Estimation, Bias and Variance

CS 3130 / ECE 3530: Probability and Statistics for Engineers

April 4, 2025

Parameters of a Distribution

All distributions we've discussed have parameters that fully describe their pdf or pmf.

- Normal: $X \sim N(\mu, \sigma^2)$ with parameters μ and σ^2
- Exponential: $X \sim \mathrm{Exp}(\lambda)$ with rate λ
- Bernoulli: $X \sim \mathrm{Ber}(p)$ with probability p

These parameters are *constants*.

Generic Notation

Notation: When making general statements, we use the Greek letter θ to denote a parameter.

Estimation of Distribution Parameters

If we assume our data follows a certain distribution (e.g., Normal via CLT), we need to estimate its parameters.

- Example: For Normal, estimate μ and σ^2
- Use sample mean $\bar{x}_n$ and variance s_n^2 as estimators

Definition: Estimator

Let $X_1, \ldots, X_n$ be iid from a distribution with parameter θ.
Estimator: $\hat{\theta} = T(X_1, \ldots, X_n)$

- "Hat" notation indicates estimation
- E.g., estimate mean μ with $\hat{\mu}$

Definition: Unbiased Estimator

An estimator $\hat{\theta}$ is **unbiased** if:

$$E[\hat{\theta}] = \theta$$

Bias:

$$\mathrm{bias}(\hat{\theta}) = E[\hat{\theta}] - \theta$$

Example: Estimating μ of a Normal

If $\hat{\mu} = \bar{X}_n$, then:

$$E[\bar{X}_n] = E[X_i] = \mu$$

Thus, $\bar{X}_n$ is an unbiased estimator of μ.

Example: Estimating σ^2 of a Normal

Estimator: $\hat{\sigma}^2 = S_n^2$
We use:

$$\mathrm{Var}(X) = E[X^2] - E[X]^2$$

And compute:

$$E[X_i^2] = \sigma^2 + \mu^2$$
$$E[\bar{X}_n^2] = \frac{\sigma^2}{n} + \mu^2$$
$$E[X_i \bar{X}_n] = \frac{\sigma^2}{n} + \mu^2$$

Putting It Together: Expected Sample Variance

$$E\left[S_n^2\right] = \sigma^2$$

If we used n instead of $n-1$, we would get a biased estimator:

$$E\left[\frac{1}{n}\sum(X_i - \bar{X}_n)^2\right] = \frac{n-1}{n}\sigma^2$$

Example: Estimating p for Bernoulli

Let $X_i \sim \mathrm{Ber}(p)$. Then:

$$E[X_i] = p \Rightarrow E[\bar{X}_n] = p$$

Conclusion: $\bar{X}_n$ is an unbiased estimator for p.

Estimator Efficiency

Suppose $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased.
Efficiency: $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$ if:

$$\text{Var}(\hat{\theta}_1) < \text{Var}(\hat{\theta}_2)$$