Overview

The purpose of this project is to conduct an empirical investigation of the theorems and ideas presented in section 5.2 of the text. In particular, you will conduct an experiment in which you generate samples of a random variable $X$ distributed binomially, with parameters $n = 20$ and $p = 0.5$. You will also generate samples of the associated proportion $\hat{p} = X/20$. Your report will be a one to two page type-written document (not including appendices) that is due on Friday, April 4, at the beginning of class.

Appendices

Prior to doing any analysis, I want you to do some grunt work, i.e. generate the data, produce a few plots, and perform some routine calculations. The results of this grunt work will be stored in appendices labeled A, B, and C, respectively. Details are as follows:

- **Appendix A: Data**
  To generate data for this project, start by looking in your wallet for a coin that seems “fair”, i.e. a coin which, if flipped, is equally likely to land heads as tails. Once you have found the coin, proceed as follows:

  1. Flip the coin 20 times, recording for each flip a 1 if the coin lands heads and a 0 if the coin lands tails.
  2. For this run of 20 flips, calculate $X$ (the total number of heads) and $\hat{p}$ (the proportion of flips that were heads.)
  3. Repeat the above two steps 10 times. You will ultimately generate 200 coin flips, and have 10 samples of $X$ and $\hat{p}$.
  4. Store your data in a table that is easy to read, and include this table as Appendix A of this report.

- **Appendix B: Plots**
  Use the data you generated above to form two plots. The first should be a histogram of the 10 values of $X$ that you generated. The second should be a histogram of the 10 values of $\hat{p}$ that you generated. Each histogram should include a “fit” normal line (this is default in Minitab.) Make sure to label your axes, give your figures titles, and include captions that indicate what each figure presents. Include these two figures as Appendix B of this report.

- **Appendix C: Calculations**
  Perform the following calculations.

  1. Calculate $\mu_X$, $\sigma_X$, $\mu_{\hat{p}}$, and $\sigma_{\hat{p}}$. (Assume your coin is fair.)
  2. Calculate the mean and standard deviation of your 10 samples of $X$. (Recall the formulas you learned in Chapter 1 for the mean and standard deviation of a bunch of concrete numbers.) Also calculate the mean and standard deviation of your 10 samples of $\hat{p}$.
  3. Calculate the exact probability that $X$ is $\leq 5$ or $\geq 15$. (You’ll need the binomial table for this.) Note that this probability is the same as the probability that $\hat{p}$ is less than $1/4$ or greater than $3/4$.
  4. Finally, use all the raw flip data (all 200 flips) to form an estimate for $p$. What is the standard deviation of this estimate?

Record the results of these calculations in Appendix C of this report.
Analysis

The main body of your report should contain the following:

• An introductory paragraph giving a brief description of the data set and how it was generated.

• Another paragraph that provides a qualitative and/or quantitative summary of your plots. This paragraph should address such issues as center, spread, and skew. You should explicitly address the issue of whether or not you think your data is roughly normal.

• A paragraph in which you describe what you calculated. You should make a clean distinction between which values are “theory” based and which are “data” based. You should specify any assumption you made in making your calculations.

• A paragraph in which you comment on how well theory is born out in practice within this experiment. Things to consider:
  – How close are the theoretical means and standard deviations to the empirical means and standard deviations?
  – What percentage of your $X$ values are within the range $[6, 14]$? Is this consistent with your third calculation?
  – How many standard deviations away from $p = 1/2$ is the estimate you got from using all 200 flips? (See calculation 4, above.) Is this reasonable? Why or why not?

• (Optional): one last paragraph in which you draw attention to anything else that is interesting and of note in your data.

Technical Specifications:
Your paper should adhere to the following guidelines:

• Use standard letter paper with reasonable margins and a reasonable font size (11 or 12 pt). Printing double-sided is fine.

• Use double or one-and-a-half line spacing to allow room for me to write comments.

• Label each table and figure (Table 1, Table 2, Figure 1, etc.) and give a reference in the text to each one (e.g. “. . . as shown in Table 2.”).