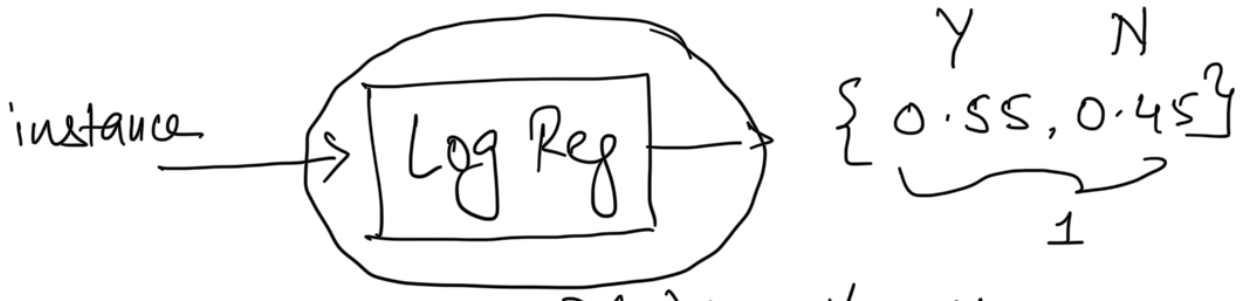
Logistic Regression :-

(Logit)

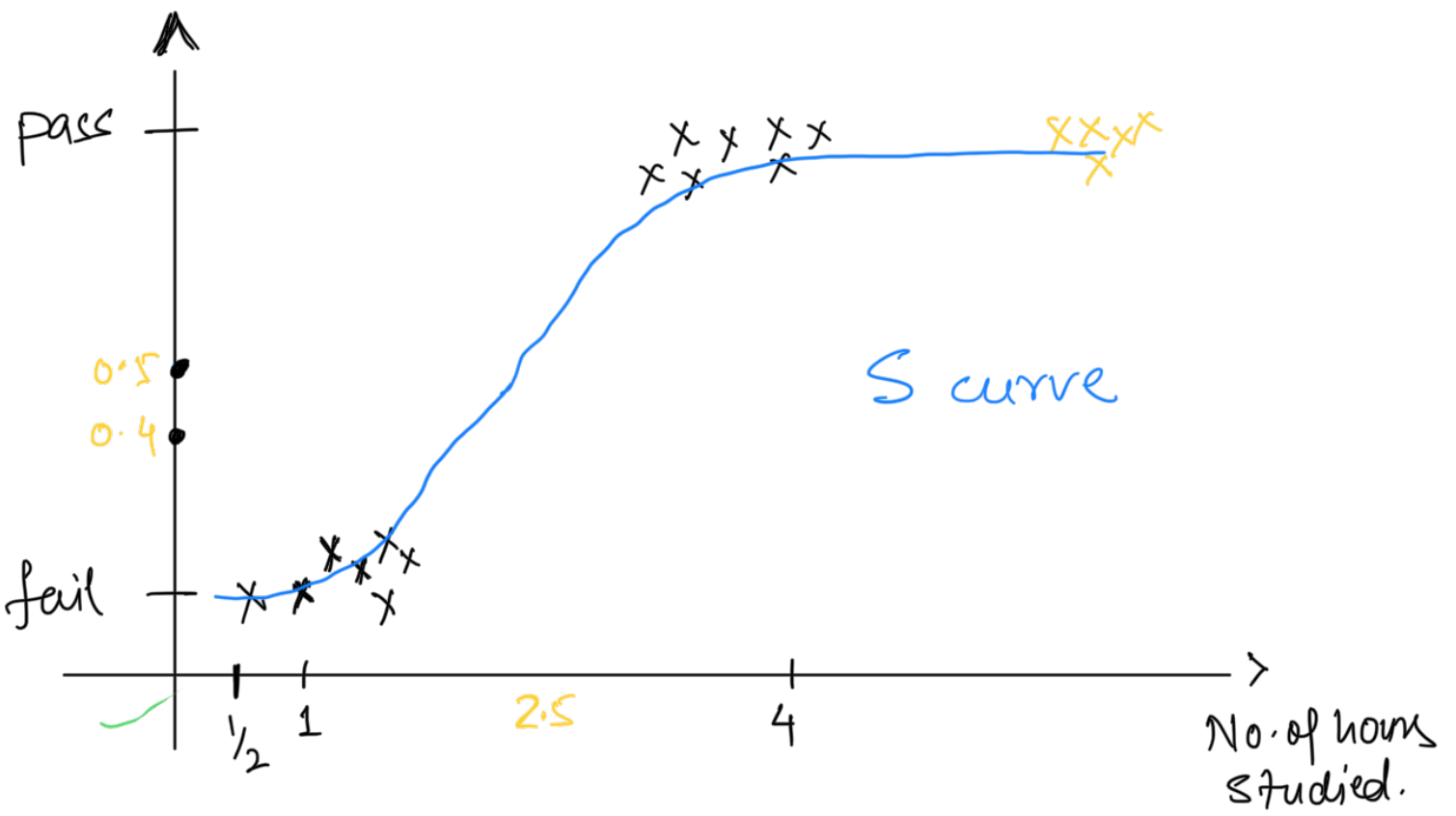
estimate the prob that an instance belongs to a certain class.

eg BANK provides loan.



P(Y) > 50% ; Yes

P(Y) < 50% ; No.



$y = f(x)$

pass/fail = f (No. of hours studied)

Scenario 1: 2.5 hours → passed.

2: 2.5 hours → fail Why ??

Because few nerdy students gave the exams.

Base $y = f(x)$

$\hookrightarrow y = \sigma(\theta^T x)$

σ : sigmoid, θ^T : parameter

① $0 < \sigma(t) = \frac{1}{1+e^{-t}} < 1$ } Log Reg.

② $\hat{y} = \begin{cases} 0 & \text{if } \hat{p} < 0.5 \\ 1 & \text{if } \hat{p} \geq 0.5 \end{cases}$

The score t is called as logit.

$$\text{logit}(p) = \log\left(\frac{p}{1-p}\right) \longrightarrow$$

Training

$y = \sigma(\theta^T x)$;

↑ learnable parameter

Set θ^T such that-

the $\sigma(\theta^T x)$ gives higher probabilities for +ve class

and $\sigma(\theta^T x)$ " lower " " -ve class

pass or fail.

No. of hours	Actual	pred (prob of passing)		valid
½	0	0.95	1	X
½	0	0.20	0	✓
5	1	0.78	1	✓
				✓

5	(1)	0.21	0	^

Cost function

$$c(\theta) = \begin{cases} -\log(\hat{p}) & \text{if } y=1 \\ -\log(1-\hat{p}) & \text{if } y=0 \end{cases}$$

(i) $-\log(\hat{p})$ ↑ when $t \to 0$; $y=1$
pred actual

(ii) $-\log(1-\hat{p})$ ↑ when $t \to 1$; $y=0$
pred actual

(iii) $-\log(\hat{p})$ ↓ when $t \to 1$; $y=1$
pred actual

(iv) $-\log(1-\hat{p})$ ↓ " $t \to 0$; $y=0$
pred actual

$$J(\theta) = -\frac{1}{m} \sum_{i=1}^{m} \left[y^{(i)} \log(\hat{p}^{(i)}) + (1-y^{(i)}) \log(1-\hat{p}^{(i)}) \right]$$

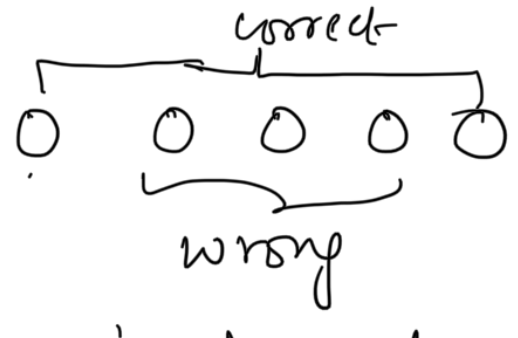
Interpretation (log of odds of passing)

How much of my chances of passing increases if I study for 1 more hour a day??

2 → 0.48
+ + ?? — slope (β_1 or m)
3 → (20%)

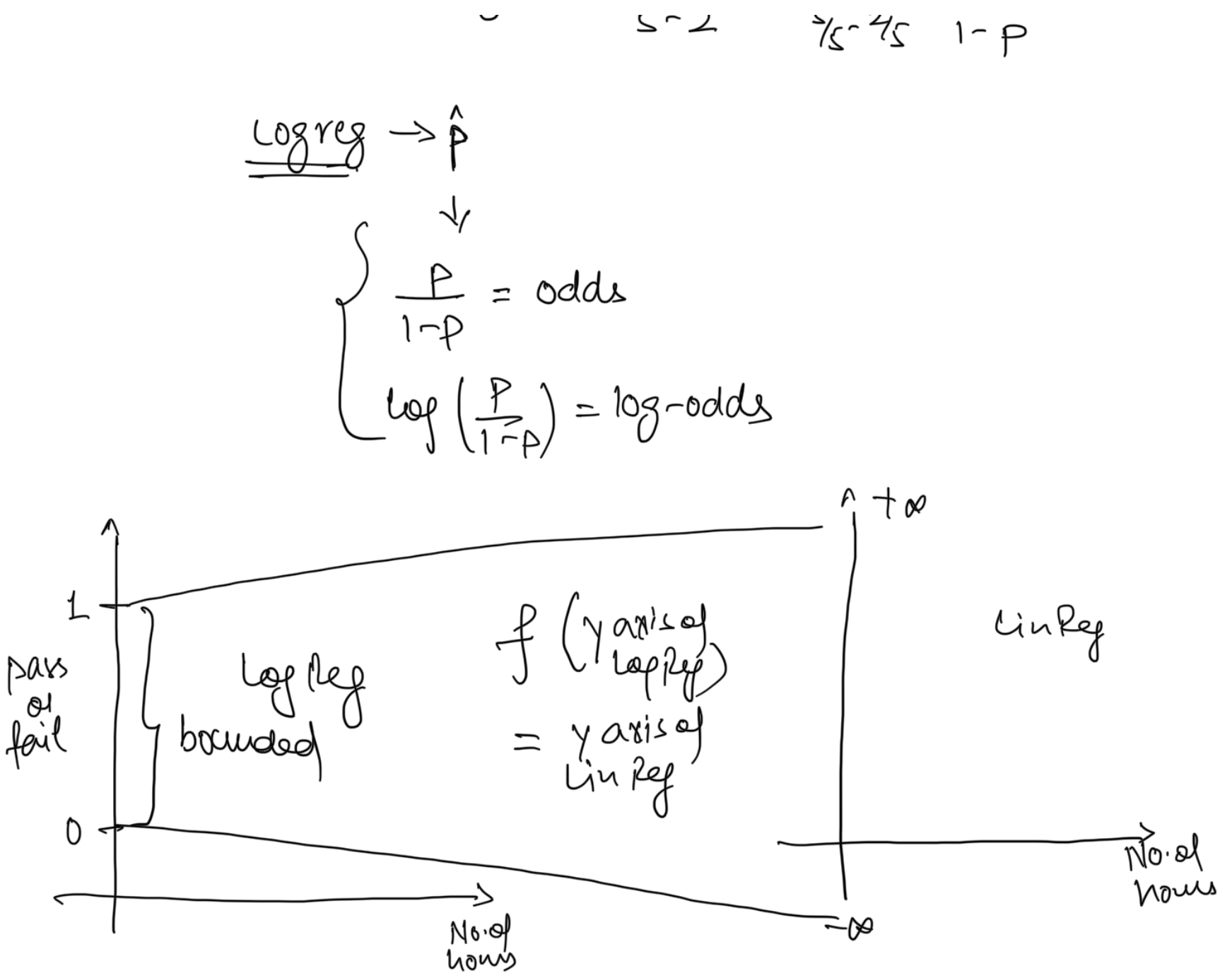
LogReg → LinReg

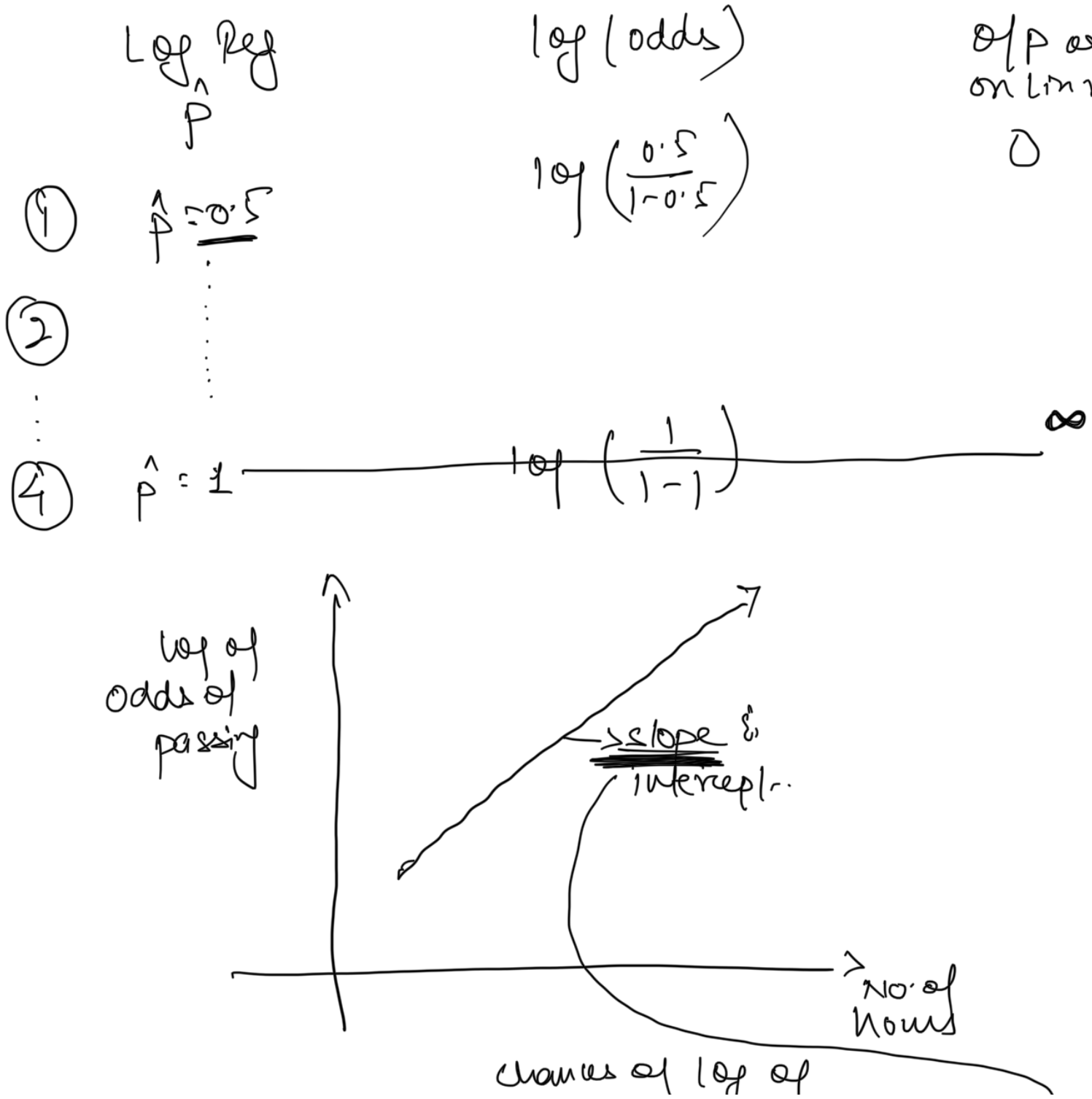
Odds



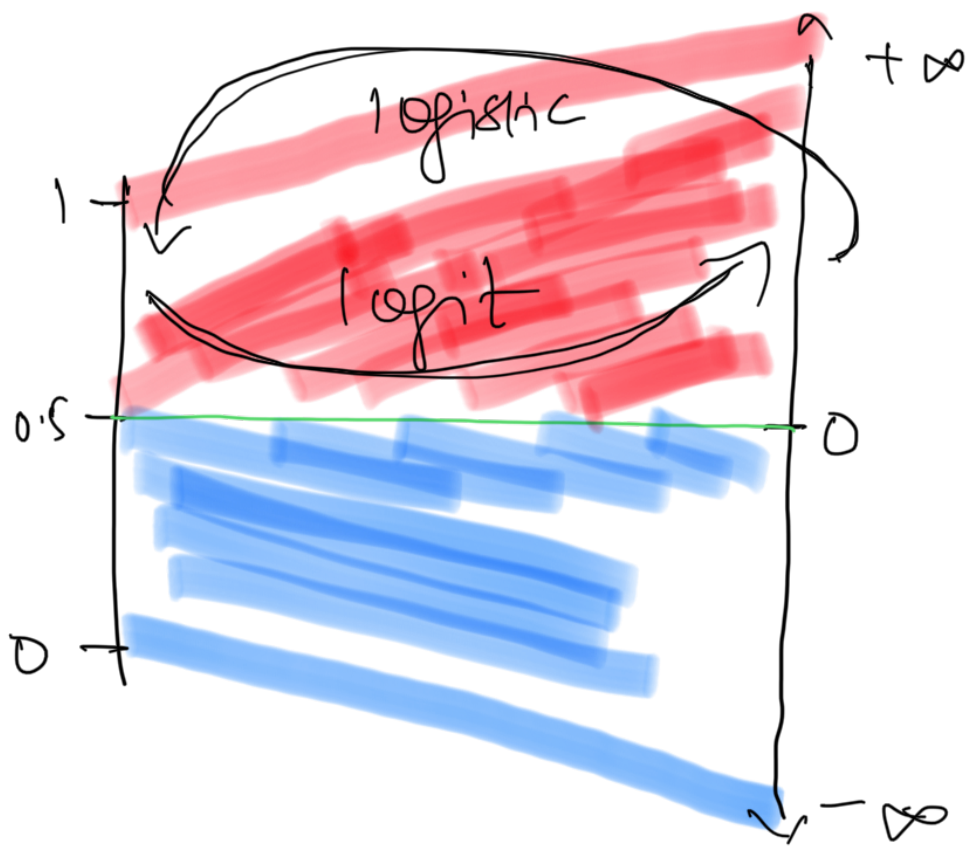
odds in favour of my answer being correct is

$$\frac{2}{3} = \frac{2}{5} = \frac{2/5}{} = \frac{P}{}$$


Logreg → p̂
p̂ / 1−p = odds
log (p̂ / 1−p) = log-odds
+∞
1
pass or fail
0
Log Reg bounded
f (y axis of LogReg) = y axis of LinReg
LinReg
No. of hours
−∞
No. of hours


Log Reg
p̂
log (odds)
O/P or plots on Lin reg axis
1
p̂ = 0.5
log (0.5 / 1−0.5)
0
2
4
p̂ = 1
log (1 / 1−1)
∞
log of odds of passing
slope & intercept
No. of hours
changes of log of

X How much of my odds of passing increases if I study for 1 more hour a day??



Evaluation Metrics

Accuracy.

Binary classification :-

Cancer detection machine

1) If test is -ve ; patient is assumed healthy
2) " " " +ve ; " will undergo additional tests.

Mistakes

1) If a healthy patient is classified as +ve, additional tests.

Consequences: Inconvenience,
Mental stress
Cost

incorrect +ve pred is called as false positive (1)

2) A sick patient classified as -ve,

No further tests or treatment [illegible]

Consequences: Serious health issues
Fatal

Incorrectly negative prediction → false negative (11)

Accuracy??

Actuals 0 0 0 0 1 0 0 0 0 0

Preds 0 0 0 0 0 0 0 0 0 0

(Naive Model All zeros)

Accuracy = 90%.

Confusion Matrix:

Actuals – 0 0 1 1 0 1 1 0 0 0

Preds – 1 0 0 1 0 0 1 1 0 0

$Acc = \frac{6}{10} = 60\%$.

Act \ Pred	0	1
0	4	2
1	2	2

Act \ Pred	+ve	-ve
+ve	TP	FN
-ve	FP	TN

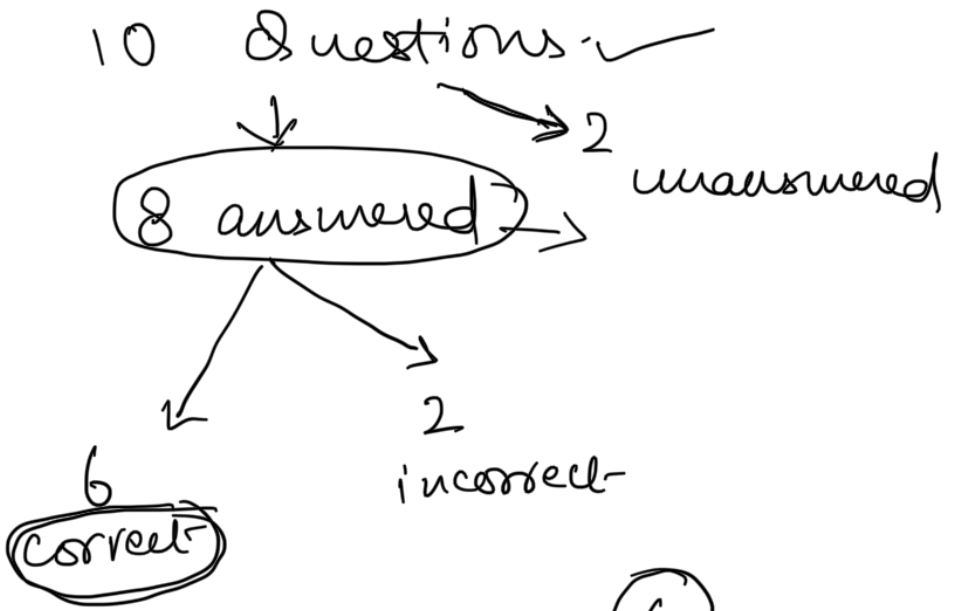
$Acc = \frac{TP + TN}{TP + FP + FN + TN}$; $Precision = \frac{TP}{TP + FP}$

limit false positives

$Recall = \frac{TP}{TP + FN}$ (sensitivity) ; $F1\ score = 2 \cdot \frac{precision \cdot recall}{precision + recall}$

limit FN

eg



$Acc = \frac{6}{10}$; $Precision = \frac{6}{8}$; $Recall = \frac{6}{10}$

At the airport; automated gun machine

identify a miscreant
and fire at miscreant.

-ve → Civilian +ve → Miscreant

① Actual = +ve
Predicts = +ve

② Actual = +ve
Predicts = -ve

③ Actual = -ve FP
Preds = +ve

④ Actual = -ve
Preds = -ve

Model 1

Act \ Pred	+	−
+	30	8
−	2	1880

Model 2

Ac \ Pr	+	−
+	35	3
−	22	980

X

FP vs FN
↓
precision

recall
Is detecting +ves more important
than detecting correct +ves ??

precision.