STA 602 HW2

Ryan Tang

September 11th 2022

1 Excerise PH 3.3

(a)

$$
\begin{aligned}
Y_{i \in n} &\overset{iid}{\sim} Poisson(\theta) \\
p(Y|\theta) &= \prod_{i=1}^{n} \frac{1}{y_i!} \theta^{y_i} e^{-\theta} \\
&= c(y_1, y_2, \ldots, y_n) \theta^{\sum y_i} e^{-n\theta}
\end{aligned}
$$

$$
\begin{aligned}
p(\theta) &\sim Gamma(a, b) \\
&= \frac{b^a}{\Gamma(a)} \theta^{a-1} e^{-b\theta}
\end{aligned}
$$

$$
\begin{aligned}
\therefore p(\theta|Y) &\propto p(Y|\theta)p(\theta) \\
&= \theta^{\sum y_i} e^{-n\theta} \theta^{a-1} e^{-b\theta} \\
&= \theta^{a-1+\sum y_i} e^{-\theta(b+n)} \\
&\sim Gamma(a + \sum_{i=1}^{n} y_i, b + n)
\end{aligned}
$$

Therefore, given $Y_A = (12, 9, 12, 14, 13, 13, 15, 8, 15, 6)$, $p(\theta_A|Y_A) \sim Gamma(237, 20)$. And similiarly, given $Y_B = (11, 11, 10, 9, 9, 8, 7, 10, 6, 8, 8, 9, 7)$, $p(\theta_B|Y_B) \sim Gamma(125, 14)$. The θ_A posterior has a mean of 11.85, variances of 0.593, and 95% confidence interval of $[10.39, 13.41]$. The θ_B posterior has a mean of 8.93, variances of 0.638, and 95% confidence interval of $[7.43, 10.56]$.

(b) As long as we keep the $12 \times n0$ term, the posterior is hard to get close to the θ_A posterior, which has a mean of 11.85, because the θ_B posterior is a weighted average between the prior and data Y_B, which has an average of 8.4 tumors instead of 11.7 from Y_A. Hence, we need a prior like $Gamma(14 \times n0, n0)$ or similar to make the θ_B posterior get closer to the θ_A posterior.

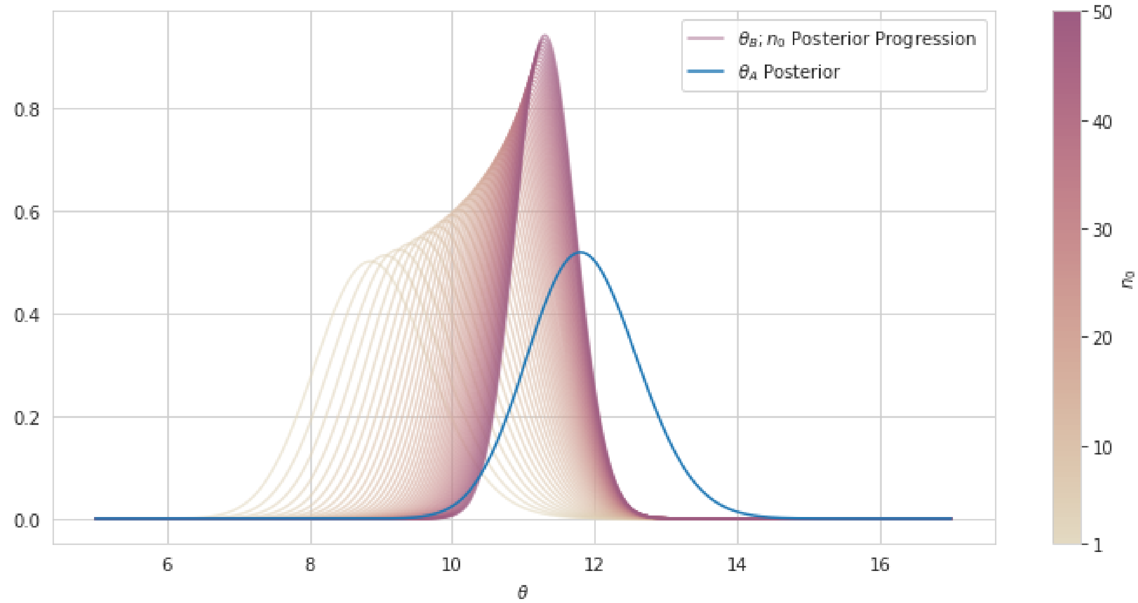

Figure 1: θ_B sensitivity to n_0

(c) The dependency between θ_A and θ_B is worth considering, and it mostly has to do with the relationship between the two strains. It is better to consult an expert on this matter. Then we can decide whether the independence assumption makes sense or not.

2 Excerise PH 3.5

3 Excerise PH 3.8

4 Excerise PH 4.3

(a) Overall the Poisson generative model matches quite well with the data, Y_A.

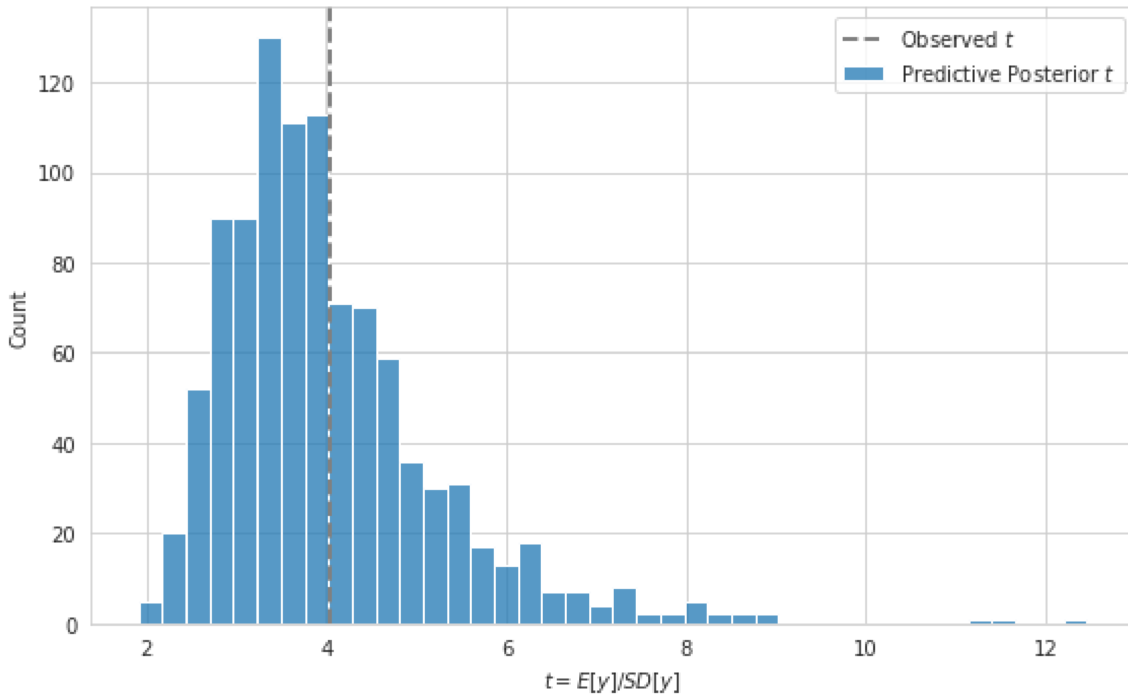

Figure 2: Posterior Predictive Model Check for Group A

(b) However, the model does not work well with the group B data, Y_B. It could be that we have a terrible prior, not enough data size, or perhaps Poisson is just not a good model for the data.

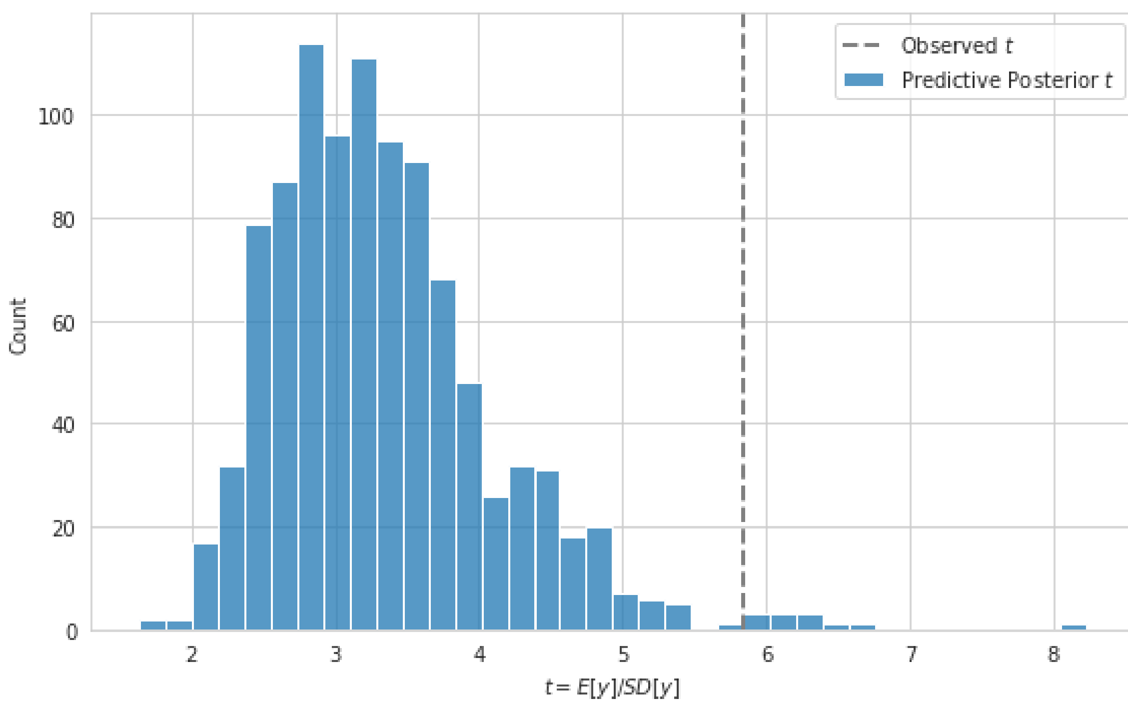

Figure 3: Posterior Predictive Model Check for Group B

5 Excerise 5

(a) Given the prior $\theta \sim Gamma(a = 10, b = 1)$ and the sampling model $X_{i \in n} \overset{iid}{\sim} Poisson(\theta)$, the posterior is also a gamma distribution. Now we know we received a total of 200 call from 10 different days, the Posterior

$$\theta | \mathbf{X} \sim Gamma(a = 210, b = 11)$$

The square error loss estimate of θ is the mean, 19.09. And the absoute error loss estimate of θ is the median, 19.06.

(b) The model is reasonable because observing 20 calls on average is a typical belief shown in the posterior. However, given only one data point, it is hard to tell the truth. The underlying each data point might will be $\{0, 0, \ldots, 0, 200\}$. In such a case, our model performs poorly instead.

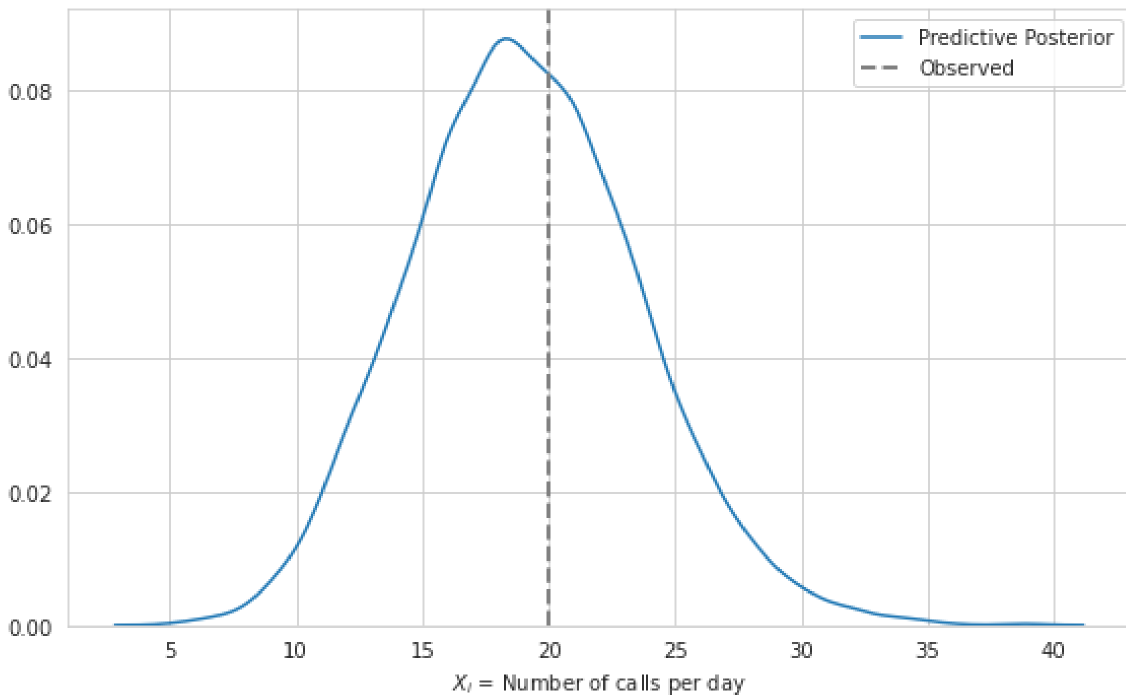

Figure 4: Posterior Predictive Model Check