Xinyu Tan

– **1. Binomial and multinomial models**

(a) marginal posterior distribution for $\alpha = \frac{\theta_1}{\theta_1+\theta_2}$.
Suppose that the prior distribution for $\theta = (\theta_1, \theta_2, \cdots, \theta_J) \sim$ Dirichlet$(a_1, \cdots, a_J)$. The posterior distribution is

$$\begin{aligned} \theta_1, \cdots, \theta_J | y &\propto \theta_1^{y_1+a_1-1} \cdots \theta_n^{y_J+a_J-1} \\ &\sim \text{Dirichlet}(y_1 + a_1, \cdots, y_J + a_J) \end{aligned}$$

From the appendix A, the marginal of $(\theta_1, \theta_2, 1-\theta_1-\theta_2) \sim$ Dirichlet$(y_1+a_1, y_2+a_2, y_r+a_r)$, where $y_r = y_3 + \cdots + y_J$ and $a_r = a_3 + \cdots + a_J$.
Let $\alpha = h(\theta_1, \theta_2) = \frac{\theta_1}{\theta_1+\theta_2}$ and $\beta = g(\theta_1, \theta_2) = \theta_1 + \theta_2$, then $\theta_1 = \alpha\beta$ and $\theta_2 = \beta(1-\alpha)$.

$$\begin{aligned} \alpha, \beta | y &\propto (\alpha\beta)^{a_1+y_1-1}(\beta(1-\alpha))^{a_2+y_2-1}(1-\beta)^{a_r+y_r-1} \begin{vmatrix} \frac{\partial h}{\partial \theta_1} & \frac{\partial h}{\partial \theta_2} \\ \frac{\partial g}{\partial \theta_1} & \frac{\partial h}{\partial \theta_2} \end{vmatrix}^{-1} \\ &= \alpha^{a_1+y_1-1}(1-\alpha)^{a_2+y_2-1}\beta^{a_1+a_2+y_1+y_2-1}(1-\beta)^{a_r+y_r-1} \\ &\sim \text{Beta}(\alpha | a_1 + y_1, a_2 + y_2)\text{Beta}(\beta | a_1 + a_2 + y_1 + y_2, a_r + y_r) \end{aligned}$$

Since α and β belong to separate factors, they are independent: $\alpha \sim$ Beta(a_1+y_1, a_2+y_2).

– **2. Comparison of two multinomial observations**

Suppose a uniform prior distribution, then the posterior distribution of $\theta_b, \theta_d, \theta_o$ for pre and post debate are

$$(\theta_{b1}, \theta_{d1}, \theta_{o1}) \sim \text{Dirichlet}(294, 307, 38) \quad (1)$$
$$(\theta_{b2}, \theta_{d2}, \theta_{o2}) \sim \text{Dirichlet}(288, 332, 19) \quad (2)$$

From Ex1, we know that the posterior distribution for θ_b, θ_b follows Beta(y_b, y_d), then

$$\alpha_1 \sim \text{Beta}(294, 307) \quad (3)$$
$$\alpha_2 \sim \text{Beta}(288, 332) \quad (4)$$

The histogram is shown 1.

The posterior probability that there is a shift toward Bush: $p(\alpha_2 > \alpha_1) \Rightarrow p(\alpha_2 - \alpha_1 > 0) = 0.1898$.

– **9. Conjugate normal model**

Suppose y is an i.i.d sample of size n from the distribution $N(\mu, \sigma^2)$. Prior for (μ, σ^2) is

$$\text{N-Inv-}\chi^2(\mu, \sigma^2 | \mu_0, \sigma_0^2/\kappa_0; \nu_0, \sigma_0^2)$$

This is derived from

$$\begin{aligned} \mu | \sigma^2 &\sim N(\mu_0, \sigma^2/\kappa_0) \\ \sigma^2 &\sim \text{Inv-}\chi^2(\nu_0, \sigma_0^2) \end{aligned}$$

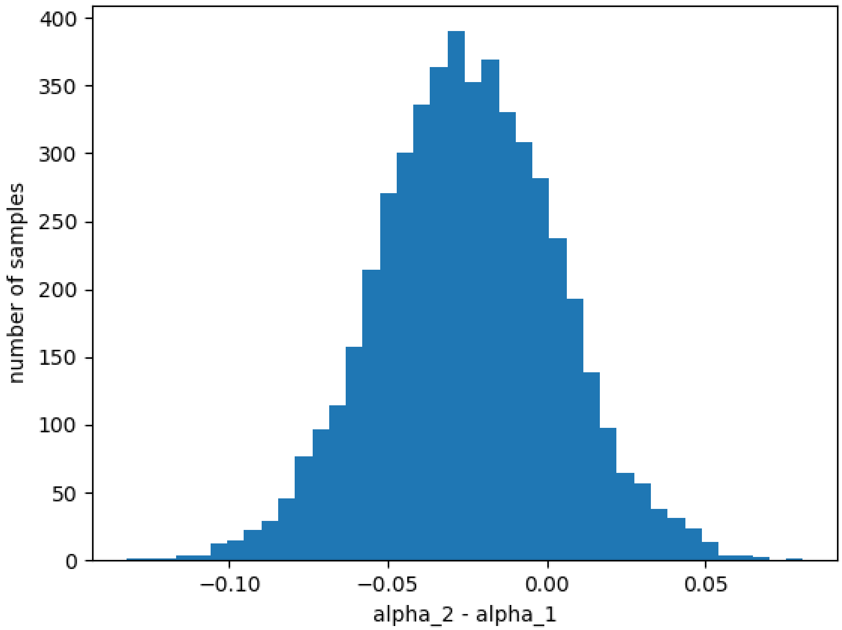

Figure 1: Histogram for Ex3

Notice the formula for Inv-$\chi^2(\nu, s^2)$ is

$$p(\theta) \propto s^\nu \theta^{-(\nu/2+1)} e^{1/(2\theta)}$$

and N-Inv-$\chi^2(\mu, \sigma^2 | \mu_0, \sigma_0^2/\kappa_0; \nu_0, \sigma_0^2)$

$$p(\mu, \sigma^2) \propto \sigma^{-1} (\sigma^2)^{-(\nu_0/2+1)} \exp\left(-\frac{1}{2\sigma^2}[\nu_0\sigma_0^2 + \kappa_0(\mu_0 - \mu)^2]\right)$$

Therefore, the posterior

$$\begin{aligned} p(\mu, \sigma^2 | y) &\propto p(\mu, \sigma^2) \prod_{i=1}^{n} p(y_i | \mu, \sigma^2) \\ &= \sigma^{-1} (\sigma^2)^{-(\nu_0/2+1)} \exp\left(-\frac{1}{2\sigma^2}[\nu_0\sigma_0^2 + \kappa_0(\mu_0 - \mu)^2]\right) \times \\ &\quad (\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}[(n-1)s^2 + n(\bar{y} - \mu)^2]\right) \end{aligned}$$

The rest arithmetic is rather dull, but I guess it's good to do it at least once. First, merge the terms outside the $\exp(\cdots)$:

$$\sigma^{-1} \sigma^{-((\nu_0+n)/2+1)}$$

Let's focus on the terms inside $\exp(\cdots)$, discarding the exponential:

$$
\begin{aligned}
& -\frac{1}{2\sigma^2}[\nu_0\sigma_0^2 + \kappa_0(\mu_0 - \mu)^2] - \frac{1}{2\sigma^2}[(n-1)s^2 + n(\bar{y} - \mu)^2] \\
&= -\frac{1}{2\sigma^2}\left(\nu_0\sigma_0^2 + (n-1)s^2 + \kappa_0(\mu_0 - \mu)^2 + n(\bar{y} - \mu)^2\right) \\
&= -\frac{1}{2\sigma^2}\left(\nu_0\sigma_0^2 + (n-1)s^2 + \kappa_0\mu_0^2 + n\bar{y}^2 + (\kappa_0 + n)\mu^2 - 2(\kappa_0\mu_0 + n\bar{y})\mu\right) \\
&= -\frac{1}{2\sigma^2}\left(\nu_0\sigma_0^2 + (n-1)s^2 + \kappa_0\mu_0^2 + n\bar{y}^2 + (\kappa_0 + n)\left(\mu - \frac{\kappa_0\mu_0 + n\bar{y}}{\kappa_0 + n}\right)^2 - \frac{(\kappa_0\mu_0 + n\bar{y})^2}{\kappa_0 + n}\right) \\
&= -\frac{1}{2\sigma^2}\left(\nu_0\sigma_0^2 + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_0 + n}(\bar{y} - \mu_0)^2 + (\kappa_0 + n)\left(\mu - \frac{\kappa_0\mu_0 + n\bar{y}}{\kappa_0 + n}\right)^2\right)
\end{aligned}
$$

Let

$$
\begin{aligned}
\mu_n &= \frac{\kappa_0}{\kappa_0 + n}\mu_0 + \frac{n}{\kappa_0 + n}\bar{y} \\
\kappa_n &= \kappa_0 + n \\
\nu_n &= \nu_0 + n \\
\nu_n\sigma_n^2 &= \nu_0\sigma_0^2 + (n-1)s^2 + \frac{\kappa_0 n}{\kappa_0 + n}(\bar{y} - \mu_0)^2
\end{aligned}
$$

, then

$$
p(\mu, \sigma^2 | y) = \textbf{N-Inv-}\chi^2(\mu, \sigma^2 | \mu_n, \sigma_n^2/\kappa_n; \nu_n, \sigma_n^2)
$$