

King of canada essay



**ASSIGN
BUSTER**

Stat 332 Sampling and Experimental Design: (Due: 12: 29pm before class, Friday, Oct 19, 2012) Assignment 2

1. In the upcoming US presidential election this November, Florida is a key swing state that is very important in determining the outcome. Currently the Democratic candidate Barack Obama and the Republican candidate Mitt Romney are virtually tied in recent polls. For this question, you can use 9 million as our population size; this is approximately the total number votes in the 2008 presidential election in Florida.

a) The latest poll asked 890 likely voters and found that the proportion of Florida residents likely to vote for the Democratic candidate is 47%, construct a 95% confidence interval for this estimate under the SRSWOR. By the way, the support for Republican is 46%, and the rest are undecided.

(b) If a news organization wants to conduct a SRSWOR survey to find out the percentage of Florida residents supporting the Democratic candidate, what is the sample size needed to achieve a result which is accurate within 0.5 percentage point, 19 out of 20 times?

2. We have 120 students in Stat332.

(a) What is the total number of possible samples of size 10 under SRSWOR?

(b) Suppose that among the 120 students of Stat332, 60 are male students and 60 are female students. We want to perform a STSRS (stratified simple random sampling) with male students and female students being two strata, and we want to survey 5 students from each stratum, what is the total number of possible samples? What is the ratio of the two numbers (STSRS/SRSWOR)?

3. Lohr (2010) Ch3, Q10 part a.

Hard shell clams may be sampled by using a dredge. Clams do not tend to be uniformly distributed in a body of water, however, as some areas provide better habitat than others. Thus, taking a simple random sample is likely to

result in a large estimated variance for the number of clams in an area.

Russell (1972) used stratified random sampling to estimate the total number of bushels of hard shell clams (*Mercenaria mercenaria*) in Narragansett Bay, Rhode Island. The area of interest was divided into four strata based on preliminary surveys that identified areas in which clams were abundant. Then n_h dredge tows were made in stratum h , for $h = 1, 2, 3, 4$. The acreage for each stratum was known, and Russell calculated that the area covered during a standard dredge tow was 0.039 acres, so that we may use $N_h = 25.6 \times \text{Area}_h$.

Here are the results from the survey taken before the commercial season. Estimate the total number of bushels of clams in the area, and give the standard error of your estimate.

| Stratum | 1 | 2 | 3 | 4 |
|--|--------|--------|--------|--------|
| Area (Acres) | 222.81 | 449.61 | 650.25 | 3197.1 |
| Tows Made | 4 | 6 | 3 | 5 |
| Average Number of Sample Bushels per Tow for Stratum | 0.44 | 0.068 | 1.17 | 0.042 |
| Variance | 3.92 | 2.146 | 1.80 | 0.794 |

4. Assume the sample variance in Q3 is a good estimate for the strata variance, and we want a stratified sample size of $n = 21$. (a) Calculate the stratum sample sizes under proportional allocation. (b) Calculate the stratum sample sizes under optimal (Neyman) allocation. (c) What are the values (approximately) of $V(\bar{y}_{st})$ with sample size allocations of (a) and (b), respectively? Compare these two variances and comment.

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. The data can be downloaded as hw2q5.txt from course website on Waterloo Learn and read into R using