Communications for Statistical Applications and Methods
2013, Vol. 20, No. 5, 347–356

DOI: http://dx.doi.org/10.5351/CSAM.2013.20.5.347

Approximate Confidence Limits for the Ratio of Two Binomial Variates with Unequal Sample Sizes

Hokwon Cho[1,a]

[a]Department of Mathematical Sciences, University of Nevada, Las Vegas

Abstract

We propose a sequential method to construct approximate confidence limits for the ratio of two independent sequences of binomial variates with unequal sample sizes. Due to the nonexistence of an unbiased estimator for the ratio, we develop the procedure based on a modified maximum likelihood estimator (MLE). We generalize the results of Cho and Govindarajulu (2008) by defining the sample-ratio when sample sizes are not equal. In addition, we investigate the large-sample properties of the proposed estimator and its finite sample behavior through numerical studies, and we make comparisons from the sample information view points.

Keywords: Approximate confidence limits, ratio of two binomial proportions, modified MLE, sample-ratio, large-sample properties.

1. Introduction

The ratio of two binomial proportions and constructing its confidence interval represents an important tool to measure risk ratio (Katz *et al.*, 1978; Bailey, 1987) or relative risk (Gart, 1985; Gart and Nam, 1988) in comparative prospective studies and in biomedical experiments. The ratio or odds ratio of two binomial proportions is also related to vaccine efficacy and attributable risk (Walter, 1976), which arises frequently in epidemiological problems (*e.g.* cohort study involving two groups).

Among sequential methods for constructing an interval for an unknown parameter based on the fixed-sample size, Ray (1957) and Starr (1966) studied the fixed-width confidence interval for the mean of a normal distribution. Khan (1969) explored a general method to determine stopping rules to obtain a fixed-width confidence interval for an unknown parameter involving some possible unknown nuisance parameters. In addition, Siegmund (1982) investigated a sequential confidence interval for the odds ratio.

The article is organized as follows. In Section 2, we begin with the notations and describe the characteristics of the problem and the proposed method. In Section 3, we study the desirable properties of the proposed estimator in terms of asymptotics. In Section 4, we examine the properties of the proposed procedure. Then, finally in Section 5, we illustrate the procedure with the Monte Carlo examples to make brief comparisons and summarize our conclusions with further remarks.

2. Formulation

Suppose that $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are two independent sequences of Bernoulli random variables with probabilities $0 < p_0 < 1$ and $0 < p_1 < 1$, respectively. Define a ratio, $\theta = p_1/p_0$. With samples of size n on X and n_1 on Y, let $R = \sum_{i=1}^{n} X_i$ and $S = \sum_{j=1}^{n_1} Y_j$, assuming that $n_1 = \kappa \cdot n$ such that $\kappa \cdot n$ is

[1] Associate Professor, Department of Mathematical Sciences, University of Nevada, Las Vegas, 4505 Maryland Parkway Box 454020, Las Vegas, NV 89154-4020, USA. E-mail: cho@unlv.nevada.edu

an integer and κ is known. We call the constant $\kappa(= n_1/n)$ the sample-ratio. We want to find optimal values of n (and n_1) and construct an interval having specified width $2d$ and confidence coefficient γ

$$P\left\{\left|\hat{\theta} - \theta\right| \leq d\right\} \geq \gamma. \tag{2.1}$$

Then, R and S are two independent binomial random variables with parameters (n, p_0) and (n_1, p_1), respectively. Since there does not exist an unbiased estimator of the ratio θ, we consider a modified estimator

$$\hat{\theta}_n = \frac{1}{\kappa}\left(\frac{S + 1/2}{R + 1/2}\right). \tag{2.2}$$

When we observe $R = r$ and $S = s$, the likelihood (of θ and p_0) is

$$L(\theta, p_0) = \binom{n}{r}\binom{n_1}{s} p_0^{r+s} (1 - p_0)^{n-r} \theta^s (1 - p_0\theta)^{n_1 - s} \tag{2.3}$$

in which $p_1 = p_0\theta$. From the log-likelihood function of θ and p_0,$l(\theta, p_0)$, we have the following maximum likelihood estimates:

$$\hat{p}_{0,n} = \frac{r}{n}$$

and since $\hat{\theta}_{mle} = (1/\kappa)(S/R) = nS/n_1R$, we have

$$\hat{\theta}_{mle} = \frac{s}{n_1\hat{p}_0}.$$

Furthermore, the Fisher information about θ is given by

$$E\left[-\frac{\partial^2 l(\theta, p_0)}{\partial\theta^2}\right] = \frac{n_1 p_0}{\theta(1 - \theta p_0)}, \tag{2.4}$$

and the information about p_0 is given by

$$E\left[-\frac{\partial^2 l(\theta, p_0)}{\partial p_0^2}\right] = \frac{n}{p_0(1 - p_0)} + \frac{n_1\theta}{p_0(1 - p_0)}. \tag{2.5}$$

Similarly, the joint information about θ and p_0 is

$$E\left[-\frac{\partial^2 l(\theta, p_0)}{\partial\theta\partial p_0}\right] = \frac{n_1(1 - p_1)}{1 - p_1} + \frac{n_1(1 - p_1)p_1}{(1 - p_1)^2} = \frac{n_1}{1 - \theta p_0}. \tag{2.6}$$

It follows from Equations (2.4)–(2.6), the information matrix about (θ, p_0) denoted by $\mathbf{I}(\theta, p_0)$, is then

$$\mathbf{I}(\theta, p_0) = \begin{bmatrix} \dfrac{n_1 p_0}{\theta(1 - \theta p_0)} & \dfrac{n_1}{1 - \theta p_0} \\ \dfrac{n_1}{1 - \theta p_0} & \dfrac{n}{p_0(1 - p_0)} + \dfrac{n_1\theta}{p_0(1 - p_0)} \end{bmatrix}, \tag{2.7}$$

and the determinant of the information matrix, $\det \mathbf{I}$ becomes

$$\det \mathbf{I} = \frac{n_1 n}{\theta (1-p_0)(1-\theta p_0)}.$$

Hence, from the inverse of $\mathbf{I}$, namely, $\mathbf{I}^{-1}$ we have the asymptotic variance of $\hat{\theta}_{mle}$

$$\begin{aligned} var\left(\hat{\theta}_{mle}\right) &= \frac{\theta (1-p_0)(1-\theta p_0)}{n_1 n}\left[\frac{n}{p_0(1-p_0)} + \frac{n_1\theta}{p_0(1-p_0)}\right] \\ &= \frac{\theta(1-p_0\theta)}{n_1 p_0} + \frac{\theta^2(1-p_0)}{np_0}. \end{aligned} \tag{2.8}$$

For the special case $n_1 = n$, Equation (2.8) reduces to

$$var\left(\hat{\theta}_{mle}\right) = \frac{\theta(1-p_0\theta) + \theta^2(1-p_0)}{np_0} = \frac{\theta(1+\theta-2\theta p_0)}{np_0},$$

which coincides with the result obtained in Cho and Govindarajulu (2008, Equation (1.10)).

3. Properties of the Estimator $\hat{\theta}_n$

In this section, we investigate the desirable properties of the modified estimator $\hat{\theta}_n$ for the proposed procedure. Even though, there is no unbiased estimator of the true ratio θ, the modified estimator is asymptotically unbiased and so $\hat{\theta}_{mle}$ is. Therefore, we must show the asymptotic equivalence of the estimators, $\hat{\theta}_n$ and $\hat{\theta}_{mle}$ through their variances.

3.1. Asymptotic unbiasedness of $\hat{\theta}_n$

Consider the expectation of $\hat{\theta}_n$. That is,

$$E\left(\hat{\theta}_n\right) = \frac{1}{\kappa}E\left(\frac{S+1/2}{R+1/2}\right) = \frac{1}{\kappa}E\left(S+\frac{1}{2}\right)E\left(\frac{1}{R+1/2}\right). \tag{3.1}$$

In order to expand, we can rewrite

$$\begin{aligned} E\left(\frac{1}{R+1/2}\right) &= E\left(\frac{1}{np_0 + R - np_0 + 1/2}\right) \\ &= \frac{1}{np_0}E\left[\left(1+\frac{R-np_0+1/2}{np_0}\right)^{-1}\right]. \end{aligned}$$

Then, after algebraic simplification

$$\begin{aligned} E\left(\frac{1}{R+1/2}\right) &= \frac{1}{np_0}E\left[1-\left(\frac{R-np_0+1/2}{np_0}\right)+\left(\frac{R-np_0+1/2}{np_0}\right)^2+\cdots\right] \\ &= \frac{1}{np_0}\left[1-\frac{1}{2np_0}+\frac{np_0(1-p_0)}{(np_0)^2}+\frac{1}{4(np_0)^2}+\cdots\right]. \end{aligned} \tag{3.2}$$

Combining Equations (3.1) and (3.2), we have

$$\begin{aligned} E\left(\hat{\theta}_n\right) &= \frac{1}{\kappa}\left(n_1 p_1 + \frac{1}{2}\right)\left\{\frac{1}{np_0}\left[1 - \frac{1}{2np_0} + \frac{np_0(1-p_0)}{(np_0)^2} + \frac{1}{4(np_0)^2} + \cdots\right]\right\} \\ &= \frac{1}{\kappa}\left\{\frac{n_1 p_1}{np_0}\left[1 - \frac{1}{2np_0} + \frac{np_0(1-p_0)}{(np_0)^2} + \frac{1}{4(np_0)^2} + \cdots\right]\right. \\ &\quad \left. + \frac{1}{2np_0}\left[1 - \frac{1}{2np_0} + \frac{np_0(1-p_0)}{(np_0)^2} + \frac{1}{4(np_0)^2} + \cdots\right]\right\} \\ &= \frac{1}{\kappa}\left\{\frac{n_1 p_1}{np_0}\left[1 - O\left(n^{-2}\right)\right] + \frac{1}{2np_0}\left[1 - O\left(n^{-2}\right)\right]\right\}. \end{aligned}$$

Therefore, for sufficiently large n

$$E\left(\hat{\theta}_n\right) \simeq \frac{1}{\kappa}\frac{(\kappa n)p_1}{np_0} = \frac{p_1}{p_0} = \theta. \tag{3.3}$$

Thus, $\hat{\theta}_n$ is an asymptotically unbiased estimator of θ.

Next, we investigate the variance of the modified estimator $\hat{\theta}_n$.

3.2. Asymptotic variance of $\hat{\theta}_n$

Now, we obtain the asymptotic variance of $\hat{\theta}_n = \kappa^{-1}(S + 1/2)/(R + 1/2)$ assuming that $\kappa n = n_1$ is an integer.

Theorem 1.

$$\lim_{n\to\infty} nVar\left(\hat{\theta}_n\right) = \frac{\theta(1 - p_0\theta)}{\kappa p_0} + \frac{\theta^2(1-p_0)}{p_0} = \frac{\theta(1-p_0\theta) + \kappa\theta^2(1-p_0)}{\kappa p_0}.$$

Proof: Noting that

$$\mathrm{var}\left(\hat{\theta}_n\right) = \frac{1}{\kappa^2}\mathrm{var}\left(\frac{S + 1/2}{R + 1/2}\right),$$

refer directly to Theorem 1.1 in Cho and Govindarajulu (2008). □

3.3. Asymptotic normality of $\hat{\theta}_n$

From the modified ratio given in Equation (2.2), we have

$$\begin{aligned} \sqrt{\kappa n}\left(\frac{S/n_1 + 1/2n_1}{R/n + 1/2n} - \theta\right) &= \sqrt{\kappa n}\left(\frac{\hat{p}_{1,n} + 1/2n_1}{\hat{p}_{0,n} + 1/2n} - \theta\right) \\ &= \sqrt{\kappa n}\left\{\frac{\hat{p}_{1,n} - p_1 + p_1 + 1/2n_1 - \theta(\hat{p}_{0,n} + 1/2n)}{\hat{p}_{0,n} + 1/2n}\right\} \\ &= \sqrt{\kappa n}\left\{\frac{\hat{p}_{1,n} - p_1 + 1/2n_1 - \theta(\hat{p}_{0,n} - p_0 + 1/2n)}{\hat{p}_{0,n} + 1/2n}\right\} \\ &= \sqrt{\kappa n}\left(\frac{\hat{p}_{1,n} - p_1}{p_0}\right) - \theta\sqrt{\kappa n}\left(\frac{\hat{p}_{0,n} - p_0}{p_0}\right) + o_p(1), \end{aligned}$$

where $\hat{p}_{0,n} = r/n$ and $\hat{p}_{1,n} = s/n$. For sufficiently large n and from Slutsky's theorem, the above modification transforms to

$$\sqrt{\kappa n}\left(\hat{\theta}_n - \theta\right) \overset{d}{\simeq} N\left\{0, \frac{\theta(1-p_1)}{p_0} + \frac{\kappa\theta^2(1-p_0)}{p_0}\right\} \equiv N\left(0, \sigma^2\right), \tag{3.3}$$

where

$$\sigma^2 = \frac{\theta\{1 - \theta p_0 + \theta\kappa(1-p_0)\}}{p_0} = \frac{\theta\{1 + \theta\kappa - \theta p_0(1+\kappa)\}}{p_0}.$$

Now we consider determining n such that

$$P\left\{\left|\hat{\theta} - \theta\right| \leq d\right\} = P\left\{\frac{\sqrt{\kappa n}\left|\hat{\theta} - \theta\right|}{\sigma} \leq \frac{d\sqrt{\kappa n}}{\sigma}\right\} \geq \gamma.$$

Thus,

$$2\Phi\left(\frac{d\sqrt{\kappa n}}{\sigma}\right) - 1 \geq \gamma,$$

or

$$\frac{d\sqrt{\kappa n}}{\sigma} \geq z_{\frac{(1+\gamma)}{2}} = z \text{ (say)},$$

for specified $d\,(> 0)$ where $\Phi\left(z_{(1+\gamma)/2}\right) = (1+\gamma)/2$.

Hence,

$$n \geq \kappa^{-1}\left(\frac{z\sigma}{d}\right)^2.$$

Moreover, the optimal fixed-sample size for the procedure becomes the smallest integer n^* such that $n \leq n^* \leq n + 1$, for estimating θ with specified d and z. That is,

$$n^* = \left\lfloor \kappa^{-1}\left(\frac{z\sigma}{d}\right)^2 \right\rfloor + 1, \tag{3.4}$$

where $\lfloor\cdot\rfloor$ indicates the greatest integer function.

However, since both θ and p_0 are unknown, we resort to the following adaptive sequential rule: We stop sampling after N observations on X, and κN observations on Y where

$$N = \inf_n \left\{n \geq m : n \geq \frac{\kappa^{-1}z^2\hat{\sigma}_n^2}{d^2}\right\}, \tag{3.5}$$

where $m\,(\geq 2)$ is the initial sample size, $\hat{\sigma}_n^2 = \hat{\theta}_n\left\{1 + \kappa\hat{\theta}_n - \hat{\theta}_n\hat{p}_{0,n}(1+\kappa)\right\}$ and $\hat{p}_{0,n} = (R + 1/2)/n$.

Upon stopping we give the $\gamma \times 100\%$ confidence interval estimate of length $2d$ for θ as

$$\left(\hat{\theta}_N - d,\ \hat{\theta}_N + d\right).$$

4. Asymptotic Properties of the Procedure

In this section we investigate the asymptotic behavior of the proposed sequential procedure and various properties of the (random) stopping time N.

4.1. Finite sure termination

Toward its finite sure termination, we have the following theorem:

Theorem 2. *Let N be the stopping time associated with the sequential procedure. Then $P\{N < \infty\}$ = 1.*

Proof: Using the stopping rule in Equation (3.5)

$$P\{N = \infty\} = \lim_{n\to\infty} P\{N > n\}$$
$$\leq \lim_{n\to\infty} P\left\{n \leq \kappa^{-1}\left(\frac{z^2\hat{\sigma}_n^2}{d^2}\right)\right\} = 0$$

since $\hat{\sigma}_n^2$ converges in probability to σ^2 as $n \to \infty$. Therefore, the proposed sequential procedure terminates finitely with probability one. □

4.2. First order asymptotics

We apply the criteria given in Chow and Robbins (1965) to establish the asymptotic efficiency and consistency of the procedure as d tends to zero.

The stopping rule given by Equation (3.5) can be written as

$$N = \inf_n \left\{n \geq m : \frac{n}{\kappa^{-1}\theta\,(z/d)^2} \geq \frac{\hat{\sigma}_n^2}{\theta}\right\}, \tag{4.1}$$

where $\hat{\sigma}_n^2 = \hat{\theta}_n\{1 + \kappa\hat{\theta}_n - \hat{\theta}_n\hat{p}_{0,n}(1+\kappa)\}/\hat{p}_{0,n}$.

From this, Equation (4.1) takes the form.

$$N = N(t) = \min_n \left\{n \geq m : Y_n \leq \frac{g(n)}{t}\right\},$$

where

$$Y_n = \frac{\hat{\theta}_n}{\theta}\left\{\frac{1 + \kappa\hat{\theta}_n - \hat{\theta}_n\hat{p}_{0,n}(1+\kappa)}{1 + \kappa\theta - \theta p_0(1+\kappa)}\right\}\left(\frac{p_0}{\hat{p}_{0,n}}\right),$$
$$g(n) = n,$$

and

$$t = \kappa^{-1}\theta\left(\frac{z}{d}\right)^2 \frac{\{1 + \kappa\theta - \theta p_0(1+\kappa)\}}{p_0}.$$

Thus, $\{Y_n\}$ is a sequence of random variables such that $Y_n > 0$ almost surely (a.s.), $\lim_{n\to\infty} Y_n = 1$ a.s. because $\hat{p}_{0,n}$ converges a.s. to p_0 and $\hat{\theta}_n/\theta$ converges a.s. to 1 as $n \to \infty$. Additionally, we see that $g(n) \to \infty$ and $g(n)/g(n-1) \to 1$ as $n \to \infty$. Since the stopping rule N is well-defined and non-decreasing as a function of t, we can apply the results of Chow and Robbins (1965) and obtain the first order asymptotics for the proposed sequential procedure.

Theorem 3.

(i) $\lim_{d\to 0} N = \infty$ *a.s.,*

(ii) $\lim_{d\to 0} N/n^* = 1$ *a.s.,*

(iii) $\lim_{d\to 0} P\left\{\left|\hat{\theta}_N - \theta\right| \le d\right\} = \gamma.$

Proof: For (i) and (ii) proceed as in Cho and Govindarajulu (2008). For the proof of (iii), since N/n^* converges in probability to one, $\sqrt{n^*}(\hat{p}_{0,N} - p_0)$ is asymptotically normal with mean zero and variance $p_0(1 - p_0)$. Furthermore since N/n^* converges in probability to one, $\sqrt{\kappa n^*}(\hat{p}_{1,\kappa N} - p_1))$ is asymptotically normal with mean zero and variance $p_1(1 - p_1)$. Then, from Anscombe's theorem (1952), it follows that $\hat{p}_{0,N}$ converges in probability to p_0. Using Slutsky's theorem, we infer that

$$\sqrt{n^*}\left(\hat{\theta}_N - \theta\right) \overset{d}{\simeq} \sqrt{n^*}\left\{\frac{(\hat{p}_{1,N} - p_1)}{p_0} - \frac{\theta\,(\hat{p}_{0,N} - p_0)}{p_0}\right\}. \tag{4.2}$$

Applying the Anscombe's condition specialized for sums of independent and identically distributed (i.i.d.) random variables on the right-hand side in (4.2), it follows that $\sqrt{n^*}\left(\hat{\theta}_N - \theta\right)$ is asymptotically $N\left(0, \sigma^2\right)$ where $\sigma^2 = \theta\left\{1 + \theta\kappa - \theta p_0(1 + \kappa)\right\}/p_0$. Therefore, we have

$$P\left\{\left|\hat{\theta}_N - \theta\right| \le d\right\} = P\left\{\frac{\sqrt{n^*}\left|\hat{\theta}_N - \theta\right|}{\sigma} \le \frac{d\sqrt{n^*}}{\sigma}\right\} = \gamma$$

as $d \to 0$, and hence, Theorem 3 is proved. □

Next, we assert the asymptotic efficiency of the proposed sequential procedure by proceeding as in Cho and Govindarajulu (2008).

5. Numerical Studies

5.1. Simulation setup

A Monte Carlo experiment is used to investigate the behavior and performance of the stopping rule in the proposed sequential procedure. The results of the experimentation are summarized in the following tables, which show the values of the parameter θ, namely $\theta = 1.0, 1.5, 2.0$ and 4.0 with selected values of p_0, p_1, and the sample-ratio κ, $0 < \kappa \le 1$. For instance, $\kappa = 1$ means that both sample sizes n and n_1 are taken equally. If $\kappa = 0.8$, the sample of X_i has taken 25% more than Y_j's, and if $\kappa = 0.5$, the sample size of X_i is two times more than Y_j's and so on. Without loss of generality (WLOG) we can assume that $p_0 \le p_1$ and hence consider only situations in which $\theta \ge 1$ because the roles of X and Y can be interchanged when $\theta \le 1$. Further since $p_0 \le p_1$, we expect to sample more from the rare population, WLOG, we can assume that $\kappa \le 1$ for simulation purposes.

In the table, every value in each row is based on 5,000 independent replications with initial sample size $m = 10$ for each experiment. Using the sample ratio $\kappa = n/n_1 = 0.8$ or 0.5, we present the coverage probability (CP) of the interval $\hat{\theta} \pm d$, and the expected stopping time and optimal sample

Table 1: θ = 1.0 with p_0 = 0.5 and p_1 = 0.5

γ	d	κ = 0.8				κ = 0.5				y_j	
		$\hat{\theta}$	CP	$E(N)$	n^*	$\hat{\theta}$	CP	$E(N)$	n^*	$E(N_1)$	n_1^*
.90	.2	1.002	.890	163.08	169	.998	.921	262.96	271	132.25	136
	.3	.999	.852	68.27	76	.996	.885	111.76	121	56.66	61
	.4	.998	.838	36.12	43	.993	.849	58.84	68	30.35	34
.95	.3	1.000	.941	234.04	241	1.001	.968	378.94	385	190.22	193
	.4	1.002	.922	100.37	107	.999	.946	162.83	171	82.17	86
	.5	.996	.879	52.69	61	.998	.915	87.14	96	44.38	49

Table 2: θ = 1.5 with p_0 = 0.4 and p_1 = 0.6

γ	d	κ = 0.8				κ = 0.5				y_j	
		$\hat{\theta}$	CP	$E(N)$	n^*	$\hat{\theta}$	CP	$E(N)$	n^*	$E(N_1)$	n_1^*
.90	.3	1.499	.882	173.54	184	1.502	.939	283.96	294	142.74	147
	.4	1.503	.832	92.43	104	1.500	.901	154.10	166	77.83	83
	.5	1.496	.804	55.24	66	1.499	.875	94.06	106	47.91	53
.95	.3	1.498	.942	250.44	260	1.503	.973	408.86	417	205.19	208
	.4	1.504	.914	136.84	147	1.504	.960	225.63	235	113.58	117
	.5	1.504	.874	83.28	94	1.503	.936	139.17	150	70.37	75

Table 3: θ = 2.0 with p_0 = 0.3 and p_1 = 0.6

γ	d	κ = 0.8				κ = 0.5				y_j	
		$\hat{\theta}$	CP	$E(N)$	n^*	$\hat{\theta}$	CP	$E(N)$	n^*	$E(N_1)$	n_1^*
.90	.4	1.999	.875	238.63	254	1.998	.937	391.29	406	196.41	203
	.5	1.994	.825	144.17	162	2.000	.915	243.56	260	122.56	131
	.6	2.000	.787	95.88	113	1.998	.879	162.17	181	81.93	90
.95	.4	2.001	.937	347.48	361	2.004	.975	566.44	578	283.98	289
	.5	2.000	.907	214.58	231	2.001	.958	354.56	370	178.05	185
	.6	1.994	.874	141.71	160	1.997	.942	239.21	256	120.38	128

Table 4: θ = 4.0 with p_0 = 0.2 and p_1 = 0.8

γ	d	κ = 0.8				κ = 0.5				y_j	
		$\hat{\theta}$	CP	$E(N)$	n^*	$\hat{\theta}$	CP	$E(N)$	n^*	$E(N_1)$	n_1^*
.90	.6	3.995	.899	613.60	638	3.998	.961	1000.16	1022	500.84	511
	.7	4.010	.874	446.07	471	3.996	.951	726.09	751	363.82	376
	.8	3.987	.856	326.70	359	4.010	.940	553.74	577	277.65	288
.95	.6	3.995	.953	884.72	906	4.000	.988	1433.33	1452	717.42	726
	.7	4.008	.946	650.37	668	4.002	.984	1047.57	1067	524.54	533
	.8	3.999	.922	483.61	511	3.999	.979	793.87	816	397.69	408

size denoted by $E(N)$ and n^* for X_i's, and $E(N_1)$ and n_1^* for Y_j's, respectively. The nominal level of confidence γ for the interval is .90 or .95 for each value of θ.

From Table 1 to Table 4, we observe that the expected stopping time $E(N)$ monotonically increases (to infinity) as the sample ratio κ becomes smaller (*i.e.*, sample of x_i's getting more) or d decreases (to zero). We observe that as d decreases the coverage probability (CP) is getting close (eventually) to the nominal probability γ, which is referred to as asymptotic consistency. It should be noted that if one takes X_i's (or Y_j's) more, then the CP is comparatively higher than the one in equal sample sizes. Therefore, the above numerical evidence indicates that the finite-sample behavior lends support to the asymptotic behavior of the proposed sequential procedure when $d \to 0$.

Increasing the starting sample size m results in the increase of both $E(N)$ and CP. Accordingly, when the CP is below the nominal level, choosing a moderate size of m is a trade-off for obtaining a

Table 5: $\theta = 2.0$ with $p_0 = 0.3$ and $p_1 = 0.6$

Sample-Ratio	$d = 0.4$	$\gamma = 0.90$	$X_i \sim \text{Ber}(p_0)$		$Y_j \sim \text{Ber}(p_1)$	
$\kappa = n_1^*/n^*$	$\hat{\theta} \pm d$	CP	$E(N)$	n^*	$E(N_1)$	n_1^*
(1) $\kappa = 1$	(1.599, 2.399)	0.841	196.41	203	196.41	203
(2) $\kappa = 0.8$	(1.599, 2.399)	0.875	238.63	254	196.41	203
(3) $\kappa = 0.5$	(1.598, 2.398)	0.937	391.29	406	196.41	203

Table 6: $\theta = 4.0$ with $p_0 = 0.2$ and $p_1 = 0.8$

Sample-Ratio	$d = 0.7$	$\gamma = 0.95$	$X_i \sim \text{Ber}(p_1)$		$Y_j \sim \text{Ber}(p_0)$	
$\kappa = n_1^*/n^*$	$\hat{\theta} \pm d$	CP	$E(N)$	n^*	$E(N_1)$	n_1^*
(1) $\kappa = 1$	(3.305, 4.705)	0.902	524.54	533	524.54	533
(2) $\kappa = 0.8$	(3.308, 4.708)	0.946	650.37	668	524.54	533
(3) $\kappa = 0.5$	(3.302, 4.702)	0.984	1047.57	1067	524.54	533

higher coverage probability. For practical purposes, the size of d can be determined from the standard error (S.E.) of the estimate $\hat{\theta}$.

5.2. Comparison: Equal-sample sizes versus Unequal-sample sizes

In this subsection we compare the results from the unequal-sample sizes with values of the sample-ratio $\kappa = 0.8$ and $\kappa = 0.5$ with the results from the equal-sample sizes ($\kappa = 1$) on x and y. For brevity, we summarize and present part of the results in the following two tables, Tables 5–6 for $\theta = 2.0$ with 90% nominal level and $\theta = 4.0$ with 95% nominal level, respectively. Each table shows three values of κ, the coverage probability (CP), and the expected stopping times $E(N)$ for each values of κ.

From the above tables, as the sample-ratio κ decreases (*i.e.*, take more samples with smaller p) we observe that the coverage probability (CP) for the interval of width $2d$ improves. For instance, when sample-ratio $\kappa = 0.8$, comparing to the equal sample sizes ($\kappa = 1$), the CPs have increased by 3.4% points for $\gamma = 0.90$ and 4.4% points for $\gamma = 0.95$, respectively. Therefore, for a more stable estimation of the ratio for two binomial variates it is reasonable to take more samples from the population having smaller probability p even though the equal-sample sizes minimize the expected stopping times..

The expected stopping time $E(N)$ is uniformly bigger than $E(N_1)$ when $\kappa < 1$. At the same time, it seems to be generally true that $\kappa < 1$ gives the higher CP than the one with equal-sample sizes; however, we need to note that eventually the CP approaches the nominal level γ as d gets smaller. From these, we surmise that reducing variability (in the denominator of the statistic used) by taking (reasonably) more samples (*i.e.*, getting 'more' information from the rare population) seems to be fair and is a better idea for a more stable estimation of the true ratio θ.

5.3. Concluding remarks

We have proposed a sequential method to obtain the approximate confidence limits for the ratio of two binomial variates that may have unequal sample sizes. The proposed method offers a relatively new perspective on the information aspects. The procedure is developed based on a modified MLE in terms of the sample-ratio κ; subsequently, the large-sample properties of the proposed estimator are investigated. The finite sample behavior was verified through numerical studies. In addition, by comparing the expected stopping times and the coverage probabilities of the intervals, it is recommended to take more samples from the rare population to have more precise intervals for the ratio parameter $\theta = p_1/p_0$.

We also note that it might be preferred to have a confidence interval based on the likelihood-ratio since it would be invariant. Future studies should include dealing with advantages and disadvantages in practical usages between two approaches even though they are not completely comparable.

Acknowledgements

The author wishes to thank the referee for the insightful review of this manuscript which resulted in a more accurate and readable paper.

References

Anscombe, F. J. (1952). Large sample theory of sequential estimation, *Proceedings of the Cambridge Philosophical Society*, **48**, 600–607.

Bailey, B. J. R. (1987). Confidence limits to the risk ratio, *Biometrics*, **43**, 201–205.

Cho, H. (2007). Sequential risk-efficient estimation for the ratio of two binomial proportions, *Journal of Statistical Planning and Inference*, **137**, 2336–2346.

Cho, H. and Govindarajulu, Z. (2008). Sequential confidence limits for the ratio of two binomial proportions, *International Journal of Statistics and Management System*, **3**, 26–42.

Chow, Y. S. and Robbins, H. (1965). On the asymptotic theory of fixed-width sequential confidence intervals for the mean, *The Annals of Mathematical Statistics*, **36**, 457–462.

Gart, J. (1985). Approximate tests and interval estimation of the common relative risk in the combination of 2×2 tables, *Biometrika*, **72**, 673–677.

Gart, J. and Nam, J. (1988) Approximate interval estimation of the ratio of binomial parameters: A review and corrections for skewness, *Biometrics*, **44**, 323–338.

Katz, D., Baptista, J., Azen, S. and Pike, M. (1978). Obtaining confidence intervals for the risk ratio in cohort studies, *Biometrics*, **34**, 469–474.

Khan, R. (1969). A general method of determining fixed-width confidence intervals, *The Annals of Mathematical Statistics*, **40**, 704–709.

Ray, D. (1957). Sequential confidence intervals for the mean of a normal distribution with unknown variance, *Journal of the Royal Statistical Society, Series B*, **19**, 133–143.

Siegmund, D. (1982). A sequential confidence interval for the odds ratio, *Probability and Mathematical Statistics*, **2**, 149–156.

Starr, N. (1966). The performance of a sequential procedure for the fixed-width interval estimation of the mean, *The Annals of Mathematical Statistics*, **37**, 36–50.

Walter, S. (1976). The estimation and interpretation of attributable risk in health research, *Biometrics*, **32**, 829–849.

Received December 15, 2012; Revised September 1, 2013; Accepted September 3, 2013