M2S1 - ASSESSED COURSEWORK 3 SOLUTIONS

(a) Solution uses methods of derivation/results from lectures.

(i) Most straightforward method uses a dummy variable ($Y_3 = X_3$ for example), to produce a 1-1 transformation, then to deduce the marginal distribution. Let matrix B be defined by

$$B = \begin{bmatrix} A \\ \underset{\sim}{a}^{\mathsf{T}} \end{bmatrix}$$

where $\underset{\sim}{a}$ is a 3×1 vector linearly independent of the rows of A; any vector will do, and such a vector is guaranteed to exist as the the rows of A themselves are linearly independent, so can set $\underset{\sim}{a}$ equal to their vector product. Now suppose $\underset{\sim}{Z} = B\underset{\sim}{X}$, with B non-singular. Note that $Z_1 = Y_1$ and $Z_2 = Y_2$.

Using the method from lectures, we have $\underset{\sim}{X} = B^{-1}\underset{\sim}{Z}$, and the exponent of the joint pdf for Z is given by

$$(B^{-1}\underset{\sim}{Z})^{\mathsf{T}}\Sigma^{-1}(B^{-1}\underset{\sim}{Z}) = \underset{\sim}{Z}^{\mathsf{T}}(B\Sigma B^{\mathsf{T}})^{-1}\underset{\sim}{Z}$$

and hence we can deduce immediately that

$$\underset{\sim}{Z} \sim N(0, B\Sigma B^{\mathsf{T}}).$$

But, using the block decomposition of Σ given in lectures

$$B\Sigma B^{\mathsf{T}} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \begin{bmatrix} B_{11}^{\mathsf{T}} & B_{21}^{\mathsf{T}} \\ B_{12}^{\mathsf{T}} & B_{22}^{\mathsf{T}} \end{bmatrix}.$$

The block of interest to us is the first block diagonal, that is

$$B_{11}\Sigma_{11}B_{11}^{\mathsf{T}} + B_{11}\Sigma_{12}B_{12}^{\mathsf{T}} + B_{12}\Sigma_{21}B_{11}^{\mathsf{T}} + B_{12}\Sigma_{22}B_{12}^{\mathsf{T}}$$

but this is merely $\Sigma_Y = A\Sigma A^{\mathsf{T}}$. Hence, by the result given in lectures concerning the marginal distribution of MVN variables, we have that

$$\underset{\sim}{Y} \sim N(0, \Sigma_Y) \equiv N(0, A\Sigma A^{\mathsf{T}}).$$

By direct calculation

$$A\Sigma A^{\mathsf{T}} = \begin{bmatrix} 86 & 45 \\ 45 & 34 \end{bmatrix}$$

[4 MARKS]

Note that this is a general result for transformed multivariate Normal random variables, and can be deduced from the results given in lectures in respect of the expectation and variance of transformed random variables.

(ii) The correlation, ρ, is available directly from Σ_Y:

$$\rho = \frac{Cov_{f_{Y_1,Y_2}}[Y_1, Y_2]}{\sqrt{Var_{f_{Y_1}}[Y_1]Var_{f_{Y_2}}[Y_2]}} = \frac{45}{\sqrt{86 \times 34}} = 0.832$$

[4 MARKS]

(b) Under the assumed model;

(i) We write

$$Y_n(x) = \sum_{i=1}^{n} I_i(X_i \leq x)$$

where $I_i(.)$ is the *indicator random variable*, with $I_i(A) = 1$ if event A occurs, and $I_i(A) = 0$ otherwise. $I_i(A)$ is a *Bernoulli* random variable with parameter θ given by

$$\theta = P[X_i \leq x] = F_X(x).$$

with $I_1 \ldots, I_n$ independent. Therefore, using the results for sums of independent *Bernoulli* random variables (or mgfs), we have that

$$Y_n(x) \sim Binomial(n, F_X(x))$$

[4 MARKS]

(ii) If $T_n(x) = Y_n(x)/n$, then as $n \longrightarrow \infty$

$$T_n(x) \xrightarrow{p} F_X(x)$$

by the Weak Law of Large Numbers. This is equivalent to saying that $T_n(x)$ *converges in probability* to $F_X(x)$, or that the limiting distribution of $T_n(x)$ is *degenerate* at $F_X(x)$.

A Normal approximation can be constructed for large but finite n using the CLT. Specifically for $Z \sim Binomial(n, \theta)$, $Z \overset{\mathcal{A}}{\sim} N(n\theta, n\theta(1-\theta))$, and hence

$$T_n(x) \overset{\mathcal{A}}{\sim} N(F_X(x)\ ,\ F_X(x)(1 - F_X(x))/n).$$

Need one or other result for full marks.

[2 MARKS]

(iii) With this cdf, $F_X(1) = 1/2$, so $Y_n(1) \sim Binomial(4, 1/2)$, and we have

y	0	1	2	3	4
$f_{Y_n(1)}(y)$	0.0625	0.2500	0.3750	0.2500	0.0625

[3 MARKS]

(iv) Here $F_X(3) = 3/4$, and

$$T_n(3) \overset{\mathcal{A}}{\sim} N(3/4\ ,\ (3/4)(1/4)/n) \equiv N(3/4, 3/(16n)).$$

[3 MARKS]