



by Dave Bock  
Cornell University  
Ithaca, New York

## Inference for Means

Whenever samples are involved, we check the Random Sample Condition and the 10 Percent Condition. Beyond that, inference for means is based on  $t$ -models because we never can know the standard deviation of the population. The theorems proving that the sampling model for sample means follows a  $t$ -distribution are based on the...

**Normal Population Assumption:** The data were drawn from a population that's Normal.

We can never know if this is true, but we can look for any warning signals. We've done that earlier in the course, so students should know how to check the...

**Nearly Normal Condition:** A histogram of the data appears to be roughly unimodal, symmetric, and without outliers.

If so, it's okay to proceed with inference based on a  $t$ -model. But what does "nearly" Normal mean? If the sample is small, we must worry about outliers and skewness, but as the sample size increases, the  $t$ -procedures become more robust. By the time the sample gets to be 30-40 or more, we really need not be too concerned.

/\* Provides a bit more detail than pg. 537 of our book \*/