Exercise 4

a)

The multivariate normal is $f(x;\mu,\Sigma) = (2\pi\det\Sigma)^{-1/2}\exp\left(-1/2\cdot(x-\mu)^T\Sigma^{-1}(x-\mu)\right)$. When $\Sigma = \sigma^2 I$, $\Sigma^{-1} = I\sigma^{-2}$ and $\det\Sigma = \sigma^{2n}$. The claim follows.

b)

Take the logarithm and find a log-likelihood proportional to $(x-\mu)^T(x-\mu)$. This is minimized by least squares.

c)

The recipe is the usual one: Take the logarithm, differentiate, and set equal to 0.

$$\log l(\widehat{\mu},\sigma) = -n/2\log(2\pi) - n\log\sigma - 2^{-1}\sigma^{-2}SSE.$$

Differentiate $-n\log\sigma - \sigma^{-2}SSE$ to get $-n\sigma^{-1} - \sigma^{-3}SSE = 0$, so that $\sigma^2 = \frac{SSE}{n}$ minimizes the expression.

d)

Since $\widehat{\sigma} = SSE/n$, we get

$$\begin{aligned}\log l(\widehat{\mu},\widehat{\sigma}) &= -n/2\log(2\pi) - n\log\widehat{\sigma} - 2^{-1}\widehat{\sigma}^{-2}SSE \\ &= -n/2\log(2\pi) - n\log\widehat{\sigma} - n2^{-1},\end{aligned}$$

and exponentiate this to verify the claim.

e)

Let SSE_0 belong to X_0 and SSE_1 belong to X_1. Using the previous result,

$$\log l(\widehat{\mu_0},\widehat{\sigma_0}) - \log l(\widehat{\mu_1},\widehat{\sigma_1}) = -n(\log\widehat{\sigma_0} - \log\widehat{\sigma_1})$$

Exponentiate these to get $\left(\widehat{\sigma_0^2}/\widehat{\sigma_1^2}\right)^{-n/2}$, which again equals $\left(\frac{SSE_1}{SSE_0}\right)^{-n/2}$. Now use that

$$\begin{aligned}\left(\frac{SSE_1}{SSE_0}\right)^{-n/2} &= \left(1 + \frac{SSE_0 - SSE_1}{SSE_0}\right)^{-n/2} \\ &= \left(1 + \frac{p_1 - p_0}{n - p_1}F\right)^{-n/2}\end{aligned}$$

by the definition of F.

Exercise 5

This is an `R` exercise. You will find it on the exercise webpage in the .Rmd file format.

Exercise 6

a)

$$\begin{aligned}
\pi_i &= P\left(Y_i = 1\right) \\
&= P\left(Y_i^* > 0\right) \\
&= P\left(\sum_{j=1}^{p} \beta_j^* x_{i,j} + \sigma\epsilon_i > 0\right) \\
&= P\left(\epsilon_i > -\sum_{j=1}^{p} \beta_j x_{i,j}\right) \\
&= 1 - P\left(\epsilon_i \leq -\sum_{j=1}^{p} \beta_j x_{i,j}\right) \\
&= 1 - F\left(-\sum_{j=1}^{p} \beta_j x_{i,j}\right) \\
&= F\left(\sum_{j=1}^{p} \beta_j x_{i,j}\right)
\end{aligned}$$

b)

Let $\eta_i = \sum_{j=1}^{p} \beta_j x_{i,j}$, then $\pi_i = F\left(\eta_i\right)$.
If $\epsilon_i \sim N(0,1)$, $\pi_i = F\left(\eta_i\right) = \Phi\left(\eta_i\right)$ where η_i is a linear predictor. This is equal to probit model. (See p.167 of the book.)

c)

If $F(z) = \frac{e^z}{1+e^z}$, $\pi_i = F\left(\eta_i\right) = \frac{e^{\eta_i}}{1+e^{\eta_i}}$. This is equal to $\log\left(\frac{\pi_i}{1-\pi_i}\right) = \eta_i$, which is logistic regression.

d)

If $\eta_i \sim$ uniform $\left(-\frac{1}{2}, \frac{1}{2}\right)$, $\pi_i = F\left(\eta_i\right) = \begin{cases} 0 & \text{if } \eta_i < -\frac{1}{2} \\ \eta_i + \frac{1}{2} & \text{if } -\frac{1}{2} \leq \eta_i < \frac{1}{2} \\ 1 & \text{if } \frac{1}{2} \leq \eta_i \end{cases}$.
This is linear probability model (p.167 of the book).

Exercise 7

a)

$$f(y, \lambda) = \frac{\lambda^y e^{-\lambda}}{y!} = \exp\left[y \log\lambda - \lambda - \log(y!)\right]$$

So, $\theta = \log\lambda$, $b(\theta) = e^\theta$, $a(\phi) = 1$, $c(y, \phi) = -\log(y!)$.

b)

With $\theta = \log\lambda$,

$$\mathrm{E}[Y] = b'(\theta) = e^\theta = \lambda$$

and

$$\mathrm{Var}(Y) = b''(\theta)a(\phi) = e^{\theta} = \lambda.$$

c)

Canonical link function $g(\cdot)$ is a link function such that $\theta = g\left(\mathrm{E}[Y]\right)$. In our case, $\theta = \log\lambda = g(\lambda)$. Thus, the canonical link function for Poission GLM is $g(\mu) = \log\mu$.

Exercise 8

a)

$$f(y,\pi) = \pi(1-\pi)^y = \exp\left[y\log(1-\pi) + \log\pi\right]$$

So, $\theta = \log(1-\pi)$, $\pi = 1-e^{\theta}$, $b(\theta) = -\log\left(1-e^{\theta}\right)$, $a(\phi) = 1$, $c(y,\phi) = 0$.

b)

$$f(y,\pi) = \binom{y+r-1}{r-1}\pi^r(1-\pi)^y = \exp\left[y\log(1-\pi) + r\log\pi + \log\binom{y+r-1}{r-1}\right]$$

So, $\theta = \log(1-\pi)$, $\pi = 1-e^{\theta}$ $b(\theta) = -r\log\left(1-e^{\theta}\right)$, $a(\phi) = 1$, $c(y,\phi) = \log\binom{y+r-1}{r-1}$.

c)

$$\mathrm{E}[Y] = b'(\theta) = r\frac{e^{\theta}}{1-e^{\theta}} = r\left(\frac{1}{\pi}-1\right)$$

and

$$\mathrm{Var}(Y) = b''(\theta)a(\phi) = r\frac{e^{\theta}}{\left(1-e^{\theta}\right)^2} = \frac{r}{\pi}\left(\frac{1}{\pi}-1\right).$$

Exercise 9

a)

i)

$$f(y,\lambda) = \lambda e^{-\lambda y} = \exp\left[-\lambda y + \log\lambda\right]$$

So, $\theta = -\lambda$, $b(\theta) = -\log(-\theta)$, $a(\phi) = 1$, $c(y,\phi) = 0$.

ii)

$$\mathrm{E}[Y] = b'(\theta) = -\frac{1}{\theta} = \frac{1}{\lambda}$$

and

$$\mathrm{Var}(Y) = b''(\theta)a(\phi) = \frac{1}{\theta^2} = \frac{1}{\lambda^2}.$$

b)

i)

$$
\begin{aligned}
f(y,\pi) &= \frac{\left(\frac{k}{\mu}\right)^k}{\Gamma(k)} y^{k-1} e^{-\frac{k}{\mu}y} \\
&= \exp\left[-\frac{k}{\mu}y + k\log\frac{k}{\mu} + (k-1)\log y - \log\Gamma(k)\right] \\
&= \exp\left[\frac{\left(-\frac{1}{\mu}y - \log\mu\right)}{\frac{1}{k}} + (k-1)\log y - \log\Gamma(k) + k\log k\right]
\end{aligned}
$$

So, $\theta = -\frac{1}{\mu}$, $\mu = -\frac{1}{\theta}$, $b(\theta) = -\log(-\theta)$, $a(\phi) = \frac{1}{k}$, $c(y,\phi) = (k-1)\log y - \log\Gamma(k)$.

ii)

$$\mathrm{E}[Y] = b'(\theta) = -\frac{1}{\theta} = \mu$$

and

$$\mathrm{Var}(Y) = b''(\theta)a(\phi) = \frac{1}{k\theta^2} = \frac{\mu^2}{k}.$$

Exercise 10

a)

The moment generating function is defined as

$$M_Y(t) = \mathrm{E}\left[e^{Yt}\right] = \begin{cases} \sum e^{yt} f(y) & \text{if } Y \text{ is discrete} \\ \int e^{yt} f(y))\,dy & \text{if } Y \text{ is continuous} \end{cases}.$$

When $t = 0$,

$$M_Y(0) = \mathrm{E}\left[1\right] = \begin{cases} \sum f(y) = 1 & \text{if } Y \text{ is discrete} \\ \int f(y))\,dy = 1 & \text{if } Y \text{ is continuous} \end{cases}.$$

And if we differentiate $M_Y(t)$ with respect to t,

$$M_Y^{(r)}(t) = \frac{d^r}{dt^r} M_Y(t) = \mathrm{E}\left[Y^r e^{Yt}\right] = \begin{cases} \sum e^{yt} y^r f(y) & \text{if } Y \text{ is discrete} \\ \int e^{yt} y^r f(y))\,dy & \text{if } Y \text{ is continuous} \end{cases}.$$

When $t = 0$, this becomes

$$M_Y^{(r)}(0) = \frac{d^r}{dt^r} M_Y(0) = \mathrm{E}\left[Y^r\right] = \begin{cases} \sum y^r f(y) & \text{if } Y \text{ is discrete} \\ \int y^r f(y))\,dy & \text{if } Y \text{ is continuous} \end{cases}.$$

When $r = 1$, we have $M_Y'(0) = \mathrm{E}\left[Y\right]$ and when $r = 2$, we have $M_Y''(0) = \mathrm{E}\left[Y^2\right]$.

b)

$$R'_Y(t) = \frac{d\log M_Y(t)}{dM_Y(t)} \cdot \frac{dM_Y(t)}{dt} = \frac{M'_Y(t)}{M_Y(t)}$$

$$R''_Y(t) = \frac{d}{dt}\left(\frac{M'_Y(t)}{M_Y(t)}\right) = \frac{M''_Y(t)M_Y(t) - M'_Y(t)^2}{M_Y(t)^2}$$

When $t = 0$, $R'_Y(0) = \frac{M'_Y(0)}{M_Y(0)} = \mathrm{E}[Y]$ and $R''_Y(0) = \frac{M''_Y(0)M_Y(0) - M'_Y(0)^2}{M_Y(0)^2} = \mathrm{E}\left[Y^2\right] - \mathrm{E}[Y]^2 = \mathrm{Var}(Y)$.

Exercise 11

a)

Put $t = a(\phi)s$ and observe that

$$\begin{aligned}
\int f(x \mid \theta, \phi) e^{tx} dx &= \int \exp\left[\frac{x(\theta + s) - b(\theta)}{a(\phi)} - c(x, \phi)\right] dx \\
&= \exp\left[-\frac{b(\theta)}{a(\phi)}\right] \int \exp\left[x\frac{(\theta + s)}{a(\phi)} - c(x, \phi)\right] dx \\
&= \exp\left[-\frac{b(\theta)}{a(\phi)}\right] \exp\left[\frac{b(\theta + s)}{a(\phi)}\right] \int f(x \mid \theta + s, \phi) dx \\
&= \exp\left[\frac{b(\theta + s) - b(\theta)}{a(\phi)}\right].
\end{aligned}$$

Since the cumulant generating function is $\log M(t)$, it equals $K(t) = \frac{b(\theta+s)-b(\theta)}{a(\phi)}$.

b)

Recall that the mean equals the first cumulant and the variance equals the second. Look it up at wikipedia if you want to.

The first derivative of the cumulant generating function is

$$\begin{aligned}
\frac{d}{dt}K(t) &= \frac{\frac{d}{dt}\exp\left[\frac{b(\theta+s)-b(\theta)}{a(\phi)}\right]}{\exp\left[\frac{b(\theta+s)-b(\theta)}{a(\phi)}\right]} \\
&= \frac{1}{a(\phi)}\frac{d}{dt}\frac{b\left(\theta + a(\phi)^{-1}t\right)}{a(\phi)} \\
&= b'\left(\theta + a(\phi)^{-1}t\right).
\end{aligned}$$

When $t = 0$, this is $b'(0)$ as claimed. As for the variance,

$$\begin{aligned}
\frac{d^2}{dt^2}K(t) &= \frac{d}{dt}b'\left(\theta + a(\phi)^{-1}t\right), \\
&= \frac{b''\left(\theta + a(\phi)^{-1}t\right)}{a(\phi)},
\end{aligned}$$

and $t = 0$ yields $b''(\theta)/a(\phi)$.