APPM 4/5720: Computational Bayesian Statistics, Spring 2018

Solutions to Problem Set One

1. The likelihood is

$$\begin{aligned} f(\vec{x}|\mu) &\stackrel{iid}{=} \prod_{i=1}^n f(x_i|\mu) \\ &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \\ &= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2}. \end{aligned}$$

The prior is

$$f(\mu) = \frac{1}{\sqrt{2\pi\sigma_0^2}} e^{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2}.$$

The posterior is

$$\begin{aligned} f(\mu|\vec{x}) &\propto f(\vec{x}|\mu) \cdot f(\mu) \\ &\propto e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2} \cdot e^{-\frac{1}{2\sigma_0^2}(\mu-\mu_0)^2} \\ &= \exp\left[-\tfrac{1}{2\sigma^2}(\textstyle\sum x_i^2 - 2\mu\sum x_i + n\mu^2) - \tfrac{1}{2\sigma_0^2}(\mu^2 - 2\mu_0\mu + \mu_0^2)\right] \\ &\propto \exp\left[-\tfrac{1}{2\sigma^2}(-2\mu\textstyle\sum x_i + n\mu^2) - \tfrac{1}{2\sigma_0^2}(\mu^2 - 2\mu_0\mu)\right] \\ &= \exp\left[-\tfrac{1}{2}\underbrace{\frac{n\sigma_0^2+\sigma^2}{\sigma^2\sigma_0^2}}_{a}\mu^2 + \underbrace{\left(\frac{\sigma_0^2\sum x_i + \mu_0\sigma^2}{\sigma^2\sigma_0^2}\right)}_{b}\mu\right] \\ &= \exp\left[-\tfrac{1}{2}a\mu^2 + b\mu\right] = \exp\left[-\tfrac{1}{2}a(\mu^2 - \tfrac{2b}{a}\mu)\right]. \end{aligned}$$

Completing the square gives us

$$\begin{aligned} f(\mu|\vec{x}) &\propto \exp\left[-\tfrac{1}{2}a(\mu - \tfrac{b}{a})^2 + \tfrac{1}{2}b^2\right] \\ &\propto \exp\left[-\tfrac{1}{2}a(\mu - \tfrac{b}{a})^2\right] \end{aligned}$$

So, we see that the posterior distribution for μ given $\vec{x}$ is $N(b/a, 1/a)$. Specifically,

$$\boxed{\mu|\vec{x} \sim N\left(\frac{\mu_0\sigma^2 + \sigma_0^2\sum x_i}{\sigma^2 + n\sigma_0^2}, \frac{\sigma^2\sigma_0^2}{\sigma^2 + n\sigma_0^2}\right)}$$

Note that we can write $\sum x_i$ as $n\overline{x}$. The mean for the posterior distribution is then

$$\frac{n\sigma_0^2}{\sigma^2 + n\sigma_0^2}\ \overline{x} + \frac{\sigma^2}{\sigma^2 + n\sigma_0^2}\ \mu_0$$

which is a weighted average of the sample mean and the mean of the prior for μ.

Since the posterior distribution is normal, the prior used was a conjugate prior for the model.

---

2. (a)

$$f_X(x) = \frac{1}{\Gamma(\alpha)} \beta^\alpha x^{\alpha-1} e^{-\beta x} \, I_{(0,\infty)}(x)$$

$$y = g(x) = 1/y \qquad \Rightarrow \qquad x = g^{-1}(y) = 1/y$$

So,

$$\begin{aligned} f_Y(y) &= f_X(g^{-1}(y)) \cdot \left| \tfrac{d}{dy} g^{-1}(y) \right| \\ &= \tfrac{1}{\Gamma(\alpha)} \beta^\alpha (1/y)^{\alpha-1} e^{-\beta/y} \, I_{(0,\infty)}(1/y) \cdot \left| -\tfrac{1}{y^2} \right| \\ &= \boxed{\tfrac{1}{\Gamma(\alpha)} \beta^\alpha y^{-\alpha-1} e^{-\beta/y} \, I_{(0,\infty)}(y)} \end{aligned}$$

Let us denote this as $Y \sim IG(\alpha, \beta)$.

(b) The likelihood is

$$\begin{aligned} f(\vec{x}|\mu) &\stackrel{iid}{=} \textstyle\prod_{i=1}^n f(x_i|\mu) \\ &= \textstyle\prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2} \\ &= (2\pi\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2}. \end{aligned}$$

The prior is

$$f(\sigma^2) = \frac{1}{\Gamma(\alpha)} \beta^\alpha (\sigma^2)^{-\alpha-1} e^{-\beta/\sigma^2} \, I_{(0,\infty)}(\sigma^2).$$

The posterior is

$$\begin{aligned} f(\sigma^2|\vec{x}) &\propto f(\vec{x}|\sigma^2) \cdot f(\sigma^2) \\ &\propto (\sigma^2)^{-n/2} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\mu)^2} \cdot (\sigma^2)^{-\alpha-1} e^{-\beta/\sigma^2} \, I_{(0,\infty)}(\sigma^2) \\ &= (\sigma^2)^{-n/2-\alpha-1} \exp\left[-\tfrac{1}{\sigma^2} \left(\tfrac{1}{2} \textstyle\sum_{i=1}^n (x_i-\mu)^2 + \beta \right) \right] \end{aligned}$$

This is another inverse gamma distribution. Thus, we see that the prior was a conjugate prior and specifically that

$$\boxed{\sigma^2|\vec{x} \sim IG\left(\alpha + \frac{n}{2}, \frac{1}{2}\sum_{i=1}^n (x_i-\mu)^2 + \beta \right).}$$

---

3. The likelihood is

$$\begin{aligned} f(\vec{x}|\lambda) &\stackrel{iid}{=} \textstyle\prod_{i=1}^n f(x_i|\lambda) \\ &= \textstyle\prod_{i=1}^n \lambda e^{-\lambda x_x} \, I_{(0,\infty)}(x_i) \\ &= \lambda^n e^{-\lambda \sum x_i} \textstyle\prod_{i=1}^n I_{(0\infty)}(x_i). \end{aligned}$$

If we look at this as a function of λ, we see a gamma looking pdf (in λ).

Let's try a gamma prior. Specifically, $\lambda \sim \Gamma(\alpha, \beta)$ for some hyperparameters $\alpha, \beta > 0$. Then

$$f(\lambda) = \frac{1}{\Gamma(\alpha)} \beta^\alpha \lambda^{\alpha-1} e^{-\beta\lambda} \, I_{(0,\infty)}(\lambda).$$

The posterior is then

$$\begin{aligned} f(\lambda|\vec{x}) &\propto f(\vec{x}|\lambda)\cdot f(\lambda) \\ &\propto \lambda^n e^{-\lambda\sum x_i}\lambda^{\alpha-1}e^{-\beta\lambda}\, I_{(0,\infty)}(\lambda) \\ &= \lambda^{\alpha+n-1}e^{-(\sum x_i+\beta)\lambda}\, I_{(0,\infty)}(\lambda) \end{aligned}$$

Thus, we see that the posterior distribution is

$$\boxed{\lambda|\vec{x} \sim \Gamma(\alpha+n, \beta+\sum x_i)}$$

which is in the same family of distributions as the prior, as desired!

---

4. There were 3 questions posed in this problem. Although they were not labeled (a), (b), and (c), I will label the solutions this way for clarity.

   Let

$$\Theta = \begin{cases} 1 & , \text{ if Judy is a heterozygote} \\ 0 & , \text{ otherwise} \end{cases}$$

   For $i = 1, 2, \ldots, n$, let

$$X_i = \begin{cases} 1 & , \text{ if child } i \text{ has brown eyes} \\ 0 & , \text{ otherwise} \end{cases}$$

   Note that the proportion of blue-eyed individuals in the population is p^2, the proportion of heterozygotes is $2p(1-p)$, and the proportion with brown eyes that are not heterozygotes (XX) is $1 - p^2 - 2p(1-p) = (1-p)^2$.

   (a) Show that among brown-eyed children of brown-eyed parents, the expected proportion of heterozygotes is $2p/(1+2p)$.

   We want to find

$$P(\text{heterozygote}|\text{brown eyes, parents brown eyes}).$$

   Temporary notation, for this part (a) only, will be $P(H|B,P)$.

$$P(H|B,P) = \frac{P(H,B,P)}{P(B,P)} = \frac{P(H,P)}{P(B,P)}$$

   The B was dropped in the numerator because it is implied by the H. That is, if a person is a heterozygote, he/she definitely has brown eyes.

   Now,

$$\begin{aligned} P(H,P) &= P(H|\underbrace{XX,XX}_{\text{parents}})P(XX,XX) + 2P(H|\underbrace{XX,Xx}_{\text{parents}})P(XX,Xx) + P(H|\underbrace{Xx,Xx}_{\text{parents}})P(Xx,Xx) \\ &= 0\cdot(1-p)^2\cdot(1-p)^2 + 2\cdot\tfrac{1}{2}\cdot(1-p)^2\cdot 2p(1-p) + \tfrac{1}{2}\cdot 2p(1-p)\cdot 2p(1-p) \\ &= 2p(1-p)^3 + 2p^2(1-p)^2. \end{aligned}$$

Also,

$$\begin{aligned}
P(B,P) &= P(B|\underbrace{XX,XX}_{\text{parents}})P(XX,XX) + 2P(B|\underbrace{XX,Xx}_{\text{parents}})P(XX,Xx) + P(B|\underbrace{Xx,Xx}_{\text{parents}})P(Xx,Xx) \\
&= 1\cdot(1-p)^2\cdot(1-p)^2 + 2\cdot 1\cdot(1-p)^2\cdot 2p(1-p) + \tfrac{3}{4}\cdot 2p(1-p)\cdot 2p(1-p) \\
&= (1-p)^4 + 4p(1-p)^3 + 3p^2(1-p)^2.
\end{aligned}$$

So,

$$\begin{aligned}
P(H|BP) &= \tfrac{2p(1-p)^3+2p^2(1-p)^2}{(1-p)^4+4p(1-p)^3+3p^2(1-p)^2} \\
&= \tfrac{2p(1-p)^2[(1-p)+p]}{(1-p)^2[(1-p)^2+4p(1-p)+3p^2]} \\
&= \tfrac{2p}{1+2p} \quad \checkmark
\end{aligned}$$

(b) Suppose that Judy, a brown-eyed child of brown-eyed parents, married a heterozygote, and they have n children, all brown-eyed. Find the posterior probability that Judy is a heterozygote.

Since Judy is a brown-eyed child of brown-eyed parents, the prior (before children) probabilities of her being a heterozygote or not are given by

$$P(\Theta=1) = \frac{2p}{1+2p} \qquad \text{and} \qquad P(\Theta=1) = 1 - \frac{2p}{1+2p} = \frac{1}{1+2p}.$$

For $i = 1, 2, \ldots, n$,

$$P(X_i = 1|\text{Judy is Xx, Judy's husband is Xx}) = 1 - P(X_i = 0|\text{Judy is Xx, Judy's husband is Xx})$$

$$= 1 - P(i\text{th child inherits } x \text{ and } x) = 1 - \frac{1}{2}\cdot\frac{1}{2} = \frac{3}{4}.$$

Generalizing, (and noting that the genes that one child inherits are independent of the genes another child inherits)

$$P(X_1 = 1, X_2 = 1, \ldots, X_n = 1|\underbrace{\text{Judy is Xx}}_{\Theta=1}, \underbrace{\text{Judy's husband is Xx}}_{\text{call this } Y=1}) = \left(\frac{3}{4}\right)^n.$$

Since Judy definitely marries a heterozygote, that information will stay constant through out the problem. It will be on the right side of any and every conditional probability statement here so we will not bother to write it. Using this convention, we know that

$$P(X_1 = 1, X_2 = 1, \ldots, X_n = 1|\Theta = 1) = \left(\frac{3}{4}\right)^n.$$

We also know that, because Judy has brown eyes, if she is not a heterozygote, she must be XX. Since her husband is Xx, each of their children will be either XX or Xx, and so all of them will be brown-eyed. That is,

$$P(X_1 = 1, X_2 = 1, \ldots, X_n = 1|\Theta = 0) = 1.$$

Now, we compute the posterior probability that Judy is a heterozygote.

$$\begin{aligned}
P(\Theta = 1|X_1 = 1, \ldots, X_n = 1) &\overset{notation}{=} P(\Theta = 1|\vec{X} = \vec{1}) \\
&= \tfrac{P(\vec{X}=\vec{1}|\theta=1)P(\Theta=1)}{P(\vec{X}=\vec{1}|\Theta=1)P(\Theta=1)+P(\vec{X}=\vec{1}|\Theta=0)P(\Theta=0)} \\
&= \boxed{\frac{\left(\frac{3}{4}\right)^n\cdot\frac{2p}{1+2p}}{\left(\frac{3}{4}\right)^n\cdot\frac{2p}{1+2p} + 1\cdot\frac{1}{1+2p}}}.
\end{aligned}$$

(c) Suppose that Judy, a brown-eyed child of brown-eyed parents, marries a heterozygote, and they have n children, all brown-eyed. Find the probability that her first grandchild has blue eyes.

I am going to be pretty abusive with notation here because there is a lot to write. Let G, C, and S be the statuses of the grandchild, child, and spouse of the child, respectively. **Assume all probabilities that I write are conditional on any and all given information.** Then,

$$\begin{aligned} P(G=xx) &= P(G=xx|C=Xx,S=Xx)\cdot P(C=Xx,S=Xx) \\ &\quad +P(G=xx|C=Xx,S=xx)\cdot P(C=Xx,S=xx) \\ &\quad +P(G=xx|C=Xx,S=Xx)\cdot P(C=xx,S=Xx) \\ &\quad +P(G=xx|C=xx,S=xx)\cdot P(C=xx,S=xx) \\ &= \tfrac{1}{4}\cdot P(C=Xx,S=Xx)+\tfrac{3}{4}\cdot P(C=Xx,S=xx) \\ &\quad +\tfrac{3}{4}\cdot P(C=xx,S=Xx)+1\cdot P(C=xx,S=xx) \end{aligned}$$

Assuming independence of the eye color of the child and his/her spouse (Who knows, maybe they met at a brown-eye club?), and using the population proportions for eye color for the spouse, we have

$$\begin{aligned} P(G=xx) &= \tfrac{1}{4}\cdot P(C=Xx)\cdot 2p(1-p)+\tfrac{3}{4}\cdot P(C=Xx)\cdot p^2 \\ &\quad +\tfrac{3}{4}\cdot P(C=xx)\cdot 2p(1-p)+1\cdot P(C=xx)\cdot p^2 \end{aligned}$$

Now, letting J be Judy's status, (and recalling (*) that her husband is a heterozygote and that we know all of her children have brown eyes)

$$\begin{aligned} P(C=Xx) &= P(C=Xx,\Theta=1)+P(C=Xx,\theta=0) \\ &= P(C=Xx,J=Xx)+P(C=Xx,J=XX) \\ &= P(C=Xx|J=Xx)\cdot P(J=Xx)+P(C=Xx|J=XX)\cdot P(J=XX) \\ &= \underbrace{\frac{2}{3}}_{(*)}\cdot q+\tfrac{1}{2}\cdot(1-q) \end{aligned}$$

where $q=P(\Theta=1|X_1=1,\ldots,X_n=1)$ is the probability computed above in part (b). Also,

$$\begin{aligned} P(C=xx) &= P(C=xx,\Theta=1)+P(C=xx,\theta=0) \\ &= P(C=xx,J=Xx)+P(C=xx,J=XX) \\ &= P(C=xx|J=Xx)\cdot P(J=Xx)+P(C=xx|J=XX)\cdot P(J=XX) \\ &= \tfrac{1}{4}\cdot q+0\cdot(1-q)=\tfrac{1}{4}\cdot q \end{aligned}$$

Putting it all together, we have

$$
\begin{aligned}
P(G = xx) &= \tfrac{1}{4} \cdot \left[\tfrac{2}{3} \cdot q + \tfrac{1}{2} \cdot (1-q)\right] \cdot 2p(1-p) \\
&+ \tfrac{3}{4} \cdot \left[\tfrac{2}{3} \cdot q + \tfrac{1}{2} \cdot (1-q)\right] \cdot p^2 \\
&+ \tfrac{3}{4}\tfrac{1}{4} \cdot q \cdot 2p(1-p) \\
&+ 1 \cdot \tfrac{1}{4} \cdot q \cdot p^2
\end{aligned}
$$

Yuck!

---

5. (a) I'll show it for continuous random variables only. Suppose that $X_1, X_2, \ldots, X_n$ are iid with pdf f. Then the joint pdf is $f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = f(x_1) \cdot f(x_2) \cdots f(x_n)$ Let $\pi = (\pi_1, \pi_2, \ldots, \pi_n)$ denote a permutation of the indices $\{1, 2, \ldots, n\}$. Then

$$
\begin{aligned}
f_{X_1,X_2,\ldots,X_n}(x_{\pi_1}, x_{\pi_2}, \ldots, x_{\pi_n}) &\overset{iid}{=} f(x_{\pi_1}) \cdot f(x_{\pi_2}) \cdots f(x_{\pi_n}) \\
&= f(x_1) \cdot f(x_2) \cdots f(x_n)
\end{aligned}
$$

since $f(x_{\pi_1}) \cdot f(x_{\pi_2}) \cdots f(x_{\pi_n})$ is just f evaluated at all of the x_i and multiplied in some order.
Thus, we see that

$$
f_{X_1,X_2,\ldots,X_n}(x_{\pi_1}, x_{\pi_2}, \ldots, x_{\pi_n}) = f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n)
$$

and so $X_1, X_2, \ldots, X_n$ are exchangeable random variables.

(b) Consider an urn containing 1 red ball and 2 white balls. Draw balls one at a time, without replacement, and note the color. Let

$$
X_i = \begin{cases} 1 & , \text{ if the } i\text{the ball drawn is red} \\ 0 & , \text{ otherwise.} \end{cases}
$$

Then X_1 and X_2 are exchangeable since

$$
P(X_1 = 1, X_2 = 0, X_3 = 0) == \frac{1}{3} \cdot 1 \cdot 1 = \frac{1}{3}
$$

$$
P(X_1 = 0, X_2 = 1, X_3 = 0) == \frac{2}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{3}
$$

and

$$
P(X_1 = 0, X_2 = 0, X_3 = 1) == \frac{2}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{3}.
$$

For any other x_1, x_2, x_3, $P(X_1 = x_1, X_2 = x_2, X_3 = x_3) = 0$.
So, $P(X_1 = x_1, X_2 = x_2, X_3 = x_3)$ is invariant under any permutation of the arguments and therefore X_1, X_2, X_3 are exchangeable.
However, they are clearly not independent!