Generalized Linear Models and Maximum Likelihood

"The trouble with normal is that it always gets worse"

Samuel Robinson, Ph.D.

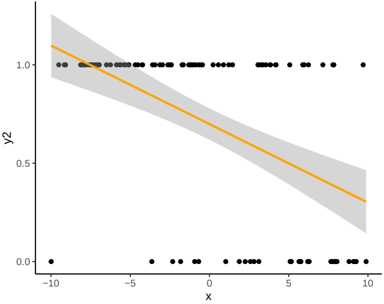
November 19, 2020

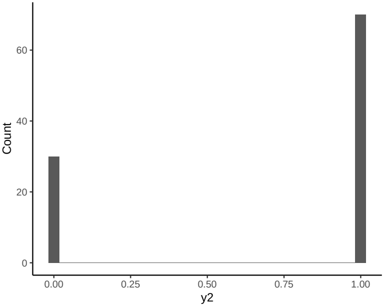
Motivation

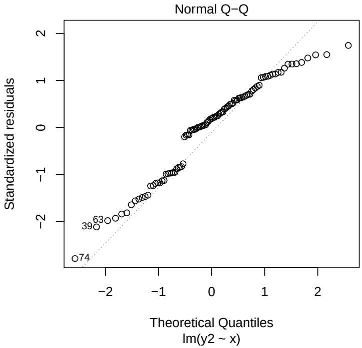
What are Generalized Linear Models? (GLMs)

- Meet the exponential family
  - Normal, t, Binomial, Poisson
  - Negative Binomial, Beta, Gamma
- Probability and Likelihood
- Tricksy hobbitses!
  - Zero-inflated models, occupancy models

Problem: not everything is normal

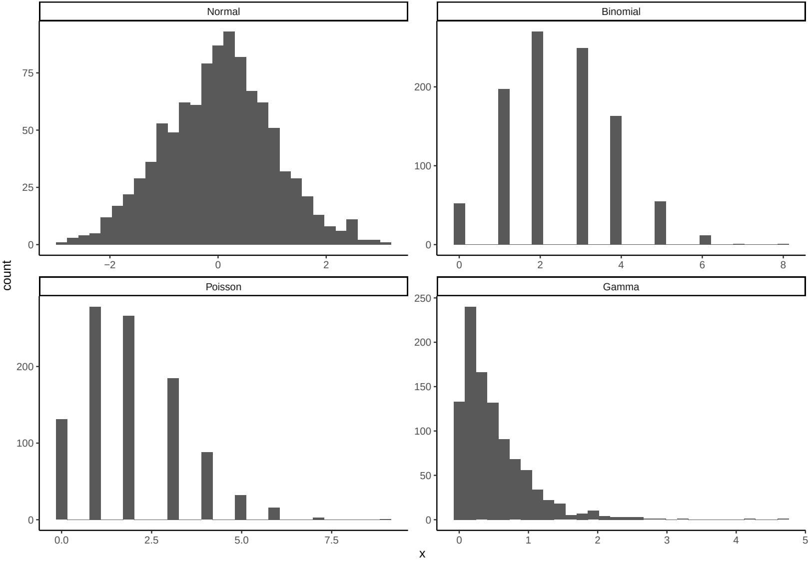






- Some types of data can never be transformed to make the residuals normal
- Solution: **use the distribution that generates the data!**

But how do I know which distribution to use?



Time to meet the family!

The Normal Distribution (aka *Gaussian*)

- Imagine many random + and - numbers added together
- If you do this *many* times:
  - Most cancel out (somewhere around 0)
  - Few are far away from 0 (tails of distribution)
- Common in nature, because of many small + and - factors adding together
  - e.g. Height is driven by many sets of genes

A Galton Board in action:

The Normal Distribution - scary math!

- 2 parameters: mean (μ) and standard deviation (σ)

$$p(x|\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2}$$

- Probability distribution function (PDF) for the Normal distribution
- Tells you about the probability of getting some number *given* μ and σ

Example: what is the probability of getting a 4, if the mean is 5 and SD is 1?

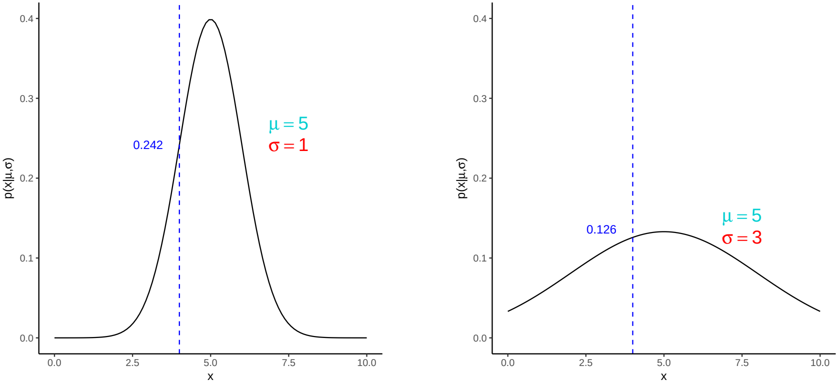
$$p(4|5,1) = \frac{1}{1\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{4-5}{1})^2}$$
$$= \sim 0.24$$

In R, this is easy:

```
#d stands for "density"
dnorm(x=4,mean=5,sd=1)
```

```
## [1] 0.2419707
```

The Normal/Gaussian Distribution



- Probability of x changes with μ and σ
- Left: $\sigma = 1$, Right: $\sigma = 3$

The Binomial Distribution

- Imagine you have 10 coins, and you flip them all
- If you do this *many* times:
  - Most will be about 5 heads, 5 tails
  - Few will be 1 head, 9 tails (or reverse)
- Common in nature where outcomes are binary
  - e.g. 10 seeds from a plant, how many will germinate?