Two-sided Tolerance Limits of Normal Distributions with Unknown Means and Unknown Common Variability

[1]Ivan Janiga, [2]Ivan Garaj

[1]Dept. of Mathematics, Mechanical Engineering Faculty, Slovak University of Technology
Námestie slobody 17, 812 31 Bratislava, Tel: 00421-7-57296-160, -350
janiga@sjf.stuba.sk
[2]Dept. of Mathematics, Faculty of Chemical and Food Technology, Slovak University of Technology, Radlinského 9, 812 37 Bratislava, Tel: 00421-7-59325 297
igaraj@chelin.chtf.stuba.sk

Abstract. *In the contribution an exact computing of the two-sided tolerance limits of normal distributions with unknown means* μ_i *and an unknown common variability* σ^2 *are introduced. The numerical output as well as an example of its usage is given.*

1 Introduction

In [1], [2] we deal with computation of tolerance factors for the two-sided tolerance limits of a normal distribution with unknown mean μ and unknown variability σ^2. Results are published in [3].

In this contribution we are interested in tolerance intervals for $m > 1$ normal distributions with different means μ_i and a common variability σ^2.

2 Tolerance intervals for m distributions with common variability

Let measurements $(x_{i1}, x_{i2}, \cdots, x_{in})$ be values of m random samples of the same sizes n drawn from m populations. We assume that values of measured characteristics X_i are normally distributed with different means μ_i and a common variability σ^2, that is $X_i \sim N(\mu_i, \sigma^2)$, $i = 1, 2, ..., m$. The parameters μ_i and σ^2 are supposed to be unknown.

We are looking for intervals, which with confidence $1-\alpha$ ($0 < \alpha < 1$) cover at least the fraction p ($0 < p < 1$) of values of the distributions $N(\mu_i, \sigma^2)$. These intervals are named $100\,p$ % tolerance intervals. The values of unbiased estimators of unknown parameters μ_i and σ^2 are computed by the formulas

$$\bar{x}_i = \frac{1}{n}\sum_{j=1}^{n} x_{ij} \qquad \text{and} \qquad s_P^2 = \frac{1}{m(n-1)}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - \bar{x}_i)^2 \tag{1}$$

where $\bar{x}_i$ is the sample mean of the *i*th random sample and s_P^2 is the estimator of the common variability (sometimes called “pooled”).

Two-sided tolerance intervals for distributions $N(\mu_i, \sigma^2)$ are

$$[\bar{x}_i - k\, s_P, \bar{x}_i + k\, s_P], \quad i = 1, 2, ..., m \tag{2}$$

for which is valid

$$P\left[P\left(\bar{x}_i - k\,s_P < X_i < \bar{x}_i + k\,s_P\right) \geq p\right] = 1-\alpha, \quad i = 1,2,\ldots,m \qquad (3)$$

The value of tolerance factor k is determined such that the tolerance intervals with the confidence $1-\alpha$ cover at least fraction p of the values of the distributions $N(\mu_i, \sigma^2)$. Number $1-\alpha$ is named confidence level.

3 Exact computation of tolerance factors k

In [3] are tables of values of exact computing tolerance factors. In the software Mathematica (see [5]) for given n, $\nu = n-1$, α and p there is given numerical solution of the integral equation (see [2], [3])

$$\sqrt{\frac{n}{2\pi}} \int_{-\infty}^{\infty} F(x,k)\, e^{-\frac{nx^2}{2}}\, \mathrm{d}x - 1 + \alpha = 0 \qquad (4)$$

where

$$F(x,k) = \int_{\frac{\nu R^2(x)}{k^2}}^{\infty} \frac{t^{\frac{\nu}{2}-1}\, e^{-\frac{t}{2}}}{2^{\frac{\nu}{2}}\, \Gamma\left(\frac{\nu}{2}\right)}\, \mathrm{d}t$$

and $R(x)$ is the solution of the equation $\Phi(x+R) - \Phi(x-R) - p = 0$.

To compute the tolerance factors $k = k(n, \nu, p, 1-\alpha)$ in (3) we use equation (4) as well. We solved this equation numerically for $\nu \neq n-1$. Tolerance factors k are computed for $1-\alpha = 0{,}95$, $\nu = 48$, $p = 0{,}50$; 0,75; 0,90; 0,925; 0,95; 0,975; 0,99; 0,995; 0,999 and $n = 2(1)$ 20; 22(2) 30; 35(5) 50; 60(10) 100; 150(50) 500; 600(100) 1000; 5000; 10000 and ∞. A part of the results is given in Table 2.

Example. Three machines are producing components. It is assumed that the output data from the machines are continuously distributed with common variance but possible different means. Random samples of 17 components are selected from each machine´s output. The results of the samples are in Table 1. We are computing 99 % two-sided tolerance interval for the measurements of the first, second and third machine, when $1-\alpha = 0{,}95$.

<u>Solution</u>. A good fit with normal distribution with unknown parameters μ and σ^2 was confirmed by W-test (*P* value of the first machine was 0,076, the second 0,870 and the third machine 0,979). The homogeneity of variances was assumed therefore we can use all 52 measurements to estimate common variability ($s_P^2 = 2{,}1498$ and $s_P = 1{,}46622$). Mean´s values are different ($H_0 : \mu_1 = \mu_2 = \mu_3$ is rejected), and so we can use only 17 measurements of individual machine´s to estimate individual means. For $n = 17$ and $\nu = 3(17-1) = 48$ we can

find in Table 2 value of $k = 3{,}2077$. Hence 99 % two-sided tolerance intervals with confidence 0,95 are as follows

1. machine: $48{,}8294 \pm 3{,}2077\sqrt{2{,}1498}$, that is (44,13; 53,53),
2. machine: $54{,}0000 \pm 3{,}2077\sqrt{2{,}1498}$, that is (49,30; 58,70),
3. machine: $59{,}6706 \pm 3{,}2077\sqrt{2{,}1498}$, that is (54,97; 64,37).

Table 1

j \ i	Machine 1	Machine 2	Machine 3
1	50,5	56,2	59,0
2	49,9	57,3	59,7
3	48,4	54,5	58,2
4	46,1	53,6	56,7
5	50,1	51,7	59,6
6	50,2	55,5	60,2
7	49,1	54,1	61,5
8	48,1	53,0	62,5
9	50,3	52,1	61,3
10	46,6	51,3	58,5
11	49,2	54,6	60,6
12	46,5	55,1	61,3
13	48,0	53,2	57,7
14	49,1	53,8	59,5
15	49,6	53,9	59,6
16	49,6	54,1	59,5
17	48,8	54,0	59,0
$\sum_{j=1}^{17}$	830,1	918,0	1014,4
$\bar{x}_i$	48,8294	54,0000	59,6706
s_i^2	1,89471	2,36625	2,18846

Acknowledgement

This contribution was supported by the Grant VEGA project No. 1/8094/01 *Statistical measurement process control* from the Science Grant Agency of the Slovak Republic.

4 Conclusion

A part from the set of computed tolerance factors is given in Table 2. The rest of results will be presented on the lecture.

Table 2

	1 - α = 0,95 p = 0,99									
n \ ν	5	10	15	20	30	48	100	500	1000	∞
2	6,2983	4,8934	4,4761	4,2746	4,0785	3,9364	3,8182	3,7333	3,7228	3,7123
3	6,0495	4,6711	4,2545	4,0500	3,8468	3,6962	3,5698	3,4802	3,4691	3,4581
4	5,9106	4,5483	4,1324	3,9261	3,7182	3,5606	3,4254	3,3301	3,3184	3,3067
5	5,8209	4,4697	4,0546	3,8472	3,6363	3,4735	3,3301	3,2284	3,2161	3,2038
6	5,7578	4,4149	4,0005	3,7926	3,5796	3,4131	3,2627	3,1542	3,1411	3,1282
7	5,7108	4,3744	3,9608	3,7525	3,5381	3,3688	3,2127	3,0972	3,0835	3,0698
8	5,6745	4,3433	3,9303	3,7218	3,5064	3,3350	3,1744	3,0521	3,0376	3,0232
9	5,6455	4,3186	3,9062	3,6976	3,4814	3,3085	3,1442	3,0153	3,0000	2,9849
10	5,6217	4,2985	3,8866	3,6780	3,4613	3,2872	3,1199	2,9849	2,9688	2,9529
11	5,6020	4,2818	3,8704	3,6618	3,4447	3,2696	3,1000	2,9593	2,9423	2,9256
12	5,5852	4,2678	3,8569	3,6482	3,4308	3,2550	3,0834	2,9374	2,9196	2,9021
13	5,5708	4,2557	3,8453	3,6367	3,4190	3,2426	3,0694	2,9187	2,8999	2,8817
14	5,5584	4,2454	3,8353	3,6268	3,4089	3,2320	3,0574	2,9024	2,8827	2,8637
15	5,5475	4,2363	3,8266	3,6181	3,4002	3,2228	3,0471	2,8882	2,8675	2,8477
16	5,5378	4,2283	3,8189	3,6105	3,3925	3,2148	3,0381	2,8757	2,8541	2,8334
17	5,5292	4,2212	3,8122	3,6038	3,3857	**3,2077**	3,0302	2,8647	2,8421	2,8206
18	5,5216	4,2149	3,8061	3,5979	3,3797	3,2014	3,0232	2,8549	2,8313	2,8090
19	5,5146	4,2092	3,8007	3,5925	3,3743	3,1958	3,0170	2,8462	2,8217	2,7984
20	5,5084	4,2041	3,7958	3,5877	3,3695	3,1908	3,0114	2,8384	2,8129	2,7888
30	5,4678	4,1712	3,7647	3,5572	3,3390	3,1595	2,9770	2,7907	2,7578	2,7246
50	5,4342	4,1444	3,7397	3,5329	3,3151	3,1354	2,9513	2,7564	2,7176	2,6688
100	5,4082	4,1240	3,7209	3,5148	3,2976	3,1182	2,9336	2,7345	2,6925	2,6239
500	5,3870	4,1076	3,7060	3,5006	3,2841	3,1051	2,9209	2,7206	2,6774	2,5861
1000	5,3843	4,1056	3,7042	3,4988	3,2825	3,1036	2,9194	2,7192	2,6759	2,5818
∞	5,3816	4,1035	3,7023	3,4971	3,2808	3,1020	2,9179	2,7178	2,6745	2,5795

References

[1] Janiga, I., Miklóš, R.: O jednom probléme výpočtu tolerančných činiteľov. In *9. medzinárodný seminár Výpočtová štatistika*. SŠDS 2000. ISBN 80-88946-09-3, s. 30-32.

[2] Janiga, I., Miklóš, R.: Statistical Tolerance Intervals for a Normal Distribution. In *Proceedings of 3rd International Conference on Measurement*. IMS SAS, Smolenice, SR, 2001. ISBN 80-967402-5-3, p. 28-31.

[3] Garaj, I., Janiga, I.: *Dvojstranné tolerančné medze pre neznámu strednú hodnotu a rozptyl normálneho rozdelenia*. Bratislava:Vydavateľstvo STU, 2002, s.147. ISBN 80-227-1779-7.

[4] Aitchison, J., Dunsmore, I.R.: *Statistical Prediction Analysis*, Cambridge: Cambridge University Press, 1975.

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

[6] Garaj, I.: W-kritérium a jeho spracovanie na počítači. In *9. medzinárodný seminár Výpočtová štatistika*. SŠDS 2000. ISBN 80-88946-09-3, s. 30-32.