BELLWORK
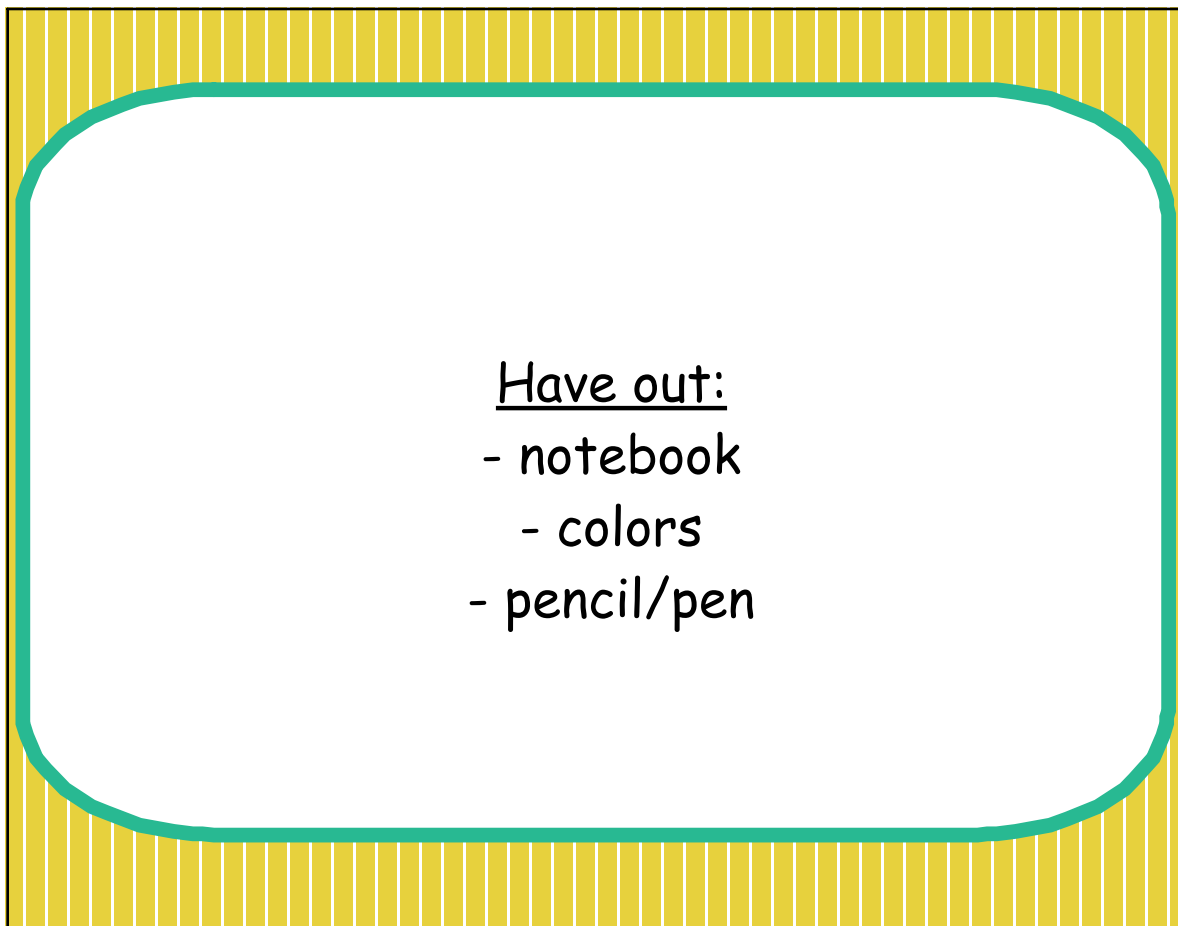
1. *flashback* The weights of adult men are normally distributed with an average of 170 and a standard deviation of 20 pounds.

    a. Where do 68% of the weights fall?

    b. How much do you have to weigh to be an outlier (in either direction)?

2. What qualifies as sampling bias?

Mar 25-12:52 PM

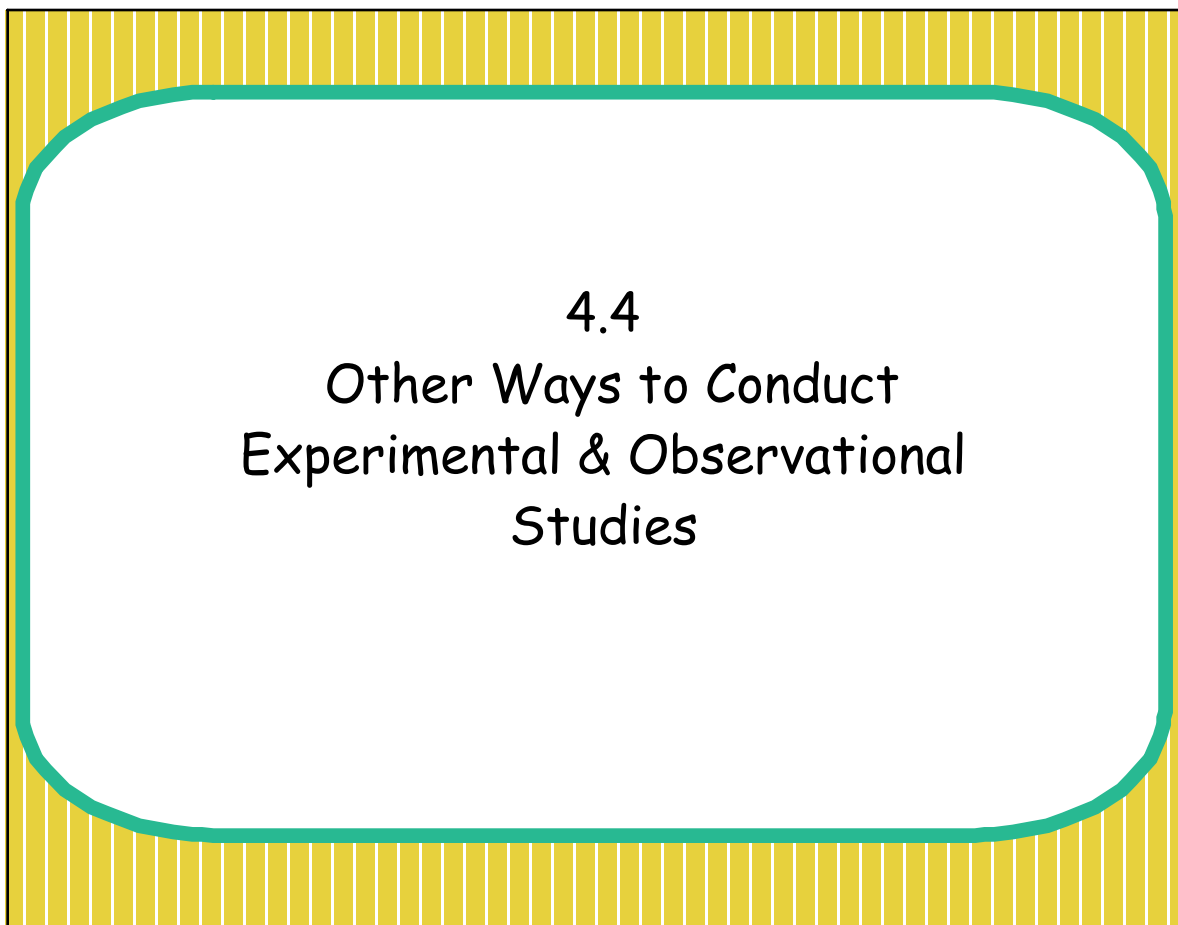Did anyone do the colored strips of paper?

Mar 25-12:52 PM

<u>Have out:</u>
- notebook
- colors
- pencil/pen

Mar 25-12:52 PM

4.4
Other Ways to Conduct
Experimental & Observational
Studies

Mar 25-12:52 PM

Learning Objectives

- Sample Surveys: Other Random Sampling Designs
- Types of Observational Studies: Prospective and Retrospective
- Multifactor Experiment
- Matched pairs design
- Randomized block design

Mar 25-12:52 PM

Random Sampling Designs

It is not always possible to conduct an experiment so it is necessary to have well designed, informative studies that are not experimental, e.g., sample surveys that use randomization

       Simple Random Sampling
       Cluster Random Sampling
       Stratified Random Sampling

Mar 25-12:52 PM

Cluster Random Sample

Steps

1. Divide the population into a large number of *clusters*, such as city blocks

2. Select a simple random sample of the clusters

3. Use the subjects in those clusters as the sample

"all from some"

Mar 25-12:52 PM

Cluster Random Sample

Examples

- city blocks
- cities
- schools
- hospitals

Mar 25-12:52 PM

Cluster Random Sample

Preferable when:

      A reliable sampling frame is unavailable
      The cost of selecting a SRS is excessive

Disadvantage:

      Usually need a larger sample size than with a SRS
      in order to achieve a particular margin of error

Mar 25-12:52 PM

Stratified Random Sample

Steps

1. Divide the population into separate groups (like age or gender), called *strata*

2. Select a simple random sample from each strata

3. Combine the samples from all strata to form complete sample

"some from all"

Mar 25-12:52 PM

Stratified Random Sample

Examples

    -Age
    -Gender
    -Year in High School
    -Counties in state

Mar 25-12:52 PM

Stratified Random Sample

Advantage:

    you can include in your sample enough subjects from
    each stratum to evaluate

Disadvantage:

    you must have a sampling frame and know the stratum
    into which each subject belongs

Mar 25-12:52 PM

Practice Example

Suppose a university has the following student demographics:

Undergraduate	Graduate	First Professional	Special
55%	20%	5%	20%

In order to insure proper coverage of each demographic, a sample of 100 students could be chosen as follows: select a SRS of 55 undergraduates, a SRS of 20 graduates, a SRS of 5 first professional students, and a SRS of 20 special students; combine these 100 students.
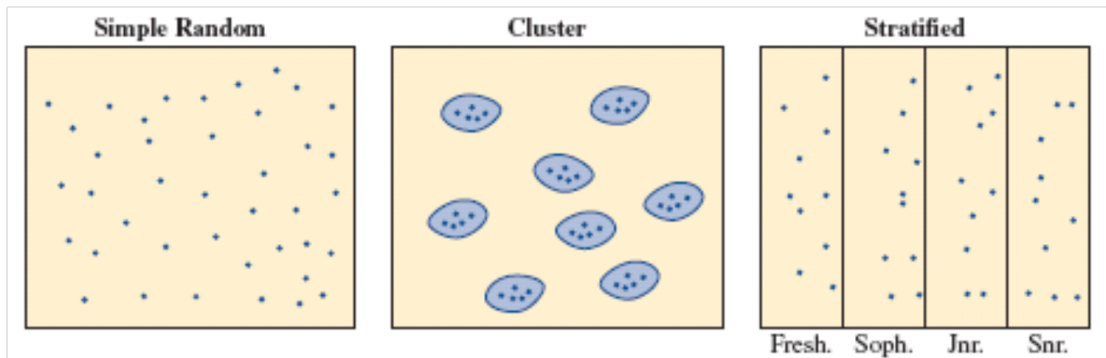
Mar 25-12:52 PM

Practice Example

Suppose that the Department of Agriculture wishes to investigate the use of pesticides by farmers in England. A sample could be taken by identifying the different counties in England. A sample of these counties would then be chosen at random, so all farmers in those counties selected would be included in the sample. Ask all farmers from the counties selected.

Mar 25-12:52 PM

Comparison

Simple Random	Cluster	Stratified

Fresh. Soph. Jnr. Snr.

Mar 25-12:52 PM

Examples

TABLE 4.2 Summary of Random Sampling Methods

Method	Description	Advantages
Simple random sample	Each possible sample is equally likely	Sample tends to be a good reflection of the population
Cluster random sample	Identify clusters of subjects, take simple random sample of the clusters	Do not need a sampling frame of subjects, less expensive to implement
Stratified random sample	Divide population into groups (strata), take simple random sample from each stratum	Ensures enough subjects in each group that you want to compare

Mar 25-12:52 PM

Types of Observational Studies

An observational study can yield useful information when an experiment is not practical.

Types of observational studies:

- *Sample Survey*: attempts to take a cross section of a population at the current time
- *Retrospective* Study: looks into the past
- *Prospective* Study: follows its subjects into the future

Causation can never be definitively established with an observational study, but well designed studies can provide supporting evidence for the researcher's beliefs

Mar 25-12:52 PM

Practice Example

Nurses' Health Study:
- Began in 1976 with 121,700 female nurses aged 30 to 55; questionnaires were filled out every two years

- Purpose was to explore the relationships among diet, hormonal factors, smoking habits, exercise habits, and the risk of coronary heart disease, pulmonary disease and stroke

- Nurses are followed into the future to determine whether they eventually develop an outcome such as lung cancer and whether certain explanatory variables are associated with it

prospective

Mar 25-12:52 PM

Knowledge Check

1. Identify the type of sampling used.
A market researcher selects 500 drivers under age 30 years of age and 500 drivers who are 30 years of age and older.

    a.   Simple random sample
    b.   Clustered random sample
    c.   Stratified random sample
    d.   Cross-sectional sample

Mar 25-12:52 PM

Knowledge Check

2. Identify the type of sampling used.
A market researcher randomly selects 500 people from each of ten randomly selected cities.

    a.   Simple random sample
    b.   Clustered random sample
    c.   Stratified random sample
    d.   Cross-sectional sample

Mar 25-12:52 PM

Knowledge Check

3.   Identify the type of observational study used.
An educational researcher used school records to determine that, in one school district, 84% of children living in two-parent homes graduated high school, while 75% of children living in single-parent homes graduated high school.

    a.   Cross-sectional
    b.   Retrospective
    c.   Prospective
    d.   Simple random

Mar 25-12:52 PM

Knowledge Check

4.   Identify the type of observational study used.
Among a group of married women who were tracked for ten years, those who worked full time were more likely to divorce than those who did not work full time.

    a.   Cross-sectional
    b.   Retrospective
    c.   Prospective
    d.   Simple random

Mar 25-12:52 PM

Homework

pg. 191 # 44,46, 47, 49

Mar 25-12:52 PM

BELLWORK

1. What is the difference between a prospective and a *future* retrospective study? *Past*

2. Cluster or <u>Stratified</u>? I have 3 different types of flower bushes in my yard. I want to test so<u>me of each</u> to see *all* which grows better with a new soil.

3. <u>Cluster</u> or Stratified? I want to see how teens in AZ are affected by alcohol, so I pick 30 high schools from the state and speak with every student there.

Mar 25-12:52 PM

WELCOME

Have out:
- Notes

Mar 25-12:52 PM

HW 4.4 Answers

44. Stratified – some from each group
46. a) i – Yes
     ii – No
    b) i – No
     ii – Yes
    c) Cluster samples the groups; stratified
       samples from within every group
47. a) You might get all small or all large accounts,
       but you want some of each
     b) Stratified – 5 from *each* group
49. Diet – social smoker perhaps? Maybe the fat causes
       the cancer and not the smoking.

Mar 25-12:52 PM

Notes continued...

Mar 25-12:52 PM

Retrospective Case-Control

A *case-control study* is a retrospective observational study in which subjects who have a response outcome of interest (the cases) and subjects who have the other response outcome (the controls) are compared on an explanatory variable

Mar 25-12:52 PM

Example

Response outcome of interest: Lung cancer
- The *cases* have lung cancer
- The *controls* did not have lung cancer

The two groups were compared on the explanatory variable smoker/nonsmoker

Lung Cancer

Smoker	Cases	Controls
yes	688	650
no	21	59
Total	709	709
Prob(smoker)	97%	92%

Mar 25-12:52 PM

Multifactor Experiments

Uses a single experiment to analyze the effects of two or more explanatory variables on the response

Categorical explanatory variables in an experiment are typically called *factors*

We are often able to learn more from a multifactor experiment than from separate one-factor experiments since the response may vary for different factor combinations

Mar 25-12:52 PM

Example

Examine the effectivness of both Zyban and nicotine patches on quitting smoking
- Two factor experiment
- 4 Treatments

Mar 25-12:52 PM

Example

How can living with another smoker affect your chances of quitting?

Mar 25-12:52 PM

Example

Subjects: a certain number of undergraduate students

All subjects viewed a 40-minute television program that included ads for a digital camera

Some subjects saw a 30-second commercial; others saw a 90-second version

The same commercial was shown either 1, 3, or 5 times during the program

There were two factors: length of the commercial (2 values) and number of repetitions (3 values)

Mar 25-12:52 PM

Example

The 6 combinations of one value of each factor form six treatments

		Factor B: Repetitions		
		1 time	3 times	5 times
Factor A: Length	30 seconds	1	2	3
	90 seconds	4	5	6

subjects assigned to Treatment 3 – see a 30-second ad five times during the program

after viewing, all subjects answered questions about: recall of the ad, their attitude toward the commercial, and their intention to purchase the product – these were the response variables.

Mar 25-12:52 PM

Matched Pairs

In a matched pairs design, the subjects receiving the two treatments are somehow matched (same person, husband/wife, two plots in the same field, etc.)

   - In a *crossover design*, the same individual is
      used for the two treatments

Aderol 30d
Rid. 30d

Randomly
      - assign the two treatments to the two matched
            subjects, or
      - randomize the order of applying the
            treatments in a crossover design

Helps to reduce effects of lurking variables
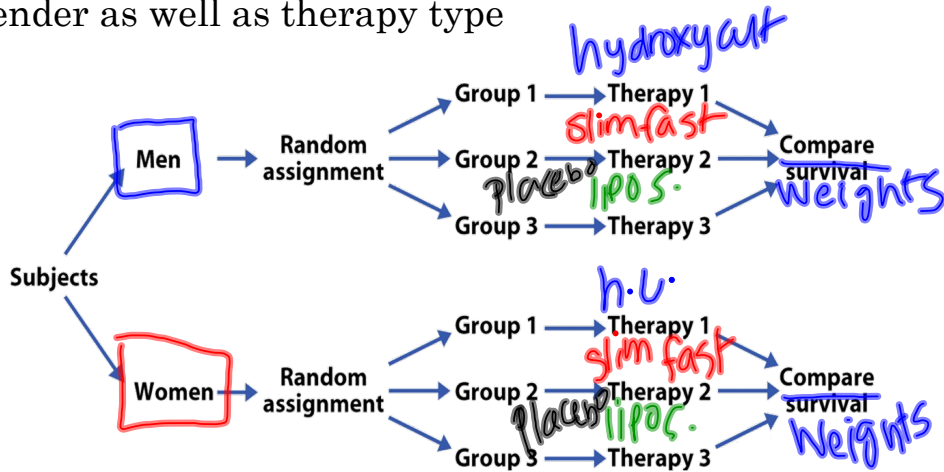
Mar 25-12:52 PM

Randomized Block

A *block* is a set of experimental units that are matched with respect to one or more characteristics (like "stratifying" it)

A *Randomized Block Design, RBD,* is when the random assignment of experimental units to treatments is carried out separately within each block

Mar 25-12:52 PM

Example

Block = gender; 3 treatments = 3 types of therapy
The men (as well as the women) are randomly assigned to the
3 treatments; differences can be compared with respect to
gender as well as therapy type

Mar 25-12:52 PM

Randomized Block Design

RBD eliminates variability in the response due to the blocking
variable; allows for better comparisons to be made among the
treatments of interest

A matched pairs design is a special case of a RBD with two
observations in each block

Mar 25-12:52 PM

Vocab Review

Simple Random Sample *(SRS)*
Cluster Random Sample
Stratified Random Sample
Prospective Study
Retrospective Study
Case-control Study
Multifactor Experiment
Matched Pairs Design
Crossover
Randomized Block Design

Mar 25-12:52 PM

Homework

Pg. 191 # 48, 50-54

Mar 25-12:52 PM

Bellwork

In a small city of 400 people, I am going to pick 50 to ask about transportation. Use the random numbers listed below to tell me the first 5 people with whom I would speak.

$$12346 \quad 50781 \quad 45983$$
$$96738 \quad 21235 \quad 29264$$

Mar 25-12:52 PM

Welcome

Have out:

- homework
- bellwork (turning in today)
- notes

Mar 25-12:52 PM

Follow up!

If I ask 2500 people my survey, what would my margin of error be? How would I interpret this if 24% of people in my survey were for a tax raise?

You better remember this!! ☺

Mar 25-12:52 PM

HW Answers

48. a) Report on *past* phone use
    b) Cases: Subjects had eye cancer
       Controls: Subjects did not
    c) i – 0.014
       ii – 0.010
50. Prospective – followed for 20 years
51. It's difficult to follow people into the future – costly – and we want answers now
52. a) Gender and type of diet versus weight loss
    b) You would have concluded that type of diet makes no difference in pounds lost
    c) Low *carb* works best for women and low *fat* works best for men

Mar 25-12:52 PM

HW Answers

53. a) Response: blood pressure, heart rate, stress
　　　Explanatory: Caffeine
　　　Units: 47 regular coffee drinkers
　　　Treatments: caffeine or placebo
　　b) Crossover – same person to both treatments
54. a) Crossover!
　　b) Double blind means that neither the subject
　　　nor the evaluator know which treatment the
　　　subject is receiving (though some executive
　　　somewhere does) – helps to eliminate bias!

Mar 25-12:52 PM

Review Answers

1. a.　1&2
　　　10 & 11
　　　8 & 9
　　　3 & 12
　　　4 & 7
　　　5 & 6

　b.　　2 & 10
　　　1 & 11
　　　8 & 12
　　　3 & 9
　　　4 & 5
　　　6 & 7

c. No - One member of each block to try each formula

Mar 25-12:52 PM

Review Answers

1.

<u>Calcium Group:</u>

Howard, Alomar, Kaplan, Zhang, Rodriguez, Rosen, Bennett, Liang, Curtis, Han

<u>Placebo Group:</u>

Clemente, Denman, Edwards, George, Guillen, Lawless, Marsden, Ogle, Thompkins, Underwood

Mar 25-12:52 PM

Review Answers

2. a) No - males and females have different ability levels

3. Age - We don't know the ages of these women, so this could be a lurking variable in the observation

4. nonresponse bias

5. units

Mar 25-12:52 PM

Review Answers

1. e	7. b
2. c	8. c
3. d	9. d
4. e	10. e
5. d	11. e
6. e	12. e

Mar 25-12:52 PM