Sampling Methods

Name _____

Task A–Stratified Samples

Suppose you were again using a sample to estimate the population mean of the areas of the rectangles shown on the activity sheet "Random Rectangles," and this time you wanted to make sure that your sample contained rectangles with both small and large widths.

Tables 1 and 2 show the set of 100 rectangles divided into two groups (strata) according to the widths of the rectangles. Note that in this division of the rectangles, *width* indicates a rectangle's horizontal dimension. One stratum contains 59 rectangles with widths less than 3 (table 1), and the other stratum contains 41 rectangles with widths greater than or equal to 3 (table 2).

Table 1

Stratum of Rectangles with Widths Less than 3, Listed by the Numbers Corresponding to Them on "Random Rectangles"

1	2	3	4	5	6	8	9	10	11	12	14	16	18	19
20	21	22	24	27	28	30	32	34	35	36	40	43	44	45
46	48	49	50	51	53	56	62	64	67	68	69	71	73	74
75	77	78	79	80	82	83	84	87	88	89	91	95	100	

Table 2.

Stratum of Rectangles with Widths Greater than or Equal to 3, Listed by the Numbers Corresponding to Them on "Random Rectangles"

7	13	15	17	23	25	26	29	31	33	37	38	39	41	42
47	52	54	55	57	58	59	60	61	63	65	66	70	72	76
81	85	86	90	92	93	94	96	97	98	99				

Stratified random sampling is a process that involves randomly selecting samples from groups like these, each of which is considered a stratum of the population.

Use the rectangles in tables 1 and 2 to explore the process and results of stratified random sampling:

1. Using some type of random number generator, randomly select 5 rectangles from each table (stratum), and then compute the mean area of each stratum of 5 rectangles. The sample mean for the combined strata is found by using the population proportion, as follows:

$$\frac{59}{100} \cdot (\text{mean of stratum from table 1}) + \frac{41}{100} \cdot (\text{mean of stratum from table 2})$$

Sampling Methods (continued)

Name _____

Repeat this process three times.

2. Compile the class data and show all sample means in a simulated sampling distribution. Describe the shape, center, and spread of the simulated sampling distribution.

3. Using the activity sheet "Data Record," sketch the simulated sampling distribution and enter a measure of center and a measure of spread (median and interquartile range, or mean and standard deviation) for the simulated sampling distribution of sample mean areas from stratified sampling. Write a brief description of this simulated sampling distribution.

Sampling Methods (continued)

Name _____

Task B—Cluster Samples

Sometimes a population of interest has so many members or the members are so dispersed that the cost of taking a simple random sample is too high. In such instances, researchers may use ***cluster sampling,*** a random selection process by which clusters of individuals are identified in the population, and the individuals in the clusters are then studied.

Table 3 shows the 100 rectangles from the activity sheet "Random Rectangles" divided into twenty clusters. Each cluster is "like" every other cluster in that it contains 5 rectangles that are relatively close together on the sheet. Each also contains as much variability as is possible in its particular "neighborhood." Each cluster is named by a Roman numeral.

Table 3.

Clusters of Rectangles Listed by the Numbers Corresponding to Them on "Random Rectangles"

I	II	III	IV	V	VI	VII	VIII	IX	X
1	3	8	7	19	18	25	40	32	39
2	4	13	11	26	20	34	41	33	47
9	5	14	12	27	21	35	42	38	48
16	6	15	22	28	30	36	43	45	49
17	10	24	23	29	31	37	44	46	50
XI	XII	XIII	XIV	XV	XVI	XVII	XVIII	XIX	XX
51	55	58	68	64	66	79	91	88	83
52	56	59	69	72	75	80	92	89	84
53	57	60	70	73	76	81	93	96	90
54	62	65	71	74	77	86	94	97	99
61	63	67	78	82	83	87	95	98	100

Use the clusters of rectangles in table 3 to explore the process and results of cluster sampling:

1. Using a random number generator, select two of the twenty clusters. Find the mean area of the your sample of 10 rectangles from the two clusters. Repeat the process three times.

Sampling Methods (continued)

Name _____

2. Compile the class data and show all sample means in a simulated sampling distribution.

3. Complete the activity sheet "Data Record," entering for this simulated sampling distribution a measure of center and a measure of spread (median and interquartile range, or mean and standard deviation). Write a description of the simulated sampling distribution, including center, spread, and shape.

4. Study the information that you have accumulated on the sheet "Data Record" and compare the sampling methods that you have studied.

   a. What can you say about the effect of sample size?

   b. What observations can you make about the three methods that you have explored: simple random sampling, stratified random sampling, and cluster sampling?