Introduction to Bayesian Data Analysis

Tutorial 2 Solutions

(1) ‘

1.3. Note: We will use "Xx" to indicate all heterozygotes (written as "Xx or xX" in the Exercise).

$$\begin{aligned}
&\Pr(\text{child is Xx}\,|\,\text{child has brown eyes \& parents have brown eyes}) \\
&= \frac{0\cdot(1-p)^4 + \frac{1}{2}\cdot 4p(1-p)^3 + \frac{1}{2}\cdot 4p^2(1-p)^2}{1\cdot(1-p)^4 + 1\cdot 4p(1-p)^3 + \frac{3}{4}\cdot 4p^2(1-p)^2} \\
&= \frac{2p(1-p)+2p^2}{(1-p)^2+4p(1-p)+3p^2} \\
&= \frac{2p}{1+2p}.
\end{aligned}$$

To figure out the probability that Judy is a heterozygote, use the above posterior probability as a prior probability for a new calculation that includes the additional information that her n children are brown-eyed (with the father Xx):

$$\Pr(\text{Judy is Xx}\,|\,n \text{ children all have brown eyes \& all previous information}) = \frac{\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n}{\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n + \frac{1}{1+2p}\cdot 1}.$$

Given that Judy's children are all brown-eyed, her grandchild has blue eyes only if Judy's child is Xx. We compute this probability, recalling that we know the child is brown-eyed and we know Judy's spouse is a heterozygote:

$$\begin{aligned}
&\Pr(\text{Judy's child is Xx}\,|\,\text{all the given information}) \\
&= \Pr((\text{Judy is Xx \& Judy's child is Xx}) \textit{ or } (\text{Judy is XX \& Judy's child is Xx})\,|\,\text{all the given information}) \\
&= \frac{\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n}{\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n + \frac{1}{1+2p}}\left(\frac{2}{3}\right) + \frac{\frac{1}{1+2p}}{\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n + \frac{1}{1+2p}}\left(\frac{1}{2}\right).
\end{aligned}$$

Given that Judy's child is Xx, the probability of the grandchild having blue eyes is 0, 1/4, or 1/2, if Judy's child's spouse is XX, Xx, or xx, respectively. Given random mating, these events have probability $(1-p)^2$, $2p(1-p)$, and p^2, respectively, and so

$$\Pr(\text{Grandchild is xx}\,|\,\text{all the given information})$$

$$= \frac{\frac{2}{3}\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n + \frac{1}{2}\frac{1}{1+2p}}{\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n + \frac{1}{1+2p}}\left(\frac{1}{4}2p(1-p) + \frac{1}{2}p^2\right)$$

$$= \frac{\frac{2}{3}\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n + \frac{1}{2}\frac{1}{1+2p}}{\frac{2p}{1+2p}\cdot\left(\frac{3}{4}\right)^n + \frac{1}{1+2p}}\left(\frac{1}{2}p\right).$$

(2) Problem 2.3 (Hoff).

(a) $p(x|y,z) = \frac{p(x,y,z)}{g(y,z)} \propto \frac{f(x,z)g(y,z)h(z)}{g(y,z)} \propto f(x,z)h(z)$

(b) $p(y|x,z) = \frac{p(x,y,z)}{f(x,z)} \propto \frac{f(x,z)g(y,z)h(z)}{f(x,z)} \propto g(y,z)h(z)$

(c) $p(x,y|z) \propto p(x|y,z)p(y|z) = p(x|z)p(y|z)$

(3) (a) Problem 2.5 (Hoff).

	Y		
X	1	0	p(x)
1	0.5×0.4	0.5×0.6	0.5
0	0.5×0.6	0.5×0.4	0.5
p(y)	0.5	0.5	1

(b) $E[Y] = \sum_y yp(y) = 1 \times 0.5 + 0 \times 0.5 = 0.5$ or

$$\begin{aligned} E[Y] = E[E[Y|X]] &= \sum_x p(x)E[Y|X=x] \\ &= \sum_x p(x)\sum_y yp(Y=y|X=x) \\ &= 0.5 \times (1 \times 0.4 + 0 \times 0.6) + 0.5 \times (1 \times 0.6 + 0 \times 0.4) \\ &= 0.5 \end{aligned}$$

(c) $Var[Y|X=0] = E[Y^2|X=0] - (E[Y|X=0])^2 = 0.6(1-0.6) = 0.24$;
$Var[Y|X=1] = E[Y^2|X=1] - (E[Y|X=1])^2 = 0.4(1-0.4) = 0.24$;
$Var[Y] = 0.5(1-0.5) = 0.25$
Var[Y] is larger because all possible values of X (which is a random variable) are considered. For $Var[Y|X=0]$ and $Var[Y|X=1]$ the range of X is restricted.

(d) $Pr(X=0|Y=1) = \frac{Pr(X=0,Y=1)}{Pr(Y=1)} = 0.3/0.5 = 3/5$

(4) Prior $\times$ likelihood

$$\begin{aligned} Pr(X=x)Pr(Y=y|X=x) &= \binom{5}{x} 0.6^x 0.4^{5-x} \binom{x}{y} 0.3^y 0.7x - y \\ &= \frac{5!}{x!(5-x)!} \frac{x!}{y!(x-y)!} \left(\frac{0.6 \times 0.7}{0.4}\right)^x 0.6^6 0.4^5 \left(\frac{0.3}{0.7}\right)^y \\ &\propto 1.05^x \frac{1}{(5-x)!(x-y)!} \end{aligned}$$

$Pr(X=0)Pr(Y=2|X=0) = 0$
$Pr(X=1)Pr(Y=2|X=1) = 0$
$Pr(X=2)Pr(Y=2|X=2) \propto 0.18375$
$Pr(X=3)Pr(Y=2|X=3) \propto 0.57881$
$Pr(X=4)Pr(Y=2|X=4) \propto 0.60775$
$Pr(X=5)Pr(Y=2|X=5) \propto 0.21271$
$Pr(Y=2) = \sum_x Pr(X=x)Pr(Y=2|X=x) \propto 1.5803$

$Pr(X=j|Y=2) = \frac{Pr(X=j)Pr(Y=2|X=j)}{Pr(Y=2)}$

j	$Pr(X=j|Y=2)$
0	0.000
0	0.000
2	0.1161
3	0.3656
4	0.3839
5	0.1344

(5) We want $p(n|y=5) = \frac{p(n)p(y=5|n)}{p(y=5)}$

Now $p(n)p(y=5|n) = 0$ for $n = 0, 1, 2, 3, 4$ and $n \geq 9$

$p(n=5)p(y=5|n=5) = 0.14/5$
$p(n=6)p(y=5|n=6) = 0.13/6$
$p(n=7)p(y=5|n=7) = 0.12/7$
$p(n=8)p(y=5|n=8) = 0.11/8$
$p(y=5) = 0.08056$

So the revised probabilities are:

n	0	1	2	3	4	5	6	7	8	≥ 9
prob	0.00	0.00	0.00	0.00	0.00	0.34756	0.26895	0.212797	0.17068	0.00

(6) (a)

$$\begin{aligned}
f(x,y) &= \frac{\partial F}{\partial x \partial y} \\
&= \frac{\partial}{\partial y}\left(\frac{\partial}{\partial x}\left(x - x\log(x) + x\log(y)\right)\right) \\
&= \frac{\partial}{\partial y}\left(1 - x \times \frac{1}{x} - \log(x) + \log(y)\right) \\
&= \frac{\partial}{\partial y}\left(-\log(x) + \log(y)\right) \\
&= \begin{cases} \frac{1}{y} & 0 < x \leq y < 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

(b)

$$\begin{aligned}
f(x) &= \int_x^1 f(x,y)dy \\
&= \int_x^1 \frac{1}{y}dy \\
&= [\log(y)]_x^1 \\
&= -\log(x) \\
&= \begin{cases} \log\left(\frac{1}{x}\right) & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}$$

$$
\begin{aligned}
f(y) &= \int_0^y f(x,y)dx \\
&= \int_0^y \frac{1}{y}dx \\
&= \frac{1}{y}[x]_0^y \\
&= \begin{cases} 1 & 0 < y < 1 \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
$$

(7) (a)

$$
\begin{aligned}
Pr(W \leq w) &= Pr(min(Y_1, Y_2) \leq w) \\
&= 1 - Pr(Y_1 > w)Pr(Y_2 > w) \\
&= 1 - \exp^{-\lambda_1 w - \lambda_2 w} \\
&= 1 - \exp^{-(\lambda_1 + \lambda_2)w}
\end{aligned}
$$

which is the CDF of an $\text{Exp}(\lambda_1 + \lambda_2)$ distribution.

Now $U_0 = \frac{X_1}{X_1+X_2} \sim Beta(1,1) \equiv U(0,1)$. So

$$
\begin{aligned}
Pr(B_0 = 1) &= Pr(Y_1 < Y_2) \\
&= Pr(\frac{X_1}{\lambda_1} < \frac{X_2}{\lambda_2}) \\
&= Pr(\frac{X_1}{X_2} < \frac{\lambda_1}{\lambda_2}) \\
&= Pr(\frac{X_1}{X_1 + X_2} < \frac{\lambda_1}{\lambda_1 + \lambda_2}) \\
&= \frac{\lambda_1}{\lambda_1 + \lambda_2}
\end{aligned}
$$

Therefore $B_0 \sim Bern(\frac{\lambda_1}{\lambda_1+\lambda_2})$.

(b)

$$
\begin{aligned}
P(U \leq p | M \geq m) &= \frac{P(U \leq p, M \geq m)}{P(M \geq m)} \\
&= \frac{P(U \leq p, U/p \geq m, \frac{1-U}{1-p} \geq m)}{P(U/p \geq m, \frac{1-U}{1-p} \geq m)} \\
&= \frac{P(U \leq p, mp < U < 1 - m(1-p))}{P(mp < U < 1 - m(1-p))} \\
&= \frac{P(mp < u < p)}{P(mp < U < 1 - m(1-p))} \\
&= \frac{p(1-m)}{(1-m)} \\
&= Pr(U \leq p) = Pr(B = 1)
\end{aligned}
$$

Also $Pr(U > p | M \geq m) = Pr(B = 0)$. Therefor, $B \perp\!\!\!\perp M$ as required.

(8) ‘

Yes, they are exchangeable. The joint distribution is

$$p(\theta_1, \dots, \theta_{2J}) = \binom{2J}{J}^{-1} \sum_p \left(\prod_{j=1}^{J} \mathrm{N}(\theta_{p(j)}|1, 1) \prod_{j=J+1}^{2J} \mathrm{N}(\theta_{p(j)}| -1, 1) \right),$$

where the sum is over all permutations p of $(1, \dots, 2J)$. The density (7) is obviously invariant to permutations of the indexes $(1, \dots, 2J)$.

5.4b. Pick any i, j. The covariance of θ_i, θ_j is *negative*. You can see this because if θ_i is large, then it probably comes from the $\mathrm{N}(1, 1)$ distribution, which means that it is more likely than not that θ_j comes from the $\mathrm{N}(-1, 1)$ distribution (because we know that exactly half of the $2J$ parameters come from each of the two distributions), which means that θ_j will probably be negative. Conversely, if θ_i is negative, then θ_j is most likely positive.

Then, by Exercise 5.5, $p(\theta_1, \dots, \theta_{2J})$ cannot be written as a mixture of iid components.

The above argument can be made formal and rigorous by defining $\phi_1, \dots, \phi_{2J}$, where half of the ϕ_j's are 1 and half are -1, and then setting $\theta_j|\phi_j \sim \mathrm{N}(\phi_j, 1)$. It's easy to show first that $\mathrm{cov}(\phi_i, \phi_j) < 0$, and then that $\mathrm{cov}(\theta_i, \theta_j) < 0$ also.

(c) In the limit as $J \to \infty$, the negative correlation between i and j approaches zero, and the joint distribution approaches iid. To put it another way, as $J \to \infty$, the distinction disappears between (1) independently assigning each j to one of two groups, and (2) picking exactly half of the j's for each group.

(9) (a) (i) Yes (ii) Yes (iii) Yes

(b) (i) Yes (ii) No (iii) No

(c) (i) Yes (ii) No (iii) Yes