

Chapter XXII

Parameter Estimation

Karl-Ernst Biebler

Ernst-Moritz-Arndt-University, Germany

ABSTRACT

Parameters are numbers which characterize random variables. They make possible the summarizing description of the observations, serve as the basis of statistical decisions and are calculated from the data. Point estimations and confidence estimations are introduced. Samples of the observed random variable are a starting point. The maximum-likelihood method for the construction of parameter estimations is introduced here. Examples concern the normal distributions and the binomial distributions. Approximate methods of the parameter estimation also can be too inaccurate at large sample sizes. This is demonstrated in an example from genetics.

INTRODUCTION

Statistical parameter estimation is a standard task in most data-mining procedures. It is presupposed that the data are a sample of the interesting random variable.

Parameters, unknown values of characterizing numbers concerning the observed random variable, can be calculated from the data. Such a calculation is called an estimate if it is carried out following statistical principles.

Examples of parameters are the expectation, the standard deviation, the median, quantiles, and so forth of a random variable. Furthermore,

parameters can be components of data models, for example, the numbers α and β in the simple model $Y = \alpha X + \beta$.

The calculation of an approximate value for the interesting parameter from the observed data is called point estimation. A confidence estimation delivers a certain region in which the interesting parameter is contained with given (high) probability. It reflects the information content of the data with respect to the parameter and the sample size better than point estimation.

Parameter-estimation procedures are developed following special principles and consid-

Parameter Estimation

ering the distribution of the random variables. They should fulfill quality requirements.

POINT ESTIMATION

Definition: Let X be a random variable, $F_{X,\alpha}(x)$ its distribution function, and (X_1, \dots, X_N) a sample of size N . A point estimator $\hat{\alpha}_N$ of the parameter α is a function $\hat{\alpha}_N = \hat{\alpha}_N(x_1, \dots, x_N)$ of the sample. A point estimation is the value of $\hat{\alpha}_N$ at given data (x_1, \dots, x_N) .

The most essential methods for the construction of point estimators are based on the maximum-likelihood principle, the Bayesian principle, the moments principle, the minimum-chi-square principle, or the least squares principle. The last named is not a statistical principle but a geometric one. Its relation to the maximum-likelihood principle is the clue of the Gauss-Markov theorem in mathematical statistics. Only one method is introduced here.

Maximum-Likelihood Estimation

The idea of the maximum-likelihood principle consists in the choice of the parameter value so that the observation gets maximal probability.

Definition: Let the parameter be $\forall 0 < \alpha < 1$, X be a continuous random variable with probability density $f_{X,\alpha}(x)$ depending on α and (x_1, \dots, x_N) be the sample data. Then:

$$L(x_1, \dots, x_N, \alpha) = \prod_{i=1}^N f_{X,\alpha}(x_i)$$

is called the likelihood function of (x_1, \dots, x_N) and α .

For a discrete random variable, the input is:

$$L(x_1, \dots, x_N, \alpha) = \prod_{i=1}^N P(X = x_i).$$

The maximum-likelihood estimator MLE $\hat{\alpha}_{MLE}$ calculates α_{MLE} from the given data so that $L(x_1, \dots, x_N, \alpha_{MLE})$ is a maximum of the likelihood function.

Examples

Normal Distribution

Let X be a normal distribution $X \approx N(\mu, \sigma^2)$ with expectation $E(X) = \mu$ and variance $V(X) = \sigma^2$. Then the parameter $\alpha = (\mu, \sigma^2)$ is an element of a two-dimensional set. The MLE of the expectation is the sample mean:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i,$$

and the MLE of the variance is:

$$s_{MLE}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2.$$

Binomial Distribution

For a binomial-distributed random variable, the relative frequency of the event in a sample of size N is an MLE of its unknown probability.

The One-Locus-Two-Allele Model in Population Genetics

A two-allele model regarding the alleles A_1 and A_2 , and $P(A_1) = p$ and $P(A_2) = 1 - p$, without any dominance relations, is described by the probabilistic model

$$M_1 = \left[\begin{array}{l} \{(A_1 A_1), (A_1 A_2), (A_2 A_2)\}; \\ P(A_1 A_1) = p^2, P(A_1 A_2) = 2p(1-p), P(A_2 A_2) = (1-p)^2 \end{array} \right]$$

The phenotypes observed are exactly the genotypes $A_i A_j$. The maximum-likelihood estimator for p based on a sample of size N with

7 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/parameter-estimation/20577

Related Content

Automated Analysis of Nursery School Observations

Jien Kato (2013). *Technological Advancements in Biomedicine for Healthcare Applications* (pp. 278-286).

www.irma-international.org/chapter/automated-analysis-nursery-school-observations/70870

Harnessing Deep Neural Architectures for Hyperspectral Imaging in Medicine: Recent Innovations and Future Directions

V. Vanitha, S. K. Mohanaveland K. Saathvick (2026). *Biomedical Applications in Deep Learning-Enhanced Hyperspectral Imaging* (pp. 119-154).

www.irma-international.org/chapter/harnessing-deep-neural-architectures-for-hyperspectral-imaging-in-medicine/406619

Elimination of Power Line Interference in ECG Signal Using Adaptive Filter, Notch Filter and Discrete Wavelet Transform Techniques

Srinivasa M.G.and Pandian P.S. (2019). *International Journal of Biomedical and Clinical Engineering* (pp. 32-56).

www.irma-international.org/article/elimination-of-power-line-interference-in-ecg-signal-using-adaptive-filter-notch-filter-and-discrete-wavelet-transform-techniques/219305

An Overview of Telemedicine Technologies for Healthcare Applications

P. S. Pandian (2016). *International Journal of Biomedical and Clinical Engineering* (pp. 29-52).

www.irma-international.org/article/an-overview-of-telemedicine-technologies-for-healthcare-applications/170460

On-body UWB Communications for Health Monitoring

Qiong Wangand Jianqing Wang (2012). *Neonatal Monitoring Technologies: Design for Integrated Solutions* (pp. 125-147).

www.irma-international.org/chapter/body-uw-b-communications-health-monitoring/65267