Tolerance Intervals Demystified

Authored by: Kevin Splinter
Gina Sigler
Michael Harman
Dr. Kyle Kolsti
31 August 2020

The goal of the STAT COE is to assist in developing rigorous, defensible test strategies to more effectively quantify and characterize system performance and provide information that reduces risk. This and other COE products are available at www.afit.edu/STAT.

STAT Center of Excellence
2950 Hobson Way – Wright-Patterson AFB, OH 45433

Table of Contents

Executive Summary

So, you have decided you might need a tolerance interval? There are three types of statistical intervals: confidence intervals, prediction intervals, and tolerance intervals. While introductory statistics courses will typically cover confidence intervals and prediction intervals, tolerance intervals are largely ignored. All three of these intervals can be very useful especially in the Department of Defense (DoD) and Department of Homeland Security (DHS) as they have applications in software, cyber, performance testing, and more. This paper will detail the differences between the statistical intervals, explain when a tolerance interval might be called for, introduce a tool that can be used for calculating sample size, and cover one particular method of calculation. For simplicity, all data is assumed to be normally distributed, as non-parametric intervals require much more complicated formulas and methods.

Keywords: tolerance interval, sample size, requirement, normal distribution

Introduction

Tolerance intervals, while probably the least known interval, are often more applicable than their confidence interval counterparts. That is because the purpose of a tolerance interval is to contain a specified proportion of a population, which often matches the goal of a requirement. However, tolerance intervals can be more mathematically difficult to calculate than confidence intervals. This means that confidence intervals are taught more widely in introductory statistics classes and thus better known. Also, statisticians have many different ideas on exactly how tolerance intervals should be calculated (Jensen 2009 compares many formulas), so there is no consensus on a single formula. The good news is that the methods yield fairly similar results as long as assumptions are met.

Differences from Other Intervals

As mentioned before, there are three types of statistical intervals: confidence intervals, prediction intervals, and tolerance intervals. While this paper focuses on tolerance intervals, it is important to note the differences between the three intervals to make sure the proper interval is chosen for the specific requirement.

Confidence intervals are used to place bounds on a single population parameter, like the population mean or variance, with a specified level of confidence. Their width is simply due to sampling error, which approaches zero as sample size increases. Confidence intervals are generated using a point estimate of the parameter and creating an interval range on that estimate. If a confidence interval seems appropriate, the STAT COE has two additional best practices focused on confidence interval interpretations (Kensler 2014) and confidence intervals for percentiles (Burke 2016). Note that confidence intervals for percentiles are related to tolerance intervals; a one-sided confidence interval for a specific quantile is equivalent to one-sided tolerance interval bound for the same proportion for the normal distribution (Meeker 2017). This is not true for two-sided intervals.

Prediction intervals place bounds on one or more future observations. Prediction intervals are always wider than confidence intervals because they are accounting for extra variance of the individual values, as well as the variance of the parameter (Meeker 2017).

Tolerance intervals are often confused with confidence intervals even though the two have completely different construction and interpretations. Tolerance intervals bound the range of values which is likely to contain a certain proportion of a population, and their width is determined not only by sampling error, but also variance in the population itself. In short, tolerance intervals place bounds for where a proportion of a population is likely to fall, while confidence intervals place bounds on parameters of the population. Tolerance intervals are most easily conceptualized by thinking about parts from a production line that must meet dimensional metrics like length; 95% of all parts must be between 26 and 34 inches in length, with 95% confidence. Tolerance intervals require a confidence level as well as a specified proportion for their calculation. Figure 1 is a visualization of two potential tolerance intervals imposed on a normal distribution with a mean $\mu = 30$ and a standard deviation $\sigma = 2$. As the sample size gets very large, the two-sided tolerance interval that contains 95% of the population would approach $\mu \pm 1.96\sigma$, or in this case [26.08, 33.92]. This interval contains the middle 95% of the area (leaving 2.5% at each tail), as shown by red shaded region. In contrast, the one-sided upper tolerance interval contains the lower 95% of the area which is bounded by the upper tolerance limit of $x = 33.29$, as shown by the diagonal hatched area. It will be shown later that for finite sample size, the tolerance intervals must be larger than those in this figure. How much larger the interval must be depends on the sample size and the confidence level. Table 1 lists all three interval types and their associated purposes for easy reference.

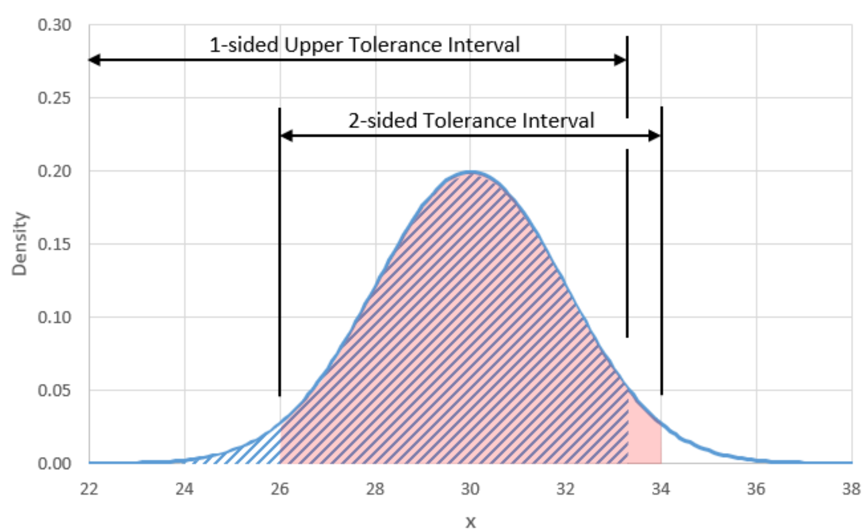

Figure 1: Visualization of one-sided and two-sided tolerance intervals

Table 1: Table of statistical intervals and their purposes

Interval Type	Purpose
Confidence Interval	Uncertainty of a parameter, like the population mean
Prediction Interval	Uncertainty of next observation(s)
Tolerance Interval	Estimate of range where a proportion of the population likely exists

Further information on each of the intervals can be found in Ortiz & Truett (2015).

Deciphering Requirements

How do you know if a tolerance interval is most appropriate for your situation? The clues can usually be found in the wording of the requirement. There are two key parts to a tolerance interval: the proportion and the confidence. Proportion refers to the proportion of the population that falls within the limits of the interval. Confidence retains the same definition as it did with confidence intervals: confidence is the likelihood that the specified interval contains the coverage proportion. Look for the mention of these two parts in the requirement. Ideally, the requirement will mention both the coverage proportion and the confidence level. However, many might skip this part merely inferring a level of confidence elsewhere in the documentation. Here are some examples of requirements that indicate needing a tolerance interval:

1. The software must respond to the cyber threat within 5 seconds 99.5% of the time.
2. 90% of the error levels must fall between 0.4 and 0.7 nm.
3. The crew shall be able to recover a person in the water within 10 minutes of the alarm initiation, 90% of the time.
4. Data latency should be no more than 30 minutes 95% of the time with 90% confidence.

Often in requirements, the confidence level is not mentioned, but the mention of a coverage proportion still indicates the need for a tolerance interval. If not otherwise specified, a 95% confidence level is often assumed. If at all possible, a specific confidence level should be discussed with the planning team before any intervals are constructed or calculated.

Before Testing- Determining Sample Size

Once it is determined that a tolerance interval is the most appropriate method for assessing the requirement, tolerance interval width predictions can be used during the planning portion of testing. As with other test sizing approaches, the goal is to find the optimal sample size that balances risk versus resources. The first step is to determine if a one- or two-sided interval is needed. A two-sided interval is used when the middle proportion of a population is of interest. A one-sided interval is used when the desired population proportion should be above or below a single bound.

Two-sided Intervals

The STAT COE has created a tool that will bound a suggested sample size for normally distributed data (email COE@AFIT.edu to obtain the Excel tools). To use the tool, the user must be able to specify a needed proportion, confidence level, estimated sample mean, and estimated standard deviation. Users also have the option to specify upper and lower requirement values, an approximate sample size (N), and increments to sigma and N. All values that can be changed by the user are indicated in blue in Figure 2. Given these values, the tool will increment the estimated standard deviation and sample size to show a table of potential upper and lower bounds. Bounds that are within the level of the requirement are marked in green (upper) and orange (lower). Since the sigma and mean estimates are used in the sample size calculation, it is good practice not to select the absolute minimum sample size. This mitigates risk in case differences exist between the estimated sample mean and standard deviation and the calculated sample mean and standard deviation.

2-Sided Tolerance Interval Calculator

Enter blue cells only

Green indicates Upper TI values meets Requirement

Orange indicates Lower TI values meets Requirement

Reference https://www.itl.nist.gov/div898/handbook/prc/section2/prc263.htm

Probability	0.99	
Confidence	0.95	
Upper Requirement	3.00	Limit Value
Lower Requirement	-3.00	Limit Value
Mean	0.50	
Sigma Increment	0.1	increases Sigma along table y axis
N Increment	2	increases N along table X axis

Sigma	N 2	4	6	8	10	12	14	16	18	20	22	24
0.100	5.531	1.341	1.081	0.991	0.944	0.916	0.896	0.882	0.871	0.862	0.854	0.848
0.100	-4.531	-0.341	-0.081	0.009	0.056	0.084	0.104	0.118	0.129	0.138	0.146	0.152
0.200	10.562	2.182	1.663	1.482	1.389	1.331	1.292	1.263	1.241	1.223	1.209	1.197
0.200	-9.562	-1.182	-0.663	-0.482	-0.389	-0.331	-0.292	-0.263	-0.241	-0.223	-0.209	-0.197
0.300	15.593	3.023	2.244	1.973	1.833	1.747	1.688	1.645	1.612	1.585	1.563	1.545
0.300	-14.593	-2.023	-1.244	-0.973	-0.833	-0.747	-0.688	-0.645	-0.612	-0.585	-0.563	-0.545
0.400	20.624	3.864	2.825	2.464	2.278	2.163	2.084	2.026	1.982	1.947	1.918	1.894
0.400	-19.624	-2.864	-1.825	-1.464	-1.278	-1.163	-1.084	-1.026	-0.982	-0.947	-0.918	-0.894
0.500	25.655	4.705	3.406	2.955	2.722	2.579	2.480	2.408	2.353	2.309	2.272	2.242
0.500	-24.655	-3.705	-2.406	-1.955	-1.722	-1.579	-1.480	-1.408	-1.353	-1.309	-1.272	-1.242
0.600	30.686	5.546	3.988	3.446	3.167	2.994	2.876	2.790	2.723	2.670	2.627	2.591
0.600	-29.686	-4.546	-2.988	-2.446	-2.167	-1.994	-1.876	-1.790	-1.723	-1.670	-1.627	-1.591
0.700	35.717	6.386	4.569	3.937	3.611	3.410	3.272	3.171	3.094	3.032	2.981	2.939
0.700	-34.717	-5.386	-3.569	-2.937	-2.611	-2.410	-2.272	-2.171	-2.094	-2.032	-1.981	-1.939
0.800	40.747	7.227	5.150	4.428	4.056	3.826	3.668	3.553	3.464	3.394	3.336	3.288
0.800	-39.747	-6.227	-4.150	-3.428	-3.056	-2.826	-2.668	-2.553	-2.464	-2.394	-2.336	-2.288
0.900	45.778	8.068	5.731	4.919	4.500	4.242	4.064	3.935	3.835	3.755	3.690	3.636
0.900	-44.778	-7.068	-4.731	-3.919	-3.500	-3.242	-3.064	-2.935	-2.835	-2.755	-2.690	-2.636

Figure 2: 2-sided tolerance interval sample size tool

One-sided Intervals

The STAT COE also has a similar tool for one-sided tolerance intervals. Users are asked to fill in their desired values in the blue highlighted cells seen in Figure 3, and the upper or lower bound of the different possible intervals is shown in the table. Values that are highlighted in green indicate the system would meet the requirement.

1-Sided Tolerance Interval Calculator (Upper or Lower)

Enter blue cells only

Green indicates TI values meets Requirement

Reference https://www.itl.nist.gov/div898/handbook/prc/section2/prc263.htm

Probability	0.99	
Confidence	0.99	
Requirement	5	Limit Value
Table Interval Choice	1	-1 for lower limit, +1 for upper limit
Mean	2.00	
Sigma Increment	0.1	increases Sigma along table y axis
N Increment	1	increases N along table X axis

	N										
Sigma	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
0.100	2.535	2.504	2.480	2.461	2.445	2.432	2.420	2.411	2.402	2.395	2.388
0.200	3.070	3.008	2.960	2.921	2.890	2.863	2.841	2.821	2.804	2.789	2.776
0.300	3.605	3.512	3.440	3.382	3.335	3.295	3.261	3.232	3.207	3.184	3.164
0.400	4.141	4.016	3.920	3.843	3.780	3.727	3.682	3.643	3.609	3.579	3.552
0.500	4.676	4.520	4.400	4.304	4.225	4.159	4.102	4.054	4.011	3.974	3.940
0.600	5.211	5.024	4.880	4.764	4.670	4.590	4.523	4.464	4.413	4.368	4.328
0.700	5.746	5.528	5.360	5.225	5.115	5.022	4.943	4.875	4.815	4.763	4.716
0.800	6.281	6.032	5.840	5.686	5.560	5.454	5.364	5.286	5.218	5.158	5.104
0.900	6.816	6.537	6.320	6.147	6.004	5.885	5.784	5.696	5.620	5.552	5.492
1.000	7.352	7.041	6.800	6.607	6.449	6.317	6.204	6.107	6.022	5.947	5.880
1.100	7.887	7.545	7.280	7.068	6.894	6.749	6.625	6.518	6.424	6.342	6.268
1.200	8.422	8.049	7.760	7.529	7.339	7.181	7.045	6.929	6.827	6.737	6.656
1.300	8.957	8.553	8.240	7.990	7.784	7.612	7.466	7.339	7.229	7.131	7.045
1.400	9.492	9.057	8.720	8.450	8.229	8.044	7.886	7.750	7.631	7.526	7.433
1.500	10.027	9.561	9.200	8.911	8.674	8.476	8.307	8.161	8.033	7.921	7.821
1.600	10.562	10.065	9.680	9.372	9.119	8.907	8.727	8.571	8.435	8.315	8.209
1.700	11.098	10.569	10.160	9.833	9.564	9.339	9.148	8.982	8.838	8.710	8.597
1.800	11.633	11.073	10.640	10.293	10.009	9.771	9.568	9.393	9.240	9.105	8.985

Figure 3: 1-sided tolerance interval sample size tool

Calculation

This section provides the mathematical formulas to calculate tolerance intervals that were employed by the Excel tool. As stated previously, this Best Practice is focusing on data from a normal distribution, which has closed form solutions for tolerance interval calculations. The formulas for one-sided and two-sided tolerance intervals require the calculation of a constant k_1 or k_2, respectively. These formulas are provided in the next two sections. The variables found in the formulas are defined in Table 2.

Table 2: Definitions of the variables used in the tolerance interval formulas

Variable	Formula	Definition
α	$\alpha = 1 - \text{confidence}$	Level of significance or Type I Error Ex: If confidence=95%, $\alpha = 0.05$
p	n/a; user's choice	Proportion of the population desired to be contained by the tolerance interval
n	n/a	Sample size; number of observations
$\bar{x}$	$\bar{x} = \frac{1}{n}\sum x_i$	Sample mean
s	$s = \frac{\sum(x_i - \bar{x})^2}{n-1}$	Sample standard deviation
ν	$\nu = n - 1$	Number of degrees of freedom
z_m	=NORM.S.INV(m) (Excel)	m^{th} quantile of the standard normal distribution
$\chi^2_{m,\nu}$	=CHISQ.INV.RT(m,ν) (Excel)	m^{th} quantile of the Chi-Square distribution with ν degrees of freedom

Two-sided Intervals

The upper and lower limits for a two-sided tolerance interval (Howe 1969) are:

$$\bar{x} \pm k_2 s \tag{1}$$

where

$$k_2 = z_{(1-p)/2}\sqrt{\frac{\nu(1 + 1/n)}{\chi^2_{1-\alpha,\nu}}} \tag{2}$$

Other methods may be found in Wheeler (1993).

One-sided Intervals

The variables in the one-sided interval calculation hold the same meaning they do in the two-sided calculation. This method may be found in Lieberman (1957). Other methods may be found in Owen (1963), Link (1985), Wheeler (1993).

The upper limit of the lower 1-sided tolerance interval is:

$$\bar{x} + k_1 s \tag{3}$$

Likewise, the lower limit of the upper 1-sided tolerance interval is:

$$\bar{x} - k_1 s \tag{4}$$

where

$$k_1 = \frac{z_p + \sqrt{z_p^2 - ab}}{a} \tag{5}$$

$$a = 1 - \frac{z_\alpha^2}{2(n-1)} \tag{6}$$

$$b = z_p^2 - \frac{z_\alpha^2}{n} \tag{7}$$

Additionally, for simplicity, a reference table has been created for specific values that can be used to calculate the tolerance intervals and is included in Appendix A. Statistical software packages such as JMP or R also include tools to calculate tolerance intervals. Please contact the STAT COE for more information on calculations of tolerance intervals.

Conclusion

Tolerance intervals are perhaps the least-known type of statistical interval, but they are directly applicable to many analyses. When a requirement calls for a specified proportion of a distribution to be above or below a value, or between two values, it may be hinting that the tolerance interval is the most appropriate approach. This Best Practice provided several ways to calculate tolerance intervals: an Excel tool, formulas, and a table. Though the methods covered here are specific to normally distributed data, there are non-parametric methods for tolerance interval calculation available as well (Hahn & Meeker 2017).

References

Burke, Sarah. “Confidence Intervals for the Median and Other Percentiles”. Scientific Test and Analysis Techniques Center of Excellence (STAT COE), 2016.

Howe, W. G. “Two-Sided Tolerance Limits for Normal Populations, Some Improvements.” Journal of the American Statistical Association, vol. 64, no. 326, 1969, pp. 610–620.

Jensen, Willis, “Approximations of Tolerance Intervals for Normally Distributed Data,” Quality and Reliability Engineering International, Vol. 25, Issue 5, pp. 571-580, July 2009.

Kensler, Jennifer and Cortes, Luis. (2014). “Interpreting Confidence Intervals.” Scientific Test and Analysis Techniques Center of Excellence (STAT COE), 2014.

Lieberman, Gerald, “Tables for One-Sided Statistical Tolerance Limits,” Technical Report No. 34, Applied Mathematics and Statistics Laboratory, Stanford University, for Office of Naval Research, 1957.

Link, Carol, “An Equation for One-Sided Tolerance Limits for Normal Distributions,” Research Paper FPL 458, Forest Products Laboratory, U.S. Department of Agriculture, 1985.

Meeker, William Q., et al. *Statistical Intervals: A Guide for Practitioners and Researchers.* John Wiley & Sons, Inc., 2017.

Ortiz, Francisco and Truett, Lenny. “Using Statistical Intervals to Assess System Performance”. Scientific Test and Analysis Techniques Center of Excellence (STAT COE), 2015.

Owen, D. B., “Factors for One-Sided Tolerance Limits and for Variables Sampling Plans,” Sandia Corporation Monograph SCR-607, March 1963.

“Tolerance Intervals for a Normal Distribution.” *NIST/SEMATECH e-Handbook of Statistical Methods*, NIST, 30 October 2013, *7.2.6.3. Tolerance Intervals for a Normal Distribution*, www.itl.nist.gov/div898/handbook/prc/section2/prc263.htm, August 2020.

Wheeler, J.T., “Statistical Computation of Tolerance Limits,” NASA Technical Memorandum 108424. Structures and Dynamics Laboratory, October 1993.

Appendix A

Table with values of k_1 and k_2 for calculating 1-sided and 2-sided tolerance intervals:

p	Confidence	n	k_2 (2-sided)	k_1 (1-sided)
0.9	80%	10	2.231	1.713
		20	1.984	1.566
		30	1.899	1.508
	90%	10	2.535	2.012
		20	2.152	1.744
		30	2.025	1.644
	95%	10	2.838	2.321
		20	2.310	1.910
		30	2.140	1.767
0.95	80%	10	2.659	2.147
		20	2.364	1.974
		30	2.263	1.906
	90%	10	3.021	2.503
		20	2.565	2.181
		30	2.413	2.064
	95%	10	3.381	2.875
		20	2.752	2.378
		30	2.550	2.209

Situation	Interval
2-sided	$[\bar{x} \pm k_2 s]$
1-sided lower limit	$[\bar{x} - k_1 s, \infty]$
1-sided upper limit	$[-\infty,\ \bar{x} + k_1 s]$