Exam 3 take home questions (ThE3)

Stat 422

Sections 1 and 10

Fall 2018

Instructions:

Since this is part of exam 3, you are to work individually on these problems. Do not discuss these questions with anyone except Renae. Work each problem on a *separate page*, bring the take home problems (and **all** intermediate calculations) with you to the exam. One or more of these problems will be handed in with the in-class exam, and information from others will be used on the in-class questions.

(1) A magical statistician has been challenged to estimate the number of Bertie Bott's Every Flavour Beans in a large container. The beans in the container are all red, blue, or yellow. The statistician took a simple random sample of 300 beans, and replaced them with purple beans (eggplant flavor) that were the same size and weight as the original beans. After stirring the container well, a second simple random sample of 500 beans is taken, of which 37 are purple.

  (a) Estimate the number of Bertie Bott's Every Flavour Beans in the container, and compute a bound for your estimate.
  (b) Now suppose that the goal was to obtain a second sample until there are 40 purple beans in the second sample, and in total, 550 beans are drawn for the second sample. Now estimate the number of Bertie Bott's Every Flavour Beans in the container, and compute a bound for your estimate.

(2) An investigator wants to estimate the proportion of Washington state residents that are under age 18. Of the 39 counties in the state of Washington, she takes a random sample of 10 counties, and is able to obtain a sample of the Census information for those counties contained in the file called 'Washres.csv'.

  (a) Based on these data, estimate the proportion of Washington state residents who are under age 18. Compute a bound for your estimate.
  (b) Based on these results, for a future study to estimate the proportion of Washington state residents that are under age 18, how many counties should be sampled to have a bound of $B = 0.03$?

(3) To estimate the average salary of players in the National Basketball Association (NBA) for the 2008-2009 season, a sample was taken. First, a simple random sample of 7 teams (out of 32 teams) was taken. For those 7 teams, a 1-in-3 systematic random sample of players was taken and the salary of selected players was recorded. Calculate an estimate of the average salary of NBA players in the 2008-2009 season, and compute a bound for your estimate. The data from the sample are available in the file called 'bball.csv'.

(4) On the question above about counties in Washington state, if we performed a study to estimate the number of unemployed adults in Washington, are there potential disadvantages to the sampling method used in question 2? Suggest one or more alternative sampling methods to use county-level data for estimation of the number of unemployed adults in Washington state, and briefly explain why your method would be better.