

Chapter 813

Nonparametric Reference Intervals for Non-Normal Data

Introduction

This routine calculates the sample size needed to obtain a specified relative width of a confidence interval about the limits of a *nonparametric* reference interval based on ranks of the data.

Consider a measurement made on a population of individuals (usually healthy patients). A **reference interval** (RI) of this measurement gives the boundaries between which a typical measurement is expected to fall. When a measurement occurs that is outside these reference interval boundaries, there is cause for concern. That is, the measurement is unusually high or low. The reference interval is often presented as percentiles of a reference population, such as the 2.5th percentile and the 97.5th percentile. Of course, the choice of the reference population is important and you would expect that there is often differences according to age, size, and so on. Hence, the population must be partitioned into homogeneous subpopulations and individual reference intervals determined for each.

This procedure calculates sample sizes based on the methods shown in Machin, Campbell, Tan, and Tan (2018). These methods are for data that is non-normal and can't be easily transformed to normal. The method is nonparametric in that only the ranks of the data are used.

Technical Details

Reference Interval

Assume that a sample of N observations of an i.i.d. random value Y are ranked from smallest to largest. Refer to these ranked values as $y[1], y[2], \dots, y[N]$. The $100(1 - \alpha)\%$ *reference interval* is given by the *reference limits* R_L and R_U

where

$$R_L = y[N\alpha/2] \text{ and } R_U = y[N(1 - \alpha/2)].$$

Actually, the limits are usually formed by some form of interpolation since $N\alpha/2$ will seldom result in an integer.

Nonparametric Reference Intervals for Non-Normal Data

Confidence Intervals of Reference Limits

The $100(1 - \gamma)\%$ confidence intervals for each of the reference limits are each estimated from the follow ranks based on quantiles q

$$r_q = Nq - \left[z_{1-\frac{\gamma}{2}} \sqrt{Nq(1-g)} \right]$$

$$s_q = 1 + Nq + \left[z_{1-\frac{\gamma}{2}} \sqrt{Nq(1-g)} \right]$$

where q is set to $\alpha/2$ for the C.I. of R_L and to $1-\alpha/2$ for the C.I. of R_U .

Relative Margin of Error

The relative width (ignoring the standard deviation) of the above confidence interval is approximated by

$$W_{CI} = 2 \eta z_{1-\frac{\gamma}{2}} / \sqrt{N}$$

where

$$\eta = \frac{\sqrt{\frac{\gamma(1-\gamma)}{2}}}{\phi}$$

$$\phi = \frac{\exp\left(-\frac{1}{2} z_{1-\frac{\gamma}{2}}^2\right)}{\sqrt{2\pi}}$$

By rearranging the sample size formula on page 343 of Machin, Campbell, Tan, and Tan (2018), the margin of error is given by

$$ME = 100 \sqrt{\frac{\sqrt{3}\eta}{N} \left(\frac{z_{1-\frac{\gamma}{2}}}{z_{1-\frac{\alpha}{2}}} \right)^2}$$

Finally, the sample size is obtained through a binary search of possible values.

Procedure Options

This section describes the options that are specific to this procedure. These are located on the Design tab. For more information about the options of other tabs, go to the Procedure Window chapter.

Design Tab

The Design tab contains most of the parameters and options that you will be concerned with.

Solve For

Solve For

This option specifies the parameter (sample size or margin of error) to be solved for from the other parameters.

Nonparametric Reference Intervals for Non-Normal Data

Percentage of the Population in the Reference Interval

RI (Reference Interval Percentage)

Enter one or more values of the reference interval percentage. This is the percentage of data values that are anticipated to fall between the reference limits. These are nonparametric limits based on the ranked data.

Recommended

Usually, this value is set at 95. Additionally, 90 and 99 are also used.

Range

$0 < RI < 100$.

Percentage Confidence Level of the CI of the RI

CL (Confidence Level Percentage)

Enter one or more values of the confidence level percentage. This is the confidence level of a confidence interval made for each reference limit. These are nonparametric limits based on the ranked data.

Recommended

Usually, this value is set to 90.

Range

$50 < CL < 100$.

Relative Margin of Error – Percentage that CI Width is of RI Width

ME (Percentage Margin of Error)

Enter one or more values of the margin of error percentage. This is an estimate of the percentage that the width of the confidence interval of the reference limits is of the width of the reference interval. This ratio is expressed as a percentage. That is, the actual ratio is multiplied by 100.

Range

Some authors have recommended that this value be between 10 and 30.

$0 < ME < 50$.

Sample Size

N (Sample Size)

Enter one or more values for the sample size. This is the number of individuals selected at random from the population to be in the study. For nonparametric reference intervals, authors often recommend a minimum 120 subjects.

You can enter a single value, a list of values, or a range of values.

Example 1 – Calculating Sample Size

Suppose a study is being planned to obtain 90%, confidence intervals of the 95% reference limits of blood pressure (BP) for a certain gender and age range. Sample sizes are need for relative margin of errors from 10% to 30%. The goal is to determine the necessary sample size for each value of the relative margin of error.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Nonparametric Reference Intervals for Non-Normal Data** procedure. You may then make the appropriate entries as listed below, or open **Example 1** by going to the **File** menu of the procedure window and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	N (Sample Size)
RI (Reference Interval Percentage).....	95
CL (Confidence Level Percentage)	90
ME (Percentage Margin of Error)	10 20 30

Annotated Output

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

Numeric Results

Numeric Results

Sample Size	Margin of Error 100 WCL/WRI	Reference Interval Percentage	Confidence Level of the C.I. of the Reference Limits	Relative Width of CL WCL	Relative Width of RI WRI
N	ME	RI	CL	WCL	WRI
258	10.0	95	90	0.433	4.330
65	19.9	95	90	0.862	4.330
29	29.8	95	90	1.291	4.330

Report Definitions

N is the total sample size of the study.

ME is the margin of error which is a measure of imprecision. It is the percentage that the width of the confidence interval of the reference limits is of the width of the reference interval.

RI is the reference interval percentage. It is the percentage of data values that are anticipated to be between the reference limits.

CL is the confidence level of a confidence interval made for each reference limit. It is a percentage.

WCL is the relative width of the confidence interval made for each reference limit. It ignores the value of the standard deviation.

WRI is the relative width of the reference interval. It ignores the value of the standard deviation.

References

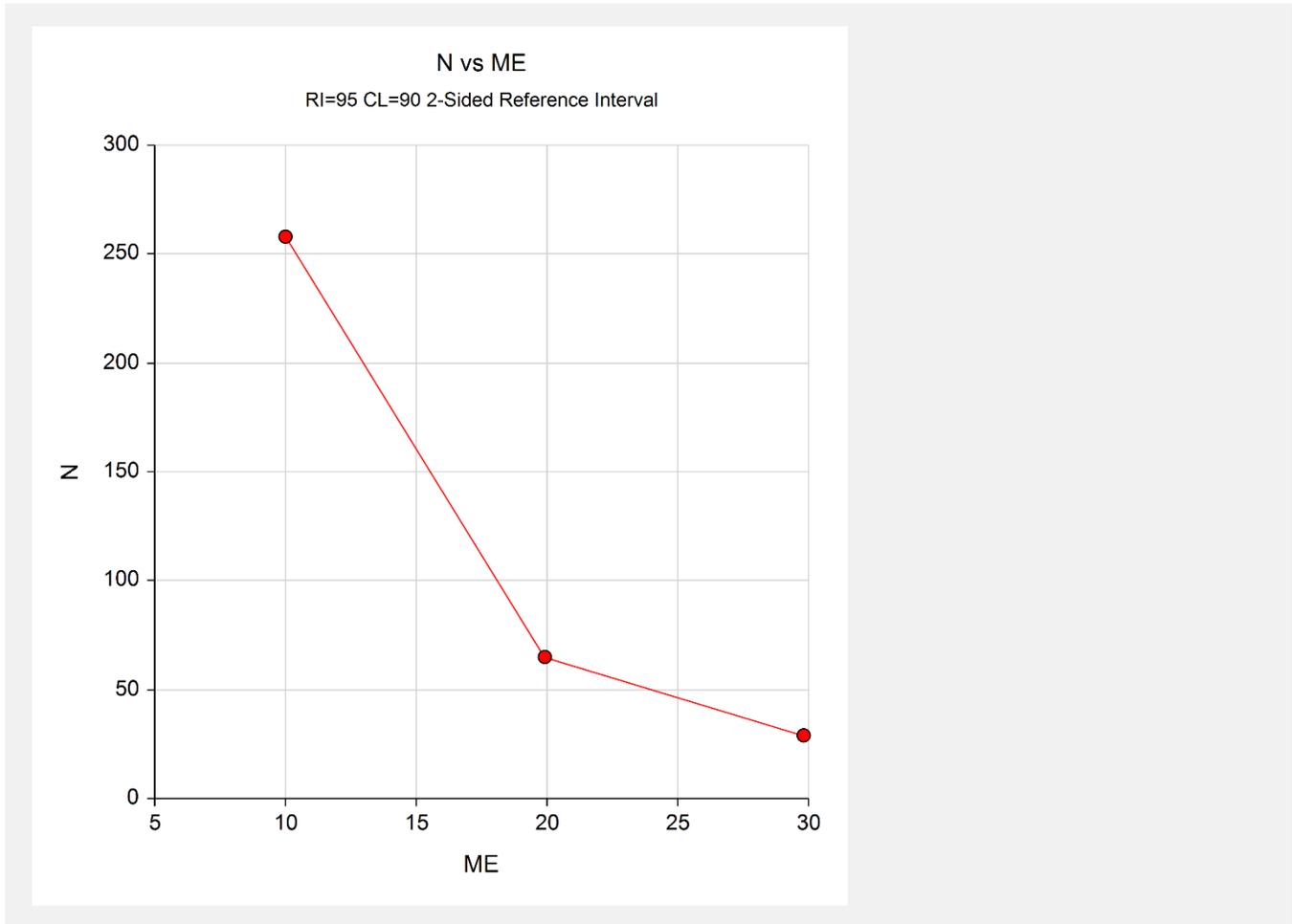
Machin, D, Campbell, M.J., Tan, S.B, Tan, S.H. 2018. 'Sample Sizes for Clinical, Laboratory and Epidemiology Studies, Fourth Edition'. John Wiley and Sons. Hoboken, New Jersey.

Summary Statements

A sample size of 258 produces two-sided, nonparametric 90% confidence intervals of the two limits of a 95% reference interval. The relative width of each of the confidence intervals is 0.433. The relative width of the reference interval is 4.330. The ratio of these widths multiplied by 100, called the margin of error, is 10.0.

Nonparametric Reference Intervals for Non-Normal Data

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

Plots Section

This plot shows the sample size versus the margin of error (MoE).

Example 2 – Validation using Machin *et al.* (2018)

Machin *et al.* (2018) page 351 give Table 21.2 which contains sample sizes for various scenarios. We will duplicate the first row of this table as a validation example. In this example RI = 90; CI = 80, 90, 95; and ME = 2.5. The sample sizes are computed as 2876, 5857, and 10512.

Setup

This section presents the values of each of the parameters needed to run this example. First, from the PASS Home window, load the **Nonparametric Reference Intervals for Non-Normal Data** procedure. You may then make the appropriate entries as listed below, or open **Example 2** by going to the **File** menu of the procedure window and choosing **Open Example Template**.

<u>Option</u>	<u>Value</u>
Design Tab	
Solve For	N (Sample Size)
RI (Reference Interval Percentage).....	90
CL (Confidence Level Percentage)	80 90 95
ME (Percentage Margin of Error)	2.5

Output

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

Numeric Results

Numeric Results						
Sample Size	Margin of Error	Reference Interval Percentage	Confidence Level of the C.I. of the Reference Limits	Relative Width of CL	Relative Width of RI	
N	ME	RI	CL	WCL	WRI	
2876	2.5	90	80	0.082	3.268	
5857	2.5	90	90	0.091	3.634	
10512	2.5	90	95	0.102	4.085	

PASS matches the first row of Table 21.2 on page 351 exactly.