We found previously that if $X_1, \ldots, X_n \overset{iid}{\sim} \text{Normal}(\mu, \sigma^2)$ (4)

the maximum likelihood estimators are

$$\hat{\mu}^{MLE} = \bar{X}$$

$$\hat{\sigma}^{2\,MLE} = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

Previous results: $\bar{X} \sim \text{Normal}(\mu, \frac{\sigma^2}{n})$

$$\frac{\sum_{i=1}^{n} (X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$

- What are the bias, variance, and MSE of $\hat{\mu}^{MLE}$?
- If $Y \sim \chi^2_d$ then $E(Y) = d$ and $Var(Y) = 2d$.

1) • What are the bias, variance, and MSE of $\hat{\sigma}^{2\,MLE}$?

2) ~~• If~~ I define a new estimator $\hat{\sigma}^{2\,NEW} = c \cdot \hat{\sigma}^{2\,MLE}$,

(a) What value of c should I pick so that $\hat{\sigma}^{2\,NEW}$ is unbiased?

(b) What value of c should I pick so that $\hat{\sigma}^{2\,NEW}$ has as small a MSE as possible?

(c) Does any of this matter if n is large?

5

1) $E\left(n \cdot \frac{\hat{\sigma}^{2\,MLE}}{\sigma^2}\right) = n-1$

$\Rightarrow \frac{n}{\sigma^2} E\left(\hat{\sigma}^{2\,MLE}\right) = n-1$

$\Rightarrow E\left(\hat{\sigma}^{2\,MLE}\right) = \frac{n-1}{n} \cdot \sigma^2$

$Var\left(n \cdot \frac{\hat{\sigma}^{2\,MLE}}{\sigma^2}\right) = (n-1)2$

$\Rightarrow \left(\frac{n}{\sigma^2}\right)^2 \cdot Var\left(\hat{\sigma}^{MLE}\right) = 2n-2$

$\Rightarrow Var\left(\hat{\sigma}^{MLE}\right) = \frac{2(\sigma^2)^2(n-1)}{n^2}$

$\Rightarrow E\left(\frac{n}{n-1}\hat{\sigma}^{2\,MLE}\right) = \sigma^2$

If $c = \frac{n}{n-1}$, $\hat{\sigma}^{2\,NEW}$ is unbiased. this is $\frac{n}{n-1} \cdot \frac{1}{n} \Sigma (x_i - \bar{x})^2$

$= \frac{1}{n-1} \Sigma (x_i - \bar{x})^2$

$Bias\left(\hat{\sigma}^{2\,MLE}\right) = E\left[\hat{\sigma}^{2\,MLE}\right] - \sigma^2$

$= \frac{n-1}{n}\sigma^2 - \sigma^2$

$= -\frac{1}{n}\sigma^2$

$MSE\left(\hat{\sigma}^{2\,MLE}\right) = \left(\frac{-1}{n}\sigma^2\right)^2 + \frac{2(\sigma^2)^2(n-1)}{n^2}$

$= \frac{\sigma^4 + 2\sigma^4(n-1)}{n^2}$

$E\left(c \cdot \hat{\sigma}^{2\,MLE}\right) = c \cdot E\left(\hat{\sigma}^{2\,MLE}\right) = c \cdot \frac{n-1}{n}\sigma^2$

$\Rightarrow Bias\left(\hat{\sigma}^{2\,NEW}\right) = c \cdot \frac{n-1}{n}\sigma^2 - \sigma^2 = \sigma^2\left(c \cdot \frac{n-1}{n} - 1\right)$

$Var\left(c \cdot \hat{\sigma}^{2\,MLE}\right) = c^2 \cdot Var\left(\hat{\sigma}^{2\,MLE}\right) = c^2 \cdot \frac{2(\sigma^2)^2(n-1)}{n^2}$

$\Rightarrow MSE\left(c \cdot \hat{\sigma}^{2\,MLE}\right) = \sigma^4\left(c \cdot \frac{n-1}{n} - 1\right)^2 + c^2 \cdot \frac{2\sigma^4(n-1)}{n^2}$

$\frac{d}{dc} MSE = 2\sigma^4\left(c \cdot \frac{n-1}{n} - 1\right) \cdot \frac{n-1}{n} + 2c\frac{2\sigma^4(n-1)}{n^2} = 0$

$\Rightarrow 2c + \frac{c(n-1)}{n} = 1 \quad \Rightarrow 2c + nc - c = n$

$\Rightarrow c(n+1) = n$

$\Rightarrow c = n/(n+1)$

$$\frac{d^2}{dc^2} MSE = 2\sigma^4\left(\frac{n-1}{n}\right)^2 + 4\sigma^4\left(\frac{n-1}{n}\right) > 0 \checkmark$$

Among all estimators of the form

$c \cdot \hat{\sigma}^2_{MLE}$, the estimator

$$\hat{\sigma}^2 = \frac{n}{n+1} \cdot \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n+1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

has lowest MSE

This is our first example of a shrinkage estimator:
we "shrank" the estimate of σ^2 toward 0
(divide by $n+1$ instead of $n-1$)
Introduces bias, but reduces variance enough to
result in an overall gain in terms of MSE.