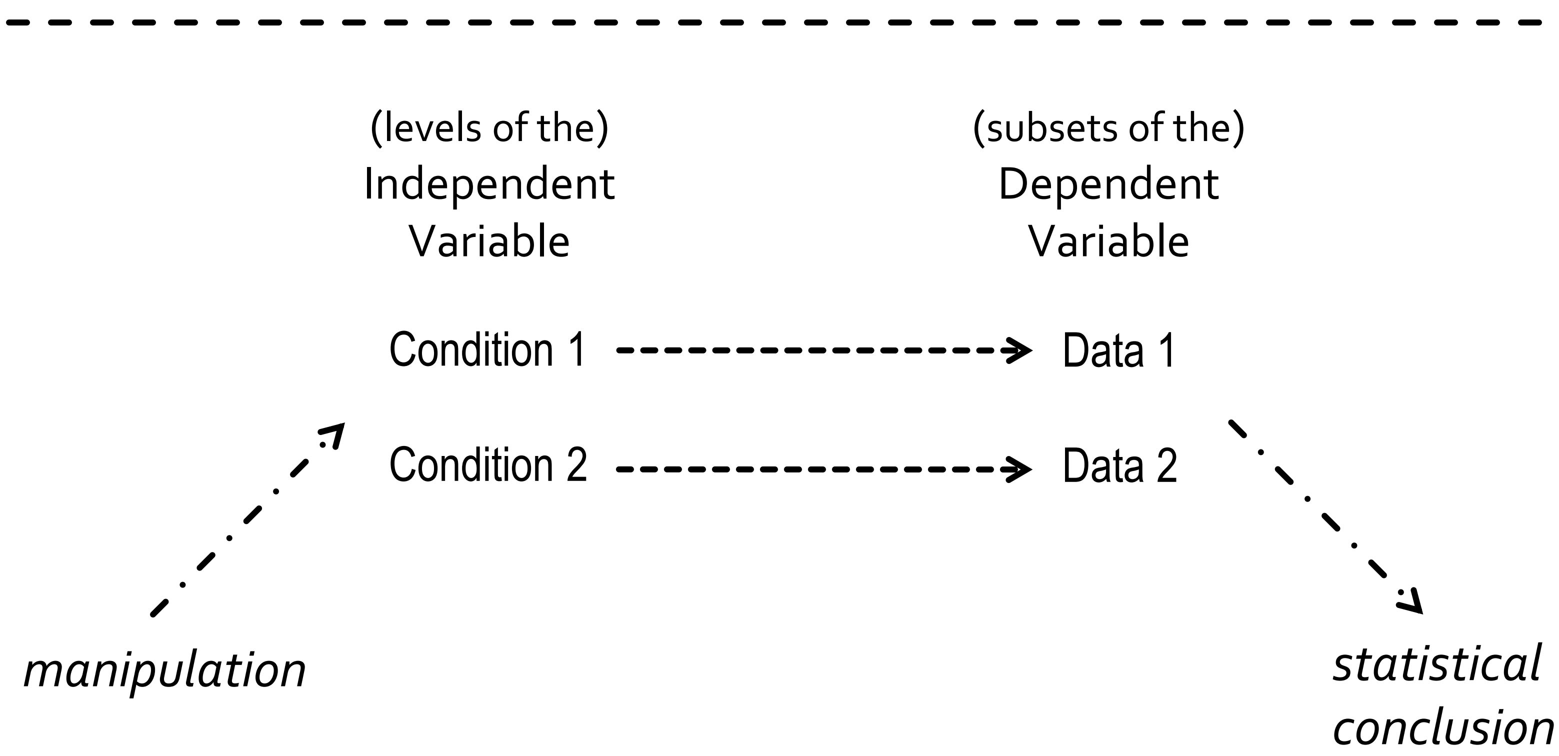
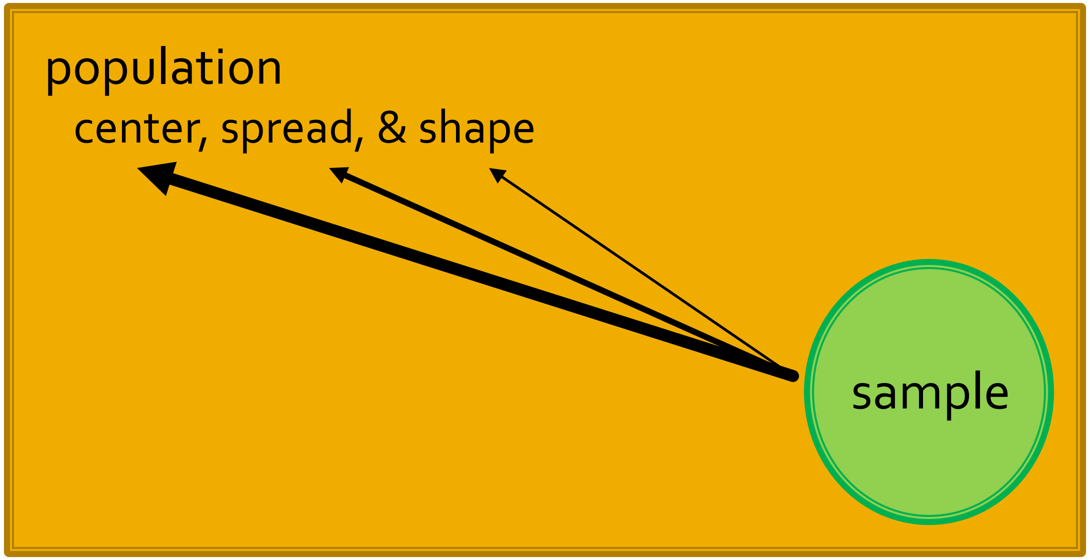
Summary Slide for Experiments



Inferential Statistics

- go beyond the actual data in hand…
  …and make a “best guess” about the population from which the sample was taken



- can be wrong … even when you do everything right

Inferential Statistics

- Step Zero: get the descriptive stats (mean & sd)
- Step One: get a best guess for whatever you're interested in

  e.g., the difference between two means

  get an estimate of how wrong you might be

- Step Two (optional; "hypothesis testing"): convert the values from Step One into a yes-or-no answer

  e.g., are the means for these conditions different in the entire sampling population ... yes or no?

Point Estimation (for the mean)

- 1a) best guess for population mean
  = sample mean

- 1b) estimate for how wrong you might be
  = sample standard deviation divided by $\sqrt{N}$
  this is called the "standard error (of the mean)"

Note: standard error has an associated value to indicate how wrong it might be (etc, etc)

this is the "degrees of freedom" = $N-1$

Point Estimation **(for the mean)**

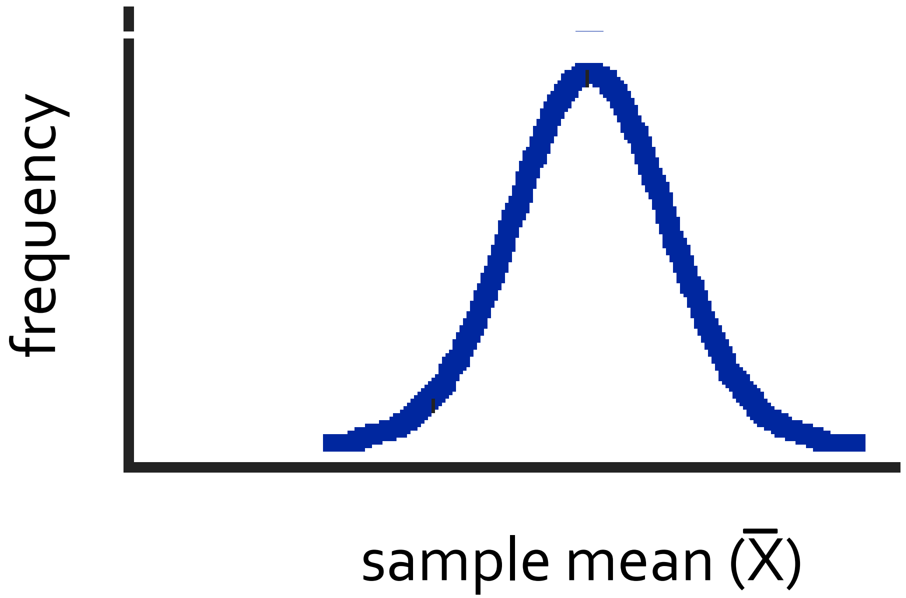
- note: there is no "validity" associated with Step One

  because we haven't made an inference yet

- to make an inferential statement about the mean, you must make some assumptions

  the most important assumption concerns the shape of the distribution of means

Point Estimation (for the mean)



Point Estimation (for the mean)

- how do you express a point estimation?

  old-fashioned: you calculate the “window” or interval that has a 95% chance of including the population value

  new standard: just provide the best guess and the estimate of how wrong you could be; i.e., $\overline{X}$ & se

Design Types

- assume an experiment with two conditions
  - (i.e., the smallest possible experiment)
- there are two ways that you can run this
  - all subjects participate in both conditions
    - within-subjects design
  - different groups of subjects in each condition
    - between-subjects design
- this is a huge decision; today you'll start to see why

Paired-samples *t*-test

- Step Two (yes-or-no answer)
- when you use a within-subjects design, you get pairs of data … Condition 1 and Condition 2

  when what you're interested in is whether there is a difference between conditions – yes or no –

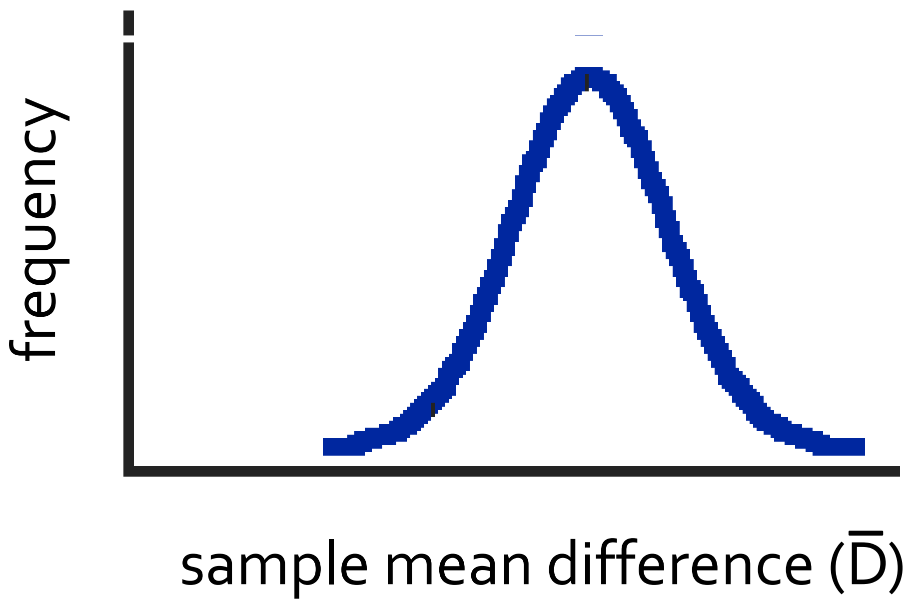
  you convert each pair to a difference score,

  and ask the yes-or-no question:

  is the mean <u>difference</u> something other than zero?

  new question: how different from zero must the sample mean be in order for you to say *yes*?

Paired-samples *t*-test



where is zero in this plot?

what is the probability of getting the observed mean difference, assuming this distribution?

rule in psychology: if the probability of getting what we found is less than 5%, then we say that the pop mean is not zero

Some More Standard Symbols

- μ (mu): population mean [unknown … usual target]
- $\bar{X}$ (X-bar): sample mean
- $\hat{\mu}$ (mu-hat): best guess for population mean = $\bar{X}$
- sd (*use full name*): standard deviation in the sample
- se (*use full name*): standard error of the mean = sd/$\sqrt{N}$
- df (*use full name*): degrees of freedom
- α (alpha): "significance" cut-off = .05 in psychology

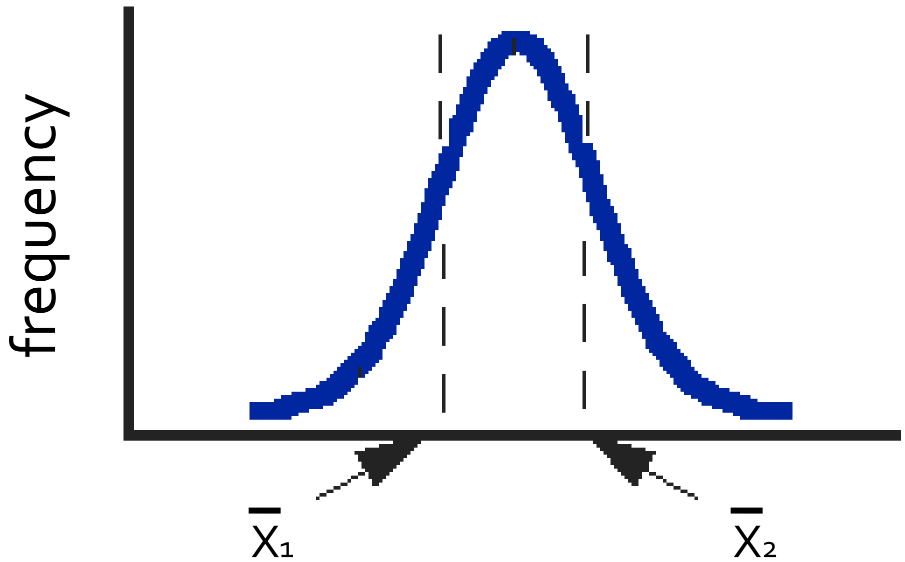
Independent-samples *t*-test

- Step Two (yes-or-no answer)
- when you use a between-subjects design, you have separate samples for Condition 1 and Condition 2

  so, the probability question (that leads to a "Yes" conclusion when the p-value is less than 5%) becomes:

  what is the probability of getting two sample means that are this far apart if we assume both samples came from one distribution?

Independent-samples *t*-test



what could go wrong?

- the answer depends on what you concluded
- if you concluded that a difference exists (in the pop)

  then you might be making a "false-alarm" error

  which is a Type-I error

- if you concluded that there isn't a difference

  then you might be making a "miss" error ... Type-II

All of the Standard Symbols

- μ (mu): population mean [unknown … usual target]
- $\bar{X}$ (X-bar): sample mean
- $\hat{\mu}$ (mu-hat): best guess for population mean = $\bar{X}$
- sd (*use full name*): standard deviation
- se (*use full name*): standard error of the mean = $sd/\sqrt{N}$
- df (*use full name*): degrees of freedom
- α (alpha): "significance" cut-off = .05 in psychology
- $1–\beta$ (one minus beta): "power"; aim for .80+ in psych

Inferential Errors (from *t*-tests)

- what causes Type-I errors?
- what causes Type-II errors?
- what are the (main) assumptions?

Power and Design-type

- what could go wrong and assumptions are the same as for both types of design
- the main difference between the two types of ***t***-test (from a statistical point of view) is the size of β

  independent-samples ***t***-tests have much less power (i.e., higher β) for a given number of subject