likelihood function

Canonical name	LikelihoodFunction
Date of creation	2013-03-22 14:27:58
Last modified on	2013-03-22 14:27:58
Owner	CWoo (3771)
Last modified by	CWoo (3771)
Numerical id	13
Author	CWoo (3771)
Entry type	Definition
Classification	msc 62A01
Synonym	likelihood statistic
Synonym	likelihood
Defines	maximum likelihood estimate
Defines	MLE
Defines	log-likelihood function

Let $\mathbf{X}$=$(X_1, \ldots, X_n)$ be a random vector and

$$\{f_{\mathbf{X}}(\boldsymbol{x} \mid \boldsymbol{\theta}) : \boldsymbol{\theta} \in \Theta\}$$

a statistical model parametrized by $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)$, the parameter vector in the *parameter space* Θ. The *likelihood function* is a map $L : \Theta \to \mathbb{R}$ given by

$$L(\boldsymbol{\theta} \mid \boldsymbol{x}) = f_{\mathbf{X}}(\boldsymbol{x} \mid \boldsymbol{\theta}).$$

In other words, the likelikhood function is functionally the same in form as a probability density function. However, the emphasis is changed from the $\boldsymbol{x}$ to the $\boldsymbol{\theta}$. The pdf is a function of the x's while holding the parameters θ's constant, L is a function of the parameters θ's, while holding the x's constant.

When there is no confusion, $L(\boldsymbol{\theta} \mid \boldsymbol{x})$ is abbreviated to be $L(\boldsymbol{\theta})$.

The parameter vector $\hat{\boldsymbol{\theta}}$ such that $L(\hat{\boldsymbol{\theta}}) \geq L(\boldsymbol{\theta})$ for all $\boldsymbol{\theta} \in \Theta$ is called a *maximum likelihood estimate*, or *MLE*, of $\boldsymbol{\theta}$.

Many of the density functions are exponential in nature, it is therefore easier to compute the MLE of a likelihood function L by finding the maximum of the natural log of L, known as the log-likelihood function:

$$\ell(\boldsymbol{\theta} \mid \boldsymbol{x}) = \ln(L(\boldsymbol{\theta} \mid \boldsymbol{x}))$$

due to the monotonicity of the log function.

Examples:

1. A coin is tossed n times and m heads are observed. Assume that the probability of a head after one toss is π. What is the MLE of π?

   Solution: Define the outcome of a toss be 0 if a tail is observed and 1 if a head is observed. Next, let X_i be the outcome of the ith toss. For any single toss, the density function is $\pi^x(1-\pi)^{1-x}$ where $x \in \{0, 1\}$. Assume that the tosses are independent events, then the joint probability density is

   $$f_{\mathbf{X}}(\boldsymbol{x} \mid \pi) = \binom{n}{\Sigma x_i} \pi^{\Sigma x_i}(1-\pi)^{\Sigma(1-x_i)} = \binom{n}{m} \pi^m (1-\pi)^{n-m},$$

   which is also the likelihood function $L(\pi)$. Therefore, the log-likelihood function has the form

   $$\ell(\pi \mid \boldsymbol{x}) = \ell(\pi) = \ln \binom{n}{m} + m\ln(\pi) + (n-m)\ln(1-\pi).$$

Using standard calculus, we get that the MLE of π is

$$\hat{\pi} = \frac{m}{n} = \overline{x}.$$

2. Suppose a sample of n data points X_i are collected. Assume that the $X_i \sim N(\mu, \sigma^2)$ and the X_i's are independent of each other. What is the MLE of the parameter vector $\boldsymbol{\theta} = (\mu, \sigma^2)$?

   Solution: The joint pdf of the X_i, and hence the likelihood function, is

   $$L(\boldsymbol{\theta} \mid \boldsymbol{x}) = \frac{1}{\sigma^n (2\pi)^{n/2}} \exp(-\frac{\Sigma(x_i - \mu)^2}{2\sigma^2}).$$

   The log-likelihood function is

   $$\ell(\boldsymbol{\theta} \mid \boldsymbol{x}) = -\frac{\Sigma(x_i - \mu)^2}{2\sigma^2} - \frac{n}{2}\ln(\sigma^2) - \frac{n}{2}\ln(2\pi).$$

   Taking the first derivative (gradient), we get

   $$\frac{\partial \ell}{\partial \boldsymbol{\theta}} = (\frac{\Sigma(x_i - \mu)}{\sigma^2}, \frac{\Sigma(x_i - \mu)^2}{2\sigma^4} - \frac{n}{2\sigma^2}).$$

   Setting

   $$\frac{\partial \ell}{\partial \boldsymbol{\theta}} = \mathbf{0} \text{ See score function}$$

   and solve for $\boldsymbol{\theta} = (\mu, \sigma^2)$ we have

   $$\hat{\boldsymbol{\theta}} = (\hat{\mu}, \hat{\sigma}^2) = (\overline{x}, \frac{n-1}{n} s^2),$$

   where $\overline{x} = \Sigma x_i / n$ is the sample mean and $s^2 = \Sigma(x_i - \overline{x})^2/(n-1)$ is the sample variance. Finally, we verify that $\hat{\boldsymbol{\theta}}$ is indeed the MLE of $\boldsymbol{\theta}$ by checking the negativity of the 2nd derivatives (for each parameter).