MA 578 HW1 Solutions

1 (BDA 1.1)

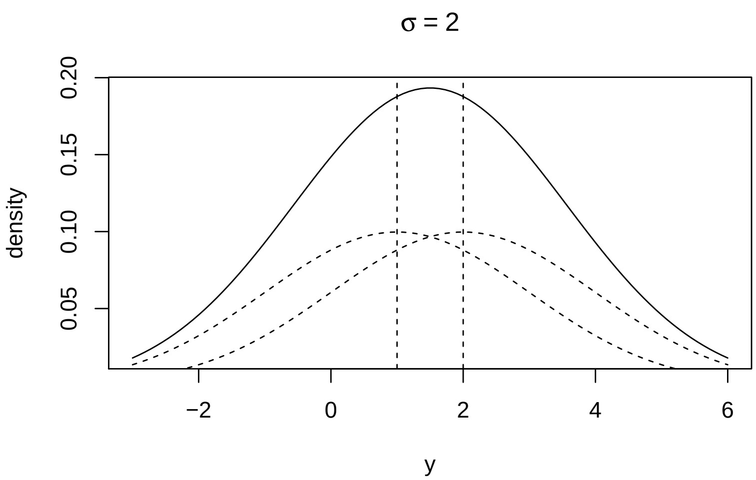
We have $y \mid \theta \sim N(\theta, \sigma^2)$, with $\mathbb{P}(\theta = i) = 1/2$ for $i \in \{1, 2\}$.

(a)

With $\sigma = 2$,

$$\mathbb{P}(y) = \sum_{i \in \{1,2\}} \mathbb{P}(y \mid \theta)\mathbb{P}(\theta) = \frac{1}{2}(2\pi\sigma^2)^{-1/2} \exp\left\{ -\frac{(y-1)^2}{2\sigma^2} \right\} + \frac{1}{2}(2\pi\sigma^2)^{-1/2} \exp\left\{ -\frac{(y-2)^2}{2\sigma^2} \right\}.$$

```
sigma <- 2
t <- seq(1 - 2 * sigma, 2 + 2 * sigma, length = 100)
plot(t, (dnorm(t, 1, sigma) + dnorm(t, 2, sigma)) / 2, type = "l",
     xlab = "y", ylab = "density", main = expression(sigma~"="~2))
lines(t, dnorm(t, 1, sigma) / 2, lty = 2)
lines(t, dnorm(t, 2, sigma) / 2, lty = 2)
abline(v = c(1, 2), lty = 2)
```



(b)

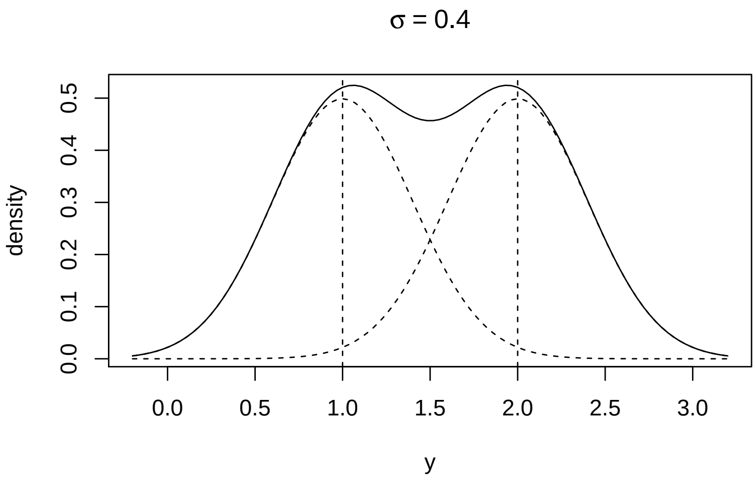
Denoting by $N(y|\mu, \sigma^2)$ the normal density at y with mean μ and variance σ^2, we have that

$$\mathbb{P}(\theta \mid y) = \frac{\mathbb{P}(y \mid \theta)\mathbb{P}(\theta)}{\sum_{i \in \{1,2\}} \mathbb{P}(y \mid \theta = i)\mathbb{P}(\theta = i)} = \frac{N(y|\theta, \sigma^2)/2}{N(y|1, \sigma^2)/2 + N(y|2, \sigma^2)/2},$$

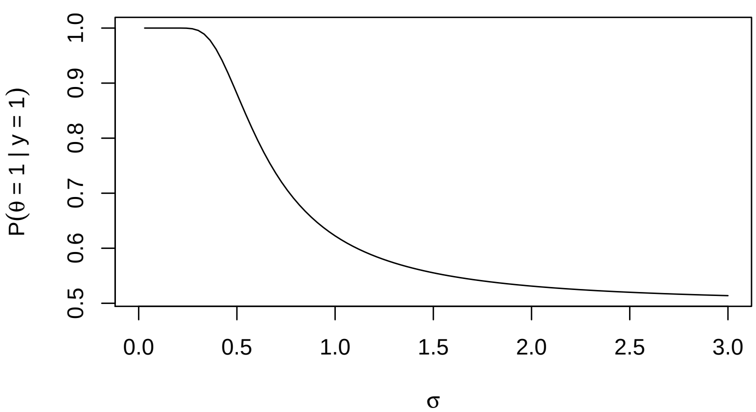
and so $\mathbb{P}(\theta = 1 \mid y = 1) = 0.53$.

(c)

As σ decreases, the contributions from each component become sharper, so the posterior becomes more bimodal. As an example, with $\sigma = 0.4$:



Clearly, as σ decreases, the likelihood becomes more informative. For instance, here's how the posterior $\mathbb{P}(\theta = 1 \mid y = 1)$ changes as a function of σ:



2 (BDA 1.3)

Let us write, for individual I, $G(I) \in \{XX, Xx, xx\}$ and $P(I) \in \{Br, Bl\}$ for the eye color genotypes and phenotypes, respectively. Then, for a child K with parents F and M, we want first $\mathbb{P}(G(K) = Xx \mid D_0)$ with $D_0 = \{P(F) = P(M) = P(K) = Br\}$. By Bayes' rule,

$$\mathbb{P}(G(K) \mid D_0) = \frac{\mathbb{P}(G(K), D_0)}{\sum_{\tilde{G} \in \{XX, Xx, xx\}} \mathbb{P}(G(K) = \tilde{G}, D_0)}.$$

But the joint can be written

$$\begin{aligned}
\mathbb{P}(G(K), D_0) &= \sum_{G(F),G(M)\in\{XX,Xx,xx\}} \mathbb{P}(G(K), G(F), G(M), D_0) \\
&= \sum_{G(F),G(M)\in\{XX,Xx,xx\}} \prod_{I\in\{F,M,K\}} \mathbb{P}(P(I) = Br \,|\, G(I))\mathbb{P}(G(K) \,|\, G(F), G(M))\mathbb{P}(G(F))\mathbb{P}(G(M)) \\
&= \sum_{G(F),G(M)\in\{XX,Xx\}} \mathbb{P}(G(K) \,|\, G(F), G(M))\mathbb{P}(G(F))\mathbb{P}(G(M)),
\end{aligned}$$

and so we get

$$\mathbb{P}(G(K) = Xx, D_0) = 2p(1-p)^3 + \frac{1}{2}[2p(1-p)]^2 = 2p(1-p)^2$$

and

$$\mathbb{P}(G(K) = XX, D_0) = (1-p)^4 + 2p(1-p)^3 + \frac{1}{4}[2p(1-p)]^2 = (1-p)^2,$$

leading to

$$\mathbb{P}(G(K) = Xx \,|\, D_0) = \frac{2p(1-p)^2}{2p(1-p)^2 + (1-p)^2} = \frac{2p}{1+2p}.$$

Next, we observe that Judy (J) has n children $K_1, \dots, K_n$ with phenotypes $D_1 = \{P(K_1) = \cdots = P(K_n) = Br\}$ and want $\mathbb{P}(G(J) \,|\, D)$ with $D = \{D_0, D_1\}$. Let us first compute, for the i-th child,

$$\begin{aligned}
\mathbb{P}(P(K_i) = Br \,|\, G(J)) &= \sum_{G(J)\in\{XX,Xx,xx\}} \mathbb{P}(P(K_i) = Br \,|\, G(K_i))\mathbb{P}(G(K_i) \,|\, G(J)) \\
&= \sum_{G(J)\in\{XX,Xx\}} \mathbb{P}(G(K_i) \,|\, G(J)) = \begin{cases} 1/2 + 1/2 = 1, & \text{if } G(J) = XX, \\ 1/4 + 1/2 = 3/4, & \text{if } G(J) = Xx \end{cases}
\end{aligned}$$

Assuming that the children phenotypes are independent,

$$\mathbb{P}(G(J), D_1 \,|\, D_0) = \prod_{i=1}^{n} \mathbb{P}(P(K_i) = Br \,|\, G(J))\mathbb{P}(G(J) \,|\, D_0),$$

and so

$$\mathbb{P}(G(J) = Xx \,|\, D) = \frac{(3/4)^n \frac{2p}{1+2p}}{(3/4)^n \frac{2p}{1+2p} + \frac{1}{1+2p}} = \frac{2p(3/4)^n}{1 + 2p(3/4)^n}.$$

For the last part, we want the probability that the first grandchild (B) of Judy has blue eyes. Let us assume that B is the offspring of K_i and A; we then want $\mathbb{P}(P(B) = Bl \,|\, D)$. Denoting by $K_{[-i]}$ all the other $n-1$ children of Judy excluding K_i, we start by finding

$$\begin{aligned}
\mathbb{P}(G(K_i) \,|\, D) &= \sum_{G(J),G(K_{[-i]})} \mathbb{P}(G(K_i), G(K_{[-i]}), G(J) \,|\, D) \\
&= \frac{\sum_{G(K_{[-i]}),G(J)} \mathbb{P}(G(K_i), G(K_{[-i]}), G(J), P(K_i), P(K_{[-i]}) \,|\, D_0)}{\sum_{G(K_i)} \sum_{G(K_{[-i]}),G(J)} \mathbb{P}(G(K_i), G(K_{[-i]}), G(J), P(K_i), P(K_{[-i]}) \,|\, D_0)}.
\end{aligned}$$

Since

$$\sum_{G(K_{[-i]}),G(J)} \mathbb{P}(G(K_i), G(K_{[-i]}), G(J), P(K_i), P(K_{[-i]}) \,|\, D_0) =$$
$$\sum_{G(J)} \mathbb{P}(P(K_i) \,|\, G(K_i))\mathbb{P}(G(K_i) \,|\, G(J)) \sum_{G(K_{[-i]})} \mathbb{P}(P(K_{[-i]}) \,|\, G(K_{[-i]}))\mathbb{P}(G(K_{[-i]}) \,|\, G(J))\mathbb{P}(G(J) \,|\, D_0),$$

we arrive at

$$\mathbb{P}(G(K_i) = Xx, D_1 \,|\, D_0) = \frac{1}{2}\left(\frac{3}{4}\right)^{n-1} \frac{2p}{1+2p} + \frac{1}{2}\frac{1}{1+2p}$$

and

$$\mathbb{P}(G(K_i) = XX, D_1 \,|\, D_0) = \frac{1}{4}\left(\frac{3}{4}\right)^{n-1} \frac{2p}{1+2p} + \frac{1}{2}\frac{1}{1+2p},$$

and thus,

$$\mathbb{P}(G(K_i) = Xx \,|\, D) = \frac{1+2p(3/4)^{n-1}}{2+3p(3/4)^{n-1}}.$$

Finally, with $\pi \doteq \mathbb{P}(G(K_i) = Xx \,|\, D)$,

$$\begin{aligned}
\mathbb{P}(P(B) = Bl \,|\, D) &= \sum_{G(B),G(A),G(K_i)} \mathbb{P}(P(B) = Bl, G(B), G(A), G(K_i) \,|\, D) \\
&= \sum_{G(B),G(A),G(K_i)} \mathbb{P}(P(B) = Bl \,|\, G(B))\mathbb{P}(G(B) \,|\, G(A), G(K_i))\mathbb{P}(G(A))\mathbb{P}(G(K_i) \,|\, D) \\
&= \sum_{G(A)\in\{Xx,xx\}} \mathbb{P}(G(B) = xx \,|\, G(A), G(K_i) = Xx)\mathbb{P}(G(A))\mathbb{P}(G(K_i) = Xx \,|\, D) \\
&= \frac{1}{4}2p(1-p)\pi + \frac{1}{2}p^2\pi = \frac{\pi p}{2}.
\end{aligned}$$

3

Let us denote by S the small stone cases, S^c, the complement of S, the large stone cases.

(a)

A direct application of Bayes' rule yields

$$\mathbb{P}(S \,|\, A) = \frac{\mathbb{P}(A \,|\, S)\mathbb{P}(S)}{\mathbb{P}(A \,|\, S)\mathbb{P}(S) + \mathbb{P}(A \,|\, S^c)\mathbb{P}(S^c)} = \frac{.24 \cdot .51}{.24 \cdot .51 + .77 \cdot .49} = 0.245,$$

and

$$\mathbb{P}(S \,|\, B) = \frac{\mathbb{P}(B \,|\, S)\mathbb{P}(S)}{\mathbb{P}(B \,|\, S)\mathbb{P}(S) + \mathbb{P}(B \,|\, S^c)\mathbb{P}(S^c)} = \frac{.76 \cdot .51}{.76 \cdot .51 + .23 \cdot .49} = 0.775.$$

Note that $\mathbb{P}(S \,|\, A) \approx \mathbb{P}(A \,|\, S)$ since $\mathbb{P}(S) \approx 1/2$.

(b)

Computing the marginals gives us

$$\begin{aligned}\mathbb{P}(\text{Succ} \mid A) &= \mathbb{P}(\text{Succ} \mid S, A)\mathbb{P}(S \mid A) + \mathbb{P}(\text{Succ} \mid S^c, A)\mathbb{P}(S^c \mid A) \\ &= .93 \cdot .245 + .73 \cdot (1 - .245) = 0.78,\end{aligned}$$

and

$$\begin{aligned}\mathbb{P}(\text{Succ} \mid B) &= \mathbb{P}(\text{Succ} \mid S, B)\mathbb{P}(S \mid B) + \mathbb{P}(\text{Succ} \mid S^c, B)\mathbb{P}(S^c \mid B) \\ &= .87 \cdot .775 + .69 \cdot (1 - .775) = 0.83.\end{aligned}$$

Thus, $\mathbb{P}(\text{Succ} \mid B) > \mathbb{P}(\text{Succ} \mid A)$ even though $\mathbb{P}(\text{Succ} \mid B, C) < \mathbb{P}(\text{Succ} \mid A, C)$ for both cases $C \in \{S, S^c\}$! The last derivation explains the problem: cases with lower performance—here, large stone cases—are assigned more frequently to treatment A, reducing its marginal rate of success.