

# Learning About a Binomial Proportion

Jim Albert

May 28, 2014

## Constructing a Beta Prior

Suppose we are interested in the proportion  $p$  on sunny days in my town. The function `bayes.select` is a convenient tool for specifying a beta prior based on knowledge of two prior quantiles. Suppose my prior median for the proportion of sunny days is .2 and my 75th percentile is .28.

```
> library(LearnBayes)
> beta.par <- beta.select(list(p=0.5, x=0.2), list(p=0.75, x=.28))
> beta.par
```

```
[1] 2.95 10.82
```

A `beta(2.95, 10.82)` prior matches this prior information

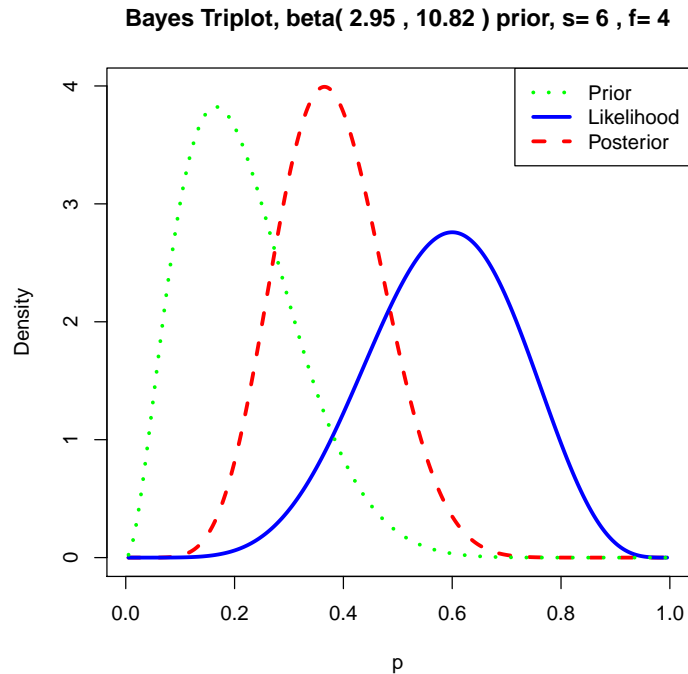
## Updating with Data

Next, I observe the weather for 10 days and observe 6 sunny days. (There are 6 “successes” and 4 “failures”.) The posterior distribution is beta with shape parameters  $2.95 + 6$  and  $10.82 + 4$ .

## Triplot

The `triplot` function shows the prior, likelihood, and posterior on the same display; the inputs are the vector of prior parameters and the data vector.

```
> triplot(beta.par, c(6, 4))
```



## Simulating from Posterior to Perform Inference

One can perform inference about the proportion  $p$  by simulating a large number of draws from the posterior and summarizing the simulated sample. Here the `rbeta` function is used to simulate from the beta posterior and the `quantile` function is used to construct a 90 percent probability interval for  $p$ .

```
> beta.post.par <- beta.par + c(6, 4)
> post.sample <- rbeta(1000, beta.post.par[1], beta.post.par[2])
> quantile(post.sample, c(0.05, 0.95))
```

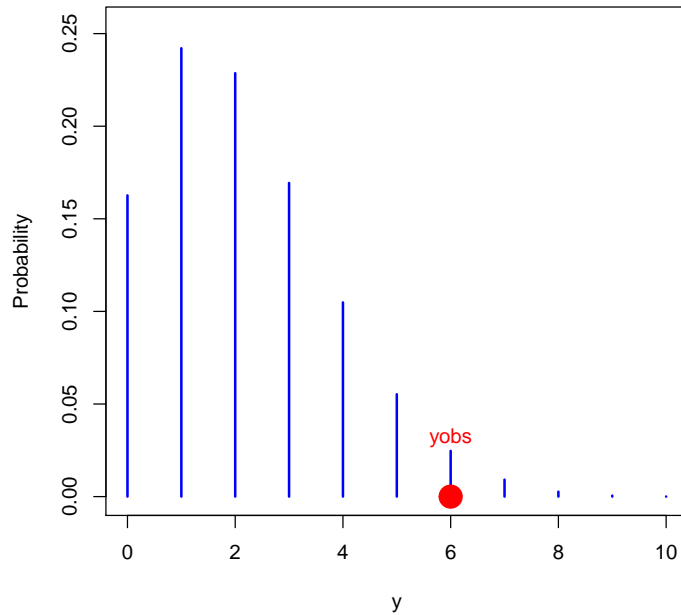
```
      5%      95%
0.2236582 0.5565551
```

## Predictive Checking

One can check the suitability of this model by means of a predictive check. The function `predplot` displays the prior predictive density for the number of successes and overlays the observed number of successes.

```
> predplot(beta.par, 10, 6)
```

Predictive Dist., beta( 2.95 , 10.82 ) prior, n= 10 , yobs= 6



The observed data is in the tail of the predictive distribution suggesting some incompatibility of the prior information and the sample.

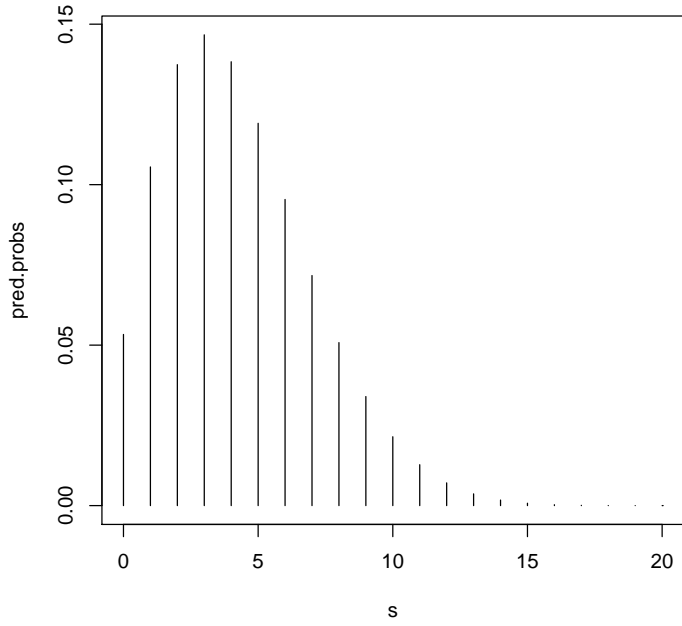
## Prediction of a Future Sample

Suppose we want to predict the number of sunny days in the future 20 days. The function `pbetap` computes the posterior predictive distribution with a beta prior. The inputs are the vector of beta prior parameters, the future sample size, and the vector of number of successes in the future experiment.

```
> n <- 20
> s <- 0:n
> pred.probs <- pbetap(beta.par, n, s)
> plot(s, pred.probs, type="h")
> discint(cbind(s, pred.probs), 0.90)
```

```
$prob
[1] 0.9181962
```

```
$set
[1] 0 1 2 3 4 5 6 7 8
```



The probability that we will observe between 0 and 8 successes in the future sample is .92.