Chapter 830

Tolerance Intervals for Normal Data

Introduction

This routine calculates the sample size needed to obtain a specified coverage of a β-content tolerance interval at a stated confidence level for data from the normal distribution and data without a specified distribution. These intervals are constructed so that they contain at least 100β% of the population with probability of at least 100(1 - α)%. For example, in water management, a drinking water standard might be that one is 95% confident that certain chemical concentrations are not exceeded more than 3% of the time.

Difference Between a Confidence Interval and a Tolerance Interval

It is easy to get confused about the difference between a *confidence interval* and a *tolerance interval*. Just remember than a *confidence interval* is usually a probability statement about the value of a distributional parameter such as the mean or proportion. On the other hand, a *tolerance interval* is a probability statement about a proportion of the distribution from which the sample is drawn.

Technical Details

This section is primarily based on results in Guenther (1977). Other useful references are Guenther (1972), Hahn and Meeker (1991), Howe (1969), Krishnamoorthy and Mathew (2009), Odeh, Chou, and Owen (1987), and Young, Gordon, Zhu, and Olin (2016).

A tolerance interval is constructed from a random sample so that a specified proportion of the population is contained within the interval. The interval is defined by two limits, L_1 and L_2, which are constructed using

$$L_1 = \bar{x} - ks, L_2 = \bar{x} + ks$$

where $\bar{x}$ is the sample mean, s is the sample standard deviation, and k is calculated as described below.

Proportion of the Population Covered

An important concept is that of *coverage*. Coverage is the proportion of the population distribution that is between the two limits. For the normal distribution, these population limits are defined by

$$\mu \pm z_{\alpha/2}\sigma$$

The coverage is the area under the standard normal distribution between these limits.

Solving for N

The tolerance factor k is constructed so that

$$\Pr(L_2 - L_1 \geq P) = 1 - \alpha$$

Guenther (1977) provides the following approximation for k in the **two-sided case**

$$k = uvw$$

where

$$u = z_{(1+P)/2}\sqrt{\frac{N+1}{N}}$$

$$v = \sqrt{\frac{N-1}{\chi^2_{N-1,\alpha}}}$$

$$w = \sqrt{1 + \frac{N - 3 - \chi^2_{N-1,\alpha}}{2(N+1)^2}}$$

Using this value of k, the minimum value of N is found by increasing N by one at each iteration and stopping when

$$\frac{\chi^2_{N-1,\alpha\prime}}{\chi^2_{N-1,\alpha}} \leq \left[\frac{z_{(1+P\prime)/2}}{z_{(1+P)/2}}\right]^2$$

Guenther (1977) also provides an algorithm for finding N in the **one-sided case**.

$$k_1 = \frac{T\prime_{N-1,1-\alpha,z_P\sqrt{N}}}{\sqrt{N}}$$

$$k_2 = \frac{T\prime_{N-1,\alpha\prime,z_{P\prime}\sqrt{N}}}{\sqrt{N}}$$

Where $T\prime$ is the noncentral T distribution.

Starting at $N = 2$, increase N by one until $k_1 \leq k_2$.

Example 1 – Calculating Sample Size

Suppose a study is planned to determine the sample size required to compute a two-sided 95% normal tolerance interval the covers 90% of the population. The researchers want to investigate using a δ of 0.01, 0.025, or 0.05 with an α' of 0.05.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 1** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Sample Size**
Interval Type ....**Two-Sided Tolerance Interval**
Proportion Covered (P) **0.9**
Confidence Level (1 – α)...............................**0.95**
Coverage Proportion Exceedance (δ)............**0.01 0.025 0.05**
α' = Pr(p ≥ P + δ)...**0.05**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Reports

Numeric Results

Solve For: Sample Size N
Interval Type: Two-Sided Tolerance Interval

| Confidence Level 1 - α | Sample Size N | Proportion of Population Covered | | | Pr(p ≥ P + δ) α' | Tolerance Factor k |
		Value P	Exceedance Margin δ	Upper Limit P + δ		
0.95	5910	0.9	0.010	0.910	0.05	1.6703
0.95	866	0.9	0.025	0.925	0.05	1.7138
0.95	179	0.9	0.050	0.950	0.05	1.8084

1 - α	Confidence Level. The proportion of studies with the same settings that produce tolerance intervals with a proportion covered of at least P.
N	The number of subjects.
P	The proportion of the population covered. It is the probability between the tolerance interval limits based on the normal distribution.
δ	Proportion Covered Exceedance Margin. The value that is added to P to set an upper bound on the coverage at P + δ.
P + δ	The upper limit of the proportion covered, P. It is a measure of the precision (closeness) of the actual coverage to P.
p	The value of P computed from a random sample.
α'	The probability that the sample value p is greater than P + δ. It is set to a small value such as 0.05 or 0.01. $\alpha' = \Pr(p \geq P + \delta)$.
k	The tolerance factor used in calculating the tolerance interval from a sample. The sample tolerance interval is Mean ± k (SD).

Summary Statements
───

A single-group design will be used to obtain a two-sided 95% tolerance interval where the target proportion of the population covered is 0.9. The underlying data are assumed to follow a normal distribution. To produce a tolerance interval where the probability that the coverage (0.9) is exceeded by more than 0.01 is 0.05, 5910 subjects will be needed. The factor k for computing the tolerance limits is 1.6703.
───

Dropout-Inflated Sample Size
───

Dropout Rate	Sample Size N	Dropout-Inflated Enrollment Sample Size N'	Expected Number of Dropouts D
20%	5910	7388	1478
20%	866	1083	217
20%	179	224	45

───

Dropout Rate	The percentage of subjects (or items) that are expected to be lost at random during the course of the study and for whom no response data will be collected (i.e., will be treated as "missing"). Abbreviated as DR.
N	The evaluable sample size at which the tolerance interval is computed. If N subjects are evaluated out of the N' subjects that are enrolled in the study, the design will achieve the stated tolerance interval.
N'	The total number of subjects that should be enrolled in the study in order to obtain N evaluable subjects, based on the assumed dropout rate. After solving for N, N' is calculated by inflating N using the formula N' = N / (1 - DR), with N' always rounded up. (See Julious, S.A. (2010) pages 52-53, or Chow, S.C., Shao, J., Wang, H., and Lokhnygina, Y. (2018) pages 32-33.)
D	The expected number of dropouts. D = N' - N.

Dropout Summary Statements
───

Anticipating a 20% dropout rate, 7388 subjects should be enrolled to obtain a final sample size of 5910 subjects.
───

References
───
Faulkenberry, G.D. and Weeks, D.L. 1968. 'Sample Size Determination for Tolerance Limits.' Technometrics, Vol.
    10, No. 2, Pages 343-348.
Guenther, William C. 1972. 'Tolerance Intervals for Univariate Distributions.' Naval Research Logistics Quarterly,
    Vol. 19, No. 2, Pages 309-333.
Guenther, William C. 1977. Sampling Inspection in Statistical Quality Control. Griffin's Statistical Monographs,
    Number 37. London.
Hahn, G. J. and Meeker, W.Q. 1991. Statistical Intervals. John Wiley & Sons. New York.
Howe, W.G. 1969. 'Two-Sided Tolerance Limits for Normal Populations-Some Improvements.' Journal of the
    American Statistical Association, Vol. 64, No. 326, Pages 610-620.
Krishnamoorthy, K. and Mathew, T. 2009. Statistical Tolerance Regions. John Wiley, New York.
Odeh, R.E., Chou, Y.M., and Owen, D.B. 1987. 'The Precision for Coverages and Sample Size Requirements for
    Normal Tolerance Intervals.' Communications in Statistics – Simulation and Computation, 16(4), Pages 969-985.
Young, D.S., Gordon, C.M, Zhu, S., and Olin, B.D. 2016. 'Sample size determination strategies for normal
    tolerance intervals using historical data.' Quality Engineering, Volume 28(3), Pages 337-351.
───

This report shows the calculated sample size for each of the scenarios.

Plots Section

Plots
───
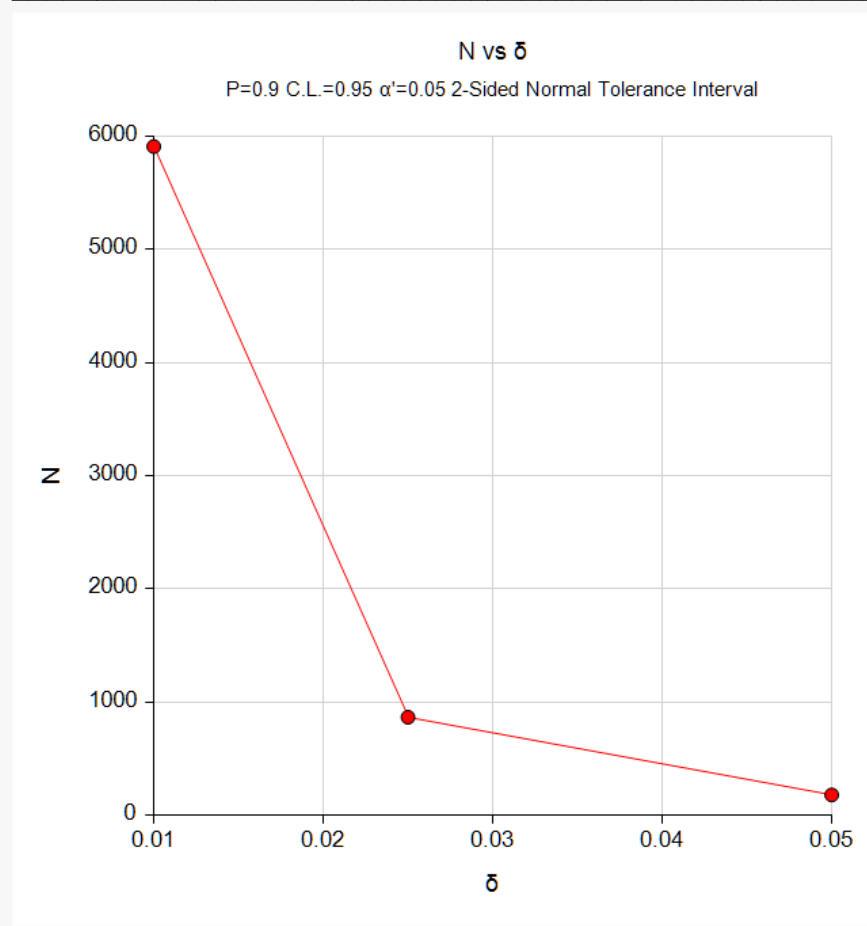
This plot shows the sample size versus the three values of δ.

Example 2 – Calculating α′

Continuing Example 1, the researchers want to show the impact of various sample sizes on α'. They decide to determine the value of α' for various value of N between 200 and 1600, keeping the other values the same except that they set δ to 0.025.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 2** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....α' = Pr(p ≥ P + δ)
Interval Type ....**Two-Sided Tolerance Interval**
N (Sample Size)..**200 to 1600 by 200**
Proportion Covered (P) **0.9**
Confidence Level (1 – α)................................**0.95**
Coverage Proportion Exceedance (δ)............**0.025**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Reports

Numeric Results

Solve For: α' = Pr(p ≥ P + δ)
Interval Type: Two-Sided Tolerance Interval

| Confidence Level 1 - α | Sample Size N | Proportion of Population Covered | | | Pr(p ≥ P + δ) α' | Tolerance Factor k |
		Value P	Exceedance Margin δ	Upper Limit P + δ		
0.95	200	0.9	0.025	0.925	0.545	1.7984
0.95	400	0.9	0.025	0.925	0.287	1.7493
0.95	600	0.9	0.025	0.925	0.140	1.7287
0.95	800	0.9	0.025	0.925	0.065	1.7168
0.95	1000	0.9	0.025	0.925	0.029	1.7088
0.95	1200	0.9	0.025	0.925	0.012	1.7029
0.95	1400	0.9	0.025	0.925	0.005	1.6984
0.95	1600	0.9	0.025	0.925	0.002	1.6948

This report shows the impact on α' of various sample sizes. Since the value of the *Tolerance Factor k* is not related to α' or δ, this report allows you to calculate *k* for use with sample data.

Plots Section
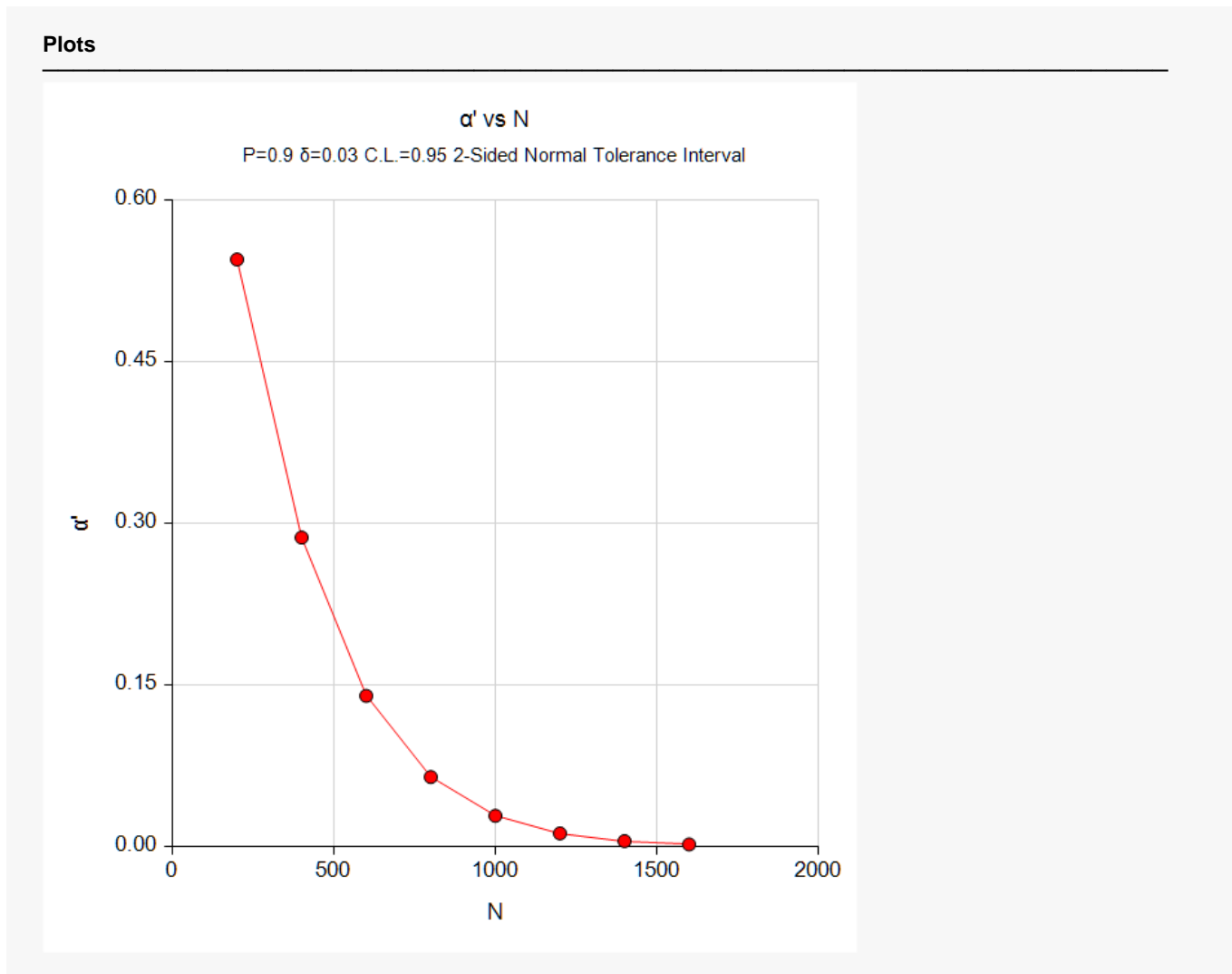
Plots
_____

This plot shows the sample size versus α'.

Example 3 – Validation using Guenther (1977)

Guenther (1977) page 174 gives an example in which P = 0.8, 1 – α = 0.9, $P + \delta$ = 0.95, and α' = 0.05. He obtains a sample size of 26 and a k of 1.613.

Setup

If the procedure window is not already open, use the PASS Home window to open it. The parameters for this example are listed below and are stored in the **Example 3** settings file. To load these settings to the procedure window, click **Open Example Settings File** in the Help Center or File menu.

Design Tab

Solve For ....**Sample Size**
Interval Type...**Two-Sided Tolerance Interval**
Proportion Covered (P) **0.8**
Confidence Level (1 – α)................................**0.9**
Coverage Proportion Exceedance (δ)............**0.15**
α' = Pr(p ≥ P + δ)...**0.05**

Output

Click the Calculate button to perform the calculations and generate the following output.

Numeric Results

Solve For: Sample Size N
Interval Type: Two-Sided Tolerance Interval

| | | Proportion of Population Covered | | | | |
Confidence Level 1 - α	Sample Size N	Value P	Exceedance Margin δ	Upper Limit P + δ	Pr(p ≥ P + δ) α'	Tolerance Factor k
0.9	26	0.8	0.15	0.95	0.05	1.6124

PASS also calculates the sample size to be 26. The value of k matches to within rounding.