35. GLMs and Neural Nets

December 5, 2025 9:31 AM

Logistic Regression: $Y \sim \text{Bernoulli}(\pi(x))$

$Y \sim \text{Binomial}(m, \pi(x))$

Modeling the odds ratio using a log linear model

$$\log\left(\frac{\pi}{1-\pi}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

we could use a different "link" function if we want.

if π is small and m is big

Then $Y \sim \text{Poisson}(\lambda(x))$

Modeling the counts (Integers) using a log linear model

$$\log(\lambda) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

What do we know about $Y \sim \text{Poisson}(\lambda)$ when λ is "big"?

The central limit theorem says that

$$\frac{Y-\lambda}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$$

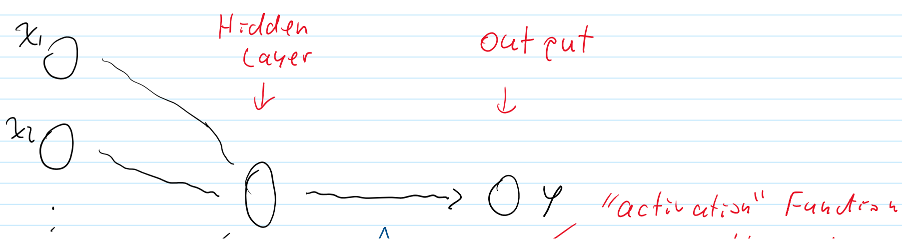
We can also think of a log linear model with Normal Data as a GLM.

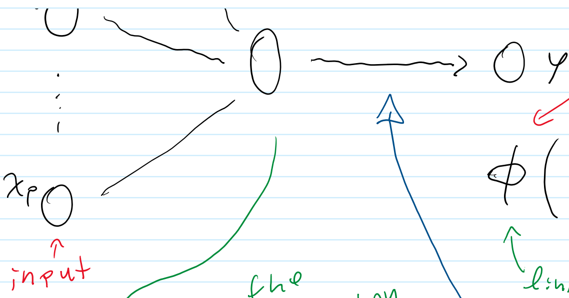
$$\log(\mu) = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p$$

$\mu(x)$

$$\mu = \mathbb{E}[Y \mid x] \text{ where } Y \sim N(\mu, \sigma^2)$$

If you see a Neural Net, it may look like this..





"activation" Function

weights bias

$$\phi(w \cdot x + b)$$

link parameters $\beta_1 \cdots \beta_p$ constant β_0

Choose ϕ to be the logistic / Sigmoid Function then this is just logistic Regression.

Key difference: We do another logistic regression with the output of the 1st one as input to the 2nd one.

- Why do this?
- Can we even fit this to data?
- is the fitted unique?
- Overfit?

so from linear to approximate Non-linear relations.