Bayesian Statistics - Lecture 12

Lecture 12: Model evaluation. Course summary

Mattias Villani

Department of Statistics
Stockholm University
and
Department of Computer and Information Science
Linköping University

OVERVIEW

- **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.

- If $p(\mathbf{y}|\theta)$ is a 'good' model, then the data actually observed should not differ 'too much' from simulated data from $p(\mathbf{y}|\theta)$.

- Bayesian: simulate data from the **posterior predictive distribution**:

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

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

- **Algorithm** for simulating from the posterior predictive density $p[T(\mathbf{y}^{rep})|\mathbf{y}]$:

1. Draw a $\theta^{(1)}$ from the posterior $p(\theta|\mathbf{y})$.
2. Simulate a data-replicate $\mathbf{y}^{(1)}$ from $p(\mathbf{y}^{rep}|\theta^{(1)})$.
3. Compute $T(\mathbf{y}^{(1)})$.
4. Repeat steps 1-3 a large number of times to obtain a sample from $p[T(\mathbf{y}^{rep})|\mathbf{y}]$.

- We may now compare the observed statistic $T(\mathbf{y})$ with the distribution of $T(\mathbf{y}^{rep})$.
- **Posterior predictive p-value**: $\Pr[T(\mathbf{y}^{rep}) \geq T(\mathbf{y})]$
- Informal **graphical analysis**.

- Ex. 1. Model: $y_1, \dots, y_n \stackrel{iid}{\sim} N(\mu, \sigma^2)$. $T(\mathbf{y}) = \max_i |y_i|$.

- Ex. 2. Assumption of no reciprocity in networks.
  $y_{ij}|\theta \stackrel{iid}{\sim} Bernoulli(\theta)$. $T(y) =$proportion of reciprocated node pairs.

- Ex. 3. ARIMA-process. $T(\mathbf{y})$ may be the autocorrelation function.

- Ex. 4. Poisson regression. $T(\mathbf{y})$ frequency distribution of the response counts. Proportions of zero counts.

Posterior predictive analysis - Normal model, max statistic