Random Variable :

A RV is function from Ω to R

Discrete RV

- Probability Mass Function (PMF)

$P_X(x) \triangleq Pr(X = x)$

Continuous RV

- Prob. density function (PDF)

i.) $f_X(x) \geq 0$

ii) $\int_{-\infty} f_X(x) = 1$

iii) $Pr\{a \leq X \leq b\} = \int_a^b f_X(x)dx$

$$E[g(X)] \triangleq \sum_x g(x) P_X(x)$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$$

$$\text{Mean} \triangleq E[X]$$

$$\text{Variance} \triangleq E[(X - E[X])^2]$$

$$\boxed{E[aX + bY] = aE[X] + bE[Y]}$$

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

Bias - Variance Tradeoff :

$y = f(x)$

f : unknown function

data x is random .

n ~~data~~ training example:

$$\underbrace{(x_1, y_1), (x_2, y_2) \cdots (x_n, y_n)}$$

$$\downarrow$$

$$\hat{F}$$

So that for new data 'x', we say our estimated value is

$$\hat{Y} = \hat{F}(x)$$

$(X_1, Y_1), \ldots (X_2, Y_2) \longrightarrow \hat{F}$ is a random function

$\downarrow$

Random

Mean Square Error (MSE) $\triangleq E\left[(\hat{Y} - y)^2\right]$

$$= E[\hat{Y}^2] - 2y E[\hat{Y}] + y^2 + (E[\hat{Y}])^2 - (E[\hat{Y}$$

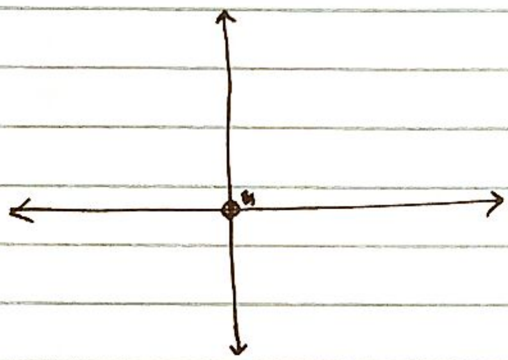
$$= (y - E[\hat{Y}])^2 + E[\hat{Y}^2] - (E[\hat{Y}])^2$$

$$= (f(x) - E[\hat{F}(x)])^2 + E\left[(\hat{F}(x))^2\right] - (E[\hat{F}(x)$$

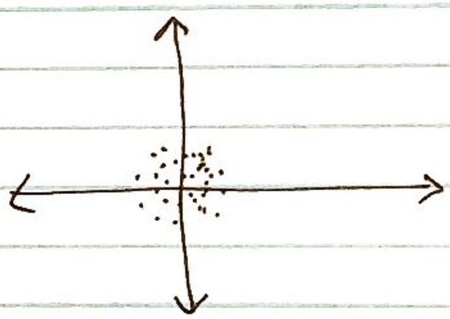
$$\text{Bias} \triangleq f(x) - E[\hat{F}(x)]$$

$$\text{variance} \triangleq E[(\hat{F}(x))^2] - (E[\hat{F}(x)])^2$$

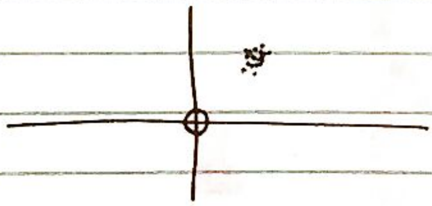
$$\text{Error} = \text{Bias}^2 + \text{Var}.$$



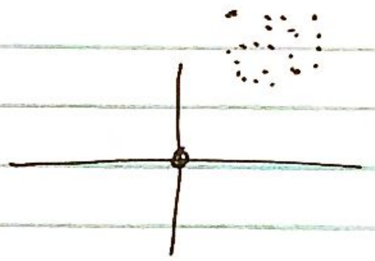
Low Bias, Low Variance



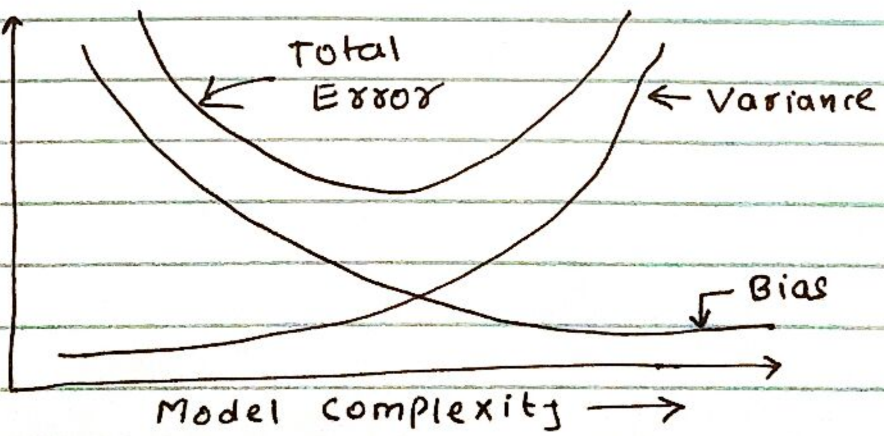
Low Bias, high variance



High Bias, Low Variance



High Bias, high variance.



Lagrange Multiplier :

optimization problem:

Find max or min of $f(x_1, \dots x_n)$ subject to constraint $g(x_1, \dots x_n) = 0$

Define
$L(x_1, \dots x_n, \lambda) \triangleq f(x_1, \dots x_n) - \lambda g(x_1, \dots x_n)$

Compute & solve
$$\frac{\partial L}{\partial x_1} = 0, \quad \frac{\partial L}{\partial x_2} = 0, \quad \dots \quad \frac{\partial L}{\partial x_n} = 0$$

$$\frac{\partial L}{\partial \lambda} = 0$$

Check the solution of above $x_1^*, x_2^*, \dots x_n^*$ for max or min.

Example :

$\cancel{f(x_1, x_2)}$ $f(x, y) = xy$ constraint to
$x + y = 20$

$L(x, y, \lambda) = xy - \lambda(x+y-20)$

$$\frac{\partial L}{\partial x} = y - \lambda, \quad \frac{\partial L}{\partial y} = x - \lambda, \quad \frac{\partial L}{\partial \lambda} = x + y - 20$$

$= 0 \Rightarrow \lambda = y$ $\qquad$ $= 0 \Rightarrow \lambda = x$ $\qquad$ $= 0 \Rightarrow$

$2\lambda = 20$

$\lambda = 10$

$x^* = y^* = \lambda = 10$

$f(x^*, y^*) = 100$