Checking for Outliers with STUDENT Residuals

To check for outliers, you can use STUDENT residuals, also known as studentized or standardized residuals. STUDENT residuals are calculated by dividing the residual values by their standard errors. This puts them on a standard deviation scale. Because the residuals are normally distributed, we can use the properties of a normal distribution to help detect outliers.

A normally distributed variable has approximately 68% of the observations within 1 standard deviation of the mean, approximately 95% of the observations within 2 standard deviations of the mean, and approximately 99% within 3 standard deviations of the mean.

Consequently, studentized residuals with a magnitude beyond 2 are unusual, and beyond 3 are very unusual.

So what's the suggested cutoff value for detecting outliers if you use STUDENT residuals? In general, studentized residuals that have an absolute value less than 2 could easily occur by chance. For a relatively small sample size, studentized residuals that are between an absolute value of 2 to 3 occur infrequently and could be outliers. Studentized residuals that are greater than an absolute value of 3 occur rarely by chance alone, and should be investigated.

Statistics 1: Introduction to ANOVA, Regression, and Logistic Regression

Close