

Construction of simultaneous tolerance intervals for several normal distributions

K. Krishnamoorthy & Saptarshi Chakraberty

To cite this article: K. Krishnamoorthy & Saptarshi Chakraberty (2022) Construction of simultaneous tolerance intervals for several normal distributions, Journal of Statistical Computation and Simulation, 92:1, 101-114, DOI: [10.1080/00949655.2021.1932885](https://doi.org/10.1080/00949655.2021.1932885)

To link to this article: <https://doi.org/10.1080/00949655.2021.1932885>



Published online: 08 Jun 2021.



Submit your article to this journal [↗](#)



Article views: 23



View related articles [↗](#)



View Crossmark data [↗](#)



Construction of simultaneous tolerance intervals for several normal distributions

K. Krishnamoorthy and Saptarshi Chakraberty

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

ABSTRACT

Methods for computing simultaneous tolerance limits and tolerance intervals (TIs) for several normal populations with a common unknown variance are proposed. We propose numerical methods that determine the confidence coefficient so that the simultaneous TIs include at least specified proportions of the populations with the intended confidence level. The methods are exact and as simple as the existing ones. They are also applicable when sample sizes are unequal and content levels of simultaneous TIs are unequal. Algorithms and R functions are provided to compute factors for constructing simultaneous one-sided tolerance limits, tolerance intervals and equal-tailed tolerance intervals. The methods are illustrated using an example with a real data set.

ARTICLE HISTORY

Received 21 December 2020
Accepted 17 May 2021

KEYWORDS

Content level; coverage probability; equal-tailed tolerance interval; tolerance factor



2010 MATHEMATICS

SUBJECT CLASSIFICATIONS
62F25; 62P30

1. Introduction

Among statistical intervals, confidence intervals (CIs) and prediction intervals are well known and commonly used in applications. A confidence interval based on a sample is constructed so that it would include a parameter of interest of the sampled population with a specified confidence. A prediction interval is used to predict the future observation from a population based on a currently available sample from the population. In some applications, one may want to find a sample interval that would include at least a proportion p of the population with confidence $1 - \alpha$. Such an interval is referred to as the $(p, 1 - \alpha)$ tolerance interval (TI). A one-sided $(p, 1 - \alpha)$ upper tolerance limit (UTL) is simply a $(1 - \alpha)$ upper confidence limit for the 100 p th percentile of the population. A $(p, 1 - \alpha)$ lower tolerance limit (LTL) is a $(1 - \alpha)$ lower confidence limit for the 100 $(1 - p)$ percentile of the population. Another type of two-sided TI (L, U) , referred to as the equal-tailed TI, is constructed so that no more than proportion $(1 - p)/2$ of the population is less than L and no more than $(1 - p)/2$ of the population is greater than U . For a good exposition of TIs and β -expectation TIs along with applications, see the books by Guttman [1] and Krishnamoorthy and Mathew [2].

Exact methods for computing all types of TIs are available for normal distributions. The methods for the normal case date back to 1940s (see [3–6] and Chapter 2 of the

CONTACT K. Krishnamoorthy  krishna@louisiana.edu Department of Mathematics,  University of Louisiana at Lafayette, Lafayette, LA 70504, USA

book by Krishnamoorthy and Mathew [2]). Following the approach for the normal case, Krishnamoorthy and Xie [7] have provided a general approach for constructing TIs for a symmetric location-scale family of distributions. Zou and Young [8] have applied bootstrap calibration to improve the coverage probabilities of two-sided parametric TIs based on one-sided TIs with Bonferroni adjustments. Hoang-Nguyen-Thuy and Krishnamoorthy [9] have proposed a method of finding exact factors for computing two-sided tolerance intervals based on modified one-sided factors. Their approach is applicable to any symmetric or asymmetric location-scale distribution. The R package by Young [10] or the *StatCalc* software that accompanies the book [11] can be used to compute the factors for finding tolerance intervals for the normal and other continuous/discrete distributions.

To define simultaneous TIs, let us suppose that there are l normal populations with means μ_1, \dots, μ_l and a common variance σ^2 , say, $N(\mu_1, \sigma^2), \dots, N(\mu_l, \sigma^2)$. Let X_{i1}, \dots, X_{in_i} denote the sample drawn from the $N(\mu_i, \sigma^2)$ distribution, $i = 1, \dots, l$. Let $N = \sum_{i=1}^l n_i$. Define the sample mean, variance and the pooled variance as

$$\bar{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}, \quad S_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_i)^2, \quad \text{and} \quad S_c^2 = \frac{1}{N - l} \sum_{i=1}^l (n_i - 1) S_i^2,$$

respectively. On the basis of \bar{X}_i and S_c^2 , we consider the problems of constructing

- (i) simultaneous one-sided lower (upper) tolerance limits (TLs) of the form $\bar{X}_i - k_i S_c$ ($\bar{X}_i + k_i S_c$) so that

$$P(\bar{X}_i - k_i S_c < \mu_i - z_{p_i} \sigma, i = 1, \dots, l) = 1 - \alpha,$$

and

$$P(\bar{X}_i + k_i S_c > \mu_i + z_{p_i} \sigma, i = 1, \dots, l) = 1 - \alpha,$$

- (ii) two-sided tolerance intervals of the form $\bar{X}_i \pm k'_i S_c$, so that

$$P_{\bar{X}_i, S_c} \{P_{X_i}(\bar{X}_i - k'_i S_c \leq X_i \leq \bar{X}_i + k'_i S_c) \geq p_i, i = 1, \dots, l\} = 1 - \alpha,$$

where X_i 's are independent random variables with $X_i \sim N(\mu_i, \sigma^2), i = 1, \dots, l$, and

- (iii) two-sided equal-tailed TIs of the form $\bar{X}_i \pm k_i^e S_c$ so that it would include the population interval $(\mu_i - z_{\frac{1+p_i}{2}} \sigma, \mu_i + z_{\frac{1+p_i}{2}} \sigma), i = 1, \dots, l$, with confidence $1 - \alpha$. That is,

$$P\left(\bar{X}_i - k_i^e S_c \leq \mu_i - z_{\frac{1+p_i}{2}} \sigma \text{ and } \mu_i + z_{\frac{1+p_i}{2}} \sigma \leq \bar{X}_i + k_i^e S_c, i = 1, \dots, l\right) = 1 - \alpha.$$

For problems (i) and (ii), Mee [12] has proposed methods for computing the factors when the sample sizes are equal. He also has suggested a method of finding approximate factors when the sample sizes are unequal. Mee has noted that his approximate method will give an actual confidence level of at least 94% (when $1 - \alpha = .95$ and $p = .99$) if the degrees of freedom (df) $n_i - 1 \geq 30$ for all i and the ratio (largest n_i)/(smallest n_i) ≤ 2 .

In this article, we propose methods for computing tolerance factors for the aforementioned three problems using the idea of Hoang-Nguyen-Thuy and Krishnamoorthy [9].

Unlike the methods by Mee [12], our methods are applicable to find exact simultaneous TIs with different content levels p_1, \dots, p_l based on unequal sample sizes. In addition, we provide an exact method of computing equal-tailed TIs. In the following section, we first describe a method for computing factors for simultaneous one-sided tolerance limits and a method for computing factors for simultaneous two-side tolerance intervals. In Section 2.3, we provide a method for computing simultaneous equal-tailed TIs. In Section 3, we illustrate the methods using an example. Some concluding remarks are given in Section 4.

2. Simultaneous tolerance intervals

In the following, we describe a method for computing factors to construct $(p_1, 1 - \alpha), \dots, (p_l, 1 - \alpha)$ simultaneous TIs for populations $1, \dots, l$, respectively. The derivations of our methods for constructing one-sided TIs and two-sided TIs are similar to the one in Mee [12], except that our methods are applicable when sample sizes are unequal, and the methods are generalization of one-sample methods for finding factors by Hoang-Thuy-Nguyen and Krishnamoorthy [9].

2.1. Simultaneous one-sided tolerance limits

Let z_p denote the p th quantile of the standard normal distribution. One-sided upper tolerance limits are constructed so that

$$\bar{X}_i + k_i S_c \geq \mu_i + z_{p_i} \sigma, \quad i = 1, \dots, l,$$

with probability $1 - \alpha$. The k_i 's are called factors, and they are to be determined such that

$$P_{\bar{X}_1, \dots, \bar{X}_l, S_c}(\bar{X}_i + k_i S_c \geq \mu_i + z_{p_i} \sigma, i = 1, \dots, l) = 1 - \alpha. \tag{1}$$

Letting $Z_i = \sqrt{n_i}(\bar{X}_i - \mu_i)/\sigma, i = 1, \dots, l$, and $U^2 = S_c^2/\sigma^2$, we can express (1) as

$$P_{Z_1, \dots, Z_l, U}(Z_i - \sqrt{n_i} z_{p_i} \geq -\sqrt{n_i} k_i U, i = 1, \dots, l) = 1 - \alpha.$$

Noticing that Z_i and $-Z_i$ are identically distributed, we can rewrite the above equation as

$$E_W \left[\prod_{i=1}^l \Phi \left(\sqrt{n_i} \left(\frac{k_i W}{\sqrt{N-l}} - z_{p_i} \right) \right) \right] = 1 - \alpha, \tag{2}$$

where $W^2 = (N - l)U^2 \sim \chi_{N-l}^2$ distribution.

The factors k_i 's should be different when the sample sizes are unequal. In this case, it is difficult to solve Equation (2) for k_1, \dots, k_l . It is also possible that, for a given (n_1, \dots, n_l) and p_1, \dots, p_l , several sets of k_1, \dots, k_l satisfy Equation (2). A unique way of determining k_i 's is to choose k_i 's as a function of γ , where γ is to be determined so that (2) holds. A choice, based on one-sample factors, is

$$k_{i;\gamma} = \frac{1}{\sqrt{n_i}} t_{n_i-1;\gamma}(z_{p_i} \sqrt{n_i}), \quad i = 1, \dots, l,$$

where $t_{m;q}(\delta)$ is the q th quantile of the noncentral t distribution with degrees of freedom m and the noncentrality parameter δ . Note that, for a given $(n_1, \dots, n_l), (p_1, \dots, p_l)$ and $(1 -$

α), $k_{i;\gamma}$ is an increasing function of γ , as a result, the coverage probability is an increasing function of γ . So the factors $k_{i;\gamma}$ determined above are unique. The value of γ is determined to satisfy the coverage requirement that

$$f(\gamma) = \frac{1}{2^{\frac{N-l}{2}} \Gamma\left(\frac{N-l}{2}\right)} \int_0^\infty \left[\prod_{i=1}^l \Phi\left(\sqrt{n_i} \left(\frac{k_{i;\gamma} \sqrt{x}}{\sqrt{N-l}} - z_{p_i}\right)\right) \right] e^{-x/2} x^{\frac{N-l}{2}-1} dx = 1 - \alpha. \quad (3)$$

Let γ^* denote the value of γ that satisfies (3). Then $k_{i;\gamma^*}$'s are the tolerance factors for computing $(p_i, 1 - \alpha)$, $i = 1, \dots, l$, simultaneous TLs. The integral above can be evaluated using the basic R function `integrate()` and the equation $f(\gamma) - (1 - \alpha) = 0$ can be solved using the basic R function `uniroot()` with root bracketing interval $(1 - \alpha - .4, 1 - \alpha)$. An R function to compute the value γ^* and the factors is given in the appendix.

For the special case of $n_1 = \dots = n_l = n$ and $p_1 = \dots = p_l = p$, it is reasonable to require that $k_1 = \dots = k_l = k^*$, and this common tolerance factor k^* is the solution of the integral equation

$$\frac{1}{2^{\frac{N-l}{2}} \Gamma\left(\frac{N-l}{2}\right)} \int_0^\infty \left[\Phi\left(\sqrt{n} \left(\frac{k^* \sqrt{x}}{\sqrt{N-l}} - z_p\right)\right) \right]^l e^{-x/2} x^{\frac{N-l}{2}-1} dx = 1 - \alpha. \quad (4)$$

The above equation follows from (3). Mee [12] has tabulated the values of k^* satisfying (4) for several values of n ranging from 2 to 1000, $l = 2, 3, 4, 5, 6, 8, 10$, $p = 0.90, 0.95, 0.99$ and $1 - \alpha = 0.90, 0.95, 0.99$. The R function given in the appendix can also be used to compute the value of k^* for equal/unequal sample sizes.

In Table 1, we report the values of γ so that $t_{n_i-1;\gamma}(z_p \sqrt{n_i})/\sqrt{n_i}$ are factors for constructing (.90, .95) simultaneous one-sided tolerance limits. The values of γ are given for $l = 2, \dots, 5$ and some sample sizes. As an example, when $(n_1, n_2, n_3) = (12, 18, 16)$, the value of γ is .9348 and the factors are

$$k_{i;\gamma} = \frac{1}{\sqrt{n_i}} t_{n_i-1;.9348}(z_{.9} \sqrt{n_i}) \quad \text{with } (k_{1;\gamma}, k_{2;\gamma}, k_{3;\gamma}) = (2.117, 1.908, 1.960). \quad (5)$$

That is, $\bar{X}_i - k_{i;\gamma} S_c$, $i = 1, 2, 3$, are (.90, .95) simultaneous lower tolerance limits and $\bar{X}_i + k_{i;\gamma} S_c$, $i = 1, 2, 3$, are (.90, .95) simultaneous upper tolerance limits. Suppose we need to compute (.80, .95), (.90, .95) and (.95, .95) upper tolerance limits, then the value of γ is .9378 and the factors are

$$k_{i;\gamma} = \frac{1}{\sqrt{n_i}} t_{n_i-1;.9378}(z_{p_i} \sqrt{n_i}) \quad \text{with } (k_{1;\gamma}, k_{2;\gamma}, k_{3;\gamma}) = (1.532, 1.919, 2.454). \quad (6)$$

Using the R function in the appendix, the factors in (5) can be computed as

```
> norm.simult.one.fac(c(12, 18, 16), rep(.90, 3), .95)
> .9348 2.117 1.908 1.960
```

Table 1. The values of γ so that $t_{n_i-1;\gamma}(z_{p_i}\sqrt{n_i})/\sqrt{n_i}$ are factors for constructing $(p_1, .95), \dots, (p_l, .95)$ simultaneous one-sided tolerance limits for l normal populations.

$l = 2$		$l = 3$		$p_1 = \dots = p_l = .90$		$l = 5$	
(n_1, n_2)	γ	(n_1, n_2, n_3)	γ	(n_1, \dots, n_4)	γ	(n_1, \dots, n_5)	γ
(4, 4)	.8169	(3, 3, 3)	.7413	(8, 8, 8, 8)	.8343	(6, 6, 6, 6, 6)	.8061
(7, 4)	.8442	(3, 5, 6)	.8112	(8, 4, 3, 6)	.8126	(6, 5, 3, 2, 8)	.8090
(7, 9)	.8556	(9, 8, 11)	.8490	(9, 11, 4, 21)	.8598	(9, 15, 13, 12, 21)	.8655
(12, 20)	.8762	(12, 18, 16)	.8695	(9, 7, 3, 5)	.8237	(9, 5, 33, 7, 18)	.8684
(22, 20)	.8815	(5, 19, 21)	.8727	(15, 25, 7, 21)	.8733	(13, 25, 12, 37, 48)	.8881
(32, 40)	.8899	(23, 45, 33)	.8894	(18, 21, 17, 31)	.8807	(20, 35, 45, 48, 50)	.8986
$(.8, .9)$		$(.8, .9, .95)$		(p_1, \dots, p_l)		$(.8, .85, .85, .9, .95)$	
(n_1, n_2)	γ	(n_1, n_2, n_3)	γ	(n_1, \dots, n_4)	γ	(n_1, \dots, n_5)	γ
(4, 4)	.9159	(3, 3, 3)	.8765	(8, 8, 8, 8)	.9246	(6, 6, 6, 6, 6)	.9135
(7, 4)	.9285	(3, 5, 6)	.9020	(8, 4, 3, 6)	.9144	(6, 5, 3, 2, 8)	.9113
(7, 9)	.9345	(9, 8, 11)	.9281	(9, 11, 4, 21)	.9351	(9, 15, 13, 12, 21)	.9388
(12, 20)	.9446	(12, 18, 16)	.9378	(9, 7, 3, 5)	.9236	(9, 5, 33, 7, 18)	.9390
(22, 20)	.9469	(5, 19, 21)	.9335	(15, 25, 7, 21)	.9429	(13, 25, 12, 37, 48)	.9485
(32, 40)	.9513	(23, 45, 33)	.9477	(18, 21, 17, 31)	.9457	(20, 35, 45, 48, 50)	.9552

Similarly, the factors in (6) can be computed using the same R function as

```
> norm.simult.one.fac(c(12, 18, 16), c(.80, .90, .95), .95)
[1] 0.9378 1.532 1.920 2.454
```

2.2. Simultaneous two-sided tolerance intervals

To construct $(p, 1 - \alpha)$ simultaneous tolerance intervals for l normal populations with common variance σ^2 , we need to determine the tolerance factors k'_1, \dots, k'_l so that

$$P_{\bar{X}_1, \dots, \bar{X}_l, S_c} \{P_{X_i}(\bar{X}_i - k'_i S_c \leq X_i \leq \bar{X}_i + k'_i S_c \mid \bar{X}_i, S_c) \geq p_i, i = 1, \dots, l\} = 1 - \alpha, \quad (7)$$

where X_1, \dots, X_l are independent random variables with $X_i \sim N(\mu_i, \sigma^2), i = 1, \dots, l$. Furthermore, $X_1, \dots, X_l, \bar{X}_1, \dots, \bar{X}_l$ and S_c are mutually independent. Let $Y_i = (\bar{X}_i - \mu_i)/\sigma, i = 1, \dots, l$ and $U^2 = S_c^2/\sigma^2$ so that

$$Y_i \sim N(0, 1/n_i) \text{ independently of } U^2 \sim \frac{\chi_{N-l}^2}{N-l}.$$

In terms of these variables, we can write (7) as

$$P_{Y_1, \dots, Y_l, U^2} \{P(Y_i - k'_i U \leq Z_i \leq Y_i + k'_i U \mid Y_i, U^2) \geq p_i, i = 1, \dots, l\} = 1 - \alpha. \quad (8)$$

Notice that, for a fixed $Y_i, \Phi(Y_i + k'_i U) - \Phi(Y_i - k'_i U) \geq p$ if and only if $k_i U \geq r_i$ or $U^2 \geq r_i^2/k_i^2$, where r_i is the solution of the equation $\Phi(Y_i + r_i) - \Phi(Y_i - r_i) = p_i$ (see Section 1.2 of [2]). For a fixed Y_i , it can be easily verified that $r_i = \chi_{1;p_i}^2(Y_i^2), i = 1, \dots, p$. Thus, we need to determine the factors k'_1, \dots, k'_l such that

$$E_{Y_1, \dots, Y_l} \left\{ P \left(U^2 > \frac{\chi_{1;p_i}^2(Y_i^2)}{k_i^2}, i = 1, \dots, l \mid Y \right) \right\} = 1 - \alpha, \quad (9)$$

or equivalently,

$$E_{Y_1, \dots, Y_l} \left\{ P \left(U^2 > \max_{1 \leq i \leq l} \left\{ \frac{\chi_{1;p_i}^2(Y_i^2)}{k_i^2} \right\} \mid Y \right) \right\} = 1 - \alpha. \tag{10}$$

As in the preceding section, we choose k'_i to be a function of γ given by

$$k'_{i;\gamma} = \frac{1}{\sqrt{n_i}} t_{n_i-1; (1+\gamma)/2} \left(\frac{z_{1+p_i}}{2} \sqrt{n_i} \right), \quad i = 1, \dots, l.$$

Thus, instead of determining k'_i 's that satisfy (10), we determine γ so that

$$E_{Y_1, \dots, Y_l} \left\{ P \left(U^2 > \max_{1 \leq i \leq l} \left\{ \frac{\chi_{1;p_i}^2(Y_i^2)}{k'^2_{i;\gamma}} \right\} \mid Y \right) \right\} = 1 - \alpha. \tag{11}$$

As the left-hand side of (11) involves an l dimensional integral, the above equation is difficult to solve numerically. However, we can estimate the left-hand side (LHS) of (11) by Monte Carlo simulation and determine the value of γ so that the LHS is equal to $1 - \alpha$. For a given (n_1, \dots, n_l) , (p_1, \dots, p_l) and $1 - \alpha$, the R function `norm.simult.two.fac()` given in the appendix can be used to find the value of γ and the factors $k'_{i;\gamma}$, $i = 1, \dots, l$.

As shown in [12], for the special case of $n_1 = \dots = n_l = n$ and $p_1 = \dots = p_l$, (11) can be expressed as an integral as follows. For the case of equal sample size, we can assume that $k_1 = \dots = k_l = k$ and write (10) as

$$E_{|Y|_{(l)}} \left\{ P \left[U^2 > \left\{ \frac{\chi_{1;p}^2(|Y|_{(l)}^2)}{k^2} \right\} \mid |Y|_{(l)} \right] \right\} = 1 - \alpha, \tag{12}$$

where $|Y|_{(l)} = \max_{1 \leq i \leq l} |Z_i|/\sqrt{n}$ and Z_i 's are independent $N(0, 1)$ random variables. Using the probability density function of $|Y|_{(l)}\sqrt{n}$, we can write expression (12) as

$$2l \int_0^\infty P \left(\chi_{N-l}^2 \geq \frac{(N-l)\chi_{1;p}^2(z^2/n)}{k^2} \right) [2\Phi(z) - 1]^{l-1} \phi(z) dz = 1 - \alpha. \tag{13}$$

The value of k' that satisfies the above equation can be obtained using numerical integration and a root finding method. For more details on computation, see [12].

Mee [12] has tabulated the values of k' satisfying (13) for several values of n ranging from 2 to 1000, $l = 2, 3, 4, 5, 6, 8, 10$, $p = .90, .95, .99$ and $1 - \alpha = 0.90, 0.95, 0.99$. The R function given in the appendix can also be used to compute the factors for the special case of $n_1 = \dots = n_l = n$ and $p_1 = \dots = p_l = p$. This R function is based on Algorithm ??.

Algorithm 1: For a given (n_1, \dots, n_l) , (p_1, \dots, p_l) and $1 - \alpha$:

- (1) If $n_1 = \dots = n_l = n$ and $p_1 = \dots = p_l = p$, then let $k'_\gamma = t_{n-1; (1+\gamma)/2}(\sqrt{n}z_{(1+p)/2})$.
- (2) Set the function

$$f(\gamma) = 2l \int_0^\infty P \left(\chi_{N-l}^2 \geq \frac{(N-l)\chi_{1;p}^2(z^2/n)}{k'^2_\gamma} \right) [2\Phi(z) - 1]^{l-1} \phi(z) dz - (1 - \alpha).$$

- (3) Evaluate the integral numerically and solve the equation $f(\gamma) = 0$ using a bisection method with root bracketing interval $(1 - \alpha - .4, 1 - \alpha)$.
- (4) Let γ^* be the root of the equation obtained in the above step. Then $k'_{i;\gamma^*} = t_{n-1;(1+\gamma^*)/2}(\sqrt{nz_{(1+p)/2}}/\sqrt{n_i})$ is the factor for constructing simultaneous TIs. If n_i 's are unequal or p_i 's are unequal, then set $z_{p_i} = \Phi^{-1}((1 + p_i)/2)$, $i = 1, \dots, l$.
- (5) Generate $Y_i \sim N(0, 1/n_i)$, $i = 1, \dots, l$.
- (6) Set $k'_{i;\gamma} = \frac{1}{\sqrt{n_i}} t_{n_i-1;(1+\gamma)/2}(z_{1+p_i}/2 \sqrt{n_i})$, $i = 1, \dots, l$.
- (7) Set $M_x = \max_{1 \leq i \leq l} \{ \frac{\chi_{1+p_i}^2(Y_i^2)}{k_{i;\gamma}^2} \}$
- (8) Compute $P_j = P(\chi_{N-l}^2 > (N - l) \max_{1 \leq i \leq l} \{ \frac{\chi_{1+p_i}^2(Y_i^2)}{k_{i;\gamma}^2} \} | Y)$
- (9) Repeat Steps 6–9, several times, say, 10,000 times
- (10) Set $f(\gamma) = \text{mean of } P_j\text{'s} - (1 - \alpha)$
- (11) Solve the equation $f(\gamma) = 0$ using a root bracketing method with bracketing interval $(1 - \alpha - .4, 1 - \alpha)$.
- (12) Let γ^* be the root obtained in the preceding step.
- (13) Then $k'_{i;\gamma^*} = \frac{1}{\sqrt{n_i}} t_{n_i-1;(1+\gamma^*)/2}(z_{1+p_i}/2 \sqrt{n_i})$, $i = 1, \dots, l$ are the factors for computing $(p_1, 1 - \alpha), \dots, (p_l, 1 - \alpha)$ simultaneous TIs.

In Table 2, we present the values of γ so that $t_{n_i-1;(1+\gamma)/2}(z_{(1+p)/2}\sqrt{n_i})/\sqrt{n_i}$ are factors for constructing (.90, .95) simultaneous TIs. The values of γ are given for $l = 2, \dots, 5$, unequal p_i 's and some sample sizes. As an example, to find (.90, .95) simultaneous TIs when $(n_1, n_2, n_3) = (12, 18, 16)$, the value of γ is 0.7012 and the factors are

$$k'_{i;\gamma} = \frac{1}{\sqrt{n_i}} t_{n_i-1; \frac{1.7012}{2}}(z_{\frac{1.9}{2}} \sqrt{n_i}) \quad \text{with } (k'_{1;\gamma}, k'_{2;\gamma}, k'_{3;\gamma}) = (2.277, 2.124, 2.163).$$

That is, $\bar{X}_i \pm k'_{i;\gamma} S_c$, $i = 1, 2, 3$, are (.90, .95) simultaneous tolerance intervals. For the same sample sizes, suppose we like to construct (.8, .95), (.9, .95) and (.95, .95) simultaneous TIs, then the value of γ from Table 2 is 0.7039 and

$$k'_{i;\gamma} = \frac{1}{\sqrt{n_i}} t_{n_i-1; \frac{1.7039}{2}}(z_{\frac{1+p_i}{2}} \sqrt{n_i}) \quad \text{with } (k'_{1;\gamma}, k'_{2;\gamma}, k'_{3;\gamma}) = (1.824, 2.124, 2.550).$$

2.3. Simultaneous equal-tailed tolerance intervals

The factors k_i^e for constructing l simultaneous equal-tailed TIs of the form $\bar{X}_i \pm k_i^e S_c$ are determined so that

$$P_{\bar{X}_{i,S}} \left(\bar{X}_i - k_i^e S_c < \mu_i - z_{\frac{1+p_i}{2}} \sigma \text{ and } \mu_i + z_{\frac{1+p_i}{2}} \sigma < \bar{X}_i + k_i^e S_c, i = 1, \dots, l \right) = 1 - \alpha. \tag{14}$$

That is, the intervals $\bar{X}_i \pm k_i^e S_c$ are constructed so that they would include $\mu_i \pm z_{(1+p_i)/2} \sigma$, $i = 1, \dots, l$, with probability $(1 - \alpha)$. After standardizing \bar{X}_i and rearranging the terms, we

Table 2. The values of γ so that $t_{n_i-1; (1+\gamma)/2}(Z_{(1+p_i)/2}\sqrt{n_i})/\sqrt{n_i}$ are factors for constructing $(p_1, .95), \dots, (p_l, .95)$ simultaneous tolerance intervals for l normal populations.

$l = 2$		$l = 3$		$p_1 = \dots = p_l = .90$		$l = 5$	
(n_1, n_2)	γ	(n_1, n_2, n_3)	γ	(n_1, \dots, n_4)	γ	(n_1, \dots, n_5)	γ
(4, 4)	.7754	(3, 3, 3)	.6844	(8, 8, 8, 8)	.6926	(6, 6, 6, 6, 6)	.6762
(7, 4)	.8102	(3, 5, 6)	.7410	(8, 4, 3, 6)	.7146	(6, 5, 3, 2, 8)	.7004
(7, 9)	.7820	(9, 8, 11)	.7219	(9, 11, 4, 21)	.7191	(9, 15, 13, 12, 21)	.6664
(12, 20)	.7711	(12, 18, 16)	.7012	(9, 7, 3, 5)	.7137	(9, 5, 33, 7, 18)	.7000
(22, 20)	.7395	(5, 19, 21)	.7211	(15, 25, 7, 21)	.6785	(13, 25, 12, 37, 48)	.6867
(32, 40)	.7223	(23, 45, 33)	.6758	(18, 21, 17, 31)	.6617	(20, 35, 45, 48, 50)	.5945
$(.8, .9)$		$(.8, .9, .95)$		$(.8, .85, .9, .95)$		$(.8, .85, .85, .9, .95)$	
(n_1, n_2)	γ	(n_1, n_2, n_3)	γ	(n_1, \dots, n_4)	γ	(n_1, \dots, n_5)	γ
(4, 4)	.7738	(3, 3, 3)	.6847	(8, 8, 8, 8)	.6902	(6, 6, 6, 6, 6)	.6757
(7, 4)	.7906	(3, 5, 6)	.7453	(8, 4, 3, 6)	.7019	(6, 5, 3, 2, 8)	.7078
(7, 9)	.7796	(9, 8, 11)	.7253	(9, 11, 4, 21)	.7343	(9, 15, 13, 12, 21)	.6726
(12, 20)	.7703	(12, 18, 16)	.7039	(9, 7, 3, 5)	.6953	(9, 5, 33, 7, 18)	.6882
(22, 20)	.7179	(5, 19, 21)	.7359	(15, 25, 7, 21)	.6726	(13, 25, 12, 37, 48)	.6589
(32, 40)	.7160	(23, 45, 33)	.6773	(18, 21, 17, 31)	.6710	(20, 35, 45, 48, 50)	.5972

see that (14) is equivalent to

$$P_{Z,S} \left(\frac{Z_i/\sqrt{n_i} + z_{\frac{1+p_i}{2}}}{S_c/\sigma} < k_i^e \text{ and } \frac{Z_i/\sqrt{n_i} - z_{\frac{1+p_i}{2}}}{S_c/\sigma} \geq -k_i^e, i = 1, \dots, l \right) = 1 - \alpha, \quad (15)$$

where $Z_i = \sqrt{n_i}(\bar{X}_i - \mu_i)/\sigma \sim N(0, 1)$. Let $\delta_i = \sqrt{n_i}z_{\frac{1+p_i}{2}}$ and $U^2 = S_c^2/\sigma^2$. In terms of these quantities, we see that (15) can be expressed as

$$P_{Z_i,U} (Z_i < -\delta_i + k_i^e\sqrt{n_i}U \text{ and } Z_i > \delta_i - k_i^e\sqrt{n_i}U, i = 1, \dots, l) = 1 - \alpha. \quad (16)$$

Notice that the inequalities in the above probability statement holds only if $\delta_i - k_i^e\sqrt{n_i}U < -\delta_i + k_i^e\sqrt{n_i}U$ or equivalently $U^2 > \frac{\delta_i^2}{(k_i^e)^2 n_i}$. Thus, (16) can be expressed as

$$E_U \left[P_Z \left(\delta_i - k_i^e\sqrt{n_i}U < Z_i < -\delta_i + k_i^e\sqrt{n_i}U \mid U^2 > \frac{\delta_i^2}{(k_i^e)^2 n_i} \right), i = 1, \dots, l \right] = 1 - \alpha. \quad (17)$$

Choosing

$$k_{i;\gamma}^e = \frac{1}{\sqrt{n_i}} t_{n_i-1; \frac{1+\gamma}{2}} \left(z_{\frac{1+p_i}{2}} \sqrt{n_i} \right), \quad i = 1, \dots, l,$$

and letting

$$G(\delta, N, l) = \max_{1 \leq i \leq l} \left\{ \frac{\delta_i^2}{(k_{i;\gamma}^e)^2 n_i} \right\},$$

we determine the value of γ so that

$$E_U \left[P_Z \left(\delta_i - k_{i;\gamma}^e\sqrt{n_i}U < Z_i < -\delta_i + k_{i;\gamma}^e\sqrt{n_i}U \mid U^2 > G(\delta, N, l) \right), i = 1, \dots, l \right] = 1 - \alpha, \quad (18)$$

where E_U denotes the expectation with respect to the distribution of U . Because $U^2 \sim \frac{\chi_{N-l}^2}{N-l}$, it follows from (18) that γ is the solution of the integral equation

$$f(\gamma) = \frac{1}{2^{\frac{N-l}{2}} \Gamma\left(\frac{N-l}{2}\right)} \int_{(N-l)G(\delta, N, l)}^{\infty} \prod_{i=1}^l \left(2\Phi\left(-\delta_i + \frac{k_{i;\gamma}^e \sqrt{n_i x}}{\sqrt{N-l}}\right) - 1 \right) e^{-x/2} x^{\frac{N-l}{2}-1} dx = 1 - \alpha, \tag{19}$$

where $\Phi(x)$ denotes the standard normal distribution function. To get (19) from (18), we have used the relation that $\Phi(x) = 1 - \Phi(-x)$. The value of γ that satisfies $f(\gamma) - (1 - \alpha) = 0$ can be found using a bisection method with root bracketing interval, say, $(1 - \alpha - .4, 1 - \alpha)$.

An R function (which uses the base R function `integrate()` to evaluate the integral in (19) and the base R function `uniroot()` to find the root of $f(\gamma) - (1 - \alpha) = 0$) to compute the factors satisfying (19) is given in the appendix. Using the R function, we have computed the value of γ so that $t_{n_i-1; (1+\gamma)/2} (z_{(1+p)/2} \sqrt{n_i}) / \sqrt{n_i}$ is a factor for computing an equal-tailed TI for the i th population. The values of γ for $l = 2, \dots, 5$ and some sample sizes are reported in Table 3. As an example, suppose it is desired to find factors for computing (.90, .95) equal-tailed TIs for three populations based on sample sizes $n_1 = 12$, $n_2 = 18$ and $n_3 = 16$. For this case, the value of gamma (see Table 3) is .8863 and the factors are

$$k_{i;\gamma}^e = \frac{1}{\sqrt{n_i}} t_{n_i-1; (1+0.8863)/2} (z_{.95} \sqrt{n_i}) \quad \text{with } (k_1^e, k_2^e, k_3^e) = (2.683, 2.416, 2.483).$$

For a given (n_1, \dots, n_l) and $(p, 1 - \alpha)$, the R function in the appendix computes both the value of γ and the factors. For the present case, we can compute the value of γ and the factors as follows:

```
> norm.simult.eqt.fac(c(12, 18, 16), .90, .95)
[1] 0.8863 2.683 2.416 2.483
```

Similarly, to find factors for constructing (.80, .95), (.90, .95) and (.95, .95) simultaneous equal-tailed TIs when $(n_1, n_2, n_3) = (12, 18, 16)$, we use the R function as

```
> norm.simult.eqt.fac(c(12, 18, 16), c(.80, .90, .95), .95)
[1] 0.8881 2.171 2.421 2.915
```

3. An example

This example and data are taken from Montgomery [13, Section 3.12]. A completely randomized experiment was conducted to assess the effective life of insulating fluids. The effective life (in hours) of four different insulating fluids at an accelerated load of 35 kV were recorded. To illustrate the construction of simultaneous TIs when the sample sizes are unequal, we randomly discarded two measurements from fluid type 1 and one measurement from fluid type 3, and present the remaining measurements in Table 4.

Table 3. The values of γ so that $t_{n_i-1;(1+\gamma)/2}(Z_{(1+p)/2}\sqrt{n_i})/\sqrt{n_i}$ are factors for constructing $(p_1, .95), \dots, (p_l, .95)$ simultaneous equal-tailed TIs for l normal populations.

(n_1, n_2)		(n_1, n_2, n_3)		$p_1 = \dots = p_l = .90$		(n_1, \dots, n_5)	
γ		γ		γ		γ	
(4, 4)	.8447	(3, 3, 3)	.7561	(8, 8, 8, 8)	.8418	(6, 6, 6, 6, 6)	.8072
(7, 4)	.8875	(3, 5, 6)	.8398	(8, 4, 3, 6)	.8381	(6, 5, 3, 2, 8)	.8286
(7, 9)	.8857	(9, 8, 11)	.8663	(9, 11, 4, 21)	.8833	(9, 15, 13, 12, 21)	.8714
(12, 20)	.9077	(12, 18, 16)	.8863	(9, 7, 3, 5)	.8454	(9, 5, 33, 7, 18)	.8906
(22, 20)	.9084	(5, 19, 21)	.8989	(15, 25, 7, 21)	.8879	(13, 25, 12, 37, 48)	.8977
(32, 40)	.9166	(23, 45, 33)	.9064	(18, 21, 17, 31)	.8905	(20, 35, 45, 48, 50)	.9030
$(.8, .9)$		$(.8, .9, .95)$		$(.8, .85, .9, .95)$		$(.8, .85, .85, .9, .95)$	
(n_1, n_2)	γ	(n_1, n_2, n_3)	γ	(n_1, \dots, n_4)	γ	(n_1, \dots, n_5)	γ
(4, 4)	.8545	(3, 3, 3)	.7628	(8, 8, 8, 8)	.8516	(6, 6, 6, 6, 6)	.8212
(7, 4)	.8967	(3, 5, 6)	.8309	(8, 4, 3, 6)	.8500	(6, 5, 3, 2, 8)	.8324
(7, 9)	.8926	(9, 8, 11)	.8692	(9, 11, 4, 21)	.8839	(9, 15, 13, 12, 21)	.8764
(12, 20)	.9135	(12, 18, 16)	.8881	(9, 7, 3, 5)	.8630	(9, 5, 33, 7, 18)	.8950
(22, 20)	.9166	(5, 19, 21)	.8925	(15, 25, 7, 21)	.8952	(13, 25, 12, 37, 48)	.8990
(32, 40)	.9241	(23, 45, 33)	.9084	(18, 21, 17, 31)	.8956	(20, 35, 45, 48, 50)	.9094

Table 4. Life (in hours) at 35 kV load.

Fluid type						
1	17.6	18.9	16.3	21.6		
2	16.9	15.3	18.6	17.1	19.5	20.3
3	21.4	23.6	19.4	18.5	20.5	
4	19.3	21.1	16.9	17.5	18.3	19.8

The summary statistics are

$$\bar{X}_1 = 18.60, \bar{X}_2 = 17.95, \bar{X}_3 = 20.68, \bar{X}_4 = 18.82 \text{ and } S_c = 1.8807.$$

Noticing that $(n_1, n_2, n_3, n_4) = (4, 6, 5, 6)$, the factors for constructing $(.90, .95)$ simultaneous one-sided tolerance limits are computed using the R function in the appendix as

```
> norm.simult.one.fac(c(4, 6, 5, 6), c(.90, .90, .90, .90), .95)
[1] 0.9004 3.1924 2.4962 2.7456 2.4962
```

That is, the adjusted confidence level $\gamma = 0.9004$, and the factors are

$$k_{i;\gamma} = \frac{1}{\sqrt{n_i}} t_{n_i-1;0.9004}(z_{p_i}\sqrt{n_i}) = (3.1924, 2.4962, 2.7456, 2.4962), \quad i = 1, \dots, 4.$$

The simultaneous lower tolerance limits $\bar{X}_i - k_{i;0.9004}S_c$ and the upper tolerance limits $\bar{X}_i + k_{i;0.9004}S_c$ are reported in Table 5.

The set of lower tolerance limits means that at least 90% fluid type 1 insulation have a minimum effective life of 12.60 h, at least 90% of fluid type 2 have a minimum effective life of 13.26 hours, at least 90% of fluid type 3 have a minimum effective life of 15.52 h and at least 90% of fluid type 4 have a minimum life of 14.12 h with confidence 95%.

The factors for computing $(.90, .95)$ simultaneous TIs when $(n_1, n_2, n_3, n_4) = (4, 6, 5, 6)$ are obtained using the R function in the appendix as

Table 5. Simultaneous (.90, .95) one-sided, two-sided and equal-tailed tolerance intervals for life hours of insulating fluids.

Fluid type	Lower TL	Upper TL	Two-sided TI	Equal-tailed TI
1	12.60	24.60	(12.35, 24.85)	(10.97, 26.23)
2	13.26	22.64	(12.81, 23.09)	(12.03, 23.87)
3	15.52	25.84	(15.13, 26.22)	(14.15, 27.21)
4	14.12	23.51	(13.68, 23.96)	(12.90, 24.73)

```
> norm.simult.two.fac(10^5, c(4, 6, 5, 6), c(.9, .9, .9, .9), .95)
[1] 0.6928 3.325 2.733 2.948 2.733
```

That is, the adjusted confidence level $\gamma = 0.6928$ and the factors are 3.325, 2.733, 2.948 and 2.733 for fluid types 1, 2, 3, and 4, respectively. The TIs $\bar{X}_i \pm k'_{i,0.6928} S_c$ are computed and presented in Table 5. This means that, with confidence 95%, at least 90% of type 1 insulating fluid have life between 12.34 and 24.86 h, at least 90% of type 2 insulating fluid have life between 81.95 and 95.13 h, at least 90% of type 3 insulating fluid have life between 15.13 and 26.22 h and at least 90% of type 4 insulating fluid have life between 13.67 and 23.96 h.

To compute (.90, .95) equal-tailed TIs, we calculated the required factors as

```
> norm.simult.eqt.fac(c(4, 6, 5, 6), c(.90, .90, .90, .90), .95)
[1] 0.8123 4.0563 3.1464 3.4695 3.1464
```

The equal-tailed TIs $(L_{ei}, U_{ei}) = \bar{X}_i \pm k^e_{i,0.8123} S_c$ are computed and reported in Table 5. The equal-tailed TIs can be interpreted as follows. We can conclude with 95% confidence that no more than 5% of fluid type i fails within L_{ei} hours and no more than 5% of fluid type 1 lasts more than U_{ei} hours, $i = 1, \dots, 4$.

4. Concluding remarks

Simultaneous statistical intervals such as simultaneous confidence intervals are often constructed from one-sample interval by adjusting the confidence coefficients using the Bonferroni approach or by Scheffé’s method. Mee [12] has developed exact numerical methods of computing simultaneous one-sided tolerance limits and simultaneous two-sided tolerance intervals. These simultaneous TIs are exact in the sense that the coverage probabilities are equal to the nominal confidence level for all parameter values. These methods, however, are applicable only when the content levels of the TIs are the same and the sample sizes are the same. Using the idea of Hoang-Thuy-Ngan and Krishanmoorthy (2020), we proposed exact numerical methods for computing $(p_1, 1 - \alpha), \dots, (p_l, 1 - \alpha)$ simultaneous one-sided tolerance limits, simultaneous TIs and also simultaneous equal-tailed TIs. For given sample sizes, content levels and confidence coefficient, the factors determined by our methods are unique. Our methods are applicable even when sample sizes are unequal. To help practitioners to find tolerance factors, we provided R code in the appendix which are simple to use.

Acknowledgments

The authors are grateful to a reviewer and an associate editor for providing useful comments and suggestions.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

- [1] Guttman I. Statistical tolerance regions: classical and Bayesian. London: Charles Griffin and Company; 1970.
- [2] Krishnamoorthy K, Mathew T. Statistical tolerance regions: theory, applications and computation. Hoboken (NJ): John Wiley & Sons; 2009.
- [3] Wilks SS. Determination of sample sizes for setting tolerance limits. *Ann Math Statist.* 1941;12:91–96.
- [4] Wilks SS. Statistical prediction with special reference to the problem of tolerance limits. *Ann Math Statist.* 1942;13:400–409.
- [5] Wald A. An extension of Wilks method for setting tolerance limits. *Ann Math Statist.* 1943;14:45–55.
- [6] Wald A, Wolfowitz J. Tolerance limits for a normal distribution. *Ann Math Statist.* 1946;17:208–215.
- [7] Krishnamoorthy K, Xie F. Tolerance intervals for symmetric location-scale distributions based on censored or uncensored data. *J Statist Plann Inference.* 2011;141:1170–1182.
- [8] Zou Y, Young DS. Improving coverage probabilities for parametric tolerance intervals via bootstrap calibration. *Stat Med.* 2020;39:2152–2166. DOI:10.1002/sim.8537
- [9] Hoang-Nguyen-Thuy N, Krishnamoorthy K. A method for computing tolerance intervals for a location-scale family of distributions. *Comput Statist.* 2021;36:1065–1092.
- [10] Young DS. Tolerance: statistical tolerance intervals and regions. R Package, Version 0.1.0; 2009.
- [11] Krishnamoorthy K. Handbook of statistical distributions with applications. 2nd ed. Boca Raton (FL): Chapman & Hall/CRC Press; 2015.
- [12] Mee RW. Simultaneous tolerance intervals for normal populations with common variance. *Technometrics.* 1990;32:83–92.
- [13] Montgomery DC. Design and analysis of experiments. 8th ed. Hoboken (NJ): Wiley; 2013.

Appendix

R code to compute factors for simultaneous one-sided tolerance limits

```
# n = vector of sample sizes; p = vector of content levels
# cl = confidence level
norm.simult.one.fac = function(n, p, cl){

  l = length(n); m = n-1; M = sum(m); zp = qnorm(p)
  fn = function(gam){
    k = qt(gam, m, zp*sqrt(n))/sqrt(n)
    f = function(x){
      prod = 1
      for(j in 1:l){
        prod = prod*pnorm(sqrt(n[j])*(k[j]*sqrt(x)/sqrt(M)-zp[j]))
      }
      y = prod*dchisq(x, M)
    }
  }
}
```

```

    return(y)}
  LowLim = 0; UppLim = 100*sqrt(2*M)
  g = integrate(f, LowLim, UppLim)[[1]]-c1
  return(g)
}
R = .9999; L = c1-.45
gams = uniroot(fn, c(L,R))[[1]]
print(c(gams,qt(gams, m, zp*sqrt(n))/sqrt(n)),4)
}
> norm.simult.one.fac(c(12,18,16),c(.90,.90,.90),.95)
[1] 0.9348 2.1171 1.9080 1.9606

```

R code to compute factors for simultaneous two-sided tolerance intervals

```

norm.simult.two.fac = function(nr, n, p, c1){
  set.seed(129)
  if(length(unique(n)) == 1 & length(unique(p)) == 1){
    l = length(n); n = n[1]; p = p[1]
    m = n-1; M = l*m; zp = qnorm((1+p)/2)
    del = sqrt(n)*zp
    fn = function(gam){
      k = qt((1+gam)/2, m, del)/sqrt(n)
      f = function(x){
        y = 2*l*(2*pnorm(x)-1)^(l-1)*dnorm(x)
        *(1-pchisq(M*qchisq(p,1,x^2/n)/k^2,M)) return(y)}
      LowLim = 0; UppLim = 50;
      g = integrate(f, LowLim, UppLim)[[1]]-c1
      return(g)
    }
  }
  else{
    l = length(n); m = n-1; M = sum(m); zp = qnorm((1+p)/2)
    y = matrix(rnorm(nr*l), nr, l); chik = y; chis = chik
    del = zp*sqrt(n)
    for(j in 1:l){
      y[,j] = y[,j]^2/n[j]}
    for(j in 1:nr){
      chis[j,] = qchisq(p, 1, y[j,])}
    fn = function(x){
      k = qt((1+x)/2, m, del)/sqrt(n)
      for(j in 1:l){
        chik[,j] = chis[,j]/k[j]^2}
      yx = apply(chik, 1, function(x) max(x))
      ps = 1-pchisq(M*yx, M)
      return(mean(ps)-c1)}
    }
  x0 = uniroot(fn, c(c1-.4,c1))[[1]]
  print(c(x0, qt((1+x0)/2, m, del)/sqrt(n)))
}
norm.simult.two.fac(10^5, c(12,18,16), c(.8,.9,.95), .95)
[1] 0.7038 1.824 2.127 2.550

```

R code to compute factors for constructing simultaneous equal-tailed tolerance intervals

```

norm.simult.eqt.fac = function(n, p, cl){
  l = length(n); m = n-1; M = sum(m); zp = qnorm((1+p)/2)
  del = sqrt(n)*zp
  fn = function(gam){
    k = qt((1+gam)/2, m, del)/sqrt(n)
    f = function(x){
      prod = 1
      for(j in 1:l){
        prod = prod*(2*pnorm(-del[j]+k[j]*sqrt(n[j]*x)/sqrt(M))-1)
      }
      y = prod*dchisq(x, M)
      return(y)}
    LowLim = M*max(del^2/k^2/n); Upplim = 50*sqrt(2*M);

    g = integrate(f, LowLim, Upplim)[[1]]-cl
    return(g)
  }
  gam = uniroot(fn, c(cl-.4,cl))[[1]]
  return(c(gam,qt((1+gam)/2, m, del)/sqrt(n)))
}
> norm.simult.eqt.fac(c(12,18,16),c(.80,.90,.95),.95)
[1] 0.8881 2.171 2.420 2.915

```