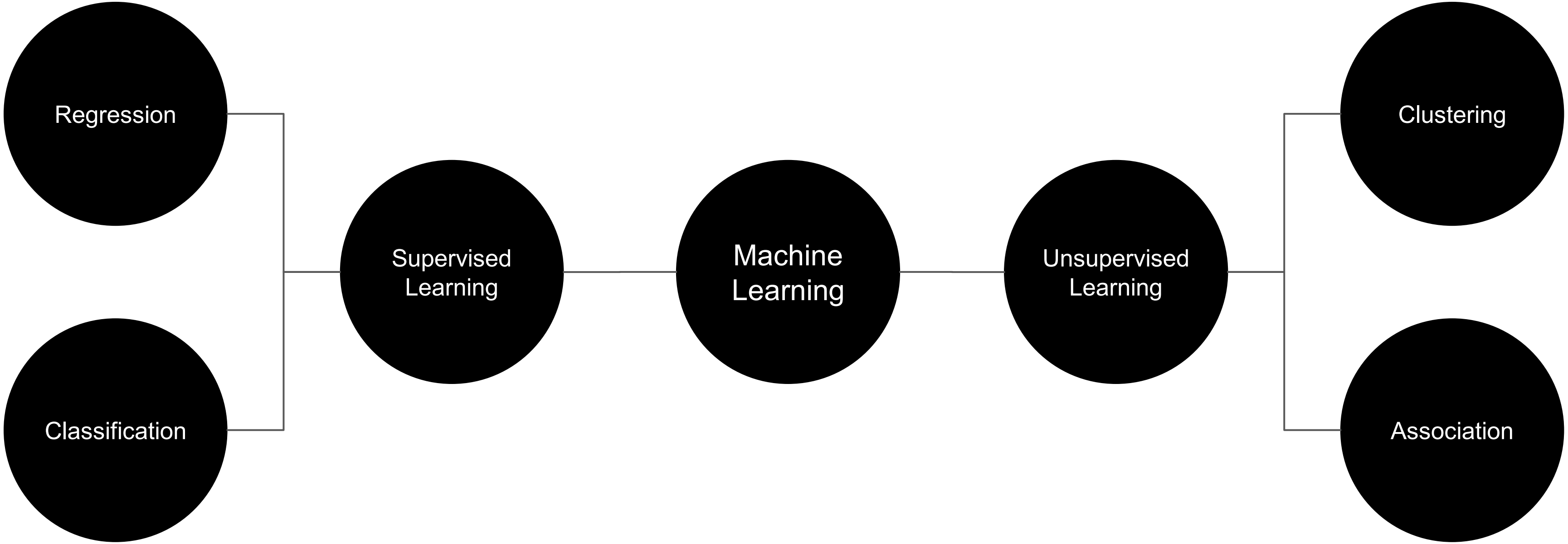
Data Prediction Model and Machine Learning

Online course #8

Logistic Regression



Regression

Classification

Supervised Learning

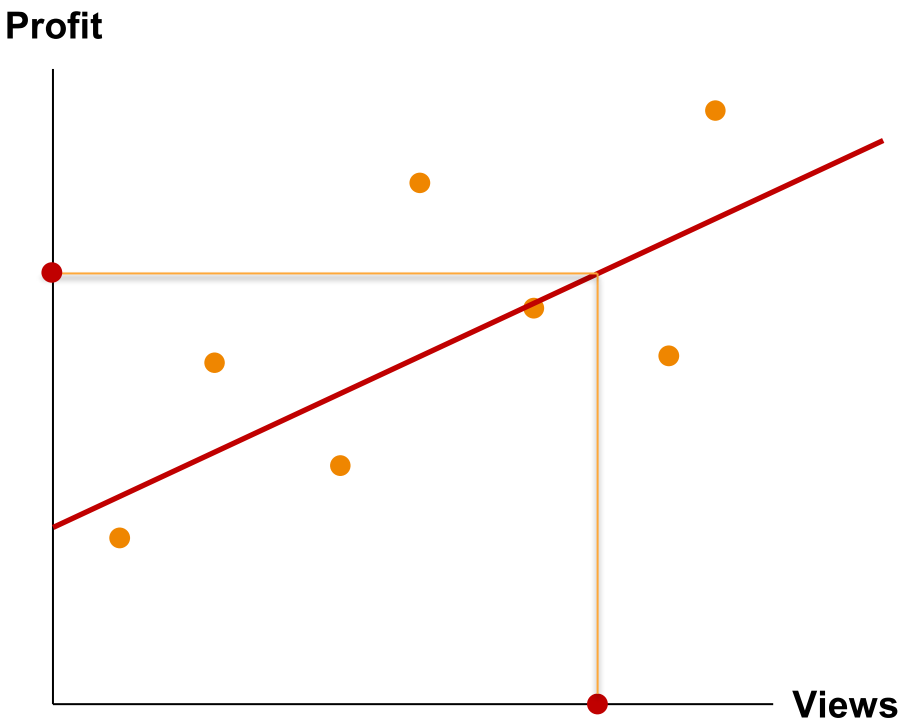
Machine Learning

Unsupervised Learning

Clustering

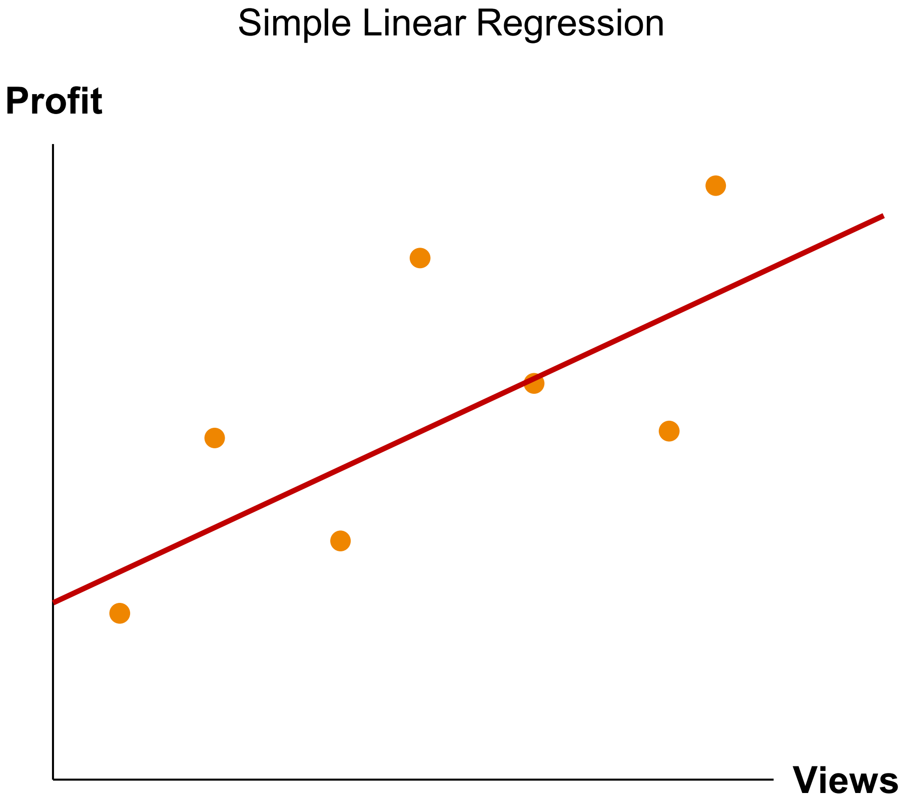
Association

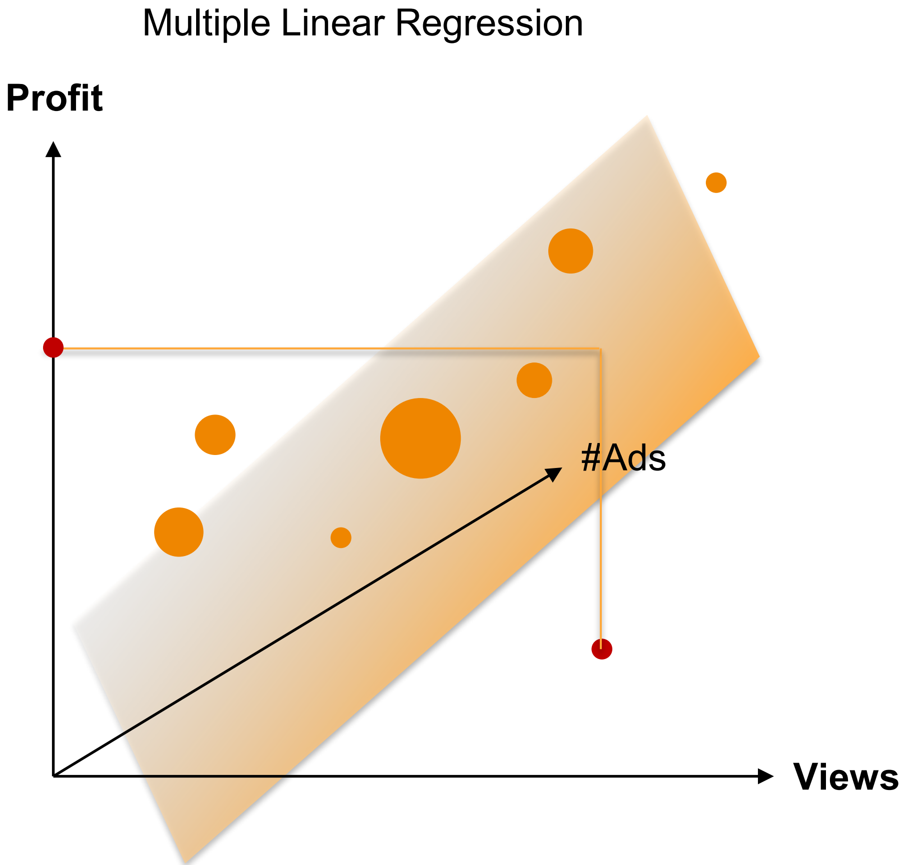
Linear Regression? Fitting a line to data



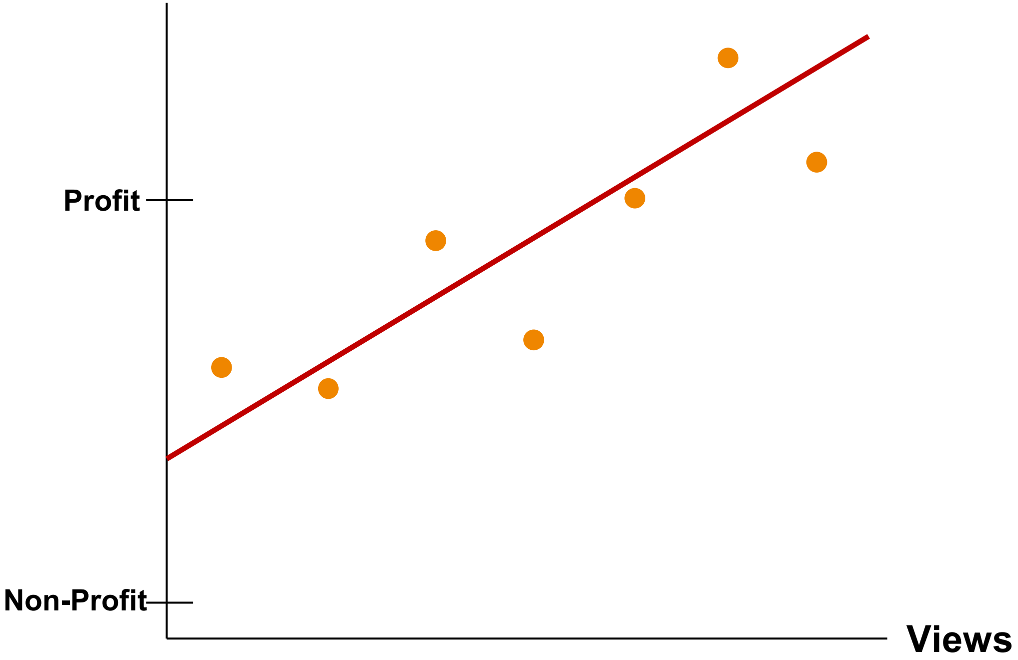
1) Find a line minimizing the SSE
2) Calculate R-squared ← significant correlation btw X and Y
3) Use the line to predict Y given X

Linear Regression? Fitting a line to data





When a line cannot be fitted in the data



$$Y = \begin{Bmatrix} 1\ when\ Profit > 0 \\ 0\ when\ Profit \leq 0 \end{Bmatrix}$$

⇩

Classification

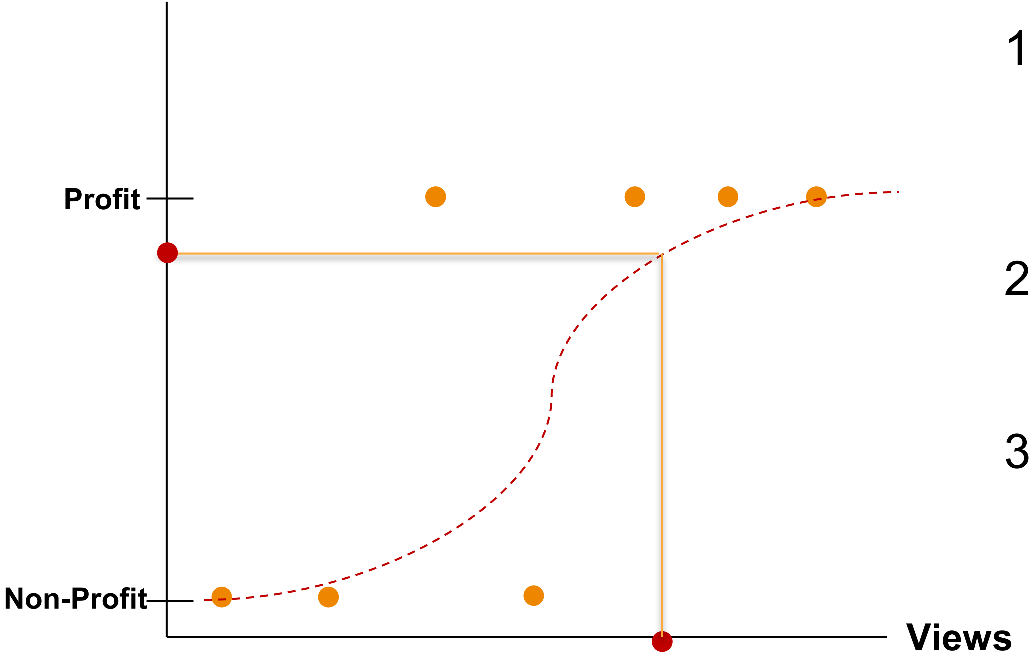
Logistic regression?

Regression ⟺ Classification

- Y = Probability [0:1]

- Y = Classification (0 or 1)
- Not allowed (classification among A,B,C)

Logistic regression: How different from LM?



Similar to linear regression **except**

1) Logistic regression predicts whether something is T or F, instead of predicting numerical Y

2) Fits a "S" shaped "logistic function" to data instead of a line

3) Curve let us know the probability

Logistic regression: for two or more features

$$Profit_i = \alpha + \beta Views_i + \varepsilon_i$$

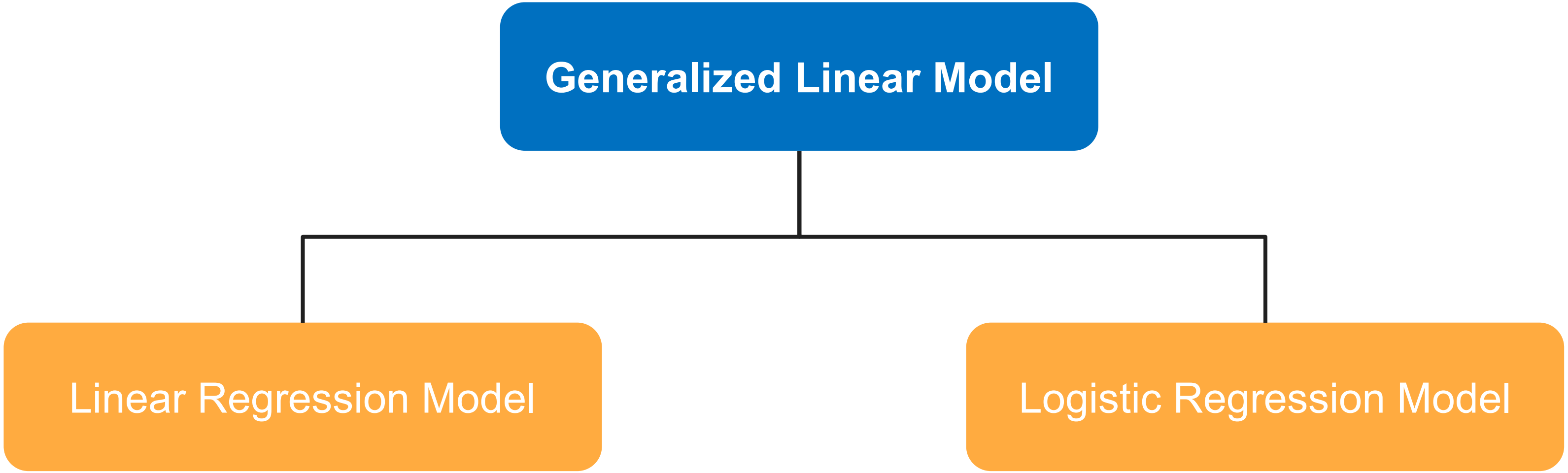
$$Profit_i = \alpha + \beta Views_i + \gamma Ads_i + \varepsilon_i$$

$$Profit_i = \alpha + \beta Views_i + \gamma Ads_i + (Youtuber's\ gender, blood\ type\ and\ so\ on \ldots)\varepsilon_i$$

Just like linear regression,

- logistic regression can work with continuous data (Views, Ads),
- and also with discrete data (like gender and blood type)

Logistic regression: as a family of GLM

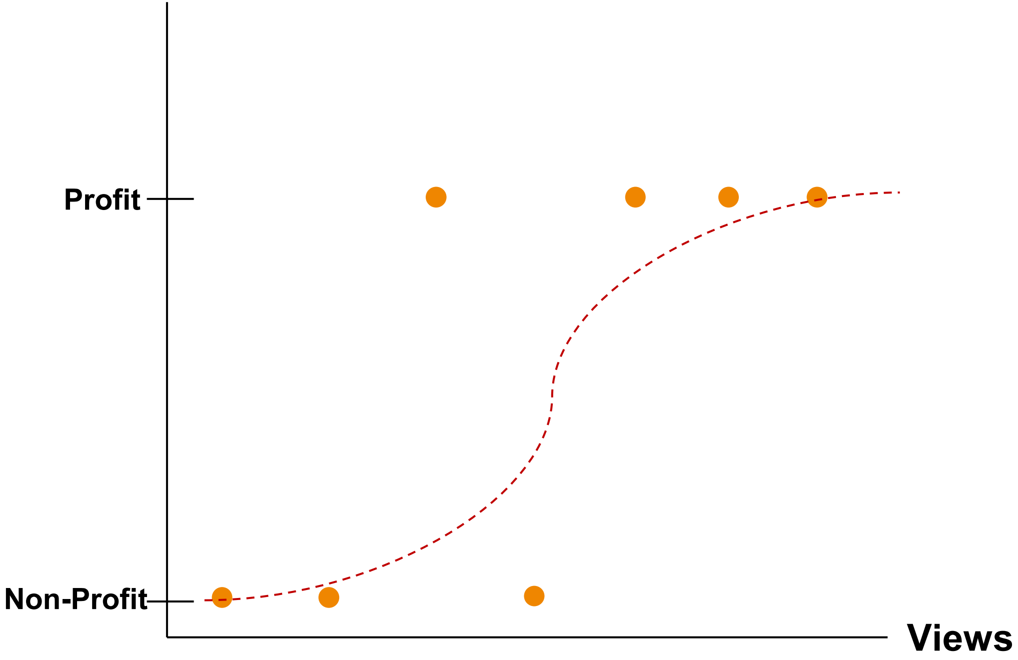


Generalized Linear Model

Linear Regression Model

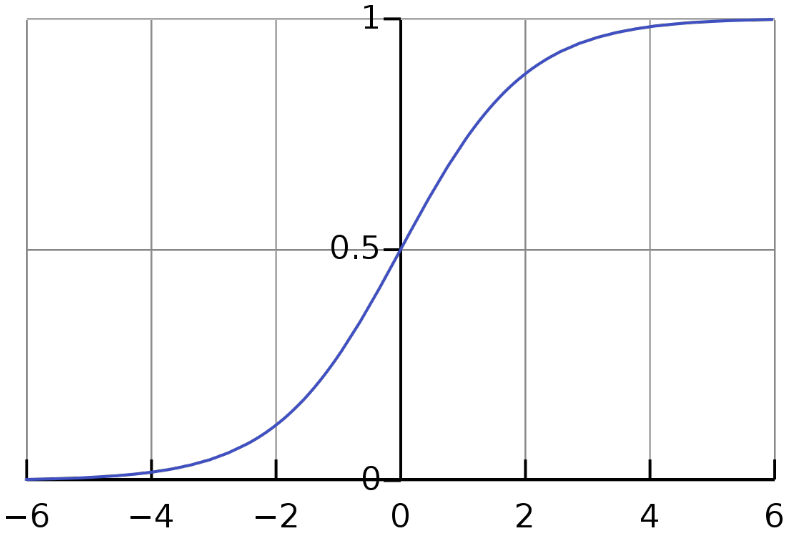
Logistic Regression Model

How to make an “S” curve?



$$S(x) = \frac{1}{1+e^{-x}}$$

$$x \in \{-inf, inf\}$$
$$S(x) \in (0,1)$$

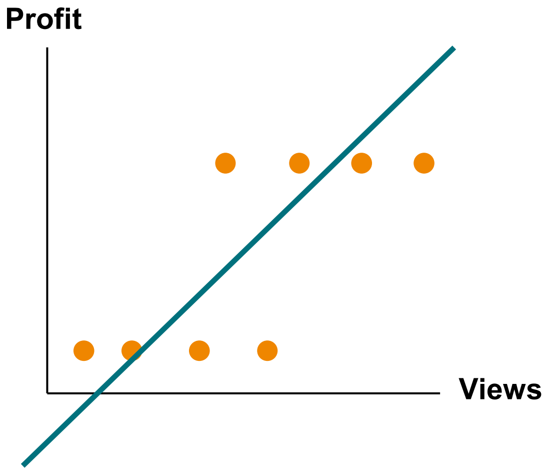


How to make an “S” curve?

Linear Model

$$Profit = f(View)$$

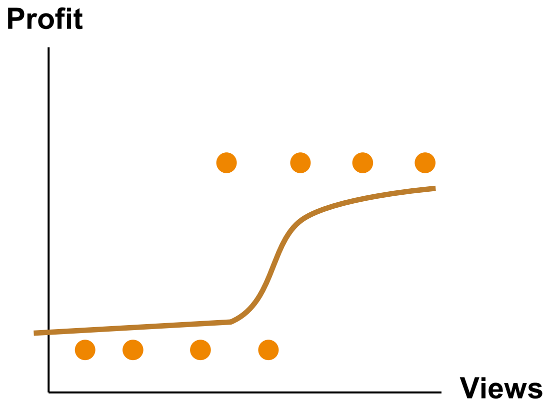
$$f(View) = \alpha + \beta\ View$$



Logistic Model

$$Profit = S(View)$$

$$S(View) = \frac{1}{1 + e^{-(\alpha + \beta\ View)}}$$

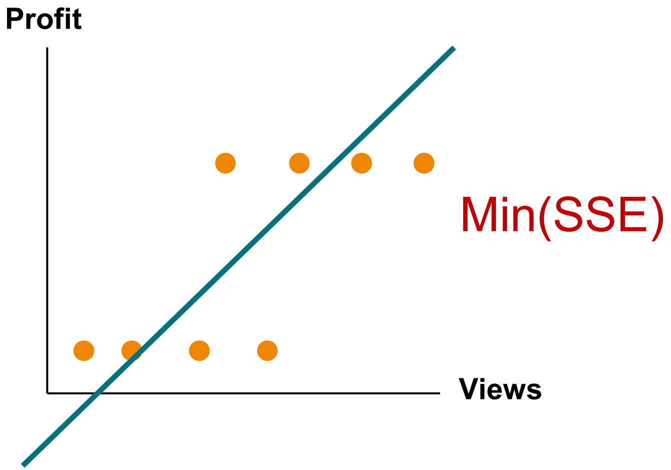


How to find a better fitting "S" curve line?

Linear Model

$$Profit = f(View)$$

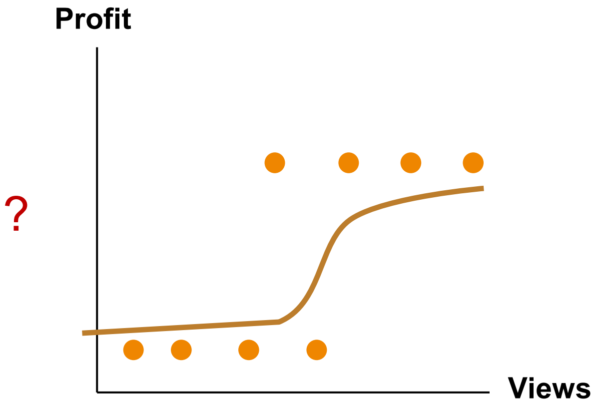
$$f(View) = \alpha + \beta\ View$$



Logistic Model

$$Profit = S(View)$$

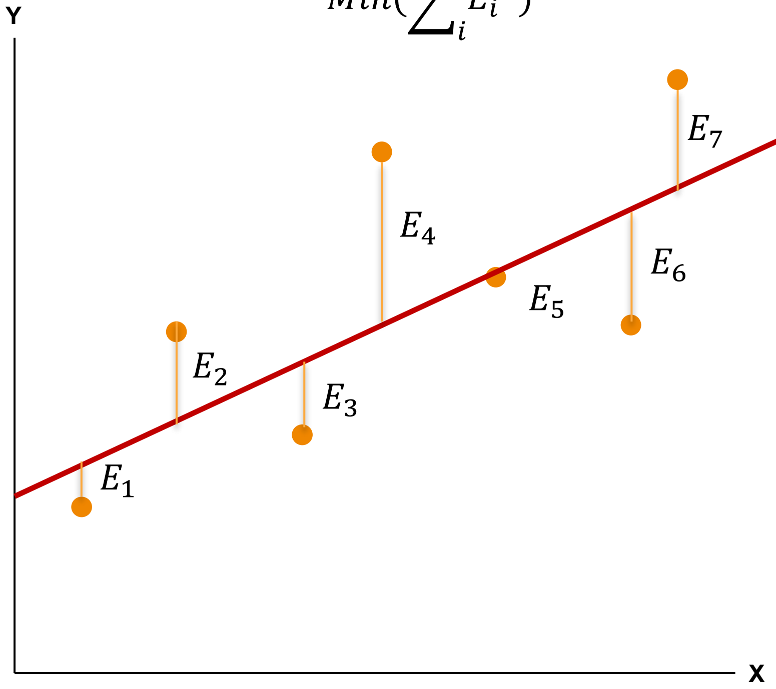
$$S(View) = \frac{1}{1 + e^{-(\alpha+\beta\ View)}}$$



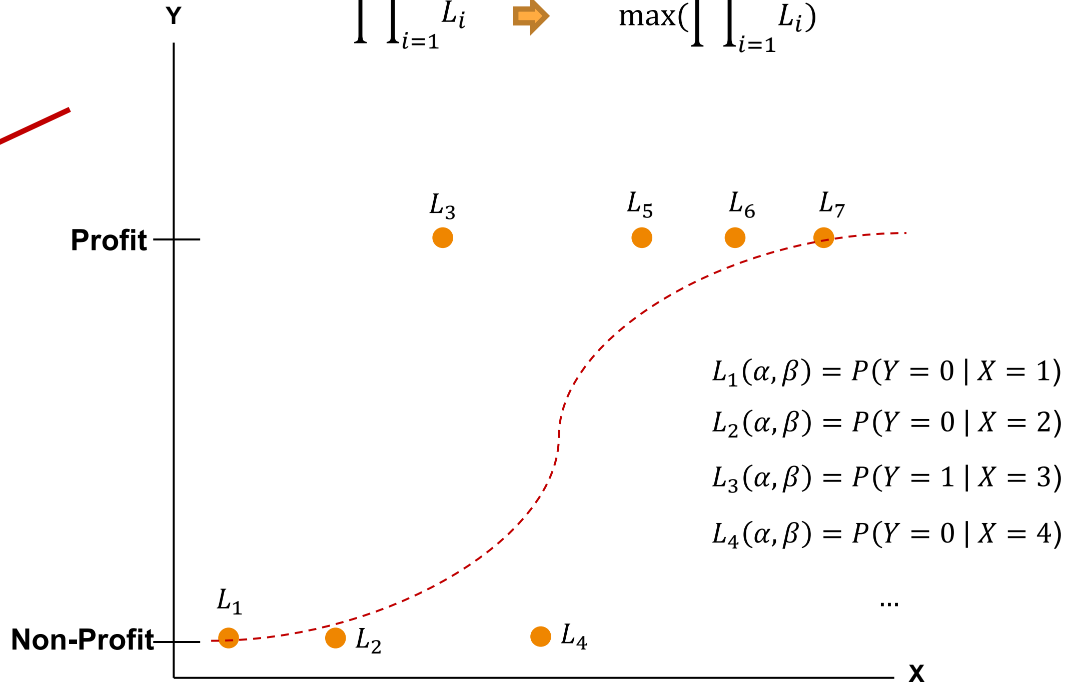
How to find a better fitting “S” curve line?

SSE vs. Likelihood

$$Min(\sum_i {E_i}^2)$$



$$\prod_{i=1}^{7} L_i \Rightarrow \max(\prod_{i=1}^{7} L_i)$$



$$L_1(\alpha, \beta) = P(Y = 0 \mid X = 1)$$

$$L_2(\alpha, \beta) = P(Y = 0 \mid X = 2)$$

$$L_3(\alpha, \beta) = P(Y = 1 \mid X = 3)$$

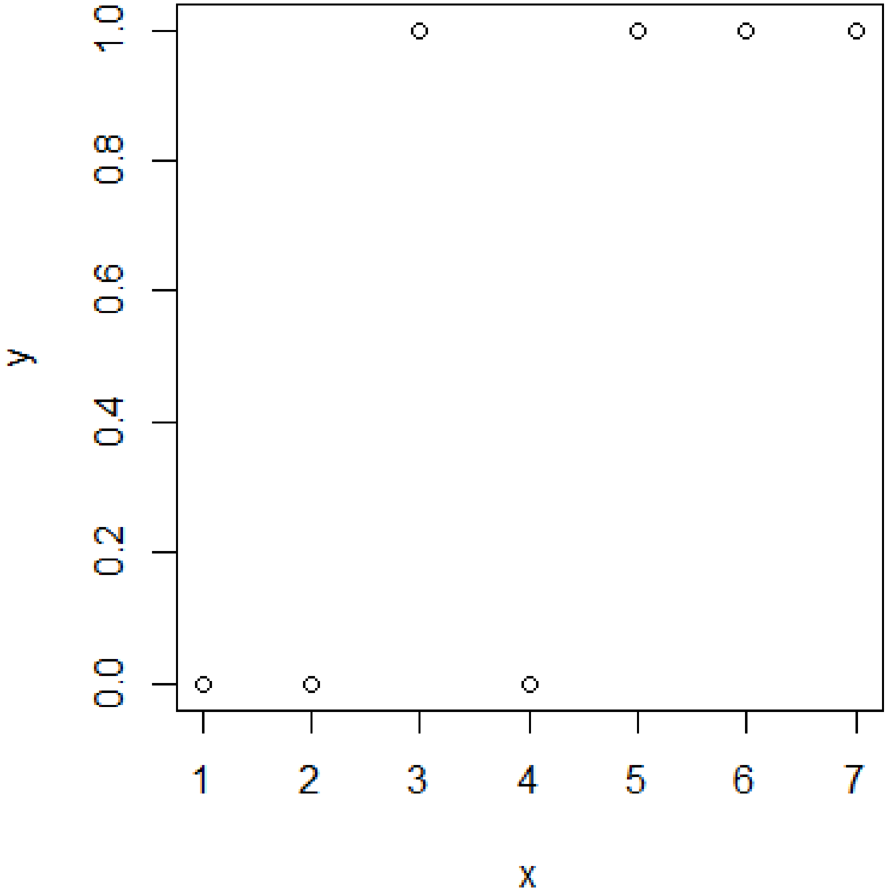
$$L_4(\alpha, \beta) = P(Y = 0 \mid X = 4)$$

…

Let’s get back to the first example

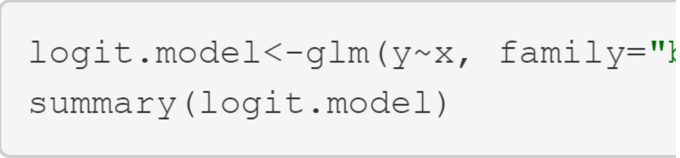
Video	Views	Profit
1	1	0
2	2	0
3	3	1
4	4	0
5	5	1
6	6	1
7	7	1

```
x=c(1:7)
y=c(0,0,1,0,1,1,1)
par(pty="s")
plot(x,y)
```

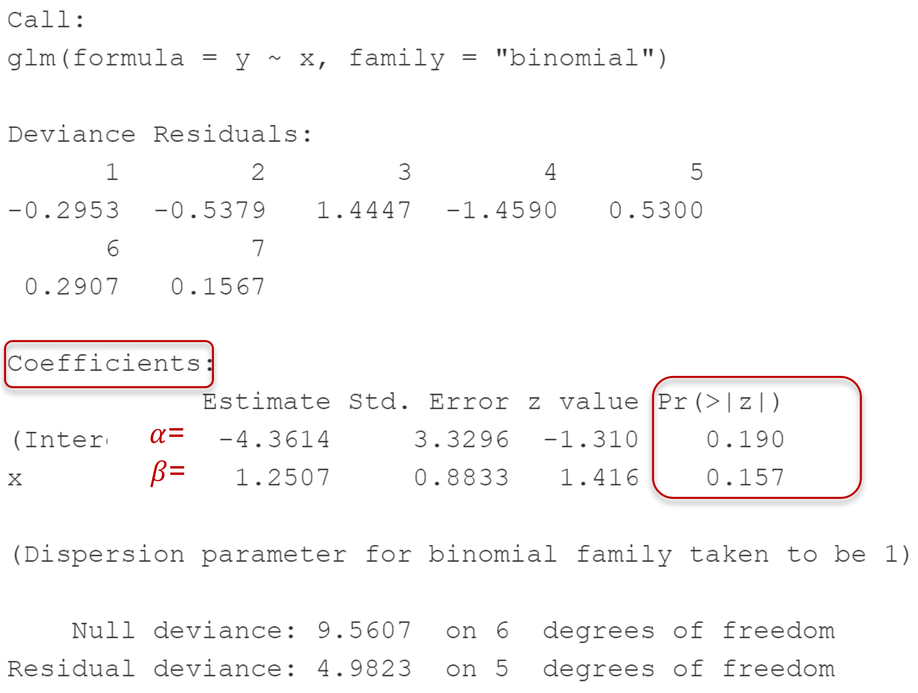


Learning model with logistic model function (glm)

Video	Views	Profit
1	1	0
2	2	0
3	3	1
4	4	0
5	5	1
6	6	1
7	7	1



```
logit.model<-glm(y~x, family="
summary(logit.model)
```



```
Call:
glm(formula = y ~ x, family = "binomial")

Deviance Residuals:
      1        2        3        4        5
-0.2953  -0.5379   1.4447  -1.4590   0.5300
      6        7
 0.2907   0.1567

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Inter  α=  -4.3614     3.3296  -1.310    0.190
x       β=   1.2507     0.8833   1.416    0.157

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 9.5607  on 6  degrees of freedom
Residual deviance: 4.9823  on 5  degrees of freedom
AIC: 8.9823
```

The logistic regression coefficients give the change in the log odds of the outcome for a one unit increase in the predictor variable.

Model interpretation

What is Odds?

Odds provide a measure of the likelihood of a particular outcome. They are calculated as the **ratio of the number of events that produce the outcome to the number that don't.** Odds are commonly used in gambling and statistics.

$$Odds = \frac{P(Event = 1)}{P(Event = 0)}$$

$$P(Event = 1) = \frac{\#\ of\ observations\ in\ which\ the\ event\ occur}{\#\ of\ observations}$$

Model interpretation

What is Odds?

Logistic Model

$$Profit = S(View)$$

$$S(View) = \frac{1}{1 + e^{-(\alpha + \beta\, View)}}$$

$$P(Profit = 1) = \frac{1}{1 + e^{-(\alpha + \beta\, View)}}$$

$$\frac{1}{P(Profit = 1)} = 1 + e^{-(\alpha + \beta\, View)}$$

$$\frac{1}{P(Profit = 1)} - 1 = e^{-(\alpha + \beta\, View)}$$

$$\ln\left(\frac{1}{P(Profit = 1)} - 1\right) = -(\alpha + \beta\, View)$$

$$\ln\left(1 - \frac{1}{P(Profit = 1)}\right) = \alpha + \beta\, View$$

Log Odds

Odds

$$\boxed{\ln\left(\boxed{\frac{P(Profit = 1)}{P(Profit = 0)}}\right)} = \alpha + \beta\, View$$

Learning model with logistic model function (glm)

Video	Views	Profit
1	1	0
2	2	0
3	3	1
4	4	0
5	5	1
6	6	1
7	7	1

```
logit.model<-glm(y~x, family="
summary(logit.model)
```

```
Call:
glm(formula = y ~ x, family = "binomial")

Deviance Residuals:
      1        2        3        4        5
-0.2953  -0.5379   1.4447  -1.4590   0.5300
      6        7
 0.2907   0.1567

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Inter   α=  -4.3614     3.3296  -1.310    0.190
x        β=   1.2507     0.8833   1.416    0.157
```

$$\ln\left(\frac{P(Profit=1)}{P(Profit=0)}\right)=\alpha+\beta\,View$$

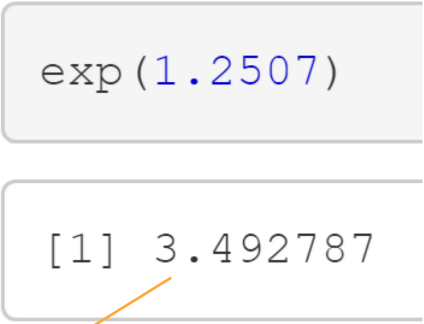
```
Residual deviance: 4.9823  on 5  degrees of freedom
AIC: 8.9823

Number of Fisher Scoring iterations: 5
```

Model interpretation

$$\ln\left(\frac{P(Profit = 1)}{P(Profit = 0)}\right) = \alpha + \beta\, View$$

$$\frac{P(Profit = 1)}{P(Profit = 0)} = e^{\alpha + \beta\, View}$$



```
exp(1.2507)
```

```
[1] 3.492787
```

- For every one unit change in Views, the Odds of Profit (versus non-profit) increases by 3.5
- P(Profit) > P(non-profit) : **Three point five times bigger**