Lecture 16: Sample Size and Power

Chapter 4.6

Two-Sided Alternative Hypothesis

Pre-specifying α

Say Dr. Q is conducting a hypothesis tests. They start with $\alpha = 0.05$.

Pre-specifying α

Say Dr. Q is conducting a hypothesis tests. They start with $\alpha = 0.05$.

They conduct the test and get p-value $= 0.09$. They then declare "having used an $\alpha = 0.10$, we reject the null hypothesis and declare our results to be significant."

Pre-specifying α

Say Dr. Q is conducting a hypothesis tests. They start with $\alpha = 0.05$.

They conduct the test and get p-value $= 0.09$. They then declare "having used an $\alpha = 0.10$, we reject the null hypothesis and declare our results to be significant."

What's not honest about this approach?

Pre-specifying α

Say Dr. Q is conducting a hypothesis tests. They start with $\alpha = 0.05$.

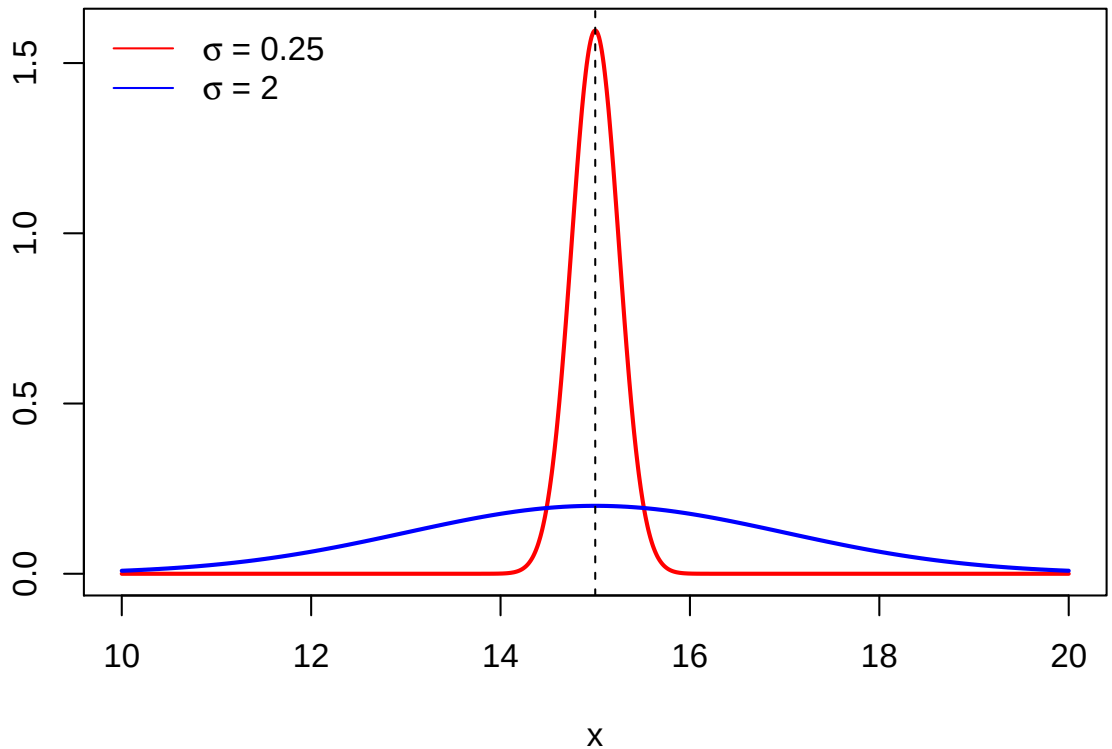
They conduct the test and get p-value $= 0.09$. They then declare "having used an $\alpha = 0.10$, we reject the null hypothesis and declare our results to be significant."

What's not honest about this approach?

Ronald Fisher, the creator of p-values, never intended for them to be used this way: http://en.wikipedia.org/wiki/P-value#Criticisms

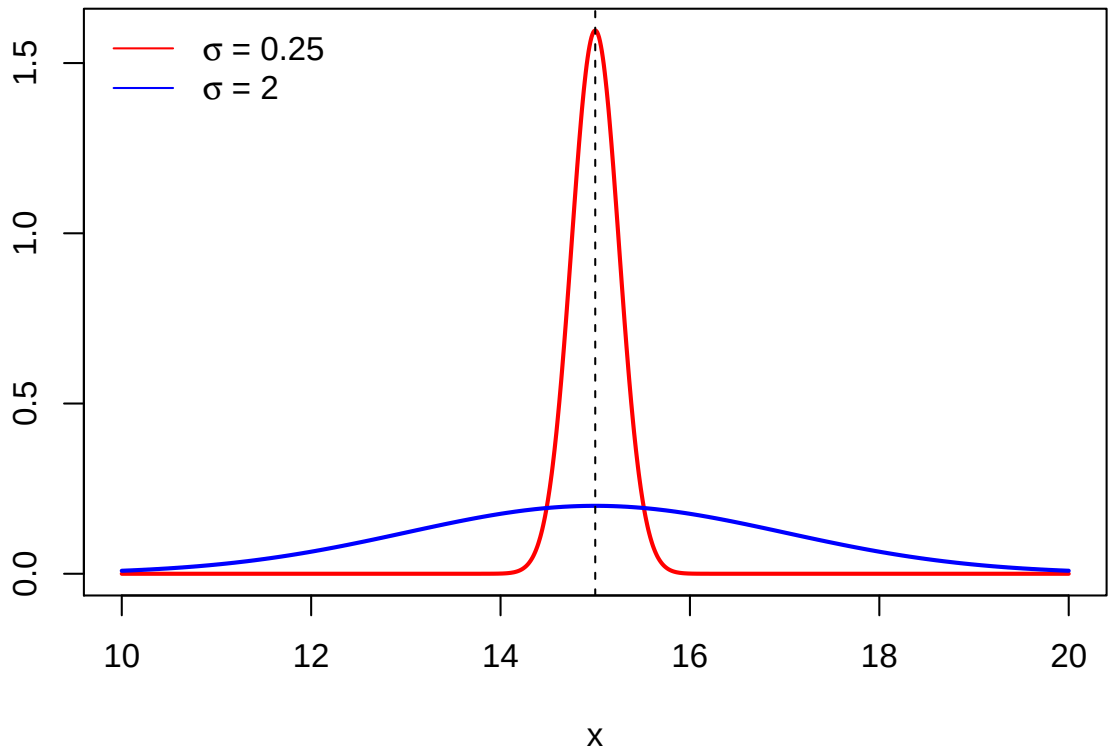
Sample Size: Thought Experiment

Say we have 2 population distributions with $\mu = 15$ but different σ:



Sample Size: Thought Experiment

Say we have 2 population distributions with $\mu = 15$ but different σ:

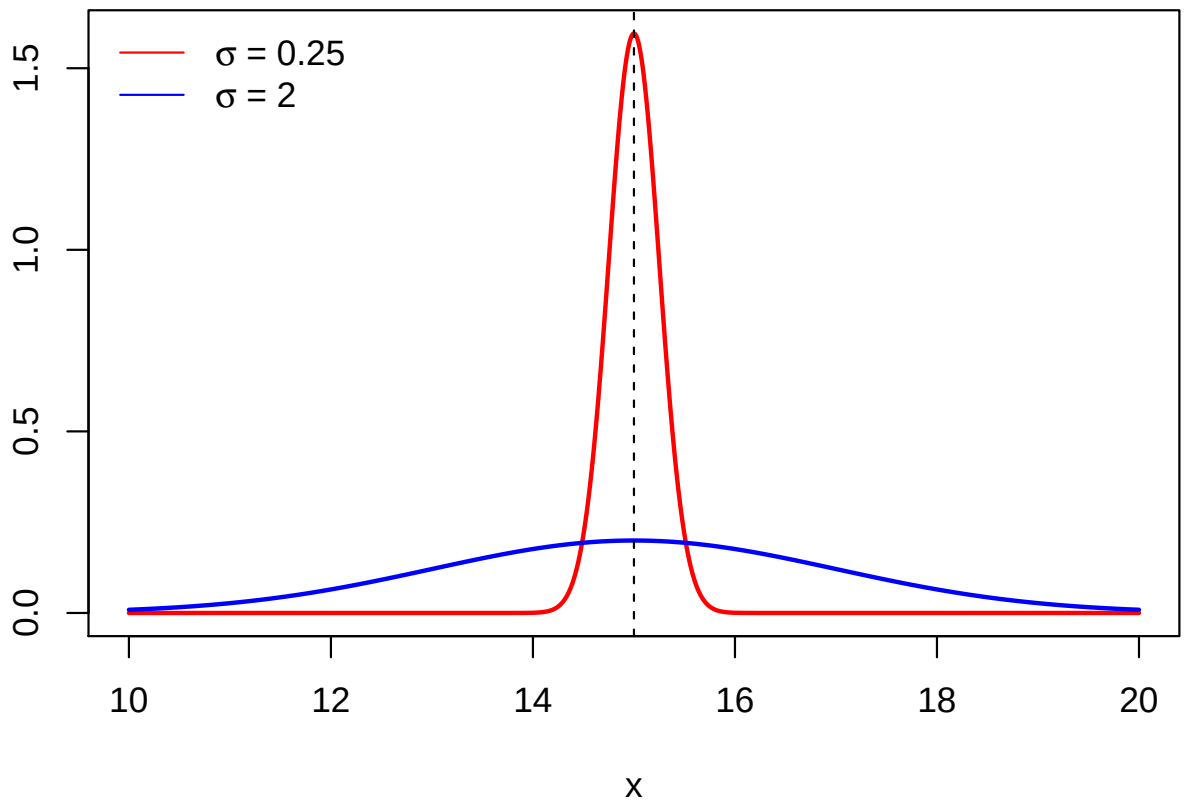


Which of the two distributions do you think will require a bigger n to estimate μ "well"?

Margin of Error

Back to Thought Experiment

For the same desired maximal margin of error m and same confidence level, we need a larger n to estimate the mean of the blue curve:



Type II Error Rate and Power