Topic 2-1: Likelihood Construction & Estimation

Univariate Models

EXST 7160
Department of Experimental Statistics
Louisiana State University

July 25, 2021

Content

Contents

Contents

Statistical Models

- A statistical model is a general functional relation between the unknown parameter(s) and the observed data.

Statistical Models

- A statistical model is a general functional relation between the unknown parameter(s) and the observed data.
- After a statistical model for the observed data has been formulated, the likelihood function of the data is the natural starting point for the inference in many statistical problems.

Statistical Models

- A statistical model is a general functional relation between the unknown parameter(s) and the observed data.
- After a statistical model for the observed data has been formulated, the likelihood function of the data is the natural starting point for the inference in many statistical problems.
- The likelihood function typically leads to essentially automatic methods of inference, including point estimation, interval estimation, and hypothesis testing.

Statistical Models

- A statistical model is a general functional relation between the unknown parameter(s) and the observed data.
- After a statistical model for the observed data has been formulated, the likelihood function of the data is the natural starting point for the inference in many statistical problems.
- The likelihood function typically leads to essentially automatic methods of inference, including point estimation, interval estimation, and hypothesis testing.
- In this topic, we will focus on constructing the likelihood functions from various types of data, including discrete, continuous, mixture of discrete and continuous, and censored data.

Statistical Models

- The likelihood is the joint density of the observed data to be analyzed.

Statistical Models

- The likelihood is the joint density of the observed data to be analyzed.
- Let the random variables $Y_1, \cdots, Y_n$ have a joint density function $f(\mathbf{Y} = (Y_1, \cdots, Y_n)^T; \boldsymbol{\theta})$ with unknown b density parameters $\boldsymbol{\theta} = (\theta_1, \cdots, \theta_b)$. Then, given observed data $\mathbf{Y} = \mathbf{y}$, where $\mathbf{y} \equiv (y_1, \cdots, y_n)^T$, the function of $\boldsymbol{\theta}$

$$L(\boldsymbol{\theta}; \mathbf{y}) = f(\mathbf{Y} = \mathbf{y}; \boldsymbol{\theta})$$

is the likelihood function.

Likeliihod for iid data

- If the random variables $Y_1, \cdots, Y_n$ are indenpendent, then the likelihood function becomes

$$L(\boldsymbol{\theta}; \mathbf{y}) = \prod_{i=1}^{n} f_i(Y_i = y_i; \boldsymbol{\theta}),$$

where $f_i(Y_i; \boldsymbol{\theta})$ is the density of Y_i.

Likeliihod for iid data

- If the random variables $Y_1, \cdots, Y_n$ are indenpendent, then the likelihood function becomes

$$L(\boldsymbol{\theta}; \mathbf{y}) = \prod_{i=1}^{n} f_i(Y_i = y_i; \boldsymbol{\theta}),$$

where $f_i(Y_i; \boldsymbol{\theta})$ is the density of Y_i.

- If the random variables $Y_1, \cdots, Y_n$ are indenpendent and indentically distributed (denoted by iid), then the likelihood function becomes

$$L(\boldsymbol{\theta}; \mathbf{y}) = \prod_{i=1}^{n} f(Y_i = y_i; \boldsymbol{\theta}),$$

where f is the distribution that all $Y_1, \cdots, Y_n$ follow.

Solving the Likelihood

- The value of $\boldsymbol{\theta}$ that maximizes the likelihood function $L(\boldsymbol{\theta}; \mathbf{y})$ is called the maximum likelihood estimator (MLE) denoted by $\hat{\boldsymbol{\theta}}_{MLE}$.

Solving the Likelihood

- The value of $\boldsymbol{\theta}$ that maximizes the likelihood function $L(\boldsymbol{\theta}; \mathbf{y})$ is called the maximum likelihood estimator (MLE) denoted by $\hat{\boldsymbol{\theta}}_{MLE}$.
- In practice, $\hat{\boldsymbol{\theta}}_{MLE}$ is usually calculated by optimizing the log likelihood function $\log(L(\boldsymbol{\theta}; \mathbf{y}))$

Solving the Likelihood

- The value of $\boldsymbol{\theta}$ that maximizes the likelihood function $L(\boldsymbol{\theta}; \mathbf{y})$ is called the maximum likelihood estimator (MLE) denoted by $\hat{\boldsymbol{\theta}}_{MLE}$.
- In practice, $\hat{\boldsymbol{\theta}}_{MLE}$ is usually calculated by optimizing the log likelihood function $\log(L(\boldsymbol{\theta}; \mathbf{y}))$
- Note: $\hat{\boldsymbol{\theta}}_{MLE}$ is "optimal" under some regularity conditions (to be discussed in a future topic).

Solving the Likelihood-differentiable likelihoods

- Denote the log likelihood function $\log(L(\boldsymbol{\theta}; \mathbf{y}))$ by $\ell(\boldsymbol{\theta})$, and assume $\ell(\boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$.

Solving the Likelihood-differentiable likelihoods

- Denote the log likelihood function $\log(L(\boldsymbol{\theta}; \mathbf{y}))$ by $\ell(\boldsymbol{\theta})$, and assume $\ell(\boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$.
- The procedure for obtaining the MLE from $\ell(\boldsymbol{\theta})$ is

Solving the Likelihood-differentiable likelihoods

- Denote the log likelihood function $\log(L(\boldsymbol{\theta};\mathbf{y}))$ by $\ell(\boldsymbol{\theta})$, and assume $\ell(\boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$.
- The procedure for obtaining the MLE from $\ell(\boldsymbol{\theta})$ is
  1. Differentiate $\ell(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ to obtain the likelihood score function $S(\boldsymbol{\theta}) = (\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_b})^T$.

Solving the Likelihood-differentiable likelihoods

- Denote the log likelihood function $\log(L(\boldsymbol{\theta}; \mathbf{y}))$ by $\ell(\boldsymbol{\theta})$, and assume $\ell(\boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$.
- The procedure for obtaining the MLE from $\ell(\boldsymbol{\theta})$ is
  1. Differentiate $\ell(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ to obtain the likelihood score function $S(\boldsymbol{\theta}) = (\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_b})^T$.
  2. Find possible MLE cadidates by solving the likelihood equations $S(\boldsymbol{\theta}) = \mathbf{0}_{b\times 1}$.

Solving the Likelihood-differentiable likelihoods

- Denote the log likelihood function $\log(L(\boldsymbol{\theta}; \mathbf{y}))$ by $\ell(\boldsymbol{\theta})$, and assume $\ell(\boldsymbol{\theta})$ is differentiable with respect to $\boldsymbol{\theta}$.
- The procedure for obtaining the MLE from $\ell(\boldsymbol{\theta})$ is
  1. Differentiate $\ell(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ to obtain the likelihood score function $S(\boldsymbol{\theta}) = (\frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_1}, \cdots, \frac{\partial \ell(\boldsymbol{\theta})}{\partial \theta_b})^T$.
  2. Find possible MLE cadidates by solving the likelihood equations $S(\boldsymbol{\theta}) = \mathbf{0}_{b \times 1}$.
  3. Check if the solution from the last step is the global maximizer of $\ell(\boldsymbol{\theta})$. If it is, then the solution is the MLE of $\boldsymbol{\theta}$.

An R Example

-

Remark

- A non-differentiable likelihood example can be found in Example 7.2.9 of Casella and Berger (2002).

Remark

- A non-differentiable likelihood example can be found in Example 7.2.9 of Casella and Berger (2002).
- An example for checking if the solution of the score function equation with two unknown parameters is a global maximizer can be found in Example 7.2.12 of Casella and Berger (2002).

Contents

Discrete IID Random Variables

Discrete IID Random Variables

Discrete IID Random Variables

Continuous IID Variables

Continuous IID Variables

Continuous IID Variables

-
-
- References: More examples for constructing the product likelihood likelihood associated with iid data can be found in Section 7.2.2 of Casella and Berger (2002).

Multinomial

Multinomial

Multinomial

-
-
- References: Examples for

Multivariate Normal

-

Multivariate Normal

-
-

Multivariate Normal

Mixture of discrete and continuous

Mixture of discrete and continuous

-
-

Mixture of discrete and continuous

Multinomial

Multinomial

Multinomial

-
-
- References: Examples for

Contents

Proportional Likelihood

-

Proportional Likelihood

Proportional Likelihood

Empirical Likelihood

Empirical Likelihood

Empirical Likelihood

Likelihood With Censored data

-

Likelihood With Censored data

Likelihood With Censored data

Contents

A general working defintion of the likelihood

-

A general working defintion of the likelihood

-
-

A general working defintion of the likelihood

Empirical Likelihood

Empirical Likelihood

Empirical Likelihood

Likelihood With Censored data

-

Likelihood With Censored data

Likelihood With Censored data

-
-
-