ECE 313: Lecture 36
Sample mean and variance of a data set, unbiased estimators (Ch 4.8, Example 4.8.7)

Reminder: $\mathrm{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])]$

a number that summarizes the relationship X & Y

(compare w. $f_{X,Y}(u, v)$ for $u, v \in \mathbb{R}$)

Allow computation $\sim \langle X, Y \rangle$

linear w.r.t each arg

E.g. $\mathrm{Cov}(aX + bY, cW + dZ) = ac\,\mathrm{Cov}(X, W) + \dots$

$\Rightarrow$ Special case $\mathrm{Cov}(X, X) = \mathrm{Var}(X)$

Application: Suppose X is a r.v. with $E[X] = \mu$, $\mu = \int u \, f_X(u) \, du$

$X \rightarrow f_X \rightarrow$ $Var[X] = \sigma^2$

$\qquad \| $

$E[(X - E[X])^2]$

In practice, we don't know f_X; but we can independently sample $X_1, X_2, \dots, X_n$ from f_X

$\Rightarrow$ Sample mean $\hat{X} = \frac{1}{n} \sum_{i=1}^{n} X_i \longrightarrow \mu$

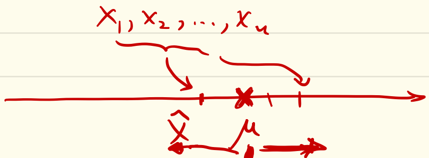
Sample variance $\hat{\sigma}^2 = \frac{1}{\cancel{n}\,(n-1)} \sum_{i=1}^{n} (X_i - \hat{X})^2 \longrightarrow \sigma$

(*) Unbiased estimator

$\hat{X}$ is unbiased estimator of $\mu \iff E[\hat{X}] = \mu$

Check: $E[\hat{X}] = E[\frac{1}{n} \sum_{i=1}^{n} X_i] = \frac{1}{n} \sum_{i=1}^{n} E[X_i]$ linear of E

$= \frac{1}{n} \cdot n\mu = \mu$



⊛ How spread is $\hat{X}$ = average $(X_1, \dots, X_n)$ around μ?

$$\mathrm{Var}[\hat{X}] = E\Big[\big(\hat{X} - \underset{\mu}{E[\hat{X}]}\big)^2\Big] = E[\hat{X}^2 - 2\mu\hat{X} + \mu^2]$$

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

$$= E\Big[\Big(\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\Big)^2\Big]$$

$$= E\Big[\Big(\frac{1}{n}\sum X_i\Big)^2\Big] - \mu^2$$

$$= \Big(\frac{1}{n}\Big)^2 E\big[(X_1 + X_2 + \dots + X_n)^2\big] - \mu^2$$

Try diff. route:

$$\mathrm{Var}[\hat{X}] = \mathrm{Var}\Big[\frac{1}{n}\sum_{i=1}^{n} X_i\Big]$$

But

$$\mathrm{Var}(aX + bY) = \mathrm{Cov}(aX + bY, aX + bY) = a^2\mathrm{Cov}(X, X) + 2\mathrm{Cov}(X,Y) + b^2\mathrm{Cov}(Y,Y)$$

$\mathrm{Cov}(X,Y) \to 0$ if X & Y are independent

$$= a^2\,\mathrm{Cov}(X,X) + b^2\,\mathrm{Cov}(Y,Y)$$

$$\mathrm{Var}[\hat{X}] = \frac{1}{n^2}\big(\mathrm{Var}(X_1) + \dots + \mathrm{Var}(X_n)\big) = \frac{n\sigma^2}{n^2} = \boxed{\frac{\sigma^2}{n}}$$

Finally, we will show

$$E[\hat{\sigma}] = E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \hat{X})^2\right] = \sigma^2$$

again, linearity of $E[\,]$ Symmetry $X_1, \dots, X_n$

$$= \frac{1}{(n-1)}\sum_{i=1}^{n} E\left[(X_i - \hat{X})^2\right] = \frac{n}{n-1} E\left[(X_1 - \hat{X})^2\right]$$

Now consider $E[(X_1 - \hat{X})^2] = Var[(X_1 - \hat{X})] + \cancel{\underbrace{(E[X_1 - \hat{X}])^2}_{= (E[X_1] - E[\hat{X}])}}^{\,0}$

$$(\underbrace{E[X_1]}_{\mu} - \underbrace{E[\hat{X}]}_{\mu})$$

$$Var[X_1 - \hat{X}] = Var\left[X_1 - \frac{1}{n}\sum_{i=1}^{n} X_i\right]$$

$$= Var\left[\frac{n-1}{n}X_1 - \frac{1}{n}X_2 - \frac{1}{n}X_3 - \dots - \frac{1}{n}X_n\right]$$

$X_1, X_2, \dots, X_n$ independent

$$= \left(\frac{n-1}{n}\right)^2 Var[X_1] + \left(\frac{1}{n}\right)^2 Var(X_2) + \dots + \left(\frac{1}{n}\right)^2 Var(X_n)$$

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