AMS 206B: Intermediate Bayesian Inference
HOMEWORK # 1 (Review and Conjugate Families)

1. Let $X \sim Exp(\lambda)$ where $E(X) = 1/\lambda$. What is the p.m.f. of $Y = \lfloor X \rfloor$ (the floor of X)? Do you recognize it as a distribution that you have studied in the past?

2. Let X_1 and X_2 be two independent random variables such that $X_i \sim Gamma(a_i, b)$ for any $a_1, a_2, b > 0$. Define $Y = X_1/(X_1 + X_2)$ and $Z = (X_1 + X_2)$.

   (a) Find the joint p.d.f. for Y and Z and show that these two random variables are independent.

   (b) Find the marginal p.d.f. of Z. Do you recognize this p.d.f. as belonging to some family that you know?

   (c) Find the marginal p.d.f. of Y. Do you recognize this p.d.f. as belonging to some family that you know?

   (d) Compute $E(Y^k)$ for any $k > 0$.

   (e) What does this result imply if $a_i = b_i = 1$?

3. Consider three independent random variables X_1, X_2 and X_3 such that $X_i \sim Gamma(a_i, b)$. Let

$$\mathbf{Y} = (Y_1, Y_2, Y_3) = \left(\frac{X_1}{X_1 + X_2 + X_3}, \frac{X_2}{X_1 + X_2 + X_3}, \frac{X_3}{X_1 + X_2 + X_3} \right).$$

   (a) Show that $\mathbf{Y} \sim Dirichlet(a_1, a_2, a_3)$, a Dirichlet distribution.

   (b) How can this result be used to generate random variables according to a Dirichlet distribution? Write a simple function in `R` or `Matlab` (your choice) that takes as inputs n, the number of trivariate vectors to be generated, and $\mathbf{a} = (a1, a2, a3)$ and generates a matrix of size $n \times 3$ whose rows correspond to independent samples from a Dirichlet distribution with parameter $(a1, a2, a3)$.

4. Y follows an inverse Gamma distribution with shape parameter a and scale parameter b ($Y \sim IG(a, b)$) if $Y = 1/X$ with $X \sim Gamma(a, b)$ (assume the Gamma distribution is parameterized so that $E(X) = ab$.

   (a) Find the density of Y.

   (b) Compute $E(Y^k)$. Do you need to impose any constrain on the problem for this expectation to exists?

5. Y follows a lognormal distribution with parameters μ and σ^2 (denotes as $Y \sim LogN(\mu, \sigma^2)$) if $Y = \exp(X)$ where $X \sim N(\mu, \sigma^2)$.

   (a) Find the density of Y.

   (b) Compute the mean and the variance of Y.

6. Let $\mathbf{X} = (X_1, X_2, \ldots, X_p)$ with $\mathbf{X} \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and set $\mathbf{Z}_1 = (X_1, \ldots, X_q)$ and $\mathbf{Z}_2 = (X_{q+1}, \ldots, X_p)$ with $1 < q < p$.

   (a) Show that

   $$\mathbf{Z}_1|\mathbf{Z}_2 \sim N_q\left(\boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}(\mathbf{Z}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Sigma}_{11} - \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right),$$

   where $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_{kl}$ denote the blocks of $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ where the rows correspond to the variables in $\mathbf{Z}_k$ and the columns to the variables in $\mathbf{Z}_l$.

   (b) If you let $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}$, show that the previous expression can also be written as

   $$\mathbf{Z}_1|\mathbf{Z}_2 \sim N_q(\boldsymbol{\mu}_1 - \boldsymbol{\Omega}_{11}^{-1}\boldsymbol{\Omega}_{12}(\mathbf{Z}_2 - \boldsymbol{\mu}_2), \boldsymbol{\Omega}_{11}^{-1}).$$

   (Hint: part (b) is a problem in linear algebra, you will likely need to consult a book for results on inverses and determinant of block-partitioned matrices).

7. Let $Y|X \sim Poisson(X)$ and let $X \sim Exponential(\lambda)$. What is the marginal distribution of Y?

8. Let $X|Y \sim Binomial(Y, \pi)$, and let $Y \sim Poisson(\lambda)$.

   (a) Show that $E(X) = E(E(X|Y)) = \lambda\pi$, and that $V(X) = E(V(X|Y)) + V(E(X|Y)) = \lambda\pi$.

   (b) Show that $X \sim Poisson(\lambda\pi)$ and $Y - X|X \sim Poisson[\lambda(1-\pi)]$.

9. Let U, V and W be independent normal random variables with $U \sim N(\mu, 1)$, $V \sim N(\mu, 1)$ and $W \sim N(0, 1)$. Let $X_1 = U + W$ and $X_2 = V + W$, i.e., a common W contaminates U and V. Show that $X = (X_1, X_2)$ has a density in the exponential family.

10. Assume that $W \sim N(\mu, 1)$. Suppose that W is physically observed only when its value is inside the interval $[a, b]$. Then, the variable X that is truly observed is a truncated normal random variable in the interval $[a, b]$. Show that $f(x|\mu)$ is in the exponential family.

11. Let $X \sim N(\mu, \sigma^2)$ with μ and σ unknown. Show that $f(x|\boldsymbol{\theta})$ with $\boldsymbol{\theta} = (\mu, \sigma)$ is in the exponential family.

12. Show that if $X \sim Exponential(\beta)$, then

    (a) $Y = X^{1/\gamma}$ has a Weibull distribution with parameters γ and β with $\gamma > 0$ a constant.

    (b) $Y = (2X/\beta)^{1/2}$ has the Rayleigh distribution.

    For both parts, derive the form of the p.d.f., verify that is a p.d.f., and calculate the mean and the variance.

13. (Robert) Let $x \sim N(\theta, \sigma^2)$, $y \sim N(\rho x, \sigma^2)$ with ρ known. Assume a prior of the form $\pi(\theta, \sigma^2) = 1/\sigma^2$. Find the predictive density of y given x.

14. (Robert) If $y \sim Binomial(n, \theta)$ and $x \sim Binomial(m, \theta)$, and $\theta \sim Beta(\alpha, \beta)$. Find the predictive distribution of y given x.

15. (Robert) Give the posterior and the marginal distributions in the following cases:

    (a) $x|\sigma \sim N(0, \sigma^2)$ and $1/\sigma^2 \sim Gamma(1, 2)$.

    (b) $x|p \sim Negative - Binomial(10, p)$, $p \sim Beta(1/2, 1/2)$.

16. Let $X|\theta \sim N(\theta, 1)$. Suppose your prior is such that θ is $N(\mu, 1)$ or $N(-\mu, 1)$ with equal probabilities. Write the prior distribution and find the posterior after observing $X = x$. Show that

$$E(\theta|x) = \frac{x}{2} + \frac{\mu}{2}\frac{1 - \exp(-\mu x)}{1 + \exp(-\mu x)}$$

    and draw a graph of $E(\theta|x)$ as a function of x.

17. Let $X = (X_1, \ldots, X_k)'$ be a random vector with a multinomial distribution with index n and probabilities $\theta_1, \ldots, \theta_k$ such that $\sum_{i=1}^k X_i = n$ and $\sum_{i=1}^k \theta_i = 1$.

    (a) Show that the p.d.f. of X (the multinomial p.d.f.) is in the exponential family. Find the natural parameterization and the natural parameters.

    (b) Find a conjugate family of distributions using the natural parameterization.

18. Let (X_1, X_2, X_3), be a vector with pmf

$$\frac{n!}{\prod_{i=1}^3 x_i!} \prod_{i=1}^3 p_i^{x_i}, \quad x_i \geq 0, \quad x_1 + x_2 + x_3 = n,$$

    where $p_1 = \theta^2$, $p_2 = 2\theta(1-\theta)$, $p_3 = (1-\theta)^2$, and $0 \leq \theta \leq 1$.

    (a) Verify if this distribution belongs to the exponential family with K parameters. If this is true, what is K?

    (b) Obtain a sufficient statistic for θ.