Likelihood and Regression

Author: Nicholas G Reich

*This material is part of the **statsTeachR** project*

Today's Lecture

- Likelihood defined
- Likelihood in the context of regression

These notes are based loosely on Michael Lavine's book Introduction to Statistical Thought, Chapters 2.3-2.4.

Parametric families of distributions

A parametric distribution

- In the analysis of real data, we often are willing to assume that our data come from a distribution whose general form we know, even if we don't know the exact distribution.
- E.g. $X \sim \mathit{Poisson}(\lambda)$ or $Y \sim N(\mu, \sigma^2)$
- Each of the above examples refer to families of distributions, defined or indexed by particular parameter(s).
- In statistics, we try to estimate or learn about the unkown parameter.

The likelihood function

Another look at a pdf

- Probability density functions (pdfs) define the probability of seeing a specific observed value of your random variable, conditional on a parameter.

$$f(X|\theta)$$

- However, we can think about this same function another way, by *conditioning* on the data and looking at the probability taken by different values of the parameter.

$$f(\theta|X) = \ell(\theta)$$

- Remember, the definition of the joint density of observations that we assume to be i.i.d.: if $X_1, X_2, ..., X_n \sim i.i.d. f(x|\theta)$ then

$$f(X_1, \ldots, X_n|\theta) = \prod f(X_i|\theta)$$

Likelihood as evidence

"A wise man … proportions his belief to his evidence."
-David Hume, Scottish philosopher

We often compare values of the likelihood function as ratios, weighing the evidence for or against particular values of θ.

$$\frac{\ell(\theta_1)}{\ell(\theta_2)} = 1$$

implies we have the same evidence to support either θ_1 or θ_2.

$$\frac{\ell(\theta_1)}{\ell(\theta_2)} > 1$$

implies we have more evidence to support θ_1 over θ_2.

A simple, canonical example: coin-flipping

Let's flip some coins! A plausible statistical model here is for the number of heads (X) when I flip a coin N times

$$X \sim Binomial(N, p)$$

where

$$f(x|\theta) = \ell(\theta) = \binom{n}{x} \cdot \theta^x \cdot (1-\theta)^{n-x}$$