

# Stats 2MB3, Tutorial 8

Mar 13<sup>th</sup>, 2015

# Confidence Interval for a Population Proportion

- Let  $p$  denote the proportion of “success” in a population. A random sample of  $n$  individuals is to be selected and  $X$  is the number of successes. The natural estimator of  $p$  is  $\hat{p} = X / n$

$E(\hat{p}) = p, \sigma_{\hat{p}} = \sqrt{p(1-p) / n}$ . Then we can get

$$P(-z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p) / n}} < z_{\alpha/2}) \approx 1 - \alpha$$

# One-sample t confidence interval

- Let  $\bar{x}$  and  $s$  be sample mean and sample standard deviation from a normal population with mean  $\mu$ . Then a  $100(1-\alpha)\%$  confidence interval for  $\mu$  is

$$\left( \bar{x} - t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2, n-1} \cdot \frac{s}{\sqrt{n}} \right)$$

# Ex 13, page 283

- An article reported that for a sample of 50 kitchens with gas cooking appliances monitored during a one-week period, the sample mean  $\text{CO}_2$  level was 654.16 (ppm) and the standard deviation was 164.43.
- a) Calculate and interpret a 95% (two-sided) confidence interval for true average  $\text{CO}_2$  level in the population of all homes from which the sample was selected.
- b) Suppose the investigators had made a rough guess of 175 for the value of  $s$  before collecting data. What sample size would be necessary to obtain an interval width of 50 (ppm) for a confidence level of 95%.

- a) For 95% confidence,  $\alpha=0.05$

$$\begin{aligned} & \left( \bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}} \right) \\ & = \left( 654.16 - 1.96 \cdot \frac{164.43}{\sqrt{50}}, 654.16 + 1.96 \cdot \frac{164.43}{\sqrt{50}} \right) \\ & = (608.58, 699.74) \end{aligned}$$

- b)  $s=175$ , and the width is  $\frac{2z_{\alpha/2}s}{\sqrt{n}}$ , then we have  $50 = \frac{2 \cdot 1.96 \cdot 175}{\sqrt{n}}$ . Hence,  $n = (13.72)^2 = 188.24$ , which should round up to 189.

## Ex 23, page 284

- The Pew Forum on Religion and Public Life reported on Dec. 9, 2009, that in a survey of 2003 American adults, 25% said they believed in astrology.
- a) Calculate and interpret a confidence interval at the 99% confidence level for the proportion of all adult Americans who believe in astrology.

- a) The confidence interval for population proportion is  $P(-z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{p(1-p)/n}} < z_{\alpha/2}) = 1 - \alpha$

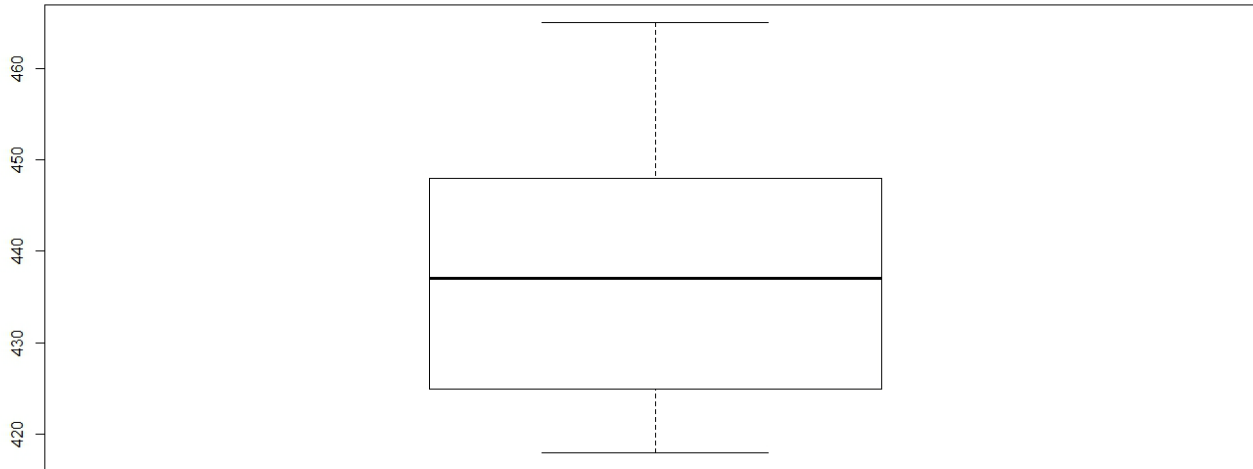
And from the page 280 on the textbook, the CI of  $p$  is approximately  $\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ . Since  $\hat{p} = 0.25$  and  $z_{\alpha/2} = z_{0.005} = 2.576$ , then 99% confidence interval is (0.225, 0.275).

# Ex 33, page 292

- An article contained the following observations on degree of polymerization for paper specimens for which viscosity times concentration fell in a certain middle range:
- 418, 421, 421, 422, 425, 427, 431, 434, 437, 439, 446, 447, 448, 453, 454, 463, 465.
- a) Construct a boxplot of the data and comment on any interesting features.
- b) Is it plausible that the given sample observations were selected from a normal distribution?
- c) Calculate a two-sided 95% confidence interval for true average degree of polymerization. Does the interval suggest that 440 is a plausible value for true average degree of polymerization? What about 450?



- a)&b)



- The boxplot has no outliers but skews to larger value. But since the median line is in the right middle of the upper and lower quantile area, it is reasonable to assume the observations are from a normal distribution.

- c)
- We construct the CI based on t distribution with degree of freedom  $n-1=16$ .
- The critical value for the 95% CI is  $t_{0.025,16}=2.120$
- Mean is 438.29, standard deviation is 15.14.
- The CI is given by

$$438.29 \pm 2.120\left(\frac{15.14}{\sqrt{17}}\right) = 438.29 \pm 7.785 = (430.51, 446.08)$$

Since 440 is inside the interval so it is a plausible value but 450 is not plausible because it lies outside the interval.