Thailand Statistician
October 2022; 20(4): 779-790
http://statassoc.or.th
Contributed paper

On the Normal Approximation for Some Special Estimators of the Ratio of Binomial Proportions

Parichart Pattarapanitchai*[a], Kamon Budsaba [a], Tran Loc Hung [b] and Andrei Volodin [c]
[a] Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Rangsit Campus, Pathum Thani, Thailand.
[b] Faculty of Basic Sciences, University of Finance and Marketing, Ho Chi Minh, Vietnam.
[c] Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, Canada.
*Corresponding author; e-mail: pattarapanitchai@gmail.com

Received: 29 July 2020
Revised: 10 February 2021
Accepted: 27 February 2021

Abstract

We focus on the normal approximation for special cases of point estimators of the ratio of Binomial proportions of two independent populations. We prove that these estimators are normally distributed, something that has not been done before. We investigate its performance in terms of bias, variance, and mean square error, using Monte Carlo simulations. The results show that the normal approximation, which is relatively simple, provides a reliable result. The normal approximation approach could be recommended on the basis of the specific values of the parameters and/or sample sizes.

Keywords: Point estimators, ratio of binomial proportions, inverse binomial sampling, direct binomial sampling, normal asymptotic of an estimator.

1. Introduction

The problem of comparing of a probabilities of success in Bernoulli trials is a topic in biological and medical investigations. In this article, we identify the sample scheme that provides the best accuracy for the point estimation for the ratio of probabilities.

A mathematical statement of the problem is as follows. Let $X_1, X_2, ...$ and $Y_1, Y_2, ...$ be two independent sequences of Bernoulli random variables with success probabilities p_1 and p_2, respectively. The observations are done according to the sequential sampling schemes with stopping times ν_1 and ν_2. Each sample may be obtained in the framework of direct or inverse binomial sampling schemes; see definitions below. From the results of observations $X^{(\nu_1)} = (X_1, ..., X_{\nu_1})$ and

$Y^{(\nu_2)} = (Y_1, ..., Y_{\nu_2}),$ it is necessary to identify the most accurate method of estimation of the ratio $\theta = \dfrac{p_1}{p_2}.$

In the previous investigations (Ngamkham et al. 2016), Chapters 1 and 2 of Ngamkham (2018), and Pattarapanitchai et al. (2020), consideration was most on the confidence estimation for the parameter θ. In this article, we are mainly interested in point estimators.

The aforementioned references give a detailed literature review pertaining to the estimation of θ. Of note, the first easy-to-calculate methods of θ estimation were suggested by Noether (1957) and Guttman (1958). A survey of these early methods can be found in Sheps (1959). Advantages of estimators for θ with the uniformly minimal risk are shown in articles by Bennett (1981) , Roberts (1993), and Lui (1996) .

In this article, we consider estimators for so-called special cases developed previously in Ngamkham et al. (2016) and Chapter 2 (Ngamkham 2018). We provide rigorous proofs that these estimators are approximately normally distributed. This has been stated in Ngamkham et al. (2016) and Chapter 2 (Ngamkham 2018), but no actual proofs were provided.

The simulation results are collected in tables that present the coefficient of skewness, coefficient of kurtosis, true (simulated) and theoretical (from normal approximation) variances, bias, and the mean square error (MSE) (quadratic risk). For each scenario, we generated 10^5 random numbers with the Bernoulli and/or Negative Binomial (Pascal) distribution with various values of parameters (success probabilities) p_1, p_2. We consider the problem of estimating of the probabilities ratio $\theta = \dfrac{p_1}{p_2}$ for the following schemes of Bernoulli trials: special case direct-inverse, and two special cases of the inverse-direct.

2. Point estimator for the ratio of Binomial proportions

The material presented in this section can be found in Ngamkham et al. (2016) and Chapter 2 of Ngamkham (2018). We present it here to fix the notation, and make the article more self-contained. For a solution of the problems stated in the introduction, we consider estimates for the ratio of Binomial proportions $\theta = \dfrac{p_1}{p_2}$. We use the following notation.

Direct binomial sampling: a random vector $X^{(n)} = (X_1, ..., X_n)$ with Bernoulli components and fixed number of observations n is observed. In the case of direct Binomial sampling, we use the statistic $\overline{X}_n = \dfrac{T}{n}$, where $T = \sum_{k=1}^{n} X_k$.

Inverse binomial sampling: a Bernoulli sequence $Y^{(\nu)} = (Y_1, ..., Y_\nu)$ is observed with a stopping time

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

That is, the components of the sequence $Y_1, Y_2, ...$ are observed until the given number m of successes appears. In the case of inverse binomial sampling, we use the statistic $\overline{Y}_m = \dfrac{\nu}{m}$.

In the following we keep the notation $X_1, X_2, \ldots$ for a Bernoulli sequence obtained by the direct sampling scheme and $Y_1, Y_2, \ldots$ for a Bernoulli sequence obtained by the inverse sampling scheme.

The following results are well known; see, for example Chapter 2 of Ngamkham (2018).

Proposition 1 *1. Random variable ν has a Pascal distribution with parameters m and p, denoted $P(m, p)$; that is, its probability mass function is*

$$P(\nu = k) = \binom{k-1}{m-1} p^m (1-p)^{(k-m)}, k = m, m+1, m+2, \ldots$$

The mean and variance are $E(\nu) = \frac{m}{p}$ and $Var(\nu) = \frac{m(1-p)}{p^2}$.

2. Random variable T has a binomial distribution with parameters n and p, denoted $B(n, p)$; that is, its probability mass function is

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

The mean and variance are $E(T) = np$ and $Var(T) = np(1-p)$.

The following result is proved in Theorem 3.2 and Corollary 3.1 of (Giang and Hung 2018). It is crucial for the proof of asymptotic normality of our estimators.

Proposition 2 *Let $U_1, U_2, \ldots$ be a sequence of independent identically distributed random variables with a mean of $\mu > 0$ and variance σ^2. Suppose that $V_1, V_2, \ldots$ is a sequence of independent Bernoulli with parameter p random variables. Additionally, assume that random variables $U_1, U_2, \ldots$ and $V_1, V_2, \ldots$ are independent. Write $S_{N_n} = \sum_{j=1}^{N_n} U_j$, where $N_n = \sum_{i=1}^{n} V_i, n \geq 1$. Then,*

$$\frac{S_{(N_n)} - np\mu}{\sqrt{n(p\sigma^2 + \mu^2 p(1-p))}} \xrightarrow{d} N(0;1) \quad \text{as} \quad n \to \infty.$$

Remark 1 Note that there is a small typographical error in the formulation of Theorem 3.2 of Giang and Hung (2018). It is said that random variables $V_1, V_2, \ldots$ should be positive. Careful analysis of the proof ensures that random variables $Y_1, Y_2, \ldots$ should be nonnegative. This is also used in Corollary 3.1 of Giang and Hung (2018), where the random variables are assumed to have a Bernoulli distribution.

3. Special Case of the Direct-Inverse Sampling Scheme

Fix two natural numbers n and m. The first sample is obtained by the scheme of direct binomial sampling with probability p_1 of success and a fixed sample size n, while the second sample is obtained by the scheme of inverse binomial sampling with the probability p_2 and a stopping time that is defined by the fixed number m of successes in the sample. Therefore, we consider two independent sequences of Bernoulli random variables $X_1,...,X_n$ and $Y_1,...,Y_\nu$, where $\nu = \min\left\{n : \sum_{k=1}^{n} Y_k \ge m\right\}$.

For the special case of the direct-inverse sampling scheme, we suggest the following procedure. The (random) sample size for the second sample depends on the choice of m. We recommend the following sampling plan for the second sample: observe until you reach the same number of successes as in the first experiment; that is, set $m = T = \sum_{k=1}^{n} X_k$. For the estimate of $1/p_2$, we consider the statistics $\bar{Y}_T = \nu_T/T$ (remind that $m = T$), where the conditional distribution of ν_T is the Pascal distribution $P(T, p_2)$ and the unconditional distribution can be obtained by a standard procedure knowing that T has the Binomial distribution. The new suggested estimate of the parameter θ is

$$\hat{\theta}_n = \bar{X}_n \bar{Y}_T = \frac{T}{n}\frac{\nu_T}{T} = \frac{\nu_T}{n}.$$

We note that there was no actual rigorous proof of the asymptotic normality of this estimator that we present in Proposition 3. This rigorous proof uses the following well known result on mean and variance of random sums. The proof of this statement can be found in Theorem 6.6.2 of Gut (2009).

Lemma 1 *Let* $Y_1, Y_2, ...$ *be independent identically distributed random variables, and* U *be a nonnegative integer-valued random variable, independent of* $Y_1, Y_2, ...$ *Set* $S_n = \sum_{i=1}^{n} Y_i$ *for* $n \ge 1$.

(a) If $E(U) < \infty$ *and* $E|Y_1| < \infty$, *then* $E(S_U) = E(U)\cdot E(Y_1)$.

(b) If, in addition $Var(U) < \infty$ *and* $Var(Y_1) < \infty$, *then* $Var(S_U) = E(U)\cdot Var(Y_1) + \left(E(Y_1)\right)^2 \cdot Var(U)$.

Proposition 3 *If* $n \to \infty$, *then the estimate* $\hat{\theta}_n$ *is asymptotically normal with a mean of* θ *and variance* $\theta^2\left(2p_1^{-1} - \theta^{-1} - 1\right)/n$.

Proof: Let $\nu_1, \nu_2, ...$ be independent random variables with the identical distribution $P(1, p_2)$ (first success distribution). We can interpret these random variables as the time between consecutive successes in the second sample. Since samples are independent, we can state that $\nu_1, \nu_2, ...$ are independent of T. By Proposition 1, $E(\nu_i) = \frac{1}{p_2}$ and $Var(\nu_i) = \frac{1-p_2}{p_2^2}$ for $i \ge 1$. Then, the random variable ν_T can be represented as the sum of a random number of random variables:

$$\nu_T = \sum_{i=1}^{T} \nu_i.$$

According to Lemma 1, $E(\nu_T) = E(\nu_1)\cdot E(T) = \frac{np_1}{p_2} = n\theta,$

$$Var(\nu_T) = E(T)\cdot Var(\nu_1) + \left(E(\nu_1)\right)^2 \cdot Var(T) = np_1\frac{(1-p_2)}{p_2^2} + \left(\frac{1}{p_2}\right)^2 np_1(1-p_1)$$
$$= n\theta^2\left(2p_1^{-1} - \theta^{-1} - 1\right).$$

Taking into consideration that $\hat{\theta}_n = \nu_T / n$, we obtain the mean and variance stated in the proposition. To prove asymptotic normality, we apply Proposition 2, with $U_i = \nu_i$ and $V_i = X_i, i \geq 1$. In this case, $p = p_1, \mu = \frac{1}{p_2}$ and $\sigma^2 = \frac{1-p_2}{p_2^2}$. Hence, $p\mu = \theta$ and $p\sigma^2 + \mu^2 p(1-p) = \theta^2 \left(2p_1^{-1} - \theta^{-1} - 1\right)$. We have that

$$\frac{\sqrt{n}\left(\hat{\theta}_n - \theta\right)}{\sqrt{\theta^2 \left(2p_1^{-1} - \theta^{-1} - 1\right)}} \xrightarrow{d} N(0,1) \quad \text{as} \quad n \to \infty.$$

4. Special Cases of the Inverse-Direct Sampling Scheme

Note that for the special case of the direct-inverse sampling scheme, the first sample is obtained by the direct Bernoulli sampling scheme, and the second sample is retrieved with the inverse sampling scheme, where the number of successes equals the number of successes in the first sample. In this section, we concentrate on special cases of the inverse-direct sampling scheme, where the first sample is obtained by the inverse binomial sampling scheme and the second sample is obtained by the direct binomial sampling scheme, where the number of trials n is the same as the number of observations in the first experiment.

The first sample is obtained by the inverse sampling method with parameters (m, p_1). We recommend the following sampling plan for the second sample. Let ν be the (random) sample size for the first sample; that is, the value when we achieve m successes. This value ν from the first sample is used in designing the second sample. For the second sample, the number of trials n is the same as the number of observations in the first experiment; that is, set $n = \nu$. Denote $T_\nu = \sum_{k=1}^{\nu} X_k$. Then the suggested estimate is

$$\hat{\theta} = \frac{(m-1)(\nu+1)}{(\nu-1)(T_\nu+1)} \approx \frac{m}{T_\nu+1}.$$

The random variable ν does not depend on $X_1, X_2, \ldots,$ so it is possible to calculate the mean value and variance of T_ν and its distribution. Since there is a typographical error in the formula for the variance in (Ngamkham 2018) for these calculations, we present the correct derivation. We also note that there was no actual rigorous proof of the asymptotic normality that we present in this article.

Proposition 4 *Statistic* T_ν *is asymptotically* $(m \to \infty)$ *normal with parameters*

$$E(T_\nu) = m\frac{p_2}{p_1} = \frac{m}{\theta} \quad \text{and} \quad Var(T_\nu) = \frac{m}{\theta^2}\left(\theta - 2p_1 + 1\right).$$

Proof: Let $X_1, X_2, \ldots$ be independent identically distributed Bernoulli random variables with the parameter p_2. Then $T_\nu = \sum_{i=1}^{\nu} X_i$ is the sum of random number of random variables. By Proposition 2 and Lemma 1

$$E(T_\nu) = E(\nu) \cdot E(X_1) = \frac{mp_2}{p_1},$$

$$Var(T_\nu) = E(\nu)\cdot Var(X_1) + (E(X_1))^2 \cdot Var(\nu) = \frac{m}{p_1} p_2(1-p_2) + p_2^2 \frac{m(1-p_1)}{p_1^2}$$
$$= m\frac{p_2}{p_1}(p_1 + p_2 - 2p_1p_2).$$

To prove asymptotic normality, we apply Proposition 2, with $n = m, U_i = X_i, i \ge 1$ and $N_m = \nu = \sum_{j=1}^{m} \nu_j$ and $\nu_j, j \ge 1$ are $P(1, p_1)$ independent random variables, similar to the random variables defined in Proposition 3. In this case $p = 1/p_1, \mu = p_2$ and $\sigma^2 = p_2(1-p_2)$. Hence, $mp\mu = \frac{m}{\theta}$ and $p\sigma^2 + \mu^2 p(1-p) = \frac{1}{\theta^2}(\theta - 2p_1 + 1)$. We have that

$$\frac{\sqrt{m}\left(\hat{\theta}_m - \frac{1}{\theta}\right)}{\sqrt{\frac{1}{\theta^2}(\theta - 2p_1 + 1)}} \xrightarrow{d} N(0,1) \quad as \quad m \to \infty.$$

4.1. Two special cases of the inverse-direct sampling scheme

1) First special case of the inverse-direct sampling scheme

Note that $\frac{T_\nu}{m}$ is an unbiased estimator of $\frac{1}{\theta}$. Hence we suggest to estimate the reciprocal

2) Second special case of the inverse-direct sampling scheme

The suggested estimate for the parameter θ is $\hat{\theta} = \frac{(m-1)(\nu+1)}{(\nu-1)(T_\nu+1)} \approx \frac{m}{T_\nu + 1}$.

In Ngamkham et al. (2016), Chapter 2 of Ngamkham (2018), and Pattarapanitchai et al. (2020), the delta method is applied to prove its asymptotically normality as $m \to \infty$ and to find the asymptotic mean and variance of $\hat{\theta}$. The main problem in the derivations in these articles is that the authors did not rigorously prove that the statistics T_ν is asymptotically normal, as we did in this article. In Pattarapanitchai et al. (2020), the following Taylor expansion of the statistic is obtained,

$$\hat{\theta} \approx \frac{\theta m}{\theta + m} - \frac{1}{\left(\frac{1}{\theta} + \frac{1}{m}\right)^2}\left(\frac{T_\nu}{m} - \frac{1}{\theta}\right) = \frac{\theta m}{\theta + m} - \left(\frac{\theta m}{\theta + m}\right)^2 \left(\frac{T_\nu}{m} - \frac{1}{\theta}\right).$$

Hence, $E(\hat{\theta}) = \frac{\theta m}{\theta + m} \approx \theta$ for $m \to \infty$, $Var(\hat{\theta}) \approx \left(\frac{\theta m}{\theta + m}\right)^4 \frac{1}{m}\left(\frac{p_2}{p_1} + \frac{p_2^2}{p_1^2}(1 - 2p_1)\right)$.

Applying asymptotic normality of the statistic T_ν provided in Proposition 4, by the classical Delta method, we can state that statistic $\hat{\theta}$ is approximately normal with a mean of θ and variance $\frac{1}{m}\left(\frac{\theta m}{\theta + m}\right)^4 \left(\frac{p_2}{p_1} + \frac{p_2^2}{p_1^2}(1 - 2p_1)\right)$.

5. Simulation Study

We use simulation studies to evaluate the properties of the normal approximation. We estimate the bias, MSE, true (simulated) variance, and asymptotic (theoretical, from normal approximation) variance through Monte Carlo simulation with the R statistical software. For the parameter configurations, we generated 10,000 random samples from two independent Bernoulli populations with parameters p_1 and p_2. Numerical results on the values of the accuracy measurements, bias, MSE, true variance, and asymptotic variance for the three estimators of θ of the ratio of two Binomial proportions for different number of trials $n = 50, 100, 200, 1000$ in the Special Scheme of Direct-Inverse sampling scheme and different values of number of successes $m = 50, 100, 200, 1000$ and different combinations of success probabilities $p_1 = 0.05, 0.1(0.1)0.9$ and $p_2 = 0.1(0.2)0.9$ are reported in the tables below.

Table 1 Special direct-inverse case: true variance (T) and asymptotic variance (A), $n = 50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	T	0.177	0.356	0.677	0.960	1.203	1.392	1.557	1.680	1.756	1.802
	A	0.185	0.360	0.680	0.960	1.200	1.400	1.560	1.680	1.760	1.800
0.3	T	0.017	0.035	0.067	0.093	0.115	0.133	0.147	0.155	0.160	0.160
	A	0.018	0.036	0.067	0.093	0.116	0.133	0.147	0.156	0.160	0.160
0.5	T	0.005	0.011	0.021	0.029	0.035	0.040	0.043	0.045	0.045	0.043
	A	0.006	0.011	0.021	0.029	0.035	0.040	0.043	0.045	0.045	0.043
0.7	T	0.002	0.005	0.009	0.012	0.015	0.016	0.017	0.017	0.016	0.015
	A	0.003	0.005	0.009	0.012	0.015	0.016	0.017	0.017	0.016	0.015
0.9	T	0.001	0.002	0.004	0.006	0.007	0.007	0.007	0.007	0.006	0.004
	A	0.001	0.002	0.004	0.006	0.007	0.007	0.007	0.007	0.006	0.004

Table 2 Special direct-inverse case: bias (B) and mean square error (M), $n = 50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	B	0.042	0.006	0.001	0.000	0.001	0.000	–0.001	–0.001	–0.003	–0.001
	M	0.179	0.356	0.677	0.960	1.203	1.392	1.557	1.680	1.756	1.802
0.3	B	0.014	0.002	0.000	0.000	–0.001	0.000	–0.001	–0.001	–0.001	–0.001
	M	0.018	0.035	0.067	0.093	0.115	0.133	0.147	0.155	0.160	0.160
0.5	B	0.008	0.001	0.000	0.000	0.000	0.000	–0.001	–0.001	–0.001	–0.001
	M	0.005	0.011	0.021	0.029	0.035	0.040	0.043	0.045	0.045	0.043
0.7	B	0.006	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	M	0.002	0.005	0.009	0.012	0.015	0.016	0.017	0.017	0.016	0.015
0.9	B	0.005	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	M	0.001	0.002	0.004	0.006	0.007	0.007	0.007	0.007	0.006	0.004

Table 3 Special direct-inverse case: kurtosis (K) and skewness (S), $n = 50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	K	5.242	3.977	3.498	3.359	3.256	3.196	3.163	3.112	3.113	3.129
	S	1.298	0.894	0.613	0.496	0.424	0.368	0.328	0.307	0.292	0.290
0.3	K	5.140	3.888	3.429	3.313	3.187	3.146	3.132	3.092	3.113	3.127
	S	1.266	0.844	0.581	0.467	0.391	0.347	0.312	0.287	0.278	0.279
0.5	K	4.844	3.762	3.379	3.257	3.161	3.117	3.099	3.090	3.098	3.129
	S	1.194	0.779	0.533	0.422	0.343	0.295	0.268	0.265	0.250	0.265
0.7	K	4.492	3.602	3.255	3.136	3.138	3.078	3.100	3.059	3.070	3.083
	S	1.099	0.686	0.443	0.329	0.294	0.237	0.213	0.192	0.194	0.233
0.9	K	3.951	3.276	3.115	3.015	3.021	3.025	2.982	3.018	3.055	3.110
	S	0.943	0.542	0.313	0.216	0.164	0.112	0.068	0.047	0.039	0.098

Table 4 First special inverse-direct case: true variance (T) and asymptotic variance (A), $m = 50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	T	0.113	0.036	0.013	0.008	0.005	0.004	0.003	0.003	0.002	0.002
	A	0.112	0.036	0.013	0.008	0.005	0.004	0.003	0.003	0.002	0.002
0.3	T	0.770	0.205	0.057	0.028	0.017	0.012	0.009	0.007	0.006	0.005
	A	0.768	0.204	0.057	0.028	0.017	0.012	0.009	0.007	0.006	0.005
0.5	T	2.010	0.502	0.126	0.056	0.031	0.020	0.014	0.010	0.008	0.006
	A	2.000	0.5	0.125	0.056	0.031	0.020	0.014	0.010	0.008	0.006
0.7	T	3.829	0.928	0.218	0.091	0.048	0.028	0.018	0.012	0.008	0.006
	A	3.808	0.924	0.217	0.090	0.047	0.028	0.018	0.012	0.008	0.006
0.9	T	6.226	1.481	0.335	0.133	0.066	0.036	0.021	0.013	0.007	0.004
	A	6.192	1.476	0.333	0.132	0.065	0.036	0.021	0.012	0.007	0.004

Table 5 First special inverse-direct case: bias (B) and mean square error (M), $m = 50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	B	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	M	0.113	0.036	0.013	0.008	0.005	0.004	0.003	0.003	0.002	0.002
0.3	B	0.001	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	M	0.770	0.205	0.057	0.028	0.017	0.012	0.009	0.007	0.006	0.005
0.5	B	0.003	0.004	0.001	0.001	0.001	0.000	0.000	0.000	0.000	0.000
	M	2.010	0.503	0.126	0.056	0.031	0.020	0.014	0.010	0.008	0.006
0.7	B	0.008	0.007	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000
	M	3.829	0.928	0.218	0.091	0.048	0.028	0.018	0.012	0.008	0.006
0.9	B	0.009	0.008	0.002	0.001	0.001	0.000	0.000	0.000	0.000	0.000
	M	6.226	1.481	0.335	0.133	0.066	0.036	0.021	0.013	0.007	0.004

Table 6 First special inverse-direct case: kurtosis (K) and skewness (S), $m = 50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	K	3.148	3.132	3.116	3.105	3.067	3.082	3.102	3.076	3.096	3.127
	S	0.295	0.291	0.285	0.305	0.305	0.327	0.333	0.346	0.360	0.367
0.3	K	3.129	3.128	3.108	3.145	3.107	3.081	3.041	3.003	3.013	3.005
	S	0.286	0.279	0.256	0.253	0.234	0.215	0.202	0.183	0.165	0.140
0.5	K	3.131	3.096	3.113	3.132	3.100	3.051	3.065	3.038	3.035	2.996
	S	0.281	0.265	0.276	0.262	0.240	0.213	0.178	0.151	0.112	0.053
0.7	K	3.134	3.107	3.115	3.122	3.142	3.130	3.059	3.059	3.017	3.034
	S	0.284	0.280	0.277	0.283	0.263	0.246	0.208	0.165	0.098	0.021
0.9	K	3.132	3.098	3.123	3.122	3.126	3.124	3.119	3.108	3.117	3.137
	S	0.285	0.276	0.282	0.292	0.297	0.286	0.280	0.254	0.208	0.101

Table 7 Second special inverse-direct case: true variance (T) and asymptotic variance (A), $m = 50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	T	0.008	0.040	0.239	0.731	1.687	3.231	5.813	9.213	13.553	18.877
	A	0.007	0.033	0.178	0.485	0.988	1.708	2.654	3.830	5.231	6.852
0.3	T	0.001	0.003	0.012	0.030	0.058	0.098	0.154	0.224	0.313	0.424
	A	0.001	0.002	0.011	0.026	0.049	0.081	0.123	0.175	0.239	0.314
0.5	T	0.000	0.001	0.003	0.008	0.013	0.021	0.029	0.040	0.051	0.065
	A	0.000	0.001	0.003	0.007	0.012	0.018	0.026	0.035	0.045	0.056
0.7	T	0.000	0.000	0.002	0.003	0.005	0.007	0.010	0.012	0.014	0.016
	A	0.000	0.000	0.001	0.003	0.005	0.007	0.009	0.011	0.013	0.015
0.9	T	0.000	0.000	0.001	0.002	0.003	0.003	0.004	0.004	0.004	0.004
	A	0.000	0.000	0.001	0.002	0.002	0.003	0.004	0.004	0.004	0.004

Table 8 Second special inverse-direct case: bias (B) and mean square error (M), $m = 50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	B	0.009	0.017	0.025	0.028	0.019	0.003	–0.016	–0.053	–0.103	–0.162
	M	0.008	0.040	0.239	0.732	1.687	3.231	5.813	9.216	13.564	18.903
0.3	B	0.003	0.006	0.008	0.009	0.005	0.000	–0.006	–0.018	–0.031	–0.047
	M	0.001	0.003	0.012	0.030	0.058	0.098	0.154	0.225	0.314	0.426
0.5	B	0.002	0.003	0.005	0.005	0.004	0.001	–0.005	–0.012	–0.019	–0.028
	M	0.000	0.001	0.003	0.008	0.013	0.021	0.029	0.040	0.052	0.066
0.7	B	0.001	0.002	0.004	0.004	0.003	0.000	–0.004	–0.008	–0.014	–0.020
	M	0.000	0.000	0.002	0.003	0.005	0.007	0.010	0.012	0.014	0.016
0.9	B	0.001	0.002	0.003	0.003	0.002	0.000	–0.003	–0.006	–0.011	–0.016
	M	0.000	0.000	0.001	0.002	0.003	0.003	0.004	0.004	0.004	0.004

6. Conclusions

We presented only some of our simulations, but have run many more. Our conclusion is based on all the results, presented or not. According to our simulations, we can conclude the following.

6.1. Special case of the direct-inverse sampling scheme

The bias of estimation increases when values of p_1 and p_2 close to 0. As expected, the bias decreases when the sample size increases for all value of the parameters p_1 and p_2. In cases where the value of p_1 is small, the bias of our approach is large (see Table 5) and decreases when the value

value of the parameter p_2 becomes larger, especially for moderate $(n=100)$ to large sample sizes $(n \geq 200)$.

Table 9 Second special inverse-direct case: kurtosis (K) and skewness (S), $m=50$

p_2	p_1	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.1	K	4.239	4.829	6.951	10.833	44.999	27.473	41.002	38.084	33.487	28.347
	S	0.783	0.933	1.326	1.760	2.921	2.832	3.709	3.877	3.846	3.688
0.3	K	3.755	3.940	4.069	4.629	4.676	5.253	5.712	6.092	7.041	8.046
	S	0.623	0.680	0.728	0.844	0.897	1.002	1.111	1.188	1.300	1.421
0.5	K	3.703	3.700	3.750	3.804	3.883	4.005	3.980	4.116	4.548	4.677
	S	0.595	0.595	0.612	0.624	0.652	0.672	0.693	0.739	0.843	0.883
0.7	K	3.646	3.627	3.555	3.552	3.480	3.441	3.461	3.537	3.680	3.780
	S	0.574	0.562	0.537	0.521	0.494	0.489	0.483	0.518	0.549	0.602
0.9	K	3.624	3.576	3.445	3.380	3.349	3.211	3.173	3.143	3.173	3.302
	S	0.566	0.545	0.491	0.448	0.402	0.347	0.302	0.259	0.243	0.297

We can say the same about the MSE and both the true and asymptotic variances; they are smaller when the value of p_1 is smaller and the value of p_2 is larger. Moreover, they are very close to zero when p_1 is close to 0 and p_2 close to 1, especially for moderate to large sample sizes.

We use the coefficients of skewness and kurtosis as descriptive statistics to show that the normal approximation is appropriated for our estimators. According to our results, the values of the skewness and kurtosis of these estimators are very close to zero and 3, respectively, when the sample size n is large. This says that the normal approximation is working well for large sample sizes, which is not surprising.

The values of skewness are all negative for the values $p_1=0.7, p_2=0.9$ for large sample sizes. This means that the data are slightly skewed to the left or negatively skewed. The skewness of the estimator is positive when p_1 and p_2 are smaller.

The estimator has a positive kurtosis with values smaller than 3 for values of p_1 and p_2 in the region $(p_1=0.7, p_2=0.9)$, which means that the estimator has a fat-tailed distribution (leptokurtic). The estimator has a negative kurtosis when p_1 and p_2 are smaller, which indicates a thin-tailed data distribution. The best result is obtained when the sample size increases so the values of skewness and kurtosis of the estimator decreases and becomes very close to zero and 3 respectively.

6.2. First special case of the inverse-direct sampling scheme

The bias of estimation is very small for all values of p_1 and p_2 and for all sample sizes. The bias increases slightly with the values of p_1 close to 0 and p_2 close to 1. As expected, the bias decreases close to 0 when the sample size increases for all value of the parameters p_1 and p_2.

We can say the same about the MSE, both true and asymptotic variances; they are smaller when the value of p_1 is larger and the value of p_2 is smaller. Moreover, they are very close to zero when p_1 is close to 1 and p_2 is close to 0, especially for moderate $(m=100)$ to large sample sizes $(m \geq 200)$.

According to our results, the values of the skewness and kurtosis of these estimators are very close to zero and 3, respectively, when the sample size is large. This means that the normal approximation is working well for large sample sizes, which is not surprising.

Overall, the skewness of the estimator is slightly positive for all values of p_1 and p_2 and for all sample sizes. This means that the data are slightly skewed to the right, or positively skewed.

The estimator has a positive kurtosis with values smaller than 3 for values of p_1 and p_2 in the region $(p_1 > 0.4, 0.1 < p_2 < 0.6)$ for large sample size, which means that the estimator is leptokurtic. The estimator has a negative kurtosis when p_1 is smalle and p_2 is extreme, which indicates a thin-tailed data distribution. The best result is obtained when the sample size increases so the values of skewness and kurtosis of the estimator decrease and become very close to zero and 3, respectively.

6.3. Second special case of the inverse-direct sampling scheme

The bias of estimation increases when the values of p_1 are close to 1 and p_2 are close to 0. As expected, the bias decreases when the sample size increases for all values of the parameters p_1 and p_2. In cases when the value of p_1 is small, the bias of our approach is small (see Table 8), and decreases when the value of parameter p_2 becomes larger, especially for moderate ($m = 100$) to large sample sizes ($m \geq 200$).

We can say the same about the MSE and both true and asymptotic variances. They are smaller when the value of p_1 is smaller and the value of p_2 is larger. Moreover, they are very close to zero when p_1 is close to 0 and p_2 is close to 1, especially for moderate to large sample sizes. When p_1 is close to 1 and p_2 is close to 0, there is a large difference between the true variance and asymptotic variance for small sample sizes.

According to our results, the value of the skewness and kurtosis of these estimators are very close to zero and 3, respectively, when the sample size n is large. This means that the normal approximation is working well for large sample sizes, which is not surprising.

Overall, the skewness of the estimator is positive for all values of p_1 and p_2 and for all sample sizes. This means that the data are skewed to the right or positively skewed. In cases where the value of p_1 is small, the skewness of the estimator is small, and decreases when the value of the parameter p_2 becomes larger especially for moderate to large sample sizes.

The estimator has a negative kurtosis with values larger than 3 for all values of p_1 and p_2, which indicates a thin-tailed distribution of the estimator. The best result is obtained when the sample size increases so the values of skewness and kurtosis of the estimator decrease and become very close to zero and 3, respectively.

6.4. A remark on the real data analysis

It is common for statistical papers to provide an example of real data analysis. Unfortunately, we cannot provide it in this article. These two Special Cases estimators were introduced in Ngamkham et al. (2016) and have not been used for a practical applications yet. According to the results presented in Pattarapanitchai et al. (2020), these estimators of the ratio of proportions are significantly more efficient in comparison with usual estimators for all four combinations of sampling schemes. The authors would like to attract the attention of practitioners to there interesting estimators.

Acknowledgements

The authors are grateful to the reviewers for carefully reading the manuscript and for offering substantial comments and suggestions which enabled them to improve the presentation. The last author research was carried out in the framework of the program of support of the Mathematical Center of the Volga Region Federal District (Project N 075-02-2020-1478).

References

Bennett BM. On the use of the negative binomial in epidemiology. Biom J. 1981; 23: 69-72.

Giang LT, Hung TL. An extension of random summations of independent and identically distributed random variables. Commun Korean Math Soc. 2018; 33(2): 605-618.

Gut A. An Intermediate course in probability. New York: Springer; 2009.

Guttman I. A note on a series solution of a problem in estimation. Biometrika. 1958; 45(3/4): 565-567.

Lui KJ. Point estimation on relative risk under inverse sampling. Biom J. 1996; 38(6): 669-680.

Ngamkham T, Volodin A, Volodin I. Confidence intervals for a ratio of Binomial proportions based on direct and inverse sampling schemes. Lobachevskii J Math. 2016; 37(4): 466-494.

Ngamkham T, Confidence interval estimation for the ratio of Binomial proportions and random numbers generation for some Statistical models. PhD [Thesis] Regina, Saskatchewan: University of Regina; 2018.

Noether GE. Two confidence intervals for the ratio of two probabilities and some measures of effectiveness. J Am Stat Assoc. 1957; 52(277): 36-45.

Pattarapanitchai P, Ngamkham T, Budsaba K, Volodin A. Comparison of accuracy properties of point estimators for the ratio of Binomial proportions with the inverse-direct sampling scheme. Lobachevskii J Math. 2020; 41(4): 686-702.

Roberts CD. Unbiased Estimation of odds ratio by using negative binomial plans. Biom J. 1993; 35: 581-587.

Sheps MC. An examination of some methods of comparing several rates of proportions. Biometrics. 1959; 15(1): 87-89.