MAXIMUM

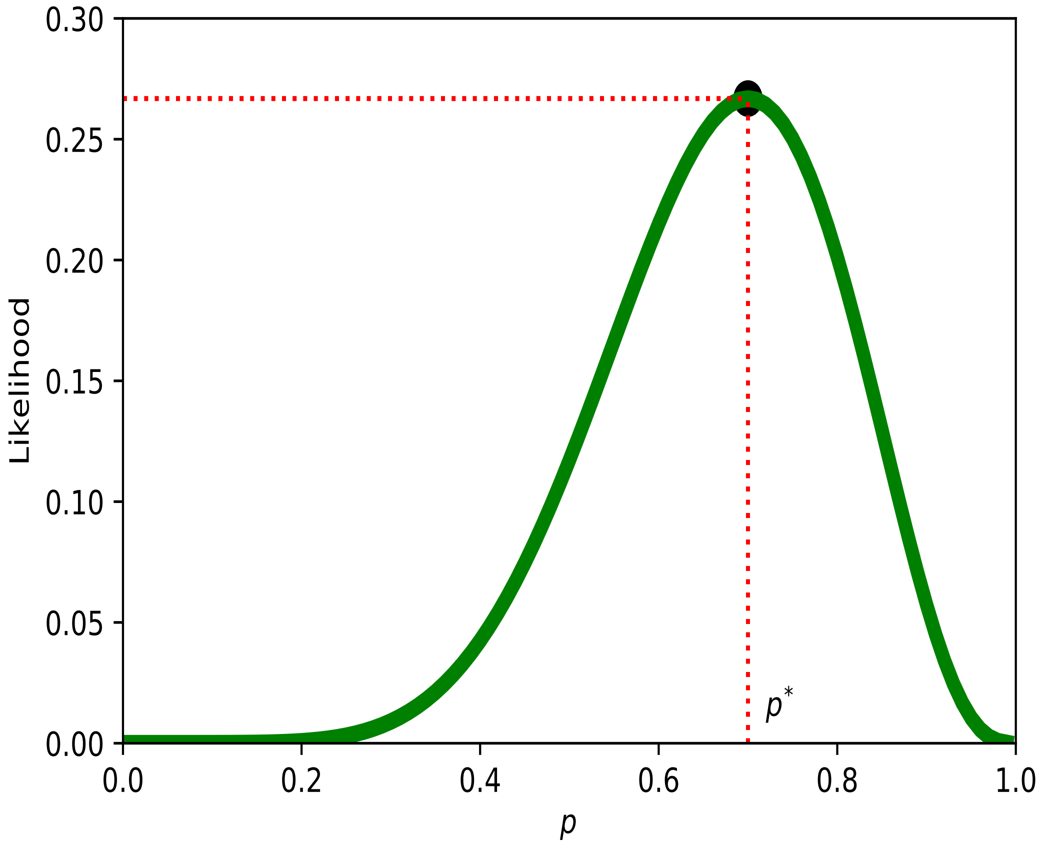
LIKELIHOOD

ESTIMATION

Overview

- find parameters to maximize the likelihood that model produced observed data
- example: toss coin 10 times
- observe 7 heads and 3 tails
- is p different from 0.5?

Solution



- find p to maximize

$$P(7\text{ heads}) = \binom{10}{7} p^7 (1-p)^3$$

Computing p from MLE

p	$\binom{10}{7}p^7(1-p)^3$
0	0
0.1	$8.75 \cdot 10^{-6}$
0.2	$7.86 \cdot 10^{-4}$
0.3	$9.00 \cdot 10^{-3}$
0.4	$4.25 \cdot 10^{-2}$
0.5	$1.17 \cdot 10^{-1}$
0.6	$2.15 \cdot 10^{-1}$
0.7	$2.69 \cdot 10^{-1}$
0.8	$2.01 \cdot 10^{-1}$
0.9	$5.74 \cdot 10^{-2}$
1.0	0

- p around 0.7

Computing p

- easier to max log-likelihood

$$\frac{\partial \log \mathrm{MLE}}{\partial p} = 0$$

$$\frac{\partial \left(7 \log p + 3 \log(1-0)\right)}{\partial p} = 0$$

$$\frac{7}{p} - \frac{3}{1-p} = 0$$

$$p^* = 0.7$$

MLE for Bernouilli

- k successes in n trials

$$f(x_i) = \binom{n}{x_i} p^{x_i}(1-p)^{n-x_i}$$

$$L(p) = \prod_{i=1}^{n} f(x_i) = \prod_{i=1}^{n} \binom{n}{x_i} p^{x_i}(1-p)^{n-x_i}$$

$$\log L(p) = \sum_{i=1}^{n} x_i \log p + \left(1 - \sum_{i=1} x_i\right) \log(1-p)$$

$$\frac{\partial \log L(p)}{\partial p} = \frac{1}{p}\sum_{i=1}^{n} x_i + \frac{1}{1-p}\left(1 - \sum_{i=1}^{n} x_i\right) = 0$$

- this gives

$$p^* = (x_1 + \cdots + x_n)/n = k/n$$

MLE for Poisson

- observations: $x_1, \ldots, x_n$

$$f(x_i) = \frac{\lambda^{x_i} e^{-\lambda}}{x_i!}$$

$$L(\lambda) = \prod_{i=1}^{n} f(x_i) = e^{-n\lambda} \cdot \frac{\lambda(x_1 + \cdots + x_n)}{x_1 x_2 \cdots x_n}$$

$$\log L(\lambda) = -n\lambda + \sum_{i=1}^{n} x_i \log \lambda - \log(x_1 \cdots x_n)$$

$$\frac{\partial \log L(\lambda)}{\partial \lambda} = -n + \frac{1}{\lambda}(x_1 + \cdots x_n) = 0$$

- this gives $\lambda^* = \big(x_1 + \cdots + x_n\big)/n$

MLE for Normal Distribution

$$f(x_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$L(\mu, \sigma) = \prod_{i=1}^{n} f(x_i) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n e^{-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$\log L(\mu, \sigma) = n \log(\sqrt{2\pi}) - n \log \sigma - \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\sigma^2}$$

$$\frac{\partial \log L(\mu, \sigma)}{\partial \mu} = \sum_{i=1}^{n} \frac{(x_i - \mu)}{\sigma^2} = 0$$

$$\mu^* = (x_1 + \cdots + x_n)/n$$

MLE for Normal Distribution

$$\frac{\partial \log L(\mu,\sigma)}{\partial \sigma} = -\frac{n}{\sigma} + \sum_{i=1}^{n} \frac{(x_i - \mu)^2}{\sigma^3} = 0$$

$$\sigma^* = \sum_{i=1}^{n} \frac{(x_i - \mu^*)^2}{n}$$

- MLE for μ and σ^2 are the mean and the variance of observations

Concepts Check:

(a) how is MLE used?

(b) MLE for Bernouilli distribution

(c) MLE for Poisson distribution

(d) MLE for Normal distribution