Chapter 7.6 Predictive Checking

Jim Albert and Monika Hu

Chapter 7 Learning About a Binomial Probability

Prior predictive checking

- In Chapter 7.5: the posterior predictive distribution is used for learning about future data
- The prior predictive density is also useful in model checking

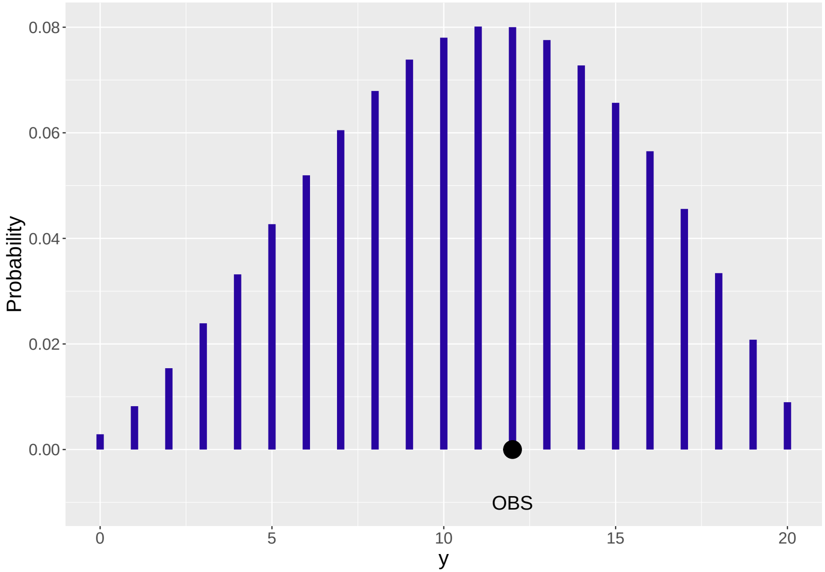
$$f(p, Y = y) = f(Y = y \mid p)\pi(p)$$

$$f(p, Y = y) = \pi(p \mid Y = y)f(Y = y)$$

Prior predictive checking cont'd

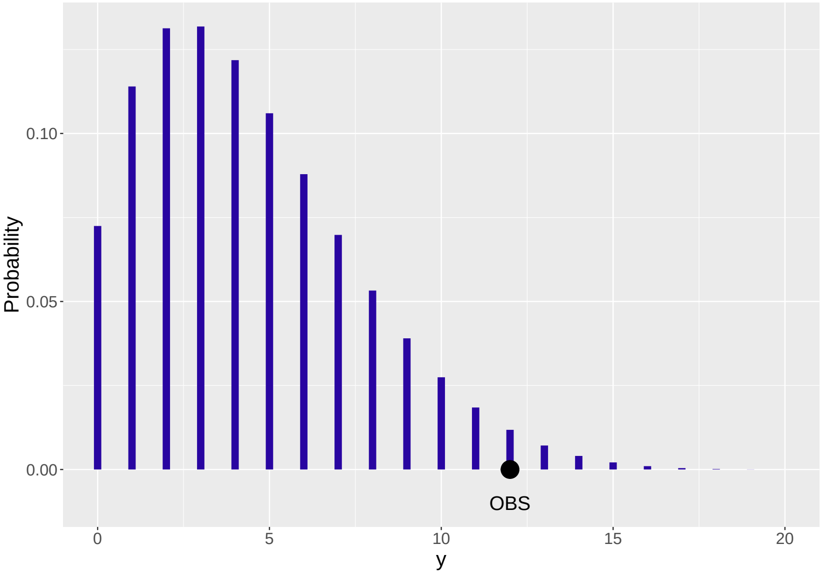
```r
df <- data.frame(y = 0:20,
        Probability = pbetap(c(3.06, 2.56), 20, 0:20))

prob_plot(df, Color = crcblue, Size = 3) +
  geom_point(data = data.frame(y = 12,
                               Probability = 0),
             size = 7) +
  increasefont() +
  annotate(geom = "text", x = 12, y = -0.01,
           label = "OBS", size = 6)
```

Prior predictive checking cont'd



Prior predictive checking cont'd

- Another prior: Beta(2.07, 7.32)



Comparing Bayesian models

- The prior predictive distribution is also useful in comparing two Bayesian models
- π_1 the owner's prior, π_2 the worker's prior
- Assume a mixture prior

$$\pi(p) = q\pi_1(p) + (1-q)\pi_2(p)$$

- The posterior density of p is proportional to:

$$\pi(p \mid Y = y) \propto \left[q\pi_1(p) + (1-q)\pi_2(p)\right] \times \binom{n}{y} p^y (1-p)^{n-y}$$

Comparing Bayesian models cont'd

- After some manupulation

$$\pi(p \mid Y = y) = q(y)\pi_1(p \mid Y = y) + (1 - q(y))\pi_2(p \mid Y = y)$$

- The quantity $q(y)$ represents the posterior probability of the owner's prior

$$q(y) = \frac{qf_1(Y = y)}{qf_1(Y = y) + (1 - q)f_2(Y = y)}$$

- The posterior odds of the model probabilities

$$\frac{P(Prior\,1 \mid Y = y)}{P(Prior\,2 \mid Y = y)} = \frac{q(y)}{1 - q(y)} = \left[\frac{q}{1 - q}\right]\left[\frac{f_1(Y = y)}{f_2(Y = y)}\right]$$

Comparing Bayesian models cont'd

$$\frac{P(Prior\,1 \mid Y=y)}{P(Prior\,2 \mid Y=y)} = \frac{q(y)}{1-q(y)} = \left[\frac{q}{1-q}\right]\left[\frac{f_1(Y=y)}{f_2(Y=y)}\right]$$

- The ratio $q/(1-q)$ represents the prior odds of the owner's prior
- The term $f_1(Y=y)/f_2(Y=y)$, the ratio of the predictive densities, is called the Bayes factor: it reflects the relative abilities of the two priors to predict the observation y

Comparing Bayesian models cont'd

- Find the Bayes factor: the function `binomial.beta.mix()`
- Inputs: the prior probabilities of the two models (priors), and the vectors of Beta shape parameters that define the owner's prior and the worker's prior

```
probs <- c(0.5, 0.5)
beta_par1 <- c(3.06, 2.56)
beta_par2 <- c(2.07, 7.32)
beta_par <- rbind(beta_par1, beta_par2)
output <- binomial.beta.mix(probs, beta_par, c(12, 8))
(posterior_odds <- output$probs[1] / output$probs[2])
```

```
## beta_par1
##  6.777823
```

Comparing Bayesian models cont'd

- The prior odds $q/(1-q)$ is equal to one
- The posterior odds is equal to the Bayes factor
- Interpretation: is that for the given observation (12 successes in 20 trials), there is 6.77 times more support for the owner's prior than for the worker's prior

Posterior predictive checking

- To simulate one replicated dataset
  - Simulate a parameter from its posterior distribution
  - Simulate new data from the data model given the simulated parameter value

Posterior predictive checking cont'd

- In the beta-binomial situation, the posterior of the proportion p is $\mathrm{Beta}(a+y, b+n-y)$
- To simulate a new data point $\tilde{Y} = \tilde{y}$
  - Simulate a proportion value $p^{(1)}$ from the beta posterior
  - Simulate a new data point $\tilde{y}^{(1)}$ from a binomial distribution with sample size n and probability of success $p^{(1)}$

Posterior predictive checking cont'd

- To obtain a sample of size S from the posterior predictive distribution

$$p^{(1)} \sim \mathrm{Beta}(a+y, b+n-y) \rightarrow \tilde{y}^{(1)} \sim \mathrm{Binomial}(n, p^{(1)})$$

$$p^{(2)} \sim \mathrm{Beta}(a+y, b+n-y) \rightarrow \tilde{y}^{(2)} \sim \mathrm{Binomial}(n, p^{(2)})$$

$$\vdots$$

$$p^{(S)} \sim \mathrm{Beta}(a+y, b+n-y) \rightarrow \tilde{y}^{(S)} \sim \mathrm{Binomial}(n, p^{(S)})$$

- The sample $\tilde{y}^{(1)}, \ldots, \tilde{y}^{(S)}$ is an approximation to the posterior predictive distribution that is used for model checking