`icomp`: Informational complexity measures

Stanislav Kolenikov, E-mail: skolenik@nes.cemi.rssi.ru

March 2000

This document provides formulae (from [3]) on which the calculations by `icomp` are based.

`icomp` calculates Akaike information criteria (AIC), Schwartz Bayesian information criteria (SBC, SBIC), and Bozdogan's index of informational complexity (ICOMP). These criteria are used to select the "best" model compromising an adequate goodness of fit and a small number of parameters by adding a penalty for overparametrization to the lack of fit measure (estimate of the maximum likelihood or the residual sum of squares). The best model minimizes the criterion. The three informational measures differ in the penalty term, SBIC penalizing more severely for larger samples, and ICOMP accounting for covariance structure of a model (and, thus, for collinearity between the factors and dependence among the parameter estimates).

$$AIC = -2 \ln L(\hat{\theta}_k) + 2k; \tag{1}$$

$$SBIC = -2 \ln L(\hat{\theta}_k) + k \ln n; \tag{2}$$

$$ICOMP = -2 \ln L(\hat{\theta}_k) + s \ln \operatorname{tr}(I^{-1}(\hat{\theta}_k)/s) - \ln |I^{-1}(\hat{\theta}_k)|, \tag{3}$$

where k is the number of parameters, n, sample size, $\hat{\theta}_k$, the parameter estimates, $I^{-1}(\hat{\theta}_k) = \widehat{Var}(\hat{\theta}_k)$, inverse of Fisher information matrix, and $s = \operatorname{rk} I^{-1}(\theta_k)$.

Empirical versions of these criteria also use

$$n \ln \frac{RSS}{n} \tag{4}$$

as a lack of fit measure plugged as the first term in (1)–(3), where RSS is the residual sum of squares. This variant is activated by option `icomp, rss`.

References

[1] Akaike, H. Information theory and an extension of the maximum likelihood principle. In B.N.Petrov and F.Csaki (eds), *Second International Symposium on Information Theory*, Academiai Kiado, Budapest, 267–281 (1973).

[2] Bozdogan, H. On the information-based measure of covariance complexity and its application to the evaluation of multivariate linear models. *Communications in Statistics, Theory and Methods*, **19**(1), 221–278 (1990).

[3] Bozdogan, H. Empirical econometric modelling of food consumption using a new informational complexity approach. *J. of Applied Econometrics*, **12**, 563–592 (1997).

[4] Kulback, S. and R. A. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, **22**, 79–86 (1951).

[5] Schwartz, G. Estimating the dimension of a model. *Annals of Statistics*, **6**, 461–464 (1978).