STAT 572
Homework 2
Shen Qu

1. If $X_1, \dots, X_n \sim$ Bernoulli(θ), then the distribution of the sum $Z = \sum_{i=1}^n X_i$ has the Binomial(n, θ) distribution. Show that with n fixed, the Binomial(n, θ) distributions form a one-parameter exponential family.

When n fixed,

$$
\begin{aligned}
P_\theta(Z=z) &= \binom{n}{z}\theta^z(1-\theta)^{n-z}\mathbf{1}_{\{z\in(0,1,2,..)\}} \\
&= \exp[\underbrace{\ln(\frac{\theta}{1-\theta})}_{\varphi(\theta)}\underbrace{z}_{t(z)} - n\underbrace{(-\ln(1-\theta))}_{\kappa(\theta)}]\underbrace{\binom{n}{z}\mathbf{1}_{\{z\in(0,1,2,..)\}}}_{h(z)}
\end{aligned}
$$

$$
\begin{aligned}
t(z) &= z \\
\varphi(\theta) &= \ln(\frac{\theta}{1-\theta}) \\
\kappa(\theta) &= -\ln(1-\theta) \\
h(z) &= \binom{n}{z}\mathbf{1}_{\{z\in(0,1,2,..)\}}.
\end{aligned}
$$

Hence, the Binomial(n, θ) distributions form a one-parameter exponential family.

2. Consider the family of generating distributions $\{\text{Poisson}(\theta) : \theta > 0\}$ for the random sample $Y_1, \dots, Y_n$ together with the conjugate family of priors $\{\text{Gamma}(a, b) : a, b > 0\}$ for θ.

- Verify that the family of posteriors for $\theta|X = x$, can be expressed in the same form as Equation (3) of Section 3 in the *Exponential Families and Conjugacy* class notes.

$$
\begin{aligned}
P(y_{1:n}|\theta) &= \theta^{\sum_i^n y_i}e^{-n\theta}(\prod_{i=1}^n \frac{1}{y_i!}\mathbf{1}_{\{y\in(0,1,2,..)\}}) \\
&= \exp[\underbrace{\ln(\theta)}_{\varphi(\theta)(\theta)}\underbrace{\sum_i^n y_i}_{t(y)} - n\underbrace{\theta}_{\kappa(\theta)}]\underbrace{\prod_{i=1}^n \frac{1}{y_i!}\mathbf{1}_{\{y\in(0,1,2,..)\}}}_{h(y)}
\end{aligned}
$$

Assuming the conjugate priors has pdf of the form:

$$
\begin{aligned}
p_{n_o,t_0}(\theta) &\propto \exp[n_0t_0\ln(\theta) - n_0\theta]\mathbf{1}_{\{\theta\in\Theta\}} \propto \theta^{n_0t_0+1-1}e^{n_0\theta}\mathbf{1}_{\{\theta\in\Theta\}} \\
&\sim \text{Gamma}(n_0t_0+1, n_0)
\end{aligned}
$$

where $n_0 > 0$ and $t_0 \in \mathbb{R}$ are values for which $p_{n_0,t_0}(\theta)$ can be normalized.

The posteriors of the form

$$\begin{aligned}
p(\theta|y) &\propto \exp[\ln(\theta)\sum_i^n y_i - n\theta]h(y)\cdot\exp[n_0 t_0 \ln(\theta) - n_0\theta]\mathbf{1}_{\{\theta\in\Theta\}}\\
&\propto \exp[(t(y)+n_0t_0)\ln(\theta) - (n+n_0)\theta]\mathbf{1}_{\{\theta\in\Theta\}}\\
&\propto \theta^{t(y)+n_0t_0+1-1}e^{(n+n_0)\theta}\mathbf{1}_{\{\theta\in\Theta\}}\\
&\sim \text{Gamma}(t(y)+n_0t_0+1, n+n_0)
\end{aligned}$$

- Show that what is referred to as $t^\star$ in the class notes can be expressed as a convex combination of the "prior guess" t_0 and the sufficient statistic $t(y)$ that you obtain for this problem.

When we have $n > 1$ observation, $t(y) = \sum \tilde{t}(y_i)$,

$$\begin{aligned}
p(\theta|y_{1:n}) &\propto \exp[(t^\star n^\star \ln(\theta) - n^\star\theta]\mathbf{1}_{\{\theta\in\Theta\}}\\
&\propto \theta^{t^\star n^\star+1-1}e^{n^\star\theta}\mathbf{1}_{\{\theta\in\Theta\}}\\
&\sim \text{Gamma}(t^\star n^\star+1, n^\star)
\end{aligned}$$

where $n^\star = n_0 + n$ and $t^\star = \frac{n_0}{n_0+n}t_0 + \frac{n}{n_0+n}\frac{1}{n}\sum\tilde{t}(y_i)$,

3. Show that for a certain choice of $t(y)$ and $h(y)$, the Gamma(a, b) distributions are in natural form with natural parameter $\theta = (a, b)^T$.

$$\begin{aligned}
p(y|\theta) &= \frac{b^a}{\Gamma a}y^{a-1}e^{-by} = \exp[(a-1)\ln y - by\ln(\frac{\Gamma a}{b^a})]\mathbf{1}_{\{y\in(0,\infty)\}}\\
&= \exp[\underbrace{a\ln y + b(-y)}_{\varphi(\theta)=\begin{bmatrix}a\\b\end{bmatrix};t(y)=\begin{bmatrix}\ln(y)\\-y\end{bmatrix}} \underbrace{-n\frac{1}{n}\ln(\frac{\Gamma a}{b^a})}_{\kappa(\theta)}]\underbrace{y^{-1}\mathbf{1}_{\{y\in(0,\infty)\}}}_{h(y)}
\end{aligned}$$

Hence, letting $\eta_1 = a$, $\eta_2 = b$, we can rewrite the expression above as in natural form as

$$p(y|\theta) = \exp[\underbrace{\eta_1\ln y + \eta_2(-y)}_{\eta=\begin{bmatrix}\eta_1\\\eta_2\end{bmatrix};t(y)=\begin{bmatrix}\ln(y)\\-y\end{bmatrix}} \underbrace{-n\frac{1}{n}\ln(\frac{\Gamma(\eta_1)}{\eta_2^{\eta_1}})}_{A(\eta)}]\underbrace{y^{-1}\mathbf{1}_{\{y\in(0,\infty)\}}}_{h(y)}$$

Should we show that $A(\eta) = \ln(\int_{\mathcal{X}} y^{-1}(\eta_1\ln y + \eta_2(-y))dy)$?

This is the natural form for the Gamma(a, b) distributions, where η corresponds to the canonical parameter.