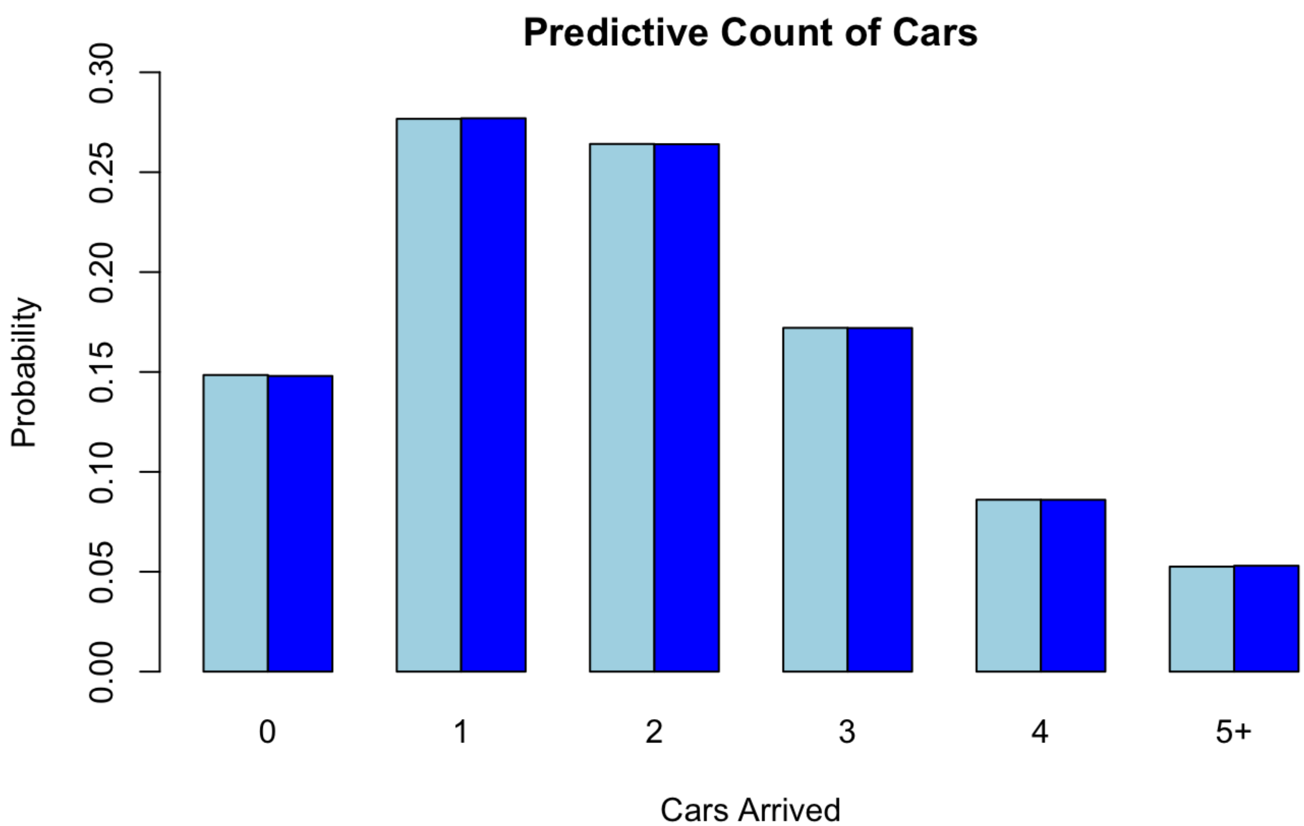
Assignment 5
Jericho McLeod

Code ▾

Problem 1: This problem concerns the automobile data from Assignments 3 and 4. As in Assignment 4, assume that counts of cars per 15-second interval are independent and identically distributed Poisson random variables with unknown mean L. Assume a uniform prior distribution for L. (This is a gamma distribution with shape 1 and scale infinity.) Using the posterior distribution from Problem 1 of Assignment 4, find the predictive distribution for the number of cars in the next 15-second interval. Name the family of distributions and the parameters of the predictive distribution. Find the predictive probability that 0, 1, 2, 3, 4, and more than 4 cars will pass the point in the next 15 seconds. Compare with your answer to Problem 1e of Assignment 3. Discuss.



These are the same (excepting minor pixel count differences stemming from the number of significant digits). This is due the fact that the poisson distribution discrete distribution, while the negative binomial is the likelihood of seeing some number of successes before a failure. We have essentially just used the negative binomial to reproduce the probabilities obtained in the poisson by using the shape and scale to find the probabililties of the discrete cuonts from the poisson.

Problem 2: In previous years, students in this course collected data on people's preferences in the two Allais gambles from Assignment 2. For this problem, we will assume that responses are independent and identically distributed, and the probability is p that a person chooses both B in the first gamble and C in the second gamble.

A. Assume that the prior distribution for π is Beta(1, 3). Find the prior mean and standard deviation for π. Find a 95% symmetric tail area credible interval for the prior probability that a person would choose B and C. Do you think this is a reasonable prior distribution to use for this problem? Why or why not?

Hide

```
print(paste0("Prior Mean:  ",prior_mean))
```

```
[1] "Prior Mean:  0.25"
```

Hide

```
print(paste0("Prior Standard Deviation:  ",prior_sd))
```

```
[1] "Prior Standard Deviation:  0.433012701892219"
```

Hide

```
print("Prior Credible Interval:")
```

```
[1] "Prior Credible Interval:"
```

Hide

```
print(cred_int)
```

```
[1] 0.008403759 0.707598226
```

Beginning from a level distribution is acceptable without any expert opinion. In this case, we know that individuals do not typically choose purely at random, but have bias towards answers. However, given that we do not currently have any quantifiable expert opinions regarding how much bias, a level distribution is reasonable.

B. In 2009, 19 out of 47 respondents chose B and C. Find the posterior distribution for the probability π that a person in this population would choose B and C. Find the posterior mean and standard deviation, and a 95% symmetric tail area credible interval for π. Do a triplot.

Hide

```
print(paste0("Posterior Mean:  ",posterior_mean))
```

```
[1] "Posterior Mean:  0.392156862745098"
```

Hide

```
print(paste0("Posterior Standard Deviation:  ",posterior_sd))
```

```
[1] "Posterior Standard Deviation:  0.488231356783872"
```

Hide

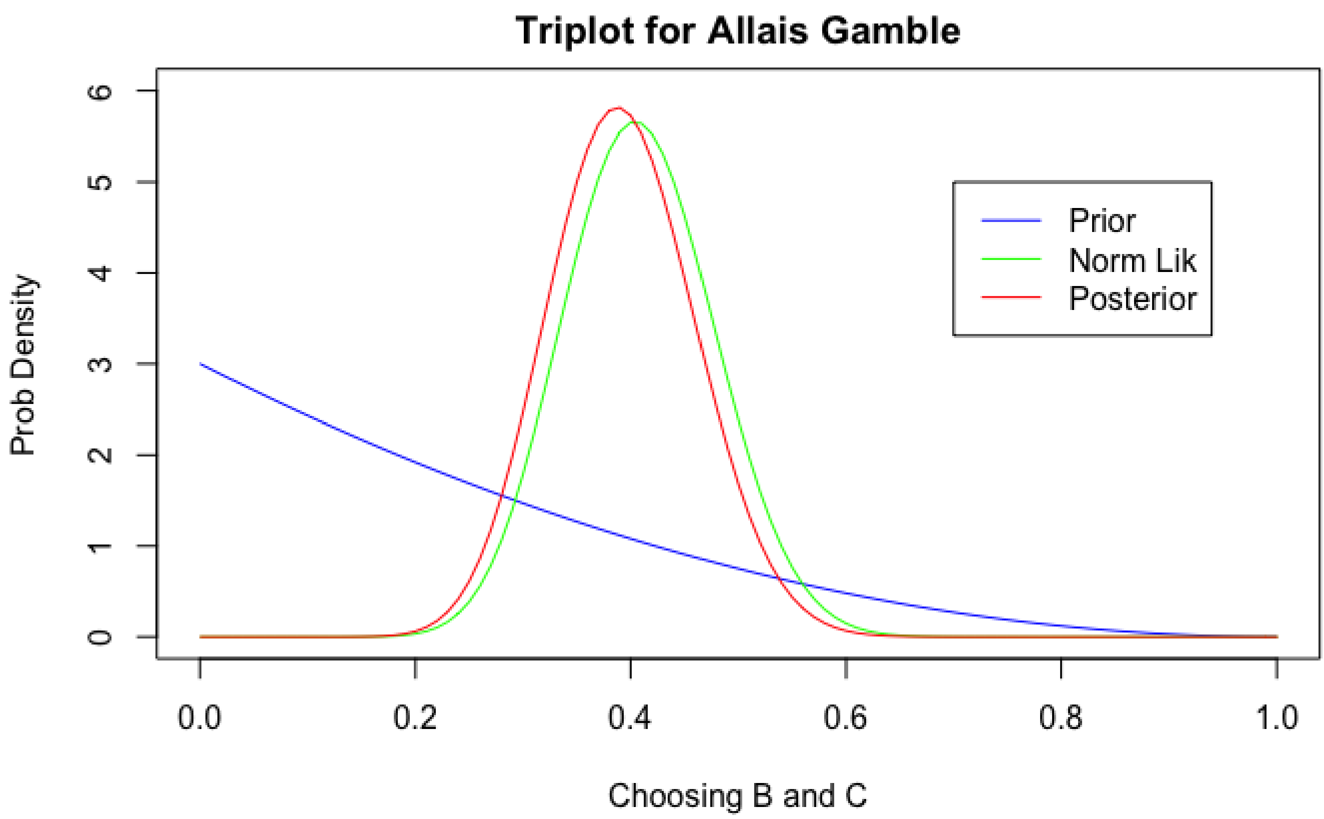
```
print("Posterior Credible Interval:")
```

```
[1] "Posterior Credible Interval:"
```

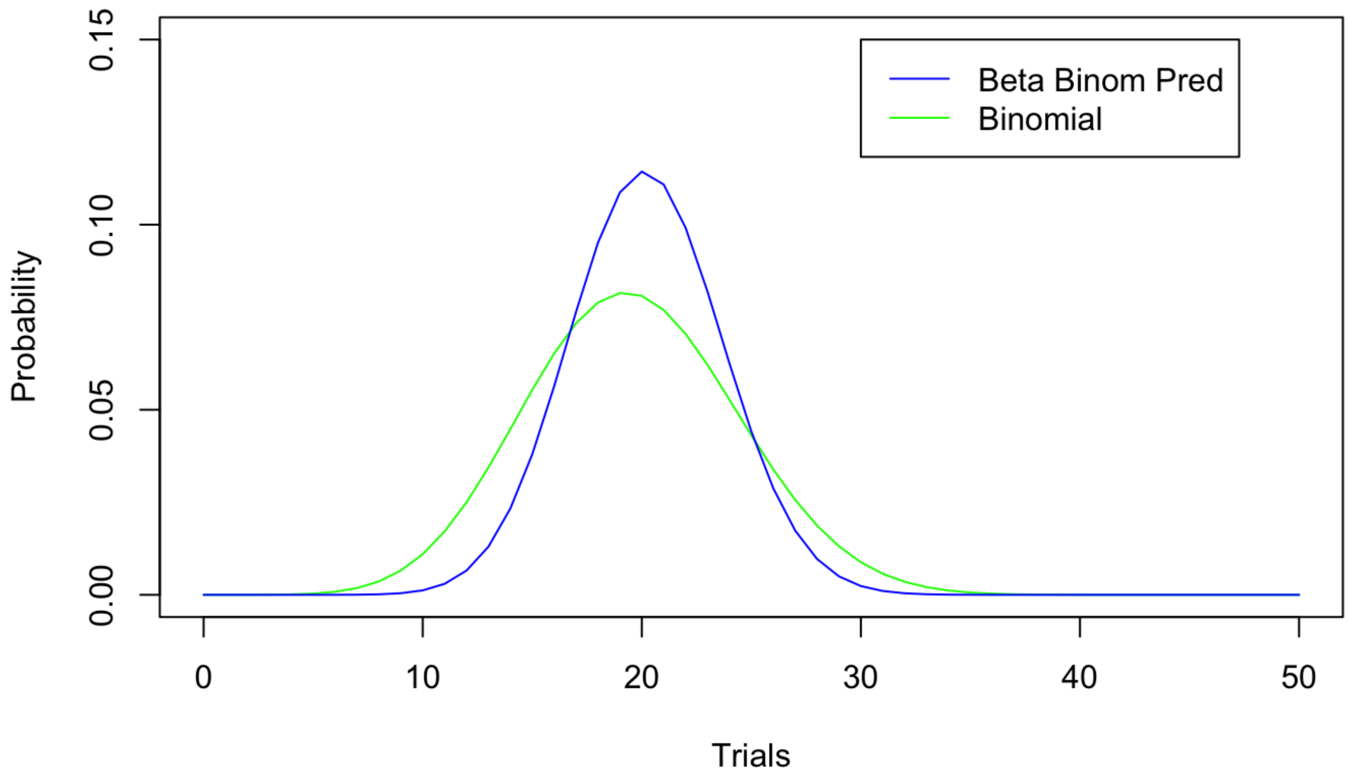
Hide

```
print(cred_int)
```

```
[1] 0.2640784 0.5282508
```



C. Find the predictive distribution for the number of B and C responses in a future sample of 50 people drawn from the same population. Compare with a Binomial distribution using a point estimate of the probability of choosing B and C.



D. Comment on your results.

The binomial point estimate is more dispersed compared to the beta binomial predictions for number of successes in the next 50 trials. This is because, while the binomial distribution is the count of successes at a given probability, the beta binomial draws successes from a beta distribution.