8.3.1 The General Framework of Interval Estimation

Let X_1, X_2, X_3, $\ldots$, X_n be a random sample from a distribution with a parameter θ that is to be estimated. Our goal is to find two estimators for θ:

1. the low estimator, $\hat{\Theta}_l = \hat{\Theta}_l(X_1, X_2, \cdots, X_n)$, and
2. the high estimator, $\hat{\Theta}_h = \hat{\Theta}_h(X_1, X_2, \cdots, X_n)$.

The interval estimator is given by the interval $[\hat{\Theta}_l, \hat{\Theta}_h]$. The estimators $\hat{\Theta}_l$ and $\hat{\Theta}_h$ are chosen such that the probability that the interval $[\hat{\Theta}_l, \hat{\Theta}_h]$ includes θ is larger than $1-\alpha$. Here, $1-\alpha$ is said to be **confidence level**. We would like α to be small. Common values for α are 0.1, $.05$, and $.01$ which correspond to confidence levels 90%, 95%, and 99% respectively. Thus, when we are asked to find a 95% confidence interval for a parameter θ, we need to find $\hat{\Theta}_l$ and $\hat{\Theta}_h$ such that

$$P\left(\hat{\Theta}_l < \theta \text{ and } \hat{\Theta}_h > \theta\right) \geq 0.95$$

The above discussion will become clearer as we go through examples. Before doing that let's formally define interval estimation.

<u>Interval Estimation</u>

Let X_1, X_2, X_3, $\ldots$, X_n be a random sample from a distribution with a parameter θ that is to be estimated. An **interval estimator** with **confidence level** $1-\alpha$ consists of two estimators $\hat{\Theta}_l(X_1, X_2, \cdots, X_n)$ and $\hat{\Theta}_h(X_1, X_2, \cdots, X_n)$ such that

$$P\left(\hat{\Theta}_l \leq \theta \text{ and } \hat{\Theta}_h \geq \theta\right) \geq 1-\alpha,$$

for every possible value of θ. Equivalently, we say that $[\hat{\Theta}_l, \hat{\Theta}_h]$ is a $(1-\alpha)100\%$ **confidence interval** for θ.

Note that the condition

$$P\left(\hat{\Theta}_l \leq \theta \text{ and } \hat{\Theta}_h \geq \theta\right) \geq 1-\alpha$$

can be equivalently written as

$$P\left(\hat{\Theta}_l \leq \theta \leq \hat{\Theta}_h\right) \geq 1-\alpha, \quad \text{or} \quad P\left(\theta \in \left[\hat{\Theta}_l, \hat{\Theta}_h\right]\right) \geq 1-\alpha.$$

The randomness in these terms is due to $\hat{\Theta}_l$ and $\hat{\Theta}_h$, not θ. Here, θ is the unknown quantity which is assumed to be non-random (frequentist inference). On the other hand, $\hat{\Theta}_l$ and $\hat{\Theta}_h$ are random variables because they are functions of the observed random variables X_1, X_2, X_3, $\ldots$, X_n.
