1. A company wants to understand what proportion (p) of customer's opt to give their change to a local charity. The plan is to take a random sample of the customers and record whether or not the customer gave their change to the charity. Define the random variable X_i for the response of the ith customer where a 0 represents that they did not give and a 1 represents that they did.

   Answer:
   The probability a random customer does donates to charity is p and the probability they do not give is $1-p$. Thus, the random variable X_i that is an indicator that the ith customer donates is a Bernoulli random variable with probability of success p where success is defined as the customer donating. Since we have a random sample of customers it is reasonable to assume that they are independent. Thus, we have $X_i \overset{ind}{\sim} Ber(p)$.

   (a) What is $E[X_i]$?

   Answer:
   By properties of the Bernoulli distribution, $E[X_i] = p$.

   (b) What is $E[\overline{X}]$?

   Answer:
   The sample mean always has the same expectation as an individual observation (when the individual observations have the same expectation) and thus $E[\overline{X}] = E[X_i] = p$.

   (c) Is $E[\overline{X}]$ unbiased for the proportion p?

   Answer:
   Yes. By part b) $E[\overline{X}] = p$ and thus $\overline{X}$ is unbiased for p.

   (d) What is the variance of X_i?

   Answer:
   By properties of the Bernoulli distribution, $Var[X_i] = p(1-p)$.

   (e) What is the standard deviation of X_i?

   Answer:

$$SD[X_i] = \sqrt{Var[X_i]} = \sqrt{p(1-p)}.$$

   (f) What is the standard error of $\overline{X}$?

   Answer:

$$SE[\overline{X}] = SD[X_i]/\sqrt{n} = \sqrt{p(1-p)}/\sqrt{n} = \sqrt{\frac{p(1-p)}{n}}$$

2. Continue the previous example and suppose $p = 0.5$.

   (a) What sample size is needed so that with (approximately) 68% probability we are within 0.1 of the true population proportion?

   Answer:
   If we assume $p = 0.5$, then the standard error is $\sqrt{0.5(1-0.5)/n} = 0.5/\sqrt{n}$. By the Empirical Rule, 68% is within 1 standard deviation of the truth and we are saying that 1 standard deviation is 0.1. Thus $0.5/\sqrt{n} = 0.1$ and thus $n = 25$.

   (b) What sample size is needed so that with (approximately) 95% probability we are within 0.1 of the true population proportion?

   Answer:
   By the Empirical Rule, 95% is within 2 standard deviation of the truth and we are saying that 2 standard deviations is 0.1. Thus $2 \times 0.5/\sqrt{n} = 0.1$ and thus $n = 100$.

   (c) What sample size is needed so that with (approximately) 99.7% probability we are within 0.1 of the true population proportion?

   Answer:
   By the Empirical Rule, 95% is within 3 standard deviation of the truth and we are saying that 3 standard deviations is 0.1. Thus $3 \times 0.5/\sqrt{n} = 0.1$ and thus $n = 225$.

   (d) How will these sample sizes change if p is closer to 0 or 1?

   Answer:
   If p is closer to 0 or 1, then $\sqrt{p(1-p)} < 0.5$. Thus a smaller sample size will be needed. For example, if $p = 0.1$ (or $p = 0.9$), then $\sqrt{p(1-p)} = 0.3$ and the sample sizes that are needed are 9, 36, and 81 respectively. Thus $p = 0.5$ can be used as a conservative estimate of the sample size that is needed.