



Data Transformation

Data Transformation:

If a measurement variable does not fit a normal distribution or has greatly different standard deviations in different groups, you should go for data transformation.

Normality Test:


- 1) Skewness & Kurtosis
- 2) Shapiro Wilk test
- 3) Histogram

Procedure to test normality in SPSS

Go to analyze, descriptive → explore, take measurement in dependent and group in factor box → tick on plots, tick on histogram and normality plots with test, continue and ok.

Results

Divide the skewness and kurtosis by std. error and see if the values lies between -1.96 to +1.96 then the data are normally distributed (a minor difference is also normally distributed). See that the Shapiro wilk test is significant or not, if it is non-significant means data is normal otherwise data is not normal.



The following are the three transformations, which are being used most commonly, in biological research

- a) Logarithmic Transformation
- b) Square root Transformation
- c) Arc sine or Angular Transformation

Logarithmic transformation for whole number counts with wide range

- This transformation is suitable for the data where the variance is proportional to the square of the mean or the coefficient of the variation ($S.D./Mean$) is constant or where effects are multiplicative.
- These conditions are generally found in the data that are whole numbers and cover a wide range of values.
- For example- number of insect per plot, number of egg mass per plant or per unit area etc.

Natural logarithm of a positive number=LN (Give cell number for which transformation to be done),

Natural logarithm is based on the constant e (2.718281828845904)

Logarithm of a positive number at base 10= Log_{10} (Give cell number for which transformation to be done), Or $\text{Log}(\text{number}, 10)$

Square root transformation

- The transformation is appropriate for the data sets where the variance is proportional to the mean.

- Here the data consists of small whole numbers, for example data obtained in counting rare events, such as number of infested plants in a plot, the number of insects caught in traps, number of weeds per plot.

Square root transformation

- These data sets generally follow the Poisson distribution and square root transformation approximates Poisson to normal distribution.

- **Square root transformation=SQRT** (Give cell number for which transformation to be done)

Arcsine Transformation for Proportions or Percentage

- The transformation is appropriate for the data on proportion i.e. data obtained from a count and the data expressed as decimal fraction and percentage.
- The distribution of percentage is binomial and this transformation makes the distribution normal.

Arcsine Transformation=ASIN (Cell Identification e.g.

$(0.05)*(180)*(7/22)$ result will be in degrees.