***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 Λ. Assume a uniform prior distribution for Λ. (As for Assignment 4, you can approximate this prior distribution by using a Gamma distribution with shape 1 and scale 10,000.) 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.

***Solution*:**

The posterior distribution for Λ from Problem 1 of Assignment 4 is a Gamma distribution with shape $\alpha = 41$ and scale $\beta = 0.0476$. The Poisson / Gamma predictive distribution is a negative binomial distribution. We are predicting just one time period, so the parameters are:

size $\alpha = 41$, and
probability $p = 1/(1+\beta) = 0.9545$

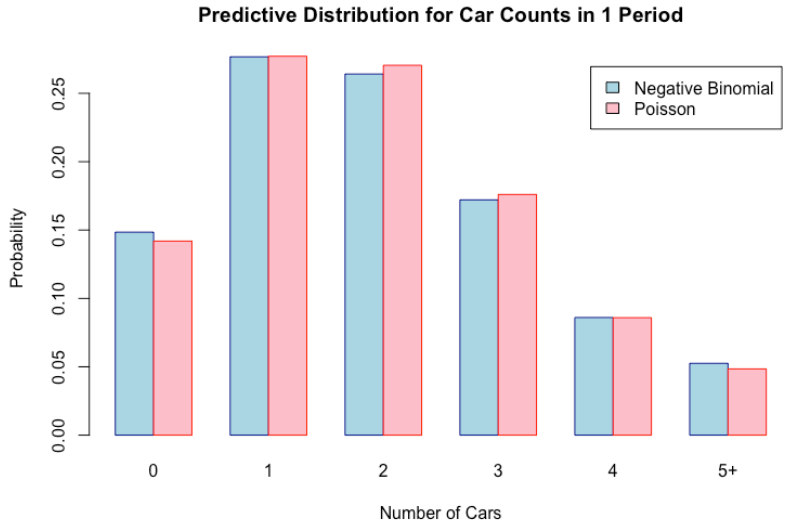
These probabilities, as computed by the R `dnbinom` function, are shown in the table. Results from Problem 1e of Assignment 3 (discretized version of this problem) and the Poisson probability mass function with mean $\alpha\beta = 1.95$ are shown for comparison.

Number of Cars	Predictive Probability	Predictive Probability (discretized)	Poisson Probability $\Lambda = 1.95$
0	0.1485	0.148	0.1419
1	0.2767	0.277	0.2771
2	0.2641	0.264	0.2705
3	0.1721	0.172	0.1760
4	0.0860	0.086	0.0859
5+	0.0526	0.053	0.0485

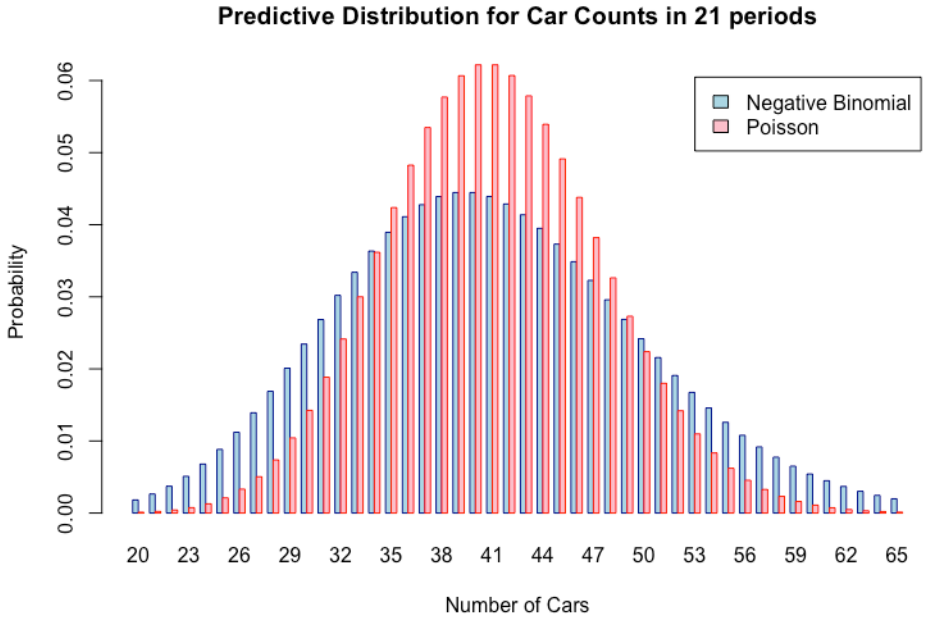
The results from the discretized analysis in Assignment 3 are identical to the results shown here to 3 decimal places.

A plot of the predictive probability mass function is shown below. The plot also shows the Poisson probability mass function with mean $\alpha\beta = 1.95$. The Bayesian predictive probabilities and the Poisson probabilities using the point estimate are very close to

each other. As expected, the Bayesian probabilities are higher at the extremes (0 and 4+ cars) and slightly lower in the middle (1-3 cars), but the differences are very small because we have enough data for accurate prediction a short time into the future.



Our uncertainty about Λ matters more if we try to predict a longer time into the future.



For example, let's look at predicting the total number of cars in the next 21 time periods (the same length of time as the data we have seen already). This was not required for the assignment, but I am including it for comparison. To predict 21 time steps, our predictive distribution is a negative binomial distribution with size 41 and

probability $1/(1+n\beta) = 0.5$. Comparing with a Poisson distribution with mean equal to our point estimate $n\alpha\beta$=41 of the number of cars in the next 21 time steps, we see a big difference between the two distributions. There is a much greater probability in the tail areas of the negative binomial predictive distribution, and the Poisson distribution assigns much higher probability to values near 41.

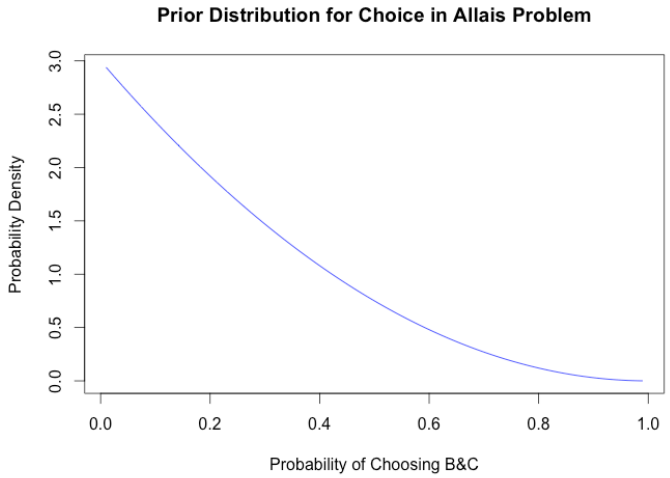
***Problem 2*:**

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 π 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?
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.
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.

***Solution*:**

Part a. It is convenient to use a prior distribution in the Beta family because it is conjugate to the Binomial likelihood function, which simplifies Bayesian updating.



A plot of the prior density function is shown at left. The prior distribution has mean $\alpha/(\alpha+\beta) = 0.25$ and standard deviation $\sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}} = 0.194$.

An expected value of 0.25 would be reasonable if we believed that people were choosing randomly among the four options A&C, B&C, A&D, and B&D. In fact, we will learn later this semester that a Beta(1,3) distribution is the marginal distribution we would obtain if we placed a uniform distribution on all probability vectors (p_1, p_2, p_3, p_4) on four categories (i.e. a uniform distribution on

four non-negative numbers that sum to 1). In other words, this is the prior distribution we might assign if all we knew was that there were four possible responses, and knew nothing about the relative likelihoods of the four responses. However, we actually do know something about the relative probabilities of the categories. According to the research literature, people choose B&C more often than 25% of the time. The center of the prior distribution does not reflect this knowledge. But the prior distribution does show a very high degree of uncertainty about the probability. A 95% symmetric tail area credible interval for the prior distribution of Θ is [0.008, 0.708]. The virtual count is small in relation to the sample size, so the data will have a large impact on the posterior distribution relative to the prior. Therefore, we won't go too far wrong using this prior distribution.

Part b. We will use the Binomial/Beta conjugate pair. The prior distribution is a member of the conjugate Beta(α, β) family, with α=1 and β=3. Therefore, the posterior distribution is a Beta(α^*, β^*) distribution, with $\alpha^*=\alpha+x$ = 1+19 = 20, and $\beta^*=\beta+n-x$ = 3+28 = 31. To summarize:

Prior distribution:	Beta(1,3)
Sample:	x = 19, n=47
Posterior distribution:	Beta(20, 31)

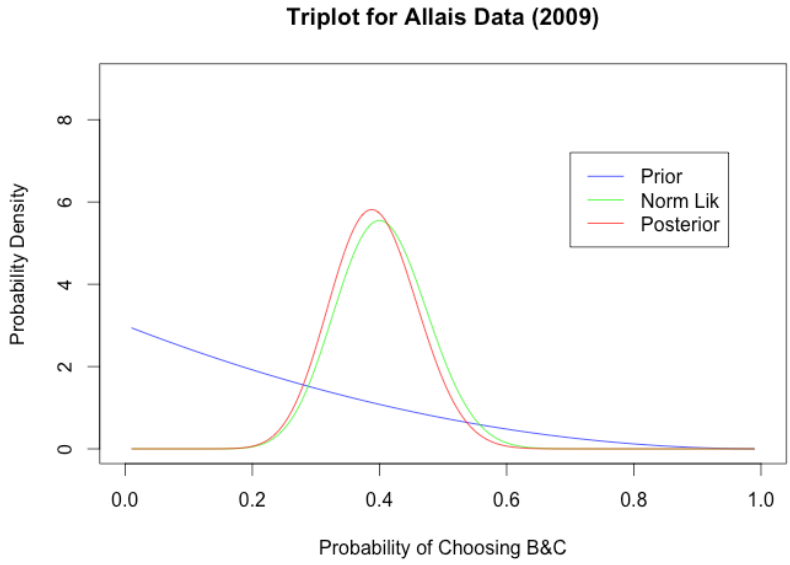
The mean and variance of a Beta(α,β) distribution are $\alpha/(\alpha+\beta)$ and $\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

To find percentiles of the posterior distribution, we use the inverse Beta cdf. In Excel, the appropriate function is BETA.INV(theta, alpha, beta). In R, we use `qbeta(prob, shape1=alpha, shape2=beta)`. Using these formulas, we find:

	Mean	Std Dev	2.5th Percentile	50th Percentile	97.5th Percentile
Prior Θ	0.25	0.194	0.0084	0.206	0.708
Posterior Θ	0.392	0.068	0.264	0.391	0.528

A 95% symmetric tail area credible interval for Θ is [0.264, 0.528]

The triplot shows the prior, normalized likelihood, and posterior distributions. The normalized likelihood is a Beta(20,29) density function. The posterior distribution is focused on slightly smaller values and is slightly narrower than the normalized likelihood. This is because the prior mean is smaller than the sample mean, and the posterior distribution has slightly more information than the prior distribution.

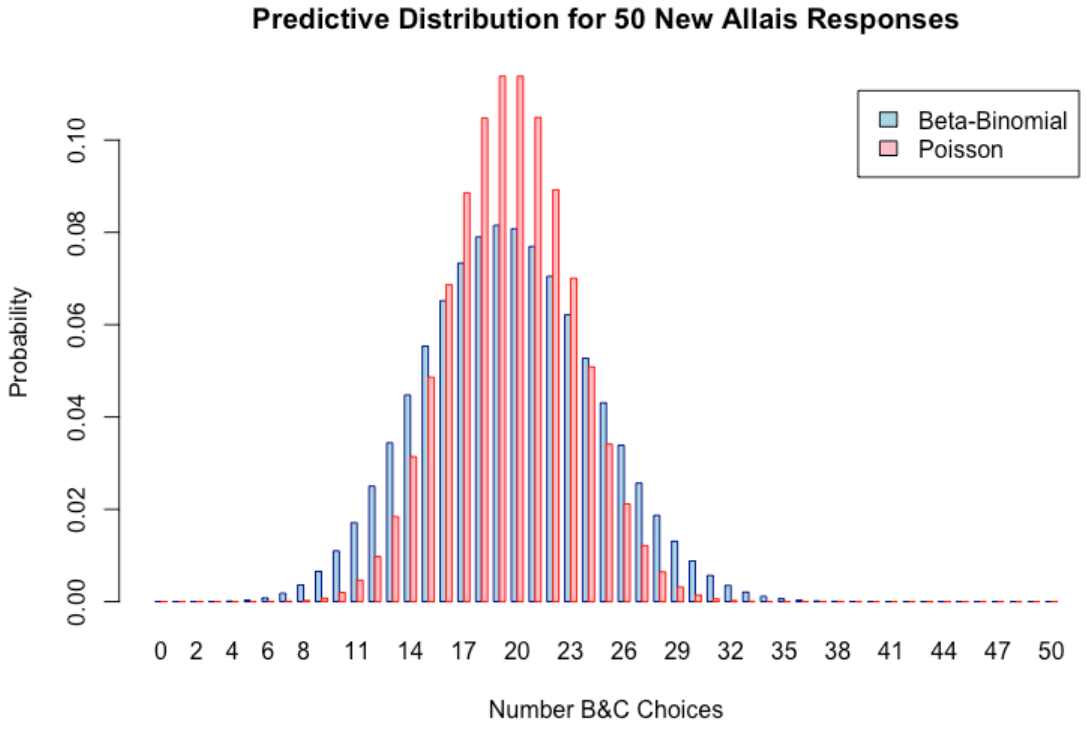


Part c. Find the predictive distribution for the number of B and C responses in a future sample of 50 people drawn from the same population. Compare with a Binomial distribution using a point estimate of the probability of choosing B and C.

The predictive distribution, as explained on page 57 of the Unit 3 notes, is a beta-binomial distribution. We will use the functions in the `rmutil` package, because it can be downloaded directly from the CRAN repository. The `rmutil` package is parameterized by probability, overdispersion, and size. As described in the Unit 3 notes, the predictive distribution is a negative binomial distribution with the following parameters:

- Size: 50
- Probability: $\frac{\alpha}{\alpha+\beta} = \frac{20}{51} = 0.392$
- Overdispersion : $\alpha + \beta = 51$

To compare with the binomial probabilities, I plotted the beta-binomial predictive distribution and a binomial distribution with size 50 and probability 0.392 on the same axes. I plotted the entire range of possible values from 0 to 50, although values below 5 and above 35 are extremely unlikely, so we could have plotted the range from 5 to 35. As with the poisson-gamma predictive distribution, we see that the beta-binomial distribution is more spread out than the binomial distribution. There is more probability on smaller and larger values, and less probability on values around the center. This is because the beta-binomial distribution includes uncertainty about the probability of choosing B and C, and not just uncertainty about how many B&C values there will be if the probability is assumed known.



To illustrate the difference between the two distributions, I found approximate 90% intervals for each distribution (we can't find exact 90% intervals because the distributions are discrete) and found the probability of each interval under both distributions. Only about 79% of the probability mass of the beta-binomial distribution is contained in a 92% interval for the binomial distribution. About 98% of the binomial probability is contained in a 90.4% interval for the beta-binomial distribution. These results emphasize that the beta-binomial predictive distribution has more variability than a binomial distribution with the same mean.

Interval	Beta-Binomial Probability	Binomial Probability
[14, 25]	0.785	0.919
[12, 27]	0.904	0.980

Part d. Friends, family and acquaintances of students in the 2009 Bayesian inference class at Mason appear less likely to choose B & C than the populations tested in other published research on the Allais paradox. The literature finds more than half of the people choose B&C, whereas about 40% of our 2009 sample did. After observing 47 trials of the experiment, the posterior distribution for the probability of choosing B&C has mean of about 0.4 and a 95% credible interval of about 0.26 to 0.53. That is, we expect somewhere between ¼ and ½ of the people in this population to choose B&C.

A Monte Carlo estimate of the posterior distribution based on a sample of 10,000 observations gave essentially the same results as our exact analysis.

The predictive distribution for a future sample of size 50 is a beta-binomial distribution, with a 90.4% probability that between 12 and 27 people will choose B&C.

I collected data on this problem several other years and found similar results.