The Method of Maximum Likelihood

James R. Graham

10/24/2006

Maximum Likelihood

- Experiments select a sample from the parent population
  - Suppose we select N_i points from a Gaussian parent distribution, with mean μ and standard deviation, σ
  - The probability of making any single observation, x_i, is

$$P_i = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu}{\sigma} \right)^2 \right]$$

Maximum Likelihood

- We do not know μ or σ *a priori*
  - μ must be derived from the data
  - Denote this estimate μ'
- What expression for μ' gives the ***maximum likelihood*** that the parent population has a particular mean given a set of data?

Using Maximum Likelihood to Estimate the Mean

- Suppose the parent population has a mean μ' and constant standard deviation $\sigma' = \sigma$
  - The probability of observing the i-th point x_i is

$$P_i(\mu') = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x_i - \mu'}{\sigma}\right)^2\right]$$

Estimating μ

- Consider all N observations
  - The probability for observing that set is the product of the individual $P_i(\mu')$

$$P(\mu') = \prod_{i=1}^{N} P_i(\mu')$$

$$= \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \exp\left[-\tfrac{1}{2}\sum_{i=1}^{N}\left(\frac{x_i - \mu'}{\sigma}\right)^2\right]$$

Estimating μ

- According to the method of ***maximum likelihood*** if we compare the $P(\mu')$ for various parent populations with different μ' (all with the same σ)
- The probability is greatest that the data were derived from a population with $\mu'=\mu$
  - We assert that the ***most likely*** parent population is the correct one

Calculating the Mean

- Using maximum likelihood the most probable value of μ' is the one which gives the maximum probability, $P(\mu')$
  - Maximize

$$P(\mu') = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^N \exp\left[-\tfrac{1}{2}\sum\left(\frac{x_i - \mu'}{\sigma}\right)^2\right]$$

or minimize, X

$$X = -\tfrac{1}{2}\sum\left(\frac{x_i - \mu'}{\sigma}\right)^2$$

Calculating the Mean

- Find the minimum of X by computing the derivative

$$\frac{\partial X}{\partial \mu'} = -\frac{1}{2}\frac{\partial}{\partial \mu'}\sum\left(\frac{x_i - \mu'}{\sigma}\right)^2 = 0$$

$$= -\frac{1}{2}\sum\frac{\partial}{\partial \mu'}\left(\frac{x_i - \mu'}{\sigma}\right)^2 = \sum\left(\frac{x_i - \mu'}{\sigma}\right) = 0$$

since the derivative of a sum is the sum of the derivatives

Calculating the Mean

- The most probable value for the mean is given by

$$\sum (x_i - \mu') = 0$$

$$\sum x_i - \sum \mu' = 0$$

$$\boxed{\mu' = \frac{1}{N} \sum x_i}$$

Weighting Data

- We assumed that all the data were from the same parent population
  - Suppose some measurements are better than others, some values are drawn from a population with smaller σ_i
- Maximize

$$P(\mu') = \prod_{i=1}^{N} \left(\frac{1}{\sigma_i \sqrt{2\pi}} \right) \exp\left[-\tfrac{1}{2} \sum \left(\frac{x_i - \mu'}{\sigma_i} \right)^2 \right]$$

The Weighted Mean

- Maximizing the probability is equivalent to minimizing the argument in the exponential

$$-\frac{1}{2}\frac{\partial}{\partial\mu'}\sum\left(\frac{x_i-\mu'}{\sigma_i}\right)^2=\sum\frac{x_i-\mu'}{\sigma_i^2}=0$$

$$\boxed{\mu'=\frac{\sum w_i x_i}{\sum w_i},\quad w_i=1/\sigma_i^2}$$

- The most probable value of the mean is the ***weighted*** (inversely by the variance) mean

How to Fit a Straight Line

- Suppose our data, y_i, are drawn from a population such that

$$y(x) = a_0 + b_0 x$$

- For any x_i we can calculate the probability of making the observation y_i as

$$P_i = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left[-\tfrac{1}{2} \left\{ \frac{y_i - y(x_i)}{\sigma_i} \right\}^2 \right]$$

Straight Line Fit

- The probability for making the observed set of measurements is the product

$$P(a_0, b_0) = \prod_{i=1}^{N} P_i$$

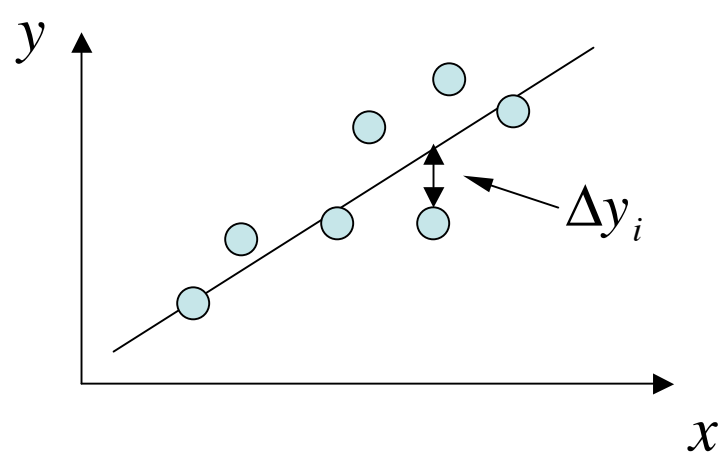
$$= \prod \left(\frac{1}{\sigma_i \sqrt{2\pi}} \right) \exp\left[-\tfrac{1}{2} \sum \left(\frac{y_i - y(x_i)}{\sigma_i} \right)^2 \right]$$

Straight Line Fit

- Similarly, the probability for making the observed set of measurements given coefficients, *a* and *b* is

$$P(a,b) = \prod \left(\frac{1}{\sigma_i \sqrt{2\pi}} \right) \exp\left[-\tfrac{1}{2} \sum \left(\frac{\Delta y_i}{\sigma_i} \right)^2 \right]$$

$$\Delta y_i = y_i - a - b x_i$$



Straight Line Fit

- The product term is a constant, independent of *a* and *b*
  - Maximizing $P(a, b)$ is equivalent to minimizing the sum of the exponential

$$\chi^2 \equiv \sum \left(\frac{\Delta y_i}{\sigma_i} \right)^2$$

$$= \sum \frac{1}{\sigma_i^2} \left(y_i - a - b x_i \right)^2$$

Minimizing χ^2

- To find *a* and *b* which corresponds to the minimum χ^2

$$\frac{\partial}{\partial a}\chi^2 = \frac{\partial}{\partial a}\left[\frac{1}{\sigma^2}\sum\left(y_i - a - bx_i\right)^2\right]$$

$$= -\frac{2}{\sigma^2}\sum\left(y_i - a - bx_i\right) = 0$$

$$\frac{\partial}{\partial b}\chi^2 = \frac{\partial}{\partial b}\left[\frac{1}{\sigma^2}\sum\left(y_i - a - bx_i\right)^2\right]$$

$$= -\frac{2}{\sigma^2}\sum x_i\left(y_i - a - bx_i\right) = 0$$

Minimizing χ^2

- These can be rearranged to find pair of simultaneous equations for *a* and *b* which corresponds to the minimum χ^2

$$\sum y_i = aN + b\sum x_i$$

$$\sum x_i y_i = a\sum x_i + b\sum x_i^2$$

Minimizing χ^2

- Solving these of simultaneous equations

$$a = \frac{1}{\Delta}\begin{vmatrix} \sum y_i & \sum x_i \\ \sum x_i y_i & \sum x_i^2 \end{vmatrix}$$

$$b = \frac{1}{\Delta}\begin{vmatrix} N & \sum y_i \\ \sum x_i & \sum x_i y_i \end{vmatrix}$$

$$\Delta = \begin{vmatrix} N & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}$$