

Comparing Outlier Tests

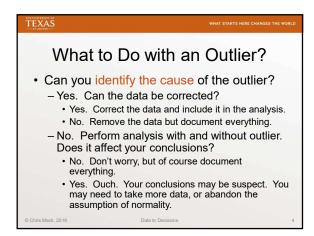
• Consider a data set where the Dixon Q-test fails to identify the extreme data point as an outlier, but the Grubbs' test does (for the same α)

- We don't think of one test as being "right" and the other "wrong"

- If we reject the null hypothesis (call the data point an outlier), then we know that our type I error rate is < α (which is set by us)

- If we fail to reject the null hypothesis, it could be because our test has insufficient power (we don't set the value of β directly)

• The Grubbs' test has more power (given its assumptions are true), which is why we prefer it



What to Do with an Outlier?

• When repeating the data analysis, there are many options of what to do with the outlier

• Delete the outlier

• Truncate (delete both the min and max data points)

• Winsorize the outlier (set its value equal to its closest neighbor)

• Replace the outlier with its expected value (from the Q-Q plot)

• Whether we delete, truncate, Winsorize, or replace the data depends on whether we identify the cause

• We always delete spurious data

• In any case, document exactly what you did

Three Types of Outlier Causation

Case 1: You notice the problem before you detect the outlier

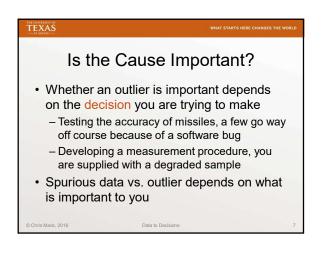
E.g., a measurement tool breaks and must be repaired, you suspect calibration will be off

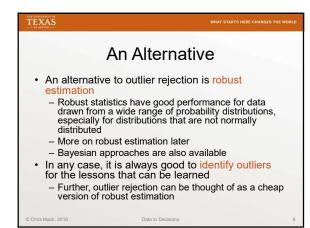
Case 2: You investigate after the outlier is observed and identify a cause

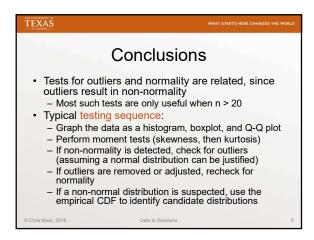
Beware of just-so stories

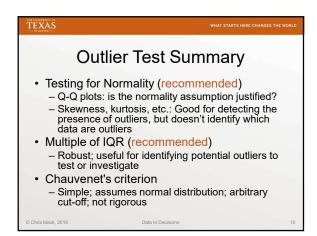
Case 3: You never find a cause

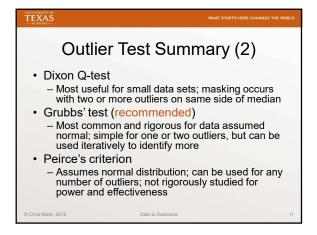
Question: when do you report the existence of outliers in your data?

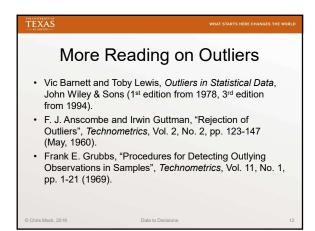












Lecture 19: What have we learned? • Why should one focus on identifying the cause of an outlier? • Name the four options for what to do with an outlier that can't be ignored • What is an important alternative to outlier testing and rejection? • Describe the recommended testing sequence for outliers