Curriculum Units by Fellows of the Yale-New Haven Teachers Institute
1985 Volume VIII: The Measurement of Adolescents

The Statistics Sampler

Curriculum Unit 85.08.04
by David B. Howell

"Do you teach statistics?" asked a sampler.

"You mean beyond graphs and charts and mean, median, and mode?" responded several math teachers covering grades 7-12.

"Yes, I mean beyond those 'basics.' But I don't mean theory."

"Then, no, I don't," answered every one of a non-random sample of ten teachers.

"Should you?"

"Well . . . , yes . . . , probably. Advertising, polls and marketing surveys keep coming at us."

"Could you?"

"I'm not so sure. Know any good resource books? I didn't study much statistics. Didn't like it, either."

"The list of good resources is growing rapidly since the NCSM 'Position paper on Basic Skills' and the NCTM *Agenda for Action* appeared. Statistics is basic. And important at all levels. That's what they said. Say, how about a unit on statistical sampling? Mo fancy theory. Just 'seat-of-the-pants' experience combined with a bit of mathematical stretching . . . "

" *That* would help. Where can I find one?"

Here . . . .

Much of our daily experience with the news media and in the marketplace is impacted by statistical information based on sampling. Newspapers and magazines regularly headline lists of favorites—most popular singers, most admired leaders, best dressed men and women, finest chocolate ice cream. The invariable economic interest behind such lists often has a direct sales emphasis. It is important that we be able to analyze the claims: On what is the claim based? Who was sampled? How many were sampled? Does the sample represent the population *I* belong to? Do the sample results support the conclusion?

Quality control in manufacturing makes extensive use of statistical sampling. A company producing light bulbs, or audio tapes, or batteries must assure that no more than some small percentage of its products are defective. It *must* sample. It cannot test every item. How many items must be sampled? If some percentage of the sample is defective the company halts production to fix the problem. How does one determine the cut-off percentage?

Advertising for products, services, and even for political candidates increasingly is based on inferences made from opinion polls and sampling. How can pollsters make their claims for millions of voters on the basis of samples of one thousand? TV ratings are based on small samples. “Look out for ___ ,” headlines a magazine ad. “It’s the ___ most often chosen #1 in the most recent . . . magazine tests . . . .” Could the claim be based on a single number one rating in one magazine in one month? Or two number one ratings in 20 magazines in two years? What was the sample? How was it chosen? What constitutes a claim we have confidence in?

It is possible, of course, to provide extensive theoretical mathematical answers to many of the questions above. possible, but not necessary for most of us. The role of probability in the theory of statistics is critical. Sampling, confidence limits, and so on depend on binomial and other probability distribution. It is not necessary, either, to have any background in theoretical probability to make reasonable analyses of many statistical claims. It is sufficient simply to have some experience with sampling and to have addressed a few of statistics’ major concepts. That is what this Unit is about.

It is the purpose of this Unit to offer students a series of lessons and activities which provide a concrete base for understanding statistical sampling. Students will conduct sampling activities. They will compare results across samples. They will predict based on samples. They will attempt to verify the predictions. They will explore the notion of “confidence.”

There are a few important skills assumed in the following lessons: counting, tallying, ratio, percent, graphing (especially histogram or bar graph), mean. Median and mode might be included in discussion, too. If students do not have all these skills in advance, the lessons may offer an interesting context in which to teach them informally. Use of calculators for converting ratios to percents and for finding averages is encouraged in order to keep the focus on the statistics rather than the arithmetic.

Let’s state the objectives for the Unit. The Unit is designed for any group of regular education students in Grades 7 through 12. Students will be able to

1. define sample of a population;

2. describe and give an example of a random sample;
3. list two or three factors which could have compromised the randomness of a given sample;
4. describe variations among several samples;
5. describe how changing sample size affects those variations;,
6. select a sample size to estimate a characteristic of a large population with confidence;

7. discuss ways to quantify confidence. Let's talk about a little bit of mathematics, *for the teacher only* , to illustrate what we're getting at.

The *Central Limit Theorem* is of fundamental importance in inferential statistics, and it is one theorem for which we are trying to build an intuitive sense based on concrete examples. The Central Limit Theorem may be summarized as follows:

Take samples of size, H, from *any* population:

1. The means of those samples have an (almost) normal distribution.

2. The mean of that distribution—i.e. the mean of the sample means—is (almost) the same as the mean of the entire population.

3. The distribution of the sample means clusters more tightly around its mean as the size, N, of the samples is increased. Example: What is the percentage of foreign-made passenger cars on the road in the U.S? By counting at pseudo-random (beginning at the next overpass) stretches of turnpike I drive, I take samples to approximate answers to that question regularly. Here are the results of four such recent samples:

TABLE I.

1	2	3	4
		Number of	percent of
Sample group	N	Foreign Cars	Foreign Cars
A	50	7	14
B	50	18	36
C	50	11	22
D	50	13	26

Let's form groups of 100 by making all possible combinations of two of the four groups above.

TABLE II.

1	2	3	4
		Number of	percent of
Sample group	N	Foreign Cars	Foreign Cars
AB	100	25	25
AC	100	18	18
AD	100	20	20
BC	100	29	29
BD	100	31	31
CD	100	24	24

The Central Limit Theorem assures us of three things.

1. The means—percent of foreign cars—of all the possible samples that could be in Table II are distributed (almost) normally.

2. The mean of Column (4) of all the possible samples in Table II is (almost) the same as the mean of the entire population. Directly, the mean of the samples listed is 24.5%. That is probably close to the mean of the entire population. Assuming my sample was random (O.K, it was taken on the Maine Turnpike in April:), approximately 24.5% of the passenger-car population in the U.S. is foreign-made.

3. The larger the sample size N, the more the means cluster centrally. The range for samples A,B,C, and D (N=50) is 36 14 = 22. The range for the six samples in Table II (N=100) is 31 18 . 13. The means are, of course, the same: 24.5%. Now, if we apply the Central Limit Theorem and if we standardize scores to the normal distribution and if we formulate null and alternate hypotheses and if we calculate variance and standard deviations and t-ratios and degrees of freedom, then we can quantify the notions of confidence intervals and confidence limits which are also among the major foundation stones of inferential statistics. But hold on: You, the teacher, may still be here, but most students in Grades 7 12 aren't.

Consider an analogy. When Geometry students reach the pons Asinorum—the theorem that says the base angles of an isosceles triangle are congruent—the reaction to the claim at least is one born of familiarity. "Sure," thinks the student, "I've measured it every year since the fifth Grade, and this teacher draws it that way every time. Of course it's true." And now we're free to focus on the logic of possible proofs.

But not so with the statistics of the foreign car population. What we need to do is take samples. Compare samples. Combine samples. Examine possible inferences from samples. Declare our faith that while we're not too sure about there being exactly a 24.5% foreign car population out there, we'd bet a whole lot that there's more than 14% and less than 36%. *That* 's confidence—intervals, limits, and all. And that's what this Unit's lessons are about.

Each of the following Lessons should take from one to three class periods depending on task efficiency in sampling, sophistication of discussion, the skill levels for charting, graphing, finding percent, combining smaller samples into larger ones, etc. The Lesson outlines deal with content, not management or individual differences or testing. All of the Lessons are described in the same format:

A. Objectives
B. The experimental question—a question which presents a statistical sampling opportunity.
C. Issues, and some possible resolutions—a series of questions which should arise, which the teacher will probably pose (with some possible student responses in parentheses); a hint of dialogue between teacher and class. There is occasionally a direct comment to the teacher in brackets. This section really defines the activity.
D. Observations and discussion to Objectives—more questions (possible student responses in parentheses) and summary relating specifically to the stated objectives.

Lesson 1.

A. Objectives—students will be able to
   1. define sample;
   2. describe and give examples of random sample;
   3. list factors compromising randomness;
   4. describe variations among samples.

B. The experimental question—What percent of the population is taller than 5’7” ?

C. Issues, and some possible resolutions— *What* population? (This class. 9th Grade. The school. Hew Haven. The world.)

Let’s start with this class. How do we find out? (Measure everybody. Ask everyone. put a mark at 5’7” and check everyone against it.)

Let’s just ask. If you don’t know, we won’t count you this time. [Record data on Worksheet 1 or board or overhead. Here is a sample showing class data for everyone who knows his/her height and just for the subgroup of boys:]

(figure available in print form)

Is that a fair picture of this whole class? (Yes. No. It depends whether we counted everyone.) Did we have to count everybody? What happens if we count only boys? (Boys are taller than girls.) Record it. [See above.] Only girls? The second row? Each row separately? A coin-toss group? [We’ll toss a coin for each person; only students matched with ‘heads’ get counted.] A coin-toss group again? [Re-toss.]

D. Observations and discussion to Objectives— Each different group we counted was a *sample* of the population. A *sample* is simply a part or subset of a group or population. A *random sample* is a sample in which every member of the group or population has the same chance of being chosen.

Which of our samples—if any—were clearly random samples? [Any coin-toss group is generally random. Other techniques might also have been used.] Which were clearly HOT random? About which aren’t you sure? What might have made a coin-toss group HOT random? (Unfair coin. Unfair toss.) How might we have made the Second Row group “more” random? (Assign seats randomly—by lot.) Did all samples give the same or almost the same results? Describe how they were the same. Describe how they were different.

Lesson 2.

A. Objectives—same as Lesson 1.
B. The experimental question—same as Lesson 1.

C. Issues, and some possible resolutions— [Try to use a larger population, such as all Grade 10 students or the entire school.]

Can we realistically count everyone? What samples could we use? How random are they? How will we do the survey? How will we record the data? Is it appropriate to use as samples other classes of mine? Other classes of yours? A table-full of students in the cafeteria?

[Record data on Worksheet 1. Have teams of two students take a sample of 10. Or use techniques suggested by the class.]

D. Observations and discussion to Objectives— Several issues have been raised. What is the possibility of membership in more than one sample? Does it matter? [This is, of course, the issue of sampling with or without replacement. For large populations and relatively small samples, there isn't much difference.] What about the impact of cliques on even the objective measure of height? Would it be likely, for example, that a group of tall people would sit together in the cafeteria? Or a group of all girls? Did everyone collect data the same way? Were different questions used, for example? Did everyone record the data the same way?

Lesson 3.

A. Objectives—students will be able to
1. define sample;
2. describe and give examples of random sample;
3. list factors compromising randomness;
4. describe variations among samples.

B. The experimental question—What percent of the population prefers pepsi to Coca-Cola?

C. Issues, and some possible resolutions—

What population? (This class. The school. Hew Haven. Kids. people over 40. Basketball players.)
How do we find out? (Ask. Give a questionnaire. Do a taste test!)

Let’s start with this class again. And let’s do a taste test. [paper cups, chilled cans of pepsi and Coca-Cola, coding on cups, pouring out of sight of tasters, etc. -here are many fair testing issues. Does it matter which is tasted first? Which is on the right? Chilled or warm? Influence of other tasters? Are the results significantly different than chance? All relate to objectives about sampling procedures and/or tests of significance. None is central to this Unit. Prepare for a long detour if the issues become important to your class. To stay on track, establish a procedure with a minimum of debate, emphasize the need for procedural consistency, and get on with it.]

Use Worksheet 1 to record data. Answer 1 is “Prefer pepsi.” [Note there are at least *two* classes of answers not recorded: “prefer Coca-Cola” and “Ho preference.”]

[Form samples in different ways within the one class. The simplest techniques to add together different combinations of small samples to get larger ones. Students who are experienced in listing combinations might list all possible combinations of 3 small samples and record their data. Remember, however, that this Lesson does HOT center on combinatorics! Compare samples as in Lesson 1.]

D. Observations and discussion to Objectives—

Did we collect data from the entire population? What is the difference between a sample and the entire population? Which of the samples were random? Which were not? What is the difference in the way they were chosen? Describe the differences in results across samples. Can any differences be attributed to whether or not the samples were chosen randomly? What about the range of results among samples? Does sample size matter?

Lesson 4.

A. Objectives—students will be able to
  1Ð4. same as Lesson 1.
  5. describe how changing sample size affects variation among several samples.
B. The experimental question—What percent of the population read “Ebony” in the last week?

C. Issues, and some possible resolutions— Let's shoot for the whole student body as our population.

What samples shall we use? (This class. My other classes. Everybody's English class. Our homerooms. All period One classes.) Some of the samples strike me as more random than others. Let's discuss those issues.

Let's take a coin-toss group of classes from a period when all grade levels have classes. We must know two things—how many people in the class and how many read "Ebony" in the last week. We might also want to record what kind of a class and what Grade students are in it so we can analyze our sample for randomness. To get the percent we'll use Worksheet 1 again.

[When the data is collected, we'll have Tables something like the foreign car Table I above. The key idea is to compare the range and mean of similar size samples and of groups of different size samples. proceed as follows: for as many samples of 20 25 as possible, calculate the mean. List the lowest and highest values. Calculate the range. Now, using another Worksheet 1, form new samples by combining the original ones into groups of 40-50. Recalculate the Ratio and Percent columns. Repeat the calculations for means, limits, and range.]

D. Observations and discussion to Objectives—

[In addition to further discussion of the randomness of the samples and of the variations among samples, focus attention on the impact of sample size. If necessary or appropriate, continue to combine samples and to examine the changes in mean (no change!), range, and limits.]

Lesson 5.

A. Objectives—students will be able to

1-5. same as Lesson 4.

6. select a sample size to estimate a population characteristic with confidence.

B. The experimental question—What percent of the population read "Ebony" in the last week?

C. Issues, and some possible resolutions— [Let's assume a student body of 500 to 1500] We have several samples, ranging in size from 20 25 up to combinations of samples totalling 200 or more. Suppose we work for "Ebony," and it costs us a lot of money to sample each additional group of 20 25. At what sample size would you be willing to stop? When will you say that your sample is large enough for you to predict what percent of the total student body reads "Ebony?" (My answer with 10 seems right. Our class sample of 23 is close. Many groups of 50 seem OK. All the 100 groups are close together—maybe one would have been enough.)

D. Observations and discussion to Objectives—

Do we need to predict exactly? Is a range of 5 or 10 percent good enough? Does your answer depend on how wide a range is sufficient? (Sure! The single class range is 40%. The 50 group range is only 20%.) Is there a sample size beyond which we don't seem to be getting any better information? (The 200s don't look much better than the 100s. Maybe we could stop at 100. What do we need to know for—printing quantities? The prediction from almost all the 50s is good enough.)

Lessons 6 and 7.

A. Objectives—students will be able to

1. define sample;
2. describe and give examples of random sample;
3. list factors compromising randomness;,
4. describe variations among samples;,
5. describe how changing sample size affects variation among several samples;
6. select a sample size to estimate a population characteristic with confidence.

B. The experimental questions—What percent of the population watches MTV? What percent of the population goes to bed before 10:30? (Or, perhaps, what percent of the students' parents have visited the school?)

C. Issues, and some possible resolutions— [It would be interesting and challenging to try to define the population as larger than the student body. (CAUTION: Hot everyone reacts positively or gracefully to being polled about almost anything! Before you prepare to have students poll outside the school population, by telephone or in any other way, get administrative approval.) If the class is up to that challenge, there will be substantial discussion about how to choose samples, how to deal with the random issue, what size sample to pick, how to poll the samples, etc. We want students to maximize the size and general nature of the population that they can reasonably sample. But predicting Connecticut's MTV audience on the basis of a sample of a math class of 14-year-old urban students will hopefully strike the students as risky! And developing a reasonably accessible, even moderately random, sample of the American population should strike students as beyond their resources.]

D. Observations and discussion to Objectives— On what basis did you choose the sample size? If we have enough different samples to "pool," let's do the same kind of analysis we did with Lesson 5. In any event, since we are dealing with a larger and more general population than our own school peers, is there any change in our feelings of confidence about predictions based on samples? Can we identify other factors that impact the change? [Difficulty with or uncertainty about randomness should be one—the larger the population and characteristics that might be studied, the more difficult it is to solve practical problems of random selection.]

Lesson 8.

A. Objectives—students will be able to
   1. define sample;
   2. describe and give examples of random sample;
   3. list factors compromising randomness;
   4. describe variations among samples;
   5. describe how changing sample size affects variation among several samples;
   6. select a sample size to estimate a population characteristic with confidence;
   7. discuss ways to quantify confidence.

B. The experimental question—What percent of the colored cubes in the box are red? (or colored paper squares)

C. Issues, and some possible resolutions— *[Materials* : A box of at least 1500 objects; some of them are red. I used a box of commercial "cube-o-grams," which are cubic centimeters of various colors. Twenty sheets of various colored construction paper cut into square inches would do just as well. So would a lot of M&Ms, though the sampling procedures would have to be sanitized. It is important that the objects be essentially identical except for color and easy to scramble so that a sampler, reaching into the box without looking, can select a random sample. It is interesting, but not essential, to have a variety of colors represented. The teacher will be more convincing if (s)he doesn't know the exact percentage of red ones, either.

At least 3 copies of Worksheet 1 per student.

At least 2 sheets of graph paper per student.

With this “pure” example we will try to make explicit the ideas developed in earlier lessons. We can examine sampling and results without subjective distractions (“I don’t care what happened; I like pepsi best!”) Hopefully, we have done enough unverifiable predicting from samples earlier that we can resist the temptation to verify here. Use of a bigger population reduces the temptation, of course.]

Here in this box are several thousand residents of the planet Colsquar. Weird little creatures . . . , they look like this [display one]. They come in many colors; it is easy to assign them to different groups based on their color just as we can assign people in America to different groups based on income, or years of school, or nationality of ancestors, or race or religion, etc. Anyway, let’s pretend that we work for a company that produces rock music hits on bright red records. Red residents of Colsquar like red records! The question is—What is our potential market? What percent of the population of Colsquar is red?

Each of you should take a sample of 10 objects. [It makes an interesting side trip to have students predict just what classes, or colors are in the total population. With my set of cubes there are at least 12 colors. How many samples of 10 need be pooled to get at least one cube of each color? Generally, about 5, or a sample of 50, is sufficient; it appears that a couple of colors each represent only 2-4% of the total.] Record your results on Worksheet 1. Now take a second sample of 10. Record it. And a third sample.

[Here is an example of a student’s worksheet at this point:]

(figure available in print form)

Let’s make a list of several samples of 10 from the class. List on the board or overhead. Here is an example:

(figure available in print form)

What would you predict is the percent of red in the total population? (10%. 5 to 40%. 20%. Less than 50%. More than 10%. How about the average of all of those?)

Let’s make a graph—a histogram, actually—of all of our samples of 10.

[Here is an example:]

(figure available in print form)

Does the histogram help you predict?

On your worksheet, add the results of each sample of 10 so you have a record for a sample of 30. [Referring to the DH samples earlier, we’d have

(figure available in print form)

Let’s graph several samples of 30.

[Here is an example:]

(figure available in print form)

What do you want to predict now? Does anyone want to predict less than 7%? More than 17%? How *sure* are you of your prediction range?

Let’s pool samples of 10 into samples of 50. List 20 or 30 groups of 50, drawn from various 10-samples from the class, on the board. How graph this.

[Results might look like this example:]

(figure available in print form)

What prediction now? How sure are you? Are there any results you want to rule out as not likely?

List the range [low to high] for your samples of 10. And the mean. How do the same for the class list of samples of 30. And the class list of 50s.

[Here is an example of the results:]

Sample size	Range (percent of red)	Mean
10	0Ð40%	13%
30	0Ð23%	13%
50	6Ð22%	13%

D. Observations and discussion to Objectives—

What happens to the range of possible results of several samples as the sample size increases? (It decreases.) Would you have expected that? Why or why not?

What happens to the mean of several samples as the sample size increases? (It doesn't change much.) Would you have expected that? Why or why not? [The Central Limit Theorem is being illustrated here.]

On the basis of *one* sample of 50, what would you be willing to predict about the percent of red cubes in the total population? How sure are you?

[In the answers to the last set of questions, we have done at an intuitive level exactly what the political pollsters and other professional samplers do. We have taken a small random sample of the population (50 of a few thousand) and predicted that the total population has a characteristic (red) at, say, the 13% level, plus or minus 5% (range of 8 to 18%) or perhaps even smaller. We have made that prediction with considerable confidence (unquantifiable just yet) based on our examples.]

The series of Lessons is concluded. Students should have met the Objectives. There are plenty of follow-up activities possible. One might wish to do more, similar, sampling activities around research questions students pose. What percent of a population is left-handed? Owns a Michael Jackson record or tape? Plans to attend college? It might be valuable to pursue other topics suggested in discussions of the activities above. One might want to return to probability, for example, to begin to build a more formal set of tools for analysis of these same activities. Wherever . . . The base of experience established here should make the task easier for both student and teacher.

Worksheet 1

(figure available in print form)

BIBLIOGRAPHY and REFERENCES

Primarily for teachers:

1. Edwards, Allen L. *Statistical Analysis* . New York. Holt, Rinehart and Winston, Inc. 1969.

Readable and without heavy theory. Good examples.

2. Mason, Robert D. *Statistical Techniques in Business and Economics* . Homewood, IL. Richard D. Irwin, Inc. 1978.

Good applications. Chapter 11 (An Introduction to Statistical Inference) takes a relatively informal approach to sampling and confidence limits.

3. National Council Of Supervisors of Mathematics. "Ten Basic Skill Areas". Milwaukee, WI. HCSM. 1977.

This is what school mathematics should be about.

4. National Council of Teachers of Mathematics. *Teaching Statistics and probability* (1981 Yearbook). Reston, VA. NCTM. 1981.

A very useful collection of course samples, activities, applications and more. Chapters 3, 7, 25, and 26 offer specific ideas related to sampling.

5.———-. *Mathematics for the Middle Grades (5-9)* (1982 Yearbook).

Excellent activities, several of which involve descriptive statistics.

6.———-. *The Agenda in Action* (1983 Yearbook).

Several articles outline statistics activities.

7.———-. *An Agenda for Action: Recommendations for School Mathematics of the 1980s* .

The recommendations for reform include statistics.

8. Runyon, Richard p. and Audrey Haber. *Fundamentals of Behavioral Statistics* . Reading, MA. Addison-Wesley publishing Company. 1984.

Readable, nontheoretical college statistics text. Good examples drawn from the behavioral sciences.

Primarily for students:

9. Herrick, Marian Cliffe, Jane Zartman and Thomas R. Conrow, Jr. *Mathematics for Achievement Individualized Course 2, Book 8 Introducing Statistics* . Boston, MA. Houghton Mifflin Company. 1972.

Excellent worksheets guiding activities in basic descriptive statistics.

10. Johnson Donovan A. et al. *Sampling and Statistics* . Glenview, IL. Scott, Foresman and Company. 1972.

Lessons 2 and 3 provide sampling exercises.

11. Srivastava, Jane Jonas. *Statistics* . New York. Thomas Y. Crowell Company. 1973.

Readable, entertaining, activity-oriented book aimed at middle grade children.

Mixed audience:

12. American Statistical Association. *Quantitative Literacy project* . Work under development, 1985.

A series of four booklets by the Joint Committee on the Curriculum in Statistics and probability (ASA-NCTM). 1. Exploring Data 2. Probability 3. Simulation 4. Information from Samples: Yes/No Populations.

13. Blakeslee, Donald W. and William G. Chinn. *Introductory Statis-* tics and Probability. Boston, MA. Houghton Mifflin Company. 1971.

Designed for a high school course, this text is a bit heavy on the mathematics. The examples, however, especially in Chapter 9 (Sampling), are helpful.

14. Fehr, Howard F, Lucas H.H. Bunt and George Grossman. *An Introduction to Sets, Probability and Hypothesis Testing* . Boston, MA. D.C. Heath and Company. 1964.

Chapters 5 and 6 develop ideas of sampling and confidence limits quite clearly. Good examples are cited. The theory is aimed at Grade 12 or college students.

15. Hooke, Robert. *How to Tell the Liars from the Statisticians* . New York. Marcel Dekker, Inc. 1983.

A collection of very short, entertaining anecdotes involving applications of statistics. Comparable to Huff's classic.

16. Huff, Darrell. *How to Lie with Statistics* . Hew York. W.W. Norton and Company, Inc. 1954.

An entertaining review of basic statistical applications.

17. Johnson, Donovan A, et al. *Activities in Mathematics, 2nd Course: Statistics* . Glenview, IL. Scott, Foresman and Company. 1971.

Solid ideas for specific lessons on many topics, especially Lesson 17: The Sampler. Aimed at Grade 7 up.
