# Biostats-202C-hw1

Jiahao Tian

2022-10-06

```
library(tidyr)
library(ggplot2)
library()
```

**Question 1**

**a)**

$$y \sim pois(\lambda), \quad \lambda \sim gamma(a, b)$$

$$f(y|\lambda) = \frac{(\lambda^y e^{-\lambda})}{y!} \quad L(y|\lambda) = \frac{\lambda^{\sum(yi)} e^{-ny}}{\Pi y!}$$

$$p(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-\lambda b}$$

$$p(\lambda|y) \propto \lambda^{\sum(yi)+a-1} e^{-\lambda(b+n)}, \lambda > 0$$

$$posterior\ follows\ gamma\ distribution\ \sim Gamma(\sum(yi) + a, b + n)$$

**b)**

$$\begin{aligned} f(y) &= \int_0^\infty \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-\lambda b} \frac{\lambda^y e^{-\lambda}}{y!} d\lambda \\ &= \frac{b^a}{\Gamma(a) y!} \int_0^\infty \lambda^{a-1} e^{-\lambda b} \lambda^y e^{-\lambda} d\lambda \\ &= \frac{b^a}{\Gamma(a) y!} \frac{\Gamma(y+a)}{(b+a)^{y+a}} \int_0^\infty \frac{(b+a)^{y+a}}{\Gamma(y+a)} \lambda^{y+a-1} e^{-\lambda(b+1)} d\lambda \\ &= \frac{b^a}{\Gamma(a) y!} \frac{\Gamma(y+a)}{(b+1)^{y+a}} \end{aligned}$$

**c)**

From part b,

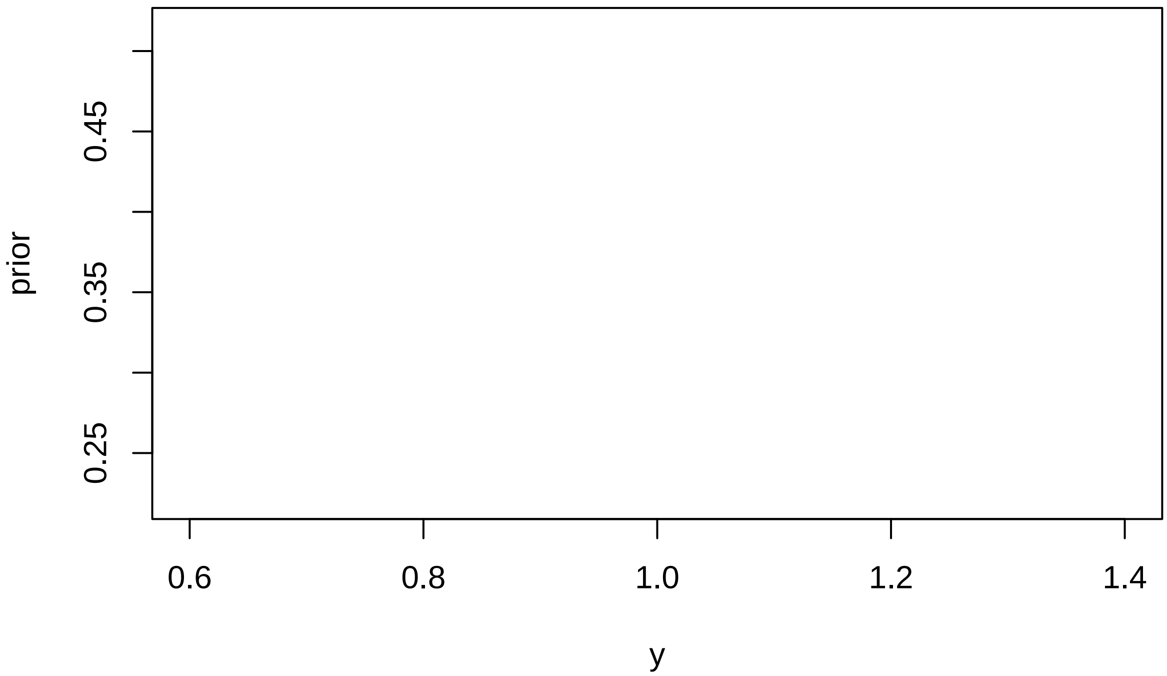
$$\begin{aligned} f(y|a, b) &= \frac{b^a}{\Gamma(a) y!} \frac{\Gamma(y+a)}{(b+1)^{y+a}} \\ &= \frac{(y+a-1)!}{(a-1)! y!} \frac{b^a}{(b+1)^{y+a}} \\ &= \binom{y+a-1}{y} b^a (b+1)^{-y-a} \end{aligned}$$

$$Which \quad f(y|a,b) \sim NB(a,b)$$

**d)**

```r
y <- 1
data <- dpois(y, 1)
prior <- dgamma(y, 1, 1)
posterior <- dgamma(y, 2, 2)

plot(y, prior, type = "l", col = "red")
lines(y, posterior, type = "l", col = "green")
lines(y, data , type = "l", col = "blue")
```




**Question 2**

**a)**

$$L(y|\lambda) = \frac{\lambda^{\sum(yi)} e^{-ny}}{\Pi y!} \quad p(\lambda) = \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-\lambda b}$$

$$p(\lambda|y) \propto \lambda^{\sum(yi)+a-1} e^{-\lambda(b+n)}, \lambda > 0$$

$$posterior \ follows \ gamma \ distribution \ \sim Gamma(\sum(yi) + a, b + n)$$

**Question 3**

**a)**

$$
\begin{aligned}
mean; \quad E(y|\tau) &= \int_0^\infty y(\frac{2}{\pi\tau})^{1/2} e^{\frac{-(y)^2}{2\tau}} dy \\
&= (\frac{2}{\pi\tau})^{1/2}[-\tau e^{\frac{-y^2}{2\tau}}|_0^\infty] \\
&= (\frac{2\tau}{\pi})^{1/2}
\end{aligned}
$$

$$
variance; \quad \tau(1-\frac{2}{\pi}) \qquad mode; \quad 0 \qquad median; \quad \tau\sqrt{2}erf^{-1}(1/2)
$$

**b)**

$$
indicator, \qquad I = \{^1_0 \quad {}^{0\,<\,yi}_{otherwise}
$$

$$
\begin{aligned}
Lf(y|\tau) &= (\frac{2}{\pi\tau})^{n/2} e^{-\frac{(\sum yi)^2}{2\tau}} I \\
logLf(y|\tau) &= \frac{n}{2}log\frac{2}{\pi\tau} - \frac{\sum yi^2}{2\tau} I \\
\nabla logLf(y|\tau) &= -\frac{n}{2\tau^3} + \frac{\sum yi^2}{2\tau^2} \;, set\ equal\ to\ 0 \\
we\ get\ , \hat{\tau} &= \frac{n}{\sum yi^2} \\
T &= \frac{n}{\sum yi^2} \;, is\ the\ sufficient\ statistics\ for\ \tau
\end{aligned}
$$

**c)**

$$
indicator, \qquad I = \{^1_0 \quad {}^{0\,<\,yi}_{otherwise}
$$

$$
Lf(y|\tau) = (\frac{2}{\pi\tau})^{n/2} e^{-\frac{(\sum yi)^2}{2\tau}} I
$$

$$
f(\tau) = \frac{(\frac{b}{2})^{\frac{a}{2}}}{\Gamma(\frac{a}{2})} \tau^{-\frac{a}{2}-1} e^{-\frac{b}{2\tau}}
$$

$$
posterior; \quad f(\tau|y) \propto (\frac{2}{\pi\tau^{a+2}})^{\frac{n}{2}} e^{\frac{-(\sum yi^2+b)}{2\tau}}, \; yi > 0
$$

$$
posterior\ follows\ halfnormal\ distribution
$$

**Question 4**

**a)**

When data is non-negative, also, collected at the individual level (binary response) or group level (proportion) might use half-normal distribution. When data shows binary response and non-negative values will lead to use half-normal distribution.

**b)**

Example: Half-normal can be use to model distance at which wildlife observe on a boat detect an animal when surveying, such as whales in the ocean.

**c)**

The competitor distribution I might use is “Poisson distribution”.

**Question 5**

**a)**

$$f(y|\sigma) = (\frac{2}{\pi\sigma})^{1/2} e^{-\frac{(y)^2}{2\sigma}},\ y > 0$$

**b)**

Yes, because half-normal distribution's variance here is unknown and inverse gamma can use as the conjugate prior of the variance.

**c)**

Yes, I thinks that it can be inverse gamma.

**Question 6**

**a)**

$$f(y|\theta) \sim beta(\theta, 1)$$
$$so\ the\ normalizing\ constant\ is;\ \frac{\Gamma(\theta+1)}{\Gamma(\theta)\Gamma(1)}$$

**b)**

$$Lf(y|\theta) = \left(\frac{\Gamma(\theta+1)}{\Gamma(\theta)\Gamma(1)}\right)^n \Pi_{i=1}^n yi^{\theta-1} I_{0<yi<1}$$

$$By\ Factorization$$
$$h(y) = \Pi_{i=1}^n yi^{-1}$$
$$k(g(y)|\theta) = \Pi_{i=1}^n yi^{\theta} I_{0<yi<1}$$
$$T = max\{yi\}, is\ the\ sufficient\ statistics\ for\ \theta$$