577	0.143	0.037	0.593	0.071	0.029
100	0.2	0.943	0.946	0.945	0.955	0.950	0.947	0.961	0.947	0.945
		3.226	0.497	0.244	1.622	0.423	0.233	1.204	0.388	0.228
		3.873	0.122	0.027	0.651	0.062	0.017	0.252	0.037	0.012
	0.5	0.942	0.946	0.949	0.956	0.950	0.952	0.963	0.954	0.945
		5.470	0.785	0.318	2.796	0.599	0.285	1.895	0.503	0.268
		5.090	0.236	0.046	1.175	0.102	0.022	0.433	0.051	0.012
	0.8	0.936	0.949	0.952	0.951	0.950	0.951	0.949	0.950	0.949
		7.985	1.062	0.349	4.018	0.740	0.281	2.561	0.557	0.241
		7.269	0.340	0.070	1.726	0.145	0.033	0.638	0.068	0.018
200	0.2	0.933	0.946	0.941	0.950	0.950	0.950	0.953	0.950	0.950
		2.532	0.378	0.170	1.305	0.306	0.158	0.925	0.271	0.153
		2.486	0.106	0.022	0.534	0.050	0.013	0.213	0.028	0.009
	0.5	0.931	0.946	0.953	0.949	0.947	0.950	0.951	0.953	0.951
		5.316	0.704	0.255	2.658	0.513	0.217	1.735	0.406	0.197
		4.965	0.212	0.044	1.111	0.096	0.021	0.430	0.046	0.010
	0.8	0.930	0.946	0.947	0.947	0.951	0.946	0.948	0.955	0.951
		8.085	1.031	0.319	4.006	0.699	0.243	2.520	0.508	0.197
		7.478	0.329	0.072	1.732	0.141	0.033	0.643	0.066	0.016

It is also recommended for the values of p_1 of the order 0.1 (see Table 8) with sample size of the first sample $n \geq 100$ without taking into consideration the parameters of the second sample. Also it can be recommended for all $p_1, p_2 \geq 0.05$, if $n \geq 200$.

Note also that for all kinds of confidence intervals (the same as in the previous sampling scheme) the coverage probability is smaller that the nominal level. It appears that this flaw may be overcame by make the confidence interval conservative, if we keep up the following rule for stopping time for the second sample.

The important part of the suggested plan realization of the estimate of θ is the choice of the number m. The (random) sample size for the second sample that depends on m and the following sampling plan for the second sample can be suggested. Repeat observations until the same number of successes as

Table 3. Coverage probability, mean width, and standard deviation for the interval (dd 3)

n_1	n_2	50			100			200		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
50	0.2	0.954	0.946	0.939	0.950	0.945	0.931	0.948	0.938	0.938
		1.730	0.502	0.283	1.360	0.464	0.277	1.184	0.446	0.273
		0.957	0.120	0.042	0.438	0.081	0.035	0.257	0.064	0.032
	0.5	0.946	0.955	0.951	0.950	0.950	0.950	0.945	0.948	0.945
		3.424	0.801	0.385	2.411	0.673	0.362	1.885	0.608	0.351
		2.002	0.194	0.044	0.778	0.099	0.024	0.377	0.055	0.015
	0.8	0.944	0.950	0.950	0.938	0.950	0.952	0.925	0.953	0.950
		5.074	1.014	0.393	3.338	0.765	0.335	2.378	0.617	0.304
		3.140	0.279	0.065	1.132	0.132	0.035	0.505	0.067	0.027
100	0.2	0.945	0.947	0.950	0.950	0.950	0.946	0.954	0.948	0.948
		1.583	0.398	0.208	1.164	0.353	0.201	0.958	0.330	0.197
		0.991	0.095	0.027	0.385	0.056	0.020	0.207	0.039	0.017
	0.5	0.945	0.947	0.952	0.948	0.954	0.949	0.950	0.954	0.950
		3.412	0.707	0.301	2.296	0.559	0.273	1.702	0.478	0.258
		2.235	0.187	0.042	0.794	0.090	0.021	0.378	0.048	0.012
	0.8	0.943	0.939	0.947	0.943	0.947	0.950	0.942	0.951	0.951
		5.236	0.970	0.341	3.401	0.710	0.277	2.359	0.543	0.238
		3.227	0.285	0.067	1.273	0.134	0.032	0.547	0.064	0.017
200	0.2	0.946	0.949	0.952	0.950	0.954	0.951	0.951	0.950	0.947
		1.509	0.328	0.156	1.033	0.276	0.147	0.799	0.248	0.142
		3.206	0.083	0.021	0.370	0.044	0.013	0.177	0.027	0.010
	0.5	0.942	0.943	0.953	0.946	0.950	0.948	0.951	0.950	0.950
		3.407	0.643	0.247	2.232	0.489	0.213	1.582	0.393	0.193
		2.245	0.183	0.041	0.820	0.086	0.020	0.361	0.044	0.010
	0.8	0.938	0.940	0.940	0.944	0.947	0.950	0.947	0.948	0.948
		5.348	0.958	0.311	3.427	0.676	0.240	2.335	0.499	0.196
		3.871	0.296	0.068	1.303	0.133	0.032	0.551	0.065	0.015

in the first experiment, that is set $m = T = \sum_{k=1}^{n} X_k$. Of course, we consider only the case when the value of T is greater than zero. Note also that if we have some prior knowledge of the type $p_2 > p_1$, when on average the second sample will have a larger sample size than n which is the sample size for the first sample. In this case for the estimate of $1/p_2$ it is natural to consider the statistics $\overline{Y}_T = \nu/T$, where the conditional distribution of ν is the Pascal distribution $P(T, p_2)$ and the unconditional distribution is obtained by taking the expectation of this distribution by the truncated at zero Binomial distribution T. The estimate of the parameter θ is $\hat{\theta}_n = \overline{X}_n \overline{Y}_T = \nu/n$.

Lemma 1. *If $n \to \infty$, then the estimate $\hat{\theta}_n$ is asymptotically normal with the mean $\mu = \theta$ and variance $\sigma^2 = \theta\left(2p_2^{-1} - \theta - 1\right)/n$.*

Table 4. Coverage probability, mean width, and standard deviation for the interval (dd 3) for small probabilities

n_1	n_2	50			100			200		
	$p_1 \setminus p_2$	0.05	0.1	0.2	0.05	0.1	0.2	0.05	0.1	0.2
50	0.05	0.902	0.901	0.887	0.898	0.899	0.893	0.897	0.900	0.887
		4.815	1.770	0.687	2.914	1.282	0.597	2.203	1.096	0.549
		6.546	2.277	0.453	2.770	0.795	0.292	1.287	0.532	0.242
	0.1	0.932	0.947	0.942	0.918	0.938	0.930	0.914	0.938	0.931
		8.376	3.004	1.088	4.924	2.048	0.906	3.488	1.694	0.829
		10.906	3.572	0.591	4.533	1.083	0.322	1.664	0.595	0.230
	0.2	0.925	0.945	0.953	0.919	0.941	0.950	0.898	0.942	0.947
		15.060	5.127	1.731	8.502	3.309	1.365	5.633	2.545	1.182
		19.144	6.186	0.906	7.979	1.975	0.436	2.645	0.797	0.258
100	0.05	0.927	0.931	0.936	0.946	0.931	0.928	0.934	0.932	0.931
		6.431	1.870	0.592	3.208	1.162	0.488	2.097	0.942	0.445
		11.073	3.539	0.348	4.146	0.848	0.184	1.145	0.361	0.130
	0.1	0.938	0.945	0.948	0.948	0.948	0.949	0.948	0.948	0.945
		11.968	3.256	0.939	5.438	1.870	0.741	3.433	1.448	0.648
		20.387	6.414	0.547	6.523	1.163	0.247	1.822	0.488	0.153
	0.2	0.939	0.942	0.952	0.935	0.946	0.948	0.929	0.944	0.952
		22.893	5.517	1.571	9.771	3.198	1.162	5.897	2.297	0.961
		38.545	10.137	0.886	12.322	1.963	0.387	3.955	0.760	0.207
200	0.05	0.895	0.933	0.949	0.945	0.946	0.946	0.953	0.948	0.943
		8.710	1.881	0.487	3.240	1.000	0.386	1.928	0.769	0.338
		19.614	5.574	0.283	6.112	0.648	0.132	1.099	0.272	0.083
	0.1	0.875	0.936	0.948	0.946	0.952	0.947	0.947	0.951	0.950
		17.695	3.489	0.853	6.229	1.695	0.612	3.322	1.246	0.512
		38.760	11.037	1.629	12.775	1.050	0.207	2.104	0.434	0.114
	0.2	0.867	0.936	0.947	0.942	0.947	0.951	0.946	0.949	0.949
		34.523	6.768	1.488	12.422	3.076	1.035	6.069	2.093	0.803
		75.401	22.292	0.933	27.438	2.038	0.359	4.552	0.738	0.176

Proof. The characteristic function of Pascal's distribution $P(m, p_2)$ (the distribution of ν given $T = m$) is $\varphi_m(t) = \lambda^m(t)$, where

$$\lambda(t) = \frac{p_2 \mathrm{e}^{\mathrm{i}t}}{1 - (1 - p_2)\mathrm{e}^{\mathrm{i}t}}.$$

Under the assumption that T has truncated at zero binomial distribution, the characteristic function of the unconditional distribution of ν takes the form

$$\varphi(t) = \frac{1}{1 - (1 - p_1)^n} \cdot \sum_{i=1}^{n} \binom{n}{i} [p_1 \lambda(t)]^i (1 - p_1)^{n-i}$$

Table 5. Coverage probability, mean width, and standard deviation for the interval (di 1)

n	m	20			50			100		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
50	0.2	0.903	0.907	0.918	0.917	0.916	0.917	0.912	0.916	0.916
		1.301	0.483	0.278	1.153	0.444	0.266	1.098	0.429	0.252
		0.368	0.121	0.053	0.234	0.081	0.039	0.187	0.066	0.036
	0.5	0.921	0.926	0.948	0.934	0.939	0.941	0.941	0.941	0.938
		2.389	0.828	0.419	1.834	0.668	0.371	1.606	0.605	0.353
		0.563	0.184	0.075	0.279	0.086	0.033	0.165	0.051	0.019
	0.8	0.931	0.926	0.936	0.937	0.946	0.954	0.942	0.949	0.945
		3.340	1.095	0.478	2.264	0.767	0.367	1.773	0.621	0.322
		0.750	0.252	0.114	0.318	0.102	0.043	0.172	0.056	0.028
100	0.2	0.914	0.928	0.932	0.929	0.929	0.936	0.932	0.932	0.934
		1.091	0.393	0.214	0.909	0.342	0.199	0.837	0.322	0.194
		0.281	0.090	0.038	0.162	0.053	0.023	0.115	0.039	0.019
	0.5	0.923	0.930	0.936	0.939	0.943	0.947	0.945	0.942	0.949
		2.198	0.734	0.346	1.574	0.551	0.287	1.306	0.475	0.265
		0.518	0.166	0.073	0.236	0.074	0.029	0.136	0.042	0.016
	0.8	0.929	0.928	0.918	0.938	0.942	0.952	0.944	0.947	0.950
		3.243	1.047	0.439	2.128	0.702	0.314	1.504	0.541	0.261
		0.731	0.241	0.117	0.301	0.101	0.043	0.161	0.051	0.022
200	0.2	0.918	0.922	0.930	0.937	0.941	0.946	0.937	0.941	0.939
		0.961	0.331	0.168	0.737	0.269	0.151	0.649	0.244	0.143
		0.238	0.076	0.032	0.121	0.038	0.016	0.081	0.026	0.011
	0.5	0.925	0.928	0.925	0.941	0.941	0.948	0.946	0.948	0.951
		2.082	0.685	0.301	1.421	0.481	0.231	1.115	0.391	0.204
		0.480	0.157	0.073	0.211	0.063	0.028	0.118	0.036	0.014
	0.8	0.924	0.923	0.905	0.936	0.939	0.943	0.950	0.942	0.950
		3.190	1.023	0.414	2.063	0.666	0.283	1.509	0.494	0.223
		0.725	0.239	0.121	0.301	0.098	0.044	0.151	0.049	0.021

$$= \frac{[p_1\lambda(t) + (1-p_1)]^n - (1-p_1)^n}{1-(1-p_1)^n}.$$

The statement of the lemma follows now from the Taylor expansion of the function $\varphi(t)$. □

The lemma immediately implies the following result.

Theorem 2. *If $n \to \infty$, the asymptotic $(1-\alpha)$-confidence interval for the parametric function θ is defined by the inequality*

$$\left|\theta - \hat{\theta}_n\right| \le \lambda_\alpha \sqrt{\frac{\theta}{n}\left(2\overline{Y}_T - \theta - 1\right)}.$$

Table 6. Coverage probability, mean width, and standard deviation for the interval (di 2)

n	m	20			50			100		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
50	0.2	0.931	0.933	0.936	0.937	0.931	0.928	0.920	0.927	0.926
		2.404	0.849	0.436	1.904	0.693	0.402	1.686	0.650	0.389
		1.116	5.041	0.189	1.058	0.243	0.110	0.485	0.177	0.095
	0.5	0.951	0.951	0.953	0.949	0.952	0.942	0.949	0.942	0.950
		3.111	1.003	0.474	2.122	0.754	0.408	1.803	0.669	0.386
		0.737	0.230	0.090	0.300	0.093	0.035	0.174	0.053	0.021
	0.8	0.949	0.948	0.936	0.952	0.950	0.954	0.949	0.948	0.944
		4.040	1.237	0.511	2.469	0.810	0.381	1.868	0.645	0.331
		0.948	0.302	0.131	0.348	0.111	0.048	0.182	0.059	0.031
100	0.2	0.951	0.953	0.949	0.950	0.945	0.949	0.948	0.946	0.943
		1.579	0.524	0.265	1.151	0.423	0.239	1.022	0.389	0.231
		0.376	0.117	0.046	0.173	0.054	0.021	0.109	0.035	0.015
	0.5	0.953	0.950	0.950	0.949	0.953	0.952	0.949	0.953	0.947
		2.708	0.853	0.374	1.746	0.598	0.303	1.401	0.504	0.277
		0.648	0.206	0.083	0.261	0.081	0.031	0.144	0.044	0.016
	0.8	0.950	0.947	0.929	0.951	0.950	0.958	0.948	0.948	0.952
		3.885	1.177	0.462	2.295	0.737	0.323	1.672	0.557	0.266
		0.910	0.290	0.131	0.332	0.107	0.045	0.168	0.054	0.023
200	0.2	0.949	0.953	0.955	0.948	0.951	0.951	0.952	0.950	0.948
		1.244	0.403	0.191	0.856	0.304	0.165	0.726	0.270	0.156
		0.305	0.095	0.038	0.136	0.042	0.017	0.082	0.026	0.011
	0.5	0.948	0.945	0.942	0.950	0.949	0.952	0.952	0.952	0.952
		2.518	0.777	0.320	1.544	0.510	0.240	1.171	0.406	0.209
		0.504	0.193	0.083	0.231	0.074	0.029	0.124	0.039	0.015
	0.8	0.951	0.943	0.920	0.948	0.947	0.947	0.949	0.952	0.948
		3.800	1.141	0.435	2.209	0.696	0.289	1.565	0.506	0.226
		0.890	0.286	0.133	0.323	0.105	0.046	0.160	0.051	0.022

The interval bounded by the points

$$\hat{\theta}_n \pm \lambda_\alpha \sqrt{\frac{\hat{\theta}_n}{n}\left(2\overline{Y}_T - \hat{\theta}_n - 1\right)} \qquad (\text{di } 1 \cdot \text{T})$$

is the asymptotically $(1-\alpha)$*-confidence interval for* θ.

Characteristics of this interval are presented in Table 9.

The results of statistical modeling presented in this table probably show that the confidence di1 · T is over conservative, but it is remarkable that its precision properties remain almost the same as for the linear interval di 1.

Table 7. Coverage probability, mean width, and standard deviation for the interval (di 3)

n	m	20			50			100		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
50	0.2	0.939	0.944	0.942	0.942	0.942	0.940	0.937	0.940	0.940
		1.407	0.518	0.293	1.223	0.470	0.280	1.156	0.451	0.276
		0.386	0.124	0.054	0.231	0.078	0.038	0.182	0.064	0.034
	0.5	0.942	0.949	0.952	0.950	0.949	0.946	0.948	0.951	0.947
		2.492	0.852	0.426	1.882	0.681	0.376	1.638	0.615	0.358
		0.594	0.189	0.076	0.279	0.087	0.032	0.166	0.051	0.020
	0.8	0.946	0.941	0.940	0.948	0.946	0.951	0.950	0.949	0.948
		3.425	1.113	0.484	2.295	0.772	0.370	1.791	0.624	0.323
		0.772	0.256	0.117	0.326	0.105	0.045	0.171	0.057	0.029
100	0.2	0.947	0.946	0.948	0.947	0.948	0.946	0.945	0.948	0.946
		1.154	0.410	0.221	0.940	0.353	0.205	0.861	0.331	0.199
		0.296	0.095	0.040	0.162	0.053	0.023	0.115	0.038	0.019
	0.5	0.948	0.940	0.948	0.946	0.951	0.950	0.950	0.951	0.951
		2.269	0.755	0.350	1.600	0.558	0.289	1.321	0.479	0.267
		0.526	0.175	0.074	0.241	0.076	0.029	0.137	0.043	0.016
	0.8	0.941	0.940	0.925	0.949	0.943	0.952	0.951	0.949	0.952
		3.328	1.073	0.442	2.156	0.707	0.316	1.615	0.544	0.252
		0.770	0.252	0.120	0.311	0.101	0.043	0.159	0.051	0.022
200	0.2	0.943	0.944	0.946	0.947	0.947	0.952	0.948	0.954	0.950
		0.993	0.341	0.171	0.754	0.274	0.152	0.660	0.248	0.145
		0.246	0.078	0.033	0.122	0.039	0.016	0.030	0.025	0.011
	0.5	0.942	0.938	0.929	0.950	0.950	0.956	0.950	0.946	0.951
		2.153	0.697	0.303	1.439	0.485	0.233	1.125	0.392	0.204
		0.500	0.164	0.074	0.211	0.063	0.028	0.118	0.037	0.014
	0.8	0.940	0.935	0.921	0.945	0.950	0.948	0.946	0.950	0.949
		3.275	1.039	0.418	2.036	0.671	0.284	1.518	0.497	0.223
		0.757	0.249	0.120	0.305	0.097	0.044	0.155	0.049	0.022

Inverse-direct. If the first sample is obtained by the inverse scheme of binomial sampling with the expected number of successes m and the second sample is obtained by the direct binomial sampling with the sample size n, then the asymptotically $(n, m \to \infty)$ $(1-\alpha)$-confidence intervals (1)–(3) for θ can be written in terms of

$$\hat{p}_1 = (m-1)/(\nu-1), \quad \hat{p}_2 = \overline{X}_n, \quad \hat{p}_2^{-1} = \frac{n+1}{n\overline{X}_n + 1}$$

as

$$\frac{m-1}{\nu-1} \cdot \frac{n+1}{T+1} \left(1 \mp \lambda_\alpha \sqrt{\frac{1-\hat{p}_1}{m} + \frac{1-\hat{p}_2}{n\hat{p}_2}} \right), \tag{id 1}$$

Table 8. Coverage probability, mean width, and standard deviation for the interval (di 3) for small probabilities

n	m	20			50			100		
	$p_1 \setminus p_2$	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15
50	0.05	0.883	0.880	0.885	0.878	0.877	0.880	0.879	0.871	0.870
		2.500	1.238	0.826	2.328	1.160	0.774	2.288	1.125	0.747
		1.214	0.603	0.392	1.012	0.505	0.333	0.939	0.478	0.320
	0.1	0.925	0.927	0.927	0.924	0.923	0.920	0.922	0.920	0.922
		3.847	1.908	1.257	3.483	1.734	1.148	3.369	1.676	1.112
		1.327	0.642	0.420	0.946	0.470	0.312	0.825	0.414	0.270
	0.15	0.940	0.943	0.943	0.936	0.940	0.939	0.943	0.935	0.936
		4.917	2.417	1.606	4.296	2.143	1.421	4.099	2.036	1.358
		1.491	0.717	0.469	0.971	0.471	0.308	0.770	0.386	0.251
100	0.05	0.926	0.930	0.933	0.924	0.926	0.928	0.923	0.925	0.925
		1.969	0.987	0.650	1.791	0.898	0.595	1.734	0.854	0.575
		0.699	0.345	0.220	0.516	0.253	0.167	0.451	0.222	0.148
	0.1	0.940	0.942	0.944	0.940	0.938	0.940	0.945	0.936	0.944
		3.061	1.508	0.993	2.630	1.305	0.864	2.478	1.233	0.820
		0.919	0.443	0.284	0.574	0.284	0.184	0.446	0.226	0.146
	0.15	0.945	0.940	0.943	0.944	0.947	0.945	0.946	0.945	0.943
		3.979	1.958	1.288	3.284	1.633	1.078	3.037	1.512	1.005
		1.105	0.540	0.351	0.640	0.307	0.203	0.458	0.233	0.152
200	0.05	0.941	0.947	0.939	0.934	0.939	0.947	0.939	0.956	0.935
		1.557	0.764	0.506	1.362	0.673	0.444	1.282	0.628	0.425
		0.469	0.225	0.156	0.301	0.151	0.096	0.248	0.116	0.079
	0.1	0.942	0.935	0.934	0.964	0.955	0.938	0.948	0.953	0.943
		2.488	1.231	0.803	2.028	0.997	0.664	1.843	0.912	0.613
		0.658	0.337	0.222	0.375	0.181	0.121	0.274	0.138	0.093
	0.15	0.943	0.944	0.956	0.959	0.949	0.947	0.961	0.943	0.936
		3.361	1.640	1.076	2.606	1.305	0.853	2.282	1.152	0.761
		0.871	0.424	0.262	0.453	0.230	0.141	0.316	0.157	0.104

$$\frac{m-1}{\nu-1}\cdot\frac{n+1}{T+1}\left(1\pm\lambda_\alpha\sqrt{\frac{1-\hat{p}_1}{m}+\frac{1-\hat{p}_2}{n\hat{p}_2}}\right)^{-1}, \qquad \text{(id 2)}$$

$$\frac{m-1}{\nu-1}\cdot\frac{n+1}{T+1}\exp\left\{\mp\lambda_\alpha\sqrt{\frac{1-\hat{p}_1}{m}+\frac{1-\hat{p}_2}{n\hat{p}_2}}\right\}. \qquad \text{(di 3)}$$

Monte-Carlo estimations of characteristics of these intervals are presented in Tables 10–13.

The same as for the previous direct-inverse sampling scheme, it is possible to make a conclusion that for all kinds of intervals the coverage probability is almost not influenced by m, the number of expected

Table 9. Coverage probability, mean width, and standard deviation for the interval (di 1 · T)

n	50			100			200		
$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
	0.969	0.963	0.944	0.979	0.966	0.959	0.984	0.974	0.963
0.2	1.548	0.558	0.304	1.103	0.396	0.217	0.781	0.281	0.154
	0.291	0.095	0.045	0.144	0.048	0.022	0.072	0.024	0.010
	0.973	0.973	0.959	0.977	0.973	0.961	0.978	0.974	0.962
0.5	2.222	0.777	0.406	1.576	0.551	0.288	1.115	0.391	0.204
	0.319	0.101	0.040	0.159	0.050	0.020	0.078	0.025	0.010
	0.959	0.962	0.962	0.965	0.959	0.963	0.967	0.962	0.960
0.8	2.474	0.826	0.388	1.750	0.584	0.275	1.237	0.413	0.195
	0.348	0.112	0.050	0.171	0.056	0.024	1.237	0.028	0.012

success in the first sample. Moreover, the same remark is equally applicable to the precision properties of the intervals.

The coverage probability by the linear interval (id 1) is significantly lower than nominal (see Table 10). We can recommend to apply this interval only if it is known that the success probability in the second sample $p_2 > 0.5$. However, the hyperbolic interval (id 2) (Table 11) has the coverage probability which is equal to the nominal for all considered sample sizes and success probabilities, but the precision properties of hyperbolic interval significantly inferior to the linear.

Precision properties of the logarithmic interval (id 3) (Table 12) are almost the same as for the linear and the coverage probability is a little bit less than for the hyperbolic, but still is within the acceptable error 0.01. The precision and reliability properties of this interval for small values of p_1 and p_2 are presented in Table 13. According to these results for small values of success probabilities, it is possible to recommend the logarithmic interval only for the sample sizes of the second sample $n \geq 100$.

It would be interesting to use the value ν from the first sample in the planning of the second sample, that is, to use the estimate

$$\hat{\theta}_\nu = \frac{(m-1)(\nu+1)}{(\nu-1)(T_\nu+1)} \approx \frac{m}{T_\nu+1}, \tag{4}$$

where

$$T_\nu = \sum_{k=1}^{\nu} X_k.$$

The random variable ν does not depend on $X_1, X_2, \ldots$, and hence it is possible to calculate the mean value and variance of T_ν and find its distribution.

Lemma 2. *The characteristic function of the statistic of the statistics T_ν is*

$$\varphi(t) = \left[\frac{p_1\left(1-p_2+p_2\mathrm{e}^{\mathrm{i}t}\right)}{1-\left(1-p_2+p_2\mathrm{e}^{\mathrm{i}t}\right)(1-p_1)}\right]^m.$$

Statistic T_ν is asymptotically ($m \to \infty$) normal with parameters

$$\mu = \mathbf{E}T_\nu = m\frac{p_2}{p_1}, \quad \sigma^2 = \mathbf{Var}T_\nu = m\frac{p_2}{p_1^2}(p_1+p_2-2p_1p_2).$$

Proof. Characteristic function $\varphi(t)$ of statistic T_ν can be obtained from the characteristic function of the binomial distribution

$$\varphi(t) = \mathbf{E}\left(1-p_2+p_2\mathrm{e}^{\mathrm{i}t}\right)^\nu$$

Table 10. Coverage probability, mean width, and standard deviation for the interval (id 1)

m	n	50			100			200		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
20	0.2	0.895	0.924	0.931	0.919	0.937	0.937	0.925	0.932	0.935
		1.399	0.385	0.206	1.118	0.349	0.201	0.957	0.331	0.198
		0.745	0.107	0.042	0.386	0.076	0.037	0.251	0.065	0.036
	0.5	0.905	0.934	0.938	0.925	0.935	0.930	0.931	0.936	0.930
		3.273	0.829	0.418	2.517	0.726	0.400	2.065	0.671	0.389
		1.727	0.202	0.056	0.845	0.121	0.042	0.457	0.082	0.035
	0.8	0.900	0.932	0.940	0.923	0.945	0.924	0.934	0.936	0.917
		4.834	1.087	0.469	3.515	0.876	0.424	2.709	0.750	0.400
		2.638	0.280	0.070	1.163	0.134	0.048	0.581	0.082	0.048
50	0.2	0.901	0.931	0.943	0.901	0.937	0.945	0.898	0.930	0.945
		1.253	0.299	0.142	1.215	0.264	0.111	1.183	0.244	0.093
		0.671	0.078	0.023	0.725	0.070	0.018	0.819	0.067	0.017
	0.5	0.898	0.931	0.945	0.899	0.931	0.948	0.893	0.929	0.944
		3.036	0.683	0.299	2.976	0.626	0.245	2.934	0.591	0.212
		1.632	0.179	0.044	1.668	0.172	0.041	1.706	0.166	0.041
	0.8	0.899	0.933	0.953	0.901	0.932	0.938	0.899	0.929	0.929
		4.684	0.982	0.371	4.658	0.938	0.326	4.641	0.914	0.303
		2.567	0.272	0.064	2.633	0.264	0.065	2.641	0.252	0.058
100	0.2	0.924	0.935	0.947	0.923	0.939	0.948	0.926	0.942	0.948
		0.943	0.253	0.133	0.875	0.211	0.100	0.841	0.186	0.079
		0.327	0.047	0.017	0.309	0.038	0.011	0.295	0.034	0.008
	0.5	0.925	0.939	0.947	0.925	0.939	0.947	0.925	0.943	0.948
		2.237	0.554	0.273	2.122	0.481	0.212	2.055	0.441	0.173
		0.761	0.095	0.026	0.734	0.087	0.022	0.728	0.083	0.021
	0.8	0.928	0.947	0.958	0.925	0.943	0.952	0.927	0.945	0.949
		3.335	0.742	0.314	3.273	0.690	0.252	3.228	0.662	0.231
		1.142	0.131	0.031	1.144	0.131	0.031	1.111	0.131	0.032

by taking expectation of the number of trails ν and the mathematical expectation is calculated according to Pascal distribution:

$$\mathbf{P}(\nu = k|m, p_1) = \binom{m-1}{k-1} p_1^m (1-p_1)^{k-m}, \quad k \geq m.$$

Therefore, the derivation of the characteristic function of T_ν reduces to a calculation of the sum

$$\mathbf{S} = \sum_{k=m}^{\infty} \binom{m-1}{k-1} a^{k-m} = \sum_{k=0}^{\infty} \binom{m-1}{k+m-1} a^k,$$

where $a = \left(1 - p_2 + p_2 e^{\mathbf{i}t}\right) p_1^m (1-p_1)$.

Table 11. Coverage probability, mean width, and standard deviation for the interval (id 2)

m	n	50			100			200		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
20	0.2	0.953	0.949	0.944	0.954	0.941	0.950	0.947	0.943	0.942
		2.284	0.501	0.250	1.644	0.432	0.238	1.250	0.401	0.234
		10.522	0.149	0.049	0.689	0.096	0.042	0.339	0.076	0.041
	0.5	0.956	0.947	0.939	0.953	0.946	0.938	0.942	0.940	0.939
		3.411	1.011	0.473	3.401	0.839	0.447	2.513	0.756	0.434
		289.453	0.274	0.063	1.457	0.140	0.042	0.583	0.085	0.033
	0.8	0.956	0.957	0.931	0.953	0.947	0.920	0.953	0.933	0.916
		10.118	1.236	0.500	4.477	0.951	0.446	3.072	0.797	0.418
		27.648	0.343	0.079	1.908	0.158	0.057	0.703	0.094	0.055
50	0.2	0.955	0.952	0.950	0.954	0.954	0.950	0.952	0.945	0.949
		2.236	0.351	0.154	2.268	0.297	0.118	2.079	0.273	0.096
		2.085	0.101	0.026	4.529	0.088	0.020	3.928	0.085	0.018
	0.5	0.950	0.950	0.951	0.950	0.951	0.950	0.952	0.943	0.947
		5.204	0.776	0.317	4.912	0.700	0.255	4.771	0.656	0.219
		4.625	0.226	0.048	4.239	0.217	0.045	4.174	0.212	0.044
	0.8	0.956	0.958	0.955	0.952	0.945	0.946	0.954	0.950	0.940
		7.728	1.087	0.383	7.554	1.031	0.337	7.530	1.002	0.311
		6.779	0.320	0.069	6.636	0.327	0.071	6.705	0.319	0.071
100	0.2	0.955	0.947	0.949	0.950	0.950	0.946	0.952	0.950	0.949
		1.220	0.282	0.143	1.101	0.227	0.104	1.028	0.198	0.081
		0.512	0.054	0.018	0.477	0.042	0.012	0.438	0.038	0.009
	0.5	0.956	0.952	0.947	0.951	0.954	0.950	0.952	0.953	0.950
		2.820	0.601	0.286	2.624	0.513	0.218	2.515	0.464	0.177
		1.176	0.106	0.027	1.147	0.095	0.022	1.130	0.091	0.021
	0.8	0.952	0.953	0.950	0.950	0.949	0.956	0.950	0.949	0.949
		4.086	0.783	0.322	3.998	0.724	0.267	3.911	0.692	0.234
		1.713	0.148	0.033	1.757	0.145	0.032	1.772	0.143	0.033

By the integral representation of Gamma-function

$$(k+m-1)! = \Gamma(k+m) = \int_0^\infty x^{k+m-1} \mathrm{e}^{-x} dx,$$

we obtain that

$$\mathbf{S} = \frac{1}{(m-1)!} \int_0^\infty x^{m-1} \mathrm{e}^{-x} \sum_{k=0}^\infty \frac{(ax)^k}{k!} dx$$

Table 12. Coverage probability, mean width, and standard deviation for the interval (id 3)

m	n	50			100			200		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
20	0.2	0.942	0.947	0.946	0.945	0.944	0.943	0.945	0.943	0.941
		1.528	0.398	0.212	1.179	0.362	0.206	0.997	0.341	0.203
		0.887	0.110	0.043	0.422	0.080	0.039	0.252	0.058	0.037
	0.5	0.938	0.944	0.945	0.950	0.945	0.940	0.948	0.940	0.939
		3.531	0.853	0.427	2.637	0.743	0.407	2.125	0.683	0.396
		2.224	0.213	0.058	0.903	0.122	0.042	0.476	0.082	0.034
	0.8	0.941	0.949	0.943	0.947	0.949	0.929	0.947	0.935	0.914
		5.240	1.105	0.474	3.679	0.886	0.429	2.759	0.757	0.403
		3.303	0.282	0.070	1.291	0.140	0.050	0.603	0.085	0.050
50	0.2	0.938	0.948	0.949	0.940	0.948	0.945	0.938	0.946	0.945
		1.343	0.308	0.144	1.305	0.270	0.112	1.255	0.248	0.093
		0.786	0.082	0.023	0.822	0.074	0.018	0.756	0.069	0.017
	0.5	0.944	0.944	0.949	0.936	0.944	0.949	0.937	0.946	0.945
		3.302	0.697	0.302	3.155	0.634	0.247	3.077	0.603	0.213
		2.081	0.186	0.045	1.986	0.176	0.042	1.807	0.172	0.042
	0.8	0.939	0.946	0.952	0.936	0.944	0.946	0.940	0.942	0.940
		4.984	0.993	0.373	4.965	0.955	0.329	5.007	0.934	0.305
		2.843	0.280	0.065	3.122	0.275	0.067	4.232	0.272	0.068
100	0.2	0.946	0.948	0.946	0.942	0.950	0.948	0.946	0.949	0.950
		0.972	0.258	0.134	0.896	0.212	0.101	0.858	0.137	0.079
		0.348	0.048	0.017	0.316	0.038	0.011	0.300	0.034	0.009
	0.5	0.946	0.954	0.949	0.944	0.947	0.948	0.945	0.944	0.947
		2.305	0.562	0.274	2.189	0.488	0.213	2.122	0.443	0.173
		0.813	0.095	0.026	0.791	0.089	0.022	0.756	0.085	0.020
	0.8	0.951	0.955	0.948	0.944	0.948	0.955	0.942	0.946	0.946
		3.413	0.750	0.315	3.357	0.694	0.252	3.317	0.668	0.232
		1.175	0.133	0.032	1.197	0.133	0.032	1.196	0.130	0.032

$$= \frac{1}{(m-1)!} \int_0^\infty x^{m-1} e^{-x(1-a)} dx = (1-a)^{-m}.$$

Expanding $\ln \varphi(t)$ by Maclaurin series we find μ and σ^2. Finally, a similar expansion of logarithm of the characteristic function of the normed statistic $(T-\mu)/\sigma$ with further taking limit $m \to \infty$ proofs the asymptotic normality of its distribution. $\square$

The statement above provides us two possibilities for a construction of a confidence interval for $\theta = p_1/p_2$. Since T_ν/m is an unbiased estimate for the ratio $\theta^{-1} = p_2/p_1$, substituting for the variance

Table 13. Coverage probability, mean width, and standard deviation for the interval (id 3) for small probabilities

m	n	50			100			200		
	$p_1 \setminus p_2$	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15
20	0.05	0.909	0.935	0.947	0.954	0.945	0.943	0.951	0.949	0.945
		4.062	1.310	0.633	2.522	0.842	0.455	1.700	0.634	0.382
		5.216	1.907	0.645	2.708	0.493	0.220	0.923	0.248	0.125
	0.1	0.906	0.936	0.932	0.945	0.943	0.948	0.948	0.940	0.952
		7.570	2.315	1.237	5.038	1.616	0.916	3.324	1.271	0.740
		9.624	2.714	1.672	5.579	0.927	0.401	1.969	0.491	0.222
	0.15	0.907	0.938	0.938	0.920	0.938	0.951	0.941	0.954	0.941
		10.277	4.030	1.781	7.206	2.465	1.323	4.812	1.859	1.099
		12.892	6.147	1.172	8.799	1.620	0.538	2.743	0.676	0.335
50	0.05	0.905	0.925	0.950	0.913	0.957	0.962	0.952	0.932	0.953
		3.729	1.115	0.594	2.282	0.731	0.398	1.485	0.527	0.297
		4.805	1.403	0.799	3.663	0.379	0.185	1.096	0.202	0.081
	0.1	0.914	0.924	0.927	0.933	0.938	0.950	0.941	0.951	0.937
		7.540	2.308	1.119	4.753	1.444	0.763	3.027	1.050	0.588
		9.469	3.438	0.855	5.872	0.856	0.323	1.869	0.372	0.166
	0.15	0.899	0.927	0.932	0.926	0.936	0.938	0.936	0.956	0.949
		10.319	3.717	1.602	7.169	2.278	1.126	4.480	1.545	0.856
		13.079	5.856	1.268	10.699	3.818	0.502	2.848	0.550	0.236
100	0.05	0.911	0.924	0.931	0.926	0.931	0.930	0.949	0.947	0.943
		3.393	1.124	0.558	2.381	0.687	0.369	1.462	0.480	0.258
		4.228	1.509	0.839	3.404	0.392	0.163	1.079	0.172	0.071
	0.1	0.925	0.930	0.933	0.927	0.961	0.943	0.940	0.948	0.944
		7.182	2.135	1.037	4.692	1.431	0.711	2.838	0.953	0.518
		9.346	2.074	0.704	6.127	0.824	0.329	1.874	0.350	0.148
	0.15	0.924	0.919	0.936	0.917	0.940	0.944	0.956	0.946	0.948
		10.547	3.265	1.563	6.737	2.090	1.068	4.318	1.417	0.782
		13.216	3.799	1.104	9.229	1.200	0.464	2.533	0.528	0.226

of T_ν/m its estimate

$$\hat{\sigma}^2 = \frac{1}{m}\left[\frac{T_\nu}{m} + \frac{T_\nu^2}{m^2}\left(1 - 2\frac{m-1}{\nu-1}\right)\right],$$

we obtain asymptotically $(1-\alpha)$-confidence interval for θ^{-1} with boundaries

$$\frac{T_\nu}{m} \pm \lambda_\alpha \hat{\sigma}.$$

Table 14. Coverage probability, mean width, and standard deviation for the interval (id $\cdot\, \nu 1$)

m	50			100			200		
$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
	0.884	0.794	0.713	0.874	0.788	0.726	0.875	0.789	0.713
0.2	0.651	0.151	0.073	0.422	0.102	0.050	0.287	0.071	0.035
	0.203	0.033	0.014	0.086	0.015	0.007	0.039	0.007	0.003
	0.946	0.948	0.949	0.945	0.945	0.948	0.948	0.946	0.951
0.5	3.148	0.625	0.294	1.811	0.416	0.200	1.180	0.286	0.139
	1.649	0.151	0.050	0.494	0.065	0.024	0.211	0.031	0.011
	0.943	0.977	0.995	0.965	0.986	0.997	0.964	0.986	0.997
0.8	9.455	1.329	0.607	4.008	0.861	0.409	2.496	0.532	0.283
	24.495	0.331	0.086	1.460	0.138	0.039	0.549	0.064	0.019

Table 15. Coverage probability, mean width, and standard deviation for the interval (id $\cdot\, \nu 2$)

m	50			100			200		
$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
	0.936	0.945	0.949	0.945	0.945	0.951	0.950	0.950	0.950
0.2	0.716	0.225	0.129	0.500	0.157	0.091	0.353	0.111	0.064
	0.169	0.034	0.015	0.081	0.016	0.007	0.040	0.008	0.003
	0.919	0.947	0.948	0.941	0.946	0.950	0.950	0.948	0.948
0.5	2.221	0.556	0.273	1.567	0.392	0.193	1.104	0.277	0.136
	0.801	0.103	0.028	0.377	0.050	0.014	0.184	0.025	0.006
	0.891	0.935	0.949	0.919	0.943	0.949	0.937	0.947	0.949
0.8	4.151	0.875	0.339	2.891	0.623	0.244	2.053	0.442	0.173
	2.041	0.217	0.050	0.879	0.106	0.025	0.439	0.052	0.012

Hence the *interval*

$$\left(\left[\frac{T_\nu}{m} + \lambda_\alpha \hat{\sigma} \right]^{-1}, \quad \left[\frac{T_\nu}{m} - \lambda_\alpha \hat{\sigma} \right]^{-1} \right) \tag{id $\cdot\, \nu 1$}$$

asymptotically ($m \to \infty$) *with the same probability* $1 - \alpha$ *covers the probabilities ratio* $\theta = p_1/p_2$.

Another method is based on the identical asymptotic distribution of the statistics $\hat{\theta}_\nu$ and m/T_ν (see formula (4)). Applying the standard delta-method (that is, expanding the statistic $\hat{\theta}_\nu$ in Taylor series), we obtain that the asymptotic ($m \to \infty$) normal distribution of the normed statistics m/T_ν is defined by the random function

$$\frac{p_1}{p_2} - \frac{p_1^2}{p_2^2} \left(\frac{T_\nu}{m} - \frac{p_2}{p_1} \right).$$

The mean value of this function equals θ and variance is

$$\sigma_1^2 = \frac{p_1^2}{m p_2^3} (p_1 + p_2 - 2p_1 p_2) = \frac{1}{m} \left[\left(\frac{p_1}{p_2} \right)^3 + \left(\frac{p_1}{p_2} \right)^2 (1 - 2p_1) \right].$$

Table 16. Coverage probability, mean width, and standard deviation for the interval (ii 1)

m_1	m_2	20			50			100		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
20	0.2	0.921	0.921	0.929	0.928	0.933	0.932	0.931	0.929	0.934
		1.105	0.398	0.219	0.925	0.349	0.204	0.855	0.329	0.200
		0.316	0.107	0.051	0.215	0.075	0.040	0.174	0.065	0.037
	0.5	0.920	0.926	0.937	0.932	0.930	0.931	0.932	0.930	0.933
		2.488	0.870	0.456	1.962	0.725	0.412	1.761	0.667	0.395
		0.625	0.204	0.086	0.336	0.111	0.048	0.237	0.078	0.038
	0.8	0.926	0.926	0.937	0.934	0.938	0.929	0.938	0.934	0.920
		3.500	1.173	0.546	2.503	0.876	0.451	2.070	0.748	0.414
		0.800	0.258	0.113	0.349	0.110	0.052	0.206	0.073	0.046
50	0.2	0.922	0.925	0.933	0.937	0.941	0.941	0.943	0.939	0.943
		0.929	0.319	0.158	0.699	0.252	0.138	0.607	0.226	0.131
		0.234	0.077	0.034	0.126	0.042	0.020	0.089	0.031	0.016
	0.5	0.923	0.927	0.937	0.939	0.942	0.946	0.942	0.941	0.942
		2.196	0.739	0.347	1.581	0.554	0.288	1.311	0.478	0.266
		0.527	0.171	0.075	0.251	0.081	0.033	0.157	0.051	0.022
	0.8	0.924	0.923	0.928	0.944	0.947	0.953	0.950	0.950	0.952
		3.314	1.072	0.462	2.218	0.743	0.348	1.715	0.591	0.301
		0.768	0.248	0.116	0.321	0.102	0.043	0.168	0.054	0.023
100	0.2	0.926	0.925	0.928	0.941	0.942	0.945	0.939	0.945	0.948
		0.863	0.287	0.132	0.606	0.211	0.107	0.496	0.178	0.097
		0.205	0.067	0.031	0.097	0.032	0.013	0.063	0.021	0.009
	0.5	0.922	0.928	0.922	0.945	0.944	0.949	0.946	0.944	0.948
		2.082	0.682	0.300	1.424	0.483	0.231	1.119	0.391	0.204
		0.482	0.157	0.074	0.213	0.070	0.029	0.125	0.040	0.016
	0.8	0.925	0.928	0.913	0.943	0.944	0.948	0.946	0.949	0.952
		3.227	1.038	0.430	2.109	0.689	0.303	1.566	0.523	0.246
		0.727	0.242	0.119	0.301	0.098	0.044	0.161	0.051	0.021

Substituting parameters p_1 and p_2^{-1} in this variance by their estimates, that is, applying instead of σ_1^2 its estimate

$$\hat{\sigma}_1^2 = \frac{1}{m}\left[\left(\frac{m}{T_\nu + 1}\right)^3 + \left(\frac{m}{T_\nu + 1}\right)^2 \left(1 - 2\frac{m-1}{\nu - 1}\right)\right],$$

we obtain *asymptotically* $(1-\alpha)$*-confidence interval for* $\theta = p_1/p_2$ as

$$\hat{\theta}_\nu \mp \lambda_\alpha \hat{\sigma}_1. \qquad (\mathrm{id}\cdot\nu 2)$$

Monte-Carlo estimates of characteristics of the intervals $(\mathrm{id}\cdot\nu 1)$ and $(\mathrm{id}\cdot\nu 2)$ are presented in Tables 14–15.

Table 17. Coverage probability, mean width, and standard deviation for the interval (ii 2)

m_1	m_2	20			50			100		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
20	0.2	0.947	0.946	0.949	0.947	0.949	0.944	0.941	0.949	0.945
		1.601	0.533	0.270	1.177	0.430	0.245	1.050	0.398	0.238
		0.467	0.150	0.063	0.259	0.089	0.046	0.208	0.075	0.041
	0.5	0.945	0.948	0.944	0.944	0.944	0.939	0.944	0.943	0.943
		3.321	1.036	0.527	2.334	0.838	0.462	2.012	0.754	0.440
		0.836	0.260	0.103	0.381	0.122	0.050	0.245	0.081	0.037
	0.8	0.951	0.953	0.948	0.942	0.941	0.931	0.938	0.932	0.916
		4.337	1.356	0.595	2.780	0.946	0.479	2.225	0.796	0.435
		1.008	0.314	0.134	0.394	0.126	0.061	0.229	0.084	0.053
50	0.2	0.953	0.947	0.945	0.934	0.940	0.942	0.938	0.942	0.947
		1.187	0.380	0.176	0.700	0.252	0.138	0.606	0.226	0.107
		0.303	0.097	0.040	0.127	0.042	0.019	0.091	0.031	0.013
	0.5	0.954	0.944	0.945	0.939	0.942	0.946	0.945	0.942	0.946
		2.722	0.857	0.378	1.577	0.554	0.289	1.312	0.478	0.267
		0.663	0.211	0.087	0.252	0.081	0.033	0.156	0.051	0.021
	0.8	0.953	0.941	0.942	0.942	0.941	0.951	0.949	0.949	0.950
		3.982	1.218	0.493	2.218	0.742	0.348	1.713	0.592	0.301
		0.926	0.302	0.131	0.322	0.104	0.043	0.168	0.053	0.023
100	0.2	0.950	0.948	0.943	0.948	0.950	0.953	0.947	0.952	0.952
		1.059	0.330	0.141	0.608	0.210	0.112	0.529	0.187	0.102
		0.261	0.082	0.034	0.098	0.032	0.015	0.067	0.022	0.010
	0.5	0.948	0.944	0.934	0.948	0.950	0.952	0.954	0.950	0.948
		2.529	0.775	0.320	1.555	0.511	0.240	1.176	0.407	0.210
		0.614	0.193	0.084	0.237	0.076	0.031	0.129	0.042	0.017
	0.8	0.951	0.944	0.920	0.950	0.950	0.954	0.952	0.949	0.953
		3.850	1.159	0.450	2.265	0.721	0.310	1.632	0.538	0.251
		0.900	0.288	0.132	0.333	0.106	0.046	0.168	0.053	0.022

According to the results from the tables, we can make a conclusion that the coverage probability for these intervals is almost free of the value of p_2, but strongly depends on p_1. Interval (id $\cdot\,\nu 1$) has good coverage probability (sometimes it may be too conservative at the expense of precision) only for $p_1 \geq 0.5$ and the interval (id $\cdot\,\nu 2$) has the coverage probability practically equal to the nominal in the region $p_1 \leq 0.5$. It is clear that with such poor performance a practical application of intervals (id $\cdot\,\nu 1$) and (id $\cdot\,\nu 2$) is quite problematic.

Inverse-inverse. For the case when both samples are obtained in the schemes of the inverse

Table 18. Coverage probability, mean width, and standard deviation for the interval (ii 3)

m_1	m_2	20			50			100		
	$p_1 \setminus p_2$	0.2	0.5	0.8	0.2	0.5	0.8	0.2	0.5	0.8
20	0.2	0.945	0.946	0.943	0.943	0.946	0.951	0.947	0.944	0.944
		1.162	0.417	0.226	0.957	0.360	0.210	0.877	0.340	0.205
		0.334	0.112	0.053	0.217	0.077	0.040	0.178	0.066	0.038
	0.5	0.944	0.940	0.941	0.944	0.943	0.936	0.939	0.940	0.940
		2.595	0.904	0.467	2.014	0.743	0.419	1.792	0.680	0.403
		0.654	0.211	0.090	0.344	0.113	0.049	0.236	0.078	0.038
	0.8	0.942	0.944	0.943	0.945	0.943	0.935	0.943	0.941	0.924
		3.618	1.201	0.553	2.550	0.886	0.455	2.089	0.756	0.418
		0.834	0.264	0.116	0.353	0.114	0.054	0.208	0.074	0.047
50	0.2	0.946	0.942	0.947	0.945	0.947	0.946	0.951	0.947	0.948
		0.963	0.327	0.161	0.713	0.257	0.140	0.614	0.229	0.132
		0.240	0.080	0.034	0.130	0.043	0.020	0.090	0.032	0.016
	0.5	0.942	0.940	0.941	0.950	0.944	0.953	0.944	0.947	0.949
		2.269	0.751	0.353	1.604	0.562	0.291	1.325	0.481	0.269
		0.541	0.177	0.077	0.254	0.083	0.033	0.159	0.051	0.021
	0.8	0.941	0.941	0.938	0.948	0.953	0.954	0.949	0.951	0.951
		3.385	1.096	0.458	2.243	0.749	0.351	1.725	0.595	0.301
		0.783	0.255	0.118	0.324	0.102	0.043	0.171	0.053	0.023
100	0.2	0.944	0.943	0.942	0.945	0.948	0.950	0.950	0.948	0.951
		0.891	0.292	0.133	0.617	0.213	0.108	0.500	0.179	0.098
		0.214	0.063	0.030	0.101	0.032	0.014	0.063	0.021	0.010
	0.5	0.947	0.940	0.936	0.947	0.946	0.949	0.947	0.951	0.950
		2.147	0.701	0.303	1.442	0.486	0.233	1.127	0.394	0.205
		0.495	0.163	0.074	0.221	0.070	0.029	0.126	0.040	0.016
	0.8	0.944	0.937	0.920	0.949	0.947	0.952	0.954	0.947	0.957
		3.313	1.055	0.432	2.132	0.694	0.305	1.578	0.526	0.247
		0.762	0.250	0.120	0.309	0.100	0.044	0.159	0.052	0.021

binomial sampling, intervals (1)–(3) can written as

$$\hat{\theta}_{m_1,m_2}\left(1 \mp \lambda_\alpha \sqrt{\frac{1-\hat{p}_1}{m_1} + \frac{1-\hat{p}_2}{m_2}}\right), \tag{ii 1}$$

$$\hat{\theta}_{m_1,m_2}\left(1 \pm \lambda_\alpha \sqrt{\frac{1-\hat{p}_1}{m_1} + \frac{1-\hat{p}_2}{m_2}}\right)^{-1}, \tag{ii 2}$$

Table 19. Coverage probability, mean width, and standard deviation for the interval (ii 3) for small probabilities

m_1	m_2	20			50			100		
	$p_1 \setminus p_2$	0.05	0.1	0.15	0.05	0.1	0.15	0.05	0.1	0.15
20	0.05	0.949	0.937	0.956	0.960	0.943	0.953	0.945	0.953	0.947
		1.293	0.651	0.415	1.048	0.525	0.344	0.960	0.480	0.318
		0.410	0.212	0.125	0.273	0.136	0.086	0.230	0.118	0.075
	0.1	0.949	0.956	0.938	0.947	0.942	0.951	0.955	0.948	0.952
		2.585	1.234	0.822	2.056	1.026	0.673	1.899	0.947	0.625
		0.771	0.363	0.261	0.523	0.262	0.163	0.428	0.213	0.145
	0.15	0.947	0.946	0.937	0.947	0.931	0.940	0.932	0.953	0.960
		3.673	1.875	1.211	3.035	1.506	0.989	2.784	1.382	0.905
		1.078	0.552	0.359	0.742	0.383	0.240	0.622	0.300	0.192
50	0.05	0.950	0.944	0.939	0.930	0.954	0.938	0.945	0.952	0.944
		1.050	0.512	0.338	0.782	0.385	0.252	0.673	0.333	0.219
		0.262	0.132	0.087	0.163	0.076	0.050	0.114	0.055	0.035
	0.1	0.940	0.928	0.927	0.943	0.940	0.945	0.953	0.959	0.946
		2.082	1.039	0.670	1.543	0.768	0.498	1.324	0.650	0.433
		0.552	0.282	0.180	0.297	0.151	0.094	0.221	0.102	0.069
	0.15	0.945	0.938	0.942	0.950	0.939	0.946	0.943	0.949	0.935
		3.102	1.533	0.997	2.286	1.121	0.742	1.949	0.964	0.636
		0.823	0.406	0.257	0.425	0.215	0.141	0.312	0.151	0.104
100	0.05	0.954	0.951	0.932	0.953	0.943	0.954	0.950	0.956	0.949
		0.981	0.475	0.309	0.672	0.330	0.216	0.545	0.270	0.177
		0.233	0.115	0.077	0.115	0.056	0.036	0.074	0.037	0.024
	0.1	0.946	0.950	0.949	0.947	0.944	0.961	0.966	0.953	0.958
		1.945	0.955	0.619	1.328	0.651	0.427	1.071	0.530	0.350
		0.485	0.232	0.151	0.228	0.110	0.068	0.139	0.072	0.046
	0.15	0.946	0.945	0.938	0.931	0.958	0.948	0.941	0.942	0.958
		2.844	1.410	0.919	1.987	0.970	0.636	1.598	0.792	0.514
		0.700	0.355	0.226	0.343	0.158	0.107	0.213	0.105	0.065

$$\hat{\theta}_{m_1,m_2} \exp\left\{\mp\lambda_\alpha\sqrt{\frac{1-\hat{p}_1}{m_1}+\frac{1-\hat{p}_2}{m_2}}\right\}, \quad \text{(ii 3)}$$

where $\hat{p}_i = (m_i - 1)/(\nu_i - 1), i = 1, 2$. The optimal unbiased estimate of $\theta = p_1/p_2$ is

$$\hat{\theta}_{m_1,m_2} = \frac{\nu_2(m_1-1)}{(\nu_1-1)m_2}.$$

Characteristics of intervals (ii 1)–(ii 3) are presented in Tables 16–19.

The precise correspondence of the coverage probability to the nominal (especially for small success probabilities) for asymptotic confidence intervals has been mentioned in papers by Lui (1995)–(2006).

Table 20. The summary of all computational results

p_1	p_2		0.2				0.8			
	sampl schem		dd (n_1, n_2)	di (n, m)	id (m, n)	ii (m_1, m_2)	dd (n_1, n_2)	di (n, m)	id (m, n)	ii (m_1, m_2)
0.2	(n, m)		(100, 100)	(100, 20)	(20, 100)	(20, 20)	(100, 50)	(100, 20)	(20, 50)	(20, 50)
	li	$\bar{p}$	0.924	0.914	0.919	0.921	0.936	0.932	0.931	0.932
		l	1.106	1.081	1.118	1.105	0.202	0.214	0.206	0.204
		σ_l	0.358	0.281	0.386	0.316	0.027	0.038	0.042	0.040
	hy	$\bar{p}$	0.955	0.951	0.953	0.947	0.945	0.944	0.944	0.944
		l	1.622	1.579	3.401	1.601	0.244	0.265	0.250	0.245
		σ_l	0.651	0.376	1.457	0.467	0.027	0.046	0.049	0.046
	lg	$\bar{p}$	0.950	0.947	0.945	0.945	0.950	0.948	0.946	0.951
		l	1.164	1.154	1.179	1.162	0.208	0.221	0.212	0.210
		σ_l	0.385	0.296	0.422	0.334	0.027	0.040	0.043	0.040
0.5	(n, m)		(100, 100)	(100, 20)	(50, 100)	(50, 20)	(100, 50)	(100, 50)	(50, 50)	(50, 50)
	li	$\bar{p}$	0.924	0.923	0.899	0.923	0.948	0.945	0.945	0.946
		l	2.222	2.198	2.976	2.196	0.299	0.287	0.299	0.288
		σ_l	0.752	0.518	1.668	0.527	0.041	0.020	0.044	0.033
	hy	$\bar{p}$	0.956	0.953	0.950	0.954	0.949	0.952	0.951	0.946
		l	2.796	2.708	4.912	2.722	0.318	0.303	0.317	0.389
		σ_l	1.175	0.648	4.239	0.667	0.046	0.031	0.048	0.033
	lg	$\bar{p}$	0.948	0.948	0.936	0.942	0.952	0.950	0.949	0.953
		l	2.296	2.269	3.155	2.269	0.301	0.289	0.302	0.291
		σ_l	0.794	0.526	1.986	0.541	0.042	0.029	0.045	0.033
0.8	(n, m)		(50, 100)	(50, 20)	(50, 100)	(50, 20)	(50, 50)	(50, 50)	(50, 50)	(50, 50)
	li	$\bar{p}$	0.906	0.931	0.901	0.924	0.948	0.954	0.953	0.953
		l	3.237	3.346	4.658	3.314	0.389	0.367	0.371	0.348
		σ_l	1.071	0.750	2.633	0.768	0.064	0.043	0.064	0.043
	hy	$\bar{p}$	0.964	0.949	0.952	0.943	0.954	0.954	0.951	0.951
		l	3.985	4.040	7.554	3.982	0.463	0.381	0.317	0.348
		σ_l	1.577	0.948	6.636	0.926	0.070	0.048	0.048	0.043
	lg	$\bar{p}$	0.938	0.546	0.936	0.941	0.950	0.951	0.949	0.954
		l	3.338	3.425	4.965	3.385	0.393	0.370	0.302	0.351
		σ_l	1.132	0.772	3.122	0.753	0.065	0.045	0.045	0.043

The results of our modeling only support the conclusions of Lui. Moreover it is difficult to give a preference to a particular interval (ii 1)–(ii 3). All are similarly good, even for small values of success probabilities (see Table 16).

Table 21. Excerpts from the summary Table

	dd (50, 100)	di (50, 20)	id (50, 100)	ii (50, 20)
$\hat{p}$	0.906	0.931	0.901	0.924
l	3.237	3.346	4.658	3.314
σ_l	1.071	0.750	2.633	0.768

Table 22. Optimal sampling schemes

p_1	p_2	0.2	0.5	0.8
0.2	n	100	100	100
	$\hat{p}$	0.979	0.966	0.944
	l	1.103	0.396	0.304
	σ_l	0.144	0.048	0.045
0.5	n	100	50	100
	$\hat{p}$	0.977	0.973	0.961
	l	1.576	0.777	0.288
	σ_l	0.159	0.161	0.020
0.8	n	50	50	50
	$\hat{p}$	0.959	0.962	0.962
	l	2.474	0.826	0.388
	σ_l	0.348	0.112	0.050

5. COMPARATIVE ANALYSIS OF THE SAMPLING SCHEMES AND RECOMMENDATIONS FOR APPLICATIONS

In the previous section we provided the analysis of precision and reliability properties of three kinds of confidence intervals for each of four possible combinations of sampling schemes. In all cases we gave a preference to the logarithmic interval. What is left to discover is which sampling scheme is best to choose in connection with the situation at hand.

We denote the sampling schemes only by their first letters: dd, di, id, and ii. In order to support our further recommendations in the best way, we constructed one additional summary table in which the characteristics of confidence intervals for different sampling schemes are compared for the same sample sizes.

For example, if we would like to compare the characteristics of linear confidence interval for dd sampling scheme with sample sizes $n_1 = 50, n_2 = 100$ with similar characteristics for different sampling schemes. Let the true values of probabilities take values $p_1 = 0.8$ and $p_2 = 0.2$. If in scheme di we choose $n_1 = 50$, $m = n_2 \cdot p_2 = 20$, in scheme id we choose $m = n_1 \cdot p_1 = 40$, $n = n_2 = 100$, and in scheme ii we choose $m_1 = n_1 \cdot p_1 = 40$, $m_2 = n_2 \cdot p_2 = 20$, then all sampling schemes can be considered as equivalent “on average” with respect to the cost for observations. We did not conduct simulations for the sample size of $m = 40$, but from a practical point of view it is sufficient to take $m = 50$ in order to use the results from the tables.

Therefore, in Table 20 we find the cell with coordinates (0.8,0.2) and in this cell we review the row li. The results presented in Table 21 clearly show poor precision and reliability properties of the linear (denoted as li) and hyperbolic (denoted as hy) intervals where the second sample is obtained by direct sampling scheme. Moreover, for all schemes it is not possible to accept that the coverage probability is close to the nominal confidence level 0.95. But if we review the results for the logarithmic interval (denoted as lg) with coordinates (0.8, 0.2), then we can conclude that for almost the same precision characteristics (width l and standard deviation of the width σ_l), the coverage probability becomes close to the nominal.

The results that were obtained for the sampling schemes di when we consider the expected number m of successes in the second sample equals to the number T of successes in the first sample, exceed our expectations (see cell (0.8, 0.2) of Table 22): $\hat{p} = 0.959, l = 2.474$ and $\sigma_l = 0.348$.

A similar comparison analysis of the properties of confidence intervals for other true values of modeling parameters and different sample sizes, provides us the following practical recommendations for preferences among the sampling schemes.

If a statistician has a possibility to plan the sample sizes for the first and second samples and already has the estimate of p_1 for the first sample that was obtained by the inverse sampling scheme, then undoubtedly the logarithmic interval (ii 3) should be used.

If the statistician has the estimate $p_1 = T/n$ for the first sample which was obtained by the direct sampling scheme, then for the second sample the inverse binomial sampling scheme should be applied with $m = T$ and the interval (di·T) should be used.

6. ACKNOWLEDGMENT

The authors would like to thank I. Kareev for his assistance in preparing the article.

REFERENCES

1. E. L. Lehmann, *Testing Statistical Hypothesis (Second Edition)* (New York: Springer-Verlag, New York, 1986).
2. E. L. Lehmann, *Theory of Point Estimation (Second Edition)* (New York: Springer-Verlag, New York, 1998).
3. V. G. Voinov and M. S. Nikulin, *Unbiased Estimators and Their Applications: Volume 1: Univariate Case (Vol. 263)* (Kluwer Academic Publishers, Dordrecht, 1993).
4. G. E. Noether, *Two confidence intervals for the ratio of two probabilities and some measures of effectiveness*, J. Amer. Statist. Assoc. **52**, 36–45 (1957).
5. I. Guttman, *A note on a series solution of a problem in estimation*, Biometrika **45**, 565–567 (1958).
6. M. C. Sheps, *An examination of some methods of comparing several rates of proportions*, Biometrics **15**, 87–89 (1959).
7. B. J. McNeil, E. Keeler, and S. J. Adelstein, *Primer on certain elements of medical decision making*, New England J. Med. **293**, 211–215 (1975).
8. D. G. Thomas and J. J. Gart, *A table of exact confidence limits for difference and ratios of two proportions and their odds ratios*, J. Amer. Statist. Assoc. **72**, 73–76 (1977).
9. T. J. Santher and M. K. Snell, *Small sample confidence intervals for $p_1 - p_2$ and p_1/p_2 in 2×2 contingency tables*, J. Amer. Statist. Assoc. **75**, 386–394 (1980).
10. D. Katz, J. Baptista, S. P. Azen, and M. C. Pike, *Obtaining confidence intervals for the risk ratio in cohort studies*, Biometrics **34**, 469–474 (1978).
11. B. J. R. Bailey, *Confidence limits to the risk ratio*, Biometrics **43**, 201–205 (1987).
12. J. Cornfield. *A statistical problem arising from retrospective studies*, Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. IV, Berkeley: University of California Press, J. Neyman (ed.), 135–148 (1956).
13. P. A. R. Koopman, *Confidence limits to the ratio of two binomial proportions*, Biometrics **40**, 513–517 (1984).
14. O. Miettinen and M. Nurminen, *Comparative analysis of two rates* Statist. Med., **4**, 213–226 (1985).
15. J. J. Gart and J. Nam, *Approximate interval estimation of the ratio of binomial parameters: A review and corrections for skewness*, Biometrics **44**, 323–338 (1988).
16. J. J. Gart and J. Nam *Approximate interval estimation of the difference in binomial parameters: Correction for skewness and extension on multiple tables*, Biometrics **46**, 637–643 (1990).
17. J. J. Gart, *Approximate tests and interval estimation of the common relative risk in the combination of 2×2 variables*, Biometrika **72**, 673–677 (1985).
18. J. Nam, *Confidence limits for the ratio of two binomial proportions based on likelihood scores: Non-iterative method*, Biom. J. *37*, 375–379 (1995).
19. E. J. Bedrick, *A family of confidence intervals for the ratio of two binomial proportions*, Biometrics **43**, 993–998 (1987).

20. A. Kinsella, *The exact bootstrap approach to confidence intervals for the risk ratio in cohort studies*, The Statistician **36**, 345–347 (1987).
21. P. K. Coe and A. C. Tamhane, *Small sample confidence intervals for the difference, ratio and odds ratio of two success probabilities*, Commun. Statist. Simul. and Comput. **22** (4), 925–938 (1993).
22. J. Nam and W.C. Blackwelder, *Analysis of the ratio of marginal probabilities in a matched-pair setting*, Statist. Med., 689–699 (2002).
23. D. J. Bonett and R. M. Price, *Confidence intervals for a ratio of binomial proportions based on paired data*, Statist. Med. **253**, 3039–3047 (2006).
24. P. Mukhopadhyay, *Exact tests and exact confidence intervals for the ratio of two binomial proportion* (North Carolina State University Libraries, Raleigh, 2003).
25. B. M. Bennett, *On the use of the negative binomial in epidemiology*, Biom. J. **23**, 69–72 (1981).
26. C. Roberts, *Unbiased estimation of odds ratio by using negative binomial plans*, Biom. J. **35**, 581–587 (1993).
27. K. J. Lui, *Point estimation on relative risk under inverse sampling*, Biom. J. **38** (6), 669–680 (1996).
28. K. J. Lui, *Confidence intervals for the risk ratio in cohort studies under inverse sampling*, Biom. J. **37**, 965–971 (1995).
29. K. J. Lui, *Estimation of rate ratio and relative difference in matched-pairs under inverse sampling*, Environmetrics **12**, 539–546 (2001).
30. K. J. Lui, *Confidence limits for the population prevalence rate based on the negative binomial distribution*, Statist. Med. **14**, 1471–1477 (1995).
31. K. J. Lui, *Notes on confidence limits for the odds ratio in case-control studies under inverse sampling*, Biom. J. **38**, 221–229 (1996).
32. K. J. Lui, *A Monte-Carlo evaluation of five interval estimators for the relative risk in sparse data*, Biom. J. **48**, 131–143 (2006).
33. W. Wu, Y. Zuo, H. Li, S. Jiang, *Estimation of relative difference under inverse sampling*, J. of Kunming Univ. Sci. and Technology (Sci. and Technology) **34** (5), 108–111 (2009). [In Chinese]
34. H. Cho, *Sequential risk-efficient estimation for the ratio of two binomial proportion*, J. Statist. Plann. and Inf. **137**, 2336–2346 (2007).