Bayesian Statistics, 22S:138
Lab 7, Nov. 17, 2006
Discrepancy Measures Based on the Posterior Predictive Distribution
With WinBUGS

Textbook example: Newcomb's speed of light data

The code for this example is on the course web page under Handouts as newcomb.bug.

```
model
{
for(i in 1:N) {
     y[i] ~ dnorm(mu, tau)
     ypred[i] ~ dnorm(mu,tau)        # predicted values
                                     # (draws from posterior predictive dist'n)
}
mu ~ dnorm(0, .00001)
tau ~ dgamma(.0001, .0001)
sigmasq <- 1 / tau
s <- 1
smallest <- ranked(ypred[], s)       #  smallest predicted value
smallthan <- step(y[6] - smallest)   #   1 if smallest
              # predicted value is smaller than smallest obs value
}

data
list(y = c(28, 26, 33, 24, 34, -44, 27, 16, 40, -2, 29, 22, 24, 21, 25,
30, 23, 29, 31, 19, 24, 20, 36, 32, 36, 28, 25, 21, 28, 29,37, 25, 28,
26, 30, 32, 36, 26, 30, 22, 36, 23,27, 27, 28, 27, 31, 27, 26, 33, 26,
32, 32, 24, 39, 28, 24, 25, 32, 25, 29, 27, 28, 29, 16, 28), N=66)

inits
list(mu=20, tau=1 )
```

After running the sampler for some burn-in iterations, which quantity will you monitor to do the same check as in the example on p. 166 of the textbook?

Checking for an outlier in the squirrel data

This example is taken from the Berry book. It concerns a study performed at the University of Minnesota in which researchers studied squirrels to determine the relationship between the amount of rainfall in the preceding 3 days and the osmolarity (concentration) of the squirrels' urine. (As Dave Barry would say, "I am not making this up"!).

The code for this example is on the course web page under Handouts as squirrel3.bug.

```
  model
  {
   x.bar <- mean(x[]) ;
   for(i in 1:N){
      Y[i] ~ dnorm(mu[i], tau)
      Ypred[i] ~ dnorm(mu[i], tau)              # predicted values
      aresid[i] <- abs(Y[i] - mu[i])            # actual residuals
      arespred[i] <- abs(Ypred[i] - mu[i])      # predicted residuals
      mu[i] <- alpha + beta * (x[i] - x.bar)
   }
   s <-8
   largest <- ranked(arespred[], s)
   largthan <- step(largest - aresid[3])
   sigma <- 1/sqrt(tau)
   alpha ~ dnorm(0, 1.0E-6)
   beta ~ dnorm(0, 1.0E-6)
   tau ~ dgamma(1.0E-3, 1.0E-3)
  }

list( Y = c(1479, 1500, 516, 1815, 1118, 1702, 1348, 1337), x = c(0.45,
0.0, 0.0, 0.24, 0.12, 0.0, 0.35, 0.44),
N = 8)

list(alpha = 0, beta = 0, tau = 1)
```

Note: Also, try the "Fit" facility under the Statistics menu.

- Enter Y in the "data" box. This is the response variable.
- If you enter *mu* in the "pred" box, you will get a plot of the pointwise confidence intervals for the mean of Y at each value of X.
- If you enter *Ypred* in the "pred" box, you will get a plot of the pointwise prediction intervals for individual observations at each value of X.
- Enter x in the "x" box. This is the predictor variable.