

22S:138
Posterior predictive checking

Lecture 21
 Nov. 19, 2008

Kate Cowles
 374 SH, 335-0727
 kcowles@stat.uiowa.edu

Model checking and sensitivity analysis

- goal: assess fit of model to
 - data
 - our substantive knowledge
- must check effects of
 - prior
 - likelihood specification
 - hierarchical structure
 - any other application-specific issues
 - * e.g. which predictor variables

- theoretically possible to set up and fit a “super model” including all possibly true models
 - but computationally infeasible
 - and really conceptually impossible
- instead we fit a feasible number of models and examine the posterior distributions that result
 - cast models as broadly as possible
 - fail to fit reality?
 - sensitive to arbitrary specifications?

Principles and methods of model-checking

- “do the model’s deficiencies have a noticeable effect on substantive inferences?”
- how to judge when assumptions of convenience can be made safely

Using the posterior distribution to check a statistical model

- compare posterior distribution of parameters to
 - substantive knowledge
 - other data
- compare posterior *predictive* distribution of future observations to substantive knowledge
 - e.g.: compare election predictions from a model to substantive knowledge
- compare posterior predictive distribution of future observations to the data that have actually occurred

Checking a model by comparing the data that we have to the posterior predictive distribution

- enables checking fit of model without any more substantive knowledge than is in existing data and model
- do datasets simulated from the model we fit “look like” the real data in ways relevant to our inference?
- requires drawing “replicated data”

Using the posterior predictive distribution to check a statistical model

- recall:
 - posterior: conditional on observed data y
 - predictive: prediction of an observable but unobserved y
 -

$$\begin{aligned}
 p(\tilde{y}|y) &= \int p(\tilde{y}, \theta|y) d\theta \\
 &= \int p(\tilde{y}|\theta, y) p(\theta|y) d\theta \\
 &= \int p(\tilde{y}|\theta) p(\theta|y) d\theta
 \end{aligned}$$

- last line holds if new data are conditionally independent of old data given model parameters

Procedure to draw a “replicated dataset” from posterior predictive distribution

- notation
 - y : observed data
 - y^{rep} : a complete simulated dataset
 - * same number of observations as in y
 - * same values of explanatory variables (if any)
 - * response variables simulated from posterior predictive distribution
 - θ : vector of all unknown model parameters, including parameters of upper stage priors if model is hierarchical

- Step 1: draw θ^* from $p(\theta|y)$
i.e. from posterior distribution of θ
- Step 2: draw y^{rep} from $p(y^{rep}|\theta^*)$
- repeat steps 1 and 2 a large number of times

Discrepancy measures or test quantities for posterior predictive checks

- intended to measure discrepancy between model and real data
- $T(y, \theta)$: scalar summary of data (and possibly parameters) used as a standard when comparing real data to data simulated from posterior predictive distribution
- choose one or more test quantities that are meaningful with respect to your research purpose

Using the test quantities: posterior predictive p-values

- compute $T(y, \theta)$ for the real data y
- compute $T(y, \theta^{rep})$ for each simulated replicate dataset
- compute the proportion of the replicated datasets for which $T(y^{rep}, \theta) \geq T(y, \theta)$
- this is an approximation to the Bayes p-value

$$\iint I_{(T(y^{rep}, \theta) \geq T(y, \theta))} p(\theta|y) p(y^{rep}|\theta) d\theta dy^{rep}$$

- that is, Bayes p-value is $Pr(T(y^{rep}, \theta) \geq T(y, \theta))$ with the probability taken over the joint posterior distribution of θ and y^{rep}

Evaluating outliers in Newcomb's speed of light data

- from GCSR textbook
- 66 measurements of speed of light; two low outliers
- what we want to evaluate: is normal density ok for likelihood?
- defined $T(y, \theta)$ as $\min(y_i)$
 - to check whether data with such extreme outliers could reasonably have come from a normal model