Isabel Muzio, Francesco Caporali
22 October 2022

Bayesian Statistics
Assignment 1

Question 1: The Galenshore distribution

Point a.

$Y|\theta \sim \text{Galenshore}\,(a, \theta)$ is such that $p(y|\theta)$ is a density in the exponential family indeed

$$p(y|\theta) = \frac{2}{\Gamma(a)}\theta^{2a}y^{2a-1}e^{-\theta^2 y^2}\mathbb{1}_{\{y>0\}}, \theta > 0, a > 0.$$

Then definining $\phi \overset{\text{def}}{=} \theta^2$ one has

$$p(y|\phi) = h(y)c(\phi)e^{\phi t(y)} \text{ with } h(y) = \frac{2y^{2a-1}}{\Gamma(a)}\mathbb{1}_{\{y>0\}}, c(\phi) = \phi^a, t(y) = -y^2,$$

hence, by the easy shape of a distribution in the exponential family, we can state that a class of conjugate priors for $p(y|\phi)$ is such that

$$p(\phi) \propto c(\phi)^{n_0}e^{\phi n_0 t_0} = \phi^{an_0}e^{\phi n_0 t_0}.$$

If ϕ has density $p(\phi)$ and we want to obtain the density of $\theta = \sqrt{\phi}$ it is sufficient to define the map $f : \mathbb{R}^+ \to \mathbb{R}^+$ such that $f(x) = \sqrt{x}$ and recall that

$$p_\theta(\theta) = p_\phi(f(\phi)) = p_\phi(f^{-1}(\theta)\left|\frac{df^{-1}(\theta)}{d\theta}\right|.$$

Observing that $\left|\frac{df^{-1}(\theta)}{d\theta}\right| = \left|\frac{d\theta^2}{d\theta}\right| = 2\theta$

$$p(\theta) \overset{\text{def}}{=} p_\theta(\theta) = p_\phi(\theta^2)2\theta \propto \theta^{2an_0}e^{\theta^2 n_0 t_0}2\theta.$$

Remark

Observing the three parameters a, n_0 and t_0 we can say

- $a > 0$ by hypotesis;
- $n_0 > 0$ because it represents the *prior sample size* ($p(\theta)$ has the same kernel of $p(y|\theta)$ after n_0 observations);
- $t_0 < 0$ because it is the *prior guess* that we make for t, with $t(y) = -\frac{y^2}{2}, \forall y \in \mathbb{R}^+ : t_0 = \frac{\sum_{i=1}^n t(y_i)}{n} = -\frac{\sum_{i=1}^n y_i^2}{n} < 0.$

Hence we have

$$an_0 + 1 > 0 \text{ and } -n_0 t_0 > 0.$$

So we can rewrite

$$p(\theta) \propto 2\theta^{2an_0+1}e^{-\left(\sqrt{-n_0t_0}\right)^2\theta^2}$$

and recognizing the kernel of a Galenshore distribution we can write explicitly

$$p(\theta) = 2\theta^{2(an_0+1)-1}e^{-\left(\sqrt{-n_0t_0}\right)^2\theta^2} \cdot \underbrace{\frac{\left(\sqrt{-n_0t_0}\right)^{2(an_0+1)}}{\Gamma(an_0+1)}}_{\text{it does not depends on }\theta} \underbrace{\mathbb{1}_{\theta>0}}_{\text{by hypotesis}} =$$

$$= \frac{2}{\Gamma(an_0+1)}\left(\sqrt{-n_0t_0}\right)^{2(an_0+1)}\theta^{2(an_0+1)-1}e^{-\left(\sqrt{-n_0t_0}\right)^2\theta^2}\mathbb{1}_{\theta>0}$$

$$\Downarrow$$

$$\theta \sim \text{Galenshore}\left(an_0+1, \sqrt{-n_0t_0}\right).$$

Finally we plot a few of these densities Galenshore $(an_0+1, \sqrt{-n_0t_0})$ sampled with the following code:

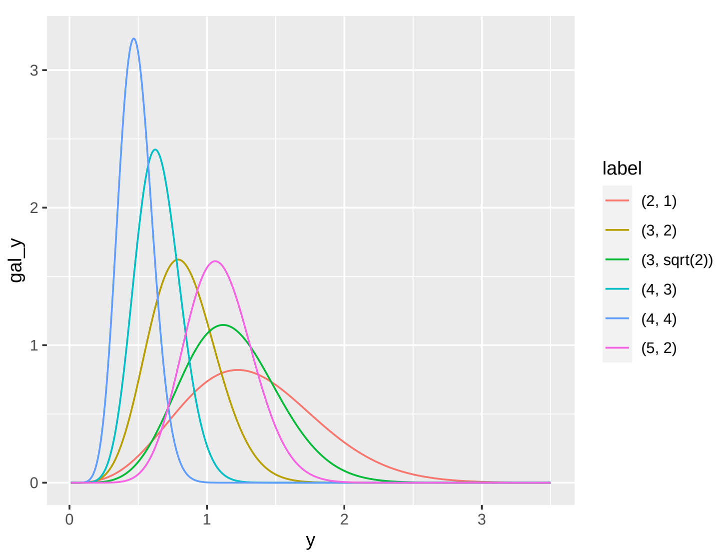
- $n_0 = 1, t_0 = -1, a = 1 \implies \text{Galenshore}\,(2, 1);$
- $n_0 = 2, t_0 = -1, a = 1 \implies \text{Galenshore}\left(3, \sqrt{2}\right);$
- $n_0 = 2, t_0 = -2, a = 1 \implies \text{Galenshore}\,(3, 2);$
- $n_0 = 2, t_0 = -2, a = 2 \implies \text{Galenshore}\,(5, 2);$
- $n_0 = 3, t_0 = -3, a = 1 \implies \text{Galenshore}\,(4, 3);$
- $n_0 = 3, t_0 = -4, a = 1 \implies \text{Galenshore}\,(4, 4).$

```
dgalenshore = function(y, a, theta) {
    (2 / gamma(a)) * theta^(2 * a) * y^(2 * a - 1) * exp(-(theta^2) * y^2)
}

y = seq(0.01, 3.5, length = 1000)
df = rbind(
    data.frame(y = y, gal_y = dgalenshore(y, 2, 1), label = "(2, 1)"),
    data.frame(y = y, gal_y = dgalenshore(y, 3, sqrt(2)), label = "(3, sqrt(2))"),
    data.frame(y = y, gal_y = dgalenshore(y, 3, 2), label = "(3, 2)"),
    data.frame(y = y, gal_y = dgalenshore(y, 5, 2), label = "(5, 2)"),
    data.frame(y = y, gal_y = dgalenshore(y, 4, 3), label = "(4, 3)"),
    data.frame(y = y, gal_y = dgalenshore(y, 4, 4), label = "(4, 4)")
)
```

Then we plot all of them at the same time:

```
ggplot(df, aes(y, gal_y, group = label, color = label)) +
geom_line() + coord_fixed(ratio = 1)
```



Point b.

Let's define $b \stackrel{\text{def}}{=} an_0 + 1$ and $c = \sqrt{-n_0 t_0}$ for coinciseness.
Recalling $\theta \sim \text{Galenshore}\,(b, c)$ and $Y_i|\theta \sim \text{Galenshore}\,(a, \theta)\,, \forall i \in 1:n$ and defining $\text{SS}\,(y_{1:n}) \stackrel{\text{def}}{=} \sum_{i=1}^n y_i^2$ we have

$$\begin{aligned} p(\theta|y_{1:n}) = p(\theta)p(y_{1:n}|\theta) \propto \\ &\propto (\theta^{2b-1}e^{-c^2\theta^2})(\theta^{2na}e^{-\theta^2\sum_{i=1}^n y_i^2}) \propto \\ &\propto \theta^{2(an+b)-1}e^{-(c^2+\text{SS}(y_{1:n}))\theta^2}. \end{aligned}$$

Hence we recognize the kernel of a Galenshore $\left(an + b, \sqrt{c^2 + \text{SS}\,(y_{1:n})}\right)$

$$\implies \theta|Y_{1:n} \sim \text{Galenshore}\left(a(n+n_0)+1, \sqrt{\text{SS}\,(y_{1:n}) - n_0 t_0}\right).$$

Point c.

$$\begin{aligned} \frac{p(\theta_a|y_{1:n})}{p(\theta_b|y_{1:n})} &= \frac{2\Gamma(a(n+n_0)+1)}{\Gamma(a(n+n_0)+1)2}\left(\text{SS}\,(y_{1:n}) - n_0 t_0\right)^{(a(n+n_0)+1)(1-1)}\left(\frac{\theta_a}{\theta_b}\right)^{2a(n+n_0)+1} e^{-(\text{SS}(y_{1:n})-n_0t_0)\left(\theta_a^2-\theta_b^2\right)} = \\ &= \left(\frac{\theta_a}{\theta_b}\right)^{2a(n+n_0)+1} e^{-\left(\sum_{i=1}^n y_i^2 - n_0 t_0\right)\left(\theta_a^2-\theta_b^2\right)}. \end{aligned}$$

Hence

$$\mathbb{P}\left(\theta \in A | Y_{1:n} = y_{1:n}\right) = \mathbb{P}\left(\theta \in A | \sum_{i=1}^n Y_i^2 = \sum_{i=1}^n y_i^2\right), \forall A$$

and then, by definition $\text{SS}\,(Y_{1:n}) \stackrel{\text{def}}{=} \sum_{i=1}^n Y_i^2$ is a sufficient statistic.

Point d.

Recalling that $\theta|Y_{1:n} \sim \text{Galenshore}\left(a(n+n_0)+1, \sqrt{\text{SS}\,(y_{1:n}) - n_0 t_0}\right)$ and that if $X \sim \text{Galenshore}\,(a, \theta) \implies \mathbb{E}\,[X] = \frac{\Gamma\left(a+\frac{1}{2}\right)}{\theta\Gamma(a)}$ we have

$$\mathbb{E}\,[\theta|y_{1:n}] = \frac{\Gamma\left(a(n+n_0) + frac32\right)}{\left(\sqrt{\text{SS}\,(y_{1:n}) - n_0 t_0}\right)\Gamma(a(n+n_0)+1)}$$

Point e.

With the usual notation $b \stackrel{\text{def}}{=} an_0 + 1$ and $c \stackrel{\text{def}}{=} \sqrt{-n_0 t_0}$:

$$\begin{aligned} p(y_{n+1}|y_{1:n}) &= \int_0^\infty p(y_{n+1}|\theta)p(\theta|y_{1:n})d\theta = \\ &= \int_0^\infty \frac{2}{\Gamma(a)}\theta^{2a}y_{n+1}^{2a-1}e^{-\theta^2 y_{n+1}^2} \cdot \frac{2}{\Gamma(an+b)}\left(c^2 + \text{SS}\,(y_{1:n})\right)^{an+b}\theta^{2(an+b)-1}e^{-(c^2+\text{SS}(y_{1:n}))\theta^2}d\theta = \\ &= \frac{4}{\Gamma(a)\Gamma(an+b)}y_{n+1}^{2a-1}\left(c^2 + \text{SS}\,(y_{1:n})\right)^{an+b}\int_0^\infty \underbrace{\theta^{2(a+an+b)-1}e^{-(c^2+\text{SS}(y_{1:n})+y_{n+1}^2)\theta^2}}_{\text{kernel of a Galenshore}\left(a+an+b, \sqrt{c^2+\text{SS}(y_{1:n})+y_{n+1}^2}\right)} d\theta \end{aligned}$$

$$\Downarrow$$

$$\int_0^\infty \theta^{2(a+an+b)-1} e^{-(c^2+\mathrm{SS}(y_{1:n})+y_{n+1}^2)\theta^2} = \frac{\Gamma(a+an+b)}{2}\left(\frac{1}{c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2}\right)^{a+an+b}$$

$$\Downarrow$$

$$\begin{aligned}
p(y_{n+1}|y_{1:n}) &= \frac{4}{\Gamma(a)\Gamma(an+b)} y_{n+1}^{2a-1}\left(c^2+\mathrm{SS}\left(y_{1:n}\right)\right)^{an+b}\frac{\Gamma(a+an+b)}{2}\left(\frac{1}{c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2}\right)^{a+an+b} = \\
&= \frac{2}{y_{n+1}}\frac{\Gamma(a+an+b)}{\Gamma(a)\Gamma(an+b)}\left(\frac{y_{n+1}^2}{c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2}\right)^{a}\left(\frac{c^2+\mathrm{SS}\left(y_{1:n}\right)}{c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2}\right)^{an+b} = \\
&= \frac{2y_{n+1}\left(c^2+\mathrm{SS}\left(y_{1:n}\right)\right)}{\left(c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2\right)^2}\cdot \\
&\quad \cdot\underbrace{\frac{1}{B(a,an+b)}\left(\frac{y_{n+1}^2}{c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2}\right)^{a-1}\left(1-\frac{y_{n+1}^2}{c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2}\right)^{(an+b)-1}}_{\text{density of } X\sim B(a,an+b) \text{ evaluated on } \frac{y_{n+1}^2}{c^2+\mathrm{SS}(y_{1:n})+y_{n+1}^2}} = \\
&= \frac{2y_{n+1}\left(c^2+\mathrm{SS}\left(y_{1:n}\right)\right)}{\left(c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2\right)^2} p_X\left(\frac{y_{n+1}^2}{c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2}\right).
\end{aligned}$$

Now one should note that this is a differentiable transformation of $\mathbb{R}^+$ of an unknown random variable. Indeed if one try to derive, with respect to y_{n+1}, the variable of the density of X in our last expression obtains

$$\begin{aligned}
\frac{d}{dy_{n+1}}\frac{y_{n+1}^2}{c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2} &= \frac{2y_{n+1}(c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2)-y_{n+1}^2 2y_{n+1}}{\left(c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2\right)^2} = \\
&= \underbrace{\frac{2y_{n+1}(c^2+\mathrm{SS}\left(y_{1:n}\right))}{\left(c^2+\mathrm{SS}\left(y_{1:n}\right)+y_{n+1}^2\right)^2}}_{>0 \text{ indeed } y_{n+1}\in\mathbb{R}^+ \text{ and the other terms are squared}}.
\end{aligned}$$

Hence if we define $f^{-1}:\mathbb{R}^+\to\mathbb{R}^+$ such that $f^{-1}(x) = \dfrac{x^2}{c^2+\mathrm{SS}\left(y_{1:n}\right)+x^2}$ we can state

$$\begin{aligned}
p(y_{n+1}|y_{1:n}) &= \left|\frac{d}{dy_{n+1}}f^{-1}(y_{n+1})\right| p_X(f^{-1}(y_{n+1}) = \\
&= p_{f(X)}(y_{n+1}).
\end{aligned}$$

Let's compute f explicitly

$$f^{-1}(x) = \frac{x^2}{c^2+\mathrm{SS}\left(y_{1:n}\right)+x^2} = 1-\frac{c^2+\mathrm{SS}\left(y_{1:n}\right)}{c^2+\mathrm{SS}\left(y_{1:n}\right)+x^2}$$

hence

$$\begin{aligned}
x = 1-\frac{c^2+\mathrm{SS}\left(y_{1:n}\right)}{c^2+\mathrm{SS}\left(y_{1:n}\right)+f(x)^2} &\iff 1-x = \frac{c^2+\mathrm{SS}\left(y_{1:n}\right)}{c^2+\mathrm{SS}\left(y_{1:n}\right)+f(x)^2} \iff \\
&\iff \frac{f(x)^2}{c^2+\mathrm{SS}\left(y_{1:n}\right)}+1 = \frac{1}{1-x} \iff \\
&\iff f(x) = \sqrt{\frac{x}{1-x}}\sqrt{c^2+\mathrm{SS}\left(y_{1:n}\right)}.
\end{aligned}$$

This leads us to conclude (substituting again $b = an_0 + 1$ and $c = \sqrt{-n_0 t_0}$) that

$$Y_{n+1}|Y_{1:n} \sim f\left(B(a, a(n+n_0)+1)\right), \text{ with } f : \mathbb{R}^+ \to \mathbb{R}^+, f(x) = \sqrt{\frac{x}{1-x}}\sqrt{\mathrm{SS}\left(y_{1:n}\right) - n_0 t_0}.$$