Zeqiu. Yu

1. $Y_i \sim \text{Poisson}(\lambda)$, $f(y_i) = \frac{\lambda^{y_i} e^{-\lambda}}{y_i!}$, $l(Y_i;\theta) = -\log(y_i!) + y_i \log\lambda - \lambda$

$\frac{\partial^2 l_1}{\partial x^2} = -\frac{y_i}{\lambda^2}$

$I_{Y_i}(\theta) = \frac{1}{\lambda}$

For $Y = (y_1, y_2, \dots, y_n)$, $I_Y(\theta) = \frac{n}{\lambda}$

Then the Jeffrey's noninformative prior is $P_{prior}(\lambda) \propto \sqrt{\frac{1}{n}\cdot\lambda} \propto \lambda^{-\frac{1}{2}}$

It corresponds to $Gamma(\frac{1}{2}, 0)$ for $f(x) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$.

2. $f(y_i;\mu,\sigma) = (2\pi\sigma^2)^{-\frac{1}{2}} e^{-\frac{(y_i-\mu)^2}{2\sigma^2}}$

$l = \log f = -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma^2 - \frac{(y_i-\mu)^2}{2\sigma^2}$

$\frac{\partial l}{\partial \mu} = \frac{1}{2\sigma^2}\cdot 2(y_i-\mu) = \frac{y_i-\mu}{\sigma^2}$

$\frac{\partial l}{\partial(\sigma^2)} = -\frac{1}{2} + \frac{1}{2}\frac{(y_i-y)^2}{\sigma^4}$

$E\left[\frac{\partial^2 l}{\partial\mu^2}\right] = -\frac{1}{\sigma^2}$ $\quad E\left[\frac{\partial^2 l}{\partial\mu\,\partial(\sigma^2)}\right] = 0$

$E\left[\frac{\partial^2 l}{\partial\sigma^2\,\partial\mu}\right] = 0$ $\quad E\left[\frac{\partial^2 l}{\partial(\sigma^2)^2}\right] = \frac{1}{2\sigma^4}$

Then we get the $J(\theta)$ with

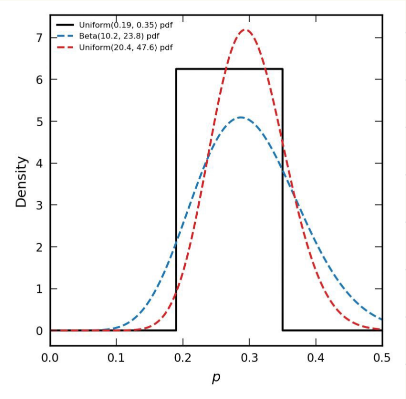
$$I(\theta) = -E[J(\theta)] = -\begin{bmatrix} -\frac{1}{\sigma^2} & 0 \\ 0 & \frac{1}{2\sigma^4} \end{bmatrix}$$

$|I(\theta)| = \frac{1}{\sigma^4}$

Then the invariant prior $Prior(\theta) \propto \sqrt{\frac{1}{\sigma^6}} = \frac{1}{\sigma^3}$, which is smaller than $p(\theta, \sigma^2) \propto \frac{1}{\sigma^2}$.

3. a. ① For the uniform distribution.

It will kill the likelihood outside the range of [0.19, 0.35]. Also this prior assumes that any value in [0.19, 0.35] is equally. If we have no prior knowledge or we can use it.



As for two Beta distribution. Both have some prior knowledge that the batting average is around 0.3 with some uncertainty.

② Beta(10.2, 23.8) has heavier tails.

③ Beta(20.4, 47.6) has much thinner tails, which means we are more confident batting average should be within [0.19, 0.35] than Beta(10.2, 23.8).

b. $P(5 \text{ hits}; p) = \binom{40}{x} \cdot p^x \cdot (1-p)^{35-x}$

For uniform(0.19, 0.35): $L(p) = \binom{40}{5} \cdot p^5 \cdot (1-p)^{35}$

$\text{posterior}(p) \propto L(p) \cdot \frac{1}{0.35-0.19} \sim Beta(6, 36)$

The mean of the posterior: $\frac{\alpha}{\alpha+\beta} = \frac{6}{6+36} = \frac{1}{7}$

The standard deviation: $\sqrt{\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta-1)}} = \sqrt{\frac{6 \cdot 36}{42^2 \cdot 41}} = 0.053$

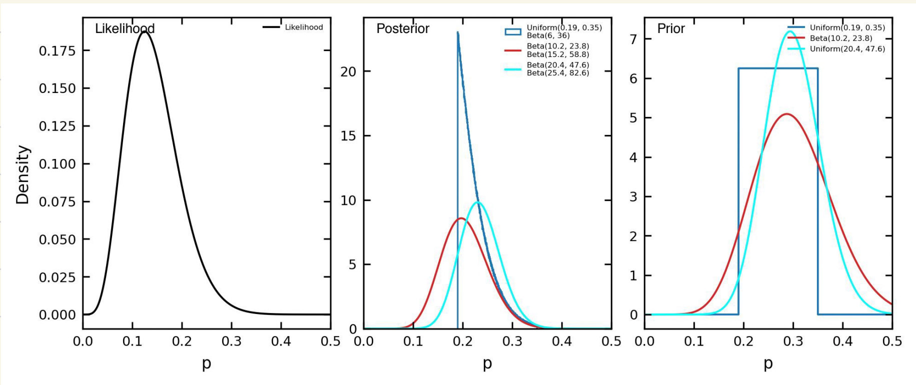
$P_r(p > 0.2) = 0.781$ HPD interval: [0.191, 0.311]

For Beta(10.2, 23.8): $L(p) = \binom{40}{5} \cdot p^5 \cdot (1-p)^{35}$

$\text{posterior}(p) \propto p^5 \cdot p^{10.2-1} \cdot (1-p)^{35} (1-p)^{23.8-1} \propto = Beta(15.2, 58.8)$

Mean: 0.205 Std: 0.047

$P_r(p > 0.2) = 0.523$ HPD interval: [0.123, 0.305]



For Beta(20.4, 47.6): $L(p) = \binom{40}{5} p^5 \cdot (1-p)^{35}$

posterior $(p) \propto p^5 \cdot p^{20.4-1} \cdot (1-p)^{35} (1-p)^{47.6-1} \propto Beta(25.4,\ 82.6)$

Mean: 0.235 Std: 0.041

$Pr(p > 0.2) = 0.803$ HPD interval: [0.161, 0.320]

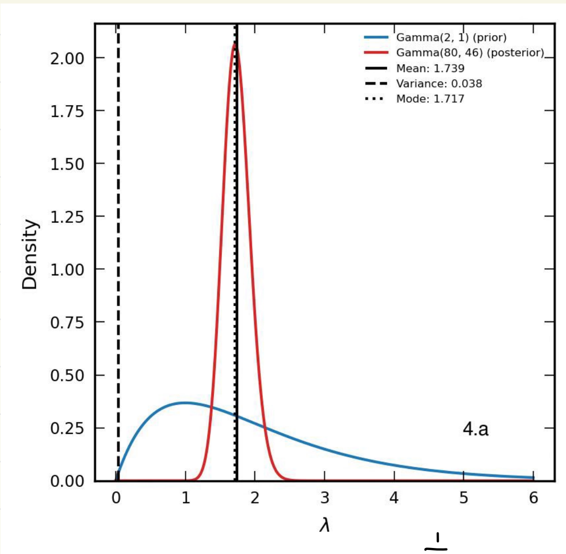
c. The proportions of success: $\frac{5}{40} = \frac{1}{8} = 0.125$

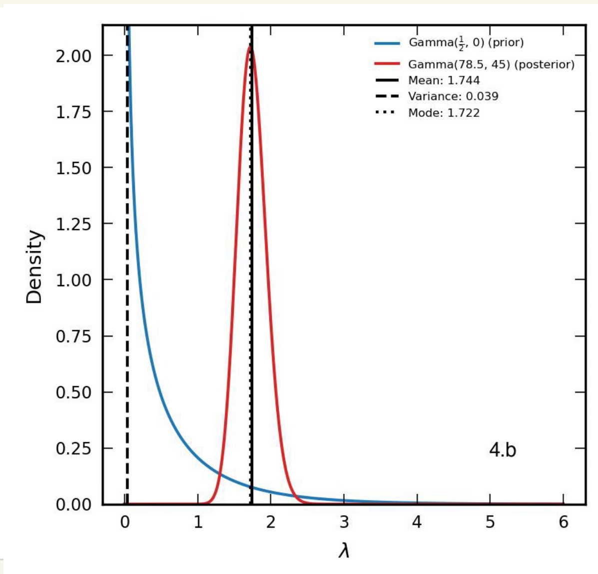
Beta prior mean: $Beta(10.2, 29.8)$: $\frac{10.2}{40} = 0.255$ $(0.125 < 0.205 < 0.255)$

$Beta(20.4, 47.6)$: $\frac{20.4}{68} = 0.3$ $(0.125 < 0.235 < 0.3$

4.a. $p(x_i; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!}$, then the likelihood $L(\lambda) = \prod_{i=1}^{45} P(x_i; \lambda) = \frac{\lambda^{\sum x_i} e^{-45\lambda}}{\prod_{i=1}^{n} x_i!}$

$p(\lambda | x) \propto L(\lambda) p(\lambda) \propto (\lambda e^{-\lambda}) \cdot \lambda^{\sum x_i} e^{-45\lambda} = \lambda^{\sum x_i + 1} e^{-46\lambda} = \lambda^{79} e^{-46\lambda} \sim Gamma(80, 46)$





b. When $p(\lambda) = Gamma(\frac{1}{2}, 0) = \lambda^{-\frac{1}{2}}$

$p(\lambda | x) \propto \lambda^{\sum x_i - \frac{1}{2}} e^{-45\lambda} = \lambda^{77.5} e^{-45\lambda} \sim Gamma(78.5, 45)$

5.

a. It's a multinomial distribution.

Then the likelihood : $L(\theta) = \binom{197}{y_1, y_2, y_3, y_4} P_1^{y_1} P_2^{y_2} P_3^{y_3} P_4^{y_4}$

$$= \frac{197!}{125!\,18!\,20!\,34!} \left(\frac{2+\theta}{4}\right)^{125} \cdot \left(\frac{1-\theta}{4}\right)^{18} \cdot \left(\frac{1-\theta}{4}\right)^{20} \left(\frac{\theta}{4}\right)^{34}$$

$$= \frac{197!}{125!\,18!\,20!\,34!} \cdot \left(\frac{1}{4}\right)^{197} \cdot (1-\theta)^{38} (2+\theta)^{125} \theta^{34}$$

b. In the same way,

$$L(\theta) = \frac{20!}{14!\,5!} \left(\frac{1}{4}\right)^{20} \cdot (1-\theta)(2+\theta)^{14} \theta^{5}$$

c. $l(\theta) = \log(A) + 38\log(1-\theta) + 125\log(2+\theta) + 34\log\theta$ with $A = \frac{197!}{125!\,18!\,20!\,34!}$

$$l'(\theta) = \frac{-38}{1-\theta} + \frac{125}{2+\theta} + \frac{34}{\theta} \qquad l''(\theta) = -\frac{125}{(2+\theta)^2} - \frac{38}{(1-\theta)^2} - \frac{34}{\theta^2}$$

$$\theta_{i+1} = \theta_i + \left(-\frac{\partial^2 l}{\partial\theta^2}\bigg|_{\theta=\theta^i}\right)^{-1} \left[\frac{\partial l}{\partial\theta}\bigg|_{\theta=\theta^i}\right]$$

$MLE(\hat{\theta}) = 0.63$

Calculate $\|\theta_{i+1} - \theta_i\|_2$ with tolerance $= 0.01$ to asses convergence.

d. For $(14, 0, 1, 5)$, repeat the process,

we get $MLE(\hat{\theta}) = 0.90$

However, during the process, for some particular initialization 0.3 and 0.6, it doesn't converge.