

We have seen how substituting  $Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$  into the equation

$$P\left(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}\right) = 1 - \alpha \quad (1)$$

leads to a  $100(1 - \alpha)\%$  confidence interval for  $\mu$ :

$$\left(\bar{x} - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right)$$

This works for a random sample from a normally distributed population with known population standard deviation  $\sigma$ . The quantity  $\frac{\sigma}{\sqrt{n}}$  is known as the *standard error* of the (sample) mean.

1. Most of the time we do not actually know the population standard deviation. Fortunately, when  $n \geq 30$  it is reasonable to use the sample variance  $s$  as an estimate of  $\sigma$ . This gives the large sample (approximate) confidence interval  $\bar{x} \pm z_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$ . It happens that I have collected the heights of 71 adult humans, more or less at random. This sample has mean  $\bar{x} = 68.28$  and standard deviation of  $s = 3.60$ . Calculate a 90% confidence interval for the mean height of an adult human.

2. When dealing with proportions (specifically a population proportion  $\theta$  and a sample proportion  $\hat{\Theta}$ ), we can sub

$$Z \approx \frac{\hat{\Theta} - \theta}{\sqrt{\frac{\hat{\Theta}(1-\hat{\Theta})}{n}}}$$

into equation 1 to find a (approximate)  $100(1 - \alpha)\%$  confidence interval for  $\theta$ . This approximation is only reasonable for large sample sizes and values of  $\theta$  close to 0.5 (page 280 of the textbook has a better approximation).

A January YouGov poll of 388 likely South Carolina Democratic primary voters found that 60% would vote for Clinton. Calculate a 95% confidence interval for the true percentage of Clinton voters in the South Carolina Democratic primary (which was held on Saturday, February 27 and in which Clinton got 73.5% of the votes).

**3.** When sampling from a normally distributed population we can also use the fact that  $T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$  has a known distribution: Student's  $t$  distribution with  $n - 1$  degrees of freedom. This is true regardless of the sample size. (The distribution is named after William Gosset who worked for Guinness early in the 20th century using statistics to improve beer, and who published under the name Student).

a) Start with the equation  $P(-t_{\frac{\alpha}{2}} < T < t_{\frac{\alpha}{2}}) = 1 - \alpha$  and find a formula for the  $100(1 - \alpha)\%$  confidence interval for  $\mu$ .

b) In a “random” sample of 16 games, the Gonzaga men’s basketball team had an average score of  $\bar{x} = 81.875$  with a standard deviation of  $s = 10.78811$ . Calculate a 98% confidence interval for the mean score (assuming that scores are normally distributed).