M2S1 - ASSESSED COURSEWORK 3

To be handed in no later than Friday, 17th December, 12.00pm.
Please hand in to the Mathematics General Office
as dictated by Departmental regulations.

1. Suppose that X_1, X_2 and X_3 are independent and identically distributed $Normal(0,1)$ variables. Denote by X the column vector $(X_1, X_2, X_3)^T$.

(a) Find the joint distribution of random variables $Y = (Y_1, Y_2, Y_3)^T$ defined by

$$Y_1 = X_1 + X_3 \qquad Y_2 = 2X_2 - X_3 \qquad Y_3 = 4X_3$$

or, in vector form

$$Y = AX \qquad \text{where} \qquad A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & -1 \\ 0 & 0 & 4 \end{pmatrix}.$$

[4 MARKS]

(b) Find the covariance between Y_1 and Y_3.

[4 MARKS]

(c) Find the marginal distribution of Y_1, where, of course,

$$Y_1 = BY$$

where B is the (1×3) matrix $(1, 0, 0)$.

[2 MARKS]

(d) The moment generating function (mgf) for vector random variable $Z = (Z_1, Z_2, Z_3)^T$ is defined as the multivariate expectation

$$M_Z(\mathbf{t}) = E_{f_Z}\left[\exp\left\{\mathbf{t}^T Z\right\}\right] = E_{f_Z}\left[\exp\left\{t_1 Z_1 + t_2 Z_2 + t_3 Z_3\right\}\right]$$

where $\mathbf{t} = (t_1, t_2, t_3)^T$ is the vector argument of the mgf (that is the mgf is a scalar function of a vector argument). Find the mgf of vector random variables X and Y as defined in part(a)

Hint: the calculation is more straightforward in vector/matrix form. Recall that X_1, X_2 and X_3 are independent.

[5 MARKS]

(e) Find the distribution (by name, or by pdf) of the random variable

$$V = Y^T \Sigma^{-1} Y$$

where $\Sigma = AA^T$.

[5 MARKS]
PTO

2. Suppose that $X_1, ..., X_n$ are independent and identically distributed continuous random variables on $\mathbb{R}^+$ with cdf F_X. The **maximum order statistic** derived from these variables is defined as

$$Y_n = \max\{X_1, ..., X_n\}$$

and it can be shown easily that

$$F_{Y_n}(y) = \{F_X(y)\}^n \qquad y > 0.$$

and is zero for $y \leq 0$.

(a) Suppose that

$$F_X(x) = \frac{(x+1)^2 - 1}{(x+1)^2} \qquad 0 < x < \infty$$

with $F_X(x) = 0$ for $x \leq 0$.

(i) Show that in this case Y_n has **no limiting distribution**, that is the limiting function defined pointwise by

$$F(y) = \lim_{n\to\infty} F_{Y_n}(y)$$

for $y \in \mathbb{R}$ is **not** a probability distribution function.

[5 MARKS]

(ii) Consider the transformed variable

$$Z_n = \frac{1}{\sqrt{n}} Y_n.$$

Show that Z_n **does** have a limiting distribution, that is, that is the limiting function defined pointwise as in (i) **is** a probability distribution function.

[5 MARKS]

(b) Now suppose that $X_1, ...X_n$ are independent *Exponential* (1) random variables.

(i) Find the distribution function for $Y_n = \max\{X_1, ..., X_n\}$.

[2 MARKS]

(ii) Find the distribution function for $Z_n = Y_n - a$ for constant a (remember to state the range of Z_n).

[2 MARKS]

(iii) Find a sequence of constants $\{a_n\}$ such that, for $z > 0$,

$$\lim_{n\to\infty} F_{Z_n}(z) = \exp\{-\exp\{-z\}\}$$

and hence find an approximation to the probability

$$P[Y_n > y_0]$$

for large n, for fixed $y_0 > 0$.

[2 MARKS]

(iv) The times between earthquakes of a certain magnitude that occur in Southern California are assumed to be independent Exponential random variables with rate parameter $\lambda = 0.01$ (per day), so that the expected time between earthquakes is 100 days. Over a given recording period, $n = 200$ earthquakes were observed, and the longest time between any two consecutive earthquakes was 1021 days. Comment on the validity of the Exponential distribution assumption in the light of this data.

[4 MARKS]