## STATISTICS
SECTION II
Part A
Questions 1-5
Spend about 65 minutes on this part of the exam.
Percent of Section II grade—75

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy of your results and explanation.

1. The goal of a nutritional study was to compare the caloric intake of adolescents living in rural areas of the United States with the caloric intake of adolescents living in urban areas of the United States. A random sample of ninth-grade students from one high school in a rural area was selected. Another random sample of ninth graders from one high school in an urban area was also selected. Each student in each sample kept records of all the food he or she consumed in one day.

   The back-to-back stemplot below displays the number of calories of food consumed per kilogram of body weight for each student on that day.

Urban		Rural
9 9 9 9 8 8 7 6	2	
4 4 3 1 0	3	2 3 3 4
9 7 6 6 5	3	5 6 6 6 7
2 0	4	0 2 2 2 4
	4	5 6 8 8 9
	5	1
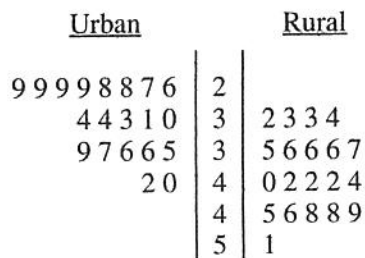
   Stem: tens
   Leaf: ones

   (a) Write a few sentences comparing the distribution of the daily caloric intake of ninth-grade students in the rural high school with the distribution of the daily caloric intake of ninth-grade students in the urban high school.

   (b) Is it reasonable to generalize the findings of this study to all rural and urban ninth-grade students in the United States? Explain.

   (c) Researchers who want to conduct a similar study are debating which of the following two plans to use.

   Plan I: Have each student in the study record all the food he or she consumed in one day. Then researchers would compute the number of calories of food consumed per kilogram of body weight for each student for that day.

   Plan II: Have each student in the study record all the food he or she consumed over the same 7-day period. Then researchers would compute the average daily number of calories of food consumed per kilogram of body weight for each student during that 7-day period.

   Assuming that the students keep accurate records, which plan, I or II, would better meet the goal of the study? Justify your answer.

GO ON TO THE NEXT PAGE.

2. A local school board plans to conduct a survey of parents' opinions about year-round schooling in elementary schools. The school board obtains a list of all families in the district with at least one child in an elementary school and sends the survey to a random sample of 500 of the families. The survey question is provided below.

A proposal has been submitted that would require students in elementary schools to attend school on a year-round basis. Do you support this proposal? (Yes or No)

The school board received responses from 98 of the families, with 76 of the responses indicating support for year-round schools. Based on this outcome, the local school board concludes that most of the families with at least one child in elementary school prefer year-round schooling.

(a) What is a possible consequence of nonresponse bias for interpreting the results of this survey?

(b) Someone advised the local school board to take an additional random sample of 500 families and to use the underline combined results to make their decision. Would this be a suitable solution to the issue raised in part (a) ? Explain.

(c) Suggest a different follow-up step from the one suggested in part (b) that the local school board could take to address the issue raised in part (a).

2. Researchers who are studying a new shampoo formula plan to compare the condition of hair for people who use the new formula with the condition of hair for people who use the current formula. Twelve volunteers are available to participate in this study. Information on these volunteers (numbered 1 through 12) is shown in the table below.

Volunteer	Gender	Age
1	Male	21
2	Female	20
3	Male	47
4	Female	60
5	Female	62
6	Male	61
7	Male	58
8	Female	44
9	Male	44
10	Female	24
11	Male	23
12	Female	46

(a) These researchers want to conduct an experiment involving the two formulas (new and current) of shampoo. They believe that the condition of hair changes with age but not gender. Because researchers want the size of the blocks in an experiment to be equal to the number of treatments, they will use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.

(b) Other researchers believe that hair condition differs with both age and gender. These researchers will also use blocks of size 2 in their experiment. Identify the volunteers (by number) that would be included in each of the six blocks and give the criteria you used to form the blocks.

(c) The researchers in part (b) decide to select three of the six blocks to receive the new formula and to give the other three blocks the current formula. Is this an appropriate way to assign treatments? If so, describe a method for selecting the three blocks to receive the new formula. If not, describe an appropriate method for assigning treatments.

with a mean of 3668 grams and a standard deviation of 511 grams.[8] In this population, the range (maximum – minimum) of birth weights is 3417 grams. We used Fathom software to take 500 SRSs of size $n = 5$ and calculate the range (maximum – minimum) for each sample. The dotplot below shows the results.

Sample range
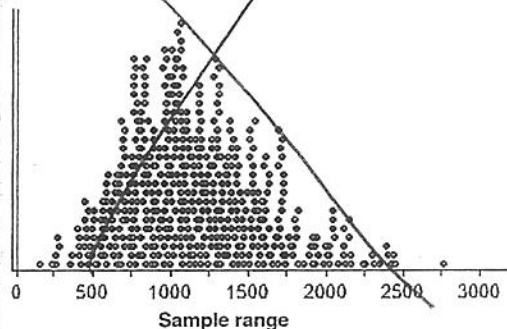
R7.2 Birth weights
(a) Sketch a graph that displays the distribution of birth weights for this population.
(b) Sketch a possible graph of the distribution of birth weights for an SRS of size 5.
(c) In the graph above, there is a dot at approximately 2750. Explain what this value represents.

R7.3 Birth weights
(a) Is the sample range an unbiased estimator of the population range? Give evidence from the graph above to support your answer.
(b) Explain how we could decrease the variability of the sampling distribution of the sample range.

R7.4 Do you jog? The Gallup Poll once asked a random sample of 1540 adults, "Do you happen to jog?" Suppose that the true proportion of all adults who jog is $p = 0.15$.
(a) What is the mean of the sampling distribution of $\hat{p}$? Justify your answer.
(b) Find the standard deviation of the sampling distribution of $\hat{p}$. Check that the 10% condition is met.
(c) Is the sampling distribution of $\hat{p}$ approximately Normal? Justify your answer.
(d) Find the probability that between 13% and 17% of a random sample of 1540 adults are joggers.

R7.5 Bag check Thousands of travelers pass through the airport in Guadalajara, Mexico, each day. Before leaving the airport, each passenger must pass through the Customs inspection area. Customs agents want to be sure that passengers do not bring illegal items into the country. But they do not have time to search every traveler's luggage. Instead, they require each person to press a button. Either a red or a green bulb lights up. If the red light shows, the passenger will be searched by Customs agents. A green light means "go ahead." Customs agents claim that the proportion of all travelers who will be stopped (red light) is 0.30, because the light has probability 0.30 of showing red on any push of the button. To test this claim, a concerned citizen watches a random sample of 100 travelers push the button. Only 20 get a red light.
(a) Assume that the Customs agents' claim is true. Find the probability that the proportion of travelers who get a red light is as small as or smaller than the result in this sample. Show your work.
(b) Based on your results in (a), do you believe the Customs agents' claim? Explain.

R7.6 IQ tests The Wechsler Adult Intelligence Scale (WAIS) is a common "IQ test" for adults. The distribution of WAIS scores for persons over 16 years of age is approximately Normal with mean 100 and standard deviation 15.
(a) What is the probability that a randomly chosen individual has a WAIS score of 105 or higher? Show your work.
(b) Find the mean and standard deviation of the sampling distribution of the average WAIS score $\bar{x}$ for an SRS of 60 people.
(c) What is the probability that the average WAIS score of an SRS of 60 people is 105 or higher? Show your work.
(d) Would your answers to any of parts (a), (b), or (c) be affected if the distribution of WAIS scores in the adult population were distinctly non-Normal? Explain.

R7.7 Detecting gypsy moths The gypsy moth is a serious threat to oak and aspen trees. A state agriculture department places traps throughout the state to detect the moths. Each month, an SRS of 50 traps is inspected, the number of moths in each trap is recorded, and the mean number of moths is calculated. Based on years of data, the distribution of moth counts is discrete and strongly skewed, with a mean of 0.5 and a standard deviation of 0.7.
(a) Explain why it is reasonable to use a Normal distribution to approximate the sampling distribution of $\bar{x}$ for SRSs of size 50.
(b) Estimate the probability that the mean number of moths in a sample of size 50 is greater than or equal to 0.6.
(c) In a recent month, the mean number of moths in an SRS of size 50 was 0.6. Based on this result, should the state agricultural department be worried that the moth population is getting larger in their state? Explain.