Activity #15: Interval Estimation (confidence intervals; bootstrap method)

Previously, we learned about point estimates – single values that best estimate unknown parameters. We demonstrated, for example, the sample mean is an unbiased point estimate (and the maximum likelihood estimator) of the population mean. We also investigated why we need (n – 1) in the denominator of our unbiased estimate of the population standard deviation.

Instead of estimating a parameter with a <u>single value</u> that, due to sampling variation, may be a lousy estimate, we may want to construct an <u>interval of plausible values</u> for the parameter. As an example, we may wish to construct an interval (a, b) such that we're 95% confident a ≤ μ ≤ b. This is the goal of *interval estimation*. In this activity, we'll learn how to construct confidence intervals for parameters using bootstrap and theory-based methods. We'll also discuss interpretations and limitations of these confidence intervals, as well as Bayesian highest density intervals.

Scenario: An internet service provider is interested in estimating the average time it takes to repair a customer's problem. The plots below display a sample of n=136 repair times (measured in minutes).

$$n = 136$$
$$\bar{X} \approx 649.1397$$
$$s \approx 1036.02$$
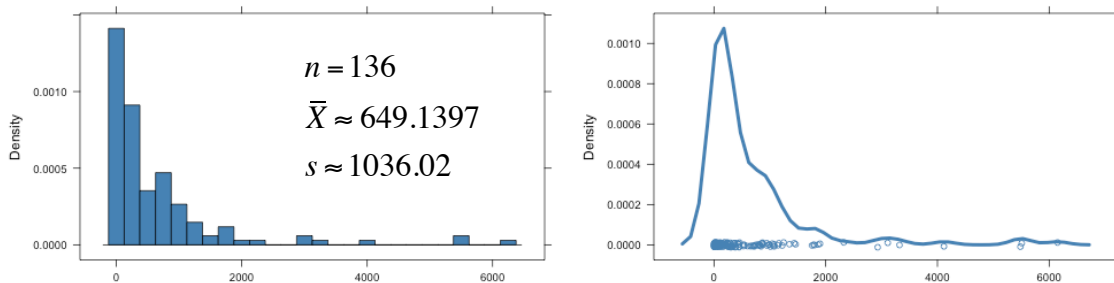
Data can be downloaded at: http://bradthiessen.com/html5/stats/m300/failures.csv

1. Suppose this sample of 136 repair times is a random sample from an unknown population of repair times.

      Based on this sample, what's our best point estimate for μ? _____

      How confident are you that this point estimate = μ? _____

      If we take another sample of 136 repair times, will we get the same point estimate? _____

Each random sample we take from our population will yield a different point estimate for μ. If we're only going to take a <u>single</u> sample of data, how could we ever judge the <u>accuracy</u> of our single point estimate?

One idea would be to <u>take a large number of samples</u> from the population and <u>calculate our point estimate for each sample</u>. Then we could see how much variation we get from all our sample means.

The problem with this idea is that it's impractical. Typically, we can only afford to take <u>one</u> large sample of data.

Could we <u>simulate</u> a large number of point estimates from different samples? While we don't know what it looks like, we do know <u>something</u> about the population. We have a random sample from this population! While we expect the population to be much larger than our sample, the values in our sample are representative of the values in the population. Let's pretend as though our sample <u>is</u> the population and <u>take a large number of samples</u>.

2. This is the logic behind <u>bootstrap methods</u>. We:

- Treat our sample of 136 values as if they represent the entire population. We then randomly sample n=136 values <u>with replacement</u> from our "population." Since we're sampling with replacement, the same values could be chosen multiple times (while other values are never chosen) in our <u>bootstrap sample</u>.

- From our bootstrap sample, we'll calculate our point estimate. In this scenario, we'll calculate $\bar{X}$.

- Repeat this process many times. We'll take 10,000 bootstrap samples of size n=136 and end up with 10,000 bootstrap estimates of $\bar{X}$.

- We can then graph those bootstrap estimates, calculate their mean, or calculate their standard deviation to get an idea of the variability in our point estimates.

Before we work with our repair time data, let's walk through the bootstrap sampling process with a small dataset.

Suppose our data consists of the numbers **1, 2, 3, 4, 5**.
To take a bootstrap sample, we sample n=5 values with replacement.

Bootstrap sample #1: 3 1 2 5 4   Mean = 3.0
Bootstrap sample #2: 3 1 1 3 1   Mean = 1.8
Bootstrap sample #3: 3 1 3 2 1   Mean = 2.0
(repeat 9,997 more times)

$\bar{X} \approx 2.99868$
$s \approx 0.64077$
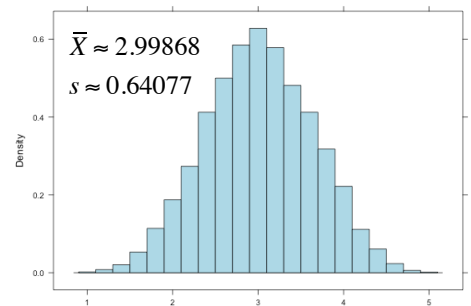
We can now examine the distribution of our 10,000 bootstrap means:

R code to generate bootstrap samples in this simple example

```
d <- c(1, 2, 3, 4, 5)                               # Create the dataset
library(mosaic)                                     # Load the mosaic package
sample <- resample(d)                               # Take one bootstrap sample with replacement
mean(sample)                                        # Calculate the mean of that bootstrap sample

bmeans <- do(10000) * mean(resample(d))             # Calculate means of 10,000 bootstrap samples
histogram(~result, data=bmeans, nint=21)            # Generate histogram of bootstrap means
mean(~result, data=bmeans)                          # Calculate mean of bootstrap means
sd(~result, data=bmeans)                            # Calculate std. dev. of bootstrap means
confint(bmeans, level=0.95, method="quantile")      # Get 95% confidence interval
```

---

3. Let's try this with our repair time dataset. We can use R or an online applet to generate our bootstrap distribution.

a) Copy this data: <u>http://bradthiessen.com/html5/stats/m300/failures.csv</u>
b) Open this applet: <u>http://lock5stat.com/statkey/bootstrap_1_quant/bootstrap_1_quant.html</u>

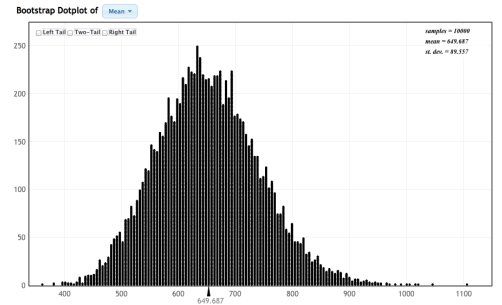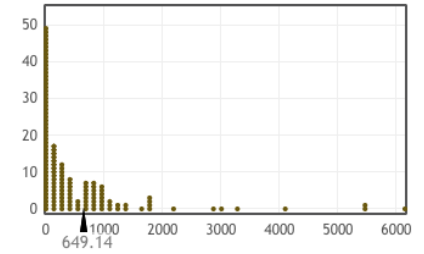R code to generate bootstrap samples for our repair time dataset

```
d <- read.csv("http://bradthiessen.com/html5/stats/m300/failures.csv")   # Load data
library(mosaic)
bmeans <- do(10000) * mean(resample(d))
histogram(~result, data=bmeans, nint=21)
mean(~result, data=bmeans)
sd(~result, data=bmeans)
confint(bmeans, level=0.95, method="quantile")
```

---

4. To use the applet:

- Click the EDIT DATA button at the top of the screen and paste the repair time data. When you click OK, you'll see the distribution of your data on the top-right of the screen (it will look just like this screenshot to the right).

Original Sample

n = 136, mean = 649.14
median = 272, stdev = 1036.02

- Click GENERATE 1 SAMPLE and the applet will take one bootstrap sample of size n=136. You can see the distribution of that sample on the bottom-right of the screen. The applet will also calculate the mean of that sample and plot that mean on the big graph in the middle of the screen.

- Go ahead and click GENERATE 1000 SAMPLES until you get at least 10,000 bootstrap samples. The distribution of your bootstrap sample means should look something like the screenshot to the right.


a) What does the lower histogram (the bootstrap distribution) represent?

b) Our original sample of data were heavily skewed to the right. Our bootstrap distribution of sample means looks approximately normal. Should we expect this bootstrap distribution to be approximately normal?

5. Remember our goal: We want to estimate the average of the population of repair times.

Our best point estimate is our single sample mean of 649.1397. Instead of relying on that single estimate, we want to construct an interval such that we're 95% confident the interval captures the population mean. How can we construct this interval using the bootstrap distribution we've generated?

95% bootstrap confidence interval for μ:

_____
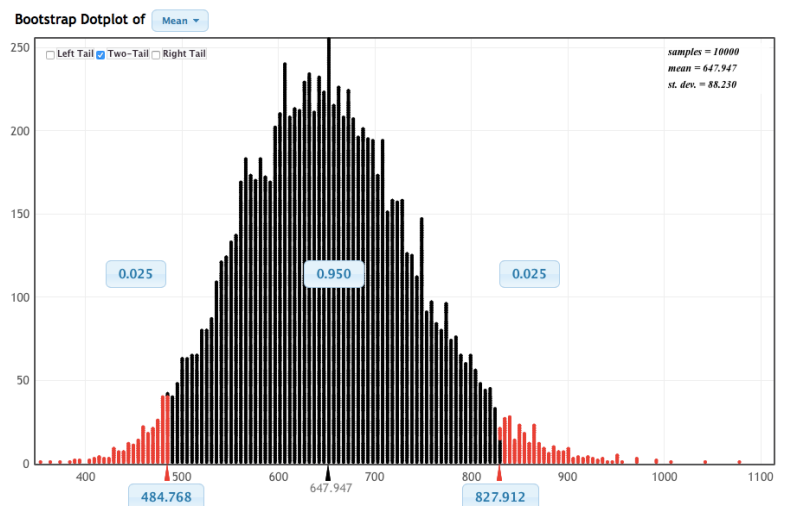
6. Record the center and width of the bootstrap confidence interval we just constructed.

    Center = _____   Width = _____

7. Load the **restaurant tips** dataset from the StatKey applet and construct a 90% bootstrap CI for μ.


    90% confidence interval: _____

8. Using the **restaurant tips** dataset, construct a 90% bootstrap CI for the population **median** and **standard deviation**.

    90% confidence interval for population **median**: _____

    90% confidence interval for population **standard deviation**: _____

9. How can we interpret these confidence intervals. Look at the interval you constructed in question #7.
    Can we say there's a 90% chance our interval captures the true population mean? Explain.

Scenario: ACT scores follow a normal distribution with μ = 21 and σ = 6.
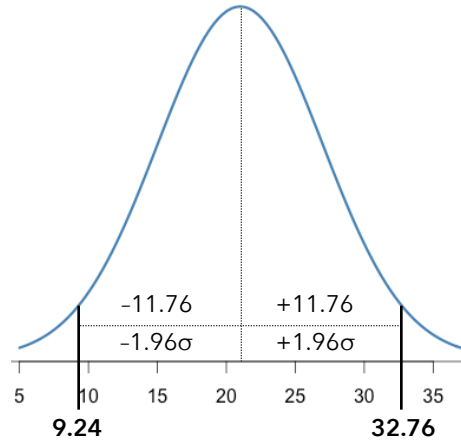
10. If we select an ACT score at random, what's the probability the score is between 9.24 and 32.76?

$$P(9.24 \le X \le 32.76) = P\left(\frac{9.24 - 21}{6} \le Z \le \frac{32.76 - 21}{6}\right) = P(-1.96 \le Z \le 1.96) = \text{normalcdf}(9.24, 32.76, 21, 6) = 0.95$$

Key idea: In a normal distribution, 95% of observations fall within +/-1.96σ of μ.

In this example, 95% of observations fall within 11.76 points of the mean (1.96 x 6 = 11.76)

11. Suppose we randomly select one ACT score (x) from this distribution. We then construct an interval by adding and subtracting 1.96 standard deviations: (x – 11.76) and (x + 11.76).

What's the probability our interval captures μ? P(x–1.96σ < μ < x+1.96σ) = P(x–11.76 < μ < x+11.76) = _____

Now suppose we repeat this process again and again. What proportion of these intervals will capture μ?
What's the probability that the <u>next</u> interval I create (from the next ACT score I sample) captures μ?
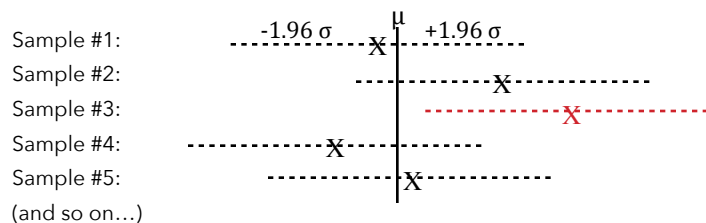
Key idea: In a normal distribution, X+/-1.96σ has a _____% chance of capturing μ.

12. Now suppose we take a random sample of n=16 ACT scores and calculate the average. If we were to do this repeatedly, described the sampling distribution of those sample means. Sketch this distribution below and calculate its mean and standard error.
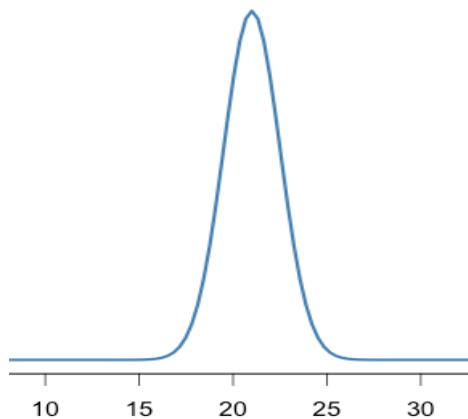
mean = _____

standard error = _____

10   15   20   25   30

13. There is a 95% chance the average ACT score from a single sample of n=16 is between _____ and _____.

14. The $\overline{X}$ symbols represent means from samples of n=16 ACT scores. Around each mean, I constructed an interval:

$$\pm 1.96\left(\sigma_{\overline{x}}\right) = \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right) = \pm 1.96\left(\frac{6}{\sqrt{16}}\right) = \pm 2.94$$

$\mu_{\overline{x}}$

Sample #1:	$\overline{X}$	Avg = 21.7	Interval: 18.76, 24.64
Sample #2:	$\overline{X}$	Avg = 19.6	Interval: 16.66, 22.54
Sample #3:	$\overline{X}$	Avg = 20.8	Interval: 17.86, 23.74
Sample #4:	$\overline{X}$	Avg = 16.8	Interval: 13.86, 19.74*
Sample #5:	$\overline{X}$	Avg = 22.6	Interval: 19.66, 25.54
(and so on…)			

I am going to take one more sample of n=16 ACT scores and calculate the mean. From this, I will construct an interval by adding and subtracting 2.94 points (1.96 standard errors). Before I take this sample, what's the probability that this interval will capture μ?

Key idea: If a sampling distribution of averages is approximately normal (when will this be the case?)

$$\overline{X} \pm 1.96\left(\sigma_{\overline{x}}\right) \text{ or } \overline{X} \pm 1.96\left(\frac{\sigma}{\sqrt{n}}\right) \text{ has a } _____\% \text{ chance of capturing μ.}$$

General form of a **confidence interval**: (estimate) +/– (critical value) (standard error)

$$100(1\text{-}\alpha)\% \text{ confidence interval for } \mu: \overline{X} \pm Z_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Let's use this formula (based on the Central Limit Theorem) to construct and interpret confidence intervals for means. In each case, evaluate whether the necessary assumptions have been satisfied.

15. Suppose we're interested in the average amount of debt of every college graduate in America*. We don't know the shape of the distribution (I imagine its positively skewed), but let's pretend we know **σ = $15,000**:

   a) Suppose we sample **one** student who graduates with **$40,000** in debt. Create a 95% confidence interval for μ.

   b) Suppose we sample **n=100** students with an average debt of **$25,000**. Create a 95% confidence interval for μ.

   c) Suppose we sample **n=10,000** students with an avg. debt of **$28,000**. Create a 95% confidence interval for μ.

16. As <u>n increases</u>, the <u>width</u> of our confidence interval _____

17. How do we interpret these confidence intervals? Look at our answers to 6a, b, and c. Is there a 95% chance that our intervals capture μ? Explain.

18. This formula for confidence intervals only works when we have a sampling distribution that is approximately normal. Is that the case here? Remember that the distribution of student debt most likely has a positive skew.

* 69% of college graduates left school with an average of $28,400 in debt (in 2013).
Source: http://www.usnews.com/news/articles/2014/11/13/average-student-loan-debt-hits-30-000

19. Suppose I want to report a 95% confidence interval for the average amount of student debt that is $2000 wide. How many college graduates would I need to sample?

20. Recall the formula for a confidence interval for μ: $\bar{X} \pm Z_{\alpha/2}\left(\dfrac{\sigma}{\sqrt{n}}\right)$. Other than increasing n, is there any other way we could <u>decrease</u> the width of a confidence interval?

21. Which would be <u>wider</u>:   90% confidence interval   99% confidence interval

22. Recall a previous dataset in which the standard deviation of body temperatures was σ = 0.75*. Suppose we sample n = 130 healthy adults and calculate their average body temperature to be 98.25 degrees.

a) Construct a 99% confidence interval for μ.

b) Construct a 95% confidence interval for μ.

c) Construct a 90% confidence interval for μ.

d) Suppose the actual population mean body temperature for a healthy adult is 98.6 degrees. Did any of our confidence intervals capture μ? How do we interpret these intervals.

* How could we know the population standard deviation if we don't know the population average? Good question.

23. Thus far, we've created confidence intervals by assuming we have σ, the population variance. If we don't know σ, we'll have to use our best estimate, s (the sample standard deviation).

$$\sigma_x = \sqrt{\frac{\sum(X-\mu)^2}{n}}$$

Population standard deviation
What we would like to use

$$s_x = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$$

Sample standard deviation
Our best (unbiased) estimate of σ

Note that we're calculating confidence intervals because we do not know μ. If we do not know μ, we cannot know σ, so it seems as though the confidence interval formula we've been using is rarely, if ever, useful.

Could we simply replace **σ** in our confidence interval formula with **s** ?

Can we replace $\mathrm{CI}:\ \bar{X} \pm Z\left(\dfrac{\sigma}{\sqrt{n}}\right)$ **with** $\mathrm{CI}:\ \bar{X} \pm Z\left(\dfrac{s}{\sqrt{n}}\right)$**?**

We know σ is going to be a single (unknown) value regardless of the sample we take. If we calculate s, it will <u>differ for each sample</u>. Since s varies for each sample, we'll have <u>more uncertainty</u> in our estimates.

We need a distribution that takes this greater uncertainty into account. Thankfully, William Gossett (a Guinness Breweries employee) published the Student's t-Distribution in 1908. The t-distribution is similar to our (standard normal) z-distribution, except it is fatter in the tails.

Note that to use the t-distribution, you need to know the *degrees of freedom*. We'll discuss degrees of freedom later. For now, in constructing a confidence interval for a population mean, we'll note **df = n - 1**.

Derivation: **Recall:** $\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}} \sim Z.$ **We want to know** $\dfrac{\bar{X}-\mu}{s/\sqrt{n}} \sim$???.

Some symbol manipulation: $\dfrac{\bar{X}-\mu}{s/\sqrt{n}} = \left(\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)\left(\dfrac{\sigma}{s}\right) = \dfrac{\left(\dfrac{\bar{X}-\mu}{\sigma/\sqrt{n}}\right)}{\left(s/\sigma\right)} = \dfrac{Z}{\sqrt{\dfrac{(n-1)s^2}{\sigma^2}/n-1}} =$

We don't discuss the chi-square distribution in this class, but it is: $\chi^2 = \dfrac{(n-1)s^2}{\sigma^2}$

Substituting the chi-square, we get: $\dfrac{\bar{X}-\mu}{s/\sqrt{n}} = \dfrac{Z}{\sqrt{\dfrac{(n-1)s^2}{\sigma^2}/n-1}} = \dfrac{Z}{\sqrt{\dfrac{\chi^2}{n-1}}} = \dfrac{Z}{\sqrt{\dfrac{\chi^2}{df}}} = T$

General form of a **confidence interval**: (estimate) +/– (critical value)(standard error)

$100(1-a)\%$ confidence interval for μ (when σ is known): $\bar{X} \pm Z_{\alpha/2}\left(\dfrac{\sigma}{\sqrt{n}}\right)$

$100(1-a)\%$ confidence interval for μ (when σ is unknown): $\bar{X} \pm t_{n-1}^{\alpha/2}\left(\dfrac{s}{\sqrt{n}}\right)$

t-distribution table: http://www.bradthiessen.com/html5/stats/m300/ttable.pdf
t-distribution calculator: http://lock5stat.com/statkey/theoretical_distribution/theoretical_distribution.html#t
confidence interval calculator: http://www.rossmanchance.com/applets/TBIA.html
confidence interval calculator: http://www.mathcelebrity.com/normconf.php

24. Construct a 90% confidence interval for the average number of hours slept by St. Ambrose students last night. Interpret the interval.

25. If we recorded the number of hours slept last night for each St. Ambrose student, what proportion of those values would fall within our intervals? What proportion of the values from students in this class fall within the interval?

26. We know the population average IQ score is 100. Suppose you find n=25 individuals who have reported seeing a UFO. You give an IQ test to these individuals and find their mean IQ is 101.5 with a standard deviation of 10. Construct a 90% confidence interval for the average IQ of all UFO observers. From this, can we conclude that UFO observers have higher IQs than all other adults?

27. In question #5, we used bootstrap methods to construct a 95% CI for the average repair time. Use the theory-based approach to construct a 95% confidence interval. Compare the center and width of the theory-based and bootstrap methods. Which method, in this scenario, should be preferred?

$n = 136$

$\bar{X} \approx 649.1397$

$s \approx 1036.02$
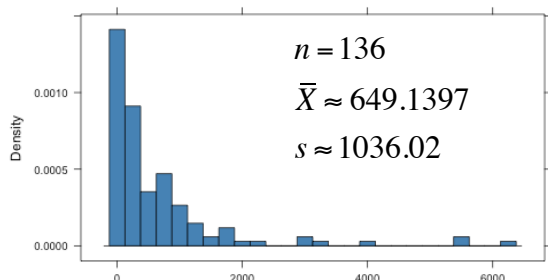
28. So far in this activity, we've constructed confidence intervals for μ. We can also calculate confidence intervals for proportions. Recall the general formula for a confidence interval:

Confidence Interval for Population Parameter: (estimate)   +/−   (critical value)   (standard error)

Back in Activity 13, we proved that the sample proportion (p) is an unbiased estimate of the population proportion.

Confidence Interval for Population Proportion: **(p)**   +/−   (critical value)   (standard error)

If we have a large sample size, we can assume the sampling distribution of proportions is approximately normal.

Confidence Interval for Population Proportion: (p)   +/−   **(Z)**   (standard error)

When we're dealing with proportions, we're counting the number of successes and dividing by the total sample size. Thus, each observation in our sample is either a success or failure. In Activity #7, we learned about Bernoulli random variables. We derived the variance of a Bernoulli random variable to be: $\mathrm{var}(x) = p(1-p)$. If we convert that variance to a standard deviation, we get $SD(x) = \sqrt{p(1-p)}$. Finally, converting a standard deviation to a standard error, we find:

Confidence Interval for Population Proportion: (p)   +/−   (Z) $\left(\sqrt{\dfrac{p(1-p)}{n}}\right)$

General form of a **confidence interval**: (estimate) +/− (critical value)(standard error)

100(1-α)% confidence interval for a proportion: $\hat{p} \pm Z_{\alpha/2}\sqrt{\dfrac{p(1-p)}{n}}$

---

29. A survey of 1000 adults finds 742 wish to abolish the penny. Construct and interpret 95% confidence interval.

30. Generate the same CI using bootstrap methods: http://lock5stat.com/statkey/bootstrap_1_cat/bootstrap_1_cat.html

31. Suppose we wanted the margin of error of our 95% confidence interval to be +/− 1%. How big of a sample do we need?

32. Suppose an alien lands on earth and wishes to estimate the percentage of all humans who are female. If the alien sampled the 100 members of the U.S. Senate, the alien would find 17 females and 83 males. With its superior intellect, the alien would then quickly calculate a 95% confidence interval:

$$0.17 \pm 1.96 \sqrt{\frac{.17(1-.17)}{100}} = (0.096,\ 0.244)$$

Interpret this interval.

33. The confidence intervals we've been calculating for proportions are called **Wald intervals**. Unfortunately, these intervals aren't very good. It turns out that it's not always the case that 95% of 95% Wald intervals contain the population proportion. The problem is that estimated standard error in the formula is a poor estimator of the actual standard error unless n is large.

So what should we do? We could use **bootstrap methods**. To do this, we can use the applet at: http://lock5stat.com/statkey/bootstrap_1_cat/bootstrap_1_cat.html.

I generated 10,000 bootstrap samples and found a 95% confidence interval for our alien example of: **(0.10, 0.24)**.

Another method designed to ensure 95% of our 95% confidence intervals contain the population proportion is the **Clopper-Pearson** method. This method, based on the binomial test we learned in activity #7, can be conducted in R with the following code:

```
d <- c(rep(0, 83), rep(1, 17))      # Enter number of wins (1) & losses (0) in the sample
binom.test(d == 1)                   # Obtain the confidence interval
```

This code yielded a 95% confidence interval of: **(0.1023, 0.2582)**.

Yet another approach (that's much easier to calculate by hand) is the **Wilson (or "plus 4") confidence interval**. If you want to learn more about this method, check out: http://www.mwsug.org/proceedings/2008/pharma/MWSUG-2008-P08.pdf.

Believe it or not, a good estimate of Wilson confidence intervals uses the same formul as Wald intervals. The only change is that we add 2 successes and 2 failures to our data before calculating the interval.

Wald formula (slightly rewritten in terms of successes (x) and failures (n-x): $p \pm z \sqrt{\dfrac{p(1-p)}{n}} = \dfrac{x}{n} \pm z \sqrt{\dfrac{\frac{x}{n}\left(1-\frac{x}{n}\right)}{n}}$

Estimate of Wilson formula: $\dfrac{x+2}{n+4} \pm z \sqrt{\dfrac{\frac{x+2}{n+4}\left(1-\frac{x+2}{n+4}\right)}{n+4}}$

Example: $\dfrac{17+2}{100+4} \pm 1.96 \sqrt{\dfrac{\frac{19}{104}\left(1-\frac{19}{104}\right)}{104}} = (0.108,\ 0.257)$

SUMMARY		
Method:	Lower,	Upper
Formula:	0.096,	0.244
Bootstrap:	0.100,	0.240
Clopper-Pearson:	0.102,	0.258
Wilson +4 CI:	0.108,	0.257

34. Finally, we can construct a highest density interval using Bayesian methods. I may not have enough time to explain Bayesian methods (or Markov Chain Monte Carlo methods), but I wanted to show you the output from a 90% HDI estimation:

Here is the R code I used to calculate the 90% HDI:

```
require(BEST)
crashes <- read.csv(url("http://bradthiessen.com/html5/stats/m300/failures.csv"))
crashvector <- crashes[,1]
BESTout1 <- BESTmcmc(crashvector)
BESTout1
plot(BESTout1)
hdi(BESTout1, credMass = 0.90)
```

Here is the output. Interpret these results:

```
Setting up the JAGS model...
Compiling model graph
   Resolving undeclared variables
   Allocating nodes
   Graph Size: 151

Initializing model
   |++++++++++++++++++++++++++++++++++++++++++++++++++| 100%
Burning in the MCMC chain...
   |**************************************************| 100%
Sampling final MCMC chain...
   |**************************************************| 100%

MCMC fit results for BEST analysis:
100002 simulations saved.
                            90%     90%
        mean      sd  median HDIlo   HDIup  Rhat n.eff
mu    226.066 39.8304 222.976 160.8 289.502 1.001 19482
nu      1.308  0.2332   1.258   1.0   1.622 1.001 16963
sigma 261.113 42.5210 257.172 191.9 328.183 1.001 15992
```
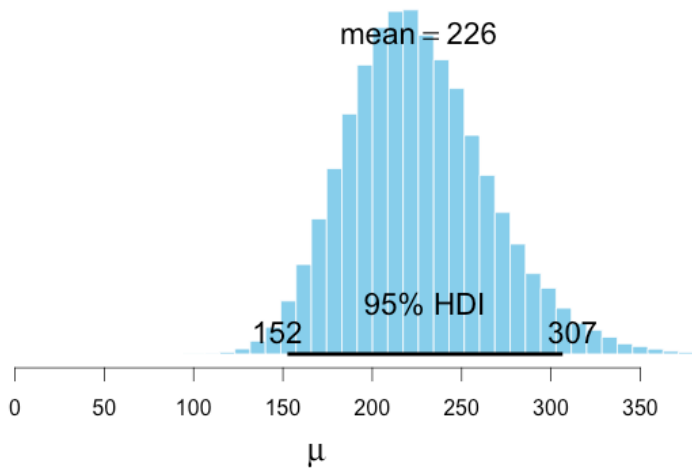

Method	90% Interval	Width	Center	Assumptions
Parametric	CI: (502.0, 796.3)	294.3	649.1	Normality, CLT
Bootstrap	CI: (508.1, 802.3)	294.2	655.2	
Bayesian	HDI: (289.5, 160.8)	128.7	225.2	Prior distribution of μ

Summary of interval estimation methods. Note: R code assumes single-column data is loaded as "d"

1. **Confidence interval for a single sample mean**
    a) Bootstrap method
        i) Applet: http://lock5stat.com/statkey/bootstrap_1_quant/bootstrap_1_quant.html
        ii) R code:
```
library(mosaic)
bmeans <- do(10000) * mean(resample(d))
confint(bmeans, level=0.95, method="quantile")
```

    b) Theory-based
        i) When σ is known: $\bar{X} \pm Z_{\alpha/2}\left(\dfrac{\sigma}{\sqrt{n}}\right)$

        z applet: http://lock5stat.com/statkey/theoretical_distribution/theoretical_distribution.html#normal

        Calculator applet: http://www.mathcelebrity.com/normconf.php

        R code:
```
observedMean = mean(d)
sigma=0.1                  # Enter population standard deviation
n=nrow(d)
confidence = 0.95     # Enter confidence level
alpha = (1-confidence)/2
stdError <- sigma/sqrt(n)
lower<- observedMean - qnorm(1-alpha)*(stdError)
upper<- observedMean + qnorm(1-alpha)*(stdError)
paste("CI: (", lower, " , ", upper, ")")
```

    ii) When σ is <u>not</u> known: $\bar{X} \pm t_{n-1}^{\alpha/2}\left(\dfrac{s}{\sqrt{n}}\right)$

        t applet: http://lock5stat.com/statkey/theoretical_distribution/theoretical_distribution.html#t

        Calculator applets: http://www.mathcelebrity.com/normconf.php
                             http://www.rossmanchance.com/applets/TBIA.html

        R code: `confint(t.test(~d, conf.level=0.95), data=d)`

2. **Confidence interval for a single sample proportion**
    a) Bootstrap method
        i) Applet: http://lock5stat.com/statkey/bootstrap_1_cat/bootstrap_1_cat.html

    b) Theory-based
        i) Wald formula: $\hat{p} \pm Z_{\alpha/2}\sqrt{\dfrac{p(1-p)}{n}}$ # This is NOT recommended!

        ii) Clopper-Pearson R code:
           R code: `d <- c(rep(0, 110), rep(1, 140))`   # Enter number of wins (1) & losses (0)
                   `binom.test(d == 1)`

        iii) Estimate of Wilson +4 confidence intervals: $\dfrac{x+2}{n+4} \pm z\sqrt{\dfrac{\dfrac{x+2}{n+4}\left(1-\dfrac{x+2}{n+4}\right)}{n+4}}$