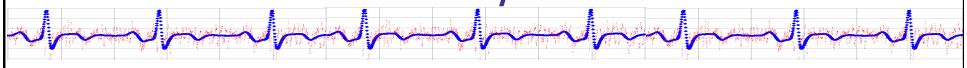
Empirical Research Methods in Information Science

IS 4800 / CS 6350

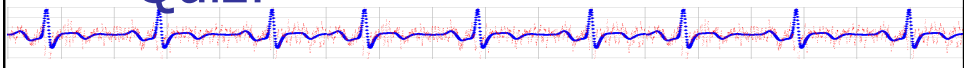


Lecture 13

The t Test for Independent Means

1

Quiz!



https://tinyurl.com/IS4800N6

Don't forget your name!

2

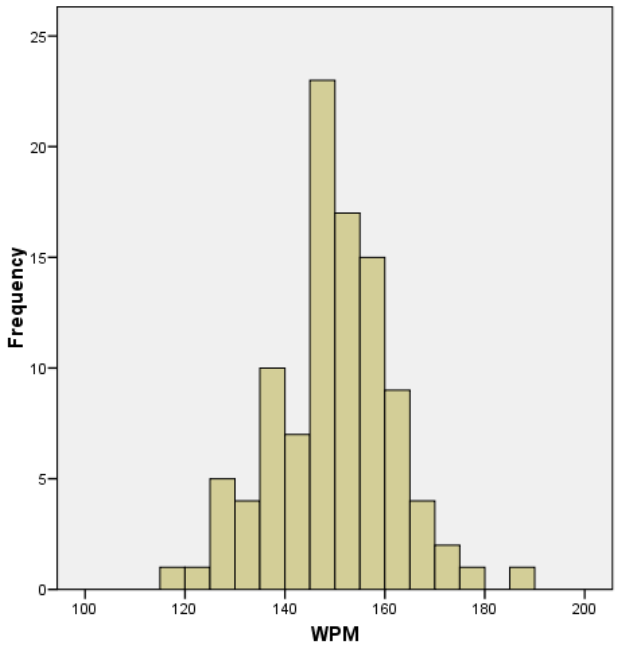
Example

- You have 100 admins in your company.
- They all use Word.
- You want to consider changing to WizziWord.
- Hypothesize it will increase their net productivity, measured as word per minute typed, averaged over an entire day.

3

Example

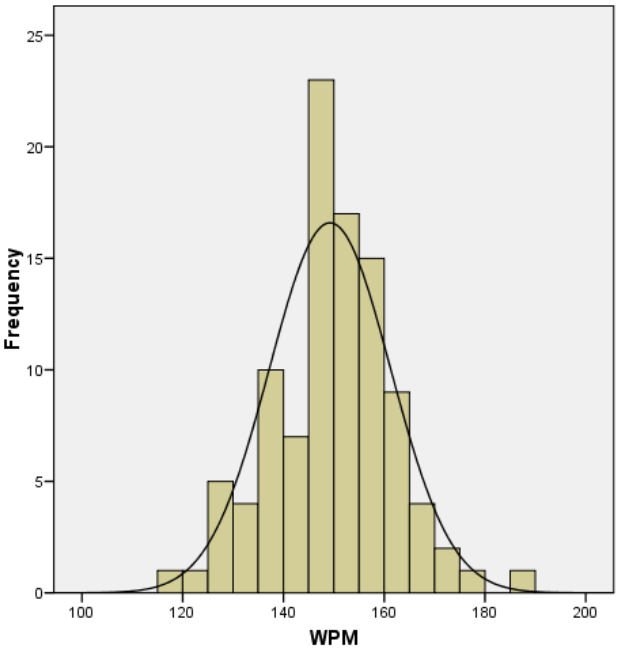
4



Productivity of
ALL admins
BEFORE WizziWord

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

5



Experiment:
Give 1 admin WW
for a month

types at 160 wpm

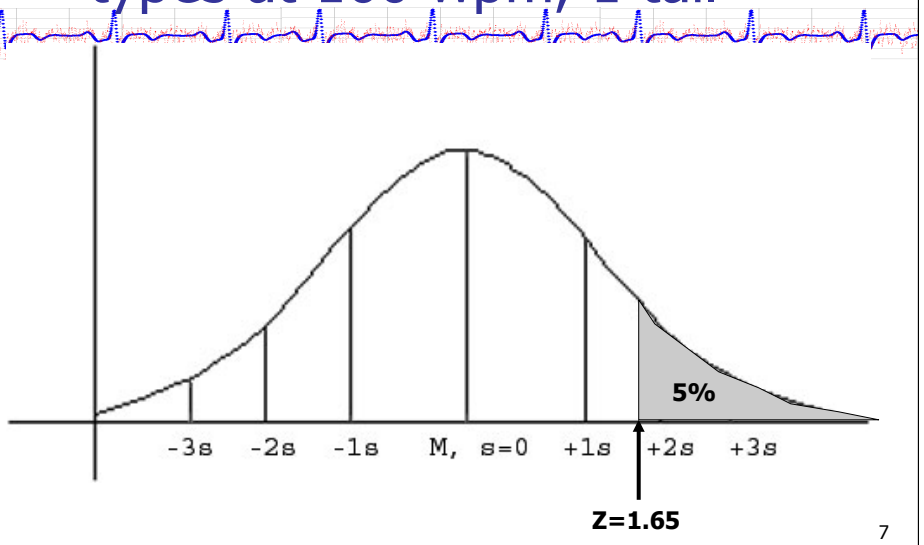
H1: WW > Word

Z=0.9

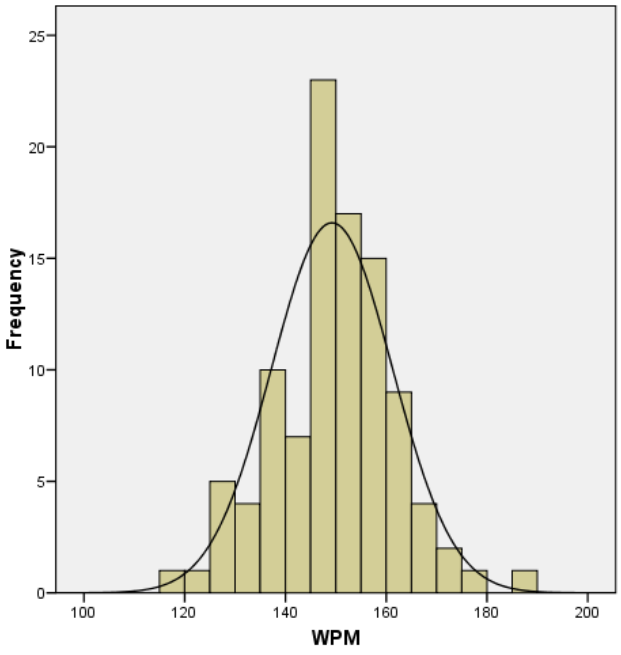
Mean =149.23
Std. Dev. =12.02
N =100

6

Example: 1 admin using WW types at 160 wpm, 1 tail



7



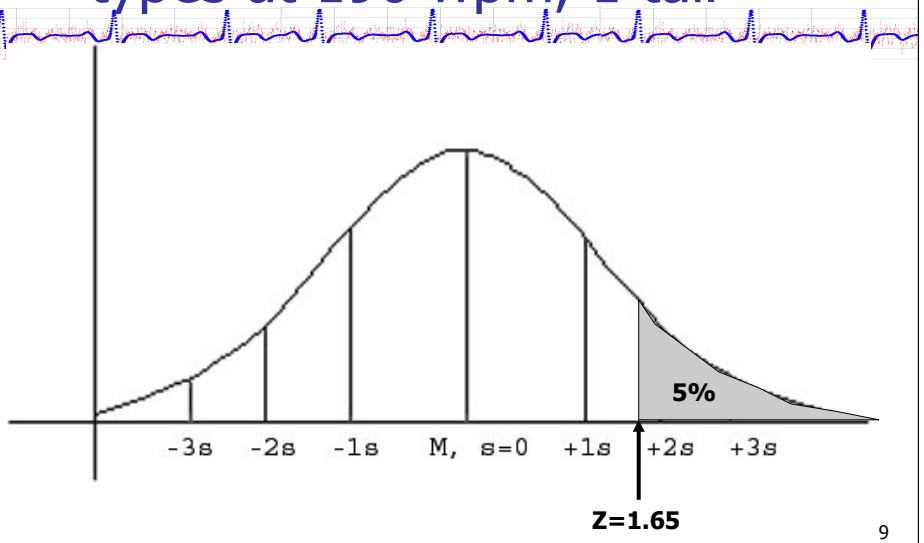
Example:
1 admin using WW types at 190 wpm
1 tail test

Z=3.4

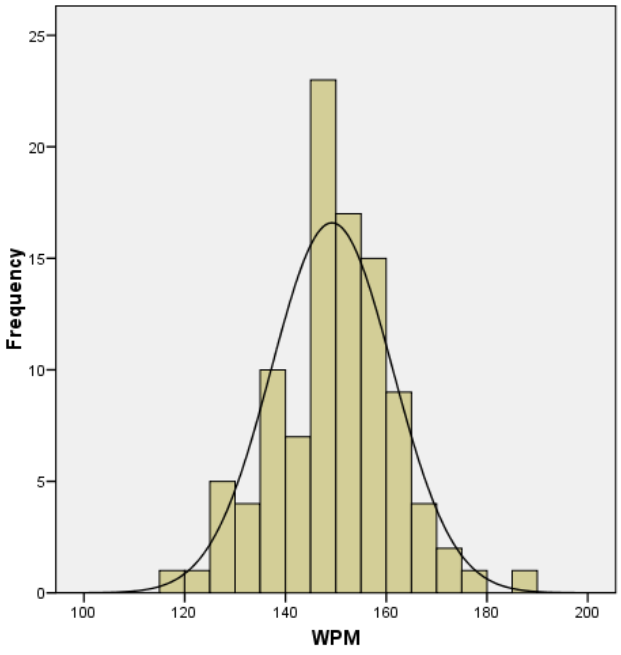
Mean =149.23
Std. Dev. =12.02
N =100

8

Example: 1 admin using WW types at 190 wpm, 1 tail



9



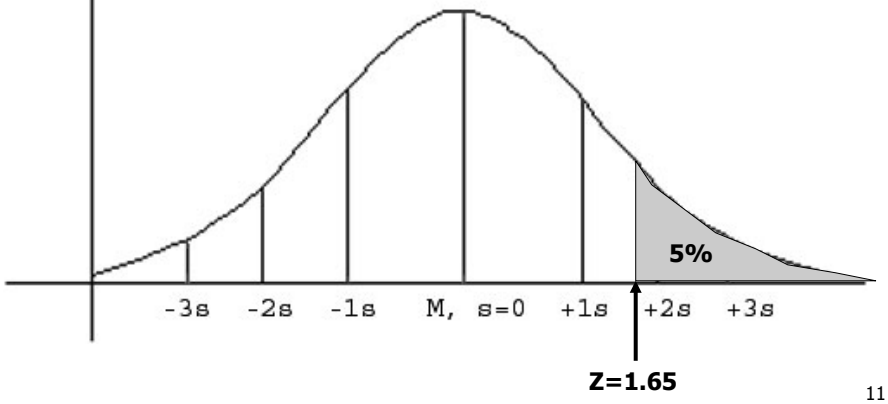
Example:
1 admin using WW types at 121 wpm
1 tail test

Z=-2.3

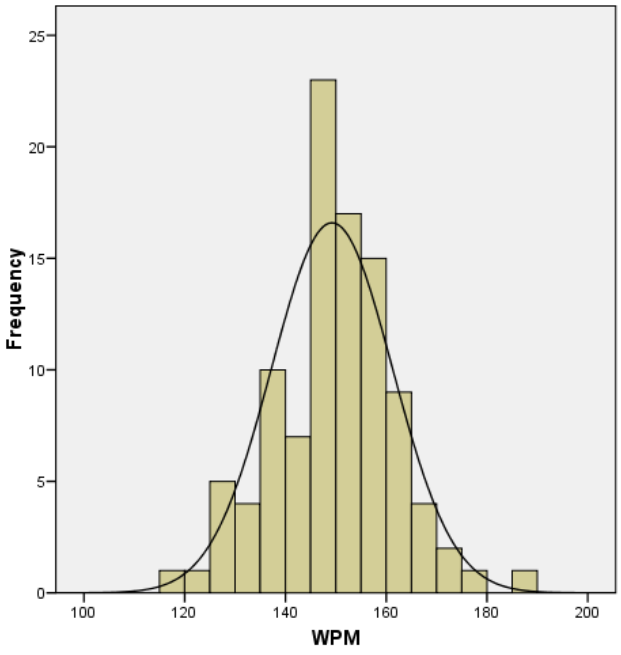
Mean =149.23
Std. Dev. =12.02
N =100

10

Example: 1 admin using WW types at 121 wpm, 1 tail



11



Example:
1 admin using WW types at 121 wpm

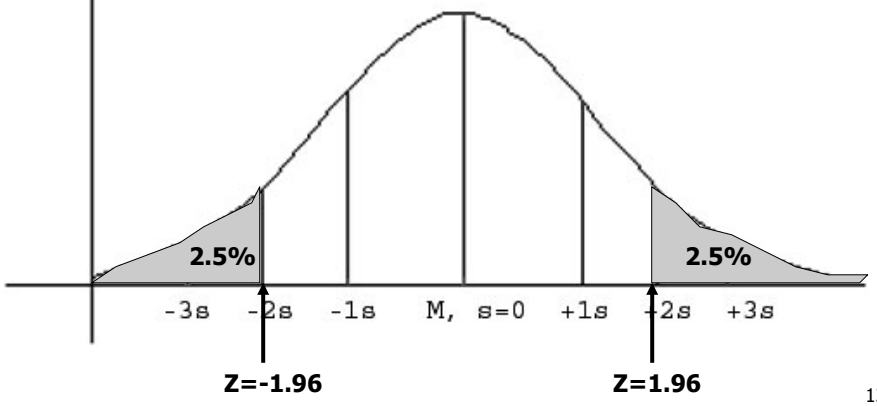
H1: WW ≠ W

Z=-2.3

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

12

Example: 1 admin using WW types at 121 wpm 2 tail test



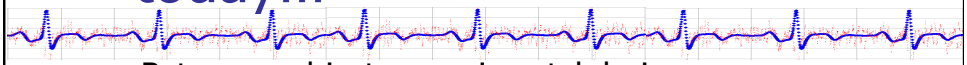
13

Don't try this at home

- You would never do a study this way.
- Why?
  - Can't control extraneous variables through randomization.
  - Usually don't know population statistics.
  - Can't generalize from an individual.

14

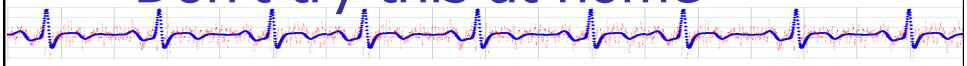
What you need to know from today…



1. Between-subjects experimental design.
   - Everything.
2. Distribution of means.
   - What it is. Formulas that relate to population distribution.
3. Distribution of the difference of means.
   - What it is. (not formulas)
4. t-distribution.
   - How different from normal. Parameters (formula for df). When/ why used.
5. t statistic/score for the difference between two means.
   - Formula.
6. Typical assumption(s) in t-test for independent means.
7. Randomization & Control conditions

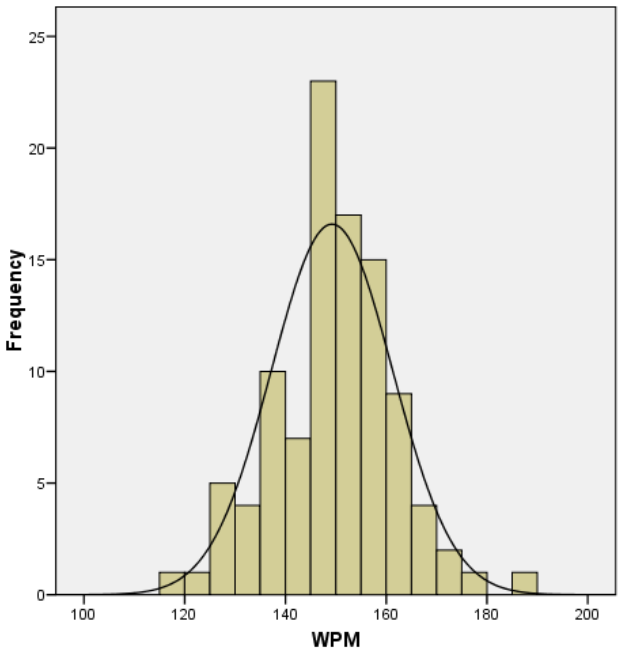
15

Don't try this at home

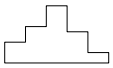


- You would never do a study this way.
- Why?
  - Can't control extraneous variables through randomization.
  - Usually don't know population statistics.
  - Can't generalize from an individual.

16



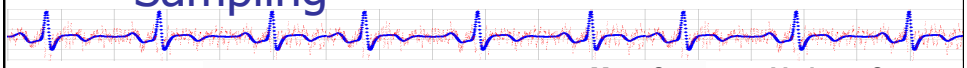
Example:
Take sample of 10 admins, give WW, assess their typing speeds



Now what?

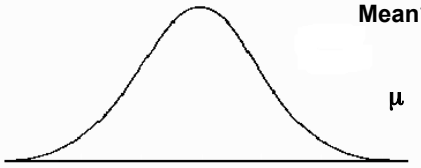
17

Sampling



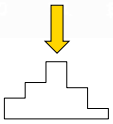
Population

Mean? **Variance?**



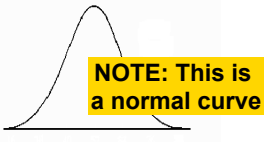
μ σ^2

Sample of size N



$M = \frac{\sum X}{N}$ $SD^2 = \frac{\sum (X - M)^2}{N - 1}$

Mean values from all possible samples of size N aka "distribution of means"

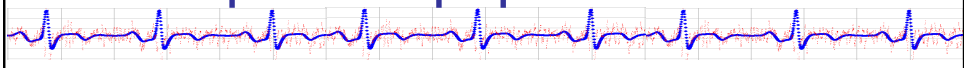


NOTE: This is a normal curve

μ $\sigma_M^2 = \frac{\sigma^2}{N}$

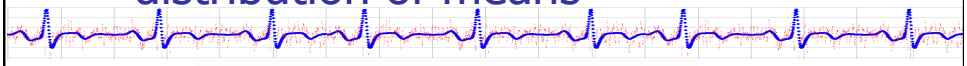
18

Single sample test, with sample size>1, and known comparison population…

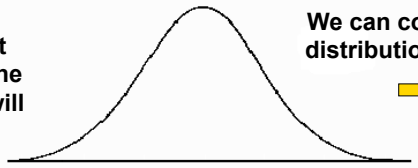


19

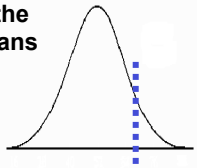
Hypothesis testing with a sample wrt distribution of means



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

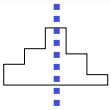


We can compute the distribution of means



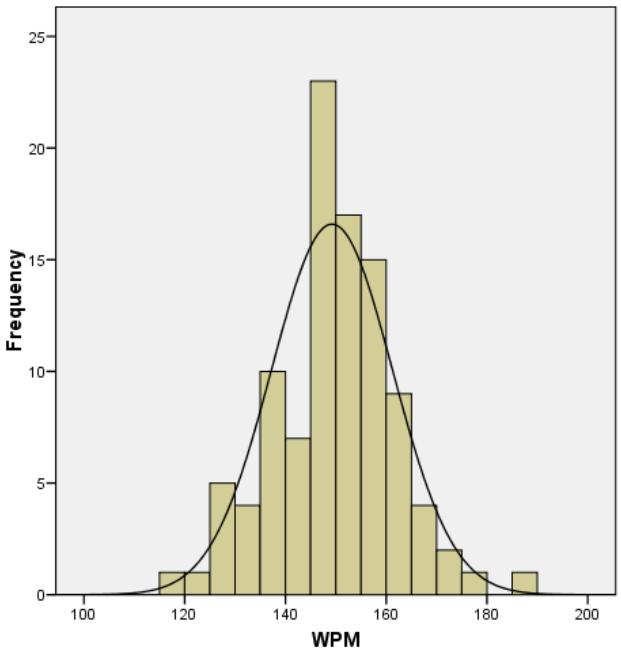
and finally determine the probability that this mean occurred by chance

Now, given a particular sample of size N

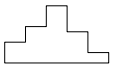


We compute its mean

20



Example:
Take sample of 10 admins, give WW, assess their typing speeds



Say M=170

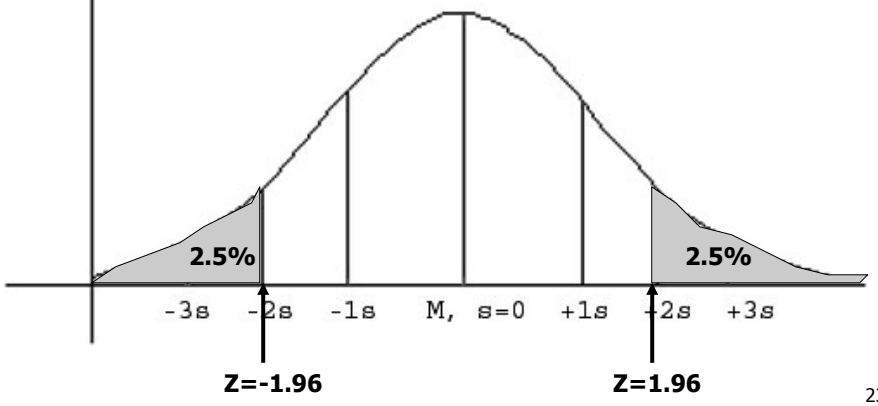
21

Comparison Distribution = Distribution of Means

- N = 10
- $\mu = 149$, $\sigma = 12$, $\sigma^2 = 144$
- $\mu_M = 149$, $\sigma_M^2 = 144/10 = 14.4$
- Z = (170 – 149) / $\sqrt{14.4} = 5.5$

22

Example: 10 admins using WW types at M=170 wpm 2 tail test



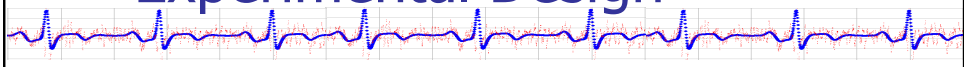
23

Don't try this at home

- You would never do a study this way.
- Why?
  - Can't control extraneous variables through randomization.
  - Usually don't know population statistics.
  - ~~Can't generalize from an individual.~~

24

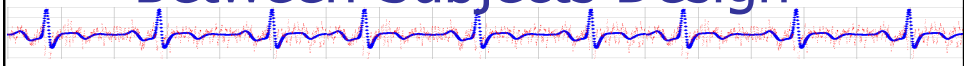
Two Group Between-Subjects Experimental Design



- B&A: "Randomized Two Group Design"
- Have two experimental conditions (treatments, levels, groups)
- Randomly assign subjects to conditions
  - Each subjects sees <u>one</u> condition
- Measure (numeric) outcome in each group

25

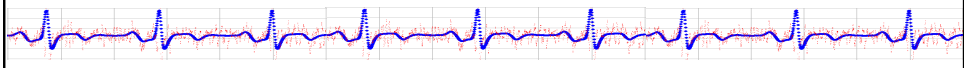
Between-Subjects Design



- Each group is a **sample** from a population
- Big question: are the populations the same (null hypothesis) or are they significantly different?

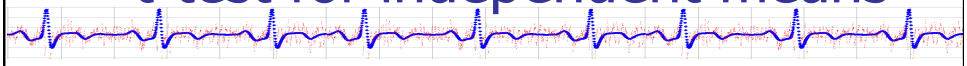
26

t-statistics, t-distributions & t-tests



28

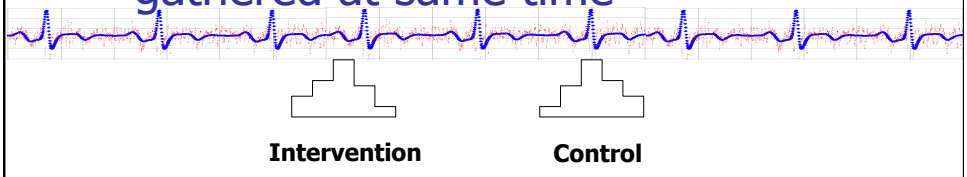
t-test for independent means



- Two unrelated samples
  - Eg as obtained in a between-subjects experiment
- No other information about comparison distribution

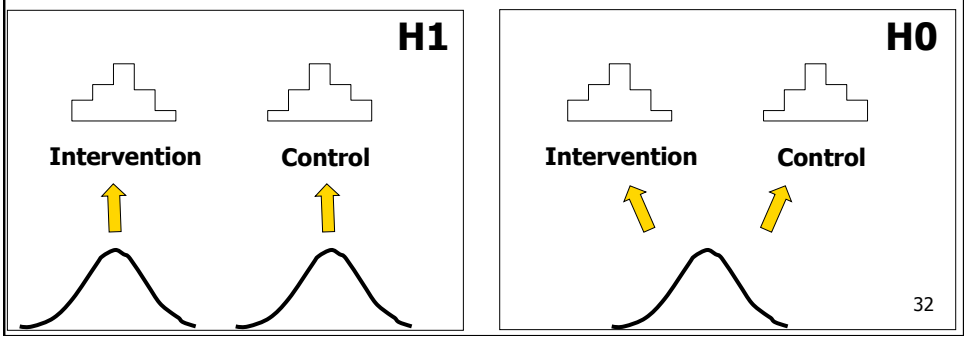
31

Solution – take two samples, gathered at same time



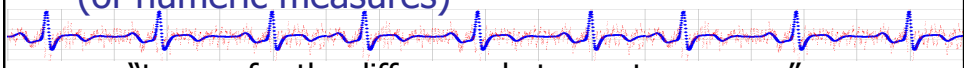
Intervention Control

The big question: which is correct?



32

Wanted: a statistic to measure how similar two samples are
(of numeric measures)



- "t score for the difference between two means"

$$t = \frac{M_1 - M_2}{S_?}$$

- If samples are identical, t=0
- As samples become more different, t increases.

- What is the comparison distribution?
  - Want to compute probability of getting a particular t score IF the samples actually came from the same distribution (what is the t score for this case?).

33

Why t?

- In this situation, you do not know the population parameters; they must be estimated from the samples.
- When you have to estimate a comparison population's variance, the resulting distribution is not normal – it is a "t distribution".
  - Looks normal, but has thicker tails (need more extreme Z score for significance)
  - As df increases, t becomes normal
- The particular kind of t distribution we are using in this case is called a "distribution of the difference of means".

34

All things t

- t distribution shape is parameterized by "degrees of freedom"
- For a distribution of the difference of means,

$$df = df_1 + df_2 = (N_1 - 1) + (N_2 - 1)$$

35

Assumptions for t

- Scores are sampled randomly from the population
- The sampling distribution of means is normal
- Variances of the two populations (whether they are the same or different) are the same.
  - *There are other forms of the statistic that do not make this assumption.*

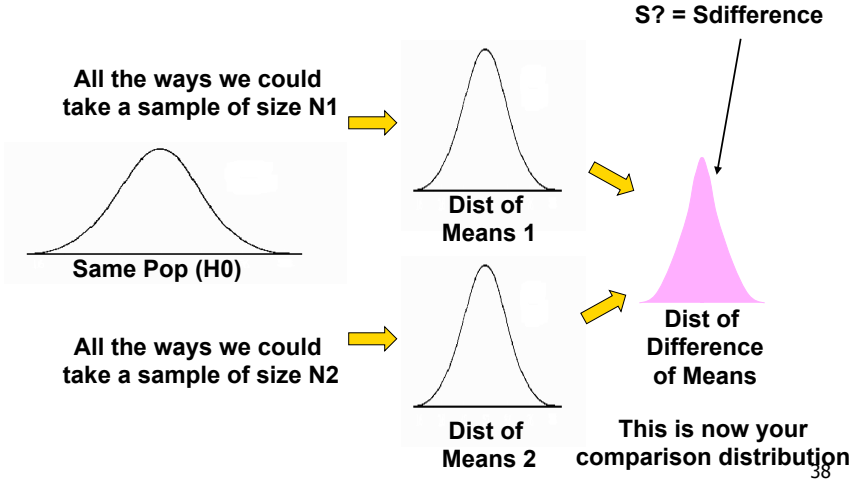
36

Only remaining loose end…

$$t = \frac{M_1 - M_2}{S_?}$$

37

Finally – the t test for independent samples



38

Reporting results

- Significant results
  t(df)=*tscore*, p<*sig*

  e.g., t(38)=4.72, p<.05

- Non-significant results
  e.g., t(38)=4.72, n.s.

- Note: usually report absolute value of t score

40

R

```
>#default 2-tail
>t.test(sample1,sample2,var.equal=TRUE)


Two Sample t-test
…
data:  sample1 and sample2
t = -2.6887, df = 12, p-value = 0.01972
```

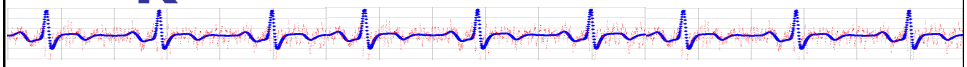
41

Group Exercise

- For each problem, write
  1. What kind of study design is it?
  2. Two populations being compared
  3. Research & Null hypotheses in English
  4. Research & Null hypotheses in terms of Pop means
  5. Test criteria
  6. Test results
     - Formal report format
     - English

42

R



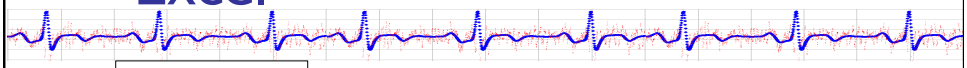
```
#1-tail
t.test(sample1,sample2,alternative="less")
t.test(sample1,sample2,alternative="greater")


#Correction for unequal variances of pops
t.test(sample1,sample2)

     Welch Two Sample t-test
…
```

43

Excel



IV	DV
I	2.3
C	1.2
I	4.7
C	0.5
…	…

44

R

```
#import frame as F
#F$IV is a factor
#F$DV is a vector

t.test(F$DV~F$IV,…)
```

45

Is my data normal?

- Eyeballing histogram is a very crude measure.
- Inspect continuous probability density function.
- Inspect Q-Q plot.
- Run statistical test.

46

R: tests for normality

```
#Plot Q-Q line
>qqline(DV)

#Shapiro-Wilk Normality test
>shapiro.test(DV)
```

47

Homework

- Review B&A Ch 14 to 440-444 (power & effect size)
- HW I5 – due 2/27
  - Work individually on this one.
  - Start Part I Now – Write a complete study proposal

48

I5 Part 1

- Write a research plan for conducting an experiment comparing WizziWord vs. CoolText word processors using admins from BigBucks, Inc (both of these are new products).
- Outcome measures to include productivity (words per day output during the 8th week after the new word processors are introduced), and satisfaction, using the ILoveWordProcessors 12-item index (Cronbach alpha=0.82, test-retest correlation of 0.93, correlation with the standard 100-item WordProcessorsAreGreat index was 0.72).
- *Power Analysis (covered on 2/24):* From studies at other sites you expect to see a difference in productivity of approximately 3,000 (SD 1,200) words per day between the products. Assume a 50% response rate to your recruitment ad, and a 85% retention rate for subjects.

49

I5 Part 1 - continued

- Be sure to include the following in your plan:
  - Hypotheses
  - Research model (the boxes and arrows diagram) and description of variables/measures
  - Human subjects issues, including eligibility criteria, recruitment procedures, and the number of potential subjects you need to reach with your recruitment ad.
  - Detailed protocol , including recruitment, sampling and randomization methods
  - Analysis plan
- Refer to sample research plan for inspiration.
- Your complete plan should be 2-3 pages long.

50