

Statistics 104 - Laboratory 5

Drawing samples from populations requires some method for random selection. In the two group activities today you will use different methods to randomly select a sample from a population. In the process you will also learn something about the variability introduced by random sampling.

1. **Activity 1:** Your group will need a sheet of 100 random rectangles for this activity. DO NOT study the sheet of random rectangles. The random rectangles represent a generic population of values. The areas of the rectangles could represent the number of people in households or the number of hours students spend on STAT 104 each week. What we wish to investigate in this activity is what can be learned about the area of the rectangles using sampling techniques.
 - a) Each individual in the group should look at the sheet of random rectangles for a few seconds and guess the mean area of the rectangles on the sheet. The unit of measure is the background square. Thus rectangle 7 has area $4 \times 3 = 12$. Compare your guess with those of the other members in your group. Discuss why you got different guesses. Record the individual guesses on your group data sheet.
 - b) Each individual should now select five rectangles that, in her/his judgment, are representative of the rectangles on the sheet. Write down the area for each of the five. Compute the mean of the five areas. How does this compare with your guess? How does it compare with the means of the judgment samples for the others in your group? Again discuss why you got different means. Record the judgment samples and means on your group data sheet.
 - c) We want to take a random sample from the 100 rectangles. We could cut out the rectangles, laminate them, put them in a bag, mix thoroughly and draw from the bag. Do you think this would give you a random sample?
 - d) You should have a 10 sided die available to use. Roll the die ten times and record the results on your answer sheet. Explain how you can use the rolls of the die to select a simple random sample of 5 rectangles.
 - e) Using your sampling plan, each individual should take a simple random sample of 5 rectangles. Record the areas on your group answer sheet. What is the mean area of each sample? Are the sample mean areas the same for all your samples? Why?
 - f) How do the random sample means compare to the judgment sample means and guesses?
 - g) Which do you think will be a better guess of the true mean area of the 100 rectangles, the random sample mean of 5 rectangles or the random sample mean of 20 rectangles? Explain briefly.

2. **Activity 2:** For this activity you could use the 10 sided die as you did in Activity 1. However, it will require many rolls of the die. Instead, you should use the table of random numbers (Table E, Appendix A of your text). This table was made in a way similar to the way you created the random numbers you used in part d) of Activity 1.

Pine forests are periodically infested with beetles and the infestation often seems to jump from tree to tree in a random manner. How can a forester estimate the total number of trees infested with beetles in a large forest? By taking a sample of plots in the forest and counting the number of infested trees in the sample, a forester will be able to use the sample data to estimate the total. To simulate this activity, imagine that Figure 1 represents a forest and the black dots represent infested trees in that forest. The forest has been divided into 100 plots of equal area.

- a) Using the table of random numbers devise a plan to produce a simple random sample of plots from the forest. Be sure to identify the population, variable of interest, and sample.
- b) Use your sampling plan to randomly select 10 plots from the forest. For each plot count the number of infested trees (dots) that are **entirely within** the plot. Do not count trees (dots) that overlap the edges of a plot. For example, the (Row 0, Column 0) plot has 2 infested trees.
- c) Calculate the mean of the 10 counts. Using this mean, estimate the total number of infested trees in the forest.
- d) Again use the random number table to select 10 plots. This time for each plot count the number of infested trees (dots) that have **any part within** the plot. For example, the (Row 0, Column 0) plot has 3 infested trees.
- e) Calculate the mean of the 10 counts. Using this mean, estimate the total number of infested trees in the forest.
- f) What would be the best estimate of the total number of infested trees in the forest? Briefly explain your choice.
- g) An alternative to sampling the plots is to randomly pick a row (or column) of plots in the forest. This row (or column) represents a transect. It would be much easier for the forester to walk through the forest along a randomly selected transect than it would be to find 10 randomly selected plots in a real forest.
 - i. Use your random number table to select a row at random and count the number of infested trees entirely within the row (transect).
 - ii. Use your random number table to select a column at random and count the number of infested trees with some part within the column (transect).
 - iii. Come up with an estimate of the total number of infested trees based on the results from your transects. Show your work.
- h) How well do your two estimates of the total number of infested trees (f and g) compare?

Statistics 104 - Laboratory 5

Group Answer Sheet

Names of Group Members: _____, _____
 _____, _____

1. Activity 1: Sampling Rectangles

Name	Guess	Areas in Judgment Sample	Judgment Sample Mean	Areas in Random Sample	Random Sample Mean

c) We could cut out the rectangles, laminate them, put them in a bag, mix thoroughly and draw from the bag. Do you think this would give you a random sample? Explain briefly.

d) Roll the die ten times and record the results on your answer sheet. Explain how you can use the rolls of the die to select a simple random sample of 5 rectangles.

e) Are the sample mean areas the same for all the random samples? Why?

f) How do the random sample means compare to the judgment sample means and guesses?

g) Which do you think will be a better guess of the true mean area of the 100 rectangles, the random sample mean of 5 rectangles or the random sample mean of 20 rectangles? Explain briefly.

2. **Activity 2:** Sampling trees in a forest

a) Sampling Plan (indicate population, variable and sample and how you will use the table of random numbers to select your sample.)

b) Identify plots selected and give counts for 10 selected plots, whole tree in plot.

Plot:										
Count:										

c) Mean number infested and estimate of the total number of infested trees in the forest.

- d) Identify plots selected and give counts for 10 selected plots, any part of tree in plot.

Plot:										
Count:										

- e) Mean number infested and estimate of the total number of infested trees in the forest.

- f) Best estimate of the total number of infested trees and explanation.

- g) Transect sampling.

i. Row (transect) selected: _____

Number of infested trees, whole tree in transect: _____

ii. Column (transect) selected: _____

Number of infested trees, any part of tree in transect. _____

iii. Estimate of total number of infested trees in the forest showing work.

- b. How do the two estimates (f and g) compare?