Discussion on "Multiscale change point inference" by Frick, Munk and Sieling

Alessio Farcomeni (Sapienza - University of Rome)

First, I would like to congratulate the authors on a very accurate and clearly written work. Their results are impressive, and the approach is undoubtedly useful and computationally efficient. I believe their contribution goes beyond what is neatly written in this paper. For instance, three major generalizations may be obtained with reasonable effort:

- *Panel data*: in many cases, we have independent replicates of Y (e.g., we may have repeated measures on m units at n occasions). In this case the local likelihood ratio statistics would be

$$T_i^j(Y,\theta_0) = \sup_{\theta\in\Theta}\left[\sum_{v=1}^{m}\sum_{l=i}^{j}\theta Y_{vl} - \psi(\theta)\right] - \sum_{v=1}^{m}\sum_{l=i}^{j}\theta_0 Y_{vl} - \psi(\theta_0)$$

  and similar adjustments can be made to the definitions of $C(q)$, $\hat{\theta}(q)$, etc. The use of independent replicates of Y should often lead to shorter confidence intervals around change point location estimates.

- *Multivariate data*: suppose now $Y_i \in \mathcal{R}^p$, with $p > 1$, arising for instance from a p-dimensional continuous exponential family. Computation of local likelihood ratio statistics seems to be straightforward also in this case. When $p > 1$, it may also be interesting to derive dimension-specific likelihood ratio statistics. In the latter case the choice of α may be driven by ideas taken from the multiple testing literature (reviewed e.g. in Farcomeni (2008)). Finally, issues with multivariate outliers may arise, suggesting the use of robust estimation methods (e.g., Cuesta-Albertos *et al.* (2008)).

- *General right-continuous functions*: the right-continuous step function θ can be offset by any known continuous function with once again minor adjustments to SMUCE. It may be of interest to to explore whether an *unknown* function within this class can be estimated with a performance as good as that of SMUCE. Some additional restrictions would perhaps be needed for consistency (e.g., that the oscillation within each interval is smaller than Δ). Similarly, if X_i is a vector of covariates, it would be interesting to estimate θ and β in the model

$$Y_{vi} \sim F_{\theta(i/n)+g(\beta' X_i)}, v = 1, \ldots, m$$

where $g(\cdot)$ is a known link function and $m \geq 1$.

References

J.A. Cuesta-Albertos, C. Matrán, and A. Mayo-Iscar (2008). Trimming and likelihood: Robust location and dispersion estimation in the elliptical model. *Annals of Statistics*, **36**, 2284–2318.

A. Farcomeni (2008). A review of modern multiple hypothesis testing, with particular attention to the false discovery proportion. *Statistical Methods in Medical Research*, **17**, 347–388.