Communications in Statistics - Simulation and Computation

ISSN: 0361-0918 (Print) 1532-4141 (Online) Journal homepage: http://www.tandfonline.com/loi/lssp20

Modified Normal-based Approximation to the Percentiles of Linear Combination of Independent Random Variables with Applications

K. Krishnamoorthy

To cite this article: K. Krishnamoorthy (2016) Modified Normal-based Approximation to the Percentiles of Linear Combination of Independent Random Variables with Applications, Communications in Statistics - Simulation and Computation, 45:7, 2428-2444, DOI: 10.1080/03610918.2014.904342

To link to this article: http://dx.doi.org/10.1080/03610918.2014.904342

Accepted author version posted online: 23 Jun 2014.
Published online: 23 Jun 2014.

Submit your article to this journal

Article views: 23

View related articles

View Crossmark data

Citing articles: 1 View citing articles

Full Terms & Conditions of access and use can be found at
http://www.tandfonline.com/action/journalInformation?journalCode=lssp20

Download by: [68.226.177.160] **Date:** 15 June 2016, At: 05:24

Communications in Statistics—Simulation and Computation®, 45: 2428–2444, 2016
Copyright © Taylor & Francis Group, LLC
ISSN: 0361-0918 print / 1532-4141 online
DOI: 10.1080/03610918.2014.904342

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Modified Normal-based Approximation to the Percentiles of Linear Combination of Independent Random Variables with Applications

K. KRISHNAMOORTHY

Department of Mathematics, University of Louisiana at Lafayette, Lafayette, Louisiana, USA

A modified normal-based approximation for calculating the percentiles of a linear combination of independent random variables is proposed. This approximation is applicable in situations where expectations and percentiles of the individual random variables can be readily obtained. The merits of the approximation are evaluated for the chi-square and beta distributions using Monte Carlo simulation. An approximation to the percentiles of the ratio of two independent random variables is also given. Solutions based on the approximations are given for some classical problems such as interval estimation of the normal coefficient of variation, survival probability, the difference between or the ratio of two binomial proportions, and for some other problems. Furthermore, approximation to the percentiles of a doubly noncentral F *distribution is also given. For all the problems considered, the approximation provides simple satisfactory solutions. Two examples are given to show applications of the approximation.*

Keywords Coverage probability; Doubly noncentral F; Fiducial approach; Modified large sample method; MOVER; Ratio of odds; Relative risk.

Mathematics Subject Classification 60E05; 62E17

1. Introduction

Approximations are commonly used in many applications where exact solutions are numerically involved, not tractable, or for simplicity and convenience. The problem of approximating linear combinations of independent random variables has been well addressed in the literature, especially for a linear combination of independent chi-square random variables (Satterthwaite, 1941, 1946; Welch, 1938, 1947) and of independent noncentral chi-square random variables (Imhof, 1961). Satterthwaite approximation is commonly used to estimate linear combinations of variance components in one-way random model and other mixed models. Furthermore, the famous Welch's "approximate degrees of freedom" solution to the Behrens-Fisher problem is based on the Satterthwaite approximation. These approximations are essentially obtained by matching moments, and they are valid only when the coefficients in the linear combination are positive.

Received June 15, 2013; Accepted March 3, 2014
Address correspondence to K. Krishnamoorthy, Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA 70508, USA; E-mail: krishna@louisiana.edu

In this article, we propose a simple method for approximating the percentiles of linear combination of independent random variables where the coefficients could be negative. The idea for our approximation stems from Graybill and Wang (1980) and Zou and Donner (2008), who proposed an approximate method of obtaining confidence intervals (CIs) for a linear combination of parameters based on independent CIs of the individual parameters. Graybill and Wang derived the result only for a linear combination variance components whereas Zou and Donner (2008) have justified that Graybill-Wang's result is valid for any parameters. Zou and Donner refer to their approach as the method of variance estimate recovery (MOVER). In a series of articles (Zou, 2013; Zou et al., 2008, 2009a,b), it has been shown that the approximate solutions based on the MOVER to many problems are very satisfactory. In view of these results, our interest here is to find approximation to percentiles of the distribution of a linear combination of independent random variables (not necessarily from the same family) based on the percentiles of the individual random variables.

To understand the similarity between the problem of estimating a linear combination of parameters, and the problem of finding the percentiles of a linear combination of independent random variables, we shall outline the solution for the former problem. Let $\widehat{\theta}_i$ be an unbiased estimate of θ_i, $i = 1, \ldots, k$. Assume that $\widehat{\theta}_i$'s are independent. Let (l_i, u_i) be a $1-\alpha$ confidence interval (CI) for θ_i, $i = 1, \ldots, k$. The approximate MOVER CI (L, U) for $\sum_{i=1}^{k} w_i\theta_i$ is given by

$$L = \sum_{i=1}^{g} w_i\widehat{\theta}_i - \sqrt{\sum_{i=1}^{g} w_i^2\left(\widehat{\theta}_i - l_i^*\right)^2}, \quad \text{with } l_i^* = \begin{cases} l_i & \text{if} \quad w_i > 0, \\ u_i & \text{if} \quad w_i < 0, \end{cases} \tag{1}$$

and

$$U = \sum_{i=1}^{g} w_i\widehat{\theta}_i + \sqrt{\sum_{i=1}^{g} w_i^2\left(\widehat{\theta}_i - u_i^*\right)^2}, \quad \text{with } u_i^* = \begin{cases} u_i \text{ if} & w_i > 0, \\ l_i & \text{if} \quad w_i < 0. \end{cases} \tag{2}$$

The approximation for the percentiles of $\sum_{i=1}^{k} w_i X_i$, where X_i's are independent continuous random variables, that we propose in the sequel is essentially the same as (1) and (2), except that the point estimates are replaced by the expected values of unobservable random variables X_i's, l_i's and u_i's are replaced by the corresponding percentiles of X_i's, $i = 1, \ldots, g$.

The rest of the article is organized as follows. In the following section, we describe the approximation for the percentiles of a linear combination of independent random variables. From the proposed approximation, we deduce an approximation to the percentiles of the ratio X/Y, where X and Y are independent random variables with Y being positive. In Section 3, we appraise the approximation by comparing the percentiles based on the approximation with the Monte Carlo (MC) estimates for the cases of beta random variables and chi-square random variables. In Section 4, we provide some applications of the proposed approximation to develop CIs for the difference between two binomial proportions, relative risk, and for the ratio of odds. We further show applications for estimating the common mean of several normal populations, coefficient of variation of a normal distribution, and for a survival probability. In Section 5, we point out applications to some other problems including approximation to the percentiles of a doubly noncentral F distribution. In Section 6, the methods are applied to find CIs for some practical problems involving

Downloaded by [68.226.177.160] at 05:24 15 June 2016

binomial distribution and normal distribution, and compared with the available ones. Some concluding remarks are given in Section 7.

2. Main Result

Let $X_1, \ldots, X_k$ be independent continuous random variables such that X_i's are approximately normally distributed under some conditions. Consider the linear combination $Q = \sum_{i=1}^{k} w_i X_i$, where w_i's are known constants. Let Q_α denote the α quantile of Q. Under the assumption that the X_i's are approximately normal, we have

$$Q_\alpha \simeq \sum_{i=1}^{k} w_i E(X_i) - z_{1-\alpha} \left[\sum_{i=1}^{k} w_i^2 \mathrm{Var}(X_i) \right]^{\frac{1}{2}}, \quad 0 < \alpha < .5, \tag{3}$$

where z_p denotes the p quantile of the standard normal distribution.

Following the MOVER, we modify the approximate quantile in (3) by replacing the $\mathrm{Var}(X_i)$ by an expression that depends on the percentile and the mean of X_i. Under the assumption on X_i, an approximate α quantile of X_i is given by

$$X_{i;\alpha} \simeq E(X_i) - z_{1-\alpha}\sqrt{\mathrm{Var}(X_i)}, \text{ for } 0 < \alpha < .5, \tag{4}$$

which implies that

$$\mathrm{Var}(X_i) \simeq \frac{(E(X_i) - X_{i;\alpha})^2}{z_{1-\alpha}^2}. \tag{5}$$

Substituting the above variance expression in (3), we obtain the approximation

$$Q_\alpha \simeq \sum_{i=1}^{k} w_i E(X_i) - \left[\sum_{i=1}^{k} w_i^2 \left(E(X_i) - X_{i;\alpha} \right)^2 \right]^{\frac{1}{2}}. \tag{6}$$

If distributions of X_i's are symmetric about their means, then $|X_{i;\alpha} - E(X_i)| = |X_{i;1-\alpha} - E(X_i)|$, and the above approximation may be satisfactory regardless of the signs of w_i's. If some of w_i's are negative and the distributions of X_i's are not necessarily symmetric, then a better approximation can be obtained as follows. Let $w_i^* = w_i$ if $w_i > 0$ and $-w_i$ if $w_i < 0$, and let $X_i^* = X_i$ if $w_i > 0$ and $-X_i$ if $w_i < 0$. As $Q = \sum_{i=1}^{k} w_i X_i = \sum_{i=1}^{k} w_i^* X_i^*$, it follows from (6) that

$$Q_\alpha \simeq \sum_{i=1}^{k} w_i^* E(X_i^*) - \left[\sum_{i=1}^{k} w_i^2 \left(E(X_i^*) - X_{i;\alpha}^* \right)^2 \right]^{\frac{1}{2}}. \tag{7}$$

Note that $\sum_{i=1}^{k} w_i^* E(X_i^*) = \sum_{i=1}^{k} w_i E(X_i)$, and if $w_i < 0$ for some i, then $X_{i;\alpha}^* = -X_{i;1-\alpha}$ and $E(X_i^*) = -E(X_i)$, and so the ith term within the square brackets of (7) is equal to $w_i^2 (X_{i;1-\alpha} - E(X_i))^2$. Using this relation, we obtain the approximation for Q_α as

$$Q_\alpha \simeq \sum_{i=1}^{k} w_i E(X_i) - \left[\sum_{i=1}^{k} w_i^2 \left(E(X_i) - X_i^l \right)^2 \right]^{\frac{1}{2}}, \text{ for } 0 < \alpha \le .5, \tag{8}$$

where $X_i^l = X_{i;\alpha}$ if $w_i > 0$, and is $X_{i;1-\alpha}$ if $w_i < 0$.

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Similarly, an approximation to an upper percentile of Q can be obtained as

$$Q_{1-\alpha} \simeq \sum_{i=1}^{k} w_i E(X_i) + \left[\sum_{i=1}^{k} w_i^2 \left[E(X_i) - X_i^u\right]^2\right]^{\frac{1}{2}}, \text{ for } 0 < \alpha < .5, \tag{9}$$

where $X_i^u = X_{i;1-\alpha}$ if $w_i > 0$, and is $X_{i;\alpha}$ if $w_i < 0$.

The standard approximation that does not require the percentiles of the individual random variables in Q is given by

$$Q_\alpha^* \simeq \sum_{i=1}^{k} w_i E(X_i) + z_\alpha \left[\sum_{i=1}^{k} w_i^2 \mathrm{Var}(X_i)\right]^{\frac{1}{2}}, \text{ for } 0 < \alpha < 1. \tag{10}$$

We shall refer to the approximations in (8) and (9) as the modified normal-based approximations (MNAs), and the standard one in (10) as the normal-based approximation (NA).

Remark 2.1. If $X_1, \ldots, X_k$ are independent normal random variables, then it can be readily verified that the approximations (8), (9), and (10) are exact. We will see in the sequel that MNAs are better than the usual approximation (10) when the distributions of X_i's are asymmetric.

2.1. *Ratio of Two Independent Random Variables*

The percentiles of the ratio of two independent random variables can be deduced from the modified normal-based approximations in (8) and (9). Let X and Y be independent random variables with mean μ_x and μ_y, respectively. Assume that Y is a positive random variable. For $0 < \alpha < 1$, let c denote the α quantile of $R = X/Y$ so that $P(X - cY \leq 0) = \alpha$. This means that c is the value for which the α quantile of $X - cY$ is zero. The approximate α quantile of $X - cY$ based on (8) is

$$\mu_x - c\mu_y - \sqrt{(\mu_x - X_\alpha)^2 + c^2(\mu_y - Y_{1-\alpha})^2},$$

where $\mu_x = E(X)$, $\mu_y = E(Y)$, and X_α and Y_α denote the α quantiles of X and Y, respectively. Equating the above expression to zero, and solving the resulting equation for c, we get an approximate α quantile for X/Y as

$$R_\alpha \simeq \begin{cases} \dfrac{\mu_x\mu_y - \left\{(\mu_x\mu_y)^2 - \left[\mu_y^2 - (\mu_y - Y_{1-\alpha})^2\right]\left[\mu_x^2 - (\mu_x - X_\alpha)^2\right]\right\}^{\frac{1}{2}}}{\left[\mu_y^2 - (\mu_y - Y_{1-\alpha})^2\right]}, & 0 < \alpha \leq .5, \\ \dfrac{\mu_x\mu_y + \left\{(\mu_x\mu_y)^2 - \left[\mu_y^2 - (\mu_y - Y_{1-\alpha})^2\right]\left[\mu_x^2 - (\mu_x - X_\alpha)^2\right]\right\}^{\frac{1}{2}}}{\left[\mu_y^2 - (\mu_y - Y_{1-\alpha})^2\right]}, & .5 < \alpha < 1. \end{cases} \tag{11}$$

It should be noted that the above expression is a generalization of the one by Li et al. (2010) for averages and that by Andreas and Hernandez (2014) for proportions.

3. Evaluation Studies

To appraise the accuracy of the MNAs, we shall compare the approximate percentiles with Monte Carlo (MC) estimates based on 100,000 runs. For simulation studies, we consider

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Table 1
Comparison of 100α percentiles of $\sum_{i=1}^{k} w_i \chi^2_{m_i}$

		$k=2$				$k=4$		
α		MC	MNA	NA		MC	MNA	NA
.01	(m_1, m_2)	−4.56	−4.30	−4.24	$(m_1, \ldots, m_4)$	0.43	0.65	−1.37
.05	(2,4)	−3.00	−2.87	−3.11	(2,2,2,3)	0.72	0.82	−0.18
.10	(w_1, w_2)	−2.31	−2.25	−2.51	$(w_1, \ldots, w_4)$	0.94	0.98	0.46
.90	(.6, −.4)	1.54	1.56	1.71	(.1,.1,.1,.7)	5.01	5.02	4.94
.95		2.37	2.33	2.31		6.08	6.14	5.58
.99		4.35	4.17	3.44		8.63	8.67	6.77
.01	(6,10)	−2.21	−2.08	−3.97	(3,4,12,5)	−2.28	−2.31	−2.36
.05	(.8,−.2)	−1.07	−1.07	−1.99	(.2,−.3,.3,−.2)	−0.93	−0.93	−1.09
.10		−0.45	−0.46	−0.93		−0.28	−0.27	−0.40
.90		6.64	6.66	6.53		4.40	4.35	4.40
.95		8.22	8.21	7.59		5.21	5.13	5.09
.99		11.54	11.58	9.57		6.80	6.80	6.36
.01	(1,20)	5.28	5.34	3.47	(8,2,1,2)	−0.37	−0.31	−0.92
.05	(.4,.6)	6.83	6.90	6.09	(.3,−.2,−.2,.3)	0.38	0.46	0.05
.10		7.80	7.85	7.48		0.75	0.82	0.57
.90		17.52	17.49	17.32		4.27	4.23	4.23
.95		19.32	19.34	18.71		4.98	4.99	4.75
.99		23.13	23.18	21.33		6.46	6.65	5.72

Note: MC, Monte Carlo simulation; MNA, based on (8) and (9); NA, based on (10).

Downloaded by [68.226.177.160] at 05:24 15 June 2016

the chi-square distribution and the beta distribution. Note that the χ^2 random variable is approximately normally distributed for large degrees of freedom (df). The beta random variable with the shape parameters a and b, denoted by $B_{a,b}$, is also approximately normally distributed for moderate and approximately equal values of a and b. The MNAs in (8) and (9) for the percentiles of these random variables are expected to be satisfactory under the aforementioned conditions. So we shall evaluate the accuracy of the approximations for cases where these distributions are skewed.

For evaluation purpose, we can choose, without loss of generality, $|w_i| < 1$ and $\sum_{i=1}^{k} |w_i| = 1$. The approximate percentiles of $Q_C = \sum_{i=1}^{k} w_i \chi^2_{m_i}$, where $\chi^2_{m_i}$ denotes the chi-square random variables with df m_i, $i = 1, \ldots, k$, and the corresponding MC estimates based on 100,000 runs are presented in Table 1 for $k = 2$ and 4. The approximations are based on the standard normal approximation (NA) given in (10), and the MNAs (8) and (9). We see in Table 1 that the MNA percentiles are comparable with the corresponding Monte Carlo estimates in most cases whereas the ones based on the usual normal approximations are not satisfactory. In particular, when the dfs are very much different, the usual normal approximations are not accurate; see the case $(m_1, m_2) = (1, 20)$. The MNAs are quite satisfactory when the dfs m_i's are not too small; see the values for $(m_1, m_2) = (6, 10)$ and $(m_1, \ldots, m_4) = (3, 4, 12, 5)$ in Table 1.

Results for $Q_B = \sum_{i=1}^{k} w_i B_{a_i,b_i}$, where B_{a_i,b_i} denotes the beta random variable with the shape parameters a_i and b_i, are given in Table 2 for $k = 2$ and 5. The results based on the MC method and the approximate method are in agreement in most cases, especially, for upper

Table 2
Comparison of 100α percentiles of $\sum_{i=1}^{k} w_i B_{a_i,b_i}$

		$k=2$				$k=5$		
α		MC	MNA	NA		MC	MNA	NA
.01	$(a_1, b_1; a_2, b_2)$	−.207	−.179	−.250	$(a_1, b_1; ..; (a_5, b_5)$	0.169	0.197	0.071
.05	(.6,3;3,12)	−.155	−.141	−.182	(2,1;2,3;5,.4;3,3;1,2)	0.212	0.215	0.171
.10	(w_1, w_2)	−.128	−.122	−.145	$(w_1, \ldots, w_5)$	0.237	0.234	0.224
.90	(.5, −.5)	0.122	0.126	0.112	(.1,.1,.1,.1,.6)	0.627	0.629	0.600
.95		0.180	0.180	0.148		0.683	0.685	0.654
.99		0.275	0.273	0.217		0.766	0.761	0.754
.01	(4,5;5,5)	−.177	−.168	−.194	(1,1;2,2;3,3;10,2;.4,5)	−.108	−.094	−.201
.05	(.6,−.4)	−.114	−.114	−.117	(.6,.1,−.1,−.1,−.1	−.065	−.066	−.081
.10		−.078	−.080	−.077		−.032	−.035	−.017
.90		0.214	0.216	0.210		0.450	0.453	0.435
.95		0.253	0.255	0.251		0.484	0.484	0.499
.99		0.323	0.318	0.327		0.528	0.511	0.619
.01	(3,4;7,8)	−.120	−.116	−.156	(2,2;3,3;4,6;5,4;10,11)	−.170	−.166	−.213
.05	(.67,−.33)	−.063	−.064	−.071	(.6,−.1,−.1,−.1,−.1)	−.117	−.118	−.120
.10		−.026	−.029	−.026		−.078	−.080	−.070
.90		0.301	0.301	0.292		0.292	0.294	0.283
.95		0.344	0.345	0.338		0.331	0.331	0.333
.99		0.419	0.415	0.422		0.386	0.379	0.427

Note: MC, Monte Carlo simulation; MNA, based on (8) and (9); NA, based on (10).

percentiles. The usual normal approximations based on (10) are unsatisfactory if some of the beta distributions are very asymmetric; see the results for $(a_1, b_1; a_2, b_2) = (.6, 3; 3, 12)$ and the first part of Table 2 under $k = 5$.

Overall, we see that the MNA percentiles are more accurate than those based on the usual approximation (10). We also compared the quantiles based on approximations with the MC estimates for $.3 \le \alpha \le .7$ for the two linear combinations considered in the preceding paragraphs, and observed that the approximations are very satisfactory for mid percentiles. These results are not reported here, because in most applications one needs the tail percentiles.

4. Applications

The tail percentiles of a linear combination of independent random variables are required in many applications. We shall point out only a few well-known problems where the MNAs can be readily applied to find satisfactory solutions.

4.1. *Binomial and Poisson Distributions*

Let $X_1, \ldots, X_g$ be independent random variables with $X_i \sim \text{binomial}(n_i, p_i)$, $i = 1, \ldots, g$. Recently, Krishnamoorthy and Lee (2010) proposed a fiducial approach using which a CI for any real valued function $f(p_1, \ldots, p_g)$ can be easily obtained. To describe

Downloaded by [68.226.177.160] at 05:24 15 June 2016

their approach, let $(k_1, \ldots, k_g)$ be an observed value of $(X_1, \ldots, X_g)$. A fiducial quantity for p_i is given by $V_i = B_{k_i+.5, n_i-k_i+.5}$. A fiducial quantity for a function is obtained by substitution, and $f(V_1, \ldots, V_g)$ is a fiducial quantity for $f(p_1, \ldots, p_g)$. The lower and upper α quantiles of $f(V_1, \ldots, V_g)$ form a $1-2\alpha$ CI for $f(p_1, \ldots, p_g)$, and these quantiles can be estimated by simulation. Closed-form approximate CIs for functions involving binomial parameters such as the relative risk p_1/p_2, the ratio of odds $[p_1/(1-p_1)/[p_2/(1-p_2)]$, and a contrast $\sum_{i=1}^{g} l_i p_i$ (Andrés et al., 2011) can be readily obtained using the approximations in the preceding sections. For instance, the fiducial quantity for the ratio of odds is given by

$$\frac{B_{k_1+.5,n_1-k_1+.5}\big(1 - B_{k_2+.5,n_2-k_2+.5}\big)}{B_{k_2+.5,n_2-k_2+.5}\big(1 - B_{k_1+.5,n_1-k_1+.5}\big)} \sim \frac{(2k_1+1)(2n_2-2k_2+1)}{(2k_2+1)(2n_1-2k_1+1)} \frac{F_{2k_1+1,2n_1-2k_1+1}}{F_{2k_2+1,2n_2-2k_2+1}}, \quad (12)$$

where $F_{m,n}$ denotes the F random variable with dfs m and n. The above result in (12) is obtained using the results that $B_{m,n} \sim \chi^2_{2m}/(\chi^2_{2m} + \chi^2_{2n})$ and $(n\chi^2_m/(m\chi^2_n) \sim F_{m,n}$, where the chi-square random variables are independent. The percentiles of (12) form a CI for the ratio of odds. Approximate percentiles of (12) can be obtained using (11). Exact coverage studies by Zhang (2013) indicated that the CI based on (12) and the approximation (11) is quite comparable with the score CI by Miettinen and Nurminen (1985). It should be noted that our CI is in closed-form and simple to calculate whereas the score CI is numerically involved. Also, see Example 1 for illustration of the approximate methods to find CIs for the relative risk, ratio of odds, and for the difference $p_1 - p_2$.

Krishnamoorthy and Lee (2010) have also proposed fiducial quantity for a Poisson mean λ. Let X denote the total number of counts based on a random sample of size n from a Poisson(λ) distribution, and let k be an observed value of X. Then $.5\chi^2_{2k+1}/n$ is a fiducial variable for λ. The appropriate percentiles of the fiducial variable form a CI for λ. Approximate solutions to some two-sample problems or estimation of a weighted average $\sum_{i=1}^{k} w_i\lambda_i$, where λ_i is the mean of the ith Poisson distribution, can be readily obtained using (8) and (9). Applications of CIs for the weighted Poisson means can be found in Dobson et al. (1991).

4.2. *Problems Involving Normal Distributions*

We shall see some applications of the approximations in Section 2 for estimating the common mean of several normal populations, coefficient of variation, and estimation of survival probability for the normal case.

4.2.1. *Estimation of the common mean of several normal populations.* Let $(\bar{X}_i, S_i^2)$ denote the (mean, variance) based on a sample of size n_i from a $N(\mu, \sigma_i^2)$ distribution, $i = 1, \ldots, k$. Fairweather (1972) proposed a method of finding exact CIs for the common mean μ based on $(\bar{X}_i, S_i^2)$, $i = 1, \ldots, k$. This exact method requires upper percentiles of $W_t = |\sum_{i=1}^{k} u_i t_{m_i}|$, where t_{m_i} denote the t random variable with df $m_i = n_i - 1$ and $u_i = [\mathrm{Var}(t_{m_i})]^{-1}/\sum_{j=1}^{k}[\mathrm{Var}(t_{m_j})]^{-1}$. Exact percentiles of W_t can be obtained numerically for the case of $k = 2$, but they are difficult to find for $k \geq 3$. Approximate percentiles of W_t can be obtained using (8) and (9). For the case of $k = 2$, the approximate ones are compared with the exact percentiles of W_t given in Table 5 of Jordan and Krishnamoorthy (1996). The results are reported in Table 3 for some values of $m_1 \geq 5$ and $m_2 \geq 5$. We observe from Table 3 that the approximate ones are quite comparable with the exact ones.

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Table 3
Exact and (approximate) 95th percentiles of W_t

	m_2				
m_1	5	6	7	8	9
5	1.814 (1.818)				
6	1.766 (1.772)	1.722 (1.730)			
7	1.733 (1.740)	1.692 (1.700)	1.663 (1.672)		
8	1.709 (1.716)	1.669 (1.678)	1.641 (1.651)	1.621 (1.631)	
9	1.690 (1.698)	1.652 (1.661)	1.625 (1.635)	1.605 (1.615)	1.590 (1.600)
10	1.675 (1.684)	1.638 (1.647)	1.612 (1.622)	1.593 (1.602)	1.578 (1.587)

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Jordan and Krishnamoorthy (1996) have proposed an alternative exact method finding CIs for the above common mean problem, which requires the upper percentiles of $W_f = \sum_{i=1}^{k} w_i F_{1,m_i}$, where F_{1,m_i} is the F random variable with dfs 1 and $m_i = n_i - 1$, and $w_i = \xi_i^{-1}/(\sum_{j=1}^{k} \xi_j^{-1})$, where $\xi_i = \mathrm{Var}(F_{1,m_i})$, $i = 1, \ldots, k$. Approximate percentiles of W_f can be obtained using (9).

4.3. *Noncentral t Distribution, Coefficient of Variation, and Survival Probability*

4.3.1. *Noncentral t percentiles.* Approximations to noncentral t percentage points are proposed in many papers including Johnson and Welch (1940), Kramer (1963) and Akahira (1995). To obtain approximations using (11), let $t_m(\delta)$ denote the noncentral t random variable with df m and the noncentrality parameter δ, and let $t_{m;\alpha}(\delta)$ denote the αth quantile of $t_m(\delta)$. Noting that

$$t_m(\delta) \sim (Z + \delta)/\sqrt{V}, \tag{13}$$

where $Z \sim N(0, 1)$ independently of $V \sim \chi_m^2/m$, we can use (11) with $X = Z + \delta$ and $Y = \sqrt{\chi_m^2/m}$ to find approximate percentiles of $t_m(\delta)$. More specifically, using $Y_\alpha = \sqrt{\chi_{m;\alpha}^2/m}$, $X_\alpha = \delta + z_\alpha$, $\mu_x = \delta$ and $\mu_y = E(\sqrt{V})$, where

$$E(\sqrt{V}) = \left(\sqrt{\frac{2}{m}}\frac{\Gamma((m+1)/2)}{\Gamma(m/2)}\right) \simeq \frac{1}{\sqrt{e}}\left(1 + \frac{1}{m}\right)^{m/2}, \tag{14}$$

we can find approximate percentiles of $t_m(\delta)$. The approximation in (14) is obtained using the approximation for the gamma function given in Wilson and Hilferty (1931). This approximation for $E(\sqrt{V})$ is very satisfactory even for $m = 2$, and we shall use this approximation for calculation of CIs and their accuracy study in the sequel.

To judge the accuracy of the approximation to the noncentral t percentiles, we evaluated the exact and approximate percentiles for some values of m and α. Because $t_{m;\alpha}(\delta) = -t_{m;1-\alpha}(-\delta)$, it is enough to compare the percentiles for $\alpha \geq .5$ and $\delta > 0$. The calculated values are given in Table 4 for smaller values of m and $\delta = .5$ and 2.5. Examination of the results in Table 4 clearly indicates that the approximations are in general satisfactory, and the accuracy improves with increasing m.

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Table 4

Approximate and (exact) percentiles of the noncentral t with df m and the noncentrality parameter δ

	$\delta = .5$				$\delta = 2.5$			
	m				m			
α	3	4	5	8	3	4	5	8
.5	0.56 (0.54)	0.55 (0.53)	0.54 (0.53)	0.52 (0.52)	2.82 (2.77)	2.73 (2.69)	2.68 (2.65)	2.61 (2.59)
.7	1.18 (1.17)	1.13 (1.13)	1.11 (1.11)	1.08 (1.08)	3.83 (3.80)	3.60 (3.58)	3.47 (3.46)	3.29 (3.29)
.8	1.63 (1.63)	1.55 (1.55)	1.50 (1.50)	1.44 (1.44)	4.67 (4.64)	4.28 (4.27)	4.07 (4.06)	3.77 (3.77)
.9	2.40 (2.43)	2.22 (2.23)	2.11 (2.13)	1.98 (1.98)	6.25 (6.24)	5.48 (5.48)	5.07 (5.08)	4.54 (4.54)
.95	3.24 (3.33)	2.89 (2.95)	2.71 (2.75)	2.47 (2.49)	8.10 (8.16)	6.79 (6.84)	6.12 (6.17)	5.27 (5.30)
.99	5.65 (6.18)	4.65 (4.95)	4.17 (4.36)	3.57 (3.66)	14.1 (14.5)	10.6 (10.8)	8.91 (9.11)	7.01 (7.13)

Table 5
Coverage probabilities of 95% one-sided lower and (upper) confidence limits for $\bar{\Phi}((t-\mu)/\sigma)$

	n					
$(t-\mu)/\sigma$	2	3	5	7	8	10
0.0	.940 (.939)	.944 (.944)	.948 (.947)	.947 (.948)	.949 (.949)	.948 (.948)
0.2	.945 (.933)	.945 (.943)	.948 (.949)	.946 (.949)	.948 (.949)	.948 (.950)
0.6	.949 (.937)	.948 (.947)	.947 (.950)	.948 (.950)	.949 (.951)	.948 (.951)
1.0	.952 (.948)	.948 (.951)	.949 (.951)	.949 (.950)	.949 (.951)	.949 (.950)
1.2	.951 (.951)	.947 (.952)	.949 (.951)	.948 (.952)	.950 (.951)	.948 (.951)
1.6	.951 (.954)	.950 (.953)	.950 (.951)	.949 (.951)	.950 (.952)	.950 (.952)
1.8	.950 (.956)	.950 (.952)	.949 (.951)	.949 (.953)	.949 (.951)	.949 (.951)
2.2	.951 (.956)	.950 (.952)	.950 (.951)	.950 (.951)	.949 (.952)	.950 (.950)
2.4	.951 (.956)	.949 (.954)	.949 (.952)	.949 (.952)	.949 (.951)	.950 (.951)

Downloaded by [68.226.177.160] at 05:24 15 June 2016

4.3.2. *Coefficient of variation and survival probability.* Let $(\bar{X}, S^2)$ denote the (mean, variance) based on a random sample of size n from a $N(\mu, \sigma^2)$ distribution. Let $(\bar{x}, s^2)$ be an observed value of $(\bar{X}, S^2)$. Several authors have proposed approximate CIs for the coefficient of variation σ/μ. Among them, the modified McKay's CI by Vangel (1996) and the recent one by Donner and Zou (2012) seem to be satisfactory. We shall find approximate CIs for the coefficient of variation and survival probability following the exact approach by Johnson and Welch (1940). These authors showed that finding exact CIs for σ/μ and the survival probability $P(X > t) = 1 - \Phi((t-\mu)/\sigma) = \bar{\Phi}((t-\mu)/\sigma)$, where $\Phi(x)$ is the standard normal distribution function and t is a specified value, involves finding the root (with respect to δ) of the equation

$$t_{m;\alpha}(\delta) = t^*, \tag{15}$$

where t^* is a function of $(\bar{x}, s)$. An approximation to the root can be readily obtained as follows. Using (13) and (15), we see that δ is the solution of the equation $P(t^*\sqrt{V} - Z > \delta) = \alpha$, and so the value of δ that satisfies (15) is the $1-\alpha$ quantile of $t^*\sqrt{V} - Z$. An approximation to this quantile can be readily obtained from (9) as

$$t^* E(\sqrt{V}) + \sqrt{t^{*2}\left(E(\sqrt{V}) - v_u^*\right)^2 + z_\alpha^2}, \tag{16}$$

where $E(\sqrt{V})$ is given in (14), $v_u^* = \sqrt{\frac{\chi^2_{m;1-\alpha}}{m}}$ if $t^* > 0$, and it is $\sqrt{\frac{\chi^2_{m;\alpha}}{m}}$ if $t^* < 0$. Similarly, it can be shown that an approximation to δ that satisfies $t_{m;1-\alpha}(\delta) = t^*$ is given by

$$t^* E(\sqrt{V}) - \sqrt{t^{*2}\left(E(\sqrt{V}) - v_l^*\right)^2 + z_\alpha^2}, \tag{17}$$

where $v_l^* = \sqrt{\frac{\chi^2_{m;\alpha}}{m}}$ if $t^* > 0$, and it is $\sqrt{\frac{\chi^2_{m;1-\alpha}}{m}}$ if $t^* < 0$.

A CI for the survival probability $\bar{\Phi}((t-\mu)/\sigma)$ can be obtained from the one for $\eta = (t-\mu)/\sigma$. Let $\widehat{\eta} = (t - \bar{x})/s$. Then the Johnson-Welch exact CI for η is given by (L, U), where L is the root of the equation $t_{n-1;1-\alpha}(\sqrt{n}L) = \sqrt{n}\widehat{\eta}$, and U is the root of

the equation $t_{n-1;\alpha}(\sqrt{n}U) = \sqrt{n}\widehat{\eta}$. Approximate roots of these equations can be readily obtained from (16) and (17), and these roots form the CI for η, given by

$$(L_\eta, U_\eta) = \left(\widehat{\eta}E\left(\sqrt{V}\right) - \sqrt{\widehat{\eta}^2\left(E\left(\sqrt{V}\right) - v_l^*\right)^2 + \frac{z_\alpha^2}{n}},\ \widehat{\eta}E\left(\sqrt{V}\right) + \sqrt{\widehat{\eta}^2\left(E\left(\sqrt{V}\right) - v_u^*\right)^2 + \frac{z_\alpha^2}{n}}\right), \tag{18}$$

where v_u^* and v_l^* are as defined in (16) and (17), respectively, with $m = n - 1$. The $1 - 2\alpha$ CI for the survival probability $\bar{\Phi}(\eta)$ is $(\bar{\Phi}(U_\eta), \bar{\Phi}(L_\eta))$.

Johnson and Welch's (1940) exact CI for $\theta = \mu/\sigma$ is (L, U), where L is the root of the equation $t_{n-1;1-\alpha}(\sqrt{n}L) = \sqrt{n}\bar{x}/s$ and U is the root of the equation $t_{n-1;\alpha}(\sqrt{n}U) = \sqrt{n}\bar{x}/s$. The interval (U^{-1}, L^{-1}) is an exact $1 - 2\alpha$ CI for the coefficient of variation σ/μ. An approximate $1 - 2\alpha$ CI for the coefficient of variation σ/μ based on (16) and (17) is given by

$$\left[\left(\widehat{\theta}E\left(\sqrt{V}\right) + \sqrt{\widehat{\theta}^2\left(E\left(\sqrt{V}\right) - v_u^*\right)^2 + \frac{z_\alpha^2}{n}}\right)^{-1},\ \left(\widehat{\theta}E\left(\sqrt{V}\right) - \sqrt{\widehat{\theta}^2\left(E\left(\sqrt{V}\right) - v_u^*\right)^2 + \frac{z_\alpha^2}{n}}\right)^{-1}\right], \tag{19}$$

where $\widehat{\theta} = \bar{x}/s$.

To judge the accuracy of the approximate CI for the survival probability, we estimated the coverage probabilities using Monte Carlo simulation with 100,000 runs. For coverage studies, we used the approximate expression (14) for $E(\sqrt{V})$ in the CI (18). We report the coverage probabilities of the CI (18) for the survival probability in Table 5. The reported coverage probabilities clearly indicate that the CI for $\Phi(\eta)$ perform satisfactorily even for sample of size two. The new approximate CI (18) maintains the coverage probabilities very close to the nominal level. We also estimated the coverage probabilities of the CI for the coefficient variation for $\sigma/\mu = .1(.1).4$. They are not reported here; however, we note that the coverage probabilities are very close to the nominal level .95 for sample sizes of four or more.

5. Other Applications

We have also checked the validity of the MNAs in some other applications. For example, the approximations in (8) and (9) can be applied to find the percentiles of a linear combination of a set of independent random variables from a location–scale family. We verified the results for the symmetric location-scale families, namely, the logistic and the Laplace (double exponential), and found that the approximations work very satisfactorily, and much better than the usual approximation in (10). The approximations also work well for the extreme-value distribution (an asymmetric location–scale family), but less accurate than the results for the symmetric case.

Percentiles of $Q_{NC} = \sum_{i=1}^{k} w_i \chi_{m_i}^2(\delta_i)$, where $\chi_{m_i}^2(\delta_i)$'s are independent noncentral chi-square random variables with df = m_i and the noncentrality parameter δ_i, are required in many applications (see Imhof, 1961; Krishnamoorthy and Mondal, 2006; Zhang, 2005). The

Downloaded by [68.226.177.160] at 05:24 15 June 2016

normal approximate percentiles based on (8) and (9) for Q_{NC} are comparable with Imhof's approximation, in particular, for upper percentiles. Both Imhof's and our approximations are not good for finding lower percentiles of Q_{NC} when the dfs associated with the noncentral chi-squares are too small.

Computation of the percentiles of a doubly noncentral F distribution is one of the challenging problems, and popular software packages do not include routine for calculating percentage points of doubly noncentral F. The software "Dataplot" by the NIST computes only the distribution function, not its inverse function. Percentiles of a doubly noncentral F can be readily approximated using (11). To see this, let $F_{m_1,m_2}(\delta_1, \delta_2)$ denote the noncentral F random variable with the numerator df m_1, the denominator df m_2, and the noncentrality parameters δ_1 and δ_2. Noting that $F_{m_1,m_2}(\delta_1, \delta_2) \sim m_2\chi^2_{m_1}(\delta)/[m_1\chi^2_{m_2}]$, where the noncentral chi-square random variables are independent, we can use the approximations (11) with $X = \chi^2_{m_1}(\delta_1)$ and $Y = \chi^2_{m_2}(\delta_2)$. Multiplying the approximate percentage points of the ratio X/Y by (m_2/m_1), we can find approximate percentiles of $F_{m_1,m_2}(\delta_1, \delta_2)$.

A commonly used simple approximation for the percentiles of the doubly noncentral F is the one based on the Patnaik approximation to the noncentral chi-square distribution given by

$$\chi^2_n(\delta) \sim \rho\chi^2_\nu, \text{ where } \rho = \frac{n+2\delta}{n+\delta} \text{ and } \nu = \frac{(n+\delta)^2}{(n+2\delta)}. \tag{20}$$

Using the above approximation, we obtain

$$F_{m_1,m_2;\alpha} \simeq \frac{\rho_1 m_2 \nu_1}{\rho_2 m_1 \nu_2} F_{\nu_1,\nu_2;\alpha}, \tag{21}$$

where ρ_i and ν_i are as defined in (20).

Bulgren (1971) provided a series expression for the cumulative distribution function, and reported the probability $P\left(F_{m_1,m_2}(\delta_1, \delta_2) \le x\right)$ for some values of $(m_1, m_2, \delta_1, \delta_2, x)$. Part of the results from Table 1 of Bulgren (1971) along with the approximate ones based on (11) and (21) are given in Table 6. Comparison of the percentiles clearly indicates that the approximate percentiles based on (11) and (21) are satisfactory, and they are in close agreement with the Bulgren's exact results when the dfs are not too small; see the second part of Table 6.

6. Examples

In the preceding sections, we have showed several applications of the MNAs. In the following, we shall illustrate the approximations for the problems involving binomial and normal distributions.

Example 6.1. The data in Table 7 are taken from Example 3.7.1 of Krishnamoorthy (2006). The interest here is to assess risk of the long-term exposure to a chemical. Data were obtained from independent samples from exposed and unexposed groups as shown in Table 7.

Let p_e and p_u denote the true proportions of adults with symptoms in the exposed and unexposed groups, respectively.

In the notation of this article, we have $k_1 = 13, n_1 = 32, k_2 = 4$, and $n_2 = 25$. For ease of presentation, let $V_i = B_{k_i+.5,n_i-k_i+.5}$, and let $V_{i;\alpha}$ denote the α quantile of V_i,

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Table 6

Approximate $100q$th percentiles based on (11)) and the exact ones of doubly noncentral F with dfs m and n and the noncentrality parameters δ_1 and δ_2

$(m_1, m_2, \delta_1, \delta_2)$	q	exact	apprx (9)	apprx (19)	$(m_1, m_2, \delta_1, \delta_2)$	q	exact	apprx (9)	apprx (19)
(2, 4, 0.5, 0.5)	.94200	6.94	7.42	6.89	(4,30,10.0, 5.0)	.93693	6.39	6.39	6.37
(2, 15, 3, 3)	.79947	3.68	3.74	3.61	(4,30,24.0, 5.0)	.05705	2.69	2.68	2.77
(2,15, 0.5, 0.5)	.92342	3.68	3.78	3.67	(4,60, 0.5, 0.5)	.92784	2.53	2.54	2.52
(2,60, 0.5, 0.5)	.91414	3.15	3.18	3.14	(8,15, 4.0, 4.0)	.92066	2.74	2.66	2.63
(2, 4, 3.0, 3.0)	.93217	6.94	7.26	6.57	(8,60, 4.0, 9.0)	.86591	2.10	2.10	2.09
(2,30,12.0, 3.0)	.18459	3.32	3.30	3.38	(8,15, 0.5, 0.5)	.94523	2.64	2.66	2.64
(4, 4, 0.5, 0.5)	.95055	6.39	6.59	6.34	(8,60, 9.0, 9.0)	.64426	2.10	2.10	2.08
(4, 4, 2.0, 2.0)	.95623	6.39	6.56	6.05	(8,60, .50, .50)	.93712	2.10	2.10	2.10
(4, 4, 5.0, 5.0)	.97115	6.39	6.48	5.68	(8,30, 9.0, 9.0)	.77072	2.27	2.27	2.26

Table 7
Exposure data

Group	Symptoms present	Symptoms absent	Totals
Exposed	13	19	32
Unexposed	4	21	25
Totals	17	40	57

$i = 1, 2$. For this example, we calculated $E(V_1) = .4091$, $E(V_2) = .1731$, $V_{1;.025} = .2502$, $V_{2;.025} = .0565$, $V_{1;.975} = .5784$, and $V_{2;.975} = .3369$. The 95% CI (L, U) for $p_e - p_u$ is given by

$$L = E(V_1) - E(V_2) - \sqrt{(E(V_1) - V_{1;.025})^2 + (E(V_2) - V_{2;.975})^2} = .0078$$

and

$$U = E(V_1) - E(V_2) + \sqrt{(E(V_1) - V_{1;.975})^2 + (E(V_2) - V_{2;.025})^2} = .4416.$$

One-sided CLs are similarly obtained using the 5th and 95th percentiles of the beta distributions. Using these above numerical values in (11), we can find a 95% CI for the relative risk.

To find the approximate CI for the ratio of odds, let $F_i = F_{2k_i+1,2n_i-2k_i+1}$, $i = 1, 2$. Noting that $E(F_{a,b}) = b/(b-2)$, we find $E(F_1) = 1.0541$, $E(F_2) = 1.0489$, $F_{1;.025} = .4821$, $F_{1;.975} = 1.982$, $F_{2;.025} = .2862$, and $F_{2;.975} = 2.4269$. Using these numbers in (12), we can find 95% CI for the odds ratio; see Table 8.

Several approximate methods for finding CIs for the difference, relative risk, and for the odds ratio are proposed in the literature (see, e.g., Bailey, 1987; Bedrick, 1987; Krishnamoorthy and Lee, 2010; Newcombe, 1998, and the references therein). Among all the methods, the Miettinen-Nurminen (1985) likelihood-score CIs appear to be satisfactory and popular. Even though the Miettinen-Nurminen (M-N) method is numerically involved, an R package "gsDesign" is available to compute the CIs.

We calculated one- and two-sided 95% CIs for the difference $p_e - p_u$, relative risk p_e/p_u, and for the odds ratio $p_e/(1-p_e)/[p_u/(1-p_u)]$, using our approximate method and the M-N methods, and reported them in Table 8. The R package "ciBinomial(k1,k2,n1,n2,alpha=al,adj=0,scale="XX")," where XX is Difference, RR or OR, was used to find M-N confidence intervals. Examination of CIs for all three problems

Table 8
95% confidence intervals for the difference between proportions of people with symptoms in the exposed and unexposed groups

	Difference			Relative risk			Odds ratio		
Method	CI	L	U	CI	L	U	CI	L	U
MNA	(.0078, .4416)	.0461	.4120	(1.027, 7.524)	1.171	6.143	(1.052,13.72)	1.267	10.86
M-N	(.0073, .4537)	.0478	.4226	(1.024, 6.824)	1.169	5.857	(1.035,12.27)	1.254	10.20

Downloaded by [68.226.177.160] at 05:24 15 June 2016

indicate that the CIs based on the approximate method and the M-N method are quite comparable. The results of this example show that the MNAs are not only simple to use, but also provide satisfactory solutions that are comparable with those based on other numerically involved methods.

Example 6.2. The data for this example are taken from Krishnamoorthy et al. (2006), and they represent air lead levels collected by the National Institute of Occupational Safety and Health (NIOSH) from a facility for health hazard evaluation purpose (HETA 89-052). The air lead levels were collected from 15 different areas within the facility.

Air Lead Levels ($\mu g/m^3$):

200, 120, 15, 7, 8, 6, 48, 61, 380, 80, 29, 1000, 350, 1400, 110

The occupational exposure limit set by the NIOSH is $120\mu g/m^3$. The interest here is to find an upper confidence limit for the exceedance probability $P(X > 120)$, where X represents the measurement at a randomly selected area within the facility. If a 95% upper CL for this exceedance probability is less than .05, then one may conclude that the facility is in compliance. Krishnamoorthy et al. (2006) showed that the log-normality assumption for the data is tenable. Since $P(X > 120) = P(\ln(X) > \ln(120))$, we can apply the approximate method of Section 4.2 to find an upper CL based on the log-transformed data. For the log-transformed data, the mean $\bar{x} = 4.333$ and $s = 1.739$. Noting that $n = 15$, we evaluated the expectation in (14) as $E(\sqrt{V}) = .9823$; the approximation in (14) is .9831. The value of $\widehat{\eta}$ in (18) is $(\ln(120) - 4.333)/1.739 = .2614$. To find a 95% upper confidence limit for $P(X > 120)$, we calculated $v_u^* = 1.3007$ and $z_{.95} = 1.645$. Using these numbers in (18), we found the lower CL for η as $-.1748$, and the upper CL for the exceedance probability is $\bar{\Phi}(-.1748) = \mathbf{.5694}$. The exact lower CL for η satisfies the equation $t_{14;.95}(\sqrt{15}\eta) = .2614\sqrt{15}$. It can be readily verified that the value of η is $-.1754$, and the exact upper CL for the survival probability is $\bar{\Phi}(-.1754) = \mathbf{.5696}$. Thus, we see that the approximate CL and the exact one for this example are practically the same.

For the sake of illustration, we shall compute 90% CI for the coefficient of variation based on log-transformed data. Noting that $\widehat{\theta} = \bar{x}/s = 2.4917$, and using the calculations in the preceding paragraph, we computed the approximate 95% CI in (19) as **(.2986, .6266)**. The exact CI is formed by the reciprocal of the roots of $t_{14,.95}(\sqrt{15}\theta) = 9.6502$, and $t_{14,.05}(\sqrt{15}\theta) = \sqrt{15} \times 9.6502$. The root of the former equation is 1.5910, and of the latter is 3.3446. The exact 90% CI for the coefficient of variation is $(1/3.3446, 1/1.5910) =$ (**.2990**, **.6285**). We once again note that the approximate and the exact solutions are in good agreement.

7. Concluding Remarks

The simple normal approximation in (10) requires only the percentiles of the standard normal distribution, and they are easy to use to approximate the percentiles of a linear function of independent random variables provided the variance expressions of the individual random variables are available. However, we have seen that the usual approximations are unsatisfactory when the distributions of the random variables are skewed. The MNA is also simple to use in situations where the expectations and percentiles of the individual random variables can be readily obtained. The proposed modified normal-based approximations provide satisfactory solutions even when a linear function involves random variables from different families. Our Monte Carlo simulation studies, comparison with exact results, and

Downloaded by [68.226.177.160] at 05:24 15 June 2016

applications to some examples clearly indicate that the modified normal-based approximations provide simple closed-form satisfactory solutions in many applications. However, like any other approximate methods, the results based on our approximations should be evaluated for their accuracy using Monte Carlo simulation or numerically before recommending for practical applications.

Acknowledgments

The author is thankful to a reviewer for providing useful comments and suggestions.

References

Akahira, M. (1995). A higher order approximation to a percentage point of the noncentral t distribution. *Communication in Statistics – Simulation and Computation* 24:595–605.

Andrés, A. M., Hernáandez, M. A. (2014). Two-tailed approximate confidence intervals for the ratio of proportions. *Statistics and Computing* 20:369–387.

Andrés, A. M., Hernáandez, M. A., Tejedor, I. H. (2011). Inference about a linear function of proportions. *Statistical Methods in Medical Research* 20:369–387.

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

Bedrick, E. J. (1987). A family of confidence intervals for the ratio of two binomial proportions. *Biometrics* 43:993–998.

Bulgren, W. G. (1971). On representations of the doubly non-central F distribution. *Journal of the American Statistical Association* 66:184–187.

Dobson, A. J., Kuulasmmaa, K., Ebrerle, E., Schere, J. (1991). Confidence intervals for weighted sums of Poisson parameters. *Statistics in Medicine* 10:457–462.

Donner, A., Zou, G. Y. (2012). Closed-form confidence intervals for functions of normal mean and standard deviation. *Statistical Methods in Medical Research* 21:347–359.

Fairweather, W. R. (1972). A method of obtaining exact confidence interval for the common mean of several normal populations. *Applied Statistics* 21:229–223.

Graybill, F. A., Wang, C.-M. (1980). Confidence intervals on nonnegative linear combinations of variances. *Journal of the American Statistical Association* 75:869–873.

Imhof, J. P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* 48:419–426.

Johnson, N. L., Welch, B. L. (1940). Application of the noncentral t-distribution. *Biometrika* 31:362–389.

Jordan, S. M., Krishnamoorthy, K. (1996). Exact confidence intervals for the common mean of several normal populations. *Biometrics* 52:77–86.

Kramer, K. H. (1963). Tables for constructing concordance limits on the multiple correlation coefficient. *Journal of the American Statistical Association* 58:1082–1085.

Krishnamoorthy, K., Lee, M. (2010). Inference for functions of parameters in discrete distributions based on fiducial approach: Binomial and Poisson cases. *Journal of Statistical Planning and Inference* 140:1182–1192.

Krishnamoorthy, K., Mathew, T., Ramachandran, G. (2006). Generalized p-values and confidence limits: A novel approach for analyzing lognormally distributed exposure data. *Journal of Occupational and Environmental Hygiene* 3:252–260.

Krishnamoorthy, K., Mondal, S. (2006). Improved tolerance factors for multivariate normal distributions. *Communications in Statistics – Simulation and Computation* 25:461–478.

Li, Y., Koval, J. J., Donner, A., Zou, G. Y. (2010). Interval estimation for the area under the receiver operating characteristic curve when data are subject to error. *Statistics in Medicine* 29:2521–2531.

Miettinen, O., Nurminen, M. (1985). Comparative analysis of two rates. *Statistics in Medicine* 4:213–226.

Downloaded by [68.226.177.160] at 05:24 15 June 2016

Newcombe, R. G. (1998). Interval estimation for the difference between independent proportions: Comparison of eleven methods. *Statistics in Medicine* 17:873–890.

Satterthwaite, F. E. (1941). Synthesis of variance. *Psychometrika* 6:309–316.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin* 2:110–114.

Vangel, M. G. (1996). Confidence intervals for a normal coefficient of variation. *The American Statistician* 50:21–26.

Welch, B. L. (1938). The significance of the difference between two means when the population variances are unequal. *Biometrika* 29:350–361.

Welch, B. L. (1947). The generalization of "student's" problem when several different population variances are involved. *Biometrika* 34:28–35.

Wilson, E. B., Hilferty, M. M. (1931). The distribution of chi-squares. *Proceedings of the National Academy of Sciences* 17:684–688.

Zhang, D. (2013). Test and interval estimation methods for binomial and Poisson distributions. Ph.D. dissertation, Department of Mathematics, University of Louisiana at Lafayette, Lafayette, LA.

Zhang, J. (2005). Approximate and asymptotic distributions of chi-squared-type mixtures with applications. *Journal of the American Statistical Association* 100:273–285.

Zou, G. Y. (2013). Confidence interval estimation for the Bland-Altman limits of agreement with multiple observations per individual. *Statistical Methods in Medical Research* 22:630–642.

Zou, G. Y., Donner, A. (2008). Construction of confidence limits about effect measures: A general approach. *Statistics in Medicine* 27:1693–1702.

Zou, G. Y., Huo, C. Y., Taleban, J. (2009a). Confidence interval estimation for lognormal data with application to health economics, *Computational Statistics and Data Analysis* 53:3755–3764.

Zou, G. Y., Taleban, J., Huo, C. Y. (2009b). Simple confidence intervals for lognormal means and their differences with environmental applications. *Environmetrics* 20:172–180.

Downloaded by [68.226.177.160] at 05:24 15 June 2016