202C HW1

Zixi Zhang

October 14, 2022

1. **Poisson-Gamma.** Datum y is distributed Poisson λ, where λ is the mean of the Poisson distribution. A priori, $\lambda \sim \text{Gamma}(a, b)$ where $a > 0$ and $b > 0$ are known scalars.
(a) What is the posterior distribution of λ given y ?
Solution:
$$\begin{aligned}
f(\lambda \mid y) &= \frac{f(y \mid \lambda) f(\lambda)}{f(y)} \\
&\propto f(y \mid \lambda) f(\lambda) \\
&\propto \frac{e^{-\lambda}\lambda^y}{y!} \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda} \\
&\propto \frac{1}{y!} \frac{b^a}{\Gamma(a)} \lambda^{(a+y)-1} e^{-\lambda(b+1)} \\
&\propto \lambda^{(a+y)-1} e^{-\lambda(b+1)} \\
\lambda \mid \mathbf{y} &\sim \text{Gamma}(a+y, b+1)
\end{aligned}$$

(b) Calculate the normalizing constant $f(y) = \int_0^\infty f(y \mid \lambda) f(\lambda) d\lambda$.
Solution:
$$\begin{aligned}
f(y) &= \int_0^\infty f(y \mid \lambda) f(\lambda) d\lambda \\
&= \int_0^\infty \frac{e^{-\lambda}\lambda^y}{y!} \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda} d\lambda \\
&= \int_0^\infty \frac{b^a}{y!\Gamma(a)} \lambda^{(a+y)-1} e^{-\lambda(b+1)} d\lambda \\
&= \frac{b^a}{y!\Gamma(a)} \int_0^\infty \lambda^{(a+y)-1} e^{-\lambda(b+1)} \frac{(b+1)^{a+y}}{\Gamma(a+y)} \frac{\Gamma(a+y)}{(b+1)^{a+y}} d\lambda \\
&= \frac{b^a \Gamma(a+y)}{y!\Gamma(a)(b+1)^{a+y}} \int_0^\infty \lambda^{(a+y)-1} e^{-\lambda(b+1)} \frac{(b+1)^{a+y}}{\Gamma(a+y)} d\lambda \\
&\overset{\star}{=} \frac{b^a \Gamma(a+y)}{y!\Gamma(a)(b+1)^{a+y}} \cdot 1 \\
&= \frac{b^a \Gamma(a+y)}{y!\Gamma(a)(b+1)^{a+y}}
\end{aligned}$$
Step $\star$ holds because
$\int_0^\infty \lambda^{(a+y)-1} e^{-\lambda(b+1)} \frac{(b+1)^{a+y}}{\Gamma(a+y)} d\lambda = \int f(\lambda; a+y, b+1) d\lambda = 1$ where $\Lambda \sim \text{Gamma}(a+y, b+1)$

(c) Thought of as a distribution in y given a and b, what distribution with what parameters is $f(y \mid a, b)$? (If you need a hint: Wikipedia Gamma-Poisson mixture).

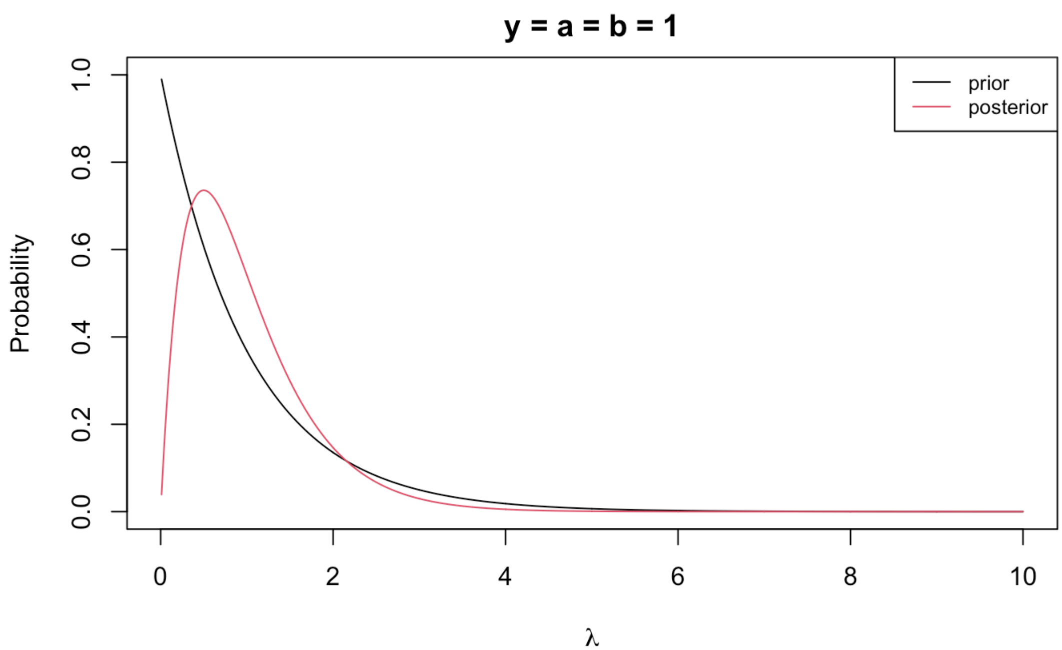
Solution: It is a negative binomial(NB) distribution

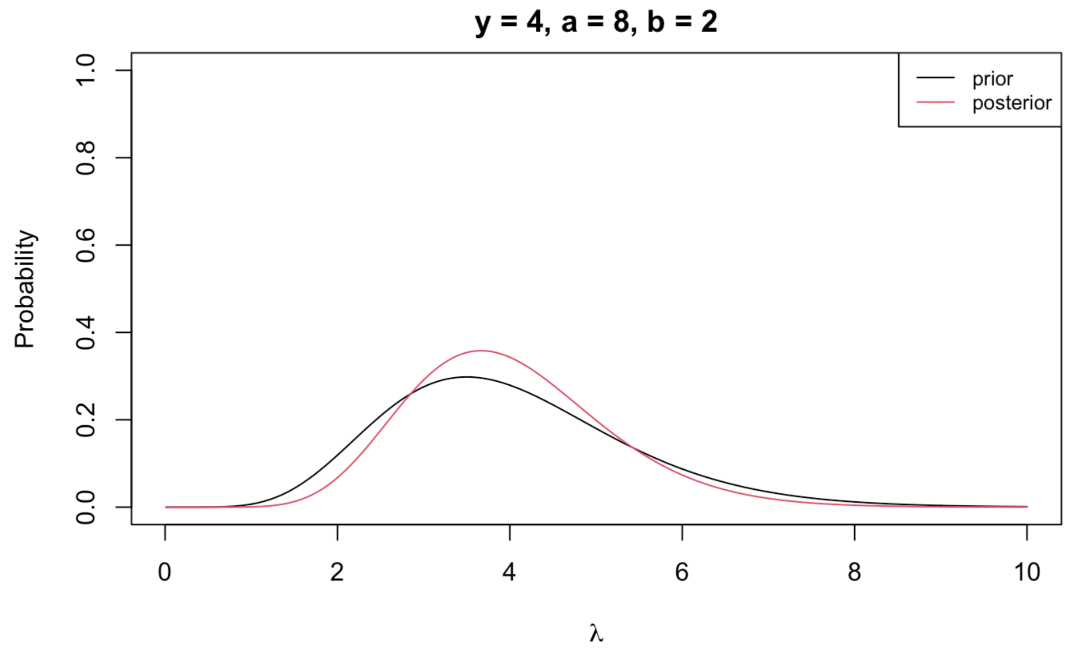
$$
\begin{aligned}
f(y \mid a, b) &= \int_0^\infty f_{\text{Poisson }(\lambda)}(y) * f_{\text{Gamma }(a,b)}(\lambda) d\lambda \\
&= \int_0^\infty \frac{\lambda^y}{y!} e^{-\lambda} \frac{\lambda^{a-1} e^{-b \cdot \lambda} b^a}{\Gamma(a)} d\lambda \\
&= \frac{b^a}{y!\Gamma(a)} \int_0^\infty \lambda^{a+y-1} \cdot e^{-(b+1)\lambda} d\lambda \\
&\stackrel{\star}{=} \frac{b^a}{y!\Gamma(a)} \frac{\Gamma(a+y)}{(b+1)^{a+y}} \\
&= \frac{\Gamma(a+y)}{y!\Gamma(a)} \left(\frac{1}{b+1}\right)^y \left(\frac{b}{b+1}\right)^a \\
&= \frac{\Gamma(a+y)}{y!\Gamma(a)} \left(\frac{1}{b+1}\right)^y \left(1 - \frac{1}{b+1}\right)^a \\
&= \frac{(a+y-1)!}{(a-1)!y!} \left(\frac{1}{b+1}\right)^y \left(1 - \frac{1}{b+1}\right)^a \\
&= \begin{pmatrix} a+y-1 \\ y \end{pmatrix} \left(\frac{1}{b+1}\right)^y \left(1 - \frac{1}{b+1}\right)^a \\
&= NB(a, \frac{1}{b+1})
\end{aligned}
$$

Step $\star$ holds because of the following equality,

$$
\int_0^\infty x^b e^{-ax} \text{ d}x = \frac{\Gamma(b+1)}{a^{b+1}}
$$

(d) Plot the prior and the posterior for λ on the same plot for $y = a = b = 1$. Repeat for $y = 4, a = 8, b = 2$. Label appropriately.





2. **Poisson-Gamma, contd.** Data $Y = (y_1, \ldots, y_n)'$ is an n-vector of observations that are independently distributed as Poisson given λ, $y_i \mid \lambda \sim \text{Poisson}(\lambda)$. A priori, $\lambda \sim \text{Gamma}(a, b)$ where $a > 0$ and $b > 0$ are known scalars.
(a) What is the posterior distribution of λ given Y ?
Solution:

$$\begin{aligned}
f(\lambda \mid y) &= \frac{f(y \mid \lambda) f(\lambda)}{f(y)} \\
&\propto f(y \mid \lambda) f(\lambda) \\
&\propto \frac{e^{-n\lambda} \lambda^{\Sigma y_i}}{\Pi y_i} \frac{b^a}{\Gamma(a)} \lambda^{a-1} e^{-b\lambda} \\
&\propto \lambda^{(a+\Sigma y_i)-1} e^{-\lambda(b+n)} \\
&\propto \lambda^{(a+\Sigma y_i)-1} e^{-\lambda(b+n)}
\end{aligned}$$

$$\lambda \mid \mathbf{Y} \sim \text{Gamma}(a + \Sigma y_i, b + n), i = 1, 2, ..., n$$

3. **Half-normal distribution.** The half-normal distribution given parameter τ has density

$$f(y \mid \tau) = \left(\frac{2}{\pi\tau}\right)^{1/2} \exp\left(-\frac{y^2}{2\tau}\right) 1\{0 < y\}$$

and let us denote this distribution as

$$y \mid \tau \sim \text{HN}(y \mid \tau).$$

Let $y_i \mid \tau \sim \text{HN}(y_i \mid \tau)$ for $i = 1, \ldots, n$ with $Y = (y_1, \ldots, y_n)'$
(a) What are the mean, median, mode and variance of the half-normal distribution?
Solution:

Mean:

$$
\begin{aligned}
E(y \mid \tau) &= \int_0^{\infty} y \cdot \left(\frac{2}{\pi\tau}\right)^{1/2} \exp\left(\frac{-y^2}{2\tau}\right) y dy \\
&= \left(\frac{2}{\pi\tau}\right)^{1/2} \int_0^{\infty} y \cdot \exp\left(\frac{-y^2}{2\tau}\right) dy \\
&= \left(\frac{2}{\pi\tau}\right)^{1/2} [-\tau e^{-\frac{y^2}{2\tau}}]_0^{\infty} \\
&= \left(\frac{2}{\pi\tau}\right)^{1/2} \cdot [0 - (-\tau)] \\
&= \left(\frac{2}{\pi\tau}\right)^{1/2} \cdot \tau \\
&= \left(\frac{2\tau}{\pi}\right)^{1/2}
\end{aligned}
$$

Median: Solve

$$
\begin{aligned}
0.5 &= \int_0^{y_{0.5}} \left(\frac{2}{\pi\tau}\right)^{1/2} \exp(\frac{-y^2}{2\tau}) dy \\
&= \left(\frac{2}{\pi\tau}\right)^{1/2} \int_0^{y_{0.5}} \exp\left(\frac{-y^2}{2\tau}\right) dy \\
&= \frac{2}{\pi^{1/2}} \int_0^{y_{0.5}/\sqrt{2\tau}} \exp\left(-t^2\right) dt \\
&= \operatorname{erf}\left(\frac{y_{0.5}}{\sqrt{2\tau}}\right)
\end{aligned}
$$

So we can get,

$$
\begin{aligned}
0.5 &= \operatorname{erf}(\frac{y_{0.5}}{\sqrt{2\tau}}) \\
\Rightarrow \operatorname{erf}^{-1}(0.5) &= \frac{y_{0.5}}{\sqrt{2\tau}} \\
\Rightarrow y_{0.5} &= \sqrt{2\tau}\, \operatorname{erf}^{-1}(0.5)
\end{aligned}
$$

Mode: Mode of the distribution is the maximum of its density, so by sloving $\frac{df}{dy} = 0$

$$
\begin{aligned}
\frac{df}{dy} &= \frac{d}{dy}\left[\left(\frac{2}{\pi\tau}\right)^{1/2} \exp\left(\frac{-y^2}{2\tau}\right)\right] \\
&= \left(\frac{2}{\pi\tau}\right)^{1/2} \frac{d}{dy}\left[\exp\left(\frac{-y^2}{2\tau}\right)\right] \\
&= \left(\frac{2}{\pi\tau}\right)^{1/2} \left(\frac{-y}{\tau}\right) \exp\left(\frac{-y^2}{2\tau}\right) = 0
\end{aligned}
$$

$$
\begin{aligned}
\Rightarrow \frac{-y}{\tau} &= 0 \\
y &= 0
\end{aligned}
$$

Variance: Let $y = |x|$, s.t. $x \sim \mathrm{N}(0, \tau)$. If $y \mid \tau \sim \mathrm{HN}(\tau)$ and $x \sim \mathrm{N}(0, \tau)$ then:

$$
\begin{aligned}
E\left(y^2\right) &= E\left(|x|^2\right) \\
&= E\left(x^2\right) \\
&= \mathrm{Var}(x) + [E(x)]^2 \\
&= \tau + 0 \\
&= \tau \\
\mathrm{Var}(y) &= E\left(y^2\right) - [E(y)]^2 \\
&= \tau - \left(\left(\frac{2\tau}{\pi}\right)^{\frac{1}{2}}\right)^2 \\
&= \tau - \frac{2\tau}{\pi} \\
&= \tau\left(1 - \frac{2}{\pi}\right)
\end{aligned}
$$

(b) What is the sufficient statistic?
Solution:

$$
\begin{aligned}
L(\tau \mid y_i) &= \prod_{i=1}^{n} \left(\frac{2}{\pi\tau}\right)^{1/2} \exp\left(\frac{-y_i^2}{2\tau}\right), i = 1, \ldots, n \\
&= \underbrace{\left(\frac{2}{\pi\tau}\right)^{n/2} \exp\left(\frac{-\Sigma_{i=1}^{n} y_i^2}{2\tau}\right)}_{g_\tau(T(y))} \underbrace{\mathbf{1}_{\min\{y_i\}>0}}_{h(y)}
\end{aligned}
$$

The sufficient statistic is $T = \Sigma_{i=1}^{n} y_i^2$

(c) The inverse gamma distribution $\tau \sim \mathrm{InverseGamma}(a/2, b/2)$ is a convenient and conjugate prior for τ. Derive the posterior of $\tau \mid Y$.
Solution:

$$
\begin{aligned}
f(\tau \mid Y) &= \frac{f(Y \mid \tau) f(\tau)}{f(y)} \\
&\propto f(Y \mid \tau) f(\tau) \\
&\propto \left(\frac{2}{\pi\tau}\right)^{n/2} \exp\left(\frac{-\sum_{i=1}^{n} y_i^2}{2\tau}\right) \frac{b/2}{\Gamma(a/2)} \tau^{-(a/2)-1} \exp\left(\frac{-b/2}{\tau}\right) \\
&\propto \left(\frac{2}{\pi}\right)^{n/2} \frac{b/2}{\Gamma(a/2)} \left(\frac{1}{\tau}\right)^{n/2} \exp\left(\frac{-\sum_{i=1}^{n} y_i^2 - b}{2\tau}\right) \tau^{-(a/2)-1} \\
&\propto \exp\left(\frac{-\left(\sum_{i=1}^{n} y_i^2 + b\right)/2}{\tau}\right) \tau^{-a/2-n/2-1} \\
&\propto \exp\left(\frac{-\left(\sum_{i=1}^{n} y_i^2 + b\right)/2}{\tau}\right) \tau^{-\left(\frac{a+n}{2}\right)-1} \\
\tau \mid \mathbf{y} &\sim \mathrm{InverseGamma}\left(\frac{a+n}{2}, \frac{\sum y_i^2 + b}{2}\right)
\end{aligned}
$$

4. **Half-normal distribution, cont'd.**
(a) What kind of data might you model using the half-normal distribution? That is, what characteristics/features of the data would lead you to use a half-normal distribution as a sampling model? (Answer with a short list of a few items.)
This requires you think about what characteristics of a data set are important and that should be reflected in choosing a sampling density. For example, would minimum daily temperatures in Fahrenheit in Fairbanks, Alaska be appropriately modeled by a half-normal density? (yes/no)

Solution: The data can use the half-normal distribution should be continuous and nonnegative. The histogram or density plot should be half of a bell shape.
No, the minimum daily temperatures in Fahrenheit in Fairbanks, Alaska can not be appropriately modeled by a half-normal density. Because the minimum daily temperatures in Fahrenheit can be negative.

(b) Give a specific example of data that might be modeled by the half-normal density. (Like the Fairbanks temp data of previous question.)

Solution: If we change the Fairbanks temp data into Kalvin temp, which means there would be no negative values, we can use half-normal.

(c) What competitor distributions might you use in place of the halfnormal distribution to model non-negative data?

Solution: log-normal, Gamma, etc.

5. **Half-normal distribution, cont'd.** Suppose you parameterize the halfnormal distribution in terms of the unknown standard deviation, $\sigma = \tau^{1/2}$.
(a) Write the sampling density of $y_i \mid \sigma$.
Solution:

$$\begin{aligned} f\left(y_i \mid \sigma\right) &= \prod_{i=1}^{n}\left[\left(\frac{2}{\pi\sigma^2}\right)^{1/2}\exp\left(\frac{-y_i^2}{2\sigma^2}\right)\right] \\ &= \left(\frac{2}{\pi\sigma^2}\right)^{n/2}\exp\left(\frac{-\sum_{i=1}^{n} y_i^2}{2\sigma^2}\right) \end{aligned}$$

(b) Is an inverse gamma prior for σ conjugate when $y_i \mid \sigma \sim \text{HN}\left(y_i \mid \sigma^2\right)$? [Yes/no.] Please explain [1 sentence].

Solution: No. The likelihood function is not an inverse gamma kernel.

6. **Power Distribution.** We define the power distribution for $\theta > 0$ and for $0 < y < 1$ to have density

$$f(y \mid \theta) \propto y^{\theta-1}\mathbf{1}\{0 < y < 1\}.$$

(a) Find the normalizing constant for $f(y \mid \theta)$ so that the right hand side is a density.

Solution: let the constant to be c, and we will get

$$\begin{aligned} \int_0^1 c y^{\theta-1} dy &= 1 \\ \left[\frac{c y^\theta}{\theta}\right]_0^1 &= 1 \\ \frac{c}{\theta} - 0 &= 1 \\ c &= \theta \end{aligned}$$

So the constant is θ

(b) Give the likelihood for a sample y_i of size $n, i = 1, \ldots, n$, with $Y = (y_1, \ldots, y_n)'$. Find the sufficient statistic. Is there a particularly convenient (i.e. interpretable) form of the sufficient statistic? What is the interpretable form?

Solution:

$$
\begin{aligned}
L\left(y \mid \theta\right) &= \prod_{i=1}^{n} \theta y_i^{\theta-1} i = 1, \ldots, n \\
&= \underbrace{\theta^n \prod_{i=1}^{n} y_i^{\theta-1}}_{g_\theta(T(y))} \underbrace{\mathbf{1}_{\min\{y_i\}>0}}_{h(y)}
\end{aligned}
$$

The sufficient statistic $T = \prod_{i=1}^{n} y_i$

$$
\begin{aligned}
L\left(y \mid \theta\right) &= \exp[nlg\theta + (\theta-1)\Sigma_{i=1}^{n} log(y_i)] \\
&= \exp[nlg\theta + \theta\Sigma_{i=1}^{n} log(y_i) - \Sigma_{i=1}^{n} log(y_i)] \\
&= \exp[nlg\theta] \cdot \exp[\theta\Sigma_{i=1}^{n} log(y_i)] \cdot \exp[-\Sigma_{i=1}^{n} log(y_i)]
\end{aligned}
$$

The interpretable form is $\Sigma_{i=1}^{n} log(y_i)$, a log-tranformed sum of data