ECE 313 / MATH 362 PROBABILITY WITH ENGINEERING APPLICATIONS

LECTURE 36 : MINIMUM MEAN SQUARE ERROR ESTIMATION PART 1

- TOPICS TO COVER (BASED ON CH 4.9)
  - → CONSTANT ESTIMATORS
  - → UNCONSTRAINED ESTIMATORS

→ CONSTANT ESTIMATORS

Y : A RANDOM VARIABLE WITH KNOWN PROBABILITY DISTRIBUTION F_Y

QUESTION : SUPPOSE Y IS NOT OBSERVED BUT WE WISH TO ESTIMATE Y.

WHAT DO YOU THINK COULD BE AN INTUITIVELY 'GOOD' CHOICE OF ESTIMATOR FOR THE UNOBSERVED Y?

ANSWER : $\hat{Y} = \underbrace{EY}_{\text{CONSTANT}}$

A NOTATION FOR AN ESTIMATOR

WE NEXT SHOW THAT THIS INTUITIVE CHOICE IS INDEED A GOOD CHOICE IN THE SENSE THAT IT MINIMIZES SOMETHING CALLED THE MEAN SQUARE ERROR (MSE).

LET δ BE A CONSTANT THAT WE WISH TO USE TO ESTIMATE THE RV Y.

THEN

ERROR IN ESTIMATION $= Y - \delta$

SQUARE(D) ERROR IN ESTIMATION $= (Y - \delta)^2$

TO TREAT UNDER- AND OVER-ESTIMATION THE SAME

MEAN SQUARE(D) ERROR IN ESTIMATION $= E(Y-\theta)^2$

EXPECTED / AVERAGE

WE WANT TO CHOOSE θ SUCH THAT THE MSE IS MIMIMIZED, I.E., SOLVE

$$\theta^* = \underset{\theta}{\arg\min}\, E(Y-\theta)^2$$

HOW?

→ DIFFERENTIATE $E(Y-\theta)^2$ AND

→ SET IT EQUAL TO ZERO

CONSIDER

$$E(Y-\theta)^2 = E(Y^2 - 2Y\theta + \theta^2)$$

$$= E(Y^2) - 2\theta E(Y) + \theta^2$$

? LINEARITY OF EXPECTATION

$$\Rightarrow \frac{d}{d\theta} E(Y-\theta)^2 = -2E(Y) + 2\theta = 0$$

$$\Rightarrow \theta^* = E(Y)$$

THIS MINIMIZES THE MSE

MSE AT $\theta^* = E(Y)$: $E(Y-\theta^*)^2 = E(Y-E(Y))^2 = Var(Y)$

⇒ MIN MSE FOR ESTIMATING Y USING A CONSTAT IS SAME AS ITS VARIANCE.

ALTERNATIVELY, CONSIDER

$$E(Y-\delta)^2 = E(Y - E(Y) + E(Y) - \delta)^2$$

$$= E(Y - E(Y))^2 + E(\underbrace{E(Y) - \delta}_{\text{CONSTANT}})^2$$

$$+ \; 2E((Y - E(Y))(\underbrace{E(Y) - \delta}_{\text{CONSTANT}}))$$

$$= Var(Y) + (E(Y) - \delta)^2 + 2(E(Y) - \delta)) \, E((Y - E(Y))$$

$$= Var(Y) + (E(Y) - \delta)^2 + 2(E(Y) - \delta)) \, (\underbrace{E(Y) - E(Y)}_{0})$$

$$= \underbrace{Var(Y)}_{\text{CONSTAT WRT } \delta} + \underbrace{(E(Y) - \delta)^2}_{\text{ALWAYS NON-NEGATIVE}}$$

HENCE TO MINIMIZE THE SUM WITH RESPECT TO δ

SET $\delta = E(Y)$

MINIMUM MSE CONSTANT ESTIMATOR OF Y

→ UNCONSTRAINED ESTIMATORS

Y : A RANDOM VARIABLE WITH KNOWN PROBABILITY DISTRIBUTION F_Y

X : A RANDOM VARIABLE WITH KNOWN PROBABILITY DISTRIBUTION F_X

QUESTION : SUPPOSE Y IS NOT OBSERVED BUT X IS AND WE WISH TO ESTIMATE Y BASED ON AN OBSERVATION FROM X.

WHAT DO YOU THINK COULD BE A INTUITIVELY 'GOOD' CHOICE OF ESTIMATOR FOR THE UNOBSERVED Y BASED ON AN OBSERVATION FROM X ?

ANSWER : RECALL YOUR ANSWER WHEN NO OBSERVATION ON X WAS THERE.

$$\hat{Y} = EY$$

NOW YOU HAVE AN OBSERVATION ON X. WHAT CAN YOU DO?

↝ UPDATE THE CONSTANT ESTIMATOR $\hat{Y} = EY$ WITH X AS:

$$\hat{Y} \mid X = \underbrace{E(Y \mid X)}_{\text{A FUNCTION OF X}} \quad : \text{CONDITIONAL EXPECTATION OF Y GIVEN X}$$

WE NEXT SHOW THAT THIS INTUITIVE CHOICE IS INDEED A GOOD CHOICE IN THE SENSE THAT IT MINIMIZES THE MEAN SQUARE ERROR (MSE).

LET $g(X)$ BE A FUNCTION OF X THAT WE WISH TO USE TO ESTIMATE THE RV Y. THEN

ERROR IN ESTIMATION $= Y - g(X)$

SQUARE(D) ERROR IN ESTIMATION $= (Y - g(X))^2$

MEAN SQUARE(D) ERROR IN ESTIMATION $= E(Y - g(X))^2$

EXPECTED / AVERAGE

WE WANT TO CHOOSE G SUCH THAT THE MSE IS MIMIMIZED, I.E., SOLVE

$$g^*(X) = \arg\min_{g(X)} E(Y - g(X))^2$$

WHY? THE DISCRETE CASE FOLLOWS SIMILARLY.

HOW?

LOTUS WHEN X AND Y CONTINUOUS

CONSIDER $E(Y - g(X))^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (v - g(u))^2 \underbrace{f_{X,Y}(u,v)} \, dv \, du$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (v - g(u))^2 \overbrace{f_{Y|X}(v|u) f_X(u)} \, dv \, du$$

$$= \int_{-\infty}^{\infty} \underbrace{\left(\int_{-\infty}^{\infty} (v - g(u))^2 f_{Y|X}(v|u) f_X(u) \, dv \right)}_{E((Y - g(X))^2 \mid X = u)} du$$

SAME FORM AS $E(Y - b)^2$ FOR A FIXED $X = u$

$\therefore$ $E((Y - g(X))^2 \mid X = u)$ IS MINIMIZED AT $g(X) = E(Y \mid X = u)$ FOR A FIXED $X = u$

$\therefore$ $g(X)$ THAT MINIMIZES THE MSE FOR AN ARBITRARY X IS $E(Y \mid X)$

$\therefore$ $g^*(X) = E(Y \mid X)$

MINIMUM MSE AT $g^*(X) = E(Y \mid X)$:

$$E(Y - g^*(X))^2 = E(Y - E(Y \mid X))^2$$

$$= E(Y^2) - E[(E(Y \mid X))^2]$$

HOW? VERY SIMILAR TO $Var(Y) = EY^2 - (EY)^2$

APPLIED USING THE CONDITIONAL DISTRIBUTION OF Y GIVEN X