ADVANCED BAYESIAN MODELING

Posterior Predictive Checking: An Example

Goal: Find any discrepancies between the fitted Bayesian model and the data.

One approach: **posterior predictive checking**

Basically, comparing the data with its posterior predictive distribution

Procedure

Basic version:

Step 1 Choose a statistic relevant to the type of discrepancy under investigation.

Step 2 Evaluate the statistic on the data.

Step 3 Find or approximate the distribution of the statistic under the posterior predictive distribution for the full data set.

Step 4 Compare the statistic value to its posterior predictive distribution. If the value is not representative (too extreme), claim evidence for a discrepancy between the Bayesian model and data.

Flint Data Example

Recall Flint water data:

$$y_i \quad = \quad \text{logarithm of first-draw lead level (ppb), for observation } i$$

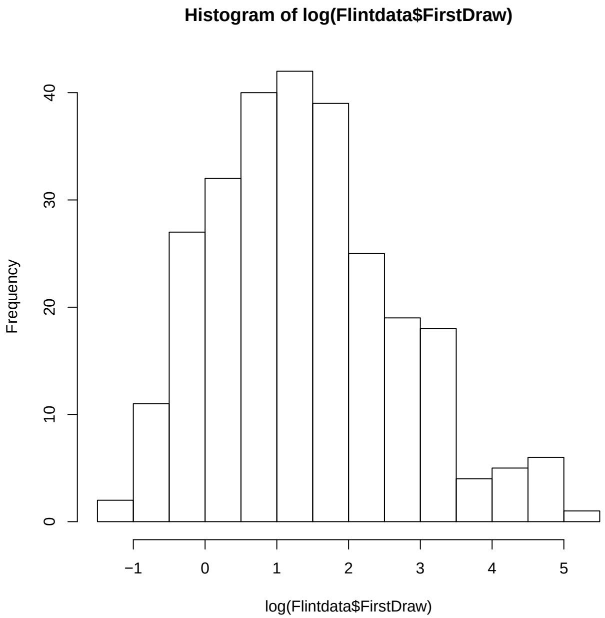
Our model assumed a normal sampling distribution:

$$y_1, \ldots, y_n \mid \mu, \sigma^2 \quad \sim \quad \text{iid } \mathrm{N}(\mu, \sigma^2)$$

But histogram suggested asymmetry:

```
> Flintdata <- read.csv("Flintdata.csv", header=TRUE, row.names=1)

> hist(log(Flintdata$FirstDraw))
```



Histogram of log(Flintdata$FirstDraw)
Frequency
0
10
20
30
40
−1
0
1
2
3
4
5
log(Flintdata$FirstDraw)

Step 1: Choose Statistic

Many ways to measure asymmetry: skewness (third moment), mean versus median, quantile-based measures, etc.

We try (compare Groeneveld & Meeden, 1984)

$$\frac{\hat{q}_{0.9} + \hat{q}_{0.1} - 2\hat{q}_{0.5}}{\hat{q}_{0.9} - \hat{q}_{0.1}}$$

where $\hat{q}_{\alpha}$ is empirical α (lower) quantile.

It's roughly zero when distribution is symmetric, and tends to be away from zero (in either direction) otherwise.

Step 2: Evaluate Statistic

```
> ( T <- sum(c(1,1,-2) * quantile(log(Flintdata$FirstDraw), c(0.9,0.1,0.5))) /
+        sum(c(1,-1) * quantile(log(Flintdata$FirstDraw), c(0.9,0.1))) )
[1] 0.1676124
```

Step 3: Find Posterior Predictive Distribution

First need a posterior distribution ...

The standard noninformative prior

$$p(\mu, \sigma^2) \;\propto\; (\sigma^2)^{-1} \qquad\qquad \sigma^2 > 0$$

led to posterior

$$\begin{aligned} \mu \mid \sigma^2, y \;&\sim\; \mathrm{N}\big(\bar{y},\, \sigma^2/n\big) \\ \sigma^2 \mid y \;&\sim\; \text{Inv-}\chi^2\big(n-1,\, s^2\big) \end{aligned}$$

based on sample statistics n, $\bar{y}$, s^2.

We characterize posterior predictive distribution using direct simulation ...

Simulate X from χ^2_{n-1}, let

$$\sigma^2_{\mathsf{sim}} \;=\; (n-1)s^2\big/X$$

then simulate

$$\mu_{\mathsf{sim}} \;\text{ from }\; \mathrm{N}(\bar{y},\, \sigma^2_{\mathsf{sim}}/n)$$

For a *replicate* data set, simulate

$$y_1^{\mathsf{rep}}, \ldots, y_n^{\mathsf{rep}} \quad \mathrm{iid}\text{ from} \quad \mathrm{N}(\mu_{\mathsf{sim}}, \sigma^2_{\mathsf{sim}})$$

Repeat the whole process independently for each replicate data set.

For 1000 replicate data sets:

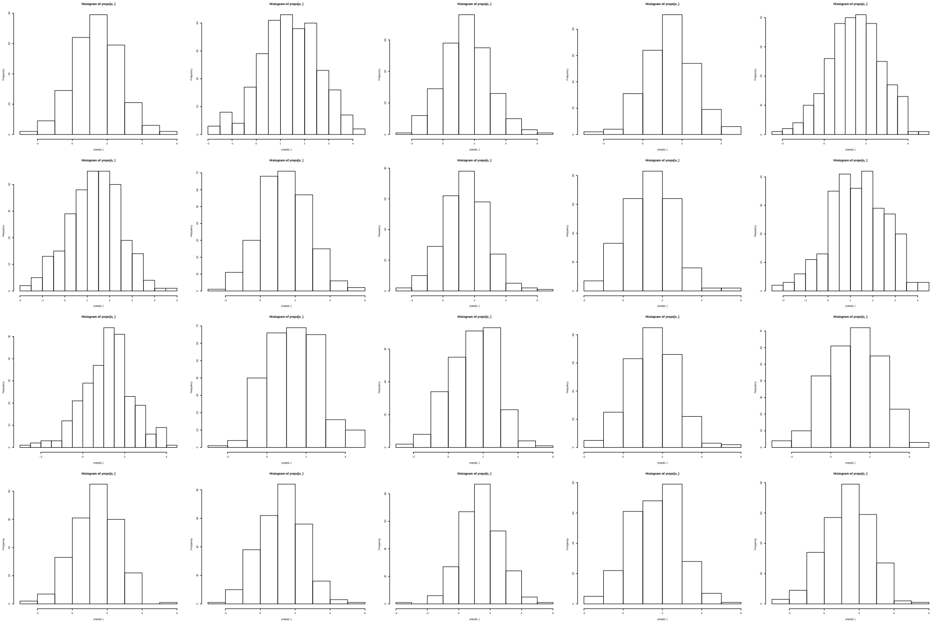
```
> n <- nrow(Flintdata)
> ybar <- mean(log(Flintdata$FirstDraw))
> s.2 <- var(log(Flintdata$FirstDraw))

> post.sigma.2.sim <- (n-1) * s.2 / rchisq(1000, n-1)
> post.mu.sim <- rnorm(1000, ybar, sqrt(post.sigma.2.sim / n))

> yreps <- matrix(NA, 1000, n)
> for(s in 1:1000)
+   yreps[s,] <- rnorm(n, post.mu.sim[s], sqrt(post.sigma.2.sim[s]))
```

For histograms of the first 20 ...

```
> par(mfrow=c(4,5))
> for(s in 1:20) hist(yreps[s,])
```


Histogram of yreps[s,]
Frequency
yreps[s,]

These histograms don't look as asymmetric as the data, but difficult to tell ...

Now compute simulated posterior predictive test statistics:

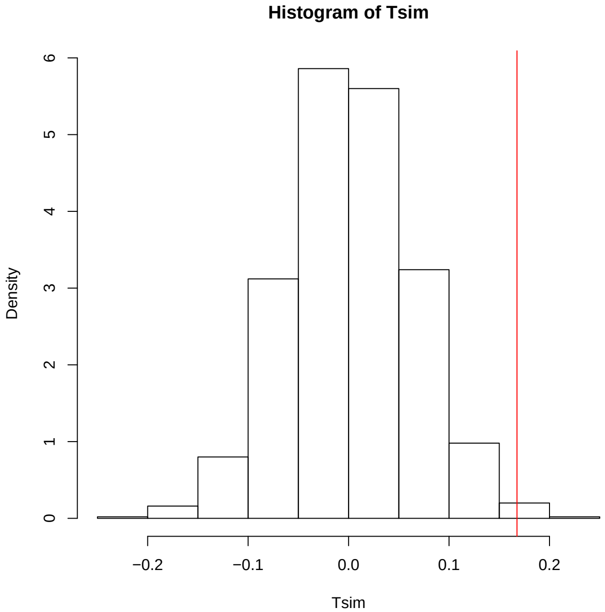
```
> Tsim <- numeric(1000)
> for(s in 1:1000)
+   Tsim[s] <- sum(c(1,1,-2) * quantile(yreps[s,], c(0.9,0.1,0.5))) /
+              sum(c(1,-1) * quantile(yreps[s,], c(0.9,0.1)))

> summary(Tsim)
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.2091577 -0.0403291  0.0006747  0.0015260  0.0433195  0.2289606
```

Step 4: Compare

How extreme is the data-computed statistic relative to its (simulated) posterior predictive distribution?

```
> hist(Tsim, freq=FALSE)
> abline(v=T, col="red")
```


Histogram of Tsim
Density
0
1
2
3
4
5
6
−0.2
−0.1
0.0
0.1
0.2
Tsim

To get something like a (two-sided) p-value:

```
> mean(abs(Tsim) >= abs(T))
[1] 0.01
```

Only 10 out of 1000 posterior predictive replicates had an asymmetry measure at least as extreme as the actual data.

Conventionally, indicates evidence of asymmetry ($p < 0.05$)