Administrative Issues

- Homeworks 3 and 4 are still being graded, but solutions to HWs 1-4 are available on Canvas. These are located on Assignment tab where you submitted the homework.
- In this course, the Normal distribution is parametrized by **mean** and **variance**. As stated previously, software packages may use a different parametrization. It is your responsibility to understand the software you are using, and how to convert between parametrizations.
- Midterm: Oct 20 12:00am ET - Oct23 11:59pm ET. The withdraw deadline is October 29, 2022 4:00 PM ET, and we plan to have your grades returned before then (however *how much* before is TBD).
- No public discussion of the midterm is permitted. If you think there is an error in the problem statement please let us know via **private message on Ed Discussion**.

1 Gibbs pdf - more algebraic steps

Example (from Gibbs.pdf): We illustrate Gibbs sampler and finding the full conditionals in the model:

$$\begin{aligned} Y_1, Y_2, \ldots, Y_n &\sim N(\mu, 1/\tau) \\ \mu &\sim N(\mu_0, 1/\tau_0) \\ \tau &\sim Ga(a, b) \end{aligned}$$

where τ, τ_0 are precision parameters (reciprocals of variances), and b is a rate parameter. Note that in the above, the normal distribution is parametrized with mean and variance, as usual. The joint distribution is

$$\begin{aligned} f(y, \mu, \tau) &= \left(\prod_{i=1}^{n} f(y_i \mid \mu, \tau)\right) \pi(\mu)\pi(\tau) \\ &= \left(\prod_{i=1}^{n} \frac{\tau^{1/2}}{\sqrt{2\pi}} e^{-0.5\tau(y_i-\mu)^2}\right) \frac{\tau_0}{\sqrt{2\pi}} e^{-0.5\tau_0(\mu-\mu_0)^2 \tau^{a-1}} e^{-b\tau} \\ &= \left(\frac{\tau^{n/2}}{(2\pi)^{n/2}} e^{-\sum_{i=1}^{n} 0.5\tau(y_i-\mu)^2}\right) \frac{\tau_0}{\sqrt{2\pi}} e^{-0.5\tau_0(\mu-\mu_0)^2} \tau^{a-1} e^{-b\tau} \\ &\propto \tau^{n/2} e^{-\sum_{i=1}^{n} 0.5\tau(y_i-\mu)^2} \tau_0 e^{-0.5\tau_0(\mu-\mu_0)^2} \tau^{a-1} e^{-b\tau} \end{aligned}$$

To find the full conditional distribution for μ we select the terms from $f(y, \mu, \tau)$ that contain μ and normalize. We have:

$$\pi(\mu \mid \tau, y) = \frac{\pi(\mu, \tau \mid y)}{\pi(\tau \mid y)} = \frac{\pi(\mu, \tau, y)}{\pi(\tau, y)} \propto \pi(\mu, \tau, y)$$

Thus,

$$\begin{aligned}
\pi(\mu \mid y, \tau) &= e^{-\sum_{i=1}^{n} 0.5\tau(y_i-\mu)^2} e^{-0.5\tau_0(\mu-\mu_0)^2} \\
&= \exp\left(-\sum_{i=1}^{n} 0.5\tau\left(y_i-\mu\right)^2 - 0.5\tau_0\left(\mu-\mu_0\right)^2\right) \\
&= \exp\left(-0.5\tau\sum_{i=1}^{n} y_i^2 + \tau\mu\sum_{i=1}^{n} y_i - 0.5n\tau\mu^2 - 0.5\tau_0\mu^2 + \tau_0\mu\mu_0 - 0.5\tau_0\mu_0^2\right) \\
&\propto \exp\left(-0.5\left(n\tau+\tau_0\right)\mu^2 + \left(\tau\sum_{i=1}^{n} y_i + \tau_0\mu_0\right)\mu\right) \\
&= \exp\left(-0.5\left(n\tau+\tau_0\right)\mu^2 + \frac{(-0.5)(n\tau+\tau_0)}{(-0.5)(n\tau+\tau_0)}\left(\tau\sum_{i=1}^{n} y_i + \tau_0\mu_0\right)\mu\right) \\
&= \exp\left(-0.5\left(n\tau+\tau_0\right)\left(\mu^2 - 2\frac{\tau\sum_{i=1}^{n} y_i + \tau_0\mu_0}{n\tau+\tau_0}\mu\right)\right) \\
&\propto \exp\left(-0.5\left(n\tau+\tau_0\right)\left(\mu^2 - 2\frac{\tau\sum_{i=1}^{n} y_i + \tau_0\mu_0}{n\tau+\tau_0}\mu + \left(\frac{\tau\sum_{i=1}^{n} y_i + \tau_0\mu_0}{n\tau+\tau_0}\right)^2\right)\right) \\
&= \exp\left(-0.5\left(n\tau+\tau_0\right)\left(\mu - \frac{\tau\sum_{i=1}^{n} y_i + \tau_0\mu_0}{n\tau+\tau_0}\right)^2\right)
\end{aligned}$$

which is the kernel of a Normal

$$N\left(\frac{\tau\sum_{i=1}^{n} y_i + \tau_0\mu_0}{n\tau+\tau_0}, \frac{1}{n\tau+\tau_0}\right)$$

distribution. We do similar (but easier) work to find the full conditional for τ :

$$\begin{aligned}
\pi(\tau \mid \mu, y) &\propto \tau^{n/2}\exp\left\{-\tau/2\sum_{i=1}^{n}\left(y_i-\mu\right)^2\right\}\tau^{a-1}\exp\{-b\tau\} \\
&\propto \tau^{n/2+a-1}\exp\left\{-\tau/2\sum_{i=1}^{n}\left(y_i-\mu\right)^2 - b\tau\right\} \\
&\propto \tau^{n/2+a-1}\exp\left\{-\tau\left(1/2\sum_{i=1}^{n}\left(y_i-\mu\right)^2 + b\right)\right\}
\end{aligned}$$

which is the kernel of a

$$Ga\left(n/2 + a, 1/2\sum_{i=1}^{n}(y_i - \mu)^2 + b\right)$$

distribution.

2 Why can we ignore normalization when we recognize the distribution?

Good discussion here:

https://stats.stackexchange.com/questions/85465/theoretically-why-do-we-not-need-to-compute-a-marginal-distribution-constant-fo

3 HW3 Q1d Analytical Answer

The Gamma distribution has the following pdf

$$f(x|a,b) = \frac{b^a}{\Gamma(a)}x^{a-1}e^{-bx} \text{ for } x, a, b > 0$$

By definition, we have:

$$\int_0^\infty \frac{b^a}{\Gamma(a)}x^{a-1}e^{-bx}dx = 1$$

Written another way

$$\int_0^\infty x^{a-1}e^{-bx}dx = \frac{\Gamma(a)}{b^a}$$

(d) As $EY = 2\sqrt{\frac{2}{\pi\theta}}$ and $\pi(\theta \mid x) \sim \mathcal{G}amma(5.5, 9.41)$, we find the predictive value for a single future observation as

$$\begin{aligned}
\widehat{y}_{n+1} &= \int_0^\infty 2\sqrt{\frac{2}{\pi\theta}}\frac{\beta^\alpha}{\Gamma(\alpha)}\theta^{\alpha-1}e^{-\beta\theta}d\theta \\
&= \int_0^\infty 2\sqrt{\frac{2}{\pi\theta}}\frac{9.41^{5.5}}{\Gamma(5.5)}\theta^{4.5}e^{-9.41\theta}d\theta \\
&= 2\sqrt{\frac{2}{\pi}}\frac{9.41^{5.5}}{\Gamma(5.5)}\int_0^\infty \theta^4 e^{-9.41\theta}d\theta \\
&= 2\sqrt{\frac{2}{\pi}}\frac{9.41^{0.5}\Gamma(5)}{\Gamma(5.5)} \\
&\approx 2.2445
\end{aligned}$$