Chapter 586

Reference Intervals

Introduction

A reference interval contains the middle 95% of measurements of a substance from a healthy population. It is a type of prediction interval. This procedure calculates one-, and two-, sided reference intervals using three different methods promoted by CLSI EP28-A3c: normal distribution, nonparametric-percentiles, or robust percentile estimators.

Horn and Pesce (2005) state that "The reference interval is the most widely used medical decision-making tool. It is central to the determination of whether or not an individual is healthy." This procedure allows one to study whether the sample meets the various assumptions needed for an accurate reference interval to be formed.

Technical Details

Let $X_1, X_2, \ldots, X_n$ be a random sample of size *n* from a population with distribution function $F(X)$. A two-sided, 100(1-α)% reference interval (R_L, R_U) for a new observation X_{new} is defined as

$$P[R_L \leq X_{new} \leq R_U] = 1 - \alpha$$

One-sided intervals are defined similarly.

CLSI EP28-A3c discusses three methods of computing these limits along with their confidence intervals. These are presented next using this document as well as Horn and Pesce (2005).

Normal - Theory Method

This method is based on traditional normal-theory. If the data are not normally distributed, you can try the *Box-Cox Transformation* procedure to determine if a power transformation will bring the distribution closer to normal.

The following formulation is given by Horne and Pesce (2005). The lower and upper limits of the reference interval are defined as

$$R_L = \bar{x} + t_{\frac{\alpha}{2}, n-1} s \sqrt{1 + \frac{1}{n}}$$

$$R_U = \bar{x} + t_{1-\frac{\alpha}{2}, n-1} s \sqrt{1 + \frac{1}{n}}$$

where $\bar{x}$ is the sample mean and *s* is the sample standard deviation.

CLSI recommends 90% confidence intervals be calculated for the two reference limits. The formulas for these confidence intervals are

$$R_L \pm z_{\gamma/2} s_{\alpha/2}$$

and

$$R_U \pm z_{\gamma/2} s_{\alpha/2}$$

where

$$s_{\alpha/2} = s\sqrt{\frac{2 + z_{\alpha/2}^2}{2n}}$$

$$\gamma = 0.90$$

Here, *z* is the standard normal variate.

Percentile Method

The following formulation for the percentile method is given by Horne and Pesce (2005). In this case, the lower and upper limits of the reference interval are defined as the $100\left(\frac{\alpha}{2}\right)$ and $100\left(1-\frac{\alpha}{2}\right)$ percentiles of the sorted data values.

There is some controversy over the definition of a percentile. NCSS provides you with five choices. CLSI recommend

$$\hat{F}(p) = (1-r)X_{(j)} + rX_{(j+1)}$$

where $Y_{(j)}$ is the j[th] ordered value, $j = [(n+1)p]$, $r = (n+1)p - j$, [z] is the integer part of z, and $X_{(n+1)} = X_{(n)}$.

CLSI recommends 90% confidence intervals be calculated for the two reference limits. The formula for the confidence interval of the lower reference limit is the interval $\left(x_{(l)}, x_{(r)}\right)$ where

$$\sum_{i=l}^{r-1} \left(\frac{\alpha}{2}\right)^i \left(1-\frac{\alpha}{2}\right)^{n-i}$$

The confidence interval for the upper reference limit is the interval $\left(x_{(n-r+1)}, x_{(n-l+1)}\right)$ were *l* and *r* are defined above.

Robust Method

The robust algorithm is given in Appendix B of CLSI EP28-A3c. This is a rather long algorithm and it is not repeated here.

Confidence intervals for the two limits are calculated using the percentile bootstrap method. This method requires a medium to large (not small!) sample size.

Data Structure

The data are contained in a single column.

Procedure Options

This section describes the options available in this procedure. To find out more about using a procedure, turn to the Procedures chapter.

Following is a list of the procedure's options.

Variables Tab

The options on this panel specify which variables to use.

Data Variables

Variables

Specify a list of one or more variables for which reference intervals are to be generated. You can double-click the field or single click the button on the right of the field to bring up the Variable Selection window.

Group Variable

You can specify an optional grouping variable. When specified, a separate line on each report is generated for each unique value of this variable.

Frequency (Count) Variable

Frequency Variable

This optional variable specifies the number of observations that each row represents. When omitted, each row represents a single observation. If your data is the result of a previous summarization, you may want certain rows to represent several observations. Note that negative values are treated as a zero count and are omitted.

Reference Interval Options

Type of Limit(s)

Specify whether a two-sided, a lower one-sided, or an upper one-sided reference interval is to be reported. The reference interval gives percentiles between which a specified percentage of the reference population is expected to lie.

Two-Sided Reference Interval

Find two limits between which a specified percentage of the reference population is expected to lie.

One-Sided Upper Reference Bound

Find the upper bound, below which a specified percentage of the reference population is expected to lie.

One-Sided Lower Reference Bound

Find the lower bound, above which a specified percentage of the reference population is expected to lie.

Reference Interval Input

Specify how the limits of the reference interval are specified.

Specify the Percentage of the Population in the Interval

This option lets you specify a single value percentage value that gives the percentage in the interval.

For example, if you want to calculate a 95% reference interval, you would enter only the 95 in the box below. The remaining 5% will be divided equally so that the resulting limits are at 2.5% and 97.5%.

Specify the Lower and Upper Percentile Limits Individually

This option lets you specify the lower and upper percentile limits directly.

Percentage in Interval

Specify the percentage of the reference population that is expected to lie between the upper and lower limits. These limits will be positioned so that the reference interval is centered between them.

The practical range is from 50 to 99.

For example, if you enter 95 here, the limits will be set to 2.5 and 97.5.

Lower Percentile Limit

Specify the lower percentile limit. This is specified as a percentage. It gives the smallest percentile that is still part of the normal (non-diseased) range.

Commonly, this value is set to 2.5. It should be between 0 and 50.

Upper Percentile Limit

Specify the upper percentile limit. This is specified as a percentage. It gives the largest percentile that is still part of the normal (non-diseased) range.

Commonly, this value is set to 97.5. It should be between 50 and 100.

Percentile Type

This option specifies which of five different methods is used to calculate the percentiles. The recommend option is ***Ave Xp(n+1)*** since it gives the common value of the median and is recommended by CLSI.

In the options below

"p" refers to the fractional value of the percentile (for example, for the 75th percentile p = .75)

"Zp" refers to the value of the percentile

"X[i]" refers to the ith data value after the values have been sorted

"n" refers to the total sample size

"g" refers to the fractional part of a number (for example, if np = 23.42, then g = .42)

Ave Xp(n+1)

This is the most commonly used option. It is recommended by CLSI EP28-A3 and Harris and Boyd (1995). The 100pth percentile is computed as

Zp = (1-g)X[k1] + gX[k2]

where k1 equals the integer part of p(n+1), k2=k1+1, g is the fractional part of p(n+1), and X[k] is the kth observation when the data are sorted from lowest to highest.

Ave Xp(n)

The 100pth percentile is computed as

Zp = (1-g)X[k1] + gX[k2]

where k1 equals the integer part of np, k2=k1+1, g is the fractional part of np, and X[k] is the kth observation when the data are sorted from lowest to highest.

Closest to np

The 100pth percentile is computed as

Zp = X[k1]

where k1 equals the integer that is closest to np and X[k] is the kth observation when the data are sorted from lowest to highest.

EDF

The 100pth percentile is computed as

Zp = X[k1]

where k1 equals the integer part of np if np is exactly an integer or the integer part of np+1 if np is not exactly an integer. X[k] is the kth observation when the data are sorted from lowest to highest. Note that EDF stands for empirical distribution function.

EDF w/Ave

The 100pth percentile is computed as

Zp = (X[k1] + X[k2])/2

where k1 and k2 are defined as follows: If np is an integer, k1=k2=np. If np is not exactly an integer, k1 equals the integer part of np and k2 = k1+1. X[k] is the kth observation when the data are sorted from lowest to highest. Note that EDF stands for empirical distribution function.

Confidence Coefficient

Specify the value of confidence coefficient of the confidence intervals of the reference limits and reference bounds. This value is specified as a percentage. CLSI recommends 90% confidence intervals for 95% reference intervals. The range is between 70 and 99.

Data Transformation to Achieve Normality

Power Transformation

Occasionally, you might want to obtain a statistical report on the square root or log of your variable. This option lets you specify an on-the-fly transformation of the variable. The form of this transformation is $X = Y^A$, where Y is the original value, A is the selected exponent (power), and X is the resulting value.

Additive Constant

Occasionally, you might want to obtain a statistical report on a transformed version of a variable. This option lets you specify an on-the-fly transformation of the variable. The form of this transformation is $X = Y+B$, where Y is the original value, B is the specified constant, and X is the value that results.

Note that if you apply both the *Power Transformation* and the *Additive Constant* the form of the transformation is $X = \left(Y + B\right)^A$.

Reports Tab

The options on this panel control the reports and plots displayed.

Select Reports

Descriptive Statistics ... Reference Intervals: Robust

Indicate whether to display the indicated reports.

If Robust Reference Intervals are checked

Bootstrap Confidence Intervals

This option provides confidence intervals for the reference limits and bounds. Accurate bootstrap require a large sample size which may in turn require a long run-time.

Samples (N)

This is the number of bootstrap samples used. We recommend using at least 3000 samples. With current computer speeds, 10000 or more samples can often be run in a short time.

Tuning Constant 1

The is the value of the robust tuning constant, c1. CLSI EP28-A3c and Horn and Pesce (2005) recommend c1 = 3.7.

Tuning Constant 2

Specify whether c2 is calculated or set by you.

Automatic

Calculate c2 = 1/(0.581734 - 0.607227 (1-α)), where α = 1 - level of a two-sided reference interval. This formula works for $0.05 \leq \alpha \leq 0.5$.

For one-sided test, replace 1-α with (2p-1) or (1-2p).

For, 90% reference interval, α = 0.1 and c2 = 28.4.

For 95% reference interval, α=0.05 and c2 = 205.4.

Custom

Enter your own value directly. You would use this when $\alpha < 0.05$ or you are using one-sided intervals.

Tuning Constant 2

Specify a value for c2. This value depends on the reference interval level and whether this is a one-sided or two-sided interval.

Common values are

90% R.I., 2-sided: c2 = 28.4.

95% R.I., 2-sided: c2 = 205.4.

98% R.I., 2-sided: c2 = 1812.

90% R.I., lower one-sided: c2 = 10.4

90% R.I., upper one-sided: c2 = 10.4

95% R.I., lower one-sided: c2 = 28.4

95% R.I., upper one-sided: c2 = 28.4

97.5% R.I., lower one-sided: c2 = 205.4

97.5% R.I., upper one-sided: c2 = 205.4

Stop Iterating When Change in Mean ≤

This option specifies a stopping value for the iteration procedure. If the percentage change in the mean is less than this amount, the iteration procedure is stopped.

If you want this option to be ignored, set it to zero.

The recommended value is 0.001

Stop Iterating When Iterations >

Specifies the maximum number of iterations allowed while finding a robust solution. If this number is reached, the procedure is terminated. The recommended value is 10.

Report Options

Precision

Specify the precision of numbers in the report. A single-precision number will show seven-place accuracy, while a double-precision number will show thirteen-place accuracy. The double-precision option only works when the Decimals option is set to General.

Note that the reports were formatted for single precision. If you select double precision, some numbers may run into others. Also note that all calculations are performed in double precision regardless of which option you select here. This is for reporting purposes only.

Variable Names

This option lets you select whether to display only variable names, variable labels, or both.

Value Labels

This option applies to the *Group Variable*. It lets you select whether to display data values, value labels, or both. Use this option if you want the output to automatically attach labels to the values (like 1=Yes, 2=No, etc.). See the section on specifying *Value Labels* elsewhere in this manual.

Report Options – Decimal Places

Reference Limits ... P-Values Decimals

Specify the number of digits after the decimal point to display on the output of values of this type. Note that this option in no way influences the accuracy with which the calculations are done.

Enter 'General' to display all digits available. The number of digits displayed by this option is controlled by whether the PRECISION option is SINGLE or DOUBLE.

Plots Tab

The options on this panel control the appearance of the histogram and probability plot.

Select Plots

Histogram and Probability Plot

Indicate whether to display these plots. Click the plot format button to change the plot settings.

Example 1 – Generating Percentile Reference Intervals

This section presents a detailed example of how to generate nonparametric-percentile reference intervals for the *Calcium* variable in the Calcium dataset. This dataset contains 120 calcium measurements from males and 120 calcium measurements from females. To run this example, take the following steps:

1 Open the Calcium dataset.

- From the **File** menu of the NCSS Data window, select **Open Example Data**.
- Click on the file **Calcium.NCSS**.
- Click **Open**.

2 Open the Reference Intervals window.

- Using the Analysis menu or the Procedure Navigator, find and select the **Reference Intervals** procedure.
- On the menus, select **File**, then **New Template**. This will fill the procedure with the default template.

3 Specify the options on the Variables tab.

- Select the **Variables tab**. (This is the default.)
- Double-click in the **Variables** text box. This will bring up the variable selection window.
- Select **Calcium** from the list of variables and then click **Ok**.
- Double-click in the **Group Variable** text box. This will bring up the variable selection window.
- Select **Gender** from the list of variables and then click **Ok**.

4 Specify the options on the Reports tab.

- Select the **Reports tab**.
- Check the **Descriptive Statistics**, **Normality**, **Quantiles**, and all three reference interval reports.

5 Run the procedure.

- From the Run menu, select **Run Procedure**. Alternatively, just click the green Run button.

The following reports and charts will be displayed in the Output window.

Descriptive Statistics

Descriptive Statistics of Calcium

Gender	Count	Mean	Median	Standard Deviation	IQR	Minimum	Maximum
Men	120	9.700	9.700	0.3272	0.500	9.000	10.600
Women	120	9.474	9.500	0.2926	0.400	8.700	10.300
Combined	240	9.587	9.600	0.3298	0.400	8.700	10.600

This report gives a statistical summary of the data. The “Combined” line gives the values for all groups combined.

Count

This is the number of nonmissing values. If no frequency variable was specified, this is the number of nonmissing rows.

Mean

This is the average of the data values.

Median

This is the median of the data values.

Standard Deviation

This is the standard deviation of the data values.

IQR

This is the interquartile range. It is the difference between the third quartile and the first quartile (between the 75th percentile and the 25th percentile). This represents the range of the middle 50 percent of the distribution. It is a very robust (not affected by outliers) measure of dispersion. In fact, if the data are normally distributed, a robust estimate of the sample standard deviation is IQR/1.35. If a distribution is very concentrated around its mean, the IQR will be small. On the other hand, if the data are widely dispersed, the IQR will be much larger.

Minimum

The smallest value in this variable.

Maximum

The largest value in this variable.

Normality Report

Normality Report of Calcium

Gender	Mean	Standard Deviation	COV	Skewness (Normal=0)	Kurtosis (Normal=3)	Anderson Darling Normality P-Value	Shapiro Wilk Normality P-Value
Men	9.700	0.3272	0.0337	0.0145	2.5092	0.0223	0.0548
Women	9.474	0.2926	0.0309	0.0481	3.0217	0.0434	0.3286
Combined	9.587	0.3298	0.0344	0.1247	2.7586	0.0029	0.0383

This report gives statistics that help you evaluation the normality assumption.

Mean

This is the average of the data values.

Standard Deviation

This is the standard deviation of the data values.

COV

The *coefficient of variation* is a relative measure of dispersion. It is most often used to compare the amount of variation in two samples. It can be used for the same data over two time periods or for the same time period but two different places. It is the standard deviation divided by the mean:

$$COV = s/\bar{x}$$

Skewness (Normal = 0)

This statistic measures the direction and degree of asymmetry. A value of zero indicates a symmetrical distribution. A positive value indicates skewness (longtailedness) to the right while a negative value indicates skewness to the left. Values between -3 and +3 indicate are typical values of samples from a normal distribution.

$$\sqrt{b_1} = \frac{m_3}{m_2^{3/2}}$$

Kurtosis (Normal = 3)

This statistic measures the heaviness of the tails of a distribution. The usual reference point in kurtosis is the normal distribution. If this kurtosis statistic equals three and the skewness is zero, the distribution is normal. Unimodal distributions that have kurtosis greater than three have heavier or thicker tails than the normal. These same distributions also tend to have higher peaks in the center of the distribution (leptokurtic). Unimodal distributions whose tails are lighter than the normal distribution tend to have a kurtosis that is less than three. In this case, the peak of the distribution tends to be broader than the normal (platykurtic). Be forewarned that this statistic is an unreliable estimator of kurtosis for small sample sizes.

$$b_2 = \frac{m_4}{m_2^2}$$

Shapiro-Wilk W Test

This test for normality has been found to be the most powerful test in most situations. It is the ratio of two estimates of the variance of a normal distribution based on a random sample of n observations. The numerator is proportional to the square of the best linear estimator of the standard deviation. The denominator is the sum of squares of the observations about the sample mean. The test statistic W may be written as the square of the Pearson correlation coefficient between the ordered observations and a set of weights which are used to calculate the numerator. Since these weights are asymptotically proportional to the corresponding expected normal order statistics, W is roughly a measure of the straightness of the normal quantile-quantile plot. Hence, the closer W is to one, the more normal the sample is.

The test was developed by Shapiro and Wilk (1965) for samples up to 20. NCSS uses the approximations suggested by Royston (1992) and Royston (1995) which allow unlimited sample sizes. Note that Royston only checked the results for sample sizes up to 5000, but indicated that he saw no reason larger sample sizes should not work.

The probability values for W are valid for samples greater than 3.

This test may not be as powerful as other tests when ties occur in your data.

Anderson-Darling Test

This test, developed by Anderson and Darling (1954), is the most popular normality test that is based on EDF statistics. In some situations, it has been found to be as powerful as the Shapiro-Wilk test.

Unfortunately, both the Shapiro-Wilk and Anderson-Darling tests have small statistical power (probability of detecting nonnormal data) unless the sample sizes are large, say over 100. Hence, if the decision is to reject, you can be reasonably certain that the data are not normal. However, if the decision is to accept, the situation is not as clear. If you have a sample size of 100 or more, you can reasonably assume that the actual distribution is closely approximated by the normal distribution. If your sample size is less than 100, all you know is that there was not enough evidence in your data to reject the normality assumption. In other words, the data might be nonnormal, you just could not prove it. In this case, you must rely on the graphics and past experience to justify the normality assumption.

Quantile Report

Quantile Report of Calcium

Gender	5th Pcntile	10th Pcntile	25th Pcntile	50th Pcntile	75th Pcntile	90th Pcntile	95th Pcntile
Men	9.105	9.200	9.500	9.700	10.000	10.100	10.200
Women	9.000	9.100	9.300	9.500	9.700	9.900	10.000
Combined	9.100	9.200	9.400	9.600	9.800	10.000	10.195

This report gives various percentiles of the data distribution.

Percentile Reference Interval

Two-Sided 95.0% Percentile Reference Interval of Calcium

		2.5% Lower Reference Limit			97.5% Upper Reference Limit		
			90% Conf Interval			90% Conf Interval	
Gender	Count	Value	Lower	Upper	Value	Lower	Upper
Men	120	9.100	9.000	9.200	10.300	10.200	10.600
Women	120	8.900	8.700	9.100	10.098	9.900	10.300
Combined	240	9.000	8.700	9.100	10.200	10.200	10.300

This report gives reference intervals and associated confidence intervals based on the percentile method.

Normal-Theory Reference Interval

Two-Sided 95.0% Normal-Theory Reference Interval of Calcium

		2.5% Lower Reference Limit			97.5% Upper Reference Limit		
			90% Conf Interval			90% Conf Interval	
Gender	Count	Value	Lower	Upper	Value	Lower	Upper
Men	120	9.049	8.965	9.133	10.351	10.267	10.435
Women	120	8.892	8.817	8.967	10.056	9.981	10.131
Combined	240	8.936	8.876	8.996	10.238	10.178	10.298

This report gives reference intervals and associated confidence intervals based on the normal-theory.

Robust Reference Interval

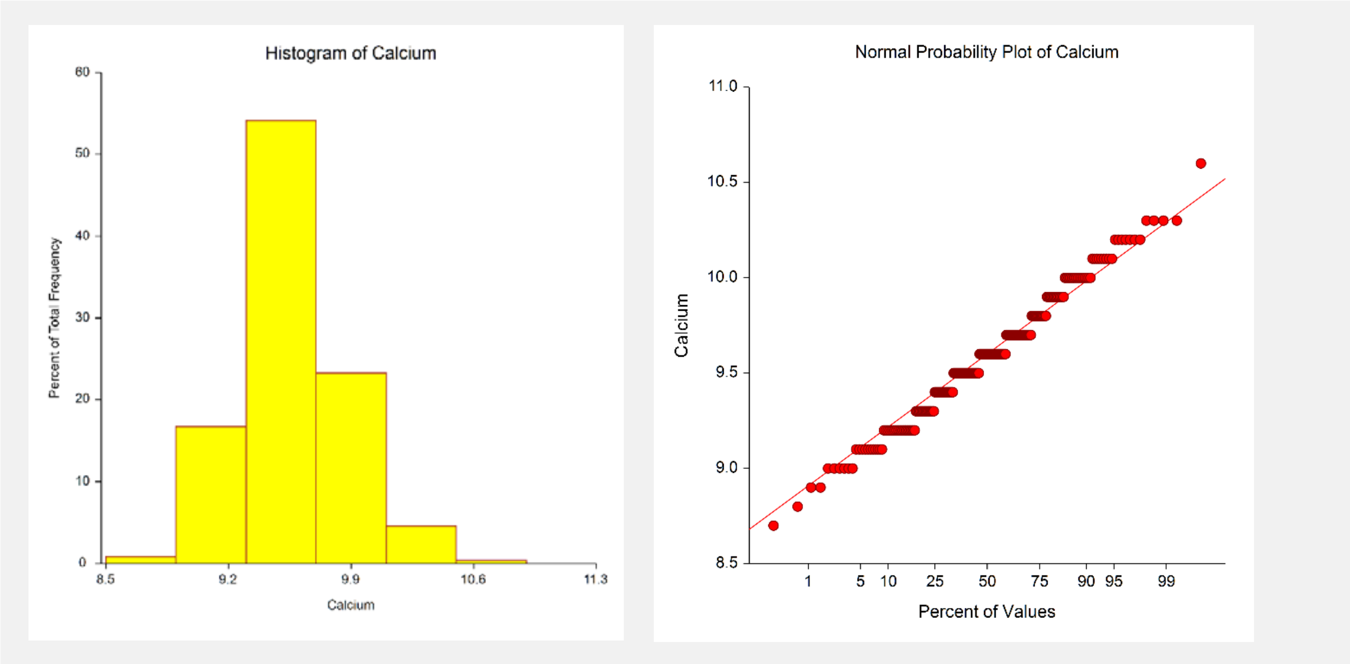
Two-Sided 95.0% Robust Reference Interval of Calcium

		2.5% Lower Reference Limit			97.5% Upper Reference Limit		
			90% Conf Interval			90% Conf Interval	
Gender	Count	Value	Lower	Upper	Value	Lower	Upper
Men	120	9.050	8.967	9.134	10.352	10.271	10.432
Women	120	8.888	8.810	8.967	10.057	9.982	10.135
Combined	240	8.927	8.867	8.987	10.230	10.173	10.291

Constants: c1 = 3.700 c2 = 205.600 MAD Scale Factor = 0.674500 Bootstrap Samples = 3000

This report gives reference intervals and associated confidence intervals based on the robust method.

Plots Section



The plots section displays a histogram and a probability plot for each line of the reports that let you assess the accuracy of the normality assumption.