

**SEMESTER SPRING 2024**

**FINAL EXAM**

**STA 2050: SURVEY DESIGN AND ANALYSIS**

**Time 1.00PM– 3.00Am (2 hours)**

**Total point (40)**

**DATE: 3rd March 2024 TIME: Time 1.00PM– 3.00PM**

**Question one**

1. A state department of agriculture desires to estimate the number of acres under corn plantation within the state. Suggest possible sampling units and frames. (2marks )
2. How do we evaluate the goodness of an estimator? (2marks)

**Question Two**

1. Foresters want to estimate the average age of trees in a stand. Determining age is cumbersome, because one needs to count the tree rings on a core taken from the tree. In general, though, the older the tree, the larger the diameter, and diameter is easy to measure. The foresters measure the diameter of all 1132 trees and find that the population mean equals 10.3. They then randomly select 20 trees for age measurement.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tree no | Diameter X | Age y | Tree no | Diameter X | Age Y |
| 1 | 12 | 125 | 11 | 5.7 | 61 |
| 2 | 11.4 | 119 | 12 | 8 | 80 |
| 3 | 7.9 | 83 | 13 | 10.3 | 114 |
| 4 | 10.5 | 85 | 14 | 12 | 147 |
| 5 | 7.9 | 99 | 15 | 9.2 | 122 |
| 6 | 9 | 117 | 16 | 8.5 | 106 |
| 7 | 7.3 | 69 | 17 | 7.0 | 82 |
| 8 | 10.2 | 133 | 18 | 10.7 | 88 |
| 9 | 11.7 | 154 | 19 | 9.3 | 97 |
| 10 | 11.3 | 168 | 20 | 8.2 | 99 |
|  |  |  |  |  |  |

1. Draw a scatterplot of y vs. x. (3 marks)
2. Estimate the population mean age of trees in the stand using ratio estimation and give an approximate standard error for your estimate. (7marks)
3. Repeat (b) using regression estimation. (7marks)
4. Label your estimates on your graph. How do they compare? (3marks)

**Question Three**

1. A language school owner takes an SRS of 10 of the 72 Introductory Spanish classes offered by the school. Each student in each of the sampled classes is given a vocabulary test and is also asked whether he or she is planning a trip to a Spanish-speaking country in the next year. The data are in file spanish.dat.
2. Estimate the total number of students planning a trip to a Spanish-speaking country in the next year, and give a 95% CI. (6 marks)
3. Estimate the mean vocabulary test score for Introductory Spanish students in the language school, and give a 95% (6 marks)
4. An industry produces customized truck bodies in 20 plants scattered throughout the country. Only a small number of truck bodies are completed at each plant on any given day. For a particular day, it is desired to estimate the average worker-hours of work for items completed that day in the industry. Discuss the relative merits of cluster sampling (plants as clusters) versus stratified random sampling (plants as strata) for estimating the desired average. Which design do you recommend? (4marks )