Confidence Intervals

Robustness of Confidence Intervals

In many inferential procedures, there are necessary conditions that must be satisfied in order for the sampling distribution of our estimators to be appropriate. An important consideration in any analysis is to examine whether or not these conditions hold. Without the necessary conditions, we can not say that a 95% confidence interval covers the true theoretical parameter for 95% of all data sets. The actual coverage might be much smaller.

There are two necessary conditions which must hold in the case of confidence interval for a proportion.

1. The n observations are independent. For example each flip of a beer cap must be independent or people surveyed for a poll should not be related in any way.
2. The sample size n is large enough so that $\hat{p}$ has approximately a Normal distribution by the Central Limit Theorem

We will now investigate the effects of the second condition on the coverage of a confidence interval using an applet, shown in Figure 1 which simulates data and calculates confidence intervals. This applet uses the notation π for the theoretical parameter we have been calling p.

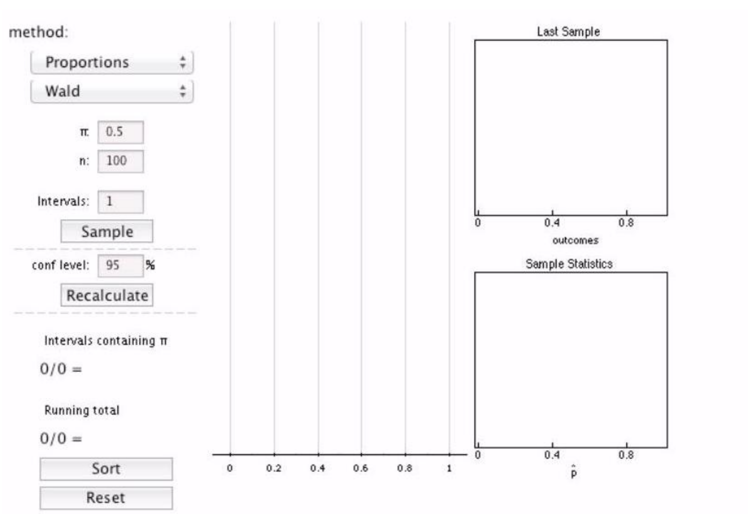

Figure 1: Applet for simulating confidence itervals for p by Allan Rossman and Beth Chance

Example 1

We want to flip a fair coin 10 times so on the applet, set $\pi = 0.5$ and $n = 10$. We get 6 successes and four failures so $\hat{p} = 0.6$.

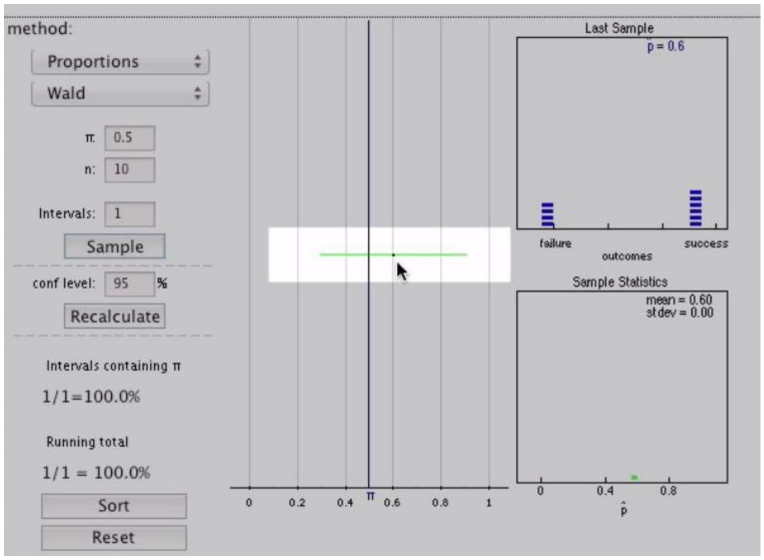

Figure 2: Highlighted confidence interval based on a sample with $n = 10$ and $\hat{p} = 0.6$

Using the 95% confidence interval formula for a proportion and substituting in the known variables:

$$\begin{aligned} 95\% \text{ CI for } p &= [\hat{p} - 1.96\sqrt{p(1-p)/n}, \hat{p} + 1.96\sqrt{p(1-p)/n}] \\ &= [0.6 - 1.96\sqrt{0.6(1-0.6)/n}, 0.6 + 1.96\sqrt{0.6(1-0.6)/n}] \\ &= [0.296, 0.904] \end{aligned}$$

In Figure 2, the vertical bar marks the theoretical parameter $\pi = 0.5$ and the horizontal green line plots the confidence interval we calculated above. We can see that the confidence interval from 0.296 to 0.904 includes 0.5.

We try another sample and again get $\hat{p} = 0.6$, shown in Figure 3.

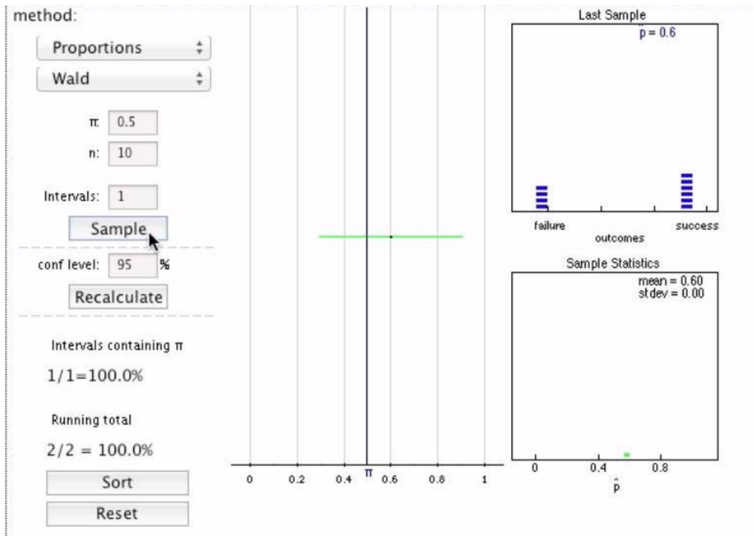

Figure 3: Sample 2, $n = 10$ and $\hat{p} = 0.6$

The third time we observe $\hat{p} = 0.8$, shown in Figure 4. Now the confidence interval for this sample does not include 0.5. (Note the horizontal line is colored red and does not cross the vertical bar.)

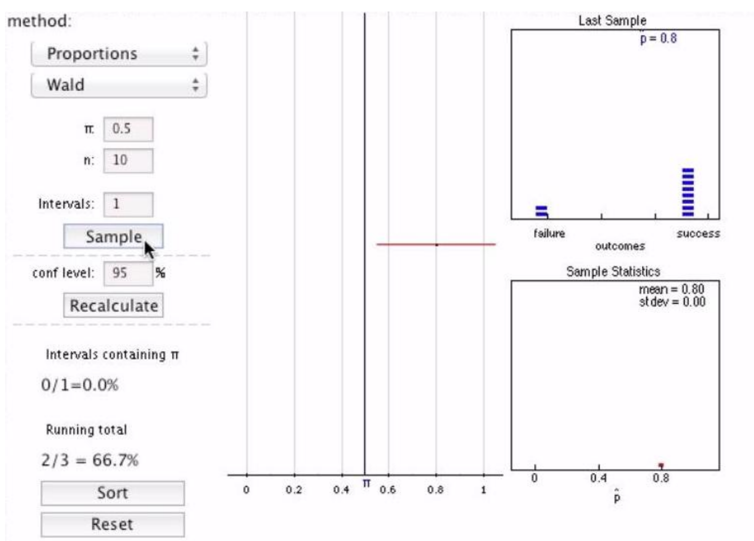

Figure 4: Sample 3, $n = 10$ and $\hat{p} = 0.8$

So far two out of three simulated confidence intervals have included the theoretical parameter. To better see how well this is working let's simulate 100 confidence intervals. We expect that 5% of the time the 95% confidence intervals will not contain the theoretical parameter. So for 100 95% confidence intervals we expect that 5 will not contain the theoretical parameter.

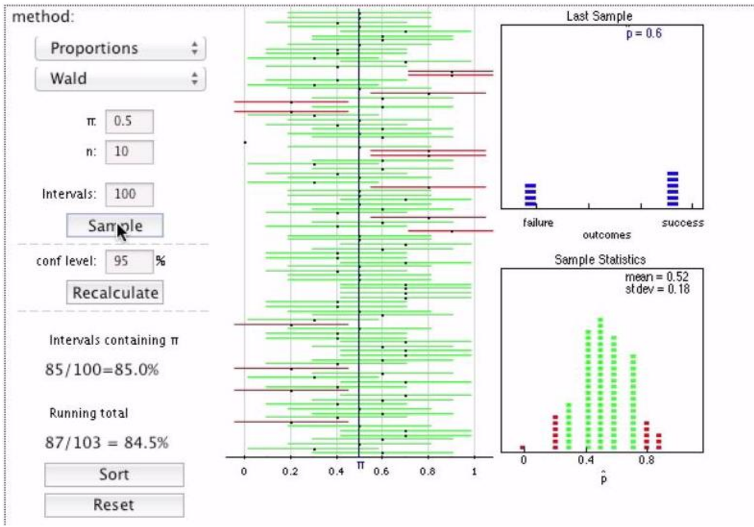

Figure 5: 100 confidence intervals based on samples with $n = 10$ and $p = 0.5$

As shown in Figure 5, 15 of the 100 confidence intervals based on simulated data did not contain 0.5. We have only captured the true value of our theoretical parameter 85% of the time instead of the 95% that the confidence interval was designed for. This is because the sample size of 10 is too small to rely on the Central Limit Theorem Normal approximation for the sampling distribution of $\hat{p}$. Since we have violated one of the assumptions of the confidence interval procedure for a proportion the results we are getting are not at the confidence level we expect.

Example 2

Next, for every experiment we will flip the fair coin 100 times. Again we will carry out 100 experiments. This time 94% of our confidence intervals included 0.5, which is very close to the 95% that we expected (Figure 20).

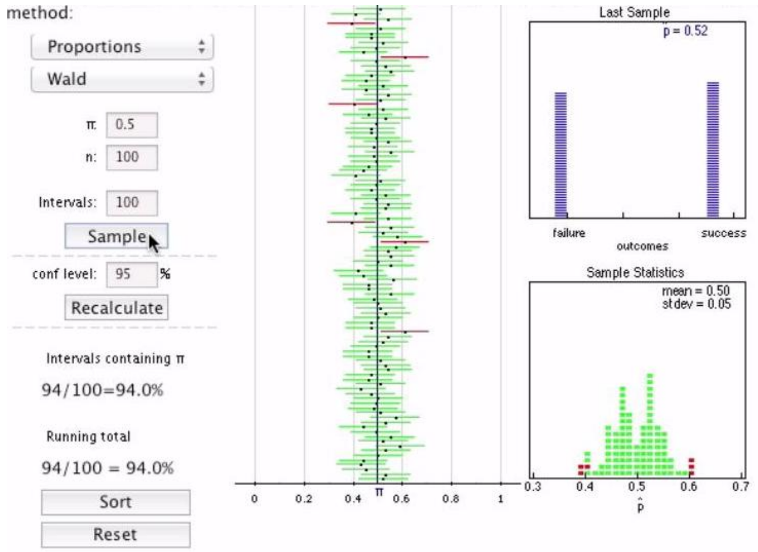

Figure 6: 100 confidence intervals based on samples with $n = 100$ and $p = 0.5$

EXAMPLE 3
When the theoretical parameter $p = 0.5$ the sampling distribution of $\hat{p}$ is symmetric with a strong central peak. A sample size of 100 provided reliable results in this case. This can be seen in Figure 7.

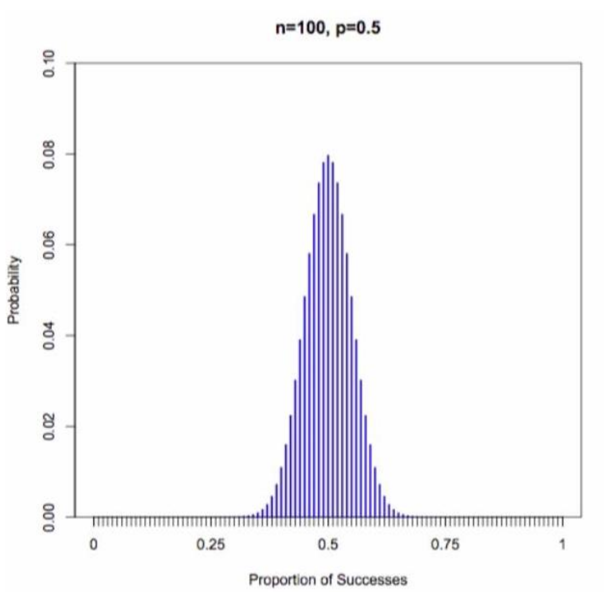

Figure 7: Sampling distribution for $\hat{p}$ when $p = 0.5$ and $n = 100$

When p is close to 0 or 1, the sampling distribution of $\hat{p}$ is skewed. For example, if $p = 0.05$, the sampling distribution is skewed to the right. The Normal distribution would not be a good approximation to this distribution. In such cases, we need a larger sample size before the normal approximation works well, as shown in Figure 8.

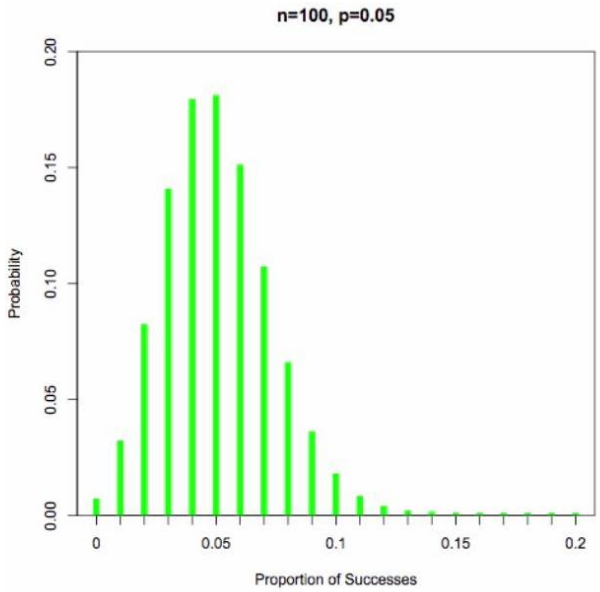

Figure 8: Sampling distribution for $\hat{p}$ when $p = 0.05$ and $n = 100$

Using the applet we will set $\pi = 0.05$ and take 100 samples each of size 100.

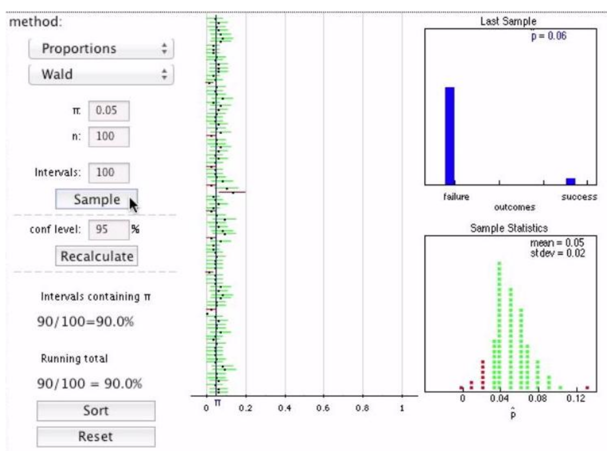

Figure 9: 100 confidence intervals based on samples with $n = 100$ and $p = 0.05$

Instead of 95% of our confidence intervals capturing 0.05, we had 90% of our confidence intervals capturing our theoretical parameter, shown in Figure 9. We repeat this experiment several times and find that each time about 90% of the confidence intervals captured the theoretical parameter. So the coverage of these intervals is not bad but it is not the 95% that we claim.

Let's try something even more extreme. This time we set the probability of success to $p = 0.01$. The sampling distribution of $\hat{p}$, shown in Figure 10, is even more sharply skewed.

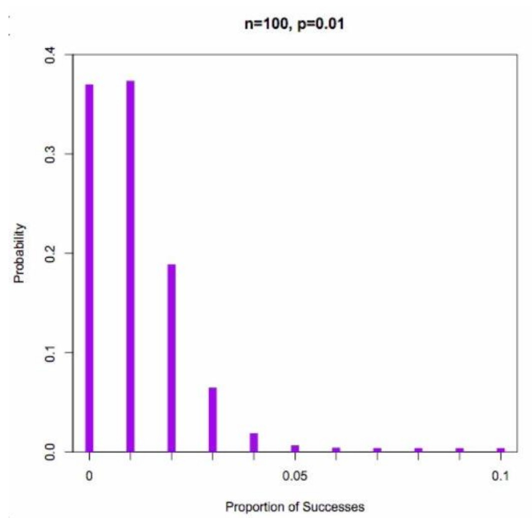

Figure 10: Sampling distribution for $\hat{p}$ when $p = 0.01$ and $n = 100$

Again we take 100 samples each of size 100. We find that only 59% of the 95% confidence intervals captured the theoretical parameter. This is shown in Figure 11.

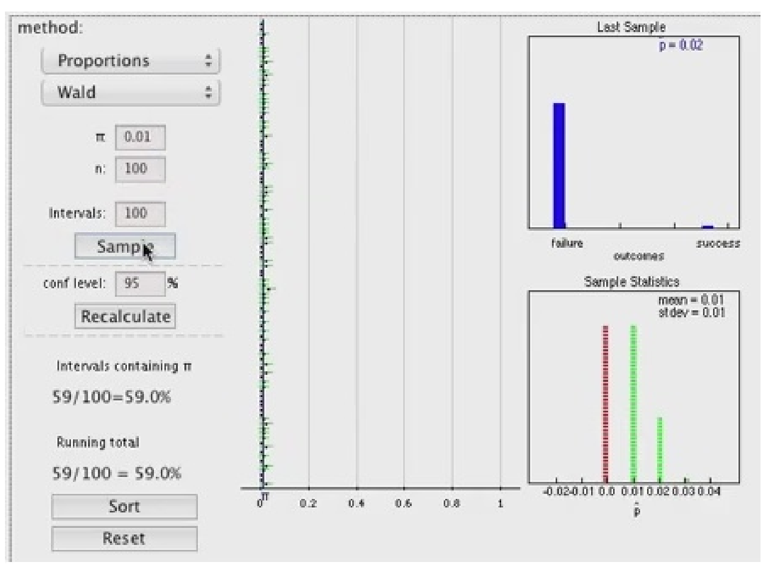

Figure 11: 100 confidence intervals based on samples with $n = 100$ and $p = 0.01$

If we repeat this procedure but this time take 100 samples each of size 1000 we find that 94% of the intervals contain the theoretical parameter used to simulate the data. This is shown in Figure 12.

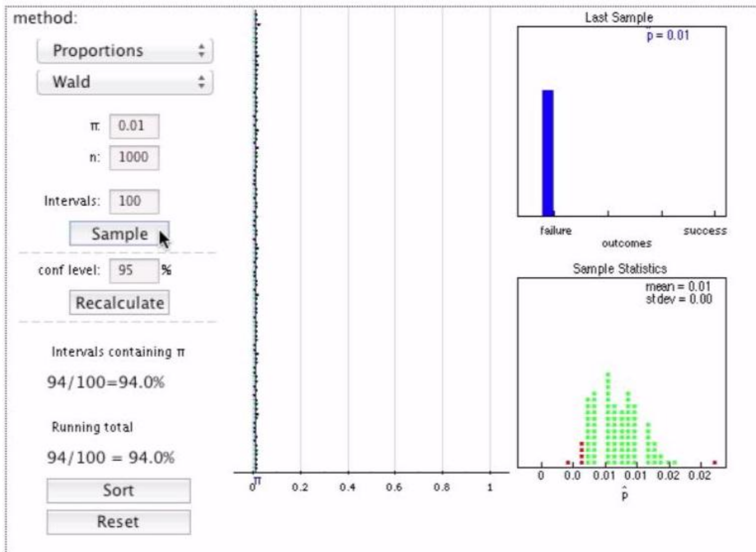

Figure 12: 100 confidence intervals based on samples with $n = 1000$ and $p = 0.01$

Now, let's look at the necessary conditions for the confidence intervals for means. When these conditions are met we are able to claim that the confidence interval contains the theoretical parameter the right percentage of time.

1. The n observations are independent
2. The sample size n is large enough so that $\overline{X}$ is approximately Normally distributed

We will focus on the second condition and use another applet to simulate data from a known distribution and compare our results under different conditions.

Example 4

The theoretical model we specify is a Normal distribution with $\mu = 500$ and $\sigma = 100$. The theoretical parameter we want to capture in the confidence interval μ is marked by the vertical blue line below the histogram in Figure 13.

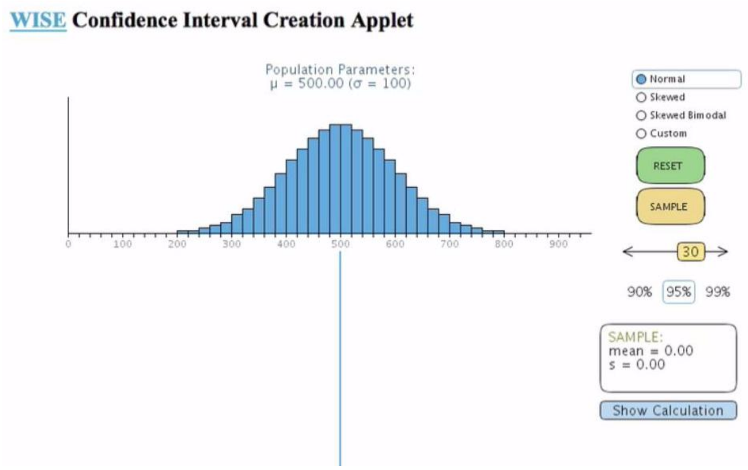

Figure 13: Theoretical distribution with $\mu = 500$ and $\sigma = 100$

We simulate a sample of size 10, shown in Figure 14. The sample mean $\overline{X} = 510$ and the sample standard deviation $s = 91.21$. Again we will be calculating 95% confidence intervals throughout these examples.

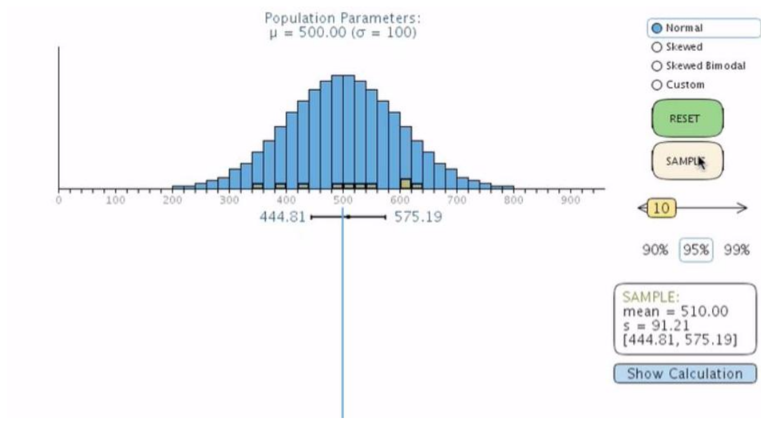

Figure 14: Confidence interval based on a sample of size $n = 10$ from a distribution with $\mu = 500$ and $\sigma = 100$

The critical value we need from the t distribution with 9 degrees of freedom for a 95% confidence interval is approximately 2.26. The confidence interval for μ based on this sample is

$$\begin{aligned} 95\% \text{ CI for } \mu &= \left[\overline{X} - t_{0.025,df=9}\frac{s}{\sqrt{n}}, \overline{X} + t_{0.025,df=9}\frac{s}{\sqrt{n}}\right] \\ &= \left[510.00 - 2.26(\frac{91.21}{\sqrt{10}}), 510.00 + 2.26(\frac{91.21}{\sqrt{10}})\right] \\ &= [441.81, 575.19] \end{aligned}$$

As can be seen in Figure 14, the confidence interval includes the theoretical parameter.

Shown in Figure 15 is when we simulate 100 samples of size 10. We find that 95 of the 95% confidence intervals captured the true mean just as we expected. The t distribution confidence interval for the mean works extremely well in terms of capturing the true mean as many times as it is supposed to, when the normal distribution is the actual distribution of the data.

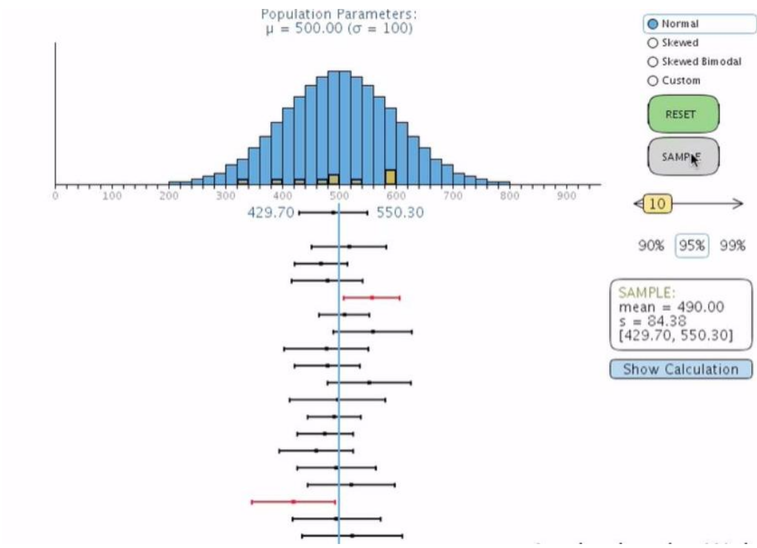

Figure 15: 100 confidence intervals based on samples of size $n = 10$ from a distribution with $\mu = 500$ and $\sigma = 100$

Example 5

Now we will simulate data from a skewed distribution instead of a Normal distribution and try to estimate the mean. The mean of the new theoretical model is $\mu = 223.1$ and $\sigma = 133.18$. This model is shown in Figure 16.

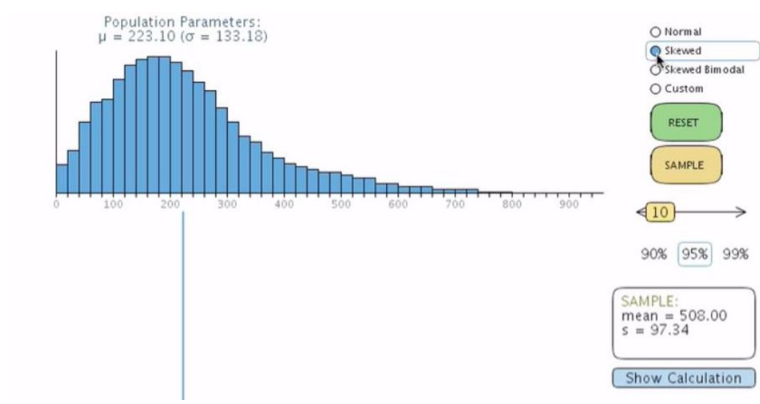

Figure 16: Skewed theoretical distribution with $\mu = 223.10$ and $\sigma = 133.18$

We take 100 samples of size 10 from this skewed distribution and find that we have 96% coverage since only 4 of the 95% confidence intervals failed to capture the theoretical mean, as shown in Figure 17.

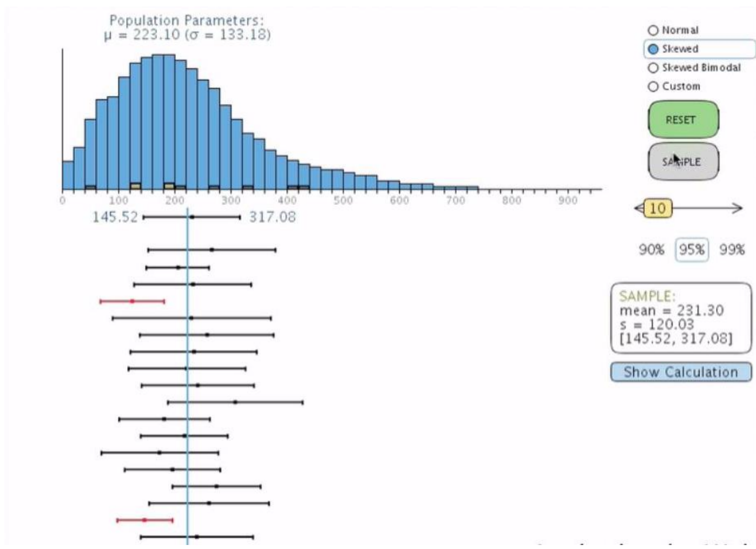

Figure 17: Confidence interval coverage of $\mu = 223.10$ based on a sample of size $n = 10$ from a skewed distribution

Even when the theoretical model from where our data is coming from is not Normal the t confidence interval procedure works extremely well at covering the mean the right percentage of times.

Example 6
For t confidence intervals for the mean to not work well, we need fairly extreme situations. Figure 18 shows one such example.

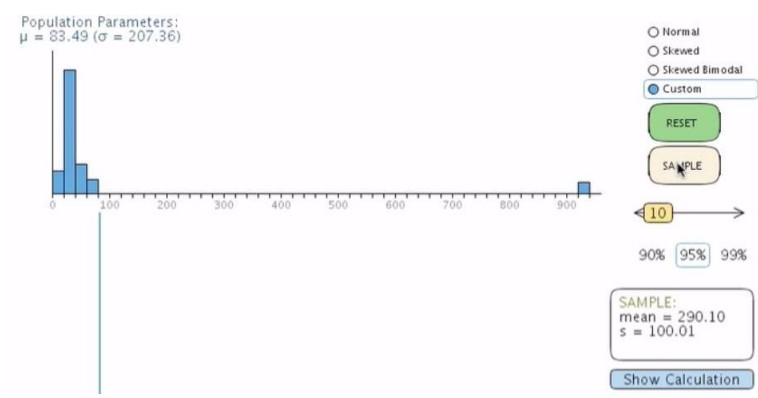

Figure 18: Extremely skewed theoretical distribution

This theoretical model puts most of the probability on values on the left, but also makes it possible to see some very large outliers. The theoretical mean is $\mu = 83.49$ and $\sigma = 207.36$. We take 20 samples of size 10 and find that 15 confidence intervals fail to capture the mean. (The coverage rate is $5/20 = 25\%$ instead of 95%). This is shown in Figure 19

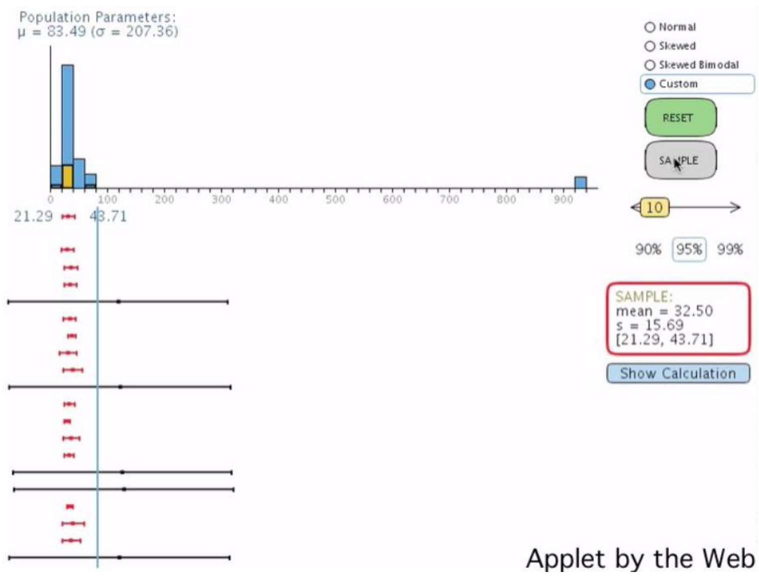

Figure 19: Confidence interval coverage of μ based on samples of size $n = 10$ from extremely skewed distribution

The Central Limit Theorem is an asymptotic result so it works better the larger the sample size n is. The sampling distribution $\overline{X}$ will be closer to Normally distributed for a larger sample size. To see the effect of increased sample size on the coverage of μ we will take 100 samples of size $n = 40$.

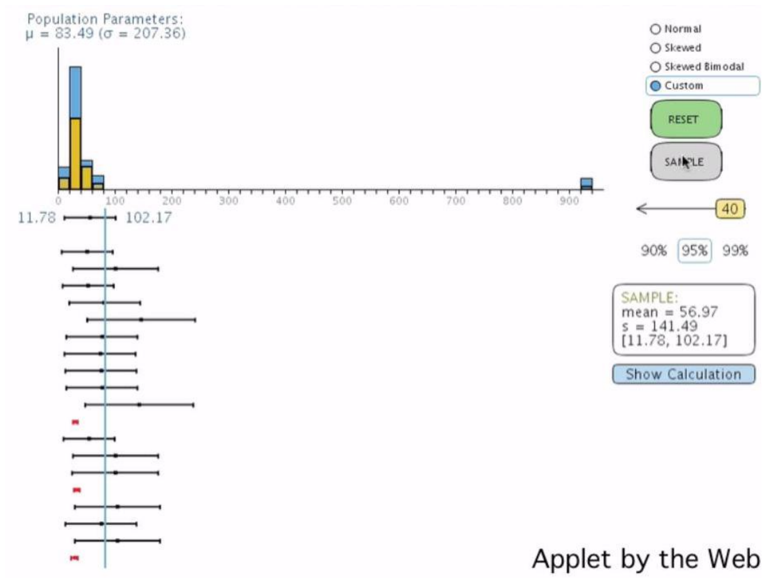

Figure 20: Confidence interval coverage of μ based on samples of size $n = 40$ from extremely skewed distribution

In the resulting 100 confidence intervals based on samples of size 40 (some shown in Figure 20), we miss the theoretical mean 10 times and obtain 90% coverage. Again, these are 95% confidence intervals, so the coverage is not quite what we claimed, but it is very close. This illustrates how well the t confidence intervals for the mean work, even with an extreme situation and a sample size of only 40.

To summarise, t confidence intervals for the mean are extremely robust. These confidence intervals match closely the coverage of the theoretical parameter claimed even when some of the necessary conditions with regard to sample size are not satisfied. Even though both confidence intervals for means and confidence intervals for proportions both rely on the Central Limit Theorem the latter tends to require much larger sample sizes. Remember that if the observations are not independent, neither of these procedures apply.