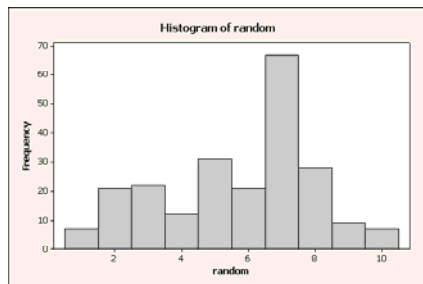


Chapters 11 and 12 – Randomness and Sample Surveys

Pick a Number...

1 2 3 4

What's Really Random?



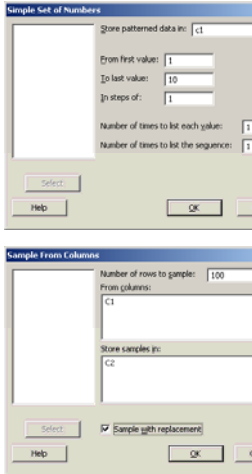
- Is asking people to pick a random number uniformly random?
- What would we want to happen if it were 'uniformly random'?

Random

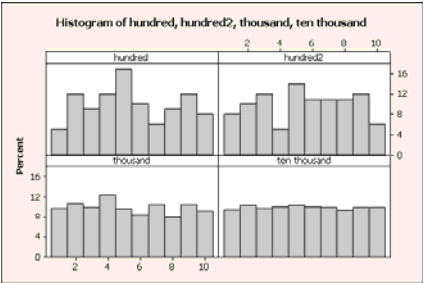
- We will call a variable **random** if we know the possible outcomes, but not the particular value a variable takes.
- How many times do you need to shuffle a deck of cards to get randomness?
- Random numbers generated by a computer are 'pseudorandom' because they use recipes to find numbers that appear random.

Minitab: Generate Random Numbers Uniformly

- First put the numbers 1 to 10 in C1
 - Minitab > Calc > Make Patterned Data > Simple Set of Numbers
- Then generate a random sample of 100 values from that column
 - Minitab > Calc > Random Data > Sample from Columns
 - Select the number of samples (100)
 - From Columns (C1)
 - Store in (C2)
 - CHECK sample with replacement
- Repeat for another 100 (store in C3), 1000 (store in C4) and 10000 (store in C5) samples. Plot histograms of C2, C3, C4, C5.



Just Checking...



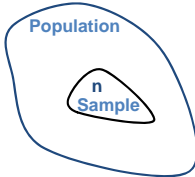
- Do you expect all of our graphs to be identical? Yours versus your neighbor's?
- What do you notice?

More About "With Replacement"

- With Replacement
 - Every time we draw a sample, we want it to be taken from the same pot as the original
- Without Replacement
 - We don't want the possibility of a repeat in the sample
- With or Without?
 1. Simulating rolling a die 500 times
 2. Taking a sample of 30 math 140 students for an opinion poll
 3. Simulating rolling 500 dice one time

Population vs. Sample

- The **population** is what we really want to study.
 - We often don't have the time or money to study the entire population.
 - A **parameter** is a function of population data
- Instead, we examine a **sample**, a smaller group of individuals selected from the population.
 - **Ex:** Tasting a bottle of wine
 - **Ex:** Blood test at the doctor's office
 - A **statistic** is a function of sample data.
- Sampling the entire population is called a _____.



Bias in Sampling

- Sampling methods that under- or over-represent a part of a population are said to be **biased**.
 - Would we determine President Obama's national approval rating based on a survey administered in San Francisco? In Texas?
 - Who should I ask to evaluate math 140? 'A' students? 'F' students?
- Some types of bias in sampling:
 - Nonresponse – do you always complete a survey you're given?
 - Guess how many of you didn't complete the teammate evaluation...
 - Voluntary response – what about the people that do?
 - Did you know you could make money for taking surveys?
 - Convenience sample – sample those who are convenient
 - Guess the worth of these...

Avoid Sampling Bias: Randomize

- Cooking soup analogy in textbook
 - Pile in ingredients and then taste? Or mix before tasting?
 - Soup for two? How much do you need to taste? Soup for 12?
- Randomizing protects us from the influences of all the features of our population by making sure the sample looks like the rest of the population (on average)
- In a **simple random sample (SRS)**, each combination of people for a given sample size has the same chance of being selected.
 - List of individuals in population: **sampling frame**
 - In Minitab, we can put the sampling frame in a column and select a sample from it (without replacement).
 - **Sample size** matters, not the fraction of the population
- Note: Sampling can get complicated: systematic, cluster, stratified, ... We will focus on analyzing SRS.

How Representative Is It?

- Each sample will be different. Recall our sample size 100 on the numbers 1 to 10...
- When we compute statistics on different samples, we get different values. Therefore, the statistics are random.
- The sample-to-sample differences are called **sampling variability**.
- We will try to quantify that using probability over the rest of the course.

So... What Else Can Go Wrong?

- From Monday, December 15, 2008 Colbert Nation



- <http://www.colbertnation.com/the-colbert-report-videos/216371/december-15-2008/great-president---or-greatest-president->

Response Bias

- **Response bias** is anything in a survey design that influences responses.
- Wording of questions
 - Do you support the President? Do you support Obama?
 - Freedom of Speech...
 - Do you think the US should forbid public speeches against democracy? 21.4% said yes
 - Do you think the US should allow public speeches against democracy? 47.8% said no
- General vs. Specific Questions.
 - How much do you usually sleep?
 - Do you sleep a lot, a little, or a moderate amount?
 - How many hours did you sleep last night?
- Phrasing Answers
 - Like Colbert
- Interviewers can introduce bias as well.

Math 140 Survey

- In part, data for Project 1 was hard to analyze because of the questions, not the content.
- So... What was so bad about the questions?
 - Year/age,
 - Glasses,
 - GPA,
 - Hair color, ...
- You will get to design your own survey for Project 2.

11. Which is your dominant hand?
 Left
 Right

12. Do you own an iPod?
 Yes
 No

13. How many hours did you sleep last night?

14. How many units are you enrolled in this semester?

15. Did you attend high school in California?
 Yes
 No

16. How many minutes does it take you to get to school?

17. What is your COC GPA?

18. Do you plan on transferring to a four year college?
 Yes
 No

19. How many hours do you work in a typical week?

20. Do you smoke?
 Yes
 No

21. How many alcoholic beverages do you drink in a week?

22. With which political party do you identify?
 Democratic
 Republican
 Other

Analyzing Exam 1 (Next Time)

- What do we want to know?
- Design Survey
 - Who should we ask?
 - What should we ask?
 - Will the questions we ask address what we want to know?
 - Is there any bias we need to worry about?

Class Work

- To get credit, it is your responsibility to get checked off.
 - Chapter 11 and 12 Handout
 - Rules for checking answers: No Pens in the Front!!!

Homework

- Textbook/Routine Homework
 - Due Next Week (25% chance of collection)
 - Read Chapter 11 and 12
 - Pg 299: #1, 3, 9
 - Pg 324: #3, 5, 7, 11, 17, 25, 29
- Project/Exploration Homework
 - Project #2 Coming Soon – Groups Optional