Choosing a model: From A(IC) to D(IC)

Giles Innocent & Matt Denwood

COST-Harmony Madrid July 2023

Why is one model better than another?

- Describes the data
- Fits the data
- As simple as possible (Occam's razor - William of Ockham)

BUT

- How good is the fit?
- How simple is simple?

Enter Akaike

- ▶ Akaike information criterion (AIC) is an estimator of prediction error and thereby relative quality of statistical models for a given set of data. (Wikipaedia)
- ▶ Based on information theory.
- ▶ How much information is in the model?
  - ▶ depends on the number of parameters estimated in the model
  - ▶ needs a coefficient to weight it equivalnt to the data
  - ▶ *usually* $2k$
- ▶ How much information is in the residuals?
  - ▶ sometimes referred to as "how surprised are we?"
  - ▶ if the model fits perfectly then there are no residuals and no information in the data beyond the model.
  - ▶ information theory indicates that we should use the deviance, i.e. $-2 * log(\hat{L})$

Bayesian MCMC

- ▶ The degrees of freedom – constraints – in a Bayesian model are difficult to define.
- ▶ Probablistic constraints, e.g. informative priors, do not have integer degrees of freedom
- ▶ Spiegelhalter *et al.* (2002) proposed the effective number of estimated parameters in a model P_D: a non-integer measure.

Issues

- There are multiple ways of calculating P_D.
- Using at least one method P_D is always less than k-1 (df in the data) *even for over specified models.*
- Using DIC alone may lead one to believe in a poorly fittng model. It is better to look at the posterior predictive P-values as well