3. Random Variables and Probability Distributions - MSE

Lai Hui Shan
m5281022

Explanation of Mean and Variance as Optimal Choices for MSE

In the context of mean squared error (MSE), we consider a random variable Y with mean μ_Y and variance σ_Y^2. When choosing a constant c to approximate Y, we define the MSE as the expected squared deviation of Y from c:

$$\mathrm{MSE} = \mathbb{E}\left\{(Y-c)^2\right\}.$$

Our goal is to find the value of c that minimizes this MSE. By minimizing this expected squared deviation, we aim to find the "best" constant approximation of Y in terms of its expected distance from Y.

Why the Mean Minimizes MSE

1. **Expanding the MSE Expression**:
   To find the optimal c, we expand the expression for MSE:

   $$\mathrm{MSE} = \mathbb{E}[(Y-c)^2] = \mathbb{E}[Y^2 - 2Yc + c^2].$$

   Since $\mathbb{E}[Y] = \mu_Y$, we can rewrite this as:

   $$\mathrm{MSE} = \mathbb{E}[Y^2] - 2c\mu_Y + c^2.$$

2. **Differentiating with Respect to c**:
   To minimize MSE, we differentiate with respect to c and set the result to zero:

   $$\frac{d}{dc}\mathrm{MSE} = -2\mu_Y + 2c = 0.$$

   Solving this equation gives $c = \mu_Y$. Therefore, the optimal choice of c to minimize MSE is the mean of Y, μ_Y.
3. **Resulting MSE as Variance**:
   When $c = \mu_Y$, the MSE becomes:

   $$\mathrm{MSE}(\mu_Y) = \mathbb{E}[(Y-\mu_Y)^2] = \sigma_Y^2,$$

   which is exactly the variance of Y.

Thus, we see that:

- **The mean μ_Y is the best constant choice** c in terms of minimizing MSE.
- **The minimum MSE achievable is equal to the variance** σ_Y^2, representing the inherent variability of Y around its mean.

Summary

The mean and variance of Y are optimal in the context of MSE because:

- The mean μ_Y minimizes the expected squared deviation, making it the most accurate constant approximation of Y.
- The variance σ_Y^2 quantifies the "best possible error" when approximating Y by its mean.

This result aligns with the general properties of expectation and variance for random variables and reinforces the interpretation of variance as a measure of dispersion around the mean.