

DSC 40A

Theoretical Foundations of Data Science I

In This Video

- What is the probability of a random sample having a certain property?

Sampling

Sampling with replacement:

1. Draw one element *uniformly at random* from list.
2. Return the element to the list.
3. Repeat

Sampling without replacement:

What does *uniformly at random* mean?

Sampling

Sampling with or without replacement:

- All samples are equally likely.
- Uniform distribution!

$$P(\text{sample having a certain property}) =$$

Sampling

Sampling with or without replacement:

- All samples are equally likely.
- Uniform distribution!

$$P(\text{sample having a certain property}) = \frac{\# \text{ samples having property}}{\# \text{ possible samples}}$$

Practice Problems

Example 5. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **with replacement**. What is the chance that a particular student is among the 5 selected students?

Practice Problems

Part 1. Denominator. If you draw a sample of size 5 at random with replacement from a population of size 20, how many different sequences of individuals could you draw?

Practice Problems

Part 2. Numerator. If you draw a sample of size 5 at random with replacement from a population of size 20, how many different sequences of individuals include a particular person?

Practice Problems

Using the complement. If you draw a sample of size 5 at random with replacement from a population of size 20, how many different sequences of individuals **do not** include a particular person?

Practice Problems

Example 5. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **with replacement**. What is the chance that a particular student is among the 5 selected students?

Practice Problems

Example 6. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **without replacement**. What is the chance that a particular student is among the 5 selected students?

Practice Problems

Example 6. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **without replacement**. What is the chance that a particular student is among the 5 selected students?

Which probability will be higher?

- A. Probability of including a particular student when sampling with replacement.
- B. Probability of including a particular student when sampling without replacement.
- C. Both probabilities are the same.

Practice Problems

Part 1. Denominator. If you draw a sample of size 5 at random without replacement from a population of size 20, how many different sequences of individuals could you draw?

Practice Problems

Part 2. Numerator. If you draw a sample of size 5 at random without replacement from a population of size 20, how many different sequences of individuals include a particular person?

Practice Problems

Using the complement. If you draw a sample of size 5 at random without replacement from a population of size 20, how many different sequences of individuals **do not** include a particular person?

Practice Problems

Example 6. There are 20 students in a class. A computer program selects a random sample of students by drawing 5 students at random **without replacement**. What is the chance that a particular student is among the 5 selected students?

Summary

- When we sample uniformly, whether with or without replacement, each possible sample is equally likely.
- Probability questions become counting questions:

$$P(\text{sample having a certain property}) = \frac{\# \text{ samples having property}}{\# \text{ possible samples}}$$

- **Next time:** combinatorics, or counting principles