

[Chapter Contents](#)
[Previous](#)
[Next](#)

Distribution Analyses

Trimmed and Winsorized Means

When outliers are present in the data, trimmed and Winsorized means are robust estimators of the population mean that are relatively insensitive to the outlying values. Therefore, trimming and Winsorization are methods for reducing the effects of extreme values in the sample. The k -times trimmed mean is calculated as

$$\bar{y}_{tk} = \frac{1}{n - 2k} \sum_{i=k+1}^{n-k} y_{(i)}$$

The trimmed mean is computed after the k smallest and k largest observations are deleted from the sample. In other words, the observations are trimmed at each end. The k -times Winsorized mean is calculated as

$$\bar{y}_{wk} = \frac{1}{n} \left\{ (k+1)y_{(k+1)} + \sum_{i=k+2}^{n-k-1} y_{(i)} + (k+1)y_{(n-k)} \right\}$$

The Winsorized mean is computed after the k smallest observations are replaced by the $(k+1)$ st smallest observation, and the k largest observations are replaced by the $(k+1)$ st largest observation. In other words, the observations are Winsorized at each end. For a symmetric distribution, the symmetrically trimmed or Winsorized mean is an unbiased estimate of the population mean. But the trimmed or Winsorized mean does not have a normal distribution even if the data are from a normal population.

The Winsorized sum of squared deviations is defined as

$$s_{wk}^2 = (k+1)(y_{(k+1)} - \bar{y}_{wk})^2 + \sum_{i=k+2}^{n-k-1} (y_{(i)} - \bar{y}_{wk})^2 + (k+1)(y_{(n-k)} - \bar{y}_{wk})^2$$

A robust estimate of the variance of the trimmed mean \bar{y}_{tk} can be based on the Winsorized sum of squared deviations (Tukey and McLaughlin 1963). The resulting trimmed t test is given by

$$t_{tk} = \frac{\bar{y}_{tk}}{\text{STDERR}(\bar{y}_{tk})}$$

where $\text{STDERR}(\bar{y}_{tk})$ is the standard error of \bar{y}_{tk} :

$$\text{STDERR}(\bar{y}_{tk}) = \frac{s_{wk}}{\sqrt{(n-2k)(n-2k-1)}}$$

A Winsorized t test is given by

$$t_{wk} = \frac{\bar{y}_{wk}}{\text{STDERR}(\bar{y}_{wk})}$$

where $\text{STDERR}(\bar{y}_{wk})$ is the standard error of \bar{y}_{wk} :

$$\text{STDERR}(\bar{y}_{wk}) = \frac{n-1}{n-2k-1} \frac{s_{wk}}{\sqrt{n(n-1)}}$$

When the data are from a symmetric distribution, the distribution of the trimmed t statistic t_{tk} or the Winsorized t statistic t_{wk} can be approximated by a Student's t distribution with $n-2k-1$ degrees of freedom (Tukey and McLaughlin 1963, Dixon and Tukey 1968). You can specify the number or percentage of observations to be trimmed or Winsorized from each end either by using the **Trimmed/Winsorized Means** options dialog or by using the **Trimmed/Winsorized Means** dialog after choosing **Tables:Trimmed/Winsorized Mean:(1/2)N** or **Tables:Trimmed/Winsorized Mean:(1/2)Percent** from the menus.