Exercises

1.Suppose we want to estimate the proportion of recipes in the Better Homes & Gardens New Cook Book that do not involve animal products. We plan to take an SRS of the N=1251 test kitchen-tested recipes, and want to use a 95% CI with margin of error 0.03. What is the required sample size?

$$n_{0,srs} = \frac{(Z_{\alpha/2})^2 \, p(1-p)}{(moe)^2} = \frac{(1.96)^2 \, 0.5(1-0.5)}{(0.03)^2} \approx 1067$$

$$n_{srs} = \frac{n_{0,srs}}{1 + \frac{n_{0,srs}}{N}}$$

$$n_{srs} = \frac{1067}{1 + \frac{1067}{1251}} = 576$$

Using fpc because sample is > 10% of the population.

2. A SRS is chosen from a population of 1000 households. Results are shown below. Estimate mean income and total income and construct a 95% confidence interval for each.

income	
Mean	60199.95
Standard Error	6671.014
Median	47526.5
Standard Deviation	42191.2
Minimum	0
Maximum	215448
Sum	2407998
Size	40

$$\bar{y} \pm 1.96 * \hat{\sigma}_{\bar{y}} = 60199.95 \pm 1.96 * 6671.014 = (47\,124.76\,, 73\,275.14)$$

$$N\bar{y} \pm 1.96 * N\hat{\sigma}_{\bar{y}}$$

3. Suppose you wanted to redo the survey above to achieve a moe (margin of error) of $8000. How large a sample size would be needed, if the population from which this sample came has 1000 members?

$$n_{0,srs} = \frac{(Z_{\alpha/2}S)^2}{(moe)^2} = \frac{(1.96 * 42191.2)^2}{(8000)^2} = 106.85 = 107$$

$$n_{srs} = \frac{n_{0,srs}}{1 + \frac{n_{0,srs}}{N}}$$

$$n_{srs} = \frac{107}{1 + \frac{107}{1000}} = 96.66 = 97$$

4. a Describe the difference between coverage error and sampling error in survey statistics.

Coverage error results from a mismatch of the target population and frame, due to undercoverage of the frame and/or ineligibles. It exists before the sample is drawn and thus is not a problem arising because we do a sample. Sampling error is the difference between the value of a statistic and the parameter value. It is deliberately introduced, and is a result of using a sample rather than a census for estimation.]

b. How can sampling error be reduced?

increasing sample size, or introducing stratification into the design.

c. How can response error be reduced?

better training of interviewers, improving questionnaire design

d. What is a probability sample? What is its advantage over a non-probability sample? Give an example of a non-probability sample.

A *probability sample* is one in which the probability of selection for every member of the sample is non-zero and known.

Probability samples allow us to use probability-based statistical procedures, such as confidence intervals and hypothesis tests, in drawing inferences about the population from which the sample was drawn.

Volunteer sample, judgment sample