

STA 711 Homework 5

Due: Friday, February 21, 12:00pm (noon) on Canvas.

Instructions: Submit your work as a single PDF. For this assignment, you may include written work by scanning it and incorporating it into the PDF. Include all R code needed to reproduce your results in your submission.

Central limit theorem with estimated variance

The central limit theorem tells us that if Y_1, Y_2, \dots is a sequence of iid random variables, then

$$\frac{\sqrt{n}(\bar{Y}_n - \mu)}{\sigma} \xrightarrow{d} N(0, 1),$$

where $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$, $\mu = \mathbb{E}[Y_i]$, and $\sigma^2 = \text{Var}(Y_i)$. This limiting distribution is useful when we want to construct confidence intervals and tests for μ , but it requires us to know σ^2 . When σ^2 is unknown, we replace it with an estimate. Two possible estimators of σ^2 are:

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2 \\ s^2 &= \frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y}_n)^2\end{aligned}$$

5. Our goal is to show that using $\hat{\sigma}^2$ or s^2 in place of σ^2 does not change our limiting normal distribution. For the purposes of this problem, suppose that Y_1, Y_2, \dots is a sequence of iid random variables, and that the moment generating function of Y_i exists in a neighborhood of 0.

- (a) Show that $\hat{\sigma}^2 \xrightarrow{p} \sigma^2$ and $s^2 \xrightarrow{p} \sigma^2$.
- (b) Show that

$$\frac{\sqrt{n}(\bar{Y}_n - \mu)}{\hat{\sigma}} \xrightarrow{d} N(0, 1)$$

and

$$\frac{\sqrt{n}(\bar{Y}_n - \mu)}{s} \xrightarrow{d} N(0, 1).$$

- (c) If both $\hat{\sigma}^2$ and s^2 can be used as estimates of the population variance σ^2 , why do we have two estimates? The reason is that $\hat{\sigma}^2$ is a *biased* estimator of σ^2 (that is, $\mathbb{E}[\hat{\sigma}^2] \neq \sigma^2$), whereas s^2 is *unbiased* (that is, $\mathbb{E}[s^2] = \sigma^2$). Later in the course we will discuss the bias of estimators in more detail.

Calculate $\mathbb{E}[\hat{\sigma}^2]$ and $\mathbb{E}[s^2]$.