different $\theta_0$ if the underlying process is not of the fitted structure, although they lead to the same $\theta_0$ if the process is of the fitted structure. Suppose we fit an AR($k$) model. Consider, for example, the Gaussian likelihood

$$L_T(\theta) = \frac{1}{T} \log \det \Sigma_\theta + \frac{1}{T} Y' \Sigma_\theta^{-1} Y$$

and alternatively an $M$ estimate

$$L_T^*(\theta) = \frac{1}{T} \sum_{t=1}^{T} \rho\left(\sum_{s=0}^{k} a_s Y_{t-s}\right)$$

$$\text{with} \quad Y_t = 0 \quad \text{if} \quad t \le 1$$

(if $\sigma^2$ is unknown the estimate has to be modified, cf. Martin and Yohai (1985)). Then,

$$EL_T(\theta) = \frac{1}{T} \log \det \Sigma_\theta + \frac{1}{T} \operatorname{tr}\{\Sigma\Sigma_\theta^{-1}\}$$

and

$$EL_T^*(\theta) \approx E\rho\left(\sum_{s=0}^{k} a_s Y_{t-s}\right).$$

If $Y_t$ is also an AR($k$) process then both $EL_T(\theta)$ and $EL_T^*(\theta)$ are minimized by the true parameter value, while in the case where $Y_t$ is not an AR($k$) process, $EL_T(\theta)$ and $EL_T^*(\theta)$ are minimized by different values. This means that one has not only to consider the quality of the estimation procedure, but also the "quality" of the estimated parameter.

In the formula below (5.8), Hannan should not compare the estimate $\hat{\Phi}_h(j)$ with $\Phi(j)$ but with the estimated parameter $\Phi_h(j)$ (in the above sense), obtained as a solution of the theoretical counterpart of equation (5.8), and then ask in a second step how good the $\Phi_h(j)$ represent the structure of the series (in fact, the finitely many $\Phi_h(j)$, $j = 1, \cdots, h$, describe the structure of the process "better" than the finitely many $\Phi(j)$, $j = 1, \cdots, h$).

It is obvious that the choice of an estimation procedure doesn't only imply an estimated parameter $\theta_0$ but also an optimal order. The results of Shibata (1980) favoring AIC are only for the case where the parameters are estimated by the Yule-Walker equations. It would be interesting to know whether using other estimation procedures (e.g., robust ones) leads to other order criteria.

## ADDITIONAL REFERENCES

DAHLHAUS, R. (1986). Small sample effects in time series analysis. I. Preprint. University of Essen.

MARTIN, R. D. and YOHAI, V. J. (1985). Robustness in time series and estimating ARMA models. In *Handbook of Statistics* (E. J. Hannan, P. R. Krishnaiah and M. M. Rao, eds.) **5** 119–155. North Holland, Amsterdam.

PARZEN, E. (1983). Autoregressive spectral estimation. In *Handbook of Statistics* (D. R. Brillinger and P. R. Krishnaiah, eds.) **3** 221–247. North Holland, Amsterdam.

# Comment

**Jorma Rissanen**

In this exceptionally lucid and comprehensive survey, Professor Hannan covers essentially all the important ideas in the theory of linear dynamic systems, both deterministic and stochastic, developed during the past twenty years or so. In addition, he describes the more recently introduced new statistical ideas for selecting such models for time series. I was particularly impressed by the apparent ease and elegance with which Professor Hannan managed to explain the rather intricate notions without any undue sacrifice in precision.

I would like to comment on two issues of a general nature raised by Professor Hannan. There have been several attempts to apply the beautiful and deep approximation theory of Adamyan, Arov and Krein in a statistical context for the purpose of obtaining an optimal low order model reduction. As explained in the paper, such a procedure begins with a high order dynamic system, arrived at, perhaps, by applying physical or chemical laws to a process, or by other means. This is then, in the second stage, reduced to a desired complexity, optimally in the sense of minimum distance in a certain norm. The point I wish to make is that because the initial system, which necessarily has the status of a model rather than any "true" system, is nonunique, the end result cannot be assigned any meaningful optimality property. Instead, it is just an optimal approximation of an arbitrary model of the data.

My remaining comments aim to amplify and, perhaps, modify some of the concluding remarks in

*Jorma Rissanen is a Member of the Research Staff, IBM Almaden Research Center, 650 Harry Road, San Jose, California 95120.*