Inference 2, 2023, lecture 6

Rolf Larsson

November 17, 2023

Today

Chap. 4. Estimation:

- Estimator
- Method of moments
- Maximum likelihood
- M estimators

Estimator

- Suppose $\mathbf{X} \sim \mathrm{P}_\theta \in \mathcal{P}$ where $\theta \in \Theta$.
- Our main interest is in $\gamma = g(\theta)$, $g : \theta \to \Gamma$. (Special case: $\gamma = \theta$.)
- Given data $\mathbf{x}$, what is a plausible value of γ?

Definition (4.1)

A function $T : \mathcal{X} \to \Gamma$ is an **estimator**. It is used to estimate $\gamma = g(\theta)$.
The value $T(\mathbf{x})$ is called the **estimate** of $g(\theta)$.
It is the realization of the random variable $T(\mathbf{X})$.

Example 1:

- Suppose $X_1, ..., X_n$ are independent, distributed as $X \sim \mathrm{Bernoulli}(p)$.
- We observe $x_1, ..., x_n$.
- Suggest an estimator of p.

Method of moments

- Let $\mathbf{X} = (X_1, ..., X_n)$ where the X_i are independent and distributed as X with distribution F and parameter(s) θ.
- If X is a continuous random variable, let the jth moment be $m_j(F) = E(X^j) = \int x^j f(x)dx$ where f is the density function.
- Similarly, if X is discrete, $m_j(F) = E(X^j) = \sum x^j p(x)$, where p is the probability function.
- Suppose that (at least) r moments of X exist.
- Suppose there is a function h such that $\gamma = h\{m_1(F), ..., m_r(F)\}$

Definition (4.2)

The **moment estimate** for γ is defined by

$$\hat{\gamma}_{\mathrm{MME}} = h(\hat{m}_1, ..., \hat{m}_r),$$

where $\hat{m}_j = \frac{1}{n}\sum_{i=1}^{n} x_i^j$ is the empirical moment of order j.

Method of moments

Example 2:

- Let $\mathbf{X} = (X_1, ..., X_n)$ where the X_i are independent and distributed as X.
- Write down the moment estimates of the parameters if X is distributed as

1. Exponential with intensity β
2. $N(\mu, \sigma^2)$
3. Uniform on $[0, \theta]$

Maximum Likelihood Estimators

- Let $\mathbf{X} = (X_1, ..., X_n)$ where the X_i are independent with distributions F_i and parameter(s) θ.
- Let $\mathbf{x} = (x_1, ..., x_n)$ be the corresponding observations.
- Define the likelihood function

$$L(\theta; \mathbf{x}) = \prod_{i=1}^{n} f_i(x_i; \theta),$$

in the continuous case, where f_i is the density function of X_i, and similarly in the discrete case.

Definition (4.3)

An estimator T is called the **maximum likelihood estimator (MLE)** of θ if

$$L\{T(\mathbf{x}); \mathbf{x}\} = \max_{\theta \in \Theta} L(\theta; \mathbf{x})$$

for all $\mathbf{x} \in \mathcal{X}$.

Maximum Likelihood Estimators

Theorem (4.1)

If $\gamma = g(\theta)$ and g is bijective, i.e. $\theta = g^{-1}(\gamma)$, then $\hat{\theta}$ is a MLE for θ if.f. $\hat{\gamma} = g(\hat{\theta})$ is a MLE for γ.

Maximum Likelihood Estimators

Definition (4.4)

The MLE of $\gamma = g(\theta)$ is defined by $\hat{\gamma}_{\mathrm{MLE}} = g(\hat{\theta}_{\mathrm{MLE}})$.

Example 3:

- Let $\mathbf{X} = (X_1, ..., X_n)$ where the X_i are independent and exponential with intensity β.
- Let $\mathbf{x} = (x_1, ..., x_n)$ be the corresponding observations.

1. Calculate the MLE of β.
2. Calculate the MLE of $\mu = 1/\beta$
   a. by maximizing the likelihood w.r.t. μ.
   b. By using 1. and theorem 4.1 (definition 4.4).

Maximum Likelihood Estimators

Example 4:

- Let $\mathbf{X} = (X_1, ..., X_n)$ where the X_i are independent and uniform on $(0, \theta)$.
- Let $\mathbf{x} = (x_1, ..., x_n)$ be the corresponding observations.
- Calculate the MLE of θ.

M estimators

- Suppose we have a sample of *i.i.d.* random variables with expectation θ. We want to estimate θ.
- The least squares estimate (LSE) is given by minimizing $\sum_{i=1}^n (x_i - \theta)^2$,
  i.e. by solving $0 = \sum_{i=1}^n (x_i - \theta)$ which yields the solution $\theta = \bar{x}$.
- The least absolute value estimate is given by minimizing $\sum_{i=1}^n |x_i - \theta|$, i.e. by solving

$$0 = -\sum_{x_i < \theta} I\{x_i - \theta < 0\} + \sum_{x_i > \theta} I\{x_i - \theta > 0\} = \sum_{i=1}^n \mathrm{sign}(x_i - \theta),$$

  which yields the median of the x_i as solution.
- The MLE is given by minimizing $-\sum_{i=1}^n l(\theta; x_i)$, i.e. (under regularity conditions) by solving $0 = \sum_{i=1}^n l'(\theta; x_i)$.

M estimators

In general:

Definition (4.5)

For a sample $X_1, \ldots, X_n$ of i.i.d. random variables **the M-estimator** $\hat{\gamma}$ w.r.t. a function $\psi : \mathcal{R} \times \Gamma \to \mathcal{R}$ is defined as the solution of

$$\sum_{i=1}^{n} \psi(X_i, \gamma) = 0.$$

Typically, $\hat{\gamma}$ minimizes a function $\sum_{i=1}^{n} \varrho(X_i, \gamma)$.

- the least squares estimator: $\psi(x_i, \gamma) = x_i - \gamma$, $\varrho(x_i, \gamma) = (x_i - \gamma)^2$.
- the least absolute value estimator: $\psi(x_i, \gamma) = \mathrm{sign}(x_i - \gamma)$, $\varrho(x_i, \gamma) = |x_i - \gamma|$.
- In these examples, the functions may be expressed as $\psi(x_i, \gamma) = \tilde{\psi}(x_i - \gamma)$, $\varrho(x_i, \gamma) = \tilde{\varrho}(x_i - \gamma)$.

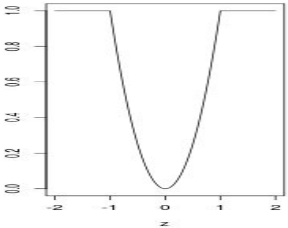
M estimators

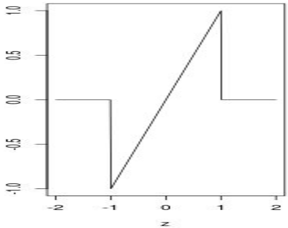
For the LSE, $\tilde{\varrho}(z) = z^2$, $\tilde{\psi}(z) = z$.

The following robust modifications of the LSE are M estimators:

The trimmed mean:

$$\tilde{\varrho}(z) = \begin{cases} z^2 & \text{if } |z| \le k, \\ k^2 & \text{if } |z| > k, \end{cases} \qquad \tilde{\psi}(z) = \begin{cases} z & \text{if } |z| \le k, \\ 0 & \text{if } |z| > k, \end{cases}$$

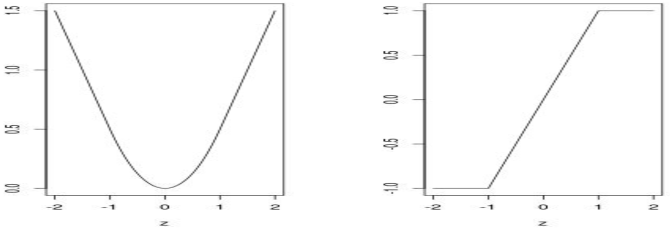




M estimators

The Winzorized mean

$$\tilde{\varrho}(z) = \begin{cases} \frac{1}{2} z^2 & \text{if } |z| \le k, \\ k|z| - \frac{1}{2}k^2 & \text{if } |z| > k, \end{cases} \qquad \tilde{\psi}(z) = \begin{cases} -k & \text{if } z < -k, \\ z & \text{if } |z| \le k, \\ k & \text{if } z > k. \end{cases}$$



M estimators

If there is more than one solution to $\sum_{i=1}^{n} \psi(x_i, \gamma) = 0$, choose the one(s) with smallest value on $\sum_{i=1}^{n} \varrho(x_i, \gamma)$.

Example 5:

- Suppose we have a random sample $1, 2, 3, 10$ from a distribution with expectation θ.
- Estimate θ by using

1. the LSE.
2. the least absolute value.
3. the trimmed mean with $k = 4$.
4. the Winzorized mean with $k = 4$.

News of today

- Estimator
- Estimate
- Method of moments
- MLE
- M estimators
    - LSE
    - Least absolute value
    - Trimmed mean
    - Winzorized mean
    - The MLE in regular cases