Tail Index Regression–Adjusted Functional Covariate

Anwar Alabdulathem[1] Miguel de Carvalho[2]

[1]s1794726@ed.ac.uk, [1] School of Mathematics, University of Edinburgh, UK, [2]miguel.decarvalho@ed.ac.uk,
[2]Reader in Statistics, School of Mathematics, University of Edinburgh, UK

Generalized Pareto distribution (GPD)

The generalized Pareto distribution (GPD) is a two parameter distribution with distribution function.

$$G_{\xi,\sigma}(x) = \begin{cases} 1-(1+\xi x/\sigma)^{-1/\xi}, & \xi \neq 0 \\ 1-\exp(-x/\sigma), & \xi = 0 \end{cases}$$

where $\sigma > 0$, and the support is $x \geq 0$ when $\xi \geq 0$ and $x \in [0, -\sigma/\xi]$ when $\xi < 0$.

Functional Covariate-Adjusted Tail Index Model:

$$1 - F(y \mid X) = y^{-\alpha_X} L(y \mid X).$$

Where $\alpha_X = \exp(\langle X, \beta \rangle)$, with $X \equiv X(t)$ is a functional covariate and $\beta \equiv \beta(t)$ is a function of the regression coefficient and both $X, \beta \in X$ being a squared–integrable function and with $\langle f, g \rangle = \int_0^1 f(t)g(t)dt$.

Coefficient Regression $\beta(t)$

Assume that $\beta(t)$ can be written as a function of spline which has the form of de Boor (2001),

$\beta(t) = \sum_{j=1}^{\eta+1} w_j B_j(t), \quad j = 1, \ldots, \eta$

where $\{B_j, j = 1, \ldots, \eta + 1\}$ is a B-spline basis.

Learning from Data

To learn about the tail index we follow Wang and Tsai (2009). Let $\{(X_i, Y_i)\}_{i=1}^n$ be a random sample where $X_i \in X$ and the Y_i are random variables and

$$\mathcal{R}_n(\mathbf{w}) = \sum_{i=1}^{n} \left\{ \exp(\langle X_i(t), \sum_{j=1}^{\eta+1} w_j B_j(t) \rangle) \log(Y_i/\nu_n) - X_i(t), \sum_{j=1}^{\eta+1} w_j B_j(t) \right\} I(Y_i > \nu_n)$$

$$\hat{\beta}(t) = \sum_{j=1}^{\eta+1} \hat{w}_j B_j(t)$$

Experiments of Artificial Data

- **Scenario 1**, Pareto(α_X).
- **Scenario 2**, Burr(1,α_X).
- **Scenario 3**, GPD(1,1,α_X).

where $X_i(t) = Z_i^2(t)$ and $Z_i(t)$ is Brownian bridge $Z = \{Z(t) : t \in [0, 1]\}$ and the true coefficient functions are, $\beta_1(t) = \cos(3t)^2$, $\beta_2(t) = \sin(2t)^2$, $\beta_3(t) = \cos(6t)^2/\exp(t)$ respectively

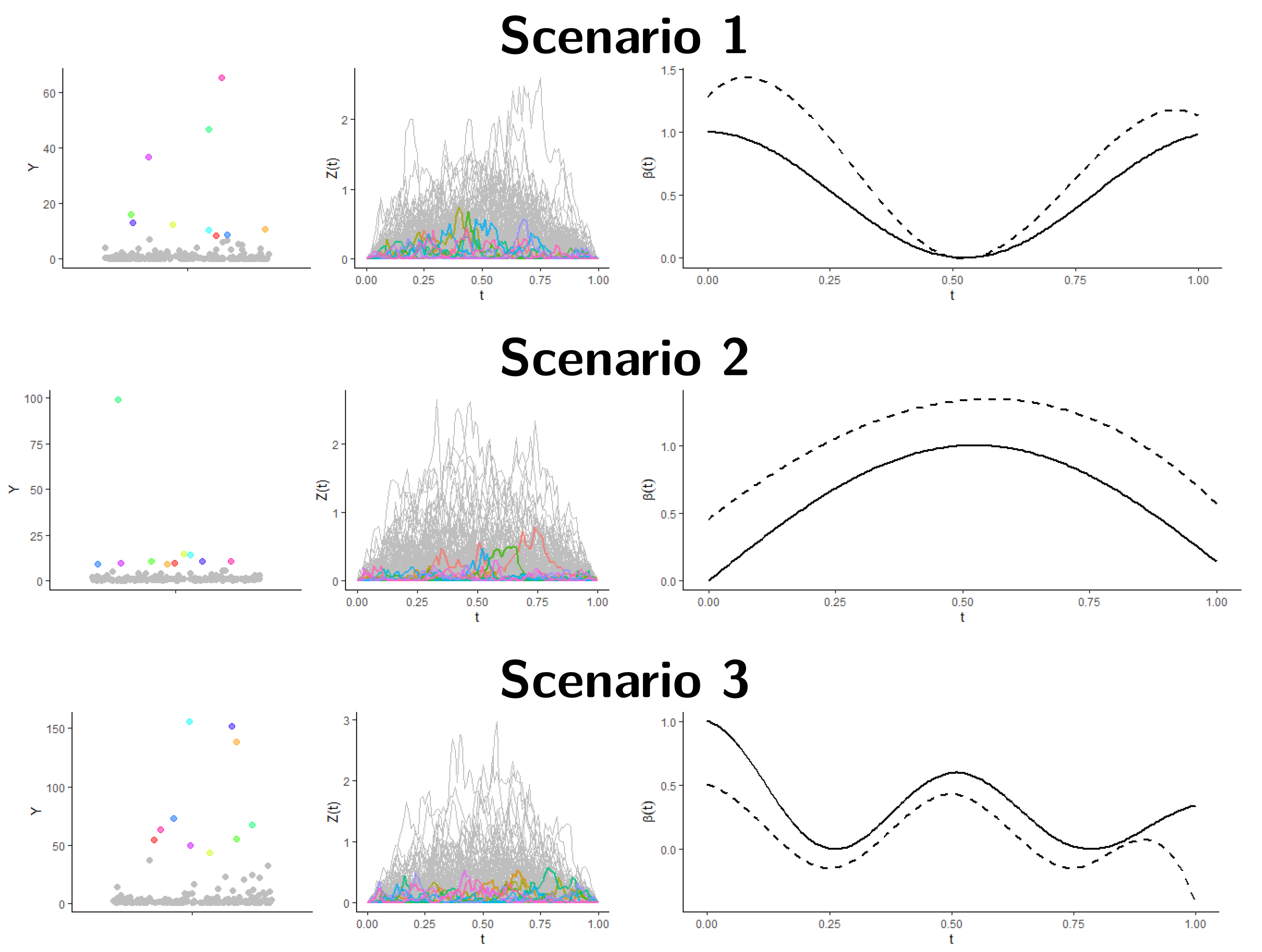

Figure 1: Response Y of sample size $n = 200$ for Scenarios 1-3;(Right) along with corresponding to trajectories $Z(t)^2$ of Brownian bridge;(Left) simulation result for a beta estimate (dashed line) against corresponding a true beta (solid line). (top) $\beta_1(t) = \cos(3t)^2$ (middle) $\beta_2(t) = \sin(2t)^2$ (bottom) $\beta_3(t) = \cos(6t)^2/\exp(t)$.

References

de Boor, C. (2001). A practical guide to splines.(rev. ed.) springer-verlag. *New York*.

Wang, H. and Tsai, C.-L. (2009). Tail index regression. *Journal of the American Statistical Association*, 104(487):1233–1240.