Parameter Estimation for the Binomial Distribution

C. Durso

Maximum Likelihood Estimation

Given data and the distribution family of the population, find the parameters that maximize the likelihood of the data.

Example

Suppose you flip a possibly biased coin 50 times and you observe 30 heads.

You decide to model this as an outcome from a binomial distribution, $Binomial(n, p)$. Here, n is 50. One way to select p is to select the value that maximizes the probability of the observed data.

Maximize $\binom{50}{30}p^{30}(1-p)^{20}$

- The probability of the observed data under $Binomial(50, p)$ is $\binom{50}{30}p^{30}(1-p)^{20}$.
- To maximize this, note that $p \in [0,1]$ and look for critical points within that interval.
  1. Differentiate: $\frac{d}{dp}\binom{50}{30}p^{30}(1-p)^{20} = \binom{50}{30}30p^{29}(1-p)^{20} - \binom{50}{30}20p^{30}(1-p)^{19}$
  2. Set $\binom{50}{30}30p^{29}(1-p)^{20} - \binom{50}{30}20p^{30}(1-p)^{19} = 0$

Solve for p

$$30p^{29}(1-p)^{20} - 20p^{30}(1-p)^{19} = 0,$$

$$p^{29}(1-p)^{19}(30(1-p) - 20p) = 0$$

$$30 - 50p = 0,$$

$$p = \frac{3}{5}$$

- This is the only critical point, and the values at $p = 0$ and $p = 1$ are smaller than $\binom{50}{30}\frac{3}{5}^{30}\left(1 - \frac{3}{5}\right)^{20}$.
- Conclude $p = \frac{3}{5}$ is the maximum likelihood estimate of p.

Maximum Likelihood Value

Formula

Given k success in n trials as data from an experiment modeled as $Binomial(n, p)$, the maximum likelihood value of value of p equals $\frac{k}{n}$.

Parameter Estimation

Normal Data

C. Durso

Normal Maximum Likelihood

Example

Suppose that you have n mutually independent observations $x_1 \dots x_n$ from $Normal(\mu, \sigma^2)$ with μ and σ^2 unknown. Select the values of μ and σ^2 that maximize the probability density function for $x_1 \dots x_n$.

The Joint Density

The density of the probability distribution for the x's is the product of the one-dimensional densities, with integration taking place in n dimensions. The density at $x_1 \dots x_n$ is

$$\Pi_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)$$

Use the Natural Log

To maximize the density, one can maximize its natural log instead:

$$\Sigma_{i=1}^{n}\left[-\frac{1}{2}\left(ln(2\pi)+ln(\sigma^2)\right)-\frac{(x_i-\mu)^2}{2\sigma^2}\right]$$

Differentiate
Replace σ^2 by v for Convenience

$$\frac{\partial}{\partial v}\left(\Sigma_{i=1}^{n}\left[-\frac{1}{2}\big(ln(2\pi)+ln(v)\big)-\frac{(x_i-\mu)^2}{2v}\right]\right)=\Sigma_{i=1}^{n}\left(-\frac{1}{2}v^{-1}+\frac{(x_i-\mu)^2}{2}v^{-2}\right)$$

$$\frac{\partial}{\partial \mu}\left(\Sigma_{i=1}^{n}\left[-\frac{1}{2}\big(ln(2\pi)+ln(v)\big)-\frac{(x_i-\mu)^2}{2v}\right]\right)=\Sigma_{i=1}^{n}\left(\frac{x_i-\mu}{v}\right)$$

Solve for μ

$$\Sigma_{i=1}^{n}\left(\frac{x_i - \mu}{v}\right) = 0$$

Solve for μ

$$\Sigma_{i=1}^{n}\left(\frac{x_i - \mu}{v}\right) = 0$$

$$\Sigma_{i=1}^{n}(x_i - \mu) = 0$$

Solve for μ

$$\Sigma_{i=1}^{n}\left(\frac{x_i - \mu}{v}\right) = 0$$

$$\Sigma_{i=1}^{n}(x_i - \mu) = 0$$

$$\Sigma_{i=1}^{n}x_i - n\mu = 0$$

Solve for μ

$$\Sigma_{i=1}^{n}\left(\frac{x_i - \mu}{\nu}\right) = 0$$

$$\Sigma_{i=1}^{n}(x_i - \mu) = 0$$

$$\Sigma_{i=1}^{n}x_i - n\mu = 0$$

$$\mu = \frac{\Sigma_{i=1}^{n}x_i}{n} = \bar{x}$$

Solve for v

$$\Sigma_{i=1}^{n}\left(-\frac{1}{2}v^{-1}+\frac{(x_i-\mu)^2}{2}v^{-2}\right)=-\frac{n}{2}v^{-1}+\frac{1}{2}v^{-2}\Sigma_{i=1}^{n}(x_i-\bar{x})^2=0$$

Solve for v

$$\Sigma_{i=1}^{n}\left(-\frac{1}{2}v^{-1}+\frac{(x_i-\mu)^2}{2}v^{-2}\right)=-\frac{n}{2}v^{-1}+\frac{1}{2}v^{-2}\Sigma_{i=1}^{n}(x_i-\bar{x})^2=0$$

$$-nv+\Sigma_{i=1}^{n}(x_i-\bar{x})^2=0$$

Solve for v

$$\Sigma_{i=1}^{n}\left(-\frac{1}{2}v^{-1}+\frac{(x_i-\mu)^2}{2}v^{-2}\right)=-\frac{n}{2}v^{-1}+\frac{1}{2}v^{-2}\Sigma_{i=1}^{n}(x_i-\bar{x})^2=0$$

$$-nv+\Sigma_{i=1}^{n}(x_i-\bar{x})^2=0$$

$$\nu=\frac{\Sigma_{i=1}^{n}(x_i-\bar{x})^2}{n}$$

Maximum Likelihood Values

Theorem

Given n mutually independent observations $x_1 \dots x_n$ from $Normal(\mu, \sigma^2)$ with μ and σ^2 unknown, the values of μ and σ^2 that maximize the probability density function for $x_1 \dots x_n$ are

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