APPM 4/5720: Computational Bayesian Statistics, Spring 2018

Solutions to Problem Set Two

1. Let $\vec{X}_1, \vec{X}_2, \ldots, \vec{X}_m$ be a random sample of m vectors each of length k from the multinomial distribution with parameters n and $\theta_1, \theta_2, \ldots, \theta_k$.

   The joint pdf is

$$\begin{aligned} f(\vec{\vec{x}}|\theta) &= \textstyle\prod_{i=1}^m f(\vec{x}_i|\theta) \\ &= \textstyle\prod_{i=1}^m \frac{n!}{x_{i1}!x_{i2}!\cdots x_{ik}!}\theta_1^{x_{i1}}\theta_2^{x_{i2}}\cdots\theta_k^{x_{ik}} \\ &= \left(\textstyle\prod_{i=1}^m \frac{n!}{x_{i1}!x_{i2}!\cdots x_{ik}!}\right)\cdot \theta_1^{\sum_{i=1}^m x_{i1}}\theta_2^{\sum_{i=1}^m x_{i2}}\cdots\theta_k^{\sum_{i=1}^m x_{ik}}. \end{aligned}$$

   The posterior is

$$\begin{aligned} f(\theta|\vec{\vec{x}}) &\propto f(\vec{\vec{x}}|\theta)\cdot f(\theta) \\ &= \theta_1^{\sum_{i=1}^m x_{i1}}\theta_2^{\sum_{i=1}^m x_{i2}}\cdots\theta_k^{\sum_{i=1}^m x_{ik}}\cdot\theta_1^{\alpha_1-1}\theta_2^{\alpha_2-1}\cdots\theta_k^{\alpha_k-1} \\ &= \theta_1^{\sum_{i=1}^m x_{i1}+\alpha_1-1}\theta_2^{\sum_{i=1}^m x_{i2}+\alpha_2-1}\cdots\theta_k^{\sum_{i=1}^m x_{ik}+\alpha_k-1} \end{aligned}$$

   which is another Dirichlet distribution with parameters

$$\alpha_j^* = \sum_{i=1}^m x_{ij} + \alpha_j$$

   for $j = 1, 2, \ldots, k$.

   Not only is this a conjugate prior but it is a natural conjugate prior as well!

---

2. (a) If Y_i and Y_j are independent, they must have covariance 0. We will show that their covariance is greater than 0, so they must be dependent. (Note: If we computed their covariance and got zero, we would not be able to make a conclusion about dependence/independence!)

   For $i \neq j$,

$$\begin{aligned} Cov(Y_i, Y_j) &= Cov(X_i + Z, X_j + Z) \\ &= Cov(X_i, X_j) + Cov(X_i, Z) + Cov(Z, X_j) + Cov(Z, Z). \end{aligned}$$

   The first three terms are zero by independence of X_i with X_j and of each with Z. Thus,

$$Cov(Y_i, Y_j) = Cov(Z, Z) = Var(Z) > 0.$$

   (The only way that $Var(Z)$ could equal 0 is if it were a constant with probability 1.)

   (b) I'm going to do this in all discrete notation. That is, I'm going to assume that the X_i and Z are discrete. You may assume either or both are continuous or make no such assumptions and use Riemann-Stiltjes integrals!

   Let $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ be any permutation of $\{1, 2, \ldots, n\}$. Then,

$$P(Y_1 = y_{\pi_1}, Y_2 = y_{\pi_2}, \ldots, Y_n = y_{\pi_n})$$

$$= \textstyle\sum_z P(Y_1 = y_{\pi_1}, Y_2 = y_{\pi_2}, \ldots, Y_n = y_{\pi_n}|Z = z)\cdot P(Z = z)$$

$$= \sum_z P(X_1 + Z = y_{\pi_1}, X_2 + Z = y_{\pi_2}, \ldots, X_n + Z = y_{\pi_n} | Z = z) \cdot P(Z = z)$$

$$= \sum_z P(X_1 = y_{\pi_1} - z, X_2 = y_{\pi_2} - z, \ldots, X_n = y_{\pi_n} - z | Z = z) \cdot P(Z = z)$$

By independence of the X_i from Z, this is

$$\sum_z P(X_1 = y_{\pi_1} - z, X_2 = y_{\pi_2} - z, \ldots, X_n = y_{\pi_n} - z) \cdot P(Z = z)$$

Since the X_i are iid, we have that

$$P(X_1 = y_{\pi_1} - z, X_2 = y_{\pi_2} - z, \ldots, X_n = y_{\pi_n} - z) = P(X_1 = y_{\pi_1} - z) \cdot P(X_1 = y_{\pi_2} - z) \cdots P(X_1 = y_{\pi_n} - z)$$

which is clearly symmetric in the arguments $y_{\pi_1}, y_{\pi_2}, \ldots, y_{\pi_n}$.
Unraveling the steps, we have that

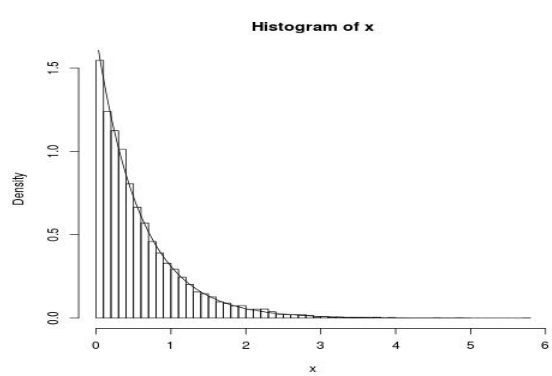
$$P(Y_1 = y_{\pi_1}, Y_2 = y_{\pi_2}, \ldots, Y_n = y_{\pi_n}) = P(Y_1 = y_1, Y_2 = y_2, \ldots, Y_n = y_n).$$

Since the permutation was arbitrary, $Y_1, Y_2, \ldots, Y_n$ are exchangeable.

---

3. Here is some R-code:

```
> lambda<-1.7
> u<-runif(10000)
> x<-(-1/lambda)*log(1-u)
> min(x)
[1] 5.788838e-05
> max(x)
[1] 5.728264
> br<-seq(0,5.8,0.1)
> hist(x,prob=T,breaks=br)
> y<-seq(0,5,0.001)
> f<-lambda*exp(-lambda*y)
> lines(y,f)
```



---

4. In the proof of deFinetti's Theorem, we saw that the cdf F used in the statement of the theorem was the limiting cdf of $\overline{X}_n$. So, we must figure out what prior to use to get the relationship

$$P(X_1 = 1, X_2 = 1, \ldots, X_k = 1, X_{k+1} = 0, \ldots, X_{n=1}) = \int_0^1 \theta^k (1-\theta)^k \, dF(\theta).$$

Note that the probability on the left-hand side is

$$\frac{R_0}{R_0+W_0} \cdot \frac{R_0+c}{R_0+W_0+c} \cdots \frac{R_0+c(k-1)}{R_0+W_0+c(k-1)} \cdot \frac{W_0}{R_0+W_0+ck} \cdot \cdot \frac{W_0+c}{R_0+W_0+c(k+1)} \cdots \frac{W_0+c(n-k-1)}{R_0+W_0+c(n-k-1)}.$$

Inspired by a similar computation in class, we are going to try to write this as a Beta function. For that we will need factorials. In order to see the terms in the numerators and denominators as increasing by 1 unit like a factorial, we will multiply the top and bottom of each fraction by $1/c$. This gives us

$$\frac{R_0'}{R_0'+W_0'} \cdot \frac{R_0'+1}{R_0'+W_0'+1} \cdots \frac{R_0'+(k-1)}{R_0'+W_0'+(k-1)} \cdot \frac{W_0'}{R_0'+W_0'+k} \cdot \cdot \frac{W_0'+1}{R_0'+W_0'+(k+1)} \cdots \frac{W_0'+(n-k-1)}{R_0'+W_0'+(n-k-1)},$$

where $R_0' = R_0/c$ and $W_0' = W_0/c$.

Now this looks exactly like the problem we did in class with R_0 replaced by R_0' and W_0 replaced by W_0'! Following that example, we see that

$$\overline{X}_n \xrightarrow{d} X$$

where $X \sim Beta(R_0', W_0') = \boxed{Beta(R_0/c, W_0/c)}$.

---

5. (a) No, a conjugate prior does not always exist. An example is the two-parameter Weibull distribution. (This is on your table of distributions with three parameters. Set $\alpha = 0$ to get the two-parameter version.)

   A proof that this does not have a conjugate prior is given in

   Soland, R. (1969), Bayesian Analysis of the Weibull Process With Unknown Scale and Shape Parameters, IEEE Transactions on Reliability Analysis, 18, 181-184.

   (b) For a one-parameter exponential family distribution, the joint pdf has the form

   $$f(\vec{x}|\theta) = a(\theta)b(\vec{x})\exp[c(\theta)d(\vec{x})].$$

   As a function of θ, this is proportional to

   $$a(\theta)\exp[c(\theta)d(\vec{x})].$$

   The natural conjugate prior would be

   $$f(\theta) \propto a(\theta)\exp[\alpha c(\theta)]$$

   for some hyperparameter α. HOWEVER, this is not guaranteed to be a proper prior. (i.e. It is not guaranteed to be integrable with respect to θ.) For more on conditions that will give a proper conjugate prior, see

   Diaconis, P., and Ylvisaker, D. (1979). Conjugate priors for exponential families. Annals of Statistics, 7, 269-281.