Normal linear

$$y_i \sim N(\mu_i, \sigma^2)$$

$$\mu_i = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ki}$$

binary logistic

$$y_i \sim \text{Bernoulli}(\theta_i)$$

$$\underbrace{\log\left[\frac{\theta_i}{1-\theta_i}\right]}_{\text{link}} = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ki}$$

Poisson

$$y_i \sim \text{Poisson}(\lambda_i)$$

$$\underbrace{\log(\lambda_i)}_{\text{link}} = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ki}$$

Neg binomial

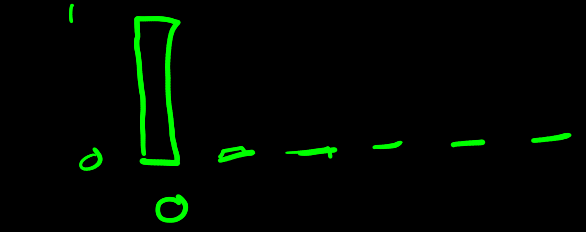
$$y_i \sim \text{NegBinom}(\mu_i, \phi)$$

$$\underbrace{\log(\mu_i)}_{\text{link}} = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ki}$$

	Z	y insta
1	0	5
2	0	2
3	0	0 ← insta user, but infreq
4	1	0 ← non insta user
⋮		
n		

$$y \quad \underline{\underline{x}} \quad z$$

$$y \sim \begin{cases} \underline{\text{Poisson}(\lambda)} & \text{if } z = 0 \\ \text{zero dist} & \text{if } z = 1 \end{cases}$$



y	x	z (latent)	
0	20	1	← non smoker
2	15	0	
10	25	0	
20	10	0	
3	12	0	
0	16	0	←
⋮	⋮		

person with 20 yrs educ

- prob of smoking : 25%
- predicted avg cigs of smoker (with 20 years educ) = 30

What is the average number of cigs for a person

$0.25 \times 30 + 0.75 \times 0$

30
smoker
25%

0
non-smoker
75%