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 by looking at samples from the population.

   a) **Each individual in the group** should look at the sheet of random rectangles for a few seconds and guess the mean area of all 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 for each member in your group on the group answer sheet.

   b) **Each individual** should now select five rectangles that, in her/his judgment, are representative of the rectangles on the sheet. Write down rectangle numbers and the areas on the lab answer sheet. Compute the mean of the five areas and record it on the group answer sheet. 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.

   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 five, without replacement, 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 group answer sheet. Explain how you can use these rolls of the die to select a simple random sample of 5 rectangles.

   e) Using your sampling plan, **each individual** should take their own simple random sample of 5 rectangles. Record the rectangle numbers and areas on your group answer sheet. What is the mean area of each sample? Are 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. Or you could 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.

Different types of trees in the U.S. are periodically threatened by invasive species. For example, ash trees in Iowa may be threatened in the near future by the Emerald Ash Borer. Many communities in Iowa are preparing for this threat by finding out how many trees in their communities are ash trees. How can a forester estimate the total number of ash trees there are in a community park? To simulate this activity, imagine that Figure 1 represents a community park. The black dots represent ash trees and the open circles represent other trees in the park.

a) Why would it be difficult to take a random sample of trees from all of the trees in the park? Hint: What would you first have to do before you could use the die or a random number table to take a random sample of trees?
b) Instead we will take an area sample. In an area sample, the sampling unit is a plot of land in the park. Note that there are 100 rectangular plots. Devise a plan to produce a simple random sample of plots from the park.
c) Use your sampling plan to randomly select 10 plots from the park. For each plot count the number of ash trees (dots) and other trees (open circles) that are **entirely within** the plot. Do not count trees (dots or open circles) that overlap the edges of the plot. For example, the (Row 0, Column 0) plot has 2 ash trees and 3 other trees.
d) Calculate the mean number of ash trees for the 10 plots. Using this mean, estimate the total number of ash trees in the park.
e) Use your data from c) to estimate the proportion of all trees in the park that are ash trees.
f) Again use your sampling plan to select a different random sample of 10 plots. This time for each plot count the number of ash trees (dots) and other trees (open circles) that have **any part within** the plot. For example, the (Row 0, Column 0) plot has 3 ash trees and 3 other trees.
g) Calculate the mean number of ash trees for the 10 plots. Using this mean, estimate the total number of ash trees in the park.
h) What would be the best estimate of the total number of ash trees in the park? Explain why and give the estimate.
i) An alternative to randomly sampling the plots is to randomly pick a row (or column) of plots in the park. This row (or column) represents a transect. It would be much easier for the forester to walk through the park along a randomly selected transect than it would be to find 10 randomly selected plots in a real park.
   i. Use your sampling scheme to select a row at random and count the number of ash trees and other trees entirely within the row (transect).
   ii. Come up with an estimate of the total number of ash trees and the proportion of ash trees in the park based on the results from your transect. How do these estimates compare to your estimates in d) and e)?
Statistics 104 - Laboratory 5
Group Answer Sheet

Names of Group Members: ____________________, ____________________
____________________, ____________________

1. **Activity 1: Sampling Rectangles**

<table>
<thead>
<tr>
<th>Name</th>
<th>Guess</th>
<th>Judgment Sample</th>
<th>Judgment Sample Mean Area</th>
<th>Random Sample</th>
<th>Random Sample Mean Area</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>#’s:</td>
<td>#’s:</td>
<td>#’s:</td>
<td>#’s:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Areas:</td>
<td>Areas:</td>
<td>Areas:</td>
<td>Areas:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>#’s:</td>
<td>#’s:</td>
<td>#’s:</td>
<td>#’s:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Areas:</td>
<td>Areas:</td>
<td>Areas:</td>
<td>Areas:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>#’s:</td>
<td>#’s:</td>
<td>#’s:</td>
<td>#’s:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Areas:</td>
<td>Areas:</td>
<td>Areas:</td>
<td>Areas:</td>
</tr>
</tbody>
</table>

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 the group answer sheet. Explain how you can use the rolls of the die to select a simple random sample of 5 rectangles.

___ ___ ___ ___ ___ ___ ___ ___ ___ ___
e) Are 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 park

   a) Why would it be difficult to take a random sample of trees from all of the trees in the park?

   b) Sampling Plan:

   c) Identify plots selected and give the number of ash trees and the number of other trees in each of the 10 selected plots, whole tree in plot.

<table>
<thead>
<tr>
<th>Plot:</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ash:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
d) Calculate the sample mean number of ash trees and estimate of the total number of ash trees in the park.

e) Use your data from c) to estimate the proportion of all trees in the park that are ash trees.

f) Identify plots selected and give the number of ash trees and the number of other trees in each of the 10 selected plots, any part of tree in plot.

<table>
<thead>
<tr>
<th>Plot:</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ash:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

g) Calculate the sample mean number of ash trees and estimate the total number of ash trees in the park.

h) What would be the best estimate of the total number of ash trees in the park? Explain why and give the estimate.

i) Transect sampling.
   i. Use your sampling scheme to select a row at random and count the number of ash trees and other trees entirely within the row (transect).

<table>
<thead>
<tr>
<th>Row</th>
<th>Ash</th>
<th>Other</th>
</tr>
</thead>
</table>

   ii. Come up with an estimate of the total number of ash trees and the proportion of ash trees in the park based on the results from your transect. How do these estimates compare to your estimates in d) and e)?