

Project 5: SAMPLING DISTRIBUTIONS

Statistics 401: Fall 2016

Due: 1:40pm on Tuesday, October 25

Your write-up must be typed. Please number your answers as the questions were numbered. Your grade will be determined by how well you answer the questions, your justification for your answers, and by the organization and clarity of our write-up.

1. Let X be the number of SPAM emails received per day per employee at a large software engineering company. Suppose that the distribution of X is:

X	0	1	2
$P(X = x)$	0.60	0.30	0.10

Compute μ_x and σ_x . See p. 9 of the Chapter 2 course notes, or pp. 105 and 107 of your textbook, for how to calculate μ and σ^2 .

2. On your shelf you have four books that you are planning to read. The numbers of pages in each of the four books are: 212, 350, 379 and 575.
 - (a) Compute μ and σ , the mean and standard deviation of the number of pages for the population of four textbooks. Use the fact that $P(x) = \frac{1}{N} = \frac{1}{4}$ in your computations. See p. 9 of the Chapter 2 course notes, or pp. 105 and 107 of your textbook, for how to calculate μ and σ^2 .
 - (b) Compute the **sampling distribution** of \bar{X} for all samples of size $n = 2$. Note that there are six ways to choose two books from four books.
 - (c) What must the center $\mu_{\bar{x}}$ of the sampling distribution be? Explain.
 - (d) What must the spread $\sigma_{\bar{x}}$ of the sampling distribution be? Why must you use the finite correction factor when calculating $\sigma_{\bar{x}}$.

Include the following two tables in your report that summarize your results from (b) and (c):

Table 1: All possible samples of size $n = 2$

Sample	\bar{x}
1. _____	_____
2. _____	_____
3. _____	_____
4. _____	_____
5. _____	_____
6. _____	_____

Table 2: Sampling Distribution of \bar{X}

Value of \bar{x}	$P(\bar{x})$
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
_____	_____
$\mu_{\bar{X}}$:	_____
$\sigma_{\bar{X}}^2$:	_____
$\sigma_{\bar{X}}$	_____

- Do problem 4.6 on page 205. The last sentence of the problem should state that “They repeated this 100 times and built a distribution of sample proportions.” Your answer to part (b) should refer to the CLT “rule of thumb for sample sizes” for sample proportions.
- A poll conducted in 2013 found that 52% of U.S. adult Twitter users get at least some news on Twitter (Pew Research Center, Washington, D.C. Twitter News Consumers: Young, Mobile and Educated, November 4, 2013). Assume that the true proportion of Twitter users who get some news from Twitter is $p = 0.52$.

A new simple random sample of $n = 100$ Twitter users were asked if they get some news from Twitter. Answer the following:

- What is the approximate sampling distribution of \hat{p} from a random sample size of $n = 100$ Twitter users? Explain why your answer is correct (i.e., consider the population size and the sample size, if necessary, in order to apply the CLT).
 - What is the probability that a random sample size of $n = 100$ Twitter users gives a \hat{p} that is less than 0.50?
- Do problem 4.38 on page 216. Explain why your answers are correct by considering the population size and the sample size, if necessary. Do you need to apply CLT?
 - Do problem 4.40 on page 216. When answering part (b), consider the population size and the sample size, if necessary. Do you need to apply CLT? Feel free to sketch part (d) by hand.