# Part IB — Statistics Example Sheet 3

Supervised by Mr. Higson
Examples worked through by Christopher Turnbull

Lent 2018

# QUESTION 1

Let $X \sim N_n(\mu, \Sigma)$, and let $A \in \mathbb{R}^{m \times n}$. $X$ has a $n$-variable normal distribution, so for every $t \in \mathbb{R}^n$, the rv. $t^T X$ has a univariate normal distribution.

Have

$$AX = \begin{pmatrix} a_{11}X_1 + \cdots + a_{1n}X_n \\ \vdots \\ \vdots \\ a_{m1}X_1 + \cdots + a_{mn}X_n \end{pmatrix}$$

Want to show that for every $s \in \mathbb{R}^m$, the variable $s^T AX$ has a univariate normal distribution. Not sure. Assuming it does, we calculate the mean and variance.

Have

$$\mathbb{E}[AX] = \begin{pmatrix} a_{11}\mathbb{E}[X_1] + \cdots + a_{1n}\mathbb{E}[X_n] \\ \vdots \\ \vdots \\ a_{m1}\mathbb{E}[X_1] + \cdots + a_{mn}\mathbb{E}[X_n] \end{pmatrix} = \begin{pmatrix} a_{11}\mu_1 + \cdots + a_{1n}\mu_n \\ \vdots \\ \vdots \\ a_{m1}\mu_1 + \cdots + a_{mn}\mu_n \end{pmatrix} = A\mu$$

and

$$\text{Var}\,[AX] = \begin{pmatrix} a_{11}^2 \text{Var}\,[X_1] + \cdots + a_{1n}^2 \text{Var}\,[X_n] \\ \vdots \\ \vdots \\ a_{m1}^2 \text{Var}\,[X_1] + \cdots + a_{mn}^2 \text{Var}\,[X_n] \end{pmatrix} = \begin{pmatrix} a_{11}^2\sigma_1^2 + \cdots + \sigma_n^2\sigma_n^2 \\ \vdots \\ \vdots \\ a_{m1}^2\sigma_1^2 + \cdots + a_{mn}^2\sigma_n^2 \end{pmatrix} = A\Sigma A^T$$

# QUESTION 2

$$
\begin{aligned}
\frac{\partial \Omega}{\partial y} &= \left(-\mu r_1 + \frac{\mu}{r_1^2}\right)\frac{\partial r_1}{\partial y} + \left(-(1-\mu)r_2 + \frac{1-\mu}{r_2^2}\right)\frac{\partial r_2}{\partial y} \\
&= \left(-\mu r_1 + \frac{\mu}{r_1^2}\right)\frac{y}{r_1} + \left(-(1-\mu)r_2 + \frac{1-\mu}{r_2^2}\right)\frac{y}{r_2} \\
&= -\mu y + \frac{\mu y}{r_1^3} - (1-\mu)y + \frac{(1-\mu)y}{r_2^3} \\
&= -y + \frac{\mu y}{r_1^3} + \frac{(1-\mu)y}{r_2^3}
\end{aligned}
$$

# QUESTION 3

Choosing $A$ be the $n_1 \times n$ matrix

$$A = \begin{pmatrix} 1 & & 0 & \\ & \ddots & & 0 \\ 0 & & 1 & \end{pmatrix}$$

Have $AX = X_1$, and from Question 1 we have

$$X_1 \sim N_{n_1}(A\mu, A\Sigma A^T)$$

And we can see that $AX_1 = \mu_1$ and $A\Sigma A^T = \Sigma_{11}$ as required.

# QUESTION 4

$$Y_i = a + bx_i + \varepsilon_i$$

This is a linear model; we have $Y_i \sim N(a + bx_i, \sigma^2)$, with likelihood

$$f_{Y_i} = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - a - bx_i)^2\right)$$

Then

$$\frac{\partial f_{Y_i}}{\partial a} \propto \sum_{i=1}^n (y_i - a - bx_i)$$

and since $\sum_{i=1}^n x_i = 0$ we have

$$\hat{a} = \frac{1}{n}\sum_{i=1}^n y_i = \bar{Y}$$

Next, we have

$$\frac{\partial f_{Y_i}}{\partial b} \propto \sum_{i=1}^n (y_i - a - bx_i)x_i$$

then

$$\hat{b} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

Now the log-likelihood is given by

$$l(\alpha, \beta, \sigma) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - a - bx_i)^2$$

Here,

$$\frac{\partial l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - a - bx_i)^2$$

Thus setting this to zero at $a = \hat{a}$, $b = \hat{b}$ yields

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^n (y_i - \hat{a} - \hat{b}x_i)^2$$

# QUESTION 5

This is a linear model with

$$y_i = \underbrace{\frac{v^2}{g}}_{\hat{\beta}} \underbrace{\sin 2\alpha}_{x_i} + \varepsilon_i$$

Here,

$$\begin{aligned}\hat{b} &= \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2} \\ &= 28156\end{aligned}$$

Hence $v = \sqrt{g\hat{\beta}}$ which is approximately 526 $\text{ms}^{-1}$.

# QUESTION 6

$$Y_{ij} = \mu_i + \varepsilon_{ij}$$

This is a linear model; we have $Y_{ij} \sim N(\mu_i, \sigma^2)$, with likelihood

$$f_{Y_{ij}} = \frac{1}{(2\pi\sigma^2)^{n_i/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{I}\sum_{i=j}^{n_i}(y_{ij}-\mu_i)^2\right)$$

Then

$$\frac{\partial f_{Y_{ij}}}{\partial \mu_i} \propto \left(2n_i\hat{\mu}_i - 2\sum_{j=1}^{n_i} y_{ij}\right)$$

thus

$$\hat{\mu}_i = \frac{1}{n}\sum_{j=1}^{n_i} y_{ij} = \hat{Y}_i$$

Now the log-likelihood is given by

$$l(\mu_i, \sigma) = -\frac{n_i}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - a - bx_i)^2$$

Here,

$$\frac{\partial l}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^2}\sum_{i=1}^{I}\sum_{i=j}^{n_i}(y_{ij}-\mu_i)^2$$

Thus setting this to zero yields

$$0 = -\frac{n_i}{2\hat{\sigma}^2} + \frac{1}{2\hat{\sigma}^4}\sum_{i=1}^{I}\sum_{i=j}^{n_i}(y_{ij}-\mu_i)^2$$

Thus

$$\hat{\sigma}^2 = \frac{1}{n_i}\sum_{i=1}^{I}\sum_{i=j}^{n_i}(y_{ij}-\mu_i)^2$$

# QUESTION 7

We can write the distribution of the linearly transformed variable

$$(\bar{X}, X_1 - \bar{X}, X_2 - \bar{X}, \cdots, X_n - \bar{X})$$

as

$$\frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=j}^{n}(x_i - \bar{x})^2\right) \exp\left(\frac{n(\bar{x}-\mu)^2}{2\sigma^2}\right)$$

Not sure how to show $\bar{X}$ and $X_i - \bar{X}$ are independent.

But if we do that, then it follows $S_{XX}$ and $\bar{X}$ are independent, as $S_{XX}$ is a function of the $X_i - \bar{X}$.

# QUESTION 8

# QUESTION 9

# QUESTION 10

# QUESTION 11

# QUESTION 12