# Asymptotic distribution theory for M-estimators in two-phase regression models

Dietmar Ferger
Technische Universität Dresden
Department of Mathematics
Helmholtzstr. 10
D-01069 Dresden, Germany
e-mail: dietmar.ferger@tu-dresden.de

Starting point of our talk is the problem in medicine to construct a diagnostic test for patients with morbus Parkinson. An appropriate mathematical model is given by a sequence $(X_i, Y_i), 1 \leq i \leq n$, of independent and identically distributed random variables in $\mathbb{R}^2$ which are copies of $(X, Y)$ with $Y = m(X) + \epsilon$. It is assumed that the domain of the regression function $m$ is divided into two regions $R_- = (-\infty, \theta]$ and $R_+ = (\theta, \infty)$ for some unknown *split-point* $\theta$. The graph of $m$ runs above the *mean level* $\mathbb{E}(Y)$ on the region $R_-$ and below that level on $R_+$, or vice versa. This includes the case that $m$ has a jump at the point $\theta$, but also the situation in which $m$ smoothly crosses the mean level, i.e. $0 = m^{'}(\theta) = \ldots = m^{(k-1)}(\theta)$ and $m^{(k)}(\theta) \neq 0$.
We motivate a nonparametric estimator $\hat{\theta}_n$ for $\theta$ which goes back to Dempfle and Stute (2002). The asymptotic behavior of the sequence $(\hat{\theta}_n)$ essentially depends on the type of change. If $m$ has has a discontinuity at $\theta$ then we can prove $n-$consistency and a weak convergence type result. The limit variable is not normal, but a maximizing point of a compound Poisson process on the real line. In case of a smooth crossing of order $k$ as described above one obtains that $n^{1/(2k+1)}(\hat{\theta}_n - \theta)$ converges in distribution to the maximizing point of a Brownian motion on $\mathbb{R}$ with a polynomial drift function of degree $k+1$. The special case $k = 1$ yields a *cube root asymptotic* as in Banerjee and McKeague (2007) or Bühlmann and Yu (2002).