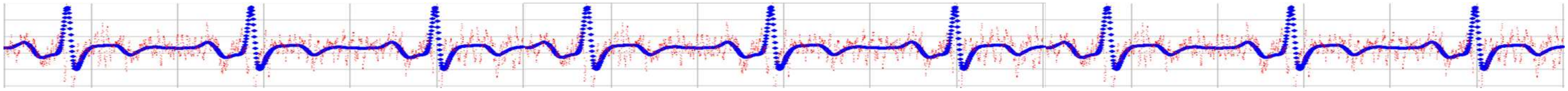
Empirical Research Methods in Information Science

IS 4800 / CS6350



Lecture 19

Outline

- Reading assessment
- T1 feedback Q&A and T2
- Study design practice
- *t* tests
- Power

Study designs practice

Example – best design?

You've developed a new web-based help system for your email client. You want to compare your system to the old printed manual.

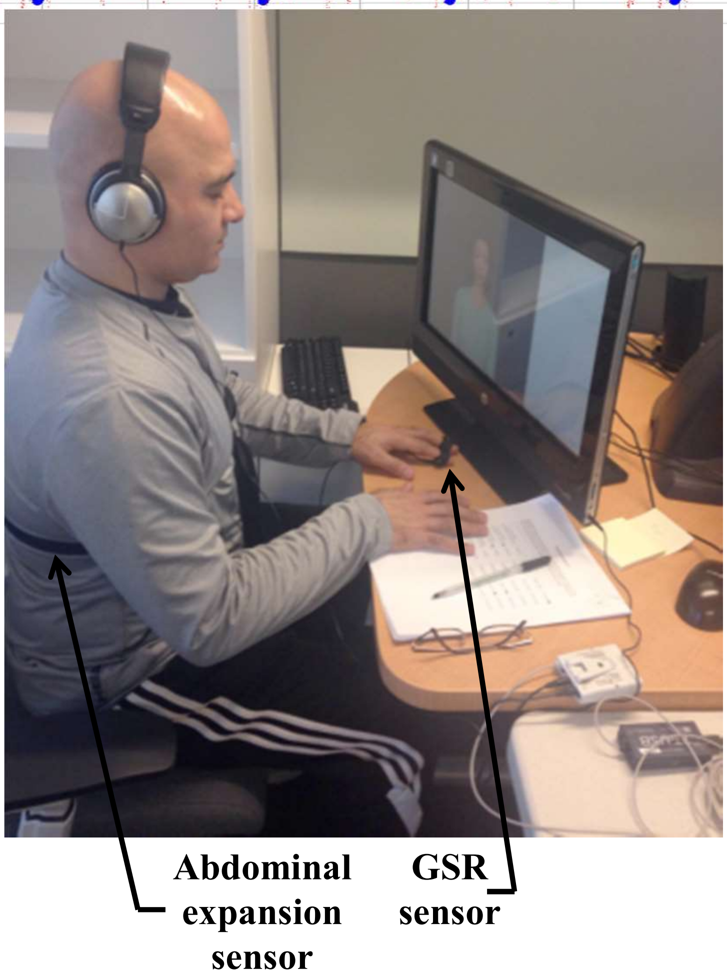
Example – best design?

You are evaluating a new customer support ticketing system and want to handle some customer calls with the new system to compare it to the old one.

Example – best design?

- Want to evaluate skype instead of face-to-face for sales calls among your international B2B salesforce
- 10x productivity difference among salespeople
- A salesperson makes 1-2 sales calls per month

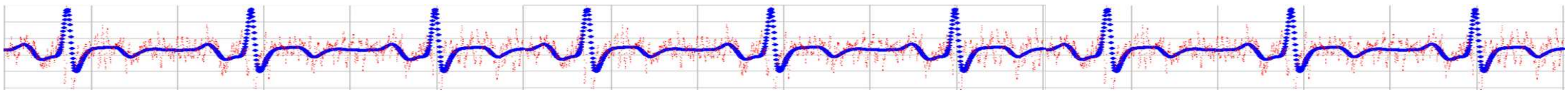
Example: Best design?



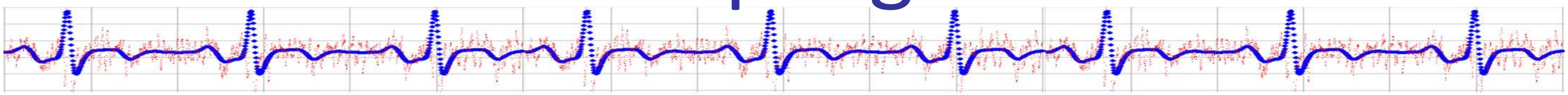
Study of Novice Programmers using Eclipse & Gild

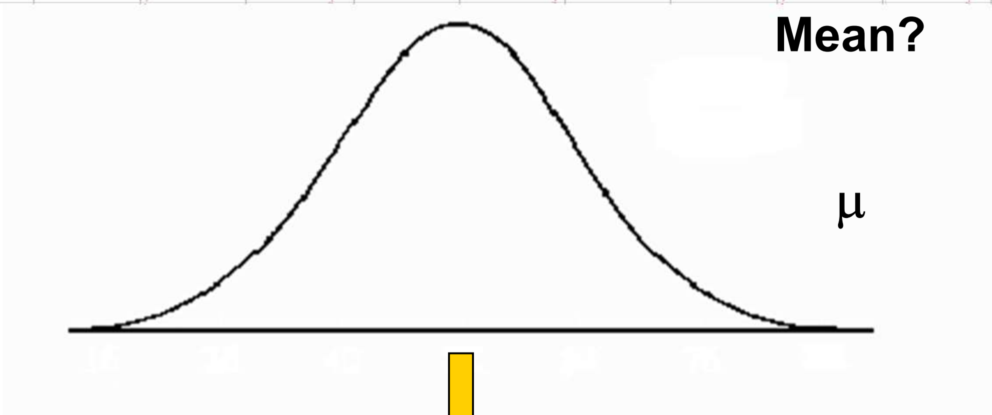
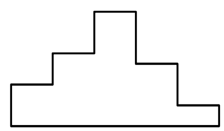
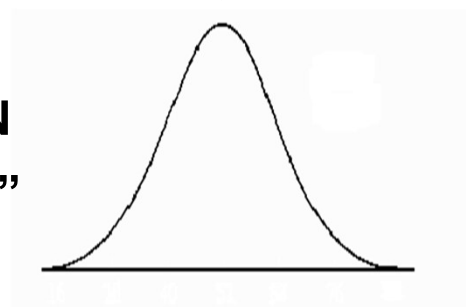
- Critique?

t-statistics,
t-distributions &
t-tests



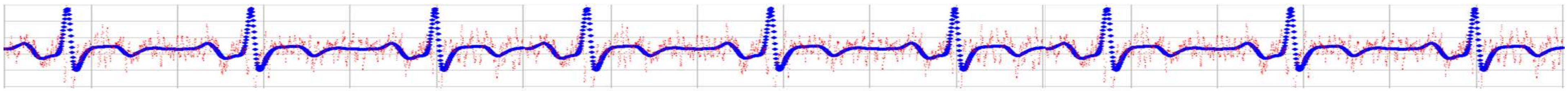
Review - sampling



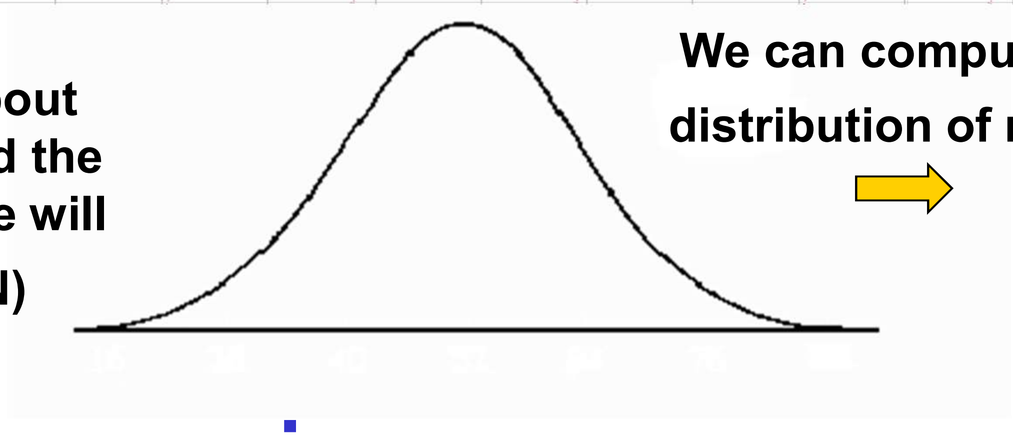
		Mean?	Variance?
Population		μ	σ^2
Sample of size N		$M = \frac{\sum X}{N}$	$SD^2 = \frac{\sum (X-M)^2}{N}$
Mean values from all possible samples of size N **aka "distribution of means"**		μ	$\sigma_M^2 = \frac{\sigma^2}{N}$

Standard error of the mean

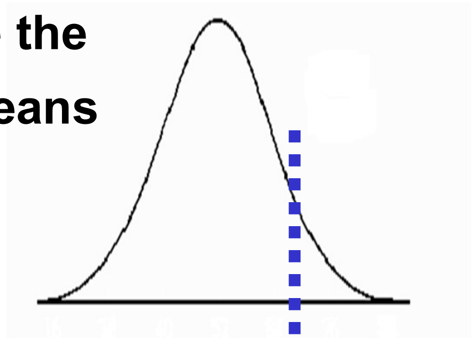
Hypothesis testing with a sample wrt distribution of means



Given info about population and the sample size we will be using (N)

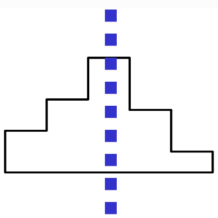


We can compute the distribution of means



and finally determine the probability that this mean occurred by chance

Now, given a particular sample of size N



We compute its mean

t-test for single sample

- College paper reports students study 17 hours per week at NU
- You think they study more at your dorm
- Take a random sample of 16 students from your dorm and find they study 21 hours per week
- What can you conclude?

t-test for single sample

- What can you conclude?
  - Compare sample to population
  - Know mean but not variance of population

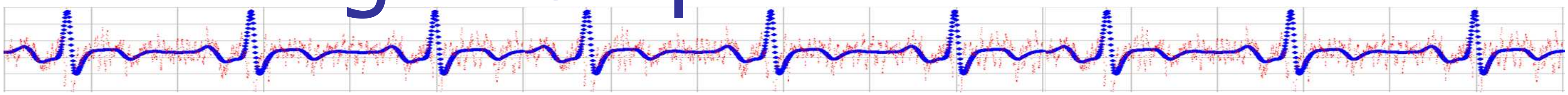
- (Note: If you can get t the mean *and* variance of population, use z-test)

- Research hypothesis: Dorm population studies more than college population

t-test for single sample

- Can’t get population variance?
- You can *estimate* the variance of the population using the sample
  - Should be similar variance
  - BUT, variance of sample usually smaller (biased estimate of population) so use variance equation with n-1

Single sample t-test



- Estimate variance from sample
  - $S^2 = \frac{\sum(X-M)^2}{N-1} = \frac{SS}{N-1} = \frac{SS}{df}$
    (Unbiased estimate of population variance)
  - $S = \sqrt{\frac{SS}{N-1}}$ (Standard deviation)
- N-1 is the degree of freedom (df)
- Variance of the distribution of means:

$$S^2{}_M = \frac{S^2}{N}$$

Symbols: S instead of σ for estimated population standard deviation

Symbols

Statistical Term	Symbol
Sample standard deviation	SD
Population standard deviation	σ
Estimated population standard deviation	S
Standard deviation of the distribution of means (based on an estimated population variance)	S_M
Sample variance	SD^2
Population variance	σ^2
Estimated population variance	S^2
Variance of the distribution of means (based on an estimated population variance)	S_M^2

Single sample t-test

- Standard deviation of the distribution of means:

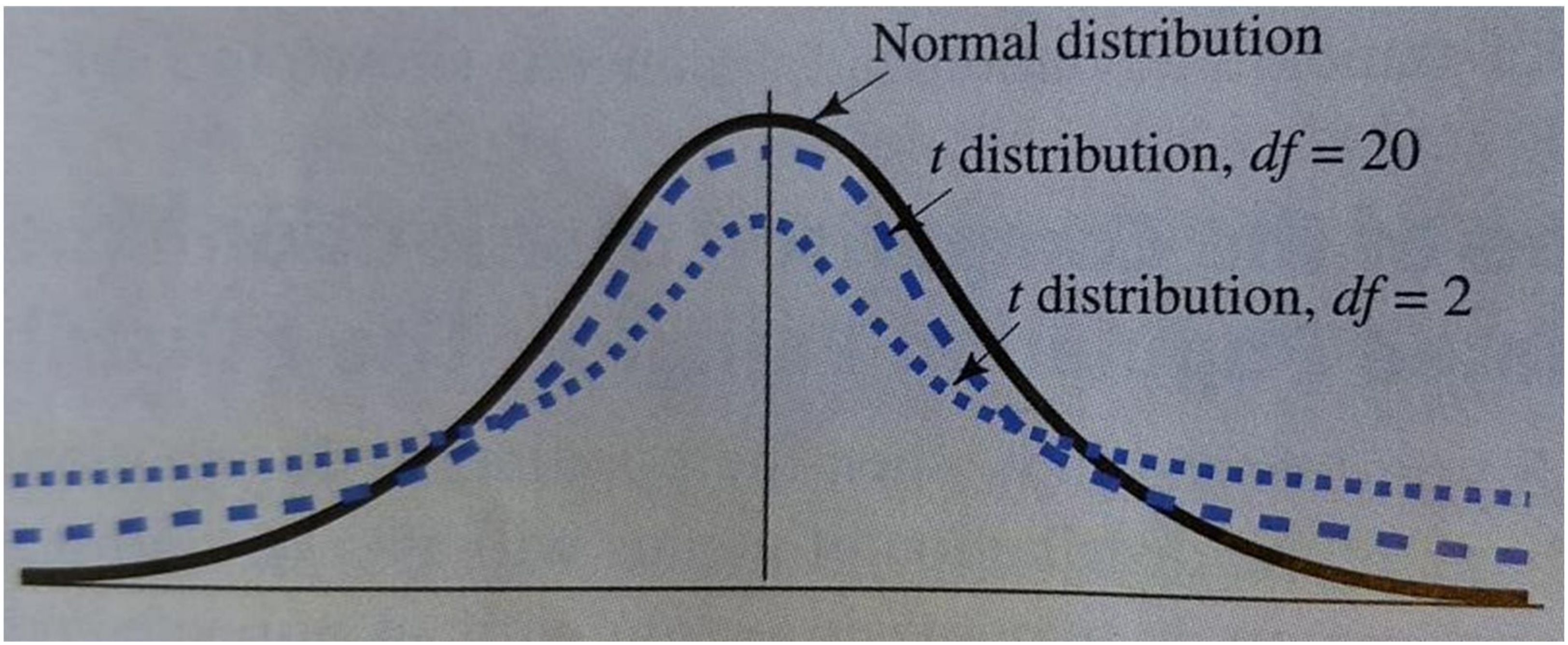
$$S_M = \frac{S}{\sqrt{N}}$$

- Shape of distribution of means not necessarily normal (because of estimated population variance)
- Comparison distribution therefore not normal: **t distribution family** (family members differ by dfs)

t distribution family

- Close to normal, but not
- Takes a slightly more extreme sample mean to get a significant result with t than nomal (e.g., z)
- Infinite sample size, t becomes normal
- By N = 30, close

t distribution

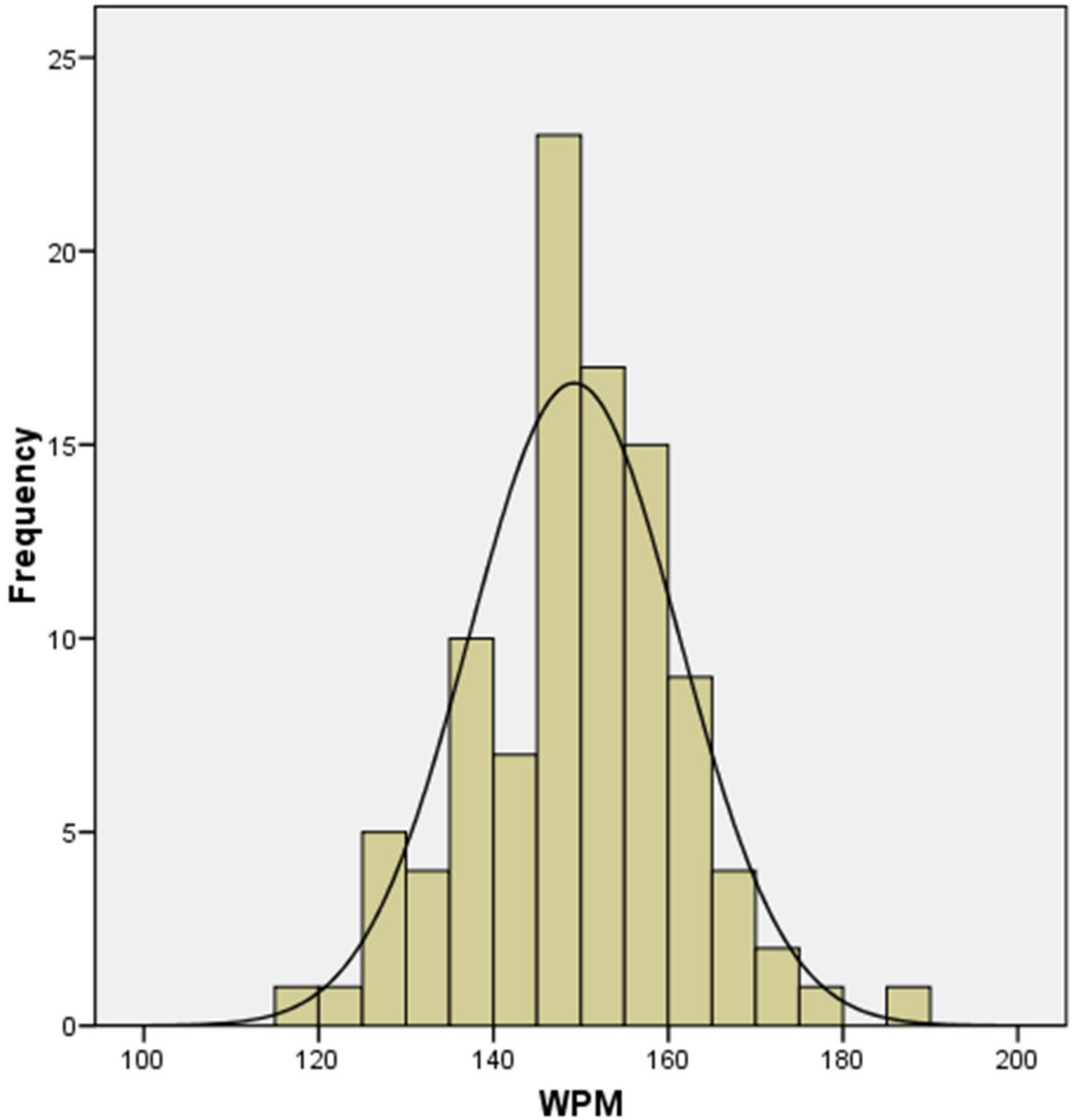


Single sample t-test

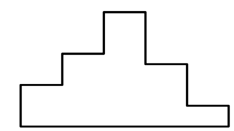
- Comparison: t-test

$$t = \frac{(M - \mu)}{S_M}$$

(Sample mean – pop mean, divided by the stddev of the distribution of means)



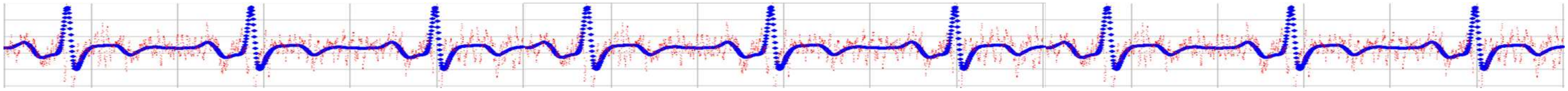
Example:
Take sample of 100 admins, give WW, assess their typing speeds



AND we know the comp population Mean but not stddev

Mean =149.23
Std. Dev. =12.02
N =100

t-test for dependent means
aka “paired sample t-test”



t-test for dependent means

- Extremely common:
  - Don't even know the population's *mean*
  - Have *two* sets of scores from each person in sample
    - E.g., measure before and after intervention

- Dependent = mean for each group of scores are dependent on each other because from the same people

t-test for dependent means: When to use

- One factor, two-level, within-subjects/repeated measures design

  -or-

- One factor, two-level, between-subjects, matched pair design

- *In general, a bivariate categorical IV and numeric DV when the DV scores are highly correlated*

- Assumes population distribution of individual scores is normal

t-test for dependent means

- Same as t-test for single sample, expect (1) use difference scores (2) assume population mean (of the difference scores) is 0

t-test for dependent means

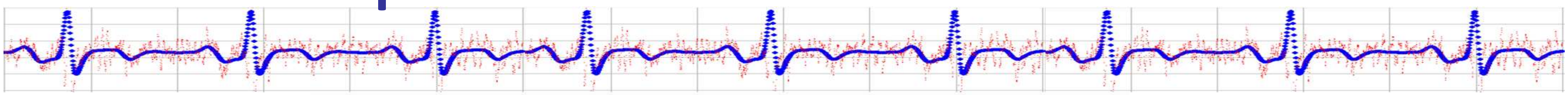
- Difference scores:
  - Make two scores into 1 by creating a change score
- Mean of zero:
  - Usually, null hypothesis in repeated measures assumes no difference between groups
  - I.e., comparison population has mean of 0

Wanted: A statistic for differences between paired measures

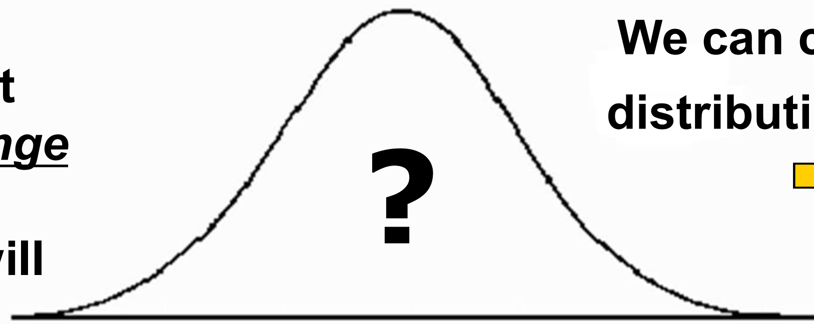
- In a repeated-measures or matched-pair design, you directly compare one subject with him/herself or another specific subject (not groups to groups).
- So, start with a sample of change (difference) scores:

Sample 1 = Mary's wpm using Wizziword –
Mary's wpm using Word

Hypothesis testing with paired samples



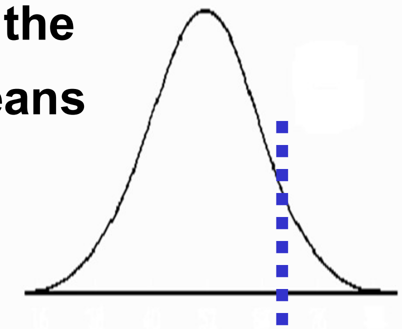
Given info about population *of change scores* and the sample size we will be using (N)



$\mu = 0$

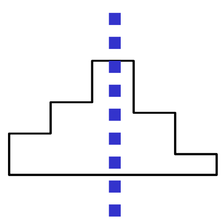
est σ^2 from sample

We can compute the distribution of means



and finally determine the probability that this mean occurred by chance

Now, given a particular sample *of change scores* of size N

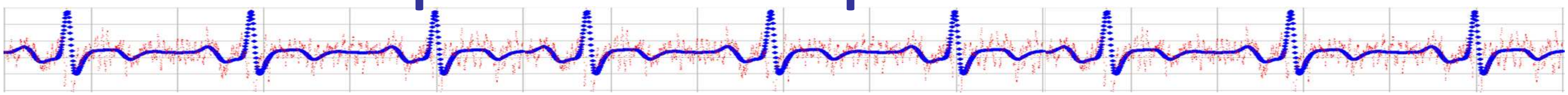


We compute its mean

$$t = \frac{(M - \mu)}{S_M} = \frac{M}{S_M}$$

df = N-1

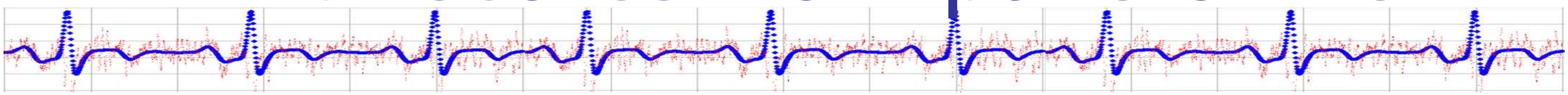
“t-test for dependent means” aka “paired sample t-test”



- Map two measures for each subject into one difference score for each
  - e.g. change due to intervention =

    after measure – before measure
- Null hypothesis (usually) no change
  - Thus mean of comparison dist is zero

t test for dependent means with scores from pairs of P's



- Scores from pairs of participants
- Consider each pair as if one person
- Figure difference between each pair

- (Similar to as if had two measurements from same person)

t test for dependent means with scores from pairs of P's

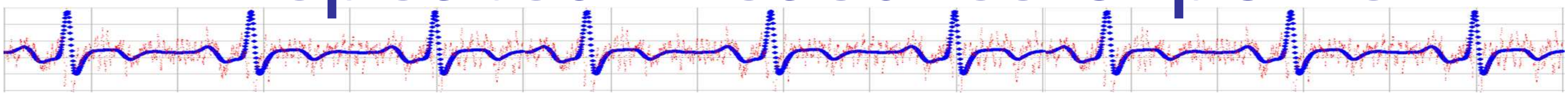
- Example:
  - Sample of 30 married couples
  - Wives do more housework than husbands?

- Example:
  - Pair task enjoyment level
  - Two people (leader/follower) complete puzzle task

Assumptions

- Normal population distribution
  - But you don't know this! Just have scores in sample!
  - In practice, (1) most distributions in psych normal, (2) t-test has been found to be robust over moderate violations of normality
    - Exception: Using a one-tailed test and population highly skewed (thus population differences highly skewed)

Repeated measures & power



- Within subjects generally has more power than between subjects tests
- Why?
  - StdDev of difference scores usually low
  - Divide by this to get effect size -> larger! (thus, increasing power)

- But … testing before/after without a control … concerns?

Reporting results

- Significant results

  t(df)=*tscore*, p<*sig*

  e.g., t(38)=4.72, p<.05 (two-tailed test)

  (If type of tail not noted, assume two-tailed; if one-sample t-test, note it (rare))

- Non-significant results

  e.g., t(38)=4.72, n.s.

- Note: usually report absolute value of t score and mean and SD of sample