Parameter estimation

The function $S(\cdot)$ is known up to the value of some parameter $\boldsymbol{\vartheta}$, i.e., $S(x) = S(\boldsymbol{\vartheta}, x)$, the observed process is

$$\mathrm{d}X_t = S(\boldsymbol{\vartheta}, X_t)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}W_t,\ X_0,\ 0 \le t \le T$$

and the likelihood ratio function

$$L(\boldsymbol{\vartheta}; X^T) = \exp\left\{\int_0^T \frac{S(\boldsymbol{\vartheta}, X_t)}{\sigma(X_t)^2}\mathrm{d}X_t - \frac{1}{2}\int_0^T \frac{S(\boldsymbol{\vartheta}, X_t)^2}{\sigma(X_t)^2}\mathrm{d}t\right\}.$$

Maximum likelihood approach. We define the MLE $\hat{\boldsymbol{\vartheta}}_T$ as one of the solutions of the equation

$$L(\hat{\boldsymbol{\vartheta}}_T, X^T) = \sup_{\boldsymbol{\vartheta}\in\boldsymbol{\Theta}} L(\boldsymbol{\vartheta}, X^T).$$

Example. Let

$$\mathrm{d}X_t = [\vartheta\ h(X_t) + g(X_t)]\ \mathrm{d}t + \sigma(X_t)\ \mathrm{d}W_t, \quad 0 \leq t \leq T,$$

where $\vartheta \in \Theta = (\alpha, \beta)$. Then the MLE

$$\hat{\vartheta}_T = \alpha\ \mathbb{I}_{\{\eta_T \leq \alpha\}} + \eta_T\ \mathbb{I}_{\{\alpha < \eta_T < \beta\}} + \beta\ \mathbb{I}_{\{\eta_T \geq \beta\}}$$

where

$$\eta_T = \frac{\int_0^T \frac{h(X_t)}{\sigma(X_t)^2} \left[\mathrm{d}X_t - g(X_t)\,\mathrm{d}t\right]}{\int_0^T \left(\frac{h(X_t)}{\sigma(X_t)}\right)^2 \mathrm{d}t}.$$

We have

$$\hat{\vartheta}_T = \vartheta + \frac{\int_0^T \frac{h(X_t)}{\sigma(X_t)} \mathrm{d}W_t}{\int_0^T \left(\frac{h(X_t)}{\sigma(X_t)}\right)^2 \mathrm{d}t} \left(1 + o(1)\right)$$

Further

$$\sqrt{T}\left(\hat{\vartheta}_T-\vartheta\right)=\frac{\frac{1}{\sqrt{T}}\int_0^T\frac{h(X_t)}{\sigma(X_t)}\mathrm{d}W_t}{\frac{1}{T}\int_0^T\left(\frac{h(X_t)}{\sigma(X_t)}\right)^2\mathrm{d}t}\left(1+o\left(1\right)\right)$$

By the LLN and CLT

$$\frac{1}{T}\int_0^T\left(\frac{h\left(X_t\right)}{\sigma\left(X_t\right)}\right)^2\mathrm{d}t\longrightarrow\mathbf{E}_\vartheta\left(\frac{h\left(\xi\right)}{\sigma\left(\xi\right)}\right)^2=\mathrm{I}\left(\vartheta\right),$$

$$\frac{1}{\sqrt{T}}\int_0^T\frac{h\left(X_t\right)}{\sigma\left(X_t\right)}\mathrm{d}W_t\Longrightarrow\mathcal{N}\left(0,\mathrm{I}\left(\vartheta\right)\right).$$

Hence

$$\sqrt{T}\left(\hat{\vartheta}_T-\vartheta\right)\Longrightarrow\mathcal{N}\left(0,\mathrm{I}\left(\vartheta\right)^{-1}\right).$$

Bayesian approach. Suppose that the unknown parameter ϑ is random variable with the density function $p\left(\theta\right), \theta \in \Theta = (\alpha, \beta)$. Then the bayesian estimator (BE) $\tilde\vartheta_T$ is defined by

$$\mathbb{E}\left(\tilde\vartheta_T - \vartheta\right)^2 = \inf_{\bar\vartheta_T} \mathbb{E}\left(\bar\vartheta_T - \vartheta\right)^2 .$$

Here

$$\mathbb{E}\left(\bar\vartheta_T - \vartheta\right)^2 = \int_\alpha^\beta \mathbf{E}_\theta \left(\bar\vartheta_T - \theta\right)^2 p\left(\theta\right)\, \mathrm{d}\theta$$

The BE is

$$\tilde\vartheta_T = \mathbf{E}\left(\vartheta | X^T\right) = \int_\alpha^\beta \theta\, p\left(\theta | X^T\right)\, \mathrm{d}\theta = \frac{\int_\alpha^\beta \theta\, p\left(\theta\right) L\left(\theta, X^T\right)\, \mathrm{d}\theta}{\int_\alpha^\beta p\left(\theta\right) L\left(\theta, X^T\right)\, \mathrm{d}\theta}$$

Example. Suppose that

$$\mathrm{d}X_t = \vartheta\, h(X_t)\,\mathrm{d}t + \mathrm{d}W_t, \quad X_0 = x, \quad 0 \le t \le T,$$

where $\vartheta \in R_+$ and $p(\vartheta) = \lambda\, \mathrm{e}^{-\lambda\vartheta}$, $\vartheta \ge 0$. Then the direct calculation provides

$$\tilde\vartheta_T = \frac{\Delta_T - \lambda}{\mathrm{I}_T} + \sqrt{\frac{2\pi}{\mathrm{I}_T}}\; \frac{\exp\left\{-\frac{(\Delta_T-\lambda)^2}{2\mathrm{I}_T}\right\}}{1-\Phi\left(\frac{\lambda-\Delta_T}{\sqrt{\mathrm{I}_T}}\right)},$$

where $\Phi(a) = \mathbf{P}\left\{\zeta < a\right\}, \quad \zeta \sim \mathcal{N}\left(0,1\right)$ and

$$\Delta_T = \int_0^T h\left(X_t\right)\,\mathrm{d}X_t, \qquad \mathrm{I}_T = \int_0^T h\left(X_t\right)^2\,\mathrm{d}t$$

Using LLN and CLT we obtain

$$\sqrt{T}\left(\tilde\vartheta_T - \vartheta\right) \Longrightarrow \mathcal{N}\left(0, \left(\mathbf{E}_\vartheta h\left(\xi\right)^2\right)^{-1}\right)$$

Minimum distance approach.

a)**EDF** Let us introduce the *empirical distribution function*

$$\hat{F}_T(x) = \frac{1}{T} \int_0^T \mathbb{I}_{\{X_t < x\}} \,\mathrm{d}t$$

and the MDE $\boldsymbol{\vartheta}_T^*$ as

$$\left\| \hat{F}_T\left(\cdot\right) - F\left(\boldsymbol{\vartheta}_T^*, \cdot\right) \right\| = \inf_{\boldsymbol{\vartheta} \in \boldsymbol{\Theta}} \left\| \hat{F}_T\left(\cdot\right) - F\left(\boldsymbol{\vartheta}, \cdot\right) \right\|$$

Where $\|\cdot\|$ is $\mathcal{L}_2\left(\mu\right)$ norm. This estimator (under regularity conditions) is consistent and asymptotically normal

$$\sqrt{T}\left(\boldsymbol{\vartheta}_T^* - \boldsymbol{\vartheta}\right) \Rightarrow \mathcal{N}\left(0, \mathbf{d}_F\left(\vartheta\right)^2\right)$$

but it is not asymptotically efficient.

The proof of the asymptotic normality is based on the representation

$$\sqrt{T}\left(\hat{F}_T(x) - F\left(x\right)\right) = \frac{2}{\sqrt{T}}\int_{X_0}^{X_T} \frac{F\left(v\wedge x\right) - F\left(v\right)F\left(x\right)}{\sigma\left(v\right)^2 f\left(v\right)}\,\mathrm{d}v$$

$$-\frac{2}{\sqrt{T}}\int_0^T \frac{F\left(X_t\wedge x\right) - F\left(X_t\right)F_S\left(x\right)}{\sigma\left(X_t\right)\ f\left(X_t\right)}\,\mathrm{d}W_t.$$

it is asymptotically normal

$$\mathcal{L}_S\left\{\sqrt{T}\left(\hat{F}_T(x) - F\left(x\right)\right)\right\} \Longrightarrow \mathcal{N}\left(0, d_F\left(\vartheta, x\right)^2\right).$$

Here, $F\left(x\right) = F\left(\vartheta, x\right), f\left(x\right) = f\left(\vartheta, x\right)$

$$d_F\left(\vartheta, x\right)^2 = 4\ \mathbf{E}_\vartheta\left(\frac{F\left(\xi\wedge x\right) - F\left(\xi\right)F\left(x\right)}{\sigma\left(\xi\right)\ f\left(\xi\right)}\right)^2 < \infty,$$

b)**LTE** Let us introduce the *local time estimator* of the density

$$\hat{f}_T(x) = \frac{2}{T\sigma\left(x\right)^2}\int_0^T \mathbb{1}_{\{X_t<x\}}\,\mathrm{d}X_t$$

and the MDE $\boldsymbol{\vartheta}_T^{**}$ as

$$\left\|\hat{f}_T\left(\cdot\right) - f\left(\boldsymbol{\vartheta}_T^{**},\cdot\right)\right\| = \inf_{\boldsymbol{\vartheta}\in\boldsymbol{\Theta}}\left\|\hat{f}_T\left(\cdot\right) - f\left(\boldsymbol{\vartheta},\cdot\right)\right\|$$

This estimator (under regularity conditions) is consistent and asymptotically normal

$$\sqrt{T}\left(\boldsymbol{\vartheta}_T^{**} - \boldsymbol{\vartheta}\right) \Rightarrow \mathcal{N}\left(0, \mathbf{d}_f\left(\vartheta\right)^2\right)$$

but it is not asymptotically efficient.

Method of moments. Let us denote $m(\vartheta) = \mathbf{E}_\vartheta g(\xi)$ and suppose that $m(\theta)$ is strictly monotone function. Then the estimator of the method of moments (EMM) $\bar{\vartheta}_T$ is defined by the relation

$$m\left(\bar{\vartheta}_T\right) = \frac{1}{T}\int_0^T g\left(X_t\right)\, \mathrm{d}t \equiv \hat{m}_T.$$

Let $h(x) = m^{-1}(x)$, then $\bar{\vartheta}_T = h(\hat{m}_T)$

$$\sqrt{T}\left(\bar{\vartheta}_T - \vartheta\right) = \frac{h'\left(m\left(\vartheta\right)\right)}{\sqrt{T}}\int_0^T \left[g\left(X_t\right) - m\left(\vartheta\right)\right] \mathrm{d}t \Longrightarrow \mathcal{N}\left(0, D\left(\vartheta\right)^2\right)$$

Example. In the case of the Ornstein–Uhlenbeck process

$$\mathrm{d}X_t = -\left(a\ X_t - b\right)\ \mathrm{d}t + \sigma\ \mathrm{d}W_t, \quad X_0,\ 0 \leq t \leq T,$$

with $\boldsymbol{\vartheta} = (a, b)$ and $a > 0$ we take $\boldsymbol{g}(x) = \left(x, x^2\right)$. We have $\mathbf{E}_{\boldsymbol{\vartheta}}\xi = b/a$ and $\mathbf{E}_{\boldsymbol{\vartheta}}\xi^2 = b^2/a^2 + \sigma^2/2a$. Hence the EMM $\bar{\boldsymbol{\vartheta}}_T = \left(\bar{a}_T, \bar{b}_T\right)$ is

$$\bar{a}_T = \frac{\sigma^2}{2\ \left(Y_2 - Y_1^2\right)}, \qquad \bar{b}_T = \frac{Y_1\ \sigma^2}{2\ \left(Y_2 - Y_1^2\right)},$$

where

$$Y_1 = \frac{1}{T}\int_0^T X_t\ \mathrm{d}t \longrightarrow \frac{b}{a}, \qquad Y_2 = \frac{1}{T}\int_0^T X_t^2\ \mathrm{d}t \longrightarrow \frac{b^2}{a^2} + \frac{\sigma^2}{2a}.$$

Therefore this estimator is consistent. Using CLT for ordinary integrals it is possible to prove its asymptotic normality.

Example. Let

$$\mathrm{d}X_t = -\mathrm{sgn}\,(X_t - \theta)\;\mathrm{d}t + \sigma\,\mathrm{d}W_t, \qquad 0 \le t \le T$$

then the invariant density $f\,(\vartheta, x) = \frac{1}{\sigma^2} e^{-\frac{2}{\sigma^2}|x-\theta|}$ and $m\,(\vartheta) = \mathbf{E}_\vartheta \xi = \vartheta$. Hence the EMM

$$\bar\vartheta_T = \frac{1}{T}\int_0^T X_t\;\mathrm{d}t \longrightarrow \vartheta$$

and is asymptotically normal

$$\sqrt{T}\left(\bar\vartheta_T - \vartheta\right) = \frac{1}{\sqrt{T}}\int_0^T \left[X_t - \theta\right]\;\mathrm{d}t \Longrightarrow \mathcal{N}\left(0, \frac{5}{4\sigma^2}\right)$$

Trajectory fitting approach. Let us introduce a family of stochastic processes

$$\hat{X}_t(\boldsymbol{\vartheta}) = X_0 + \int_0^t S(\boldsymbol{\vartheta}, X_s)\,\mathrm{d}s,\ 0 \le t \le T, \quad \boldsymbol{\vartheta} \in \boldsymbol{\Theta}$$

and define the trajectory fitting estimator (TFE) as

$$\boldsymbol{\vartheta}_T^\star = \arg\inf_{\boldsymbol{\vartheta}\in\boldsymbol{\Theta}} \int_0^T \left[X_t - \hat{X}_t(\boldsymbol{\vartheta})\right]^2 \mathrm{d}t.$$

This estimator (under regularity conditions) is consistent and asymptotically normal. Important condition is

$$\mathbf{E}_{\boldsymbol{\vartheta}} S(\boldsymbol{\vartheta}_1, \xi) \neq 0, \qquad \text{if} \qquad \boldsymbol{\vartheta} \neq \boldsymbol{\vartheta}_1.$$

Example. In the linear case

$$\mathrm{d}X_t = -\left(a\, X_t - b\right)\, \mathrm{d}t + \sigma \mathrm{d}W_t, \qquad X_0,\ 0 \le t \le T$$

the TFE of the parameter $\boldsymbol{\vartheta} = (a, b)$ is not consistent because

$$\mathbf{E}_{\boldsymbol{\vartheta}}\left(a_1 \xi - b_1\right) = \frac{a_1\, b}{a} - b_1 = 0, \quad \text{for} \quad \boldsymbol{\vartheta}_1 = (\kappa a, \kappa b)$$

and all $\kappa > 0$. But if $\vartheta = a$ and $b \neq 0$ is known, then the TFE

$$a_T^\star = -\frac{\int_0^T \left(X_t - X_0 - bt\right)\, Y_t\, \mathrm{d}t}{\int_0^T Y_t^2 \mathrm{d}t}, \qquad Y_t = \int_0^t X_s\, \mathrm{d}s$$

can have good properties. In particular, we have

$$a_T^\star = a - \sigma\, \frac{\int_0^T W_t\, Y_t\, \mathrm{d}t}{\int_0^T Y_t^2 \mathrm{d}t}, \qquad Z_t = \frac{1}{t} \int_0^t X_s\, \mathrm{d}s \longrightarrow \mathbf{E}_a \xi = \frac{b}{a}.$$