Estimators

Bias Example #1

Let $Y_i \mid \theta, \sigma^2 \overset{iid}{\sim} \text{Normal}(\theta, \sigma^2)$

Let $\hat{\theta}_e = \bar{Y} = \frac{1}{n} \sum_{i=1}^{n} Y_i$

$$\text{Bias}(\hat{\theta}_e \mid \theta = \theta_0) = \mathbb{E}[\hat{\theta}_e \mid \theta_0] - \theta_0$$

$$= \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}[Y_i \mid \theta_0] - \theta_0$$

$$= \frac{1}{n} \cdot n\theta_0 - \theta_0$$

$$= 0$$

The ~~sample~~ mean $\hat{\theta}_e$ is an unbiased estimator of θ.

Bias Example #2

Let $X_i \overset{iid}{\sim} \text{exponential}(\theta)$

Let $\hat{\theta} = \frac{1}{\bar{X}}$ where $\bar{X} = \frac{1}{n} \Sigma X_i$

Is $\hat{\theta}$ biased?

$$\mathbb{E}\, \hat{\theta} \mid \theta = \theta_0 = \mathbb{E}\, \frac{1}{\bar{X}} \mid \theta_0 = \mathbb{E}\, \frac{n}{\Sigma X_i} \mid \theta_0$$

Notice $p(\Sigma X_i \mid \theta) \sim \text{gamma}(n, \theta)$ (sum of exponentials w/ identical rate follow a gamma)

Let $y = \Sigma X_i$.

$$\Rightarrow \mathbb{E}\, \frac{n}{y} \mid \theta_0 = \int_0^\infty \frac{n}{y} \frac{\theta^n}{\Gamma(n)} y^{n-1} e^{-\theta y}\, dy$$

$$= \frac{n\theta^n}{\Gamma(n)} \underbrace{\int_0^\infty y^{(n-1)-1} e^{-\theta y}\, dy}_{\text{kernel of gamma}(n-1,\, \theta)}$$

using kernel trick to integrate

$$= \frac{n\theta_0^n}{\Gamma(n)} \cdot \frac{\Gamma(n-1)}{\theta_0^{n-1}} = \boxed{\frac{n}{n-1}\theta_0}$$

$\boxed{\hat{\theta} = \frac{1}{\bar{X}} \text{ is a biased estimator of } \theta}$

VARIANCE EXAMPLE

AGAIN $Y_i | \theta, \sigma^2 \overset{iid}{\sim} \text{normal}(\theta, \sigma^2)$; $\hat{\theta}_e = \bar{y}$

$$\text{var}(\hat{\theta}_e | \theta_0) = \frac{1}{n^2} \sum_{i=1}^{n} \text{var}(y_i | \theta_0)$$

$$= \boxed{\frac{\sigma^2}{n}}$$

MSE EXAMPLE:

$$\text{MSE}(\hat{\theta}_e | \theta_0) = \underbrace{\frac{\sigma^2}{n}}_{\text{variance}} + \underbrace{0^2}_{\text{bias}^2}$$

EXERCISE #3

Let $\hat{\theta}_b = w\bar{y} + (1-w)\mu_0$

$$\text{MSE}(\hat{\theta}_b | \theta_0) = E[(\hat{\theta}_b - \theta_0)^2 | \theta_0]$$

WE COULD COMPUTE MSE = VARIANCE + BIAS², BUT HERE IS A TRICK TO COMPUTE DIRECTLY:

TRICK: $(\hat{\theta}_b - \theta_0)^2 = (\hat{\theta}_b - \underbrace{(w\theta_0 + (1-w)\theta_0)}_{\theta_0})^2$

So we have:

$E[(\hat{\theta}_b - (w\theta_0 + (1-w)\theta_0))^2 | \theta_0]$. Plug in $\hat{\theta}_b$ & collect terms:

$$E[(w(\bar{y} - \theta_0) + (1-w)(\mu_0 - \theta_0))^2 | \theta_0]$$

$$= E\, w^2(\bar{y} - \theta_0)^2 + 2w(1-w)(\mu_0 - \theta_0)\underbrace{E(\bar{y} - \theta_0)}_{0} + \underbrace{E(1-w)^2(\mu_0 - \theta_0)^2}_{\text{constant}}$$

$$=$$

EXERCISE #3 CONTINUED

$$= w^2 \underbrace{\mathbb{E}(\bar{y} - \theta_0)^2 | \theta_0}_{var(\bar{y} | \theta_0)} + (1-w)^2 (\mu_0 - \theta_0)^2$$

$MSE(\hat{\theta}_w | \theta_0) < MSE(\hat{\theta}_e | \theta_0)$ if

$$w^2 \, var(\bar{y} | \theta_0) + (1-w)^2 (\mu_0 - \theta_0)^2 < var(\bar{y} | \theta_0)$$

Rearranging:

$$(\mu_0 - \theta_0)^2 < \frac{1+w}{1-w} var(\bar{y} | \theta_0)$$

$\uparrow$

i.e. if is small enough.

In words: If our prior guess μ_0 is "close" to θ_0, then our Bayesian estimator will have smaller MSE.