Robust, sequential design strategies

D.P. Wiens[1]

[1] Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, CANADA T6G 2G1

Keywords: Nonlinear regression, Clinical trials, Spatial studies.

We discuss the problem of designing an experimental strategy, when the response variable obeys an *approximate* regression model. Observations are to be made in such a way that the resulting predictions are robust against this approximate knowledge. Specific applications might be:

(i) Nonlinear regression. Here one observes

$$Y = E[Y|\mathbf{x}] + error,$$

where the regression response $E[Y|\mathbf{x}]$ is thought to be approximately of the form $f(\mathbf{x};\theta)$ for some function $f(\mathbf{x};\cdot)$, nonlinear in θ. For instance $f(\mathbf{x};\theta)$ might be the exponential response $\theta_0 e^{-\theta_1 x}$ and $E[Y|\mathbf{x}]$ might instead be an unknown member of a class containing both $f(\mathbf{x};\theta)$ and the Michaelis-Menten response $\theta_0 x/(\theta_1 + x)$. The aim is to choose design points $\mathbf{x}_1, ..., \mathbf{x}_n \in \mathcal{S}$ ("design space") so as to minimise the Integrated Mean Squared Error

$$IMSE = \int_{\mathcal{S}} E\left[\left(f\left(\mathbf{x};\hat{\theta}\right) - E[Y|\mathbf{x}]\right)^2\right] d\mathbf{x}.$$

This requires the *estimation* of $d(\mathbf{x};\theta) = f(\mathbf{x};\theta) - E[Y|\mathbf{x}]$, calling for a *sequential* design strategy.

(ii) Clinical trials. Subjects arrive and are to be assigned to one of p treatments. The response is assumed to depend on the treatment administered, and as well on one or more covariates which are observed prior to the assignment. The relationship of response to covariates is assumed to be *approximately linear* in an appropriate set of regressors. If $f_i(\mathbf{x})$ denotes the difference between the true and assumed regression response in the i^{th} treatment group, and $\rho_i(\mathbf{x})$ the probability that a subject with covariates $\mathbf{x}$ is assigned to the i^{th} group, then the loss considered is

$$\mathcal{L}(\rho_1, ..., \rho_p) = \max_{f_1,...,f_p} \lim_{n\to\infty} \left| n\mathbf{MSE}\left(\mathbf{W}_0\hat{\theta}\right)\right|,$$

where $\mathbf{W}_0\hat{\theta}$ is a complete set of orthogonal contrasts in the treatment effects and **MSE** is the mean squared error matrix. A sequential strategy, leading to the attainment of those optimal assignment probabilities which minimise $\mathcal{L}(\rho_1, ..., \rho_p)$, is given.

If time allows, work in progress relating to the design of *spatial studies* may be discussed.

Please fill in this form and mail it together with your abstract.

My abstract fits best to topic number 19

List of Topics:

1. Algorithms
2. Applications
3. Biostatistics
4. Computing and graphics
5. Data analysis
6. Data mining
7. Economics, finance
8. Efficiency and robustness
9. Functionals and bias
10. Fuzzy statistics
11. Geostatistics
12. Inference for robust methods, model testing
13. Location depth and regression depth
14. Multivariate methods
15. Neural networks
16. Rank-based methods
17. Regression quantiles, trimming
18. Robust covariance
19. Robust designs
20. Robust regression
21. Time series analysis
22. Wavelets
23. Other (please specify)