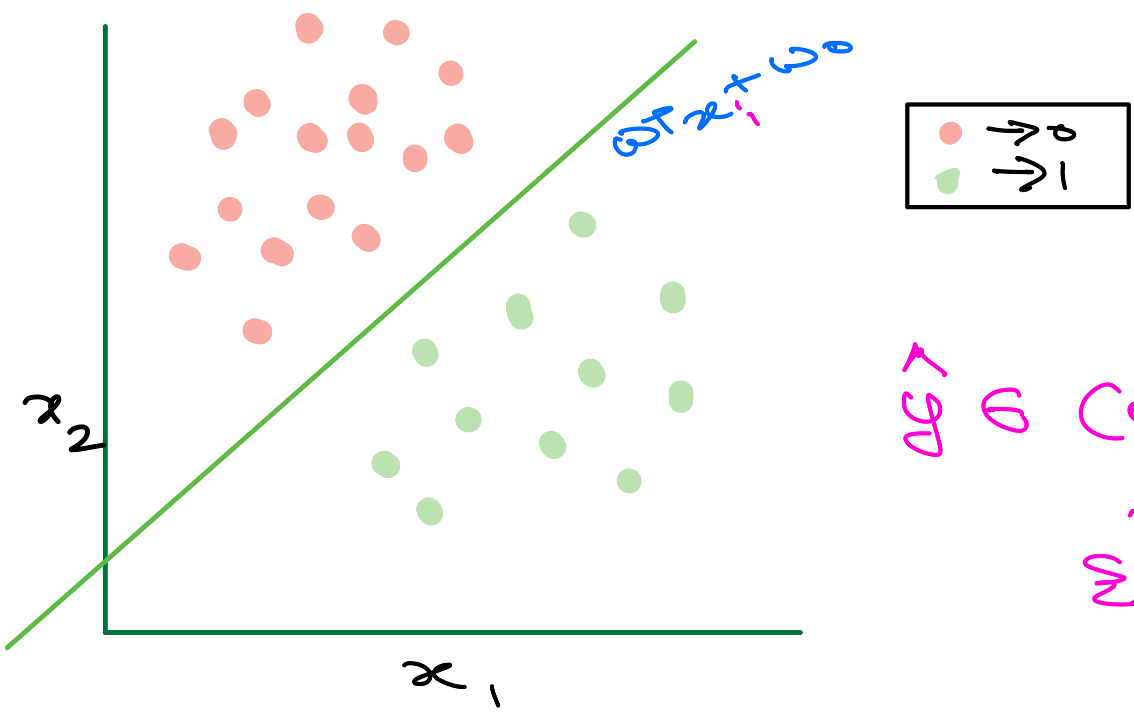
- ⇒ Recap
- ⇒ Regularization in Logistic Regression
- ⇒ Odds interpretation of Hyperplane
- ⇒ Impact of outliers
- ⇒ Multi-Class Classification

RECAP

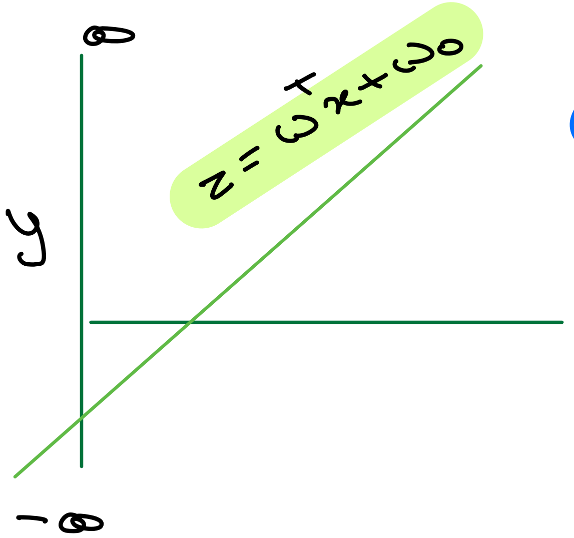


$\hat{y} \in (0, 1)$

$\downarrow$ threshold

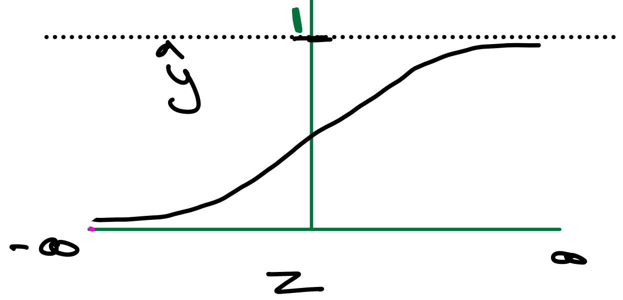
$\{0, 1\}$

$$z = w^T x + w_0 \Rightarrow \sigma(z) \to (0, 1)$$



$\sigma(z) \Rightarrow$

$\frac{1}{1+e^{-z}} \Rightarrow \frac{1}{1+e^{-(w^T x + w_0)}}$



$\hat{y}_i = P(y_i = 1)/x_i \Rightarrow 1 - P \leftarrow P_{y_i = 0}$

$$\boxed{\text{Loss}^i \Rightarrow -y^{(i)} \cdot \log(\hat{y}^{(i)}) - (1 - y^{(i)}) \times \log(1 - \hat{y}^{(i)})}$$

Negative Log Likelihood

Binary Crossentropy

Optional: Derivative of Loss

$$loss \Rightarrow \frac{1}{n}\sum_{i=1}^{n} \underbrace{-y^{(i)}\cdot \log(\hat{y}^{(i)})}_{A} - \underbrace{(1-y^{(i)})\times \log(1-\hat{y}^{(i)})}_{B}$$

$$\frac{\partial L_A}{\partial w_j} \Rightarrow -y(1-\hat{y})\times x_j$$

$$\frac{\partial L_B}{\partial w_j} \Rightarrow (1-y)\times \hat{y}\times x_j$$

Ⓐ $-y(1-\hat{y})\times x_j$
$+$
Ⓑ $(1-y)\times\hat{y}\times x_j$

$(-y+\hat{y}\times y)\times x_j$
$+$
$(\hat{y}-y\times\hat{y})\times x_j$

$x_j(-y+\cancel{y\times\hat{y}}$
$+\hat{y}-\cancel{y\hat{y}})$

$x_j(\hat{y}-y)$

$$\text{Total Loss} \Rightarrow \frac{\partial L_A}{\partial w_j}+\frac{\partial L_B}{\partial w_j}$$

$$\boxed{\frac{\partial L}{\partial w_j} \Rightarrow (\hat{y}-y)\times x_j}$$

Accuracy

0.25

threshold=0.5

x_1 x_2 ... x_d	y	$\hat{y}$	Pred	
x_1	1	0.8	1	✓
x_2	0	0.3	0	✓
x_3	1	0.2	0	✗
	0	0.51	1	✗
x_4	1	0.90	1	✓

$\frac{3}{5}$

Hyperparameter

$$\boxed{C = 1/\lambda}$$

$$-y^{(i)} \cdot \log(\hat{y}^{(i)}) - (1-y^{(i)}) \times \log(1-\hat{y}^{(i)}) + C(w_j)^2$$

Liner Regression

Higher Values of λ $\uparrow$ Underfitting

Logistic Regression

Higher Values of $C \rightarrow 1/\lambda$ $\uparrow$ Overfitting

If you want to increase Regularization strength ; reduce C

Interpretation of 'z' in terms of log Odd

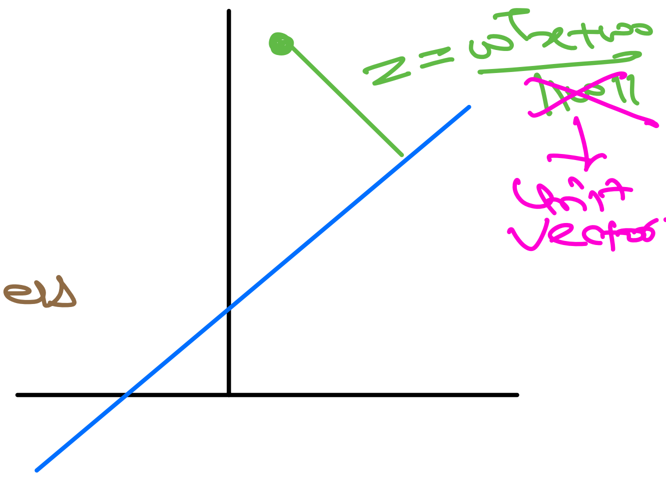
Horse H : 4 : 1

5 races

$$\text{Odds} \Rightarrow \frac{\text{Probability of Success}}{\text{Probability of Failures}}$$

$$P_H(\text{win}) \Rightarrow \frac{4}{5} \Rightarrow \frac{\frac{4}{\cancel{5}}}{\frac{1}{\cancel{5}}} \Rightarrow 4:1$$

$$P_H(\text{Lose}) \Rightarrow \frac{1}{5}$$



$$P(y=1 \mid x) \Rightarrow P_{success}$$

$$P_{failure} \Rightarrow 1 - P_{success}$$

$$\text{Odds} = \frac{P}{1-P}$$

The logistic regression model predicts:

4 options

Active Duration(Most preferred: 30 seconds)

Appears for	30 Secs

A	Probabilities ✓
B	Class labels
C	Continuous values ✓ $(0, 1)$
D	Ordinal values

$$P = \sigma(z) = \frac{1}{1+e^{-z}} = \sigma(w^T \underline{x} + w_0)$$

$$= \frac{1}{1+\frac{1}{e^z}}$$

$$P = \frac{1}{\frac{e^z+1}{e^z}} = \frac{e^z}{1+e^z}$$

$$1-P = 1 - \frac{e^z}{1+e^z} = \frac{1+\cancel{e^z}-\cancel{e^z}}{1+e^z} = \frac{1}{1+e^z}$$

$$\text{Odds} = \frac{P}{1-P} = \frac{\frac{e^z}{\cancel{1+e^z}}}{\frac{1}{\cancel{1+e^z}}} = e^z$$

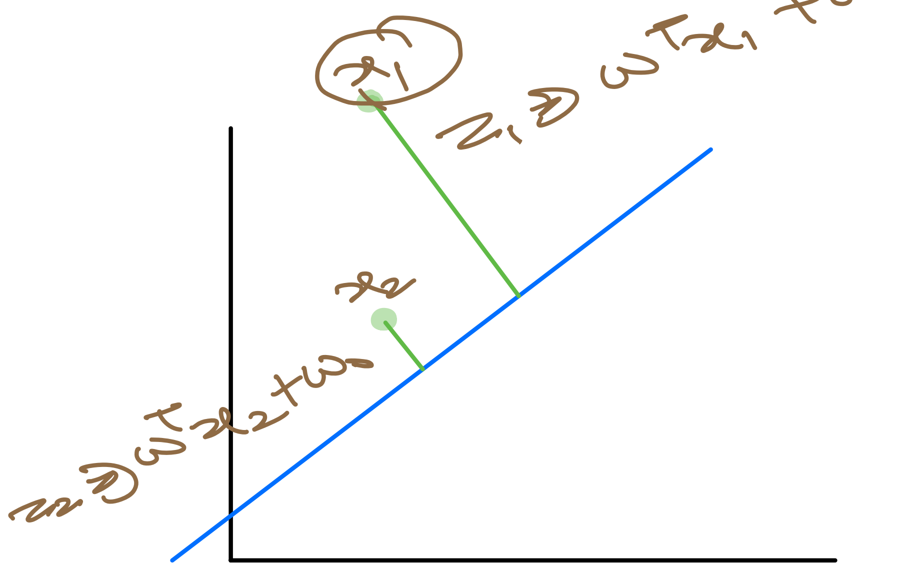
Log Odds Interpretation

$$Odd = \frac{P}{1-P} = e^{Z}$$

$$\downarrow \log_e$$

$$\log_e(odds) = \log_e\left(\frac{P}{1-P}\right) \Rightarrow Z \underbrace{\log_e e}_{1}$$

$$\log(odd) \Rightarrow Z$$



if $d_1 > d_2$

$\log odds\ x_1 > \log odd\ x_2$

$odd\ x_1 > odds\ x_2$

Linear Regression

$$y \Rightarrow w^T x + w_0 \quad \text{(Direct Relation)}$$

Logistic Regression

$$\log\left(\frac{P}{1-P}\right) \Rightarrow w^T x + w_0 \quad \text{(Non Linear Relation)}$$

Sigmoid

How are log odds transformed into probabilities in logistic regression?

4 options

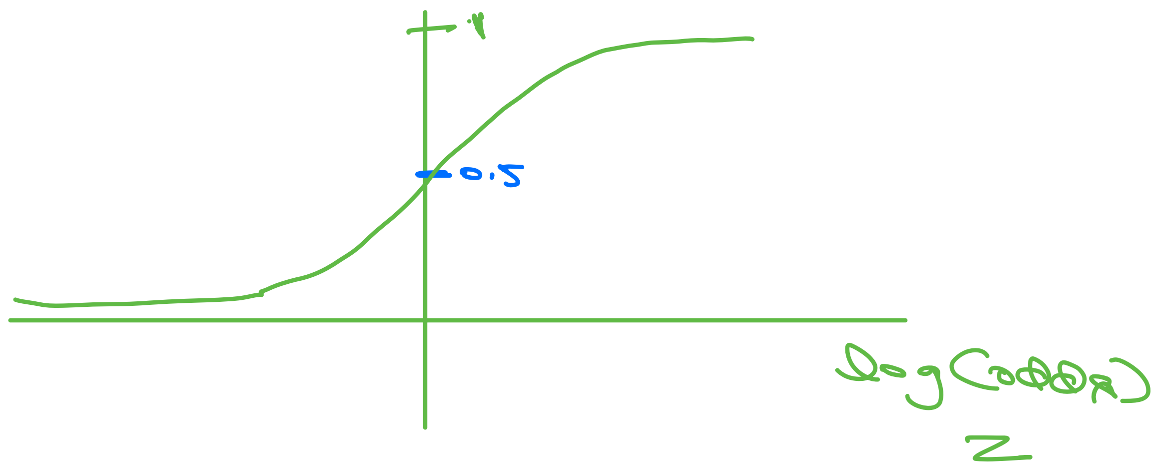
Active Duration(Most preferred: 30 seconds)

Appears for	30 Secs

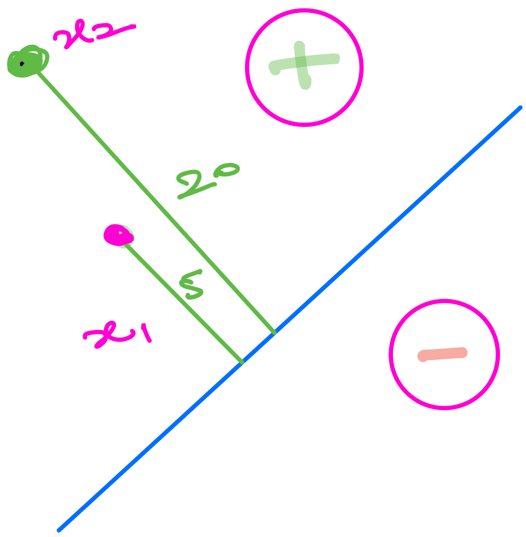
A	By applying the sigmoid function
B	By taking the exponential function
C	By dividing by the odds ratio

$$P \quad \frac{1}{1+e^{-w^T x + w_0}}$$

Impact of Outliers on Logistic Regression



Case-1 : when the outlier is on correct side



$$\text{Loss}^{i} \Rightarrow -y^{(i)} \cdot \log(\hat{y}^{(i)}) - (1-y^{(i)}) \times \log(1-\hat{y}^{(i)})$$

$$\quad A \qquad\qquad B$$

$$-1 \times A \quad - \cancel{(1-1)} \times B \qquad y_i = 1$$

$$0$$

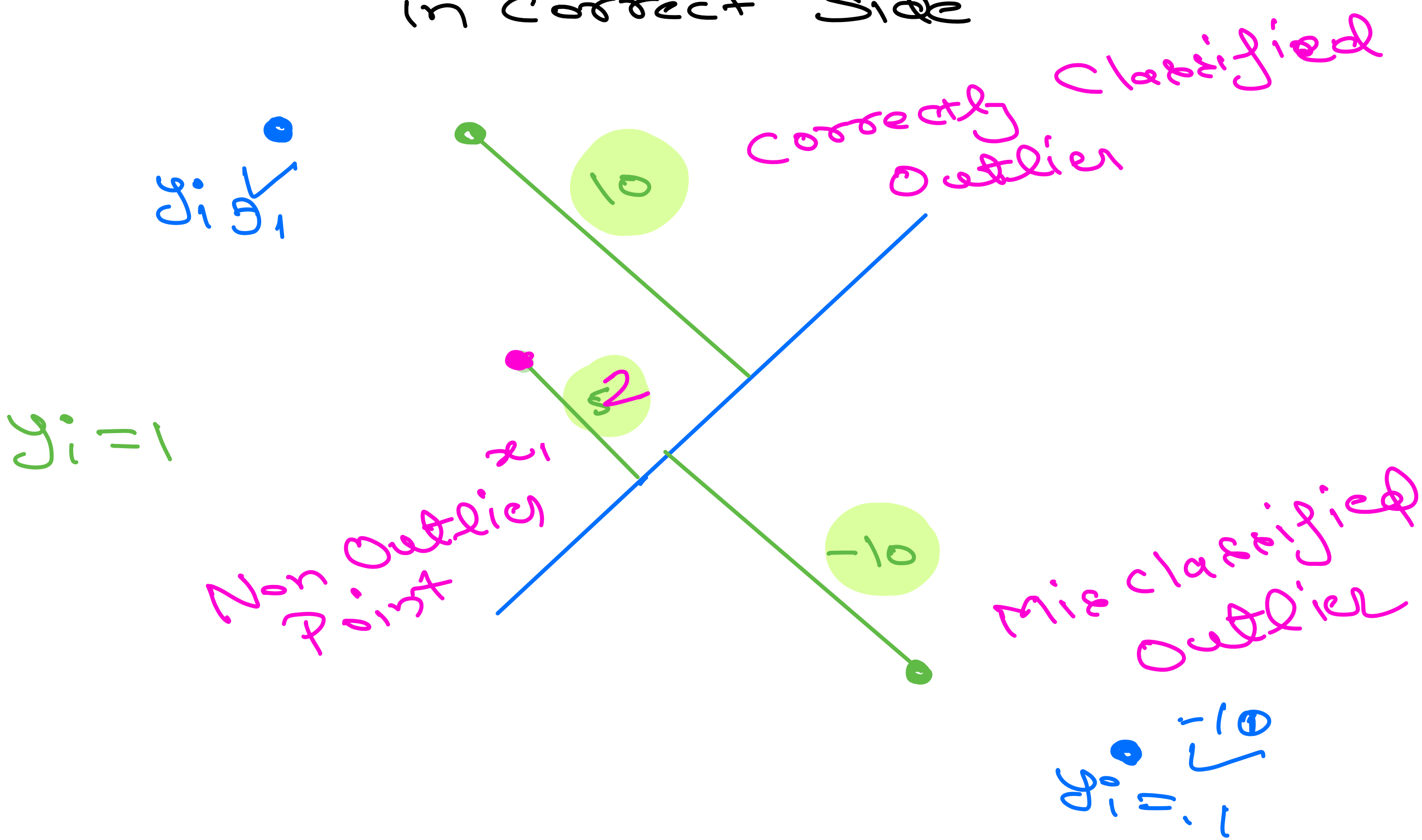
$$-y^{(i)} \cdot \log(\hat{y}^{(i)}) - 0$$

Let's say

$\hat{y}_{i=1} \Rightarrow 0.7 \quad x_1 \quad \odot \quad 0.35$

$\hat{y}_{i=2} \; 0.99 \quad x_2 \quad \odot \quad 0.01$ $\qquad 0.34$

Case-2: when the outlier is on in correct side



Misclassified Outlier will have very Large impact and Large Loss

what type of Outlier?

How do outliers affect the classification boundaries in logistic regression?

25 users have participated

A	Outliers shift the classification boundaries closer to the outlier values	52%
B	Outliers have no effect on the classification boundaries	20%
C	Outliers widen the gap between the classification boundaries	16%
D	Outliers make the classification boundaries more sensitive to minor changes	12%

Correctly Outliers → Not Much Impact

miss Classifie Outliers → Huge Impact

Multi-Class Classification

Orange

Apple → 3 Classes

Grape

	y_i
	0
	1
	2
	2
	0

$\{0, 1, 2\}$

3 Category

y_cap → $(0, 1) \neq \{0, 1, 2\}$

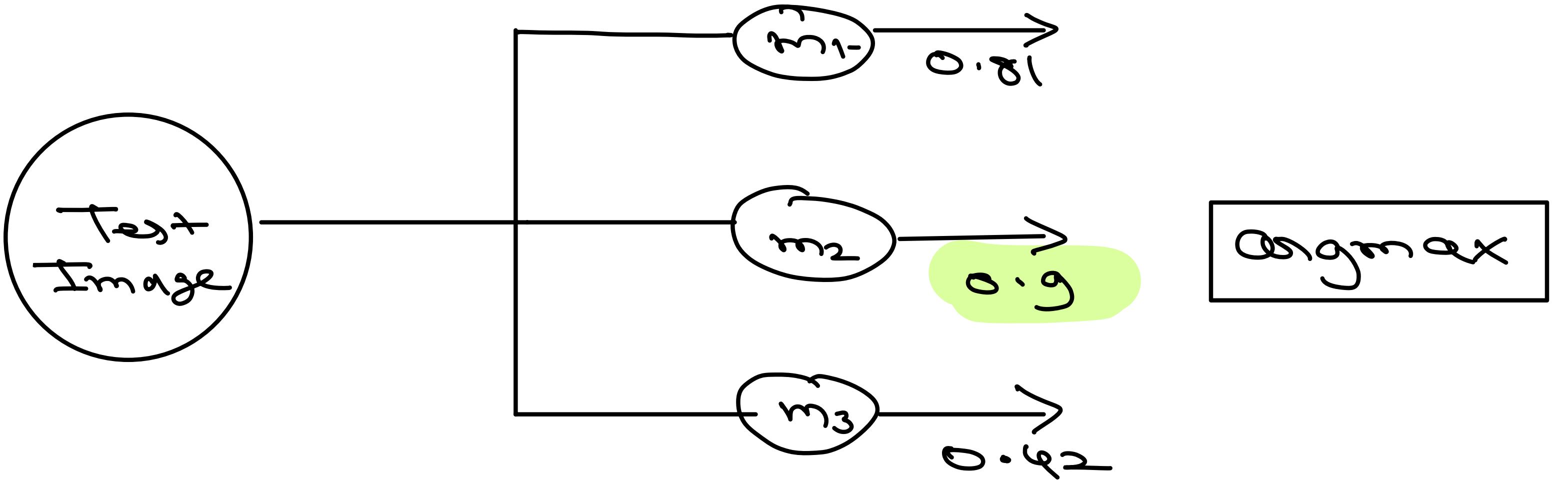
$1 = P(\text{orange}) + P(\text{not\_orange})$

M1 ⇒ Orange or Not_Orange

M2 ⇒ Apple or Not_apple

M3 ⇒ Grape or Not_Grape

$P(O) + P(A) + P(G) \neq 1$



Probs ⇒ [0.81 0.9 0.42]

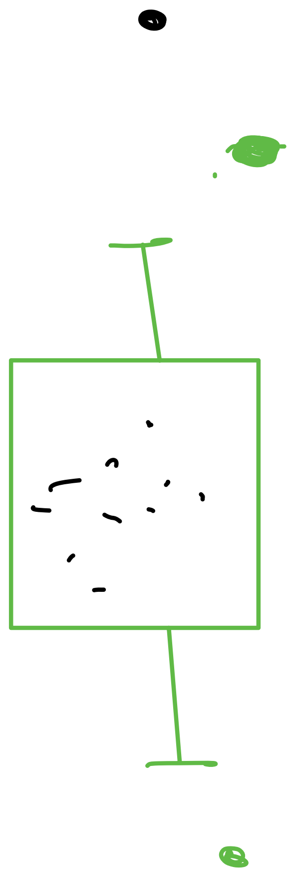
argmax ⇒ 1 ← Class 1

O v R

(One vs Rest Classifier)

call_time

x_1



outlier

$\hat{y} = 0.99 \Rightarrow 1$

compare with $y_i = 1$

① outlier