5. Bayesian Diagnostics

Residuals

We can extend the idea of residuals to a Bayesian setting.

Standardised (Pearson) residuals suppose that we observe $\mathbf{y}$ but not $\boldsymbol{\theta}$, then define

$$r_i(\boldsymbol{\theta}) = \frac{y_i - \mathbb{E}(Y_i|\boldsymbol{\theta})}{\sqrt{\mathrm{Var}\,(Y_i|\boldsymbol{\theta})}}$$

These quantities will have posterior distributions.

As in the frequentist setting, a very large or small residual indicates a point where the model does not fit well, and patterns in the residuals point to systemmatic problems with the model.

Residuals

Deviance residuals Recall that for a binomial sample $y_i \sim \text{bin}(m_i, p_i)$ the (scaled) deviance is

$$D(\mathbf{p}) = 2\sum_i \left(y_i \log \frac{y_i/m_i}{p_i} + (m_i - y_i) \log \frac{1 - y_i/m_i}{1 - p_i} \right) = \sum_i d_i(\mathbf{p})$$

and the i-th deviance residual is

$$r_i^D(\mathbf{p}) = \text{sign}(y_i - \mathbb{E}(Y_i|\mathbf{p}))\sqrt{d_i(\mathbf{p})}$$

Similarly if $y_i \sim \text{Poisson}(\lambda_i)$ then

$$r_i^D(\boldsymbol{\lambda}) = \text{sign}(y_i - \mathbb{E}(Y_i|\boldsymbol{\lambda}))\sqrt{2\,(y_i \log(y_i/\lambda_i) - y_i + \lambda_i)}.$$

Residuals

Because Bayesian residuals have (posterior) distributions and not just values, we can't just plot them (in order, or against the observations, or against a predictor variable).

Easiest is to represent each residual in the plot using a "boxplot", where the boxplot is based on the posterior distribution, rather than a sample of values.

Analogously to regular boxplots, the box gives the interquartile range. We can use the 95% credible interval.

Posterior predictive distribution

Is the model consistent with data? If the model fits, then replicated data generated under the model should look similar to observed data. To put it another way, the observed data should look plausible under the posterior predictive distribution.

Posterior predictive distribution:

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

We form 95% credible interval for y^{rep}. If the model fits, we would expect the observation y to fall in this interval.