Lesson 20: Memory test

Goals

- Describe (orally) connections between sampling and probability.
- Generate a random sample, and use it to make inferences (in writing) about the population.
- Justify (orally and in writing) whether a given method produces a random sample.

Learning Targets

- I can compare two groups by taking a random sample, calculating important measures, and determining whether the populations are meaningfully different.

Lesson Narrative

This lesson is optional. It gives students a chance to use the material they have learned in the unit with the final goal of comparing two populations, but may be shortened or skipped due to time constraints.

In this lesson, students apply what they have learned about probability, sampling, and comparing populations to analyse two data sets. Half of the class works with one data set while the other half of the class works with another. Students choose their own tools for selecting a sample at random and calculate the mean, range, and a proportion to summarise their sample. Then students compare their results with a partner that had the other data set to construct an argument for whether there is a meaningful difference between the sets.

Building On

- Relating the choice of measures of centre and variability to the shape of the data distribution and the context in which the data were gathered.

Addressing

- Analyse proportional relationships and use them to solve real-world and mathematical problems.
- Use random sampling to draw inferences about a population.
- Understand that statistics can be used to gain information about a population by examining a sample of the population; generalisations about a population from a sample are valid only if the sample is representative of that population. Understand that random sampling tends to produce representative samples and support valid inferences.
- Use data from a random sample to draw inferences about a population with an unknown characteristic of interest. Generate multiple samples (or simulated samples) of the same size to gauge the variation in estimates or predictions. For example,

estimate the mean word length in a book by randomly sampling words from the book; predict the winner of a school election based on randomly sampled survey data. Gauge how far off the estimate or prediction might be.

- Use measures of centre and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations. For example, decide whether the words in a chapter of a year 8 science book are generally longer than the words in a chapter of a year 5 science book.
- Develop a uniform probability model by assigning equal probability to all outcomes, and use the model to determine probabilities of events. For example, if a student is selected at random from a class, find the probability that Jane will be selected and the probability that a girl will be selected.

Instructional Routines

- Compare and Connect
- Discussion Supports

Required Materials

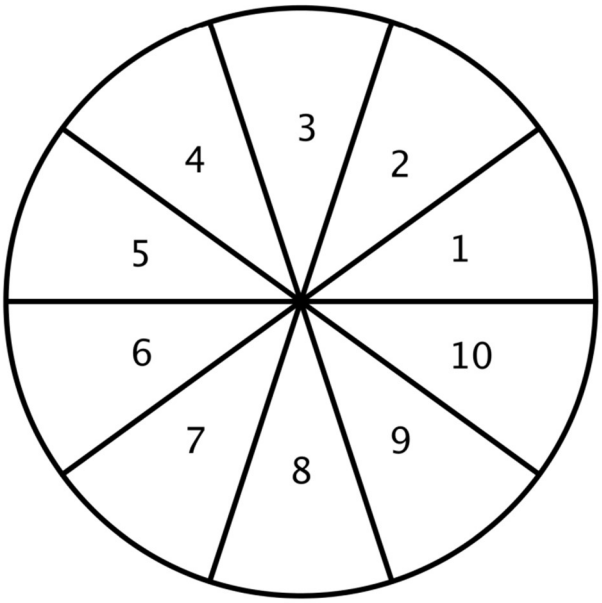
Copies of blackline master

Data Set 1

	column 1	column 2	column 3	column 4	column 5	column 6	column 7	column 8	column 9	column 10
row 1	35	47	42*	33	42	57	40	39	33	37*
row 2	58	33*	55	42	66	10	47	44	45*	54
row 3	45	59	44	51	57	52*	42	42	43	22*
row 4	63	35	37	81	37*	11*	27	43	35*	43
row 5	95*	60*	41*	70*	64	47	34	63	40	41
row 6	47	39	57	71	46	53	63	57*	62	36
row 7	53	44	45	30	73*	39	54	48	30	39
row 8	29	36	27*	69	45	55	40	58*	62	42
row 9	34	49	61	18	40	40	52	54	45	26
row 10	65	35*	63	49	76	62	46*	42*	59	32

Data Set 2

	column 1	column 2	column 3	column 4	column 5	column 6	column 7	column 8	column 9	column 10
row 1	26	39	10	42*	52	34	39	71	30	54
row 2	69	49	45	42	40	46*	47	46	30	18
row 3	22*	47	34	54	48	57*	42	58*	65	27
row 4	40	53	51	63	62	55	37*	41	59	40
row 5	33*	45	36	11*	35*	42	39	63	36	57
row 6	49	63	45	44	70*	42	29	73*	35*	59
row 7	40	32	40	47	57	42*	27*	35	64	57
row 8	54	42	45*	33	60*	63	53	43	35	33
row 9	44	43	43	62	41*	61	44	95*	52*	39
row 10	58	81	37*	45	55	37	62	76	66	47



Dice
cubes with sides numbered from 1 to 6

Paper bags
Paper clips

Required Preparation

Print the Collecting a Sample blackline master. Provide one data set and one spinner for each student. If the spinners are used to select a random sample during the Sample Probabilities activity, provide a paper clip and sharpened pencil to use with the spinners. If

possible, provide access to other tools for selecting a random sample from a 10-by-10 grid such as a 10-sided polyhedra or coloured cubes and paper bags.

Student Learning Goals

Let's put it all together.

20.1 Collecting a Sample

Optional: 5 minutes

In this activity, students review methods of obtaining samples that are fair and random.

Launch

Arrange students in groups of 2. Each group gets both sets of data from the blackline master, one data set for each partner. Students will not need the spinners from the blackline master for this activity, but the spinners are included for use later in the lesson. Partners may work together to answer the questions, but should not share their data set with one another until told to do so in a later activity.

Student Task Statement

You teacher will give you a paper that lists a data set with 100 numbers in it. Explain whether each method of obtaining a sample of size 20 would produce a random sample.

Option 1: A spinner has 10 equal sections on it. Spin once to get the row number and again to get the column number for each member of your sample. Repeat this 20 times.

Option 2: Since the data looks random already, use the first two rows.

Option 3: Cut up the data and put them into a bag. Shake the bag to mix up the papers, and take out 20 values.

Option 4: Close your eyes and point to one of the numbers to use as your first value in your sample. Then, keep moving one square from where your finger is to get a path of 20 values for your sample.

Student Response

1. This would produce a random sample since each row and column has an equal chance of being selected.

2. This would not produce a random sample since all of the values do not have an equal chance of being selected.

3. This would produce a random sample since all the papers are the same size and each value has an equal chance of being selected.

4. This would not produce a random sample since the path limits the values you can get in your sample. For example, the four corners could not all be in the same sample of 20.

Activity Synthesis

The purpose of the discussion is to help students solidify their understanding of methods for selecting random samples.

Consider these questions for discussion:

- "Can you think of other methods for selecting a random sample that are not listed here?" (Roll a polyhedron with 10 equal faces showing the numbers 1 through 10 to get the row and again to get the column.)
- "What do you need to look for when determining if a sample is random?" (Are all values equally likely to be included in the random sample?)

20.2 Sample Probabilities

Optional: 10 minutes

In this activity, students begin by practising their understanding of proportions and probabilities by examining the data set they have available. In the fourth problem, students obtain a sample from the population using tools they choose and examine the sample they selected to compare it to the expected proportions and probabilities calculated in the first 3 problems.

The problems are intended for students to use their own data set to answer. Although they are kept in pairs for the entire lesson, this activity should be done individually.

Instructional Routines

- Compare and Connect

Launch

Keep students in the same groups of 2. Give students 5–7 minutes of quiet work time followed by a whole-class discussion.

If possible, allow students to use their chosen method of random sampling to obtain a sample of 10 for this activity. Have items such as paper clips, scissors, 10-sided polyhedra, and other materials available for student use. The blackline master for the first activity in this lesson contains accurate spinners that could be used to select a random sample.

Action and Expression: Internalise Executive Functions. Chunk this task into more manageable parts. After students have solved the first 2-3 problems, check-in with either select groups of students or the whole class. Invite students to share the strategies they have used so far as well as any questions they have before continuing.
Supports accessibility for: Organisation; Attention

Student Task Statement

Continue working with the data set your teacher gave you in the previous activity. The data marked with a star all came from students at Springfield Middle School.

1. When you select the first value for your random sample, what is the probability that it will be a value that came from a student at Springfield Middle School?
2. What proportion of your entire sample would you expect to be from Springfield Middle School?
3. If you take a random sample of size 10, how many scores would you expect to be from Springfield Middle School?
4. Select a random sample of size 10.
5. Did your random sample have the expected number of scores from Springfield Middle School?

Student Response

1. $\frac{20}{100} = \frac{1}{5} = 0.2$
2. The proportion I would expect is about 0.2.
3. From a random sample of 10, I expect there to be around 2 from Springfield Middle School.
4. Answers vary. Sample response: 53, 46, 63, 42, 41, 33, 71, 39, 40, 18.
5. Answers vary. Sample response: No, I only had 1 value come from Springfield Middle School.

Activity Synthesis

The purpose of this discussion is to connect the ideas of probability and random sampling from the unit.

Consider these questions for discussion:

- "How is selecting a sample at random connected to probability?" (A random sample should give each value an equal chance of being chosen. Therefore, each value has a $\frac{1}{100}$ probability of being chosen.)
- "How could we simulate the probability of getting at least 2 values in the sample of 10 from Springfield Middle School?" (Since 20% of the values come from Springfield Middle School, we could put 10 blocks in a bag with 2 coloured red to represent Springfield Middle School. Draw a block from the bag, and if it is red, it represents a score from Springfield Middle School; replace the block and repeat. Get a sample of 10 and see if the sample has at least 2 red blocks. Repeat this process many times and use

the fraction of times there are at least 2 red blocks as an estimate for the probability that a random sample will have at least 2 scores from Springfield Middle School.)

Representing, Speaking: Compare and Connect. Invite students to create a visual display of their random sample of size 10 and response to the question: "Did your random sample have the expected number of scores from Springfield Middle School?" Invite students to investigate each other's work and compare their responses. Listen for the language students use to describe a random sample and assign a probability of each value being chosen. This will help students connect the ideas of probability and random sampling through discussion.
Design Principle(s): Optimise output (for representation); Cultivate conversation

20.3 Estimating a Measure of Centre for the Population

Optional: 10 minutes

In this activity, students practise estimating a measure of centre for the population using the data from a sample. The variability is also calculated to be used in the following activity to determine if there is a meaningful difference between the measure of centre for the population they used to select their sample and the measure of centre for another population.

Instructional Routines

- Discussion Supports

Launch

Keep students in groups of 2. Students should work with their partner for the first question, then individually for the last 2 problems. Follow up with a whole-class discussion.

Student Task Statement

1. Decide which measure of centre makes the most sense to use based on the distribution of your sample. Discuss your thinking with your partner. If you disagree, work to reach an agreement.
2. Estimate this measure of centre for your population based on your sample.
3. Calculate the measure of variability for your sample that goes with the measure of centre that you found.

Student Response

Answers vary. Sample response:

1. We chose to use mean since there were no values far from the centre of the data.
2. Mean: 44.6
3. Range: 35

Activity Synthesis

The purpose of the discussion is for students make clear their reasoning for choosing a particular measure of centre and reiterate the importance of variability when comparing groups from samples.

Consider these questions for discussion:

- "Which measure of centre did your group choose? Explain your reasoning." (A median should be used if there are a few values far from the centre that overly influence the mean in that direction. If the data is not approximately symmetrical, a median should be used as well. In other cases, the mean is probably a better choice for the measure of centre.)
- "Why is it important to measure variability in the data when estimating a measure of centre for the population using the data from a sample. If there is small variation, then the samples may have come from a population that also has a small variation, so differences among groups may be more clearly defined.)

Speaking: Discussion Supports. Use this routine to support whole-class discussion. For each response that is shared, ask students to restate and/or revoice what they heard using precise mathematical language. Consider providing students time to restate what they hear to a partner, before selecting one or two students to share with the class. Ask the original speaker whether their peer was accurately able to restate their thinking. Call students' attention to any words or phrases that helped clarify the original statement. This will provide more students with an opportunity to produce language as they interpret the reasoning of others.
Design Principle(s): Support sense-making

20.4 Comparing Populations

Optional: 5 minutes

In this activity, students use the values computed in the previous activity to determine if there is a meaningful difference between two populations. Following the comparison of the groups, students are told that the populations from which they selected a sample were identical, although shuffled.

Launch

Keep students in the same groups of 2 established at the beginning of this lesson. Allow students 3 minutes of partner work time followed by a whole-class discussion.

Student Task Statement

Using only the values you computed in the previous two activities, compare your sample to your partner's.

Is it reasonable to conclude that the measures of centre for each of your populations are meaningfully different? Explain or show your reasoning.

Student Response

Answers vary. Sample response: They are not meaningfully different. The difference in means is 4.9, but the larger of the two ranges is 41, so that is small compared with the range.

Activity Synthesis

Ask each group to share whether they found a meaningful difference.

Tell students, "With your partner, compare the starred data for the two groups. What do you notice?"

Tell students that the two populations are actually identical, but rearranged. Ask, "Did any groups get different means for your samples? Explain why that might have happened, even though the populations are the same." (Two random samples from the population will usually not contain the same values, so different means are probably expected.)

Lesson Synthesis

Key learning points:

- Probability and random samples are connected through the equal likelihood of individuals from the population being selected.
- It is important to select samples through a random process in order to compare two populations.

Consider asking these discussion questions:

- "Why was it important to select a random sample from the population data you had?" (A random sample gives us the best chance of being representative of the population.)
- "A scientist has access to data for the high temperature in London for each day of every year since 1945. Describe a process the scientist could use to compare the temperatures from 1963 and 2003." (Select a random sample of temperatures from each year. Determine the correct measure of centre and variation. Use our general rule to compare the measure of centre for each year based on the sample characteristics.)