CCSS Advanced Algebra 4 AA8_3: Sampling

1. **Polls**

A recent poll stated that 42% of Americans approve of the job that Donald Trump is doing as President. This poll was based on a random sample of $1,500$ adults living in the United States.

The 42% corresponds to a proportion of 0.42, and is called a sample proportion. It is an estimate of the proportion of all adults who would say they approve of the job that Donald Trump is doing as President.

a. If you were to take a random sample of 20 Americans, how many would you predict would say that they approve of the job Donald Trump is doing as President?

$.42 \cdot 20 = 8.4$ So, about 8 people

b. If you were to take a random sample of 500 Americans, how many would you predict would say they they approve of the job Donald Trump is doing as President?

$.42 \cdot 500 = 210$ about 210 people

c. Polls from the 2016 election did not predict the final results of the election. With your table groups, discuss reasons the polls from the 2016 election could have gone wrong.

Sampling bias. Voluntary response.

2. **Random Sampling**

One of the issues with collecting data from an entire population is the cost and difficulty with reaching every person in an entire population. For example, consider the population of Cleveland High School. How difficult would it be to conduct a survey, observational study or experiment on every student currently enrolled?

a. Most data is collected using a random sample. What is meant by a sample of a population?

A part, not the whole.

b. What is meant by a random sample of a population?

Each member has equal chance

c. Indicate whether the following are random samples from the given population, and explain why or why not.

i. *Population:* All students in school; the sample includes every fifth student in the hall outside of class. Not random. Just kids in the hall

ii. *Population:* Students in your class; the sample consists of students who have the letter s in their last names. Not random. Just kids with "s", not all.

iii. *Population:* Students in your class; the sample is selected by putting their names in a hat and drawing the sample from the hat.

Yes random. Each has equal chance.

iv. *Population:* People in your neighborhood; the sample includes those outside in the neighborhood at 6:00 p.m. Not random. Some on inside.

v. *Population:* Everyone in a room; the sample is selected by having everyone toss a coin, and those that result in heads are the sample.

Yes random. Each has equal chance

d. Formulate a plan that might allow you to find a random sample of 50 students from Cleveland High School. Be specific and detailed. Student ID #s in a hat.

3. **Using Samples**

M&Ms come in 6 colors: Brown, Yellow, Red, Orange, Green and Blue. Do M&M bags contain equal numbers of each color or are some colors manufactured more than others?

a. How could you use a sample to decide the distribution of M&M colors manufactured?
Open 25 packs and count colors.

b. Each group is given some M&Ms. Does this represent a random sample of all M&Ms made? Explain why or why not. No. Could only be from one M&M factory.

c. Do you expect most bags will have equal numbers of each color? Explain why or why not.
No. Some sample to sample variability

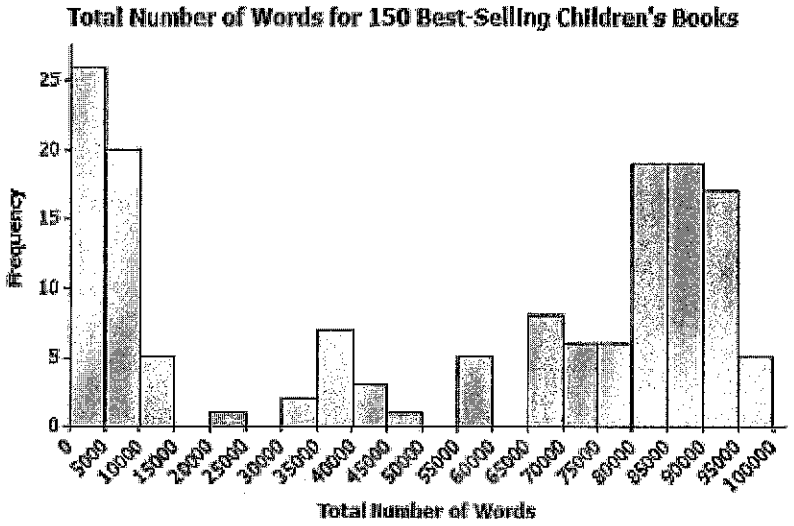
d. Use your sample to predict the number of orange M&Ms in the 42 ounce bag sitting at the front of the room. SKIP

e. M&M/Mars reports that the colors are distributed as follows: 24% blue, 14% brown, 16% green, 20% orange, 13% red, and 14% yellow. How do these proportions compare to what you found in your samples? What could you do to get a sample that might be more representative of the proportions reported by M&M/Mars? SKIP

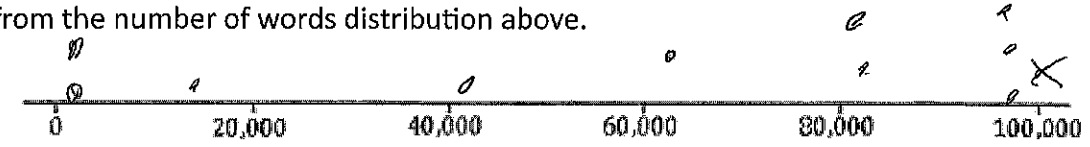
4. **Sampling Distributions**

What is the longest book you have ever read? The Hobbit has 95,022 words, and The Cat in the Hat has 830 words. Popular books vary in the number of words they have—not just the number of different words but the total number of words. The table below shows the total number of words in some of those books. The histogram displays the total number of words in 150 best-selling children's books with fewer than 100,000 words.

Book	Words		Book	Words		Book	Words
Black Beauty	59,635		*Charlie and the Chocolate Factory*	30,644		*The Hobbit*	95,022
The Catcher in the Rye	73,404		*Old Yeller*	35,968		*Judy Moody Was in a Mood*	11,049
The Adventures of Tom Sawyer	69,066		*The Cat in the Hat*	830		*Treasure Island*	66,950
The Secret Garden	80,398		*Green Eggs and Ham*	702		*Magic Tree House Lions at Lunchtime*	5,313
The Mouse and the Motorcycle	22,416		*Little Bear*	1,630		*Harry Potter and the Sorcerer's Stone*	77,325
The Wind in the Willows	58,424		*The Red Badge of Courage*	47,180		*Harry Potter and the Chamber of Secrets*	84,799
My Father's Dragon	7,682		*Anne Frank: The Diary of a Young Girl*	82,762		*Junie B. Jones and the Stupid Smelly Bus*	6,570
Frog and Toad All Year	1,727		*Midnight for Charlie Bone*	65,006		*White Mountains*	44,763
Book of Three	46,926		*The Lion, The Witch and the Wardrobe*	36,363		*Double Fudge*	38,860



a. Put dots on the number line below that you think would represent a random sample of size 10 from the number of words distribution above.



b. The data for the number of words in the 150 best-selling children's books are listed below. Select a random sample of the number of words for 10 books.

Books 1–10	59,635	82,762	92,410	75,340	8,234	59,705	92,409	75,338	8,230	82,768
Books 11–20	73,404	65,006	88,250	2,100	81,450	72,404	88,252	2,099	81,451	65,011
Books 21–30	69,066	36,363	75,000	3,000	80,798	69,165	75,012	3,010	80,790	36,361
Books 31–40	80,398	95,022	71,200	3,250	81,450	80,402	71,198	3,252	81,455	95,032
Books 41–50	22,416	11,049	81,400	3,100	83,475	22,476	81,388	3,101	83,472	11,047
Books 51–60	58,424	66,950	92,400	2,750	9,000	58,481	92,405	2,748	9,002	66,954
Books 61–70	7,682	5,313	83,000	87,000	89,170	7,675	83,021	87,008	89,167	5,311
Books 71–80	1,727	77,325	89,010	862	88,365	1,702	89,015	860	88,368	77,328
Books 81–90	46,926	84,799	88,048	927	89,790	46,986	88,042	926	89,766	84,796
Books 91–100	30,644	6,570	90,000	8,410	91,010	30,692	90,009	8,408	91,015	6,574
Books 101–110	35,968	44,763	89,210	510	9,247	35,940	89,213	512	9,249	44,766
Books 111–120	830	8,700	92,040	7,891	83,150	838	92,037	7,889	83,149	8,705
Books 121–130	702	92,410	94,505	38,860	81,110	712	94,503	87,797	81,111	92,412
Books 131–140	1,630	88,250	97,000	7,549	8,245	1,632	97,002	7,547	8,243	88,254
Books 141–150	47,180	75,000	89,241	81,234	8,735	47,192	89,239	81,238	8,739	75,010

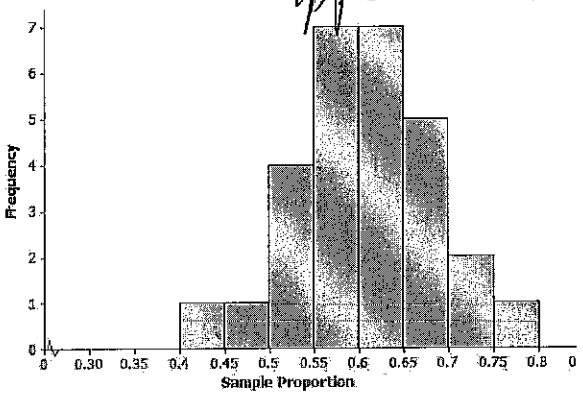
c. If you were to compare your samples to your classmates' samples, do you think you would find the same average number of words? Why or why not?

Nope. Sample to Sample variability

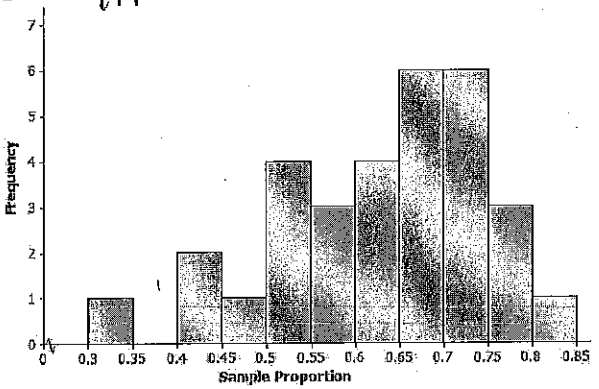
Problem Set:

1. A group of eleventh graders wanted to estimate the proportion of all students at their high school who suffer from allergies. Each student in one group of eleventh graders took a random sample of 20 students, while each student in another group of eleventh graders each took a random sample of 40 students. Below are the two sampling distributions (shown as histograms) of the sample proportions of high school students who said that they suffer from allergies. Which histogram is based on random samples of size 40 ? Explain. Histogram A has bigger samples b/c it appears more like a normal curve.

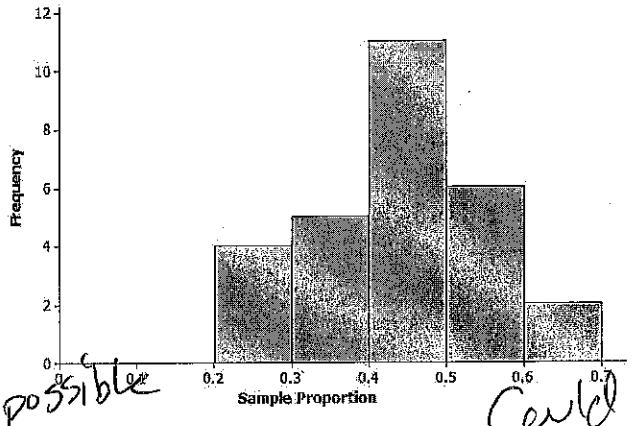
Histogram A



Histogram B



2. The nurse in your school district would like to study the proportion of all high school students in the district who usually get at least eight hours of sleep on school nights. Suppose each student in your class takes a random sample of 20 high school students in the district and each calculates their sample proportion of students who said that they usually get at least eight hours of sleep on school nights. Below is a histogram of the sampling distribution.



.75 is unlikely, but possible. Could be .4

a. Do you think that the proportion of all high school students who usually get at least eight hours of sleep on school nights could have been 0.4 ? Do you think it could have been 0.55 ? Could it have been 0.75 ? Justify your answers based on the histogram. .55 is possible

b. Suppose students had taken random samples of size 60 . How would the distribution of sample proportions based on samples of size 60 differ from those of size 20

It would be more tightly clustered and "normal".