# EE306: Signals and Systems II

## Lecture 9

### Basics of parameter estimation

* Consider a random variable X with unknown parameter $\theta$. Assume we have a random sample $\underline{X}_n = (X_1, X_2, \dots, X_n)$ having iid versions of X.

A parameter estimator of $\theta$ is given by:

$$\hat{\theta}(\underline{X}_n) = g(X_1, X_2, \dots, X_n) = \hat{\theta}_n .$$

* Evaluating an estimator:

→ Estimator bias

$$B(\hat{\theta}) = E(\hat{\theta}) - \theta .$$

Thus, for unbiased estimator $E(\hat{\theta}) = \theta$.

→ Estimator consistency

An estimator is consistent if

$$\lim_{n\to\infty} P\left[\,|\hat{\theta}-\theta| > \varepsilon\,\right] = 0 .$$

That is, $\hat{\theta}$ converges to $\theta$ in probability.

(gets better with increasing sample size n).

Q Is an unbiased estimator always consistent?

Q Is a consistent estimator always unbiased?

Ex1 Sample mean: Suppose $X_1, X_2, \ldots, X_n$ are iid RVs with mean $E(X) = \mu$ and variance $V(X) = \sigma^2$.
Let $M_n = \frac{1}{n}\sum_{i=1}^{n} X_i$. Find $E(M_n)$ and $V(M_n)$.
Is $M_n$ an unbiased (consistent) estimator of $\mu$?

Sol. $E(M_n) = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}(n\mu) = \mu$.

$$V(M_n) = V\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\sum_{i=1}^{n} V(X_i) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}.$$

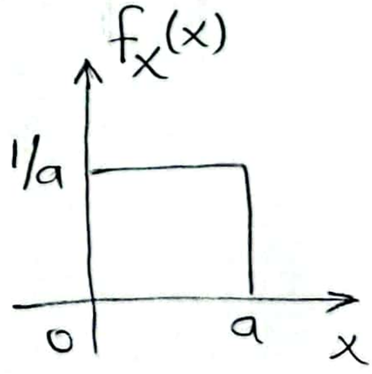
$E(M_n) = \mu \Rightarrow M_n$ is unbiased.

$\lim_{n\to\infty} M_n = \mu \Rightarrow M_n$ is consistent.

Ex2 Suppose $X$ is uniformly distributed in $[0, a]$, where $a$ is unknown. Let $X_1, X_2, \ldots, X_n$ be iid versions of $X$. We choose the estimator
$Y_n = \max\{X_1, X_2, \ldots, X_n\}$ for $a$.

(a) Find $F_{Y_n}(y)$. (b) Find $E(Y_n)$ and $V(Y_n)$.

(c) Evaluate the estimator $Y_n = \hat{a}$.




Sol. The PDF $f_X(x)$ is shown.

(a) $F_{Y_n}(y) = P(Y_n \leq y)$

$= P(\max\{X_1, X_2, \ldots, X_n\} \leq y)$

$= P(X_1 \leq y)\, P(X_2 \leq y) \cdots P(X_n \leq y)$.

Scanned with CamScanner

Therefore, $F_{Y_n}(y) = [P(X \leq y)]^n = [F_X(y)]^n$.

$$F_X(y) = \int_{x=0}^{y} \frac{1}{a}\,dx = \frac{y}{a}, \quad 0 \leq y \leq a.$$

Thus, $F_{Y_n}(y) = \left(\frac{y}{a}\right)^n, \; 0 \leq y \leq a, \quad F_{Y_n}(y) = \begin{cases} 0, & y<0 \\ 1, & y>a \end{cases}$.

(b) $$f_{Y_n}(y) = \frac{dF_{Y_n}(y)}{dy} = \frac{n}{a^n} y^{n-1}, \quad 0 \leq y \leq a.$$

Therefore, $E(Y_n) = \int_y y f_{Y_n}(y)\,dy$

$$= \int_{y=0}^{a} y \frac{n}{a^n} y^{n-1}\,dy = \int_{y=0}^{a} \frac{n}{a^n} y^n\,dy = \frac{n}{a^n} \frac{y^{n+1}}{(n+1)}\Bigg|_{y=0}^{a}$$

$$= \frac{n}{n+1} \frac{a^{n+1}}{a^n} = \frac{n}{n+1} a.$$

$$E(Y_n^2) = \int_y y^2 f_{Y_n}(y)\,dy = \int_{y=0}^{a} y^2 \frac{n}{a^n} y^{n-1}\,dy = \int_{y=0}^{a} \frac{n}{a^n} y^{n+1}\,dy$$

$$= \frac{n}{a^n} \frac{y^{n+2}}{(n+2)}\Bigg|_{y=0}^{a} = \frac{n}{n+2} a^2.$$

Thus, $V(Y_n) = E(Y_n^2) - [E(Y_n)]^2$

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

$$\Rightarrow V(Y_n) = \frac{n(n+1)^2 - n^2(n+2)}{(n+2)(n+1)^2} a^2 = \frac{\cancel{n^3} + \cancel{2n^2} + n - \cancel{n^3} - \cancel{2n^2}}{(n+2)(n+1)^2} a^2$$

$$= \frac{n}{(n+2)(n+1)^2} a^2.$$

(c) The estimator does have non-zero bias since $E(Y_n) = E(\hat{a}) \neq a$.

Instead, $B(Y_n) = E(Y_n) - a = \frac{n}{n+1}a - a$

$\Rightarrow B(Y_n) = \frac{-1}{n+1}a.$

However, $\lim_{n\to\infty} E(Y_n) = \lim_{n\to\infty} \frac{1}{1+1/n}a = a.$

Moreover, $\lim_{n\to\infty} V(Y_n) = \lim_{n\to\infty} \frac{O(n)}{O(n^3)}a^2 = 0.$

This means as $n\to\infty$, $Y_n$ converges to a constant $= a. \Rightarrow$ consistent and good estimator.

Minimum MSE linear estimator [RV estimation]

* We want to estimate $X$ via $g(Y) = \hat{X}$, where $Y$ is the observation.

$\rightarrow$ The MSE: $e = E[(X - g(Y))^2]$. (eval. metric).

Case 1 $g(Y) = a$ (const.) $\Rightarrow e = E[(X-a)^2]$

$\Rightarrow e = E(X^2) - 2aE(X) + a^2 \Rightarrow \frac{\partial e}{\partial a} = -2E(X) + 2a$

$\Rightarrow \frac{\partial e}{\partial a} = 0\Big|_{a=a^*} \Rightarrow a^* = E(X) \Longrightarrow \hat{X} = E(X).$ $Z$

Case 2 $g(Y) = aY + b \Rightarrow e = E[(\overbrace{X - aY}^{} - b)^2]$

$\Rightarrow b^* = E(Z) = E(X) - aE(Y)$

$\Rightarrow e = E[((X - E(X)) - a(Y - E(Y)))^2]$ Orthogonality $\downarrow$

$\Rightarrow \frac{\partial e}{\partial a} = -2E[\underbrace{((X - E(X)) - a(Y - E(Y)))}(Y - E(Y))] = 0\Big|_{a=a^*}$

$\Rightarrow a^* = \frac{CV(X,Y)}{V(Y)} \Rightarrow \hat{X} = g(Y) = \rho_{X,Y}\frac{\sigma_X}{\sigma_Y}(Y - E(Y)) + E(X).$