Semiparametric Transformation Models

Chris A.J. Klaassen
Korteweg-de Vries Instituut voor Wiskunde
Universiteit van Amsterdam

We will discuss a quite general class of semiparametric transformation models. Let (T, Z) be a random vector with Z having a distribution on some measurable space and with $p_0(t; z, \theta)$, $\theta \in \Theta \subset \mathbb{R}^d$, the conditional density of T at t given $Z = z$ with respect to Lebesgue measure on $(\mathbb{R}, \mathcal{B})$. We observe i.i.d. copies of (Y, Z) with $\psi(Y) = T$, where $\psi : \mathbb{R} \to \mathbb{R}$ is an unknown transformation. If ψ is absolutely continuous with respect to Lebesgue measure with derivative $\psi' \geq 0$, then the conditional density $p(y; z, \theta, \psi)$ of Y given $Z = z$ becomes

$$p(y; z, \theta, \psi) = p_0\left(\psi(y); z, \theta\right) \psi'(y) \, .$$

Note that, for

$$\psi(y) = \Lambda(y) = \int_0^y \lambda(x) dx \, , \; y \geq 0 \, ,$$

and

$$p_0(t; z, \theta) = e^{\theta z} \exp\{-e^{\theta z} t\} \mathbf{1}_{(0,\infty)}(t) \, , \qquad d = 1 \, , \; z \in \mathbb{R} \, ,$$

this model leads to

$$p(y; z, \theta, \Lambda) = e^{\theta z} \lambda(y) \exp\left\{-e^{\theta z} \Lambda(y)\right\} \, , \; y \geq 0 \, .$$

This is the density of the Cox proportional hazards model with regression parameter θ and baseline hazard rate λ. Since the parameter λ varies over a function space, this is called a semiparametric model. Another important model that fits into this framework is a semiparametric generalization of the Box-Cox model with $p_0(t; z, \theta) = \varphi(t - \theta z)$, where φ is the standard normal density. So, if ϵ has a standard normal distribution we observe i.i.d. copies of (Y, Z) , where $\psi(Y) = T = \theta Z + \epsilon$ holds with θ and ψ unknown.

In this lecture we will study efficient estimation of the parameter of interest θ in the introduced general transformation model, thus illustrating some fundamental issues in semiparametric statistics.