Josef Machek
Normal tolerance limits base on mean range

Acta Universitatis Carolinae. Mathematica et Physica, Vol. 5 (1964), No. 2, 21--32

1964 ACTA UNIVERSITATIS CAROLINAE—MATHEMATICA ET PHYSICA 2 PAG. 21–32

NORMAL TOLERANCE LIMITS BASED ON MEAN RANGE

JOSEF MACHEK

Department of Mathematical Statistics, Charles University, Prague

Summary. Consider a normal population with mean μ and standard deviation σ. Let $\overline{R}$ be the average of ranges of k mutually independent samples of size n from this population. Further let $\hat{\mu}$ be an estimate of μ, independent of $\overline{R}$ and normally distributed about μ with variance σ^2/N. In this paper we present tables of auxiliary quantities for the computation of coefficients T_1 and T_2 which posess the following properties:

(1) the expectation of the proportion of the underlying population covered by the interval $\hat{\mu} - T_1\overline{R}$, $\hat{\mu} + T_1\overline{R}$ is approximately equal to a predetermined value P, i. e.

$$\mathsf{E}\left\{ \int_{\hat{\mu}-T_1\overline{R}}^{\hat{\mu}+T_1\overline{R}} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \mathrm{d}x \right\} = P;$$

(2) the probability that the interval $\hat{\mu} - T_2\overline{R}$, $\hat{\mu} + T_2\overline{R}$ covers at least the predetermined proportion P of the population is approximately equal to a given confidence coefficient γ, i. e.

$$\mathsf{P}\left\{ \int_{\hat{\mu}-T_2\overline{R}}^{\hat{\mu}+T_2\overline{R}} (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \mathrm{d}x \geqq P \right\} \doteq \gamma$$

Introduction. The following problem is often encountered in statistics, especially in industrial applications: to estimate on the basis of a random sample an interval which covers a given proportion of the population with some prescribed degree of assurance. This degree of assurance refers to the performance of the estimation procedure in the long run. Most frequently it is expressed either by the requirement that the relative frequency of intervals covering less than the desired proportion of the population shall not exceed a given small quantity, or by the requirement that the proportion of the population covered by the interval be on the average equal to the given value. The intervals posessing some of the above properties are called tolerance intervals. For normal populations there are well known methods of constructing tolerance intervals of either type mentioned. In those methods the end-points of the intervals are linear functions of sample mean and sample standard deviation. However, since in routine work in industry only means or medians and ranges of small samples are frequently recorded, it would be advatageous to have a method for the construction of tolerance intervals from this type of data. In the present paper an approximate solution is proposed and some auxiliary tables given.

Survey of existing results. Let X be a random variable with a normal distribution with mean μ and variance σ^2. In the sequel we shall denote such a distribution briefly by $N(\mu, \sigma^2)$. Further let $X_1, X_2, \ldots, X_n$ denote a sample from $N(\mu, \sigma^2)$, i. e. n mutually independent observations of X. By a tolerance interval we mean an interval whose end points are functions of the sample, say L and U, which posesses either the property

$$\mathbf{E}\left\{\int_L^U (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \mathrm{dx}\right\} = P, \tag{1}$$

or the property

$$\mathbf{P}\left\{\int_L^U (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right] \mathrm{dx} \geqq P\right\} = \gamma \tag{2}$$

where P and γ are given numbers in the interval (0, 1), $\mathbf{E}$ denotes the expectation of the random variable shown in brackets and $\mathbf{P}\{\ldots\}$ is the probability of the event shown in brackets.

Intervals (L, U) computed so as to satisfy (1) will thus in repeated samples have the following property: the long run average of areas under the frequency curve of $N(\mu, \sigma^2)$ bounded by the ordinates in the points L and U will be equal to P. Briefly we may say that the intervals satisfying (1) will cover on the average $100P\%$ of the population. Intervals (L, U) satisfying (2) provide an assurance that in a long series of samples approximately $100\gamma\%$ samples will yield values L and U such that the area under the frequency curve between ordinates in L and U will be at least equal to P. Briefly stated, approximately $100\gamma\%$ of intervals will cover at least $100P\%$ of the population.

To fix terms, we shall call intervals fulfilling condition (1) simply "$100P\%$ *tolerance intervals*", and intervals fulfilling condition (2) "$100P\%$ tolerance intervals to $100\gamma\%$ *confidence level*" or "*to confidence level* γ".

Formulas for $100P\%$ tolerance intervals have been given by Wilks [1]. His intervals have the form

$$\bar{x} - k_1 s, \quad \bar{x} + k_1 s \tag{3}$$

where $\bar{x}$ denotes the sample mean,

s the sample standard deviation,

$$s = \left\{\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{x})^2\right\}^{\frac{1}{2}} \tag{4}$$

and the coefficient k_1 is given by

$$k_1 = t_{1-P}(n-1)\sqrt{\frac{(n+1)}{n}} \tag{5}$$

In equation (5), $t_{1-P}(n-1)$ is the $100\,(1-P)\%$ critical point of Student's distribution with $n-1$ degrees of freedom, defined by

$$(6) \qquad 2\int\limits_{0}^{t_{1-P}(n-1)} \left(1+\frac{z^2}{n-1}\right)^{-\frac{n}{2}} dz = P.\Gamma\left(\frac{n-1}{2}\right)\frac{\sqrt{(n-1)\pi}}{\Gamma\left(\frac{n}{2}\right)}.$$

Details of the derivation of (5) have been described by Proschan, [2].

Wald and Wolfowitz, [3], have suggested an approximate method for the construction of tolerance intervals to a given confidence level γ. Their intervals have the same form as those of Wilks, viz.

$$(7) \qquad \bar{x}-k_2 s, \quad \bar{x}+k_2 s$$

where

$$(8) \qquad k_2 = r\sqrt{\frac{(n-1)}{\chi^2_\gamma\,(n-1)}},$$

r is the root of the equation

$$(9) \qquad \Phi\,(n^{-\frac{1}{2}}+r)-\Phi\,(n^{-\frac{1}{2}}-r)=P.$$

In (8) and (9) $\Phi(u)$ denotes the distribution function of $N(0,1)$, and $\chi^2_\gamma\,(n-1)$ is defined by the relation

$$(10) \qquad \int\limits_{\chi^2_\gamma(n-1)}^{\infty} e^{-\frac{x}{2}}\,x^{\frac{n-1}{2}-1}\,dx = \gamma.2^{\frac{n-1}{2}}.\,\Gamma\left(\frac{n-1}{2}\right).$$

The approximation involved in the method is based essentially on following considerations. It is easily seen that the following relation holds for μ known

$$(11) \qquad P\left\{\Phi\left(\frac{ts}{\sigma}\right)-\Phi\left(-\frac{ts}{\sigma}\right)\geqq P\right\}=\gamma$$

if

$$(12) \qquad t = u_{\frac{1+P}{2}}\cdot\sqrt{\frac{(n-1)}{\chi^2_\gamma(n-1)}}$$

where $u_{\frac{1+P}{2}}$ is the $100\,\frac{1+P}{2}\%$ fractile of $N\,(0,1)$ and $\chi^2_\gamma(n-1)$ is defined by (10). If the mean of the underlying normal population were μ' instead of μ and the tolerance interval were nevertheless constructed symmetrically about μ, it would be necessary to increase the coefficient t at s, viz., we would have to put t equal to

$$(13) \qquad t' = u'\sqrt{\frac{(n-1)}{\chi^2_\gamma\,(n-1)}},$$

where u' is the root of

$$(14) \qquad \Phi\left(\frac{|\mu-\mu'|}{\sigma}+u'\right)-\Phi\left(\frac{|\mu-\mu'|}{\sigma}-u'\right)=P.$$

The intervals (7) are computed symmetrically about $\bar{x}$, the true value of μ being unknown. The approximation consists in replacing $\frac{|\mu - \mu'|}{\sigma}$ in (14) by $\frac{1}{\sqrt{n}}$. Wald and Wolfowitz have investigated the accuracy of the approximation and have found that it worked very well for samples of size ten and more. The values of the coefficients k_2 given by (8) have been tabled by Bowker, [4], for some selected values of P and γ and a large scale of sample sizes. Later, Weissberg and Beatty, [5], tabled the factors r and $\sqrt{\frac{(n-1)}{\chi_\gamma^2(n-1)}}$ of (8) separately to enable the computation of tolerance intervals to a given confidence level also for more complicated experiments than a simple random sample. In such cases the intervals have the form

$$(15) \qquad \hat{\mu} - k_3\hat{\sigma}, \quad \hat{\mu} = k_3\hat{\sigma},$$

where $\hat{\mu}$ and $\hat{\sigma}$ are mutually independent estimates of μ and σ resp., $\hat{\mu}$ is supposed to be distributed normally according to $N\left(\mu, \frac{\sigma^2}{N}\right)$, $\hat{\sigma}$ is assumed to distributed as $\sigma\sqrt{\frac{\chi_f^2}{f}}$, i. e. as σ-times second square root of a chi-square variate divided by the corresponding number of degrees of freedom. The coefficient k_3 is then computed from the formula

$$(16) \qquad k_3 = r_{N,P}\sqrt{\frac{f}{\chi_\gamma^2(f)}}$$

where $r_{N,P}$ is the root of

$$(17) \qquad \Phi(N^{-\frac{1}{2}} + r_{N,P}) - \Phi(N^{-\frac{1}{2}} - r_{N,P}) = P$$

and $\chi_\gamma^2(f)$ is defined as in (10). In the paper quoted Weissberg and Beatty give extensive tables of $r_{N,P}$ and $\sqrt{\frac{f}{\chi_\gamma^2(f)}}$. As has been stated in the introduction, the problem of estimating tolerance intervals arises often in industrial statistics, where a favourite statistic for estimating variability is sample range or mean range of several samples. Thus e. g. means and ranges or maxima and minima of samples or medians and ranges are frequently recorded on control charts. To enable the subsequent utilisation of such routinely kept data to the estimation of tolerance intervals we shall consider in the next section the problem of estimating tolerance intervals on the basis of mean range.

Tolerance intervals based on mean range. The construction of tolerance intervals for normal distribution with mean range as a statistic for variability may be based on following two facts. First, it is easily seen that the special choice of $\bar{x}$ and s as statistics for mean and variance is not essential in any of the methods so far discussed, but that both Wilks's and Wald-Wolfowitz's procedures may be modified for use in connection with any pair of mutually independent estimates for μ and σ, $\hat{\mu}$ and $\hat{\sigma}$ say, where $\hat{\mu}$ is normally distributed and $\hat{\sigma}$ as a constant multiple of a chi-variate. Second, it has been shown by Patnaik, [6], that the distribution of the mean of ranges of k independent samples of size n is very closely approximated by the Γ-type distribution whose parameters are determined so as to make the first two moments of this distribution coincide with the first two moments of mean range.

Therefore, let $\bar{R}$ denote the mean range of k samples of size n from $N(\mu, \sigma^2)$ and $\hat{\mu}$ an unbiassed estimate of μ, normally distributed with variance $\frac{\sigma^2}{N}$ and independent of $\bar{R}$. According to Patnaik, [6], the density of $\bar{R}$ is very close to

$$p_{k,n}(y) = \frac{\exp\left(\frac{\nu y^2}{2c^2}\right) \cdot y^{\nu-1} \left(\frac{\nu}{c^2}\right)^{\frac{\nu}{2}}}{2^{\frac{\nu}{2}-1}\, \Gamma\left(\frac{\nu}{2}\right)}, \tag{18}$$

i. e. to the density of $c\sqrt{\frac{\chi^2}{\nu}}$, where χ^2 is a chi-square variate with ν degrees of freedom and c and ν are determined so as to satisfy

$$\frac{c\sqrt{2}\, \Gamma\left(\frac{\nu+1}{2}\right)}{\left[\sqrt{\nu}\, \Gamma\left(\frac{\nu}{2}\right)\right]} = \mathbf{E}\{\bar{R}\} \tag{19}$$

and

$$\frac{c^2}{\nu}\left[\frac{\nu - 2\Gamma^2\left(\frac{\nu+1}{2}\right)}{\Gamma^2\left(\frac{\nu}{2}\right)}\right] = \mathrm{D}\{\bar{R}\}. \tag{2}$$

Of course, the solution of (19) and (20) as a rule yields a non-integral value for ν, so that a "fractional number of degrees of freedom" results. For selected values of k and n the solutions of (19) and (20) are tabled in [7]. In this paper we give a more extensive table.

Thus if the estimate $\hat{\mu}$ is distributed according to $N\left(\mu, \frac{\sigma^2}{N}\right)$ it may be shown in the same way as in [2] that the following relation holds approximately

$$\mathbf{E}\left\{\Phi\left(\frac{z\bar{R}}{\sigma}\right) - \Phi\left(-\frac{z\bar{R}}{\sigma}\right)\right\} \doteq 2G_\nu\left(\frac{cz\sqrt{N}}{\sqrt{N+1}}\right) - 1, \tag{21}$$

where $G_\nu(t)$ denotes the distribution function of Student's distribution with (non-integral) number of degrees of freedom ν, ν and c being determined by (19) and (20). If it is required that the expectation (21) be equal to a prescribed number P, we must clearly put

$$G_\nu\left(cz\sqrt{\frac{N}{(N+1)}}\right) = \frac{P+1}{2} \tag{22}$$

whence

$$z = \frac{t_{1-P}(\nu)}{c}\sqrt{\frac{N+1}{N}}, \tag{23}$$

where, in correspondence with current usage, $t_{1-P}(\nu)$ denotes the value of a t-variate corresponding to the two sided tail area $1 - P$, viz.

$$2(1 - G_\nu(t_{1-P}(\nu))) = 1 - P. \tag{24}$$

Thus the $100P\%$ tolerance interval is given by

$$\hat{\mu} - \frac{t_{1-P}(\nu)}{c}\sqrt{\frac{N+1}{N}}\,\overline{R},\quad \hat{\mu} + \frac{t_{1-P}(\nu)}{c}\sqrt{\frac{N+1}{N}}\,\overline{R}. \tag{25}$$

If one-sided $100P\%$ tolerance limit is required, we find it from the formula

$$\hat{\mu} \pm \frac{t_{2(1-P)}(\nu)}{c}\sqrt{\frac{N+1}{N}}\,\overline{R}; \tag{26}$$

the sign "+" is used for the upper and the sign "—" for the lower limit.

To determine the $100P\%$ tolerance interval to a given confidence level γ we first note that the probability that the interval $\mu \pm z\overline{R}$ with μ known will cover at least $100R\%$ of the population is

$$\boldsymbol{P}\left\{\Phi\left(\frac{z\overline{R}}{\sigma}\right) - \Phi\left(-\frac{z\overline{R}}{\sigma}\right) \geqq P\right\} \doteq P\left\{\frac{\overline{R}}{\sigma} \geqq \frac{u_{(1+P)/2}}{z}\right\}, \tag{27}$$

where $u_{(1+P)/2}$ is the $100\,\frac{1+P}{2}\,\%$ fractile of $N(0,1)$. If we replace the distribution of $\frac{\overline{R}}{\sigma}$ by the distribution (18) with the parameters given by (19) and (20), we obtain for the probability (27) the approximation

$$\boldsymbol{P}\left\{\Phi\left(\frac{z\overline{R}}{\sigma}\right) - \Phi\left(-\frac{z\overline{R}}{\sigma}\right) \geqq P\right\} \doteq \int_{\frac{\nu\mu^2(1+P/2)}{c^2z^2}}^{\infty} \frac{\exp\left(\frac{-x}{2}\right)x^{\nu/2-1}}{2^{\nu/2}\,\Gamma\left(\frac{\nu}{2}\right)}\,\mathrm{dx}. \tag{28}$$

If the probability (28) is required to be equal to a given confidence level γ, z must be equal to

$$z = \frac{u_{(1+P/2)}}{c}\sqrt{\frac{\nu}{\chi^2_\gamma(\nu)}} = u_{(1+P/2)}\,z(\gamma, k, n), \tag{29}$$

where $\chi^2_\gamma(\nu)$ is defined similarly as in (10) and may be found by interpolation in some standard table of percentage points of χ^2. Since, however, the tolerance interval has $\hat{\mu}$ and not μ as its centre, the coefficient $u_{(1+P)/2}$ in (28) and (29) is replaced by the root $r_{(N,P)}$ of equation (17). The coefficients $r_{(N,P)}$ are extensively tabled in [5]. Thus we obtain finally as approximate $100P\%$ tolerance intervals to confidence level γ

$$\hat{\mu} - r_{N,P}z(\gamma, k, n)\,\overline{R},\quad \hat{\mu} + r_{N,P}z(\gamma, k, n). \tag{30}$$

Corresponding one-sided tolerance limits to a given confidence level γ may be obtained from an analogue of the non-central t-distribution, based on mean range. For it may be easily seen that the probability that the interval $(-\infty,$ $\hat{\mu} + t\overline{R})$ will cover at last $100P\%$ of the population $N(\mu, \sigma^2)$ is equal to

(31) $$P\left\{\Phi\left(\frac{tR+\hat{\mu}}{\sigma}\right) \geqq P\right\} \doteq P\left\{\frac{\xi + u_P\sqrt{N}}{w_{k,n}} \leqq t\sqrt{N}\right\},$$

where ξ and $w_{k,n}$ are mutually independent random variables, the former being distributed according to $N(0,1)$ and the latter as the mean range of k independent samples of size n from a normal population with unit variance. If this probability is required to be equal to a specified confidence level γ, the coefficient t must be chosen to be

(32) $$t = \frac{1}{\sqrt{N}}\, t_\gamma\,(k, n, u_P\sqrt{N}),$$

where $t_\gamma\,(k, n, u_P\sqrt{N})$ denotes the $100P\%$ quantile of the statistic

(33) $$\frac{(\xi + \mu_P\sqrt{N})}{w_{k,n}}$$

If we replace — following Patnaik, [6] — the variable $w_{k,n}$ by the variable $c\sqrt{\frac{\chi^2}{\nu}}$, where x^2 has ν "degrees of freedom", c and ν being determined according to (19) and (20), we may regard (33) as a $\frac{1}{c}$ — multiple of a non-central t-variable with ν degrees of freedom and the parameter of non-centrality equal to $u_P\sqrt{N}$. Using the notation of Johnson and Welch, [8] we then may approximately put the upper $100P\%$ tolerance limit to confidence level γ equal to

(34) $$\hat{\mu} + \frac{1}{c\sqrt{N}}\,.\,t\,(\nu, u_P\sqrt{N}, 1-\gamma)\,\bar{R},$$

where c and ν depend on the number k and size n of subsamples for $\bar{R}$. Alternatively, we may use the tables of A. Žaludová [9] for the determination of $t_\gamma\,(k, n, u_P\sqrt{N})$

The tables. The main part of this paper are the tables which have been computed to facilitate the determination of tolerance intervals. Table I. contains the constants ν and c for the approximation to the distribution of mean range. This table has been computed by numerically solving equations (19) and (20) on a desk calculating machine. This table may be used in connection with formulas (25) and (26); the critical point of t-distribution corresponding to the fractional "number of degrees of freedom" ν, is found by interpolation in some standard table of Student's distribution. Alternatively, we may use a table of a mean-range analogue of Student's statistic (e. g. Lord, [10].)

Tables IIa, IIb and IIc contain the coefficients

(35) $$z\,(\gamma, k, n) = \frac{1}{c}\sqrt{\frac{\nu}{\chi^2_\gamma\,(\nu)}}$$

appearing in formula (30). Three confidence levels are included, viz. $\gamma = 0{,}90$; $\gamma = 0{,}95$ and $\gamma = 0{,}99$. The tables have been computed in the following manner: the constants ν and c have been taken from Table I., the $\chi^2_\gamma(\nu)$ values have been found for $\nu \leqq 30$ by interpolation in the table [11] by means of Lagrangian four point formula, for $30 < \nu \leqq 100$ by linear interpolation, and for $\nu > 100$ computed according to Goldberg's and Levine's formulas [12]. The computations have been carried out by Mrs M. KRŠŇÁKOVÁ in the computing laboratory of the Institute of Mathematical Statistics at the Charles' University, Praha.

An example. To demonstrate the application of the formulas developed in the preceding section let us consider the following example. Assume that we have at our disposal the records of twenty samples from a normal distribution of size five and that the statistics recorded are the sample medians and sample

Table I

Constants ν and c for the approximation of the distribution of mean range of k samples of size n by the χ^2-distribution

k \ n	5	6	7	8	9	10	11	12	15	20
1	3.829 2.481	4.679 2.672	5.486 2.830	6.252 2.963	6.983 3.078	7.668 3.179	8.349 3.269	8.990 3.350	10.673 3.554	13.373 3.805
2	7.472 2.405	9.160 2.604	10.768 2.768	12.296 2.906	13.753 3.024	15.146 3.129	16.480 3.221	17.760 3.305	21.319 3.513	26.515 3.770
3	11.103 2.379	13.634 2.581	16.040 2.747	18.332 2.886	20.516 3.006	22.603 3.112	24.604 3.205	26.524 3.289	31.862 3.499	39.653 3.759
4	14.729 2.366	18.103 2.570	21.311 2.736	24.364 2.877	27.276 2.997	30.060 3.103	32.727 3.197	35.284 3.282	42.401 3.492	52.791 3.753
5	18.355 2.358	22.571 2.563	26.579 2.730	30.397 2.871	34.035 2.992	37.514 3.098	40.849 3.192	44.047 3.277	52.941 3.488	65.931 3.749
6	21.980 2.353	27.039 2.558	31.848 2.726	36.428 2.867	40.795 2.988	44.969 3.095	48.970 3.189	52.808 3.274	63.480 3.486	79.064 3.747
7	25.608 2.349	31.505 2.555	37.116 2.723	42.459 2.864	47.554 2.986	52.424 3.092	57.092 3.187	61.569 3.272	74.020 3.484	92.200 3.745
8	29.228 2.346	35.971 2.552	42.384 2.720	48.490 2.862	54.312 2.984	59.878 3.090	65.213 3.185	70.325 3.270	84.559 3.482	105.336 3.744
9	32.852 2.344	40.438 2.550	47.652 2.719	54.521 2.860	61.072 2.982	67.333 3.089	73.334 3.184	79.090 3.269	95.100 3.481	118.485 3.743
10	36.475 2.342	44.904 2.549	52.920 2.717	60.552 2.859	67.830 2.981	74.786 3.088	81.456 3.183	87.850 3.268	105.639 3.480	131.617 3.742
12	43.722 2.339	53.836 2.546	63.455 2.715	72.614 2.857	81.347 2.979	89.694 3.086	97.696 3.181	105.372 3.266	126.717 3.479	157.888 3.741
15	54.592 2.337	67.234 2.544	79.257 2.713	90.705 2.855	101.622 2.977	112.057 3.084	122.058 3.179	131.655 3.265	150.793 3.478	179.293 3.740
20	72.716 2.334	89.564 2.541	105.593 2.711	120.849 2.853	135.413 2.976	149.330 3.083	162.665 3.178	175.457 3.263	211.033 3.476	262.978 3.739
30	108.94 2.331	134.21 2.539	158.27 2.709	181.19 2.851	203.00 2.974	223.87 3.081	243.88 3.176	263.06 3.262	316.42 3.475	394.35 3.737
60	217.63 2.329	268.20 2.537	316.28 2.706	362.14 2.849	405.75 2.972	447.47 3.079	487.47 3.174	525.87 3.260	632.59 3.473	788.39 3.736

Table IIa

Coefficients $z(\gamma, k, n)$ for the determination of tolerance limits to confidence level $\gamma = 0{,}90$ with mean range as an estimator of the standard deviation

k \ n	5	6	7	8	9	10	11	12	15	20
1	0.794	0.671	0.595	0.546	0.508	0.478	0.455	0.435	0.394	0.352
2	0.641	0.562	0.510	0.473	0.445	0.418	0.405	0.390	0.358	0.325
3	0.589	0.522	0.477	0.445	0.420	0.401	0.385	0.373	0.343	0.312
4	0.562	0.500	0.459	0.430	0.407	0.389	0.374	0.362	0.334	0.306
5	0.544	0.487	0.448	0.420	0.398	0.381	0.367	0.355	0.329	0.301
6	0.532	0.477	0.440	0.413	0.392	0.375	0.362	0.350	0.325	0.298
7	0.523	0.470	0.434	0.408	0.387	0.371	0.358	0.347	0.322	0.296
8	0.519	0.464	0.429	0.403	0.384	0.368	0.355	0.344	0.319	0.294
9	0.510	0.460	0.425	0.400	0.381	0.365	0.352	0.342	0.317	0.292
10	0.505	0.456	0.422	0.397	0.378	0.363	0.350	0.340	0.316	0.291
12	0.498	0.454	0.417	0.393	0.374	0.359	0.347	0.336	0.313	0.288
15	0.490	0.443	0.411	0.388	0.370	0.355	0.343	0.333	0.311	0.287
20	0.481	0.436	0.405	0.383	0.365	0.351	0.339	0.329	0.307	0.283
30	0.471	0.428	0.398	0.376	0.359	0.346	0.335	0.325	0.303	0.281
60	0.458	0.418	0.390	0.369	0.352	0.339	0.329	0.319	0.299	0.277

Table IIb

Coefficients $z(\gamma, k, n)$ for the determination of tolerance limits to confidence level $\gamma = 0{,}95$ with mean range as an estimator of the standard deviation

k \ n	5	6	7	8	9	10	11	12	15	20
1	0.971	0.800	0.694	0.625	0.575	0.538	0.508	0.483	0.431	0.382
2	0.729	0.628	0.563	0.514	0.484	0.458	0.437	0.420	0.382	0.344
3	0.650	0.569	0.516	0.478	0.450	0.428	0.409	0.394	0.361	0.327
4	0.610	0.538	0.491	0.456	0.431	0.410	0.394	0.380	0.349	0.318
5	0.585	0.519	0.475	0.443	0.419	0.399	0.384	0.371	0.342	0.312
6	0.568	0.505	0.463	0.433	0.410	0.392	0.377	0.364	0.336	0.307
7	0.555	0.495	0.455	0.426	0.404	0.386	0.372	0.359	0.332	0.304
8	0.549	0.487	0.445	0.420	0.399	0.382	0.367	0.356	0.329	0.302
9	0.537	0.481	0.443	0.416	0.395	0.378	0.364	0.352	0.327	0.301
10	0.530	0.476	0.439	0.412	0.391	0.375	0.361	0.350	0.326	0.299
12	0.520	0.468	0.432	0.406	0.386	0.370	0.357	0.346	0.321	0.295
15	0.509	0.458	0.425	0.399	0.383	0.365	0.352	0.341	0.318	0.292
20	0.499	0.449	0.416	0.392	0.374	0.359	0.348	0.336	0.313	0.288
30	0.483	0.440	0.407	0.385	0.366	0.352	0.340	0.330	0.308	0.284
60	0.466	0.424	0.396	0.374	0.357	0.344	0.333	0.323	0.302	0.279

ranges. Then we may take as an estimate of the population mean the average of the twenty medians,

$$\hat{\mu} = \frac{1}{20} \sum_{i=1}^{20} \tilde{x}_i .$$

This estimate will be very nearly normally distributed with mean μ and variance

Table IIc

Coefficients z (γ, k, n) for the determination of tolerance limits to confidence level $\gamma = 0{,}99$ with mean range as an estimator of the standard deviation

k \ n	5	6	7	8	9	10	11	12	15	20
1	1.515	1.149	0.952	0.833	0.746	0.680	0.633	0.595	0.518	0.444
2	0.917	0.791	0.692	0.652	0.577	0.540	0.510	0.487	0.436	0.386
3	0.795	0.678	0.604	0.552	0.514	0.485	0.461	0.442	0.400	0.358
4	0.720	0.622	0.559	0.515	0.482	0.456	0.436	0.418	0.381	0.343
5	0.676	0.589	0.532	0.492	0.462	0.438	0.419	0.404	0.369	0.333
6	0.645	0.566	0.513	0.476	0.448	0.426	0.408	0.393	0.360	0.327
7	0.624	0.549	0.500	0.464	0.438	0.417	0.400	0.385	0.354	0.322
8	0.612	0.536	0.489	0.455	0.430	0.410	0.393	0.379	0.349	0.318
9	0.594	0.526	0.481	0.448	0.423	0.404	0.388	0.374	0.345	0.314
10	0.583	0.517	0.474	0.442	0.418	0.399	0.383	0.370	0.342	0.312
12	0.566	0.504	0.463	0.433	0.410	0.391	0.376	0.364	0.336	0.307
15	0.548	0.490	0.451	0.423	0.401	0.383	0.369	0.357	0.332	0.303
20	0.529	0.475	0.438	0.412	0.391	0.374	0.361	0.350	0.324	0.297
30	0.507	0.458	0.424	0.399	0.380	0.364	0.352	0.341	0.317	0.292
60	0.483	0.448	0.407	0.384	0.366	0.352	0.340	0.330	0.308	0.284

$$\mathrm{D}\,(\hat{\mu}) = \frac{0{,}28683\sigma^2}{20},$$

where 0,28683 is the variance of the median of the sample of size five from N (0,1); the value has been taken from [13]. Thus the factor N appearing in our formulas is $N = 69{,}73$. From table I we have as the constants for the approximation of the distribution of mean range $\nu = 72{,}716$ and $c = 2{,}331$. The average 90% tolerance interval then is obtained from formula (25) as

$$\hat{\mu} - 0{,}720\overline{R}, \quad \hat{\mu} + 0{,}720\overline{R}.$$

A 90% tolerance interval to, say, confidence level $\gamma = 0{,}99$ is obtained as follows: from Weissberg's and Beatty's table [5] (or from a table of the normal distribution computing by trial and error) we find $r_{69.73;\,0.9} = 1{,}656627$. Then from Table IIc we get z (0,99; 20; 5) $= 0{,}529$. Thus the 90% tolerance interval to confidence level 0,99 is

$$\hat{\mu} - 0{,}876\overline{R}, \quad \hat{\mu} + 0{,}876\overline{R}.$$

REFERENCES

[1] Wilks S. S., Determination of sample sizes for setting tolerance limits, Annals of mathematical statistics, 4 (1941), 91—96.

[2] Proschan F., Confidence and tolerance intervals for the normal distribution, Journal of the American statistical association, 48 (1953), 550—564.

[3] Wald A., Wolfovitz J., Tolerance limits for a normal distribution, Annals of mathematical statistics 17 (1946), 208—215.

[4] Bowker A., Chapter 2. in "Eisenhart C., Hastay M. W., Wallis W. A., Selected techniques of statistical analysis", New York 1947.

[5] Weissberg A., Beatty G., Tables of tolerance limit factors for normal distributions, Technometrics 2 (1960), 483—500.

[6] PATNAIK, P. B., The use of mean range as an estimator of variance in statistical tests, Biometrika 37, 1950, 78—87.
[7] PEARSON, E. S.-HARTLEY, H. O., Biometrika Tables for Statisticians, Camb. Univ. Press 1954.
[8] JOHNSON N. L., WELCH B. L., Applications of the non-central t-distribution, Biometrika 31, 1940, 362—389.
[9] ŽALUDOVÁ A., BILWACHS M., Research Report No 5337 of The Research Institute for Heat Engineering, Praha.
[10] LORD E., The use of range in place of standard deviation in the t-test, Biometrika 34 (1947), 41—67.
[11] HALD A.-SINKBAEK S. A., A table of percentage points of the χ^2-distribution, Skandinavisk Aktuarietidskrift 33, (1950) 168—175.
[12] GOLDBERG A.-LEVINE H., Approximate formulas for the percentage points and normalization of t and χ^2, Annals of Math. Stat., 17, 1946, 216—225.
[13] GODWIN H. J., Some low moments of order statistics, Annals of Mathematical Statistics 20, (1949), 279—285.

TOLERANČNÍ MEZE PRO NORMÁLNÍ ROZDĚLENÍ

Souhrn

Označme $\overline{R}$ průměrné rozpětí k navzájem nezávislých náhodných výběrů rozsahu n ze souboru s normálním rozdělením se střední hodnotou μ a s rozptylem σ^2, $\hat{\mu}$ nestranný odhad parametru μ nezávislý na $\overline{R}$ a rozdělený normálně s rozptylem σ^2/N, kde N je známé číslo. V článku jsou tabelovány pomocné veličiny pro stanovení tolerančních intervalů, založených na $\hat{\mu}$ a $\overline{R}$. (V tabulce I. jsou sestaveny konstanty c a ν pro aproximaci rozdělení veličiny $\overline{R}/\sigma$ (viz seznam literatury, [6]). Interval tvaru

$$\hat{\mu} - \frac{t_{1-P}(\nu)}{c} \sqrt{\frac{N+1}{N}}\,\overline{R}, \quad \hat{\mu} + \frac{t_{1-P}(\nu)}{c} \sqrt{\frac{N+1}{N}}\,\overline{R}$$

pokrývá v průměru 100P% základního souboru. ($t_{1-P}(\nu)$ je 100(1-P)% kritická hodnota Studentova rozdělení s ν stupni volnosti, definovaná vztahem

$$2\int_{t_{1-P}(\nu)}^{\infty} \left(1+\frac{\mu^2}{\nu}\right)^{-\frac{\nu+1}{2}} du = \frac{(1-P)\,\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}}{\Gamma\left(\frac{\nu+1}{2}\right)}.)$$

Tabulky IIa, IIb a IIc obsahují veličiny $z\,(\gamma, k, n)$ takové, že interval

$$\hat{\mu} - r_{N,P}\, z\,(\gamma, k, n)\,\overline{R}, \quad \hat{\mu} + r_{N,P}\, z\,(\gamma, k, n)\,\overline{R}$$

pokryje s pravděpodobností γ aspoň 100P% základního souboru. ($r_{N,P}$ je kořen rovnice

$$\frac{1}{\sqrt{2\pi}} \int_{1/\sqrt{N}-r}^{1/\sqrt{N}+r} e^{-\frac{u^2}{2}}\, du = P.)$$

ТОЛЕРАНТНЫЕ ПРЕДЕЛЫ ДЛЯ НОРМАЛЬНОГО РАСПРЕДЕЛЕНИЯ

Резюме

Пусть $\overline{R}$ средний размах k взаимно независимых выборок объема n из нормально распределенной совокупности с математическим ожиданием μ и с дисперсией σ^2, $\hat{\mu}$ несмещенная оценка параметра μ, распределенная нормально с дисперсией $\sigma^2 N$, где N - известное число. В настоящей статье табулированы вспомогательные величины для вычисления толерантных интервалов, основанных на $\hat{\mu}$ и $\overline{R}$. В таблице I приведены постоянные c и ν, отвечающие разным сочетаниям k и n и обладающие свойством: интервал

$$\hat{\mu} - \frac{t_{1-P}(\nu)}{c}\sqrt{\frac{N+1}{N}}\,\overline{R}, \quad \hat{\mu} + \frac{t_{1-P}(\nu)}{c}\sqrt{\frac{N+1}{N}}\,\overline{R},$$

где $t_{1-P}(\nu)$ — 100P -перцентное критическое значение распределения Стьюдента определяемое соотношением

$$2\int_{t_{1-P}(\nu)}^{\infty}\left(1+\frac{u^2}{\nu}\right)^{-\frac{\nu+1}{2}} du = \frac{(1-P)\,\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}}{\Gamma\left(\frac{\nu+1}{2}\right)}.$$

покрывает в среднем долю $100P\%$ генеральной совокупности. В таблице II приведены величины $z(\gamma, k, n)$, обладающие свойством: интервал

$$\hat{\mu} - r_{N,P} \cdot z(\gamma, k, n)\,\overline{R},\ \hat{\mu} + r_{N,P}\, z(\gamma, k, n)\,\overline{R},$$

где $r_{N,P}$ — корень уравнения

$$\frac{1}{\sqrt{2\pi}}\int_{1/\sqrt{N}-r}^{1/\sqrt{N}+r} e^{-\frac{u^2}{2}} du = P,$$

покроет с вероятностью γ по меньшей мере $100P\%$ совокупности. Таблица IIa предназначена для коэффициента доверия $\gamma = 0,90$, таблица IIb для $\gamma = 0,95$ и IIc для $\gamma = 0,99$.