One-sided Tolerance Factors of Normal Distributions with unknown mean and variability

I. Janiga, I. Garaj

Dept. of Mathematics, Mechanical Engineering Faculty, Slovak University of Technology
Nám. slobody 17, 812 31 Bratislava, Tel: 00421-7-57296-160, ivan.janiga@stuba.sk
Dept. of Mathematics, Faculty of Chemical and Food Technology, Slovak University of Technology
Radlinského 9, 812 37 Bratislava, Tel: 00421-7-59325 297, ivan.garaj@stuba.sk

Abstract. *In the contribution are given some approximate formulas as well as a derivation of exact formula for computing of the one-sided tolerance factors of normal distribution with unknown mean and unknown variability. It is also shown that the exact computation of the factors is closely related with non-central t-distributions.*

Keywords: tolerance intervals, mean, variability

1. Introduction

In [1] we deal with computation of tolerance factors for the two-sided tolerance limits of a normal distribution with unknown parameters. Results are published in [2]. In [3] are given the results for more than one normal distribution with different means and common variability. All the above mentioned parameters are unknown. In this contribution we are interested in one-sided tolerance limits of a normal distribution with unknown parameters.

2. One-sided tolerance intervals

Let random sample $X_1, X_2, \dots, X_n$ be taken from distribution $N(\mu, \sigma^2)$. Parameters μ and σ^2 are unknown.

We will find intervals, which with confidence $1-\alpha$ $(0<\alpha<1)$ cover at least the fraction p $(0<p<1)$ of values of the distribution $N(\mu, \sigma^2)$. Such intervals are called $100\,p\,\%$ tolerance intervals. In the following text we are only dealing with one-sided tolerance intervals which can be right-hand or left-hand.

For right-hand tolerance interval $\left(-\infty, \overline{X}+kS\right)$ is valid

$$P\left[P(X<\overline{X}+kS)\geq p\right]=1-\alpha \tag{1}$$

and

for left-hand tolerance interval $(\overline{X}-kS, \infty)$ is valid

$$P\left[P\left(X>\overline{X}-kS\right)\geq p\right]=1-\alpha \tag{2}$$

where X is a normally distributed variable with mean μ and variability σ^2 from which the random sample $X_1, X_2, \dots, X_n$ was taken, $\overline{X}=n^{-1}\sum_{i=1}^{n}X_i$ and $S^2=(n-1)^{-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$.

The value of the factor k is determined so that the tolerance intervals with the confidence $1-\alpha$ cover at least fraction p of the values of the distribution $N(\mu,\sigma^2)$. Number $1-\alpha$ named the confidence level and constant k is the one-sided tolerance factor.

3. Non-central *t*-distribution

Let us consider random variables

$$Y = \frac{\overline{X} - \mu}{\sigma}\sqrt{n}\,, \qquad V = \frac{\nu S^2}{\sigma^2}\,, \qquad T = \frac{Y+\delta}{\sqrt{V}}\sqrt{\nu} \tag{3}$$

The random variable Y has a standard normal distribution $N(0,1)$. It is independent of the random variable V, which has χ^2-distribution with $\nu = n-1$ degrees of freedom. The random variable T has a non-central *t*-distribution with ν degrees of freedom and non-centrality parameter δ $(-\infty < \delta < \infty)$, which is denoted by $T \sim T(\nu, \delta)$. Its probability density function [4] is

$$f_\nu(t,\delta) = \frac{2^{-\frac{\nu-1}{2}}\, \mathrm{e}^{-\frac{\nu\delta^2}{2(\nu+t^2)}}}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\nu}\left(1+\frac{t^2}{\nu}\right)^{\frac{\nu+1}{2}}} \int_0^\infty y^\nu \mathrm{e}^{-\frac{1}{2}\left(y-\frac{t\delta}{\sqrt{\nu+t^2}}\right)^2} dy \qquad -\infty < t < \infty \tag{4}$$

where $\Gamma(\nu) = \int_0^\infty x^{\nu-1}\,\mathrm{e}^{-x}\,dx$ is the gamma function $(\nu > 0)$.

Let us denote the α^{th} quantile $(0 < \alpha < 1)$ of the distribution $T(\nu, \delta)$ by $t_\alpha(\nu, \delta)$. It follows from (4) that for any real numbers t and δ holds

$$t_\alpha(\nu,\delta) = -t_{1-\alpha}(\nu,-\delta) \tag{5}$$

The most widespread tables of percentiles of non-central *t*-distributions can be found in the monograph [5], where they are computed with accuracy to five decimal places for $\alpha =$ 0,01; 0,025; 0,05; 0,10; 0,20; 0,30; 0,70; 0,80; 0,90; 0,95; 0,975; 0,99, $\nu = 1\,(1)\,60$ and $\delta = 0{,}1\,(0{,}1)\,8{,}0$.

4. Approximate computation of one-sided tolerance factors

There exist a great number of approximations of the one-sided tolerance factors [7] and [8]. The Wallis´s approximation [9] is the best known and up till now the most used. It was derived on the basis of the approximation of the statistic $\overline{X} + kS$ by the normal distribution $N\left(\mu + k\sigma, \frac{\sigma^2}{n} + \frac{\sigma^2 k^2}{2n-2}\right)$. The one-sided tolerance factor k is given by the relationship

$$k \approx \frac{u_p + \sqrt{u_p^2 - AB}}{A} \tag{6}$$

where $A = 1 - \frac{u_{1-\alpha}^2}{2(n-1)}$, $B = u_p^2 - \frac{u_{1-\alpha}^2}{n}$ and u_p , $u_{1-\alpha}$ are quantiles of $N(0,1)$. Slightly better results are given by Jennett´s and Welch´s approximation [10] by means of the quantiles of the non-central *t*-distribution

$$t_\alpha(v, \delta) \approx \frac{\delta b_v + u_\alpha \sqrt{b_v^2 + \left(1 - b_v^2\right)\left(\delta^2 - u_\alpha^2\right)}}{b_v^2 - u_\alpha^2\left(1 - b_v^2\right)} \tag{7}$$

where $b_v = \left(\Gamma((\nu+1)/2)/\Gamma(\nu/2)\right) \times \sqrt{2/\nu}$. For a small δ is convenient the approximation by van Eeden [8], which was derived by means of Cornish-Fisher expansion [11] $(v = n-1)$

$$t_\alpha(v,\delta) \approx u_\alpha + \frac{u_\alpha^3 + u_\alpha}{4v} + \frac{5u_\alpha^5 + 16u_\alpha^3 + 3u_\alpha}{96v^2} +$$

$$+\delta\left(1 + \frac{2u_\alpha^2 + 1}{4v} + \frac{u_\alpha}{4v}\delta + \frac{4u_\alpha^4 + 12u_\alpha^2 + 1}{32v^2} + \frac{u_\alpha^3 + 4u_\alpha}{16v^2}\delta - \frac{u_\alpha^2 - 1}{24v^2}\delta^2 - \frac{u_\alpha}{32v^2}\delta^3\right) \tag{8}$$

Neither of the above mentioned approximations can be considered to give good results in general. This non-availability does not contain Akahira´s approximation [12]. The quantiles of the non-central *t*-distribution $x = t_\alpha(v,\delta)$ can be found by solving the following equation

$$\frac{b_v x - \delta}{\sqrt{1 + x^2\left(1 - b_v^2\right)}} = u_\alpha - \frac{x^3\left(u_\alpha^2 - 1\right)}{24\sqrt{\left[1 + x^2\left(1 - b_v^2\right)\right]^3}}\left(\frac{1}{v^2} + \frac{1}{4v^3}\right) \qquad (v = n-1) \tag{9}$$

where b_v is the same as given in (7) and u_α is the α^{th} quantile of the standard normal distribution. The approximation (9) yields good results even if v is small. When $v \geq 200$ there are many cases where the error is less than 10^{-4}. There is a disadvantage that the equation (9) can be solved for unknown x only by using numerical methods.

5. Exact computation of tolerance factors

For the right-sided tolerance interval $\left(-\infty, \overline{X} + kS\right)$ is

$$Z = P\left(X < \overline{X} + kS\right) = \Phi\left(\frac{\overline{X} + kS - \mu}{\sigma}\right) \tag{10}$$

where Φ is the distribution function of a standard normal distribution $N(0,1)$. For given p $(0 < p < 1)$ the inequality $Z \geq p$ is equivalent to

$$\frac{\overline{X} + kS - \mu}{\sigma} \geq u_p \tag{11}$$

where u_p is the p^{th} quantile of $N(0,1)$. The tolerance factor k is determined so that for the given confidence coefficient $1-\alpha$ is fulfilled the condition $P(Z \geq p) = 1-\alpha$. By using consecutive modifications and utilizing (3) we obtain

$$P\left(\overline{X} + kS \geq \mu + \sigma u_p\right) = P\left(\frac{(\overline{X} - \mu)\sqrt{n}/\sigma - u_p\sqrt{n}}{S/\sigma} \geq -k\sqrt{n}\right) =$$

$$= P\left(T \geq -k\sqrt{n}\right) = 1 - \alpha \tag{12}$$

The random value $T = \frac{\overline{X} - \mu - \sigma u_p}{S}\sqrt{n}$ has a distribution $T\left(n-1, -u_p\sqrt{n}\right)$. From relations (5) and (12) it follows that

$$k = -\frac{t_\alpha\left(n-1, -u_p\sqrt{n}\right)}{\sqrt{n}} = \frac{t_{1-\alpha}\left(n-1, u_p\sqrt{n}\right)}{\sqrt{n}} \tag{13}$$

For the left-sided tolerance interval $\left(\overline{X} - kS, \infty\right)$ we will also get $P\left(T \leq k\sqrt{n}\right) = 1 - \alpha$, where T has a distribution $T\left(n-1, u_p\sqrt{n}\right)$ and k is given by (13) as well.

6. Conclusion

The exact computation of the one-sided tolerance factors k is closely related with the exact computation of quantiles of non-central *t*-distribution. One-sided tolerance intervals can be successfully applied in statistical quality control. For those who are interested in the mentioned field we recommend [13], [14].

Acknowledgement

This contribution was supported by the Grant VEGA project No. 1/1247/04 Advanced statistical techniques and decision making in the quality improvement process from the Science Grant Agency of the Slovak Republic.

References

[1] JANIGA, I., MIKLÓŠ, R. Statistical tolerance intervals for a normal distribution. In *Measurement Science Review*. ISSN 13, 2001, vol. 1, no. 1, p. 29-32.

[2] GARAJ, I., JANIGA, I. *Dvojstranné tolerančné medze pre neznámu strednú hodnotu a rozptyl normálneho rozdelenia*. Bratislava: Vydavateľstvo STU, 2002. 147 s. ISBN 80-227-1779-7. [Two sided tolerance limits of normal distribution with mean and variability unknown].

[3] GARAJ, I., JANIGA, I. *Dvojstranné tolerančné medze normálnych rozdelení s neznámymi strednými hodnotami a s neznámym spoločným rozptylom. Two sided tolerance limits of normal distributions with unknown means and unknown common variability.* Bratislava: Vydavateľstvo STU, 2004. 218 s. ISBN 80-227-2019-4.

[4] JOHNSON, N. L., WELCH, B. L. Applications of the non-central t-distribution. In *Biometrika*, 1940, vol. 31, p. 362-389.

[5] BAGUI, S. C. *CRC Handbook of Percentiles of Non-Central t-Distributions*. CRC Press, Boca Raton, Florida, 1993. 400 p. ISBN 0-8493-8669-1.

[6] WOLFRAM, S. *The Mathematica Book*. 3rd ed. Wolfram Media/Cambridge University Press, 1996. 1403 p. ISBN 0-521-58889-8.

[7] SAHAI, H., OJEDA, M. M. A comparison of approximations to percentiles of the non-central t-distribution. In *Revista Investigacion Operacional*, 2000, vol. 21, no. 2, p.121-144.

[8] EEDEN, van C. Some approximations to the percentage points of the non-central t-distribution. In *Revue de l′ Institut International de Statistique* 29, 1961, p. 4-31.
[9] WALLIS, W. A. *Use of Variables in Acceptance Inspection for Percent Defective. Selected Techniques of Statistical Analysis* (EISENHART, C., HASTAY, M. W., WALLIS, W. A. eds.). New York, McGraw Hill Book Company, 1947. p. 7-93.
[10] JENNETT, W. T., WELCH, B. L. The control of proportion defective as judged by a single quality characteristic varying on a continuous scale. In *Journal of the Royal Statistical Society*, B 6, 1939, p. 80-88.
[11] CORNISH, E. A., FISHER, R. A. Moments and cumulants in the specification of distributions. In *Revue de l′ Institut International de Statistique*, 5, 1937, p. 307-320.
[12] AKAHIRA, M. A higher order approximation to a percentage point of the non-central t-distribution. In *Communications in Statistics*, *Part B: Simulation and Computation*, 1995, 24(3), p. 595-605.
[13] TEREK, M., HRNČIAROVÁ, Ľ *Štatistické riadenie kvality*. Bratislava: Vyd. IURA EDITION, 2004. 234 s. ISBN 80-89047-97-1.
[14] TEREK, M., HRNČIAROVÁ, Ľ. *Analýza spôsobilosti procesu*. Vydavateľstvo EKONÓM, Ekonomická univerzita v Bratislave, 2001. 205 s. ISBN 80-225-1443-8.