Applied Statistic Modeling Assignment 1

Jiaming Deng 22302794

(1) Likelihood is $L(\theta) = \prod \left(\frac{e^{-\theta} \theta^{x_i}}{x_i!} \right)$

Log-likelihood is $\log L(\theta) = -n\theta + \sum (x_i \log(\theta) - \sum(\log(x_i!))$

(2)

In order to prove that $c \sim Gamma(a,b)$ is a conjugate distribution for θ, I need to show that the posterior distribution $p(\theta | y, a, b)$ is also Gamma and to prove that. I need to show the prior $p(\theta | a, b)$ is also a Gamma distribution.

As the probability function of the Poisson distribution is $f(x) = \theta^x \frac{\exp\{-\theta\}}{x!}$. It can also write as $p(x_i|\theta) = \frac{e^{-\theta} \theta^{x_i}}{x_i!}$ where $i = 1, 2, \cdots 20$

Let prior distribution of θ be gamma distribution with parameters a, b. $\theta \sim Gamma(a,b)$. So probability function is $p(\theta|a,b) = \frac{b^a \theta^{a-1} e^{-b\theta}}{\Gamma(a)}$ $\Gamma(a)$ is Gamma function

Using Bayes' theorem

$p(\theta | x_1 \ldots x_{20}, a, b) \propto p(x_1, \ldots x_{20} | \theta) \cdot p(\theta | a, b)$

so apply the formula above we get

$p(\theta | x_1 \cdots x_{20}, a, b) \propto \prod \left(\frac{e^{-\theta} \cdot \theta^{x_i}}{x_i} \right) \cdot \frac{b^a \theta^{a-1} e^{-b\theta}}{\Gamma(a)}$

Then

$\log p(\theta \mid x_1, \cdots x_{20}, a, b) = (a-1)\log(\theta) - b\theta + \Sigma(x_i)\log(\theta)$

$- \Sigma(\log(x_i!) - \log(ba)) + \varepsilon$

ε is a constant

So this is the probability function of a gamma function with parameters $(a + \Sigma(x_i))$ and $(b+20)$.

So $\theta \sim$ Gamma(a, b) is a conjugate prior distribution for θ.

For (1) if θ=7.3

likelihood L(7.3) = $\prod (\frac{7.3x_i}{x_i}) e^{7.3}$

Log-likelihood log L(7.3) = $\sum log(\frac{7.3x_i}{x_i!}) - 7.3$

(3)

The a posteriori parameters a_n and b_n, respectively, are updated versions of the hyperparameters a and b. The posterior parameter a_n is equal to the sum of the observed data's mean and the hyperparameter a's sum, whereas the posterior parameter b_n is equal to the sum of the observed data's mean and the hyperparameter b's sum.
The prior distribution can be updated to the posterior distribution using the mean of the data summary (n and x), which provides details about the observed data. The updated values of the shape and scale parameters of the Gamma distribution are represented by the posterior parameters an and bn, which are obtained by combining information from the data.

(4)

For Colleague A, use the Gamma(2.5, 0.5). The hyperparameters (α = 2.5, β = 0.5) have weak a priori beliefs and low confidence estimates for large values of due to their small a priori shape parameters and rates, respectively. This shows that Colleague A's viewpoint is unconvinced and supports a lower mean flow rate.
The Gamma(80, 10) refers to Colleague B. Large a priori shape parameters and large rate parameters, which correspond to large a priori beliefs, large values, and high confidence estimates, are present in the hyperparameters (α = 80, β = 10). This suggests that Colleague B's opinion is very confident and supports a higher average flow rate.
The Gamma(50, 5) is designated for Colleague C. A moderate shape parameter and a moderate rate parameter are present in the hyperparameters (α = 50, β = 5), which indicate a moderate level of prior beliefs for larger values and a moderate level of confidence estimates for the prior. This suggests that viewpoint C's confidence level is moderate and that viewpoint C supports a moderate mean flow rate.

(5)

1) Discuss which prioor you think would be best to use in this case.
My colleague B's advice seems to me to be the most sensible and reliable, as he is the only one with relevant working experience in the field of network traffic. Therefore, it may be reasonable to use his advice as the basis for prior assignment.
2) Using your preferred prirr, and the data summary provided, compute the posterior parameters a_n and b_n.

Through the analysis in question (3), I learned that of the three gamma priors recommended by the three colleagues, gamma (50, 5) is probably the most appropriate prior to reflect colleague B's opinion. In this case, with α = 50 and β = 5, the mean and variance were calculated as follows.

mean = $\frac{\alpha}{\beta}$ = 10

variance = $\frac{\alpha}{\beta^2}$ = 2

corresponds to a mean of 10 and a variance of 2. These two figures are close to colleague B's suggestion of 8 visits per minute to the site.
Next I calculate the posterior parameters an and bn as follows.

a_n = α + n = 50 + 20 = 70

b_n = β + ∑xi = 5 + 20 * 7.3 = 151

3) Sketch and interpret the resultant posterior distribution.
To compare two gamma distributions, we can look at their shape and scale parameters. The gamma distribution with parameters α = 50 and β = 5 is a relatively On the other hand, the posterior distribution we computed with parameters αn = 70 and βn = 149 has a higher shape parameter and a higher scale parameter, which indicates a broader and more spread out distribution. The mean and variance of the posterior parameters were calculated as follows.

mean = $\frac{\alpha}{\beta}$ = 0.4698

variance = $\frac{\alpha}{\beta^2}$ = 0.00316

Thus, the mean of the posterior distribution is lower than the original gamma distribution, and the variance of the posterior distribution is significantly smaller than the original gamma distribution.
4) Compare this posterior to those produced using the other priors.

For the prior (80, 10), α_n = α + n = 80 + 20 = 100, β_n = β + n × mean = 10 + 20 × 7.3 = 154. mean = $\frac{\alpha}{\beta}$ = 100/154 = 0.6494, variance = $\frac{\alpha}{\beta^2}$ = 100/23716 = 0.00421.

For the prior (2.5, 0.5), α_n = α + n = 2.5 + 20 = 22.5, β_n = β + n × mean = 0.5 + 20 × 7.3 = 147.5. mean = $\frac{\alpha}{\beta}$ = 22.5/147.5 = 0.1525, variance = $\frac{\alpha}{\beta^2}$ = 22.5/21752.25 = 0.00000103

Compared to the posterior with the (2.5, 0.5) prior and the posterior we computed earlier, the posterior with the (80, 10) prior has a higher mean and larger variance. The posterior with the (2.5, 0.5) prior has the lowest mean and the smallest variance of the three. The variance of the posterior distribution with the (80, 10) prior is larger than the variance of the posterior with the (2.5, 0.5) prior and the posterior we computed earlier.
Also, we can compare the shapes of the posterior distributions. The posterior with the (50, 5) prior is narrow and peaked. While the posterior with the (2.5, 0.5) prior is much wider and flatter, the posterior with the (80, 10) prior is also fairly narrow and

peaked.

5) How much does the choice of prior affect the resulting posterior?

The choice of prior can have a significant impact on the resulting posterior distribution. This is because, before we see any data, the prior distribution represents our presumptions or beliefs about the parameter. If the prior distribution assigns high probability to a narrow range of parameter values, then the resulting posterior distribution will also be narrow. On the other hand, if the prior distribution is comparatively flat, the data will have a greater impact on the final posterior distribution.

I observe that the choice of prior can influence the mean, variance, and shape of the resulting posterior distribution in the case of the gamma distribution with the provided parameters. In comparison to the posterior with the (2.5, 0.5) prior and the posterior we computed earlier, the posterior with the (80, 10) prior has a higher mean and a larger variance. The posterior with the (2.5, 0.5) prior has the lowest mean and the smallest variance of the three.