SYST 664: HW Assignment 4

Jake Whalen

February 19, 2018

# Problem 1

This problem continues analysis of the automobile traffic data from Assignment 3. Transforming the arrival times to counts of cars in each 15-second interval gives the following table of counts:

|  |  |
| --- | --- |
| Number.of.Cars | Number.of.Occurences |
| 0 | 3 |
| 1 | 5 |
| 2 | 7 |
| 3 | 3 |
| 4 | 3 |
| 5 | 0 |

The problem says to assume a Poisson likelihood and a uniform prior distribution for the unknown rate lambda. The footnote says to create a uniform prior by using a gamma distribution with a Shape equal to 1 and Scale equal to 10^{4}.

To find the shape for the posterior distribution first find the total number of cars.

sum(cars \* occurences)

[1] 40

Next add the priors’ shape to the total number of cars to get the posterior shape.

gamma\_shape + sum\_xi

[1] 41

Then to find the scale calculate the total number of time periods observed.

sum(occurences)

[1] 21

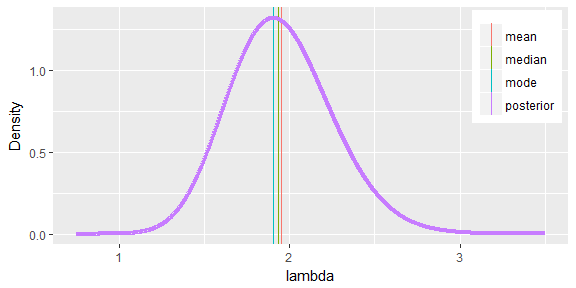
Divide the prior scale by the product of the prior scale and number of time periods plus one. This results in the posterior scale parameter.

gamma\_scale / (1 + N \* gamma\_scale)

[1] 0.04761882

Last use the dgamma() function in R to calculate the posterior distribution over a range for lambda.

Posterior Plot



The next step is to find the mean, standard deviation, median and mode of the posterior distribution as well as a 95% symmetric tail area credible interval for lambda. The table below shows these statistics as well as the values from the previous assignments solution.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Problem | Mean | StDev | Median | Mode | Lower | Upper |
| HW 3 Problem 2 | 1.950000 | 0.3050000 | 2.000000 | 2.000000 | 1.400000 | 2.600000 |
| HW 4 Problem 1 | 1.952372 | 0.3049092 | 1.936522 | 1.904753 | 1.401056 | 2.593733 |

The differences between the results from assignment 3 and this problem are minimal. Both processes for finding the posterior distribution resulted in similar distributions.

# Problem 2

Suppose a highway engineer provided the following prior judgments about the rate of traffic on the stretch of highway (before seeing the data).

|  |  |
| --- | --- |
| Parameters | Cars |
| 10th Percentile | 1.50 |
| 50th Percentile | 3.75 |
| 90th Percentile | 7.00 |

Find a Gamma prior distribution that matches these judgments reasonably well.

### Deriving Parameters

First find the parameters of this Gamma distribution using the engineers judgement. To find a shape value I hold the scale equal to 1 and adjust the shape value. I continue adjusting the shape in order to match the ratio (90th / 50th percentiles) as close as possible to the ratio below using the engineers judgement.

p2\_q90 / p2\_q50

[1] 1.866667

Below is a function for calculating the ratio for shape k.

f <- function(k){qgamma(.9, k, 1)/qgamma(.5, k, 1)}

The code below finds the best value for the shape. It does so by minimizing the difference between the new and original ratios as described above. It inspects all values between 0.1 and 4 for the shape.

p2\_alpha <- uniroot(function(k){f(k)-p2\_q90 / p2\_q50}, c(0.1, 4))$root

After finding the optimal shape value we can calculate the scale using the equation below. This equation divides the engineers judgement of the median by the median calculated using the new shape and a scale of 1.

p2\_beta <- p2\_q50 / qgamma(.5, p2\_alpha, 1)

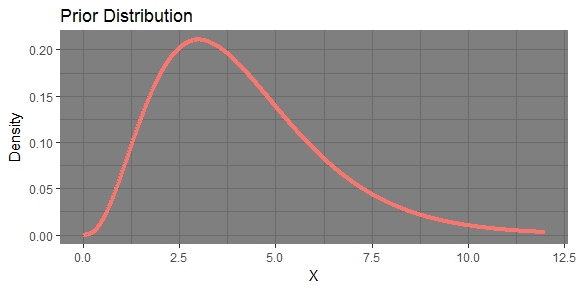
Below is a table showing the results of the analysis done above deriving the scale and shape of the gamm distribution.

|  |  |
| --- | --- |
| Parameters | Cars |
| Shape | 3.668046 |
| Scale | 1.122540 |

Comparing the derived Gamma prior to the engineers judgement below there is only one differing value. The derived distribution matches the engineers judgement very well for the median and 90th percentile. The difference in the 10th percentile however is only 0.2.

|  |  |  |
| --- | --- | --- |
| Parameters | Engineer | Gamma |
| 10th Percentile | 1.50 | 1.71 |
| 50th Percentile | 3.75 | 3.75 |
| 90th Percentile | 7.00 | 7.00 |

Below is a plot of the derived prior gamma distribution.



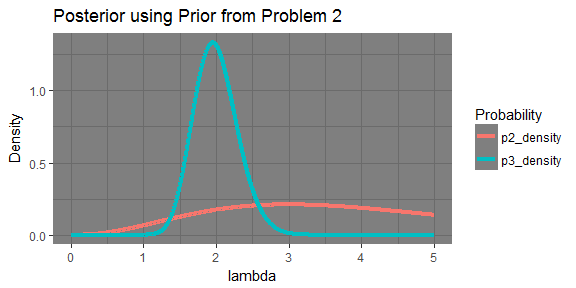
The table below compares this derived prior to the resulting distribution percentiles when the data is know (Problem 1).

|  |  |  |
| --- | --- | --- |
| Percentile | Data | Derived |
| 0.1 | 1.573224 | 1.712415 |
| 0.2 | 1.692233 | 2.289772 |
| 0.3 | 1.781719 | 2.782678 |
| 0.4 | 1.860656 | 3.258157 |
| 0.5 | 1.936522 | 3.750000 |
| 0.6 | 2.014424 | 4.289473 |
| 0.7 | 2.100049 | 4.921244 |
| 0.8 | 2.203279 | 5.734731 |
| 0.9 | 2.351901 | 7.000000 |

The derived distribution from the engineer judgement does not match the results of the distribution based on the data very well. However using this distribution would be better then using the uniform distribution.

# Problem 3

Problem 3 simply repeats the steps from Problem 1 using the prior distribution from Problem 2. Below is the resulting posterior plot.

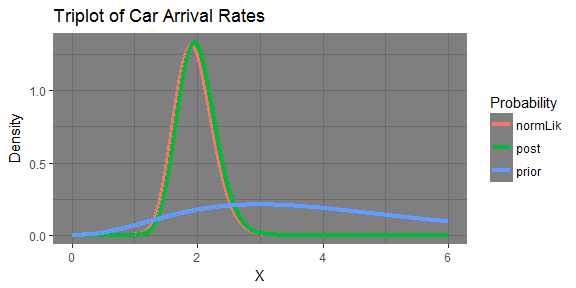


The pink line ‘p2\_density’ is the prior from problem 2. The blue line represents the posterior that was just calculated for problem 3. The table below shows the stats from the three problems involving this data (including last assignments).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Problem | Mean | StDev | Median | Mode | Lower | Upper |
| HW 3 Problem 2 | 1.950000 | 0.3050000 | 2.000000 | 2.000000 | 1.400000 | 2.600000 |
| HW 4 Problem 1 | 1.952372 | 0.3049092 | 1.936522 | 1.904753 | 1.401056 | 2.593733 |
| HW 4 Problem 3 | 1.994809 | 0.3018697 | 1.979603 | 1.949128 | 1.447534 | 2.628475 |

The distributions statistics have changed slightly from problem to problem. None have changed drastically. Overall the distribution seems to be remaining steady in all areas.

# Problem 4



The plot above shows the triplot of the prior, normalized likelihood and posterior from problem 3. The plot shows that the posterior matched the data (normalized likelihood) well. The prior, derived from the engineers judgement, resulted in a posterior that followed the normalized likelihood. Using the engineers judgement helped achieve a posterior that matched the real world well.