Inverse Problems in Statistics
Chapter 4: Some Methods of Estimation

Ghazale Jazayeri, Alessia Fenaroli

7. Mai 2007

1 Plug-in Estimators

The plug-in estimator $\widetilde{\theta}_n$ of $\theta = Q(F)$ is defined by $Q(F_n)$, where

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{\{X_i = x\}}$$

is the empirical distribution function, X_i are indipendent identically F-distributed.

1.1 Wicksell's problem I

We observe a sample from the density g.
The object of interest is the distribution function F, which can be expressed as :

$$F(x) = 1 - \frac{V(x)}{V(0)}$$

where

$$V(x) = \int_x^\infty (z - x)^{-\frac{1}{2}} g(z) \mathrm{d}z$$

Plug-in estimator for V:

$$\widetilde{V_n}(x) = \frac{1}{n} \sum_{i=1}^{n} (Z_i - x)^{-\frac{1}{2}} 1_{(x,\infty)}(Z_i)$$

2 Isotonic inverse estimators

With isotonic inverse estimators we can correct the nonmonotonicity.
Method 1: first the plug-in estimator is computed and afterwards the estimator is projected on the set of monotone functions.
Method 2: project the nonparametric estimator of the sampling distribution on the set of possible sampling distributions, then the inversion of this projected estimator yields a distribution function as estimator.

Lemma 2.1 *Let $0 < M < \infty$ and ϕ some a.e. continuous nonnegative integrable function on $[0, M]$. Denote by Φ the primitive*

$$\Phi(x) = \int_0^x \phi(y)\mathrm{d}y$$

of ϕ on $[0, M]$. Let Φ^* be the least concave majorant of Φ on $[0, M]$, and ϕ^* its right derivative. If ϕ^* is buonded, then, for each bounded decreasing right continuous function σ on $[0, M]$,

$$\int_0^M (\sigma^2(x) - 2\sigma(x)\phi(x))\mathrm{d}x \geq \int_0^M (\phi^*(x)^2 - 2\phi^*(x)\phi(x))\mathrm{d}x$$

Consequently, if ϕ is square integrable,

$$\int_0^M (\sigma(x) - \phi(x))^2\mathrm{d}x \geq \int_0^M (\phi^*(x) - \phi(x))^2\mathrm{d}x$$

for all such σ.

2.1 Wicksell's problem II

$\widetilde{U_n}(x)$ is the primitive of $\widetilde{V_n}$,

$$\widetilde{U_n}(x) = 2\int_{z=0}^\infty \sqrt{z}\mathrm{d}G_n(z) - 2\int_x^\infty \sqrt{z-x}\mathrm{d}G_n(z)$$

$\widetilde{U_n}^*$ is the concave majorant of $\widetilde{U_n}$, and the right continuous evaluation of the derivative of $\widetilde{U_n}^*$ is the projected inverse estimator for V (the isotonic inverse estimator for V).

3 Maximum Likelihood estimators

The likelihood function is defined by

$$l(g) = \sum_{i=1}^{n} \log g(z_i)$$

where $z_1, ..., z_n$ is a given realized sample from density g.
The MLE is defined as the maximizer of $l(g)$ over some class of densities.
The MLE is also an isotonic inverse estimator. We can have two kinds of complications with the MLE: the next two subsections show cases where MLE isn't well defined and the last subsection shows the case of a not uniquely defined MLE.

3.1 Wicksell's problem III

Relation between the density g of the observable data Z and the distribution function V:

$$g(z) = -\frac{1}{\pi} \int_z^{\infty} \frac{\mathrm{d}V(x)}{\sqrt{x-z}}$$

Restriction for V:

$$\int_0^{\infty} \sqrt{x}\mathrm{d}V(x) = -\frac{\pi}{2}$$

The likelihood function can be made infinity, so we change it as following to solve this problem and be able to maximize it:

$$l(V) = \frac{1}{n} \sum_{i=0}^{n-1} \log \left(\sum_{j=i+1}^{n} \frac{V(z_i) - V(z_j)}{\sqrt{z_j - z_i}} \right)$$

New restriction for V:

$$\sum_{i=1}^{n} \sqrt{z_i}(V(z_{i-1}) - V(z_i)) = \frac{\pi}{2}$$

3.2 Double censoring

Density of the observable data Z:

$$g(y, \delta, \gamma) = (F(y)h_L(y))^{\delta}((H_L(y) - H_R(y))f(y))^{\gamma}((1 - F(y))h_R(y))^{1-\gamma-\delta}$$

In this case the likelihood function isn't well defined, a way out of the problem is to use the empirical likelihood $\widetilde{l}$:

$$\widetilde{l}(F) = \sum_{i=1}^{n} \delta_i \log F(y_i) + \gamma_i \log F(\{y_i\}) + (1 - \delta_i - \gamma_i) \log(1 - F(y_i))$$

3.3 Interval censoring case 1 problem

Density g of the observable data Z:

$$g(t,\delta) = F(t)^{\delta}(1 - F(t))^{1-\delta}h(t)$$

The likelihood function

$$l(F) = \sum_{i=1}^{n} \delta_i \log(\hat{F}_n(t_i) + (1-\delta_i)\log(1 - \hat{F}_n(t_i)$$

has multiple maximizers, so we restrict the class of distribution functions to those that are constant between successive time points to have a unique MLE.