ADVANCED BAYESIAN MODELING

WAIC: A More Bayesian Approach?

Scores Reconsidered

Bayesian methods for predicting $\tilde{y}$ don't use a "plug-in" estimate, as in

$$p\big(\tilde{y} \mid \hat{\theta}\big)$$

but rather use a posterior predictive density

$$p_{\text{post}}(\tilde{y}) \;=\; \int p(\tilde{y} \mid \theta)\, p(\theta \mid y)\, d\theta$$

Suggests a different way to define the score for a Bayesian model:

$$\log\, p_{\text{post}}(\tilde{y})$$

May seem reasonable to just substitute data y for $\tilde{y}$ and then add a correction, as for DIC (or AIC).

BDA3 recommends something slightly different:

Consider this **expected log pointwise predictive density**

$$\text{elppd} \;=\; \sum_{i=1}^{n} \mathrm{E}_{f_i}\big(\log\, p_{\mathsf{post}}(\tilde{y}_i)\big)$$

totaled over n separate individual predictions of $\tilde{y}_i$.

Here, f_i is the "true" density of $\tilde{y}_i$.

When $\tilde{y}_i$ is a replication of observation y_i, elppd can be estimated by substitution:

$$\mathrm{lppd} \;=\; \sum_{i=1}^{n} \log p_{\mathsf{post}}(y_i)$$

Since the same data y is being used twice (for the posterior and for substitution), this will tend to overestimate elppd.

Needs a correction ...

WAIC

In BDA3, the **Watanabe-Akaike Information Criterion (WAIC)** is

$$\mathrm{WAIC} \;=\; -2\,\mathrm{lppd} \;+\; 2p_{\mathrm{WAIC}}$$

where effective number of parameters p_{WAIC} could be

$$p_{\mathrm{WAIC1}} \;=\; 2\sum_{i=1}^{n}\Big(\log\big(\mathrm{E}_{\textsf{post}}\, p(y_i \mid \theta)\big) \;-\; \mathrm{E}_{\textsf{post}}\big(\log p(y_i \mid \theta)\big)\Big)$$

$$p_{\mathrm{WAIC2}} \;=\; \sum_{i=1}^{n} \mathrm{var}_{\textsf{post}}\big(\log p(y_i \mid \theta)\big)$$

BDA3 recommends second form.

Notes:

- Smaller values of WAIC are preferred (as with AIC or DIC).
- p_{WAIC1} and p_{WAIC2} can be approximated using a posterior Monte Carlo sample: Replace posterior means and variances with sample means and variances.
- WAIC can estimate elppd:

$$\widehat{\text{elppd}}_{\text{WAIC}} \;=\; -\frac{1}{2}\,\text{WAIC} \;=\; \text{lppd} \,-\, p_{\text{WAIC}}$$