JOHNS HOPKINS BLOOMBERG SCHOOL of PUBLIC HEALTH

JOHNS HOPKINS
BLOOMBERG
SCHOOL of PUBLIC HEALTH

Lecture 3f: Practice Problem Solutions

John McGready
Johns Hopkins University

Estimating a 95% Confidence Interval

1. Suppose you are interested in estimating the proportion of employed Baltimore residents who use public transportation to get to their workplace on a regular basis. You apriori hypothesize this proportion to be roughly 20%. Suppose this is the (unknown) truth, and you do a study to estimate this proportion. How precise (within what boundaries) will you be able to estimate a 95% confidence interval for this proportion if you take a single random sample based on each of the following sizes?
   - The 95% CI would be formed by the formula $\hat{p} \pm 2 \times SE(\hat{p})$
   - The portion $2 \times SE(\hat{p})$ is sometimes called the 95% error bound or margin of error
   - This is what we are asked to compute
   - To do so we will need to employ the formula $SE(\hat{p}) = \sqrt{\frac{p \times (1 - p)}{n}}$
   - We will need to use the assumed value of p, 20% or .20
   - This is a classic conundrum of study design (we will cover more in SR2)

Estimating a 95% Confidence Interval

1. So here we go with a table of the results:

n	$SE(\hat{p}) = \sqrt{\frac{p \times (1-p)}{n}}$	$2 \times SE(\hat{p})$
120	$\sqrt{\frac{.2 \times .8}{120}} \approx .037$	0.074 or 7.4%
600	$\sqrt{\frac{.2 \times .8}{600}} \approx .016$	0.032 or 3.2%
1200	$\sqrt{\frac{.2 \times .8}{1200}} \approx .011$	0.022 or 2.2%

Estimating a 95% Confidence Interval

2. Suppose your hypothesized estimate of the proportion of residents taking public transportation to work was changed to 50%. How precise (within what boundaries) will you be able to estimate a 95% confidence interval for this proportion if you take a single random sample based on each of the following sizes?
   a) n = 120
   b) n = 600
   c) n = 1,300

Estimating a 95% Confidence Interval

2. So here we go with a table of the results: (using p =.50 instead)

n	$SE(\hat{p}) = \sqrt{\frac{p \times (1-p)}{n}}$	$2 \times SE(\hat{p})$
120	$\sqrt{\frac{.5 \times .5}{120}} \approx .046$	$0.092 \text{ or } 9.2\%$
600	$\sqrt{\frac{.5 \times .5}{600}} \approx 0.020$	$0.040 \text{ or } 4.0\%$
1200	$\sqrt{\frac{.2 \times .8}{1200}} \approx .014$	$0.028 \text{ or } 2.8\%$

Estimating a 95% Confidence Interval

3. FYI: here is a graph of the 95% error bound for all choices of p = 0.01 to 0.99 for n = 120, n = 600, and n = 1,200