# Model Checking (Chp 6)

- Need tools to assess (a) fit and (b) prediction

Main Bayesian Tool: Posterior Prediction Check

- Sample "replicated" data $y^{rep}$ and compare to $y^{obs}$

Rmk: $y^{rep}$ has identical explanatory variables (contrasted to $\tilde{y}$ predictive data)

Has same posterior

$\mathbb{P}(y^{rep}|y)$

$$\mathbb{P}(y^{rep}) = \int \mathbb{P}(y^{rep}|\theta)\mathbb{P}(\theta)\,d\theta \qquad \text{(Prior)}$$

$$\mathbb{P}(y^{rep}|y) = \int \mathbb{P}(y^{rep}|\theta)\mathbb{P}(\theta|y)\,d\theta \qquad \text{(Post.)}$$

Then to sample $y^{rep}$ we just need to

(a) Sample $\theta^{(s)} \sim \theta | y$

(b) Sample $y^{rep}_s \sim y^{rep} | \theta = \theta^s$

For the actual check, need a test function $T(y, \theta)$ – discrepancy measure

$T$ chosen to capture aspects not covered by the model

e.g. Ranks/Quantiles and Correlations

Freq Stat: $p_F = \mathbb{P}(T(y^{rep}; \theta_0) \geq T(y, \theta_0))$

Bayesian Stat: $p_B = \mathbb{P}(T(y^{rep}, \theta) \geq T(y, \theta) \,|\, y)$ – Prob over $(y^{rep}, \theta)$

$$= \mathbb{E}_{y^{rep}, \theta | y}\left[\mathbb{1}\left(T(y^{rep}, \theta) \geq T(y, \theta)\right)\right]$$

$$= \mathbb{E}_{\theta|y}\left[\mathbb{E}_{y^{rep}}\left[I\left(T(y^{rep},\theta) \geq T(y,\theta)\right)\right]\right]$$

$$= \int\int I\left(T(y^{rep},\theta) \geq T(y,\theta)\right) \mathbb{P}(y^{rep}|\theta)\, dy^{rep}\, \mathbb{P}(\theta|y)\, d\theta$$

Estimate $p_B$ by

1. Sample $\theta^{(s)} \sim \theta|y$
2. Sample $y_s^{rep} \sim y^{rep}|\theta = \theta^{(s)}$
3. $\hat{p}_B = \frac{1}{S}\sum_{s=1}^{S} \mathbb{1}\left[T(y_s^{rep},\theta_s) \geq T(y,\theta^{(s)})\right]$

Good values: $\boxed{p_B = 1/2}$. B.C. we want $T(y^{rep},\theta^{(s)}) \doteq T(y,\theta^{(s)})$

Q: How do you measure discrepancy? May just be better to directly compare dist.

Rmk: Not conducting a test.

Ex: $y_j|\theta_j \overset{ind}{\sim} Binom(n_j,\theta_j)$ $\qquad$ $\mathbb{E}[\theta_j] = p$

$\theta_j|k \overset{iid}{\sim} Beta(pk,(1-p)k)$ $\qquad$ $Var[\theta_j] = \frac{p(1-p)}{k+1}$

$k \sim Gamma(\beta,\beta)$ $\qquad$ $\mathbb{E}[k] = 1$, $Var(k) = 1/\beta$