# MATH3823 - Solutions to Chapter 3 Exercises

Exercise 3.1   The estimate of p is simply $\hat{p} = \bar{y}/m$, were $m = 62$, that is $\hat{p} = 0.5866935$. Hence, here our model is $Y \sim \mathrm{Bin}(62, 0.59)$ with the corresponding fitted probability mass function added to the density-scaled histogram. The fitted model is a very bad fit to the data and is clearly inappropriate.

```
y = c(6,13,18,28,52,53,61,60)
m = 62

hist(y, main="", ylim=c(0,0.1), probability=T)

p.hat = mean(y)/m
fitted = dbinom(0:m, m, p.hat)
lines(0:m, fitted, type='h', lwd=3, col="blue")
```
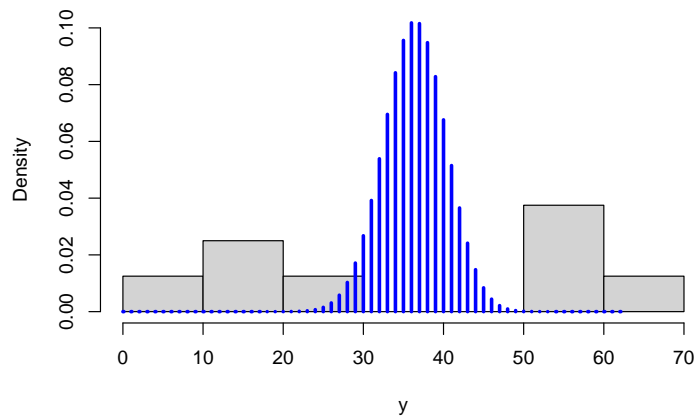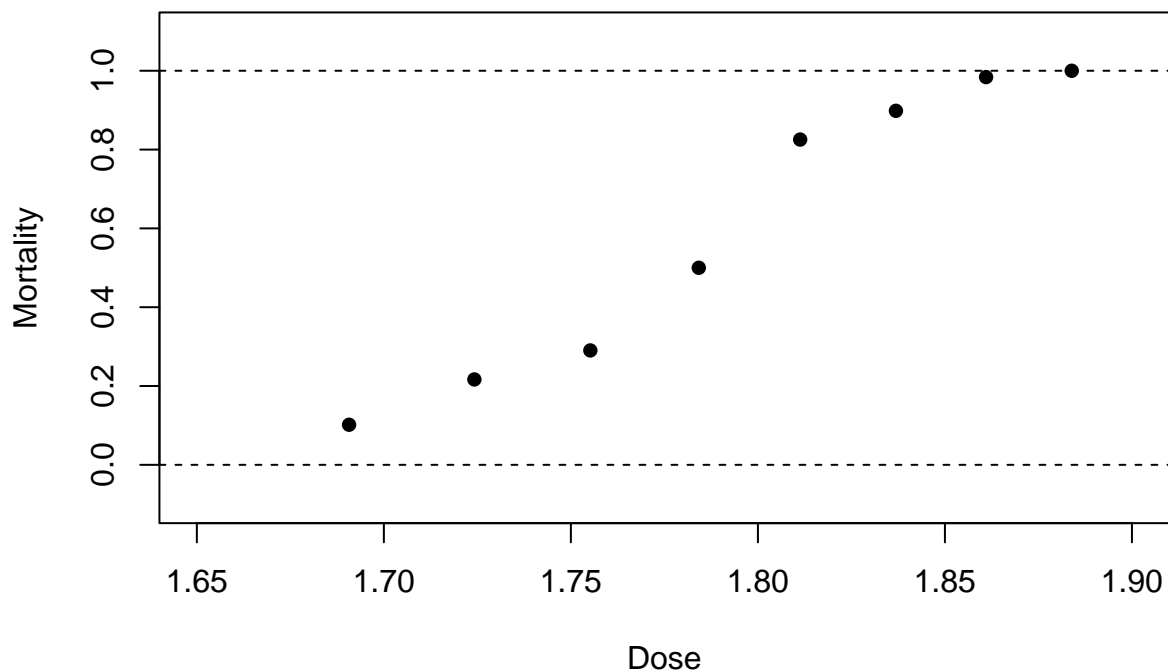

Although not asked in the question, consider the problem more generally. Recall the original situation, with mortality against dose:

```
beetle = read.table("https://rgaykroyd.github.io/MATH3823/Datasets/beetle.txt", header=T)

dose = beetle$dose
mortality = beetle$died/beetle$total

plot(dose, mortality, pch=16,
     xlim=c(1.65, 1.90), xlab ="Dose",
     ylim=c(-0.1, 1.1),  ylab="Mortality")
abline(h=c(0,1), lty=2)
```

We see that although the *average* mortality is likely to be around 0.59, it is very clear that the actually mortality is strongly dependent on dose. Here, the model misfit is caused by an important variable not being involved in the modelling – a hidden variable.

Exercise 3.2   Solving the first equation for x yields the following steps:

$$q = \frac{1}{1 + e^{-x}}$$

$$\Leftrightarrow \qquad q(1 + e^{-x}) = 1$$

$$\Leftrightarrow \qquad e^{-x} = \frac{1 - q}{q}$$

$$\Leftrightarrow \qquad x = \log\left(\frac{q}{1 - q}\right),$$

as required.

For the second part, starting from the above but with $x = \theta$ and $q = p$ then we know that $\theta = \log\left(p/(1 - p)\right)$

implies that $p = 1/(1 + e^{-\theta})$ and hence

$$
\begin{aligned}
-m \log(1 - p) &= -m \log \left(1 - \frac{1}{1 + e^{-\theta}} \right) \\
&= -m \log \left(\frac{(1 + e^{-\theta}) - 1}{1 + e^{-\theta}} \right) \\
&= -m \log \left(\frac{1}{e^{\theta} + 1} \right) \\
&= m \log \left(1 + e^{\theta} \right)
\end{aligned}
$$

as required.

Exercise 3.3 This can be interpreted as a regular exponential family with natural parameter λ/α and scale parameter $1/\alpha$. Start by writing the log pdf (probability density function) as

$$
\begin{aligned}
\log f(y) &= \frac{-(\lambda/\alpha)y + \log \lambda}{1/\alpha} + (\alpha - 1) \log y - \log \Gamma(\alpha) \\
&= \frac{-(\lambda/\alpha)y + \log \lambda/\alpha}{1/\alpha} + (\alpha - 1) \log y - \log \Gamma(\alpha) + \alpha \log \alpha \\
&= \frac{\theta y - b(\theta)}{\phi} + c(y, \phi),
\end{aligned}
$$

for $y > 0$, where
$$
\theta = -\lambda/\alpha \ (< 0), \quad b(\theta) = -\log(-\theta), \quad \phi = 1/\alpha \ (> 0),
$$

and

$$
\begin{aligned}
c(y, \phi) &= (\alpha - 1) \log y - \log \Gamma(\alpha) + \alpha \log \alpha \\
&= \left(\frac{1 - \phi}{\phi} \right) \log y - \log \Gamma(\phi^{-1}) - \phi^{-1} \log \phi.
\end{aligned}
$$

Note the extra term $\alpha \log \alpha$ which appears in $c(y, \phi)$ compared to the first equation.

Exercise 3.4 To express this in exponential form we rearrange the probability mass function:

$$
\begin{aligned}
f(y) &= \exp\{(y - 1) \log(1 - p) + \log p\} \\
&= \exp\{y \log(1 - p) + \log \{p/(1 - p)\} \}.
\end{aligned}
$$

Hence, we have that $\theta = \log(1 - p)$ and hence $p = 1 - e^{\theta}$ and $1 - p = e^{\theta}$. Further, $b(\theta) = \log \{p/(1 - p)\} = \log \{e^{-\theta} - 1\}$, $c(y, \theta) = 0$ and $\phi = 1$.

Exercise 3.5 For a discrete random variable, starting with the property that all probability mass functions sum to 1, we have

$$
1 = \sum_{y \in \Omega_Y} \exp \left\{ \frac{y\theta - b(\theta)}{\phi} + c(y, \phi) \right\}
$$

and then differentiating both sides with respect to θ gives

$$0 = \sum \left[\frac{y - b'(\theta)}{\phi} \right] \exp \left\{ \frac{y\theta - b(\theta)}{\phi} + c(y, \phi) \right\}.$$

Next, using the definition of the exponential family to simplify the equation gives

$$0 = \sum \left[\frac{y - b'(\theta)}{\phi} \right] f(y; \theta)$$

and expanding the brackets leads to

$$0 = \frac{1}{\phi} \left(\sum y f(y; \theta) - b'(\theta) \sum f(y; \theta) \right).$$

The first sum is simply the expectation of Y and the second is the integral of the probability mass function of Y, and hence

$$0 = \frac{1}{\phi} \left(\mathrm{E}[Y] - b'(\theta) \right)$$

which implies that

$$\mathrm{E}[Y] = b'(\theta),$$

which proves the first part of the proposition.

Differentiating a second time, by parts and then using the definition of the exponential family to simplify again, yields

$$0 = \sum \left\{ -\frac{b''(\theta)}{\phi} + \left[\frac{y - b'(\theta)}{\phi} \right]^2 \right\} f(y; \theta)$$

and using the just prooved result about expectation gives,

$$0 = -\frac{b''(\theta)}{\phi} + \sum \left[\frac{y - \mathrm{E}[Y]}{\phi} \right]^2 f(y; \theta)$$

and

$$0 = -\frac{b''(\theta)}{\phi} + \frac{\mathrm{Var}[Y]}{\phi^2}$$

which implies that

$$\mathrm{Var}[Y] = \phi \, b''(\theta).$$

which proves the second part of the proposition for discrete random variables.

Now apply the result to the Poisson and binomial distributions.

For the Poisson, since $b(\theta) = e^\theta$ then $b'(\theta) = e^\theta$ and $b''(\theta) = e^\theta$. With $\theta = \log \lambda$ we get the usual, $\mathrm{E}[Y] = \lambda$ and $\mathrm{Var}[Y] = \lambda$.

For the binomial, $Y \sim \mathrm{Bin}(m, p)$, $b(\theta) = m \log(1 + e^\theta)$ and hence

$$b'(\theta) = m \frac{e^\theta}{(1 + e^\theta)}$$

and

$$b''(\theta) = m \left(\frac{e^\theta}{(1 + e^\theta)} - \frac{(e^\theta)^2}{(1 + e^\theta)^2} \right) = m \left(\frac{e^\theta}{(1 + e^\theta)^2} \right).$$

From $\theta = \mathrm{logit}(p)$ we get $e^\theta = p/(1 - p)$ and $1 + e^\theta = 1/(1 - p)$. These give,

$$\mathrm{E}[Y] = m \frac{p/(1 - p)}{1/(1 - p)} = mp$$

and

$$\text{Var}[Y] = m\frac{p/(1-p)}{1/(1-p)^2} = mp(1-p)$$

– as expected.

Exercise 3.6 For the $Y \sim \text{Geom}(p)$, the mean and variance of Y are, of course, well-known and can also be found by differentiating $b(\theta)$,

$$E[Y] = b'(\theta) = 1 + \frac{e^\theta}{1-e^\theta} = \frac{1}{1-e^\theta} = 1/p,$$

$$\text{Var}[Y] = b''(\theta) = \frac{e^\theta}{(1-e^\theta)^2} = (1-p)/p^2.$$

For the gamma distribution, $y \sim \text{Gamma}(\alpha, \lambda)$,

Then

$$E[Y] = b'(\theta) = \alpha/\lambda \quad \text{Var}[Y] = \phi \, b''(\theta) = \alpha/\lambda^2.$$

(Take care differentiating $b(\theta)$.) These moments for Y are also well-known for the gamma distribution.

Exercise 3.7 For the $\text{Bin}(m, p)$ distribution, we have $\mu = b'(\theta) = m/(1 + e^{-\theta})$. Hence

$$1 + e^{-\theta} = m/\mu$$
$$-\theta = \log(m/\mu - 1)$$
$$\theta = \log\left(\frac{1}{m/\mu - 1}\right) = \log\left(\frac{\mu/m}{1 - \mu/m}\right)$$
$$\theta = \text{logit}(\mu/m) = g(\mu).$$

For the $\text{Gamma}(\alpha, \lambda)$ distribution, $\mu = b'(\theta) = -1/\theta$, so

$$\mu = -1/\theta$$
$$\theta = -1/\mu = g(\mu).$$

End of Solutions to Chapter 3 Exercises