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-$\alpha$)% 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

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

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

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

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 \( 100(\alpha/2) \) and \( 100(1 - \alpha/2) \) 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 \( (x_{(l)}, x_{(r)}) \) 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 \( (x_{(n-r+1)}, x_{(n-l+1)}) \) 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.
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:

Setup

To run this example, complete the following steps:

1  **Open the Calcium example dataset**
   - From the File menu of the NCSS Data window, select *Open Example Data*.
   - Select Calcium and click **OK**.

2  **Specify the Reference Intervals procedure options**
   - Find and open the Reference Intervals procedure using the menus or the Procedure Navigator.
   - The settings 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.

<table>
<thead>
<tr>
<th>Variables Tab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variables.................................................Calcium</td>
</tr>
<tr>
<td>Group Variable...........................................Gender</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reports Tab</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive Statistics.................................Checked</td>
</tr>
<tr>
<td>Normality ...................................................Checked</td>
</tr>
<tr>
<td>Quantiles ...................................................Checked</td>
</tr>
<tr>
<td>Reference Intervals using the Percentile Method ...........Checked</td>
</tr>
<tr>
<td>Reference Intervals using the Normal-Theory Method ......Checked</td>
</tr>
<tr>
<td>Reference Intervals using the Robust Method...............Checked</td>
</tr>
<tr>
<td>Random Seed..............................................3883571 (for reproducibility)</td>
</tr>
</tbody>
</table>

3  **Run the procedure**
   - Click the Run button to perform the calculations and generate the output.
Descriptive Statistics

Descriptive Statistics of Calcium

<table>
<thead>
<tr>
<th>Gender</th>
<th>Count</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
<th>IQR</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>120</td>
<td>9.700</td>
<td>9.700</td>
<td>0.3272</td>
<td>0.500</td>
<td>9.000</td>
<td>10.600</td>
</tr>
<tr>
<td>Women</td>
<td>120</td>
<td>9.474</td>
<td>9.500</td>
<td>0.2926</td>
<td>0.400</td>
<td>8.700</td>
<td>10.300</td>
</tr>
<tr>
<td>Combined</td>
<td>240</td>
<td>9.587</td>
<td>9.600</td>
<td>0.3298</td>
<td>0.400</td>
<td>8.700</td>
<td>10.600</td>
</tr>
</tbody>
</table>

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

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

This report gives statistics that help you evaluate 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:

\[
COV = \frac{s}{\bar{x}}
\]

### 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

<table>
<thead>
<tr>
<th>Gender</th>
<th>5th Pctl</th>
<th>10th Pctl</th>
<th>25th Pctl</th>
<th>50th Pctl</th>
<th>75th Pctl</th>
<th>90th Pctl</th>
<th>95th Pctl</th>
</tr>
</thead>
</table>

This report gives various percentiles of the data distribution.

Percentile Reference Interval

Two-Sided 95% Percentile Reference Interval of Calcium

<table>
<thead>
<tr>
<th>Gender</th>
<th>Count</th>
<th>Value</th>
<th>Lower</th>
<th>Upper</th>
<th>Value</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>120</td>
<td>9.100</td>
<td>9.000</td>
<td>9.200</td>
<td>10.300</td>
<td>10.200</td>
<td>10.600</td>
</tr>
<tr>
<td>Combined</td>
<td>240</td>
<td>9.000</td>
<td>8.700</td>
<td>9.100</td>
<td>10.200</td>
<td>10.200</td>
<td>10.300</td>
</tr>
</tbody>
</table>

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

Normal-Theory Reference Interval

Two-Sided 95% Normal-Theory Reference Interval of Calcium

<table>
<thead>
<tr>
<th>Gender</th>
<th>Count</th>
<th>Value</th>
<th>Lower</th>
<th>Upper</th>
<th>Value</th>
<th>Lower</th>
<th>Upper</th>
<th>Value</th>
<th>Lower</th>
<th>Upper</th>
<th>Value</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
</table>

This report gives reference intervals and associated confidence intervals based on the normal-theory.
# Robust Reference Interval

## Two-Sided 95% Robust Reference Interval of Calcium

<table>
<thead>
<tr>
<th>Gender</th>
<th>Count</th>
<th>Value</th>
<th>Lower</th>
<th>Upper</th>
<th>Value</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Women</td>
<td>120</td>
<td>8.888</td>
<td>8.809</td>
<td>8.969</td>
<td>10.057</td>
<td>9.981</td>
<td>10.132</td>
</tr>
</tbody>
</table>

Robust Method Constants: c1 = 3.700, c2 = 205.600, MAD Scale Factor = 0.674500
Bootstrap C.I. Estimation: Number of Samples = 3000, User-Entered Random Seed = 3883571

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

## Plots Section

**Histogram of Calcium**

**Normal Probability Plot of Calcium**
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.