

# Maximum likelihood estimation

Ciaran Evans

## Fitting a *logistic* regression model?

Linear regression: minimize  $\sum_{i=1}^n (Y_i - \beta_0 - \beta_1 X_{i1} - \cdots - \beta_k X_{ik})^2$

**Question:** Should we minimize a similar sum of squares for a *logistic* regression model?

## Motivation: likelihoods and estimation

Let  $Y \sim Bernoulli(p)$  be a Bernoulli random variable, with  $p \in [0, 1]$ . We observe 5 independent samples from this distribution:

$$Y_1 = 1, Y_2 = 1, Y_3 = 0, Y_4 = 0, Y_5 = 1$$

The true value of  $p$  is unknown, so two friends propose different guesses for the value of  $p$ : 0.3 and 0.7. Which do you think is a “better” guess?

## Likelihood

**Definition:** Let  $\mathbf{y} = (Y_1, \dots, Y_n)$  be a sample of  $n$  observations, and let  $f(\mathbf{y}|\theta)$  denote the joint pdf or pmf of  $\mathbf{y}$ , with parameter(s)  $\theta$ . The *likelihood function* is

$$L(\theta|\mathbf{y}) = f(\mathbf{y}|\theta)$$

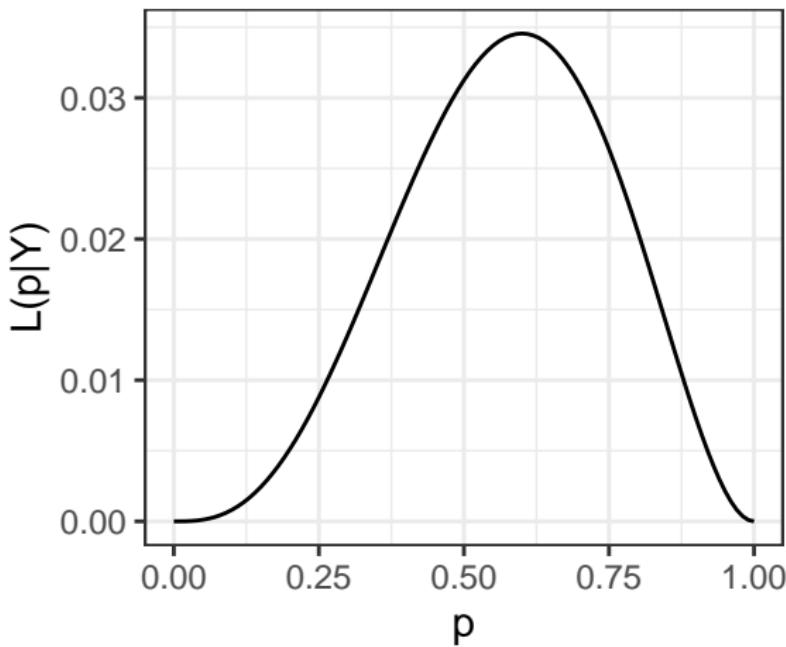
## Example: Bernoulli data

## Example: Bernoulli data

$Y_1, \dots, Y_5 \stackrel{iid}{\sim} \text{Bernoulli}(p)$ , with observed data

$$Y_1 = 1, Y_2 = 1, Y_3 = 0, Y_4 = 0, Y_5 = 1$$

$$L(p|\mathbf{y}) = p^3(1-p)^2$$



## Maximum likelihood estimator

**Definition:** Let  $\mathbf{y} = (Y_1, \dots, Y_n)$  be a sample of  $n$  observations. The *maximum likelihood estimator* (MLE) is

$$\hat{\theta} = \operatorname{argmax}_{\theta} L(\theta|\mathbf{y})$$

Example:  $Bernoulli(p)$

Example:  $N(\theta, 1)$

## Example: $Uniform(0, \theta)$

Let  $Y_1, \dots, Y_n \stackrel{iid}{\sim} Uniform(0, \theta)$ , where  $\theta > 0$ . We want the maximum likelihood estimator of  $\theta$ .

Discuss with your neighbors what the MLE of  $\theta$  might be. *Hint: focus on finding and sketching the likelihood function  $L(\mathbf{y}|\theta)$*