(1) Machine learning
- framing problems

(2) Loss Functions
- Squared error loss.
- 0-1 Loss.

(3) Conditional Expectation
- Binary variables

(4) Regression as M.L.

$$f(x) = \beta_0 + \beta_1 x$$

or $\text{logit}(f(x)) = \beta_0 + \beta_1 x$.

(5) Naive Bayes.

$$f(x) = P(Y=1 \mid X=x)$$

(6) Confusion Matrix

truth (Y)

pred. ($\hat{Y}$)	0	1
0	TN	FN
1	FP	TP

$$\boxed{(Y-\hat{Y})^2}$$

$$\hat{Y} \in \mathbb{R}$$

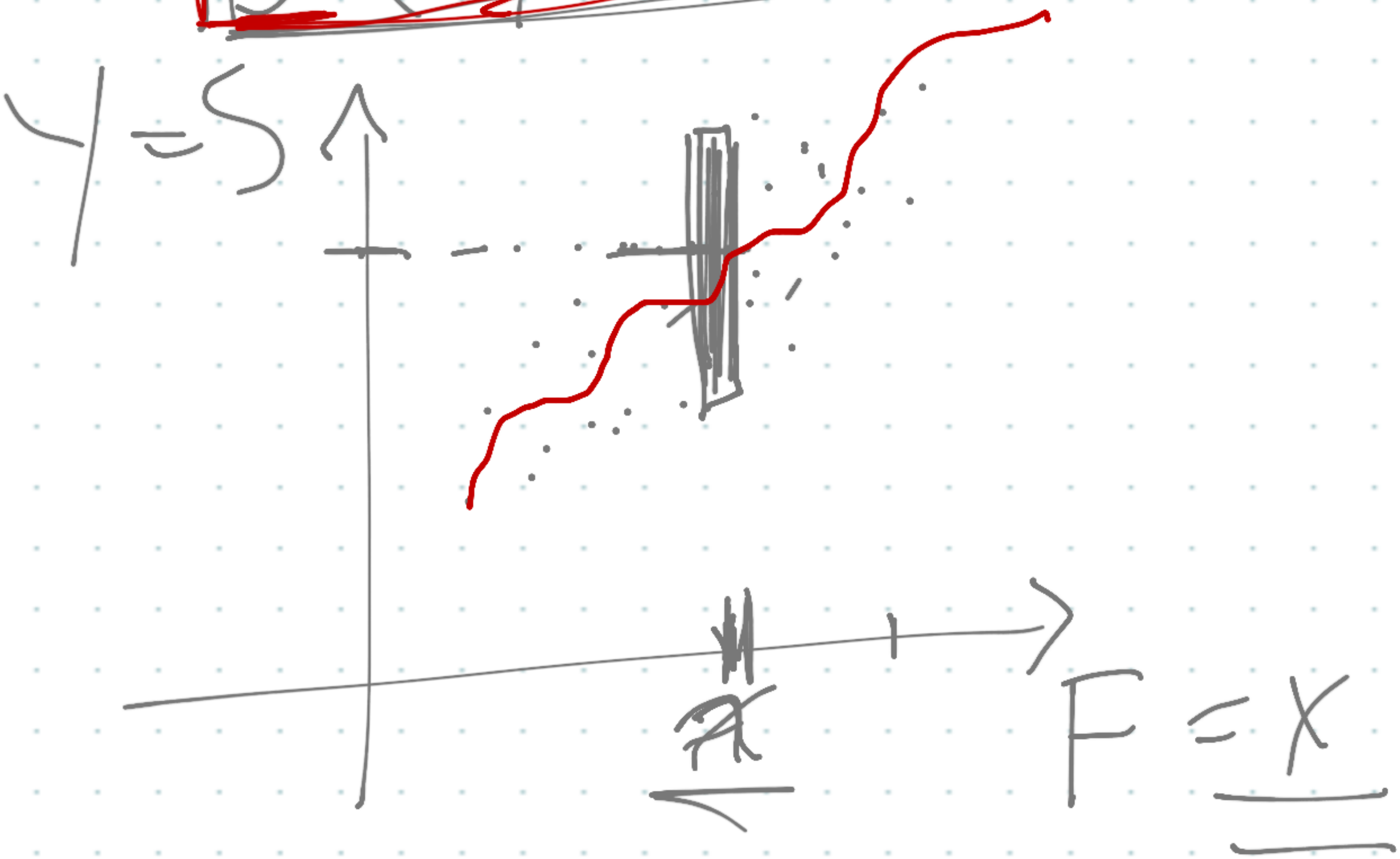
$$(\hat{Y} - Y)^2$$

$$\left(Y - (\beta_0 + \beta_1 X)\right)^2$$

0-1 Loss

$$w \cdot \mathbb{1}\{Y = \hat{Y}\}$$

$$\boxed{E(Y \mid X = x) = f(x)}$$



$$f(x) = \beta_0 + \beta_1 x$$

$$\mathbb{E}(\underset{\sim}{Y} \mid \underline{X = x})$$

$$Y = \underset{\sim}{0} \text{ or } \underset{\sim}{1}$$

$$= \mathbb{P}(Y = 1 \mid X = x)$$

Back to Regresion

$$f(x) = \mathbb{E}(Y \mid X = x)$$

Linear:

$$\boxed{f(x) = \underline{\beta_0} + \underline{\beta_1} x}$$

Logistic Regression

$\text{logit}(P) = \log\left(\boxed{\dfrac{P}{1-P}}\right)$

log-odds

P = prob of success,
prob of label 1

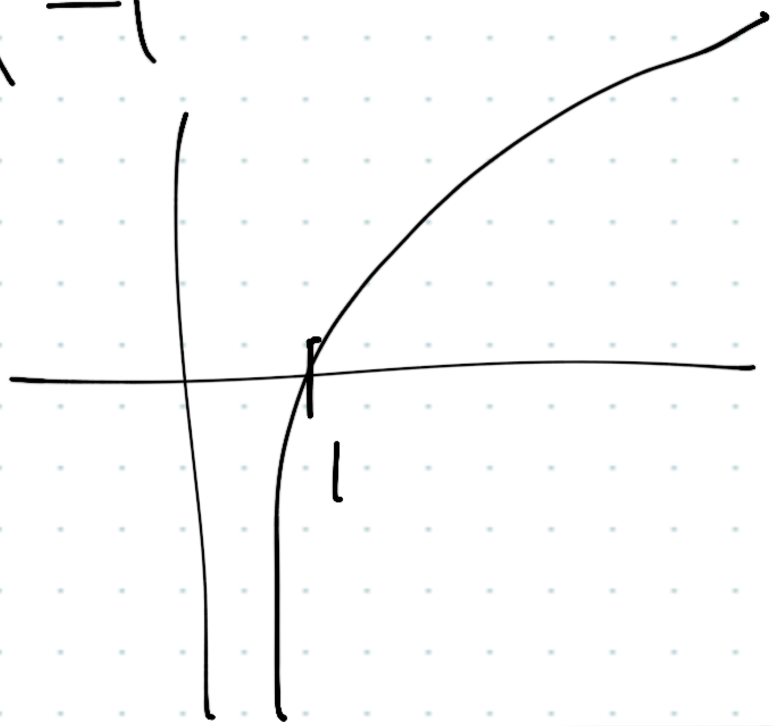
<u>if</u> P small, $1-P \approx 1$
$\simeq 0$

$\dfrac{P}{1-P} \simeq 0$

<u>if</u> $P \simeq 1$ $\quad 1-P \simeq 0$

$\dfrac{P}{1-P}$ large $\approx \infty$

$$\frac{p}{1-p} \in (0, \infty)$$



$$\boxed{\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x}$$

$$p = \frac{\exp(\boxed{\beta_0} + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)}$$

$$\boxed{f(x)} = \exp(\beta_0 + \beta_1 x) \,/\, (1 + \dots)$$

Naive Bayes

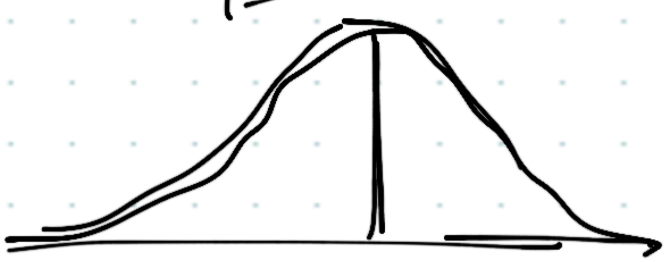
$Y = 0$ or 1

M F

$$\boxed{f(x)} = E(Y \mid X = x)$$

$$= P(Y = 1 \mid X = x)$$

$$= \frac{\boxed{P(X = x \mid Y = 1)}\,\boxed{\pi}}{(1-\pi)\,P(X = x \mid Y = 0) + \pi\, P(X = x \mid Y = 1)}$$



$$\boxed{\pi = P(Y = 1)}$$

$$1 - \pi = P(Y = 0)$$