1. Exercise 5.40 (page 180)

   Solution: In a population of flatworms in a pond, twenty percent are adult. The sample proportion $\hat{p}$ of adults in a sample of size 20 is an estimate of the population proportion. The count of adults in the sample is Y and has a binomial distribution with $n = 20$ and $p = 0.2$.

   (a) $\Pr\{\hat{p} = p\} = \Pr\{Y = 4\} = 0.2182$.
   (b) $\Pr\{0.15 \le \hat{p} \le 0.25\} = \Pr\{3 \le Y \le 5\} = 0.5981$.

2. Exercise 5.41(b) (page 180) except use R to compute the probability exactly. Then, do the normal approximation. Note that $\Pr\{0.15 \le \hat{p} \le 0.25\} = \Pr\{3 \le Y \le 5\}$. Find the area between 2.5 and 5.5 under a normal curve whose mean and standard deviation agree with the mean and standard deviation of the approriate binomial distribution.

   Solution: The first part of this problem gives the answer to part (b) of the previous problem. There are at least two ways to do this in R.

```
> sum(dbinom(3:5, 20, 0.2))

[1] 0.5981231

> pbinom(5, 20, 0.2) - pbinom(2, 20, 0.2)

[1] 0.5981231
```

   For the second part of the problem, the normal curve that best approximates the binomial distribution will have mean $\mu = 20 \times 0.2$ and standard deviaiton $\sqrt{20(0.2)(0.8)}$. The discrete probability of exactly three successes is approximated by the continuous probability associated with the area from 2.5 to 3.5 under the normal curve. Putting these together for 3, 4, and 5, we want the area from 2.5 to 5.5 for the approximation. We can use R to make the normal approximation.

```
> mu <- 20 * 0.2
> sigma <- sqrt(20 * 0.2 * 0.8)
> mu

[1] 4

> sigma

[1] 1.788854

> pnorm(5.5, mu, sigma) - pnorm(2.5, mu, sigma)

[1] 0.5982644
```

3. Exercise 5.47 (page 181).

   Solution: Skull breadths in a population of animals are normally distributed with a standard deviation of 10 mm. $\bar{Y}$ is the sample mean in a sample of size 64 and μ is the population mean.

   (a) Suppose $\mu = 50$ mm. Find $\Pr\{\bar{Y}$ is within ± 2 mm of $\mu\}$.

```
> pnorm(52, 50, 10/sqrt(64)) - pnorm(48, 50, 10/sqrt(64))
[1] 0.8904014
```

   (b) Suppose $\mu = 50$ mm. Find $\Pr\{\bar{Y}$ is within ± 2 mm of $\mu\}$.

```
> pnorm(102, 100, 10/sqrt(64)) - pnorm(98, 100, 10/sqrt(64))
[1] 0.8904014
```

   (c) If μ were unknown, we could still find the answer because the z-score does not depend on μ. In particular, for the upper endpoint

$$z = \frac{(\mu + 2) - \mu}{10/\sqrt{64}} = \frac{2}{10/\sqrt{64}}.$$

   The lower endpoint has the negative of this z-score.

4. Exercise 6.4 (page 188).

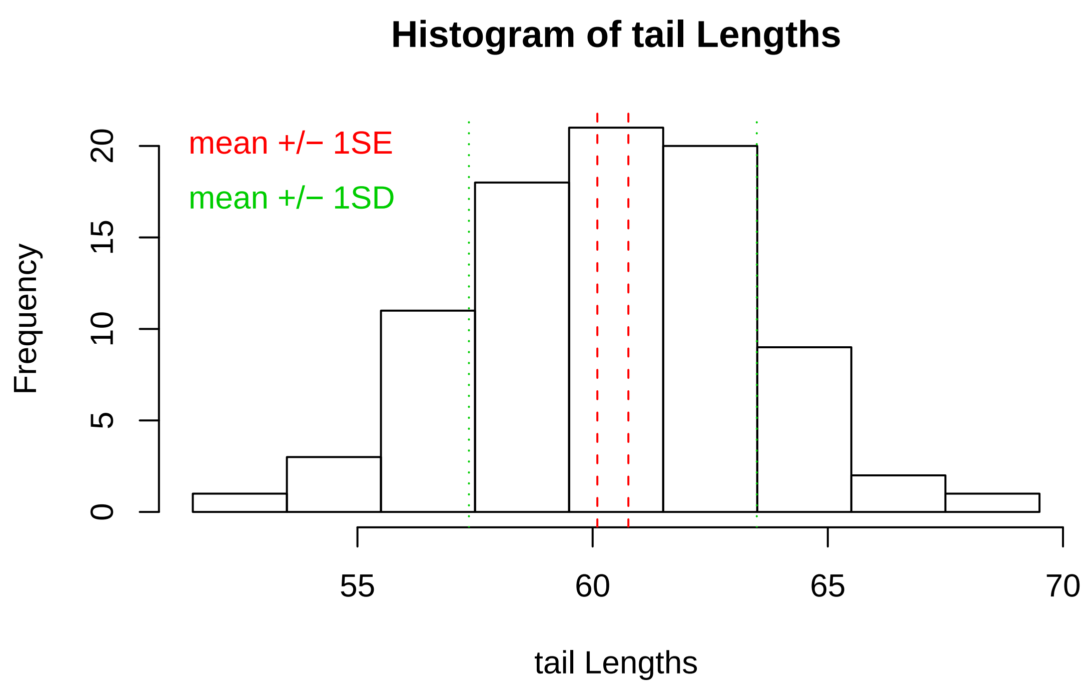
Solution: The tail lengths of one-year old individuals of the deermouse *Peromyscus* in a sample of size 86 have mean 60.43 mm and stnadard deviation 3.06 mm.

(a) The standard error of the mean is $s/\sqrt{n} = 3.06/\sqrt{86} \doteq 0.33$.

(b) Construct a histogram and indicate the intervals $\bar{y} \pm \mathrm{SD}$ and $\bar{y} \pm \mathrm{SE}$ on the histogram.

I'll do this in R as an example for those who want to learn how to modify plots. (You can do it by hand, of course.) The R function `hist` first calculates counts and breaks and then calls the function `plot.histogram` to make the graph. Because we have the counts and breaks already, I will use `plot.histogram` directly. Also, the function `abline` adds a line to an existing plot. I will use it to add vertical lines (with option `v=`) to indicate the intervals.

```
> ybar <- 60.43
> sd <- 3.06
> se <- 3.06/sqrt(86)
> counts <- c(1, 3, 11, 18, 21, 20, 9, 2, 1)
> breaks <- seq(51.5, 69.5, by = 2)
> xname <- "tail Lengths"
> plot.histogram(list(counts = counts, breaks = breaks, xname = xname))
> abline(v = ybar - se, lty = 2, col = 2)
> abline(v = ybar + se, lty = 2, col = 2)
> abline(v = ybar - sd, lty = 3, col = 3)
> abline(v = ybar + sd, lty = 3, col = 3)
> text(51, 20, "mean +/- 1SE", col = 2, pos = 4)
> text(51, 17, "mean +/- 1SD", col = 3, pos = 4)
```



5. Exercise 6.5 (page 188).

Solution: Refer to the previous exercise. Suppose there were 500 additional measurements.

(a) Predict the stnadard deviaiton of the new 500 measurements.

The sample standard deviation s is an estimate of the population standard deviation σ. I would predict that the new standard deviation would be close to the first, 3.06.

(b) Predict the new SE.
This number is different because there is a larger sample size. I would predict $3.06/\sqrt{500} \doteq 0.1368$.

6. Exercise 6.9 (page 198).

Solution: The weights of thymus glands in five chick embryos after 14 days of incubation have a mean of 31.7 mg and a standard deviation of 8.7 mg.

(a) Calculate the standard error of the sample mean.
$s/\sqrt{n} = 8.7/\sqrt{5} \doteq 3.9$.

(b) Construct a 90% confidence interval for μ.
There are 4 degrees of freedom. Form the table, the t-multiplier is 2.132. The confidence interval is $31.7 \pm 2.132 \times (8.7/\sqrt{5})$, or $(23.4, 40)$.

(c) Construct a 95% confidence interval for μ.
There are 4 degrees of freedom. Form the table, the t-multiplier is 2.776. The confidence interval is $31.7 \pm 2.776 \times (8.7/\sqrt{5})$, or $(20.9, 42.5)$.

You could also use R to find the appropriate multipliers for these two problems.

```
> qt(0.95, 4)

[1] 2.131847

> qt(0.975, 4)

[1] 2.776445
```

7. Exercise 6.10 (page 198).

Solution: This exercise deals with the blood serum concentration (μg/ml) of Gentamicin 1.5 hours after injection at a dosage of 10mg/kg body weight in a sample of six healthy three-year-old female Suffolk sheep. The sample mean and standard deviation are 28.7 and 4.6, respectively.

(a) Construct a 95% confidence interval for the population mean μ.
There are five degrees of freedom. $28.7 \pm 2.571 \times 4.6/\sqrt{6}$, or $(23.9, 33.5)$.

(b) Define in words the population mean.
The population mean μ is the mean blood serum concentration in μg/ml of Gentamicin 1.5 hours after injection at a dosage of 10mg/kg body weight in healthy three-year-old female Suffolk sheep.

(c) The fact that the 95% confidence interval for μ contains nearly all the observations is mainly due to the small sample size. For much larger samples, confidence in the location of μ is much more concentrated and the interval will be much tighter.

8. Exercise 6.11 (page 198).

Solution: A sample of 86 individuals of the deermouse *Peromyscus* has a sample mean tail length of 60.43 mm and standard deviation of 3.06 mm leading to a 95% confidence interval of (59.77,61.09) for the population mean.

(a) The statement "We are 95% confident that the average tail length of the 86 individuals in the sample is between 59.77 and 61.09" is incorrect, because we are 100% confident that the sample mean is 60.43 which is in the interval.

(b) The statement "We are 95% confident that the average tail length of all the individuals in the populaiton is between 59.77 and 61.09" is correct.