

BAYESIAN STATISTICS - LECTURE 12

LECTURE 12: MODEL EVALUATION AND COURSE SUMMARY

MATTIAS VILLANI

DEPARTMENT OF STATISTICS

STOCKHOLM UNIVERSITY

AND

DEPARTMENT OF COMPUTER AND INFORMATION SCIENCE

LINKÖPING UNIVERSITY

- **Model evaluation - Posterior predictive analysis**
- **Course summary and discussion**

- We now know how to **compare** models.
- But how do we know if any given model is 'any good'?
- George Box: '**All models are false, but some are useful**'.

WHAT IS YOUR MODEL FOR REALLY?

■ **Prediction.**

- Interpretation not a concern
- Black-box approach may be ok.
- Extrapolation?
- Model averaging may be a good idea.

■ Abstraction to **aid in thinking** about a phenomena.

- Prediction accuracy of less concern.
- Model averaging may be a bad idea.

■ Model as a **compact description of a complex phenomena.**

- Computational cost of model evaluation may be a concern.
- Online/real-time analysis.

POSTERIOR PREDICTIVE ANALYSIS

- If $p(y|\theta)$ is a 'good' model, then the data actually observed should not differ 'too much' from simulated data from $p(y|\theta)$.
- Bayesian: simulate data from the **posterior predictive distribution**:

$$p(y^{rep}|y) = \int p(y^{rep}|\theta)p(\theta|y)d\theta.$$

- Difficult to compare y and y^{rep} because of dimensionality.
- Solution: compare **low-dimensional statistic** $T(y, \theta)$ to $T(y^{rep}, \theta)$.
- Evaluates the full probability model consisting of both the likelihood *and* prior distribution.

- **Algorithm** for simulating from the posterior predictive density $p[T(y^{rep})|y]$:
 - 1 Draw a $\theta^{(1)}$ from the posterior $p(\theta|y)$.
 - 2 Simulate a data-replicate $y^{(1)}$ from $p(y^{rep}|\theta^{(1)})$.
 - 3 Compute $T(y^{(1)})$.
 - 4 Repeat steps 1-3 a large number of times to obtain a sample from $T(y^{rep})$.
- We may now compare the observed statistic $T(y)$ with the distribution of $T(y^{rep})$.
- **Posterior predictive p-value:** $\Pr[T(y^{rep}) \geq T(y)]$
- Informal **graphical analysis**.

- Ex. 1. Model: $y_1, \dots, y_n \stackrel{iid}{\sim} N(\mu, \sigma^2)$. $T(y) = \max_i |y_i|$.
- Ex. 2. Assumption of no reciprocity in networks.
 $y_{ij}|\theta \stackrel{iid}{\sim} \text{Bernoulli}(\theta)$. $T(y)$ = proportion of reciprocated node pairs.
- Ex. 3. ARIMA-process. $T(y)$ may be the autocorrelation function.
- Ex. 4. Poisson regression. $T(y)$ frequency distribution of the response counts. Proportions of zero counts.

POSTERIOR PREDICTIVE ANALYSIS - NORMAL MODEL, MAX STATISTIC

