

Robustness of t -statistics

- ▶ t -distribution of $\frac{\bar{x} - \mu}{SE_{\bar{x}}}$ assumes variable has normal distribution for the population
- ▶ in practice, must judge normality of population distribution based on
 - ▶ evidence from sample distribution (from histogram and/or normal quantile plot)
 - ▶ experience with similar data
- ▶ inference with t -distribution model can work well even if normality assumption is not true (*robustness*)
- ▶ guidelines (based on simulation and theory):
 - ▶ small samples ($n < 15$): use t -distribution methods if there is strong evidence of normality in the sample distribution
 - ▶ medium samples ($15 < n < 40$): use t -distribution methods unless there are outliers or evidence of strong skew in the sample distribution
 - ▶ large samples ($40 < n$): use t -distribution methods unless there are outliers