SM Exercise 11.1.3

Kristine Villaluna

19/02/2023

(a) SM Exercise 11.1.3.

Suppose $x_1, \ldots, x_n$ is an i.i.d. sample from a Poisson distribution with expected value θ:

$$f(x_i \mid \theta) = \frac{\theta^{x_i} e^{-\theta}}{x_i!}, \quad x = 0, 1, 2, \ldots; \quad \theta > 0.$$

Suppose that we use a Gamma prior for θ:

$$\pi(\theta) = g(\theta; \alpha, \lambda) = \frac{\lambda^\alpha \theta^{\alpha-1}}{\Gamma(\alpha)} \exp(-\lambda\theta).$$

1. Show that the posterior density is $g(\theta; \alpha + \Sigma x_i, \lambda + n)$, and find conditions under which the posterior density remains proper (i.e. integrates to 1) as $\alpha \downarrow 0$, even though the prior density becomes improper in that limit.

We are given that $x_1, \ldots, x_n$ is an i.i.d. sample from a Poisson distribution, $f(x_i \mid \theta)$, with expected value θ. Let us denote $(x_1, \ldots, x_n) = x$, and $p(\theta|x)$ as the joint posterior probability function.

Then, since $x_1, \ldots, x_n$ is i.i.d, we can write:

$$\begin{aligned}
f(x|\theta) &= \prod_{i=1}^{n} f(x_i|\theta) \\
&= \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} \\
&= \frac{\theta^{\sum_{i=1}^n x_i} e^{-n\theta}}{\prod_{i=1}^n x_i!}
\end{aligned}$$

Next, we can calculate the posterior probability. The posterior probability is proportional to the likelihood times the prior as x is constant w.r.t θ. This can be shown as follows:

$$\begin{aligned}
p(\theta|x) &\propto \frac{f(x|\theta)\pi(\theta)}{f(x)} \\
&\propto \pi(\theta) f(x|\theta) \\
&= \left[\frac{\lambda^\alpha \theta^{\alpha-1}}{\Gamma(\alpha)\exp(-\lambda\theta)}\right]\left[\frac{\theta^{\sum_{i=1}^n x_i} e^{-n\theta}}{\prod_{i=1}^n x_i!}\right] \\
&\propto \left[\theta^{\alpha-1}\exp(-\lambda\theta)\right]\left[\theta^{\sum_{i=1}^n x_i} e^{-n\theta}\right] \\
&= \theta^{(\alpha+\sum_{i=1}^n x_i)-1} \exp\{-(\lambda+n)\theta\}
\end{aligned}$$

Which is the probability density function of Gamma$(\alpha + \Sigma x_i, \lambda + n)$.

Therefore, the posterior density Gamma$(\alpha + \Sigma x_i, \lambda + n)$, i.e., $g(\theta; \alpha + \Sigma x_i, \lambda + n)$.

Next, we must find conditions under which the posterior density remains proper (i.e. integrates to 1) as $\alpha \downarrow 0$, even though the prior density becomes improper in that limit.

First, we know that since $x_1, \ldots, x_n$ is i.i.d prom a Poisson distribution, $x_i \geq 0, i = 1, \ldots, n$. Next, the parameter space for the posterior density can be written as $\alpha + \sum_{i=1}^n x_i > 0, \lambda + n > 0$. Therefore, $\alpha + \sum_{i=1}^n x_i > 0$ as $\alpha \downarrow 0$.

Assuming $n > 0$, $\lambda + n > 0 \implies \lambda > -n$. Therefore, the conditions are: $\lambda > -n$ and $x_i \geq 0, i = 1, \ldots, n$.

2. Find the expected value of θ under the prior and under the posterior, and hence give an interpretation of the prior parameters.

First, let us find the expected value of θ under the prior. This can be written as:

$$
\begin{aligned}
E_\pi(\theta) &= \int_0^\infty \theta \frac{\lambda^\alpha \theta^{\alpha-1}}{\Gamma(\alpha)} \exp(-\lambda\theta) d\theta \\
&= \int_0^\infty \frac{\lambda^\alpha \theta^\alpha}{\Gamma(\alpha)} \exp(-\lambda\theta) d\theta \\
&= \int_0^\infty \frac{\alpha}{\lambda} \frac{\lambda^{\alpha+1}\theta^\alpha}{\alpha\Gamma(\alpha)} \exp(-\lambda\theta) d\theta \\
&= \frac{\alpha}{\lambda} \int_0^\infty \frac{\lambda^{\alpha+1}\theta^\alpha}{\Gamma(\alpha+1)} \exp(-\lambda\theta) d\theta \\
&= \frac{\alpha}{\lambda}
\end{aligned}
$$

Where we know that $\Gamma(n+1) = n\Gamma(n)$ and $\frac{\lambda^{\alpha+1}\theta^\alpha}{\Gamma(\alpha+1)} \exp(-\lambda\theta)$ is the pdf of Gamma$(\theta; \alpha + 1, \lambda)$.

Next, we can find the expected value of θ under the posterior:

$$
\begin{aligned}
E_p(\theta) &= \int_0^\infty \theta \cdot \theta^{(\alpha + \sum_{i=1}^n x_i) - 1} \exp\{-(\lambda + n)\theta\} d\theta \\
&= \int_0^\infty \theta \frac{(\lambda+n)^{\alpha + \sum_{i=1}^n x_i} \theta^{\alpha + \sum_{i=1}^n x_i - 1}}{\Gamma(\alpha + \sum_{i=1}^n x_i)} \exp(-(\lambda+n)\theta) d\theta \\
&= \int_0^\infty \frac{(\lambda+n)^{\alpha + \sum_{i=1}^n x_i} \theta^{\alpha + \sum_{i=1}^n x_i}}{\Gamma(\alpha + \sum_{i=1}^n x_i)} \exp(-(\lambda+n)\theta) d\theta \\
&= \frac{\alpha + \sum_{i=1}^n x_i}{\lambda + n} \int_0^\infty \frac{(\lambda+n)^{\alpha + \sum_{i=1}^n x_i + 1} \theta^{\alpha + \sum_{i=1}^n x_i}}{(\alpha + \sum_{i=1}^n x_i)\Gamma(\alpha + \sum_{i=1}^n x_i)} \exp(-(\lambda+n)\theta) d\theta \\
&= \frac{\alpha + \sum_{i=1}^n x_i}{\lambda + n} \int_0^\infty \frac{(\lambda+n)^{\alpha + \sum_{i=1}^n x_i + 1} \theta^{\alpha + \sum_{i=1}^n x_i}}{\Gamma(\alpha + \sum_{i=1}^n x_i + 1)} \exp(-(\lambda+n)\theta) d\theta \\
&= \frac{\alpha + \sum_{i=1}^n x_i}{\lambda + n}
\end{aligned}
$$

Again, where we know that $\Gamma(n+1) = n\Gamma(n)$ and $\frac{(\lambda+n)^{\alpha + \sum_{i=1}^n x_i + 1} \theta^{\alpha + \sum_{i=1}^n x_i}}{\Gamma(\alpha + \sum_{i=1}^n x_i + 1)} \exp(-(\lambda+n)\theta)$ is the pdf of Gamma$(\theta; \alpha + 1, \lambda)$.

Now, we can try to interpret our prior parameters. α can be thought of as the number of events in λ time units, as θ is the mean parameter of the Poisson(θ) distribution.

3. Find Jeffreys' prior $\pi_J(\theta) \propto i^{1/2}(\theta)$. (The proportionality is meant to convey the impropriety of the density.)

First, we note that

$$i(\theta) = -E\left[\frac{d^2\ell(\theta; x)}{d\theta^2}\right]$$

Thus, in order to find Jeffreys' prior, we must first find the likelihood and log-likelihood functions.

The likelihood function is as follows:

$$\begin{aligned}
\mathcal{L}(\theta; x) &= \prod_{i=1}^{n} f(x_i|\theta) \\
&= \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} \\
&= \frac{\theta^{\sum_{i=1}^{n} x_i} e^{-n\theta}}{\prod_{i=1}^{n} x_i!}
\end{aligned}$$

Next, we can take the log to find the log-likelihood:

$$\begin{aligned}
\ell(\theta; x) &= \log(\mathcal{L}(\theta; x)) \\
&= \log\left(\theta^{\sum_{i=1}^{n} x_i}\right) + \log(e^{-n\theta}) - \log\left(\prod_{i=1}^{n} x_i!\right) \\
&= \left(\sum_{i=1}^{n} x_i\right) \log\left(\theta\right) - n\theta \log(e) - \log\left(\prod_{i=1}^{n} x_i!\right) \\
&= \left(\sum_{i=1}^{n} x_i\right) \log\left(\theta\right) - n\theta - \log\left(\prod_{i=1}^{n} x_i!\right)
\end{aligned}$$

Taking the first derivative, we obtain:

$$\begin{aligned}
\frac{d}{d\theta}\ell(\theta; x) &= \frac{d}{d\theta}\left[\left(\sum_{i=1}^{n} x_i\right) \log\left(\theta\right) - n\theta - \log\left(\prod_{i=1}^{n} x_i!\right)\right] \\
&= \left(\sum_{i=1}^{n} x_i\right) \frac{1}{\theta} - n
\end{aligned}$$

Taking the second derivative, we obtain:

$$\begin{aligned}
\frac{d^2}{d\theta^2}\ell(\theta; x) &= \frac{d}{d\theta}\left[\frac{d}{d\theta}\ell(\theta; x)\right] \\
&= \frac{d}{d\theta}\left[\left(\sum_{i=1}^{n} x_i\right) \frac{1}{\theta} - n\right] \\
&= -\left(\sum_{i=1}^{n} x_i\right) \frac{1}{\theta^2}
\end{aligned}$$

Next, we can plug this into the formula for $i(\theta)$:

$$\begin{aligned} i(\theta) &= -E\left[\frac{d^2\ell(\theta;x)}{d\theta^2}\right] \\ &= -E\left[-\left(\sum_{i=1}^{n} x_i\right)\frac{1}{\theta^2}\right] \\ &= \frac{1}{\theta^2}\sum_{i=1}^{n} E(x_i) \\ &= \frac{1}{\theta^2}(n\theta) \\ &= \frac{n}{\theta} \end{aligned}$$

Thus, Jeffreys' prior is given by

$$\begin{aligned} \pi_J(\theta) &\propto i^{1/2}(\theta) \\ &= \left(\frac{n}{\theta}\right)^{1/2} \\ &\propto \sqrt{\frac{1}{\theta}} \end{aligned}$$

4. Let X_{new} be a new observation independent of $x_1, \ldots, x_n$,but following the same Poisson(θ) distribution. Find its posterior predictive density. To what density does this converge as $n \to \infty$?

The posterior predictive density is given by:

$$p(X_{new}|x) = \int_{\theta=0}^{\infty} p(X_{new}|\theta)p(\theta|x)d\theta$$

where

$$p(X_{new}|\theta) = f(X_{new}|\theta) = \frac{\theta^{X_{new}}e^{-\theta}}{X_{new}!}$$

and

$$p(\theta|x) = \frac{(\lambda+n)^{\alpha+\sum_{i=1}^{n} x_i}\theta^{\alpha+\sum_{i=1}^{n} x_i - 1}}{\Gamma(\alpha+\sum_{i=1}^{n} x_i)}\exp(-(\lambda+n)\theta)$$

Thus, we can find the posterior predictive density of X_{new} as follows:

$$
\begin{aligned}
p(X_{new}|x) &= \int_{\theta=0}^{\infty} \frac{\theta^{X_{new}} e^{-\theta}}{X_{new}!} \frac{(\lambda+n)^{\alpha+\sum_{i=1}^n x_i} \theta^{\alpha+\sum_{i=1}^n x_i - 1}}{\Gamma(\alpha+\sum_{i=1}^n x_i)} \exp(-(\lambda+n)\theta) d\theta \\
&= \frac{(\lambda+n)^{\alpha+\sum_{i=1}^n x_i}}{X_{new}!\Gamma(\alpha+\sum_{i=1}^n x_i)} \int_0^{\infty} \theta^{\alpha+\sum_{i=1}^n x_i + X_{new} - 1} \exp(-(\lambda+n+1)\theta) d\theta \\
&= \frac{\Gamma(\alpha+\sum_{i=1}^n x_i + X_{new})}{X_{new}!\Gamma(\alpha+\sum_{i=1}^n x_i)} \frac{(\lambda+n)^{\alpha+\sum_{i=1}^n x_i}}{(\lambda+n+1)^{\alpha+\sum_{i=1}^n x_i + X_{new}}} \times \\
&\int_0^{\infty} \frac{(\lambda+n+1)^{\alpha+\sum_{i=1}^n x_i + X_{new}}}{\Gamma(\alpha+\sum_{i=1}^n x_i)+X_{new}} \theta^{\alpha+\sum_{i=1}^n x_i + X_{new} - 1} \exp(-(\lambda+n+1)\theta) d\theta \\
&= \frac{\Gamma(\alpha+\sum_{i=1}^n x_i + X_{new})}{X_{new}!\Gamma(\alpha+\sum_{i=1}^n x_i)} \frac{(\lambda+n)^{\alpha+\sum_{i=1}^n x_i}}{(\lambda+n+1)^{\alpha+\sum_{i=1}^n x_i + X_{new}}} \\
&= \frac{(\alpha+\sum_{i=1}^n x_i + X_{new} - 1)!}{X_{new}!(\alpha+\sum_{i=1}^n x_i - 1)!} \left(\frac{\lambda+n}{\lambda+n+1}\right)^{\alpha+\sum_{i=1}^n x_i} \left(\frac{1}{\lambda+n+1}\right)^{X_{new}} \\
&= \binom{(X_{new})+(\alpha+\sum_{i=1}^n x_i)-1}{(\alpha+\sum_{i=1}^n x_i)-1} \left(\frac{\lambda+n}{\lambda+n+1}\right)^{\alpha+\sum_{i=1}^n x_i} \left(\frac{1}{\lambda+n+1}\right)^{X_{new}}
\end{aligned}
$$

Where we know $\Gamma(n) = (n-1)!$ and $\int_0^{\infty} \frac{(\lambda+n+1)^{\alpha+\sum_{i=1}^n x_i + X_{new}}}{\Gamma(\alpha+\sum_{i=1}^n x_i)+X_{new}} \theta^{\alpha+\sum_{i=1}^n x_i + X_{new} - 1} \exp(-(\lambda+n+1)\theta) d\theta = 1$.

Looking at this result, we can recognize this as the negative binomial distribution,

$$
f(X_{new}; r; p) = \binom{X_{new+r-1}}{r-1} (1-p)^r p^{X_{new}}
$$

where $r = \alpha + \sum_{i=1}^n x_i = \alpha + n\bar{x}$ and $p = \frac{1}{\lambda+n+1}$.

Next, we need to find what density this converges to as $n \to \infty$. For convenience, let us use the r, p, and $k = X_{new}$. By doing this, evaluating the density as $n \to \infty$ is equivalent to evaluating the density which uses k, r, and p as $r \to \infty$ as α and $\bar{x}$ are constant under this parameterization.

First, recall that the mean (m) of a $k \sim$ Negative Binomial(r, p) random variable is given by $m = \frac{pr}{(1-p)}$.

Then, rearranging the formula yields:

$$
\begin{aligned}
pr &= m - mp \\
p &= \frac{m}{r} - \frac{mp}{r} \\
p\left(1 + \frac{m}{r}\right) &= \frac{m}{r}
\end{aligned}
$$

Thus, we can write p and $(1-p)$ as:

$$
p = \frac{m}{m+r}
$$

and

$$
\begin{aligned}
1 - p &= \frac{r+m}{r+m} - \frac{m}{r+m} \\
&= \frac{r}{r+m}
\end{aligned}
$$

Substituting these values back into the density we get:

$$\begin{aligned}
f(k;r;p) &= \frac{\Gamma(k+r)}{k!\Gamma(r)}p^k(1-p)^r \\
&= \frac{\Gamma(k+r)}{k!\Gamma(r)}\left(\frac{m}{m+r}\right)^k\left(\frac{r}{r+m}\right)^r \\
&= \frac{m^k}{k!}\left(\frac{\Gamma(k+r)}{\Gamma(r)(r+m)^k}\right)\left(\frac{r}{r+m}\right)^r \\
&= \frac{m^k}{k!}\left(\frac{\Gamma(k+r)}{\Gamma(r)(r+m)^k}\right)\left(\frac{r/r}{(r+m)/r}\right)^r \\
&= \frac{m^k}{k!}\left(\frac{\Gamma(k+r)}{\Gamma(r)(r+m)^k}\right)\left(\frac{1}{1+\frac{m}{r}}\right)^r
\end{aligned}$$

Taking the limit as $r \to \infty$:

$$\begin{aligned}
\lim_{r\to\infty} f(k;r;p) &= \left(\frac{m^k}{k!}\right)(1)\left(\frac{1}{e^m}\right) \\
&= \left(\frac{m^k}{k!}\right)\left(\frac{1}{e^m}\right) \\
&= \frac{m^k e^{-m}}{k!}
\end{aligned}$$

which is the probability mass function of a Poisson random variable with mean m.

Thus,

$$\begin{aligned}
m &= \frac{pr}{(1-p)} \\
&= \frac{\frac{1}{\lambda+n+1}(\alpha+n\bar{x})}{\frac{\lambda+n}{\lambda+n+1}} \\
&= \frac{\alpha+n\bar{x}}{\lambda+n}
\end{aligned}$$

Where $r = \alpha + \sum_{i=1}^{n} x_i = \alpha + n\bar{x}$, $p = \frac{1}{\lambda+n+1}$, and $1-p = \frac{\lambda+n}{\lambda+n+1}$.

Therefore, the posterior predictive density converges to a

$$\text{Poisson}\left(\frac{\alpha+n\bar{x}}{\lambda+n}\right)$$