Homework 6

1. Flipping a fair coin N times.

   a. $P(h \in [495000, 505000])$ given that $N = 10^6$

   If we change this into terms of a random variable F where $F = \frac{h-500000}{500}$.

   Then we get:

   $P(F \in [-\ 10, 10])$

   Which equals

   $$\int_{-10}^{10} \frac{1}{\sqrt{2\pi}} e^{-\frac{F^2}{2}} dF$$

   b. $P(h > 9000)$ given that $N = 10^4$

   using $Z = \frac{9000-500}{50}$

   We get

   $P(Z > 80)$

   Then:

   $$\int_{80}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{F^2}{2}} dF$$

   c. $P(h < 40\ or\ h > 60)$ given that $N = 10^2$

   using $Z = \frac{40-50}{5}$

   We get

   $P(Z <- 2)$

And if we use $Z = \frac{60-50}{5}$

We get

$P(Z > 2)$

To find the probability we add the two:

$$\int_{-\infty}^{-2} \frac{1}{\sqrt{2\pi}} e^{-\frac{F^2}{2}} dF + \int_{2}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{F^2}{2}} dF$$

2. Textbook Problem 6.3

a.

Standard error can be calculated using:

$$SE = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Substitute the known values:

$$SE = \sqrt{\frac{0.3(1-0.3)}{10}} = \sqrt{0.021} \approx 0.144$$

b. B

To estimate that we just need to set SE = 0.5 and solve the SE equation for n:

$$0.05 = \sqrt{\frac{0.3(1-0.3)}{n}}$$

$$0.0025 \times n = 0.21$$

$$n = \frac{0.21}{0.0025} = 84$$

Therefore you would have to repeat it 84 times.

3.

a. Textbook Problem 6.4

i. Give 68% confidence interval

First lets find the standard error

$$SE = \frac{75}{\sqrt{40}} \approx \frac{75}{6.32} \approx 11.84\ grams$$

To find the confidence interval of 68% we need to use the mean $\pm$ 1 standard deviation. So,

$$CI = 340 \pm 1 \times 11.84$$

$$Lower\ Limit = 340 - 11.84 \approx 328.16$$

$$Upper\ Limit = 340 + 11.84 \approx 351.84$$

Therefore the confidence interval is

$\approx (328.16, 351.84)\ grams$

ii. Give 80% confidence interval

The first thing we need to calculate is how many standard deviations 80% is. Using a Z table we find it is 1.28. Then:

$$CI = 340 \pm 1.28 \times 11.84$$

$$Lower\ Limit = 340 - 15.14 \approx 324.86$$

$$Upper\ Limit = 340 + 15.14 \approx 355.14$$

Therefore the confidence interval is

$\approx (324.86, 355.14)\ grams$

b. Kind of textbook problem 6.5

We could use a t-distribution, the degrees of freedom for a t-distribution is calculated with this equation:

$$df = n - 1$$

Where n is the sample size, so in our case df = 9

4. Textbook Problem 6.6

Before we continue let's calculate the proportion of female births and male births.

$$\hat{p}_f = \frac{x_f}{n} = \frac{1026}{2009} \approx 0.510$$

$$\hat{p}_m = \frac{x_m}{n} = \frac{983}{2009} \approx 0.490$$

a.

The SE for female births is:

$$SE = \sqrt{\frac{0.510(1-0.510)}{2009}} = \sqrt{\frac{0.2499}{2009}} \approx 0.0111$$

We can use that SE to calculate the CI for female births:

$$CI_f = 0.510 \pm 2.576 \times 0.0111$$

So,

$$CI_f \approx (0.4814, 0.5386)$$

b. B

The SE for malebirths is:

$$SE = \sqrt{\frac{0.489(1-0.489)}{2009}} = \sqrt{\frac{0.2499}{2009}} \approx 0.0111$$

We can use that SE to calculate the CI for female births:

$$CI_m = 0.489 \pm 2.576 \times 0.0111$$

So,
$CI_m \approx (0.4604, 0.5176)$

c. C
Yes, the intervals do overlap. This suggests that our estimates might not be entirely accurate and we might need a larger sample to be accurate. The male and female births might be closer to each other.

5. Textbook Problem 6.8

a. A

$$\bar{x} = \frac{\sum flavanoids}{n}$$

$$std_1 = \sqrt{\frac{\sum(x_i - \bar{x}_1)^2}{n-1}}$$

The confidence interval is

2.7 ± 0.165

so,

$CI_1 \approx (2.535, 2.865)$

b. B

$$\bar{x} = \frac{\sum flavanoids}{n}$$

$$std_1 = \sqrt{\frac{\sum(x_i - \bar{x}_1)^2}{n-1}}$$

The confidence interval is

1.9 ± 0.195

so,

$CI_1 \approx (1.751, 2.049)$

c. C

They do not overlap meaning there is a difference between the flavanoids in the two regions.