

Group Work - Week 7

1 In the wild, black cherry trees have a mean height of 82 feet with a standard deviation of 7.2 feet.

- (a) For samples drawn from this the wild population with sample sizes of 31, what is the mean and standard deviation of the sampling distribution of mean heights?

$$\mu_{\bar{x}} = \mu = 82$$
$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} = \frac{7.2}{\sqrt{31}} = 1.29$$

- (b) Suppose the built-in data set `trees` represented a sample of black cherry trees from a certain region. What is a 95% confidence interval for the population mean height based on this sample? What is a 90% confidence interval?

```
> x.bar <- mean(trees$Height)
> se <- sd(trees$Height) / sqrt(nrow(trees))
> z.95 <- 1.96
> z.90 <- 1.645
>
> ci.95 <- x.bar + c(-1, 1) * z.95 * se
> ci.95
[1] 73.75695 78.24305
>
> ci.90 <- x.bar + c(-1, 1) * z.90 * se
> ci.90
[1] 74.11744 77.88256
```

- (c) Do either of the confidence intervals contain the population mean height listed above? What conclusions might we draw from the answer?

Neither confidence interval contains 82 feet. Thus, it is likely that the sample was drawn from a population that is different than the general population of black cherry trees.

2 From the 2015 Youth Risk Behavior Survey found that the probability of a teen driver texting or emailing while driving is 0.404. Assume this is the population proportion for teens in the U.S.

- (a) For samples of U.S. teens with sample sizes of 50, what is the mean and standard deviation of the sampling distribution of the proportion of teens who text or email while driving?

$$p_{\hat{p}} = p = 0.404$$
$$\sigma_{\hat{p}} = \frac{\sigma}{\sqrt{n}} = \frac{pq}{\sqrt{n}} = \frac{0.404 \times 0.596}{\sqrt{50}} = 0.0341$$

- (b) Suppose a survey of 50 teenagers from a single high school find the proportion of drivers who text or email while driving is 0.38. What is a 95% confidence interval for the population proportion based on this sample? What is a 98% confidence interval?

```
> p.hat <- 0.38
> se <- p.hat * (1-p.hat) / sqrt(50)
> z.95 <- 1.96
> z.90 <- 1.645
>
> ci.95 <- p.hat + c(-1, 1) * z.95 * se
> ci.95
[1] 0.314695 0.445305
>
> ci.90 <- p.hat + c(-1, 1) * z.90 * se
> ci.90
[1] 0.3251905 0.4348095
```

- (c) Do either of the confidence intervals contain the population proportion listed above? What conclusions might we draw from the answer?

Both confidence intervals contain the population value of 0.404. Thus, there is no evidence that the students from this high school are different in terms of texting or emailing while driving from the general population.