

Problems from 9th edition of *Probability and Statistical Inference* by Hogg, Tanis and Zimmerman:

- Section 7.1, Exercises 2, 4, 7
- Section 7.3, Exercises 1, 6, 8(a,b)

Additional Problems.

1. Sample Standard Deviation. Let X_1, X_2, \dots, X_n be independent samples from an underlying population with mean μ and variance σ^2 . We have seen that the sample mean $\bar{X} = (X_1 + X_2 + \dots + X_n)/n$ is an *unbiased estimator* for the population mean μ because

$$E[\bar{X}] = \mu.$$

The most obvious way to estimate the population variance σ^2 is to use the formula

$$V = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})^2.$$

Unfortunately, you will show that this estimator is **biased**.

- Explain why $E[X_i^2] = \mu^2 + \sigma^2$ for each i .
- Use the linearity of expectation together with part (a) and the fact that $\sum X_i = n\bar{X}$ to show that

$$\begin{aligned} E[V] &= \frac{1}{n} \left(E[\sum X_i^2] - 2E[\bar{X} \sum X_i] + E[n\bar{X}^2] \right) \\ &= \frac{1}{n} \left(n(\mu^2 + \sigma^2) - nE[\bar{X}^2] \right) \\ &= \mu^2 + \sigma^2 - E[\bar{X}^2] \end{aligned}$$

- Use the formula $\text{Var}(\bar{X}) = E[\bar{X}^2] - E[\bar{X}]^2$ to show that

$$E[\bar{X}^2] = \mu^2 + \sigma^2/n.$$

- Put everything together to show that

$$E[V] = \frac{n-1}{n} \cdot \sigma^2 \neq \sigma^2,$$

hence V is a **biased** estimator for σ^2 .

It follows that the weird formula

$$S^2 = \frac{n}{n-1} \cdot V = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$$

satisfies

$$E[S^2] = E \left[\frac{n}{n-1} \cdot V \right] = \frac{n}{n-1} \cdot E[V] = \frac{\cancel{n}}{\cancel{n}-1} \cdot \frac{\cancel{n}-1}{\cancel{n}} \cdot \sigma^2 = \sigma^2$$

and hence S^2 is an **unbiased** estimator for σ^2 . We call it the *sample variance* and we call its square root S the *sample standard deviation*. It is a sad fact that S is a **biased** estimator for σ but you will have to take more statistics courses if you want to learn about that.