The Sample Moments Integrating Normal Kernel (SMINK) density estimator

Let $X_1, ..., X_n$ be a random sample from a k-variate absolutely continuous distribution with density $f(x)$, expectation μ, and non-singular variance matrix Σ. Let $x^{(i)}$ be the i-th component of x, and $X_{i,j}$ the i-th component of X_j. The SMINK density estimator of $f(x)$ takes the form:

$$\widehat{f}(x|\gamma) = \frac{1}{n}\sum_{j=1}^{n}\left(\prod_{i=1}^{k} I(x^{(i)} \neq X_{i,j})\right)\widehat{f}_{n,j}(x|\gamma),$$

where

$$\begin{aligned}\widehat{f}_{n,j}(x|\gamma) &= \left(\left(\gamma\sqrt{2\pi}\right)^{k}\det\left(\widehat{\Sigma}\right)\right)^{-1} \\ &\times \exp\left[-\frac{1}{2}\left(x - \sqrt{1-\gamma^2}X_j - \left(1-\sqrt{1-\gamma^2}\right)\overline{X}\right)'\widehat{\Sigma}^{-1}\right. \\ &\times \left.\left(x - \sqrt{1-\gamma^2}X_j - \left(1-\sqrt{1-\gamma^2}\right)\overline{X}\right)/\gamma^2\right],\end{aligned}$$

with $\overline{X} = (1/n)\sum_{j=1}^{n} X_j$ and $\widehat{\Sigma} = (1/n)\sum_{j=1}^{n}\left(X_j - \overline{X}\right)\left(X_j - \overline{X}\right)'$.

For $\gamma \in (0,1]$ we have

$$\int x\widehat{f}(x|\gamma)dx = \overline{X}, \quad \int xx'\widehat{f}(x|\gamma)dx = \frac{1}{n}\sum_{j=1}^{n} X_jX_j'.$$

Let $\xi_n \in (0,1]$ be a sequence of non-random numbers such that $\lim_{n\to\infty}\xi_n = 0$, $\lim_{n\to\infty}\sqrt{n}\xi_n^k = \infty$. In particular, let

$$\xi_n = \left(\sqrt{n}\right)^{-\alpha/k}, \ 0 < \alpha < 1.$$

EasyReg International will ask you to specify α (the default value is 0.5).

Moreover, let γ_n be a sequence of (random) numbers such that $\gamma_n \in [\xi_n, 1]$ (a.s), $(p)\lim_{n\to\infty}\gamma_n = 0$. Then $\widehat{f}(x|\gamma_n)$ is uniformly consistent:

$$p\lim_{n\to\infty}\left|\widehat{f}(x|\gamma_n) - f(x)\right| = 0. \quad (1)$$

Furthermore, if we choose

$$\gamma_n = \arg\min_{\gamma\in[\xi_n,1]}\widehat{Q}(\gamma),$$

where

$$\widehat{Q}(\gamma) = \int \widehat{f}(x|\gamma)^2 dx - 2\frac{1}{n}\sum_{j=1}^{n} \widehat{f}(X_j|\gamma),$$

then $\int \left(\widehat{f}(x|\gamma_n) - f(x)\right)^2 dx$ is (approximately) minimal, and (1) carries over.

Reference: Bierens, H.J. (1983), "Sample Moments Integrating Normal Kernel Estimators of Multivariate Density and Regression Functions", *Sankhya*, **45**, *Series B*, 160-192.