Imperial College
London

[E 2.9 (Maths 4) 2011]

B.ENG. AND M.ENG. EXAMINATIONS 2011

PART II Paper 4 : MATHEMATICS (ELECTRICAL ENGINEERING)

Date Thursday 9th June 2011 2.00 - 4.00 pm

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO.

Answer FOUR questions.

Please answer questions from Section A and Section B in separate answer-books.

A mathematical formulae sheet is provided.

Statistical data sheets are provided.

[Before starting, please make sure that the paper is complete; there should be 6 pages, with a total of 6 questions. Ask the invigilator for a replacement if your copy is faulty.]

SECTION A

1. The normalized eigenvectors $\boldsymbol{a}_i$ corresponding to the distinct eigenvalues of an $n \times n$ real symmetric matrix A are put together to form the $n \times n$ matrix

$$P = \left(\boldsymbol{a}_1\ \boldsymbol{a}_2\ \boldsymbol{a}_3 \ldots \boldsymbol{a}_n\right).$$

Use the orthonormality property

$$\boldsymbol{a}_i^T \boldsymbol{a}_j = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

to show that P satisfies

$$P^{-1} = P^T.$$

Find the eigenvalues and normalized eigenvectors of the matrix

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Using these, show that the matrix

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \\ \sqrt{2} & 0 & 0 \end{pmatrix}$$

diagonalises A such that

$$P^{-1}AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

2. Show that the quadratic form

$$Q = 4x_1^2 + 4\,x_1x_2 + x_2^2 + 4x_3^2$$

can be written as

$$Q = \boldsymbol{x}^T A \boldsymbol{x}\,,$$

where $\boldsymbol{x} = (x_1, x_2, x_3)^T$ and A is a real symmetric matrix, which is to be found.

Hence show that Q can be re-expressed in the diagonal form

$$Q = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \lambda_3 y_3^2\,,$$

where the λ_i are to be determined, by finding a matrix P that satisfies $\boldsymbol{x} = P\boldsymbol{y}$ where $\boldsymbol{y} = (y_1, y_2, y_3)^T$.

Find y_1, y_2 and y_3 in terms of x_1, x_2 and x_3 from the matrix P.

SECTION B

3. A screening test has been developed for a medical condition which occurs in 1 in every 1000 of the population. The result of the test may be positive, negative or inconclusive.

   For someone with the condition, the probability of a positive test result is 0.94 and of a negative result 0.05. For someone without the condition, the probability of a positive result is 0.02 and of a negative result 0.93.

   (i) We test a random person from the population. What is the probability their test result is positive? What is the probability the test is inconclusive?

   (ii) Someone is tested and found positive. Find the probability that they have the condition.

   (iii) Everyone whose test result is inconclusive is tested again. Tests are repeated as often as necessary until a clear (positive or negative) result is obtained. What is the distribution of X, the number of tests required? Assume that the outcomes of different tests are independent.

   (iv) In a population of 1 million, how many people do you expect will require more than three tests?

4. The continuous random variable Y has probability density function (PDF) given by

$$f_Y(y) = \begin{cases} \theta^{-2} y e^{-y/\theta} & y > 0 \\ 0 & \text{otherwise} \end{cases} .$$

   (i) Verify that $f_Y(y)$ is a valid PDF. *(Hint: You need to integrate by parts.)*

   (ii) Find the moment generating function of Y and hence, or otherwise, compute $E(Y)$ and $\mathrm{Var}(Y)$.

   (iii) Suppose that we draw a random sample $Y_1, Y_2, \ldots, Y_n$ from this distribution. Find the method of moments estimator of θ. Is it unbiased?

   (iv) Show that the maximum-likelihood estimator of θ is $\hat{\theta} = \bar{Y}/2$ and compute its mean square error.

PLEASE TURN OVER

5. (i) Two football teams, Team X and Team Y, are about to play a match against each other. Let X and Y denote the number of goals scored by Team X and Team Y, respectively. The joint probability function of X and Y is:

		Y			
	Goals	0	1	2	3
X	0	0.25	0.12	0.05	0.02
	1	0.10	0.09	0.06	0.02
	2	0.07	0.06	0.04	0.03
	3	0.02	0.02	0.03	0.02

(a) What is the probability that Team X wins the match? What is the probability of a draw?

(b) Find the marginal distributions of X and Y, and compute $E(X)$, $E(Y)$.

(c) Compute $\mathrm{Var}(X)$, $\mathrm{Var}(Y)$ and $\mathrm{Cov}(X, Y)$. Are X and Y uncorrelated?

(d) Write down the conditional distribution of Y given that Team X does not score. Are X and Y independent?

(ii) A system consists of k independent components in parallel, i.e. it functions as long as at least one component functions. The components are unreliable: they fail with probability 0.20.

(a) What is the probability that the system functions?

(b) What is the minimum number of components required to ensure that the probability the system fails is less than 1 in 1000 (i.e. 0.1%)?

(iii) Let T be a random variable with range $[1, \infty)$ and hazard rate

$$h(t) = \frac{\alpha}{t}$$

for some $\alpha > 0$.

Find the cumulative hazard function, the cumulative distribution function (CDF) and the probability density function (PDF) of T.

6. Nitrogen oxide (NO_x) emissions from cars in inner cities are known to contribute to the formation of ozone smog, which is an important cause of health problems. For this reason, European standards define acceptable NO_x exhaust emission limits for all new cars sold in EU member states. The current limit for passenger cars is set at 0.97 g/km.

A car manufacturer wishes to release a new model on the European market, but needs to prove that the car meets emission standards. In 12 independent measurements, the NO_x emissions of the new model were (in g/km):

0.93 0.85 0.95 0.89 1.05 0.83 1.04 1.09 0.83 1.02 0.91 0.87

Some useful summary statistics:

$$\sum_{i=1}^{12} x_i = 11.26\,, \quad \sum_{i=1}^{12} x_i^2 = 10.657,$$

where x_i is the ith measurement.

(i) Compute the sample mean, median and sample variance of this dataset.

(ii) Let X denote the emissions (in g/km) of a randomly chosen car. Construct a 95% confidence interval for the mean of X. Is it necessary to assume that X is normally distributed?

(iii) On the basis of this confidence interval, would you say that the car meets the emission standards?

After a modification to the catalytic converter, 13 new measurements are taken:

0.80 0.92 0.83 0.84 0.84 0.82 0.97 0.78 1.09 0.75 0.88 0.75 0.77

Some useful summary statistics:

$$\sum_{i=1}^{13} y_i = 11.04\,, \quad \sum_{i=1}^{13} y_i^2 = 9.489,$$

where y_i is the ith new measurement.

(iv) Does the modification seem to reduce emissions? Conduct a hypothesis test at two levels of significance. Assume equal variances for the distributions of the x's and the y's.

END OF PAPER

MATHEMATICS DEPARTMENT

MATHEMATICAL FORMULAE

1. VECTOR ALGEBRA

$$\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} = (a_1,\, a_2,\, a_3)$$

Scalar (dot) product: $\mathbf{a}\,.\,\mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3$

Vector (cross) product:

$$\mathbf{a}\times\mathbf{b} \;=\; \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

Scalar triple product:

$$[\mathbf{a},\,\mathbf{b},\,\mathbf{c}] = \mathbf{a}\,.\,\mathbf{b}\times\mathbf{c} = \mathbf{b}\,.\,\mathbf{c}\times\mathbf{a} = \mathbf{c}\,.\,\mathbf{a}\times\mathbf{b} = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

Vector triple product: $\mathbf{a}\times(\mathbf{b}\times\mathbf{c}) = (\mathbf{c}\,.\,\mathbf{a})\mathbf{b} - (\mathbf{b}\,.\,\mathbf{a})\mathbf{c}$

2. SERIES

$$(1+x)^\alpha = 1 + \alpha\,x + \frac{\alpha(\alpha-1)}{2!}\,x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}\,x^3 + \ldots \quad (\alpha \text{ arbitrary},\; |x|<1)$$

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots\;,$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots + (-1)^n\frac{x^{2n}}{(2n)!} + \ldots\;,$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n\frac{x^{2n+1}}{(2n+1)!} + \ldots\;,$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^n\frac{x^{n+1}}{(n+1)} + \ldots (-1 < x \le 1)$$

3. TRIGONOMETRIC IDENTITIES AND HYPERBOLIC FUNCTIONS

$$\sin(a+b) = \sin a\,\cos b + \cos a\,\sin b\;;$$

$$\cos(a+b) = \cos a\,\cos b - \sin a\,\sin b\,.$$

$$\cos iz = \cosh z\;;\quad \cosh iz = \cos z\;;\quad \sin iz = i\sinh z\;;\quad \sinh iz = i\sin z\,.$$

4. DIFFERENTIAL CALCULUS

(a) Leibniz's formula:

$$D^n(fg) = f\,D^n g + \binom{n}{1}\,Df\,D^{n-1}\,g + \ldots + \binom{n}{r}\,D^r f\,D^{n-r} g + \ldots + D^n f\;g\,.$$

(b) Taylor's expansion of $f(x)$ about $x = a$:

$$f(a+h) = f(a) + hf'(a) + h^2 f''(a)/2! + \ldots + h^n f^{(n)}(a)/n! + \epsilon_n(h)\,,$$

where $\quad \epsilon_n(h) = h^{n+1} f^{(n+1)}(a+\theta h)/(n+1)!, \quad 0 < \theta < 1\,.$

(c) Taylor's expansion of $f(x,\,y)$ about $(a,\,b)$:

$$f(a+h,\,b+k) = f(a,\,b) + [hf_x + kf_y]_{a.b} + 1/2!\,\left[h^2 f_{xx} + 2hkf_{xy} + k^2 f_{yy}\right]_{a.b} + \cdots$$

(d) Partial differentiation of $f(x,\,y)$:

i. If $y = y(x)$, then $f = F(x)$, and $\dfrac{dF}{dx} = \dfrac{\partial f}{\partial x} + \dfrac{\partial f}{\partial y}\dfrac{dy}{dx}\,.$

ii. If $x = x(t),\ y = y(t)$, then $f = F(t)$, and $\dfrac{dF}{dt} = \dfrac{\partial f}{\partial x}\dfrac{dx}{dt} + \dfrac{\partial f}{\partial y}\dfrac{dy}{dt}\,.$

iii. If $x = x(u,\,v),\ y = y(u,\,v)$, then $f = F(u,\,v)$, and

$$\frac{\partial F}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}\,,\quad \frac{\partial F}{\partial v} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial v} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial v}\,.$$

(e) Stationary points of $f(x,\,y)$ occur where $f_x = 0,\ f_y = 0$ simultaneously.
Let $(a,\,b)$ be a stationary point: examine $D = [f_{xx}f_{yy} - (f_{xy})^2]_{a.b}$.
If $D > 0$ and $f_{xx}(a,\,b) < 0$, then $(a,\,b)$ is a maximum;
If $D > 0$ and $f_{xx}(a,\,b) > 0$, then $(a,\,b)$ is a minimum;
If $D < 0$ then $(a,\,b)$ is a saddle-point.

(f) Differential equations:

i. The first order linear equation $dy/dx + P(x)y = Q(x)$ has an integrating factor $\quad I(x) = \exp[\int P(x)(dx]$, so that $\dfrac{d}{dx}(Iy) = IQ$.

ii. $P(x,\,y)dx + Q(x,\,y)dy = 0$ is exact if $\partial Q/\partial x = \partial P/\partial y\,.$

5. INTEGRAL CALCULUS

(a) An important substitution: $\tan(\theta/2) = t$:
$\sin\theta = 2\,t/(1+t^2)$, $\cos\theta = (1-t^2)/(1+t^2)$, $d\theta = 2\,dt/(1+t^2)$.

(b) Some indefinite integrals:

$$\int (a^2 - x^2)^{-1/2} dx = \sin^{-1}\left(\frac{x}{a}\right), \quad |x| < a\,.$$

$$\int (a^2 + x^2)^{-1/2} dx = \sinh^{-1}\left(\frac{x}{a}\right) = \ln\left\{\frac{x}{a} + \left(1 + \frac{x^2}{a^2}\right)^{1/2}\right\}.$$

$$\int (x^2 - a^2)^{-1/2} dx = \cosh^{-1}\left(\frac{x}{a}\right) = \ln\left|\frac{x}{a} + \left(\frac{x^2}{a^2} - 1\right)^{1/2}\right|.$$

$$\int (a^2 + x^2)^{-1} dx = \left(\frac{1}{a}\right)\tan^{-1}\left(\frac{x}{a}\right).$$

6. NUMERICAL METHODS

(a) Approximate solution of an algebraic equation:

If a root of $f(x) = 0$ occurs near $x = a$, take $x_0 = a$ and
$x_{n+1} = x_n - [f(x_n)/f'(x_n)]$, $n = 0, 1, 2 \ldots$

(Newton Raphson method).

(b) Formulae for numerical integration: Write $x_n = x_0 + nh$, $y_n = y(x_n)$.

i. Trapezium rule (1-strip): $\int_{x_0}^{x_1} y(x)dx \approx (h/2)\,[y_0 + y_1]$.

ii. Simpson's rule (2-strip): $\int_{x_0}^{x_2} y(x)dx \approx (h/3)\,[y_0 + 4y_1 + y_2]$.

(c) Richardson's extrapolation method: Let $I = \int_a^b f(x)dx$ and let I_1, I_2 be two

estimates of I obtained by using Simpson's rule with intervals h and $h/2$.

Then, provided h is small enough,

$$I_2 + (I_2 - I_1)/15\,,$$

is a better estimate of I.

7. LAPLACE TRANSFORMS

Function	Transform	Function	Transform
$f(t)$	$F(s) = \int_0^\infty e^{-st} f(t)dt$	$af(t) + bg(t)$	$aF(s) + bG(s)$
df/dt	$sF(s) - f(0)$	d^2f/dt^2	$s^2F(s) - sf(0) - f'(0)$
$e^{at}f(t)$	$F(s-a)$	$tf(t)$	$-dF(s)/ds$
$(\partial/\partial\alpha)f(t, \alpha)$	$(\partial/\partial\alpha)F(s, \alpha)$	$\int_0^t f(t)dt$	$F(s)/s$
$\int_0^t f(u)g(t-u)du$	$F(s)G(s)$		
1	$1/s$	$t^n (n = 1, 2\ldots)$	$n!/s^{n+1}$, $(s > 0)$
e^{at}	$1/(s-a)$, $(s > a)$	$\sin\omega t$	$\omega/(s^2+\omega^2)$, $(s > 0)$
$\cos\omega t$	$s/(s^2+\omega^2)$, $(s > 0)$	$H(t-T) = \begin{cases} 0, & t < T \\ 1, & t > T \end{cases}$	e^{-sT}/s, $(s, T > 0)$

8. FOURIER SERIES

If $f(x)$ is periodic of period $2L$, then $f(x+2L) = f(x)$, and

$$f(x) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi x}{L} + \sum_{n=1}^{\infty} b_n \sin\frac{n\pi x}{L}, \text{ where}$$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}dx, \quad n = 0, 1, 2, \ldots, \text{ and}$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}dx, \quad n = 1, 2, 3, \ldots$$

Parseval's theorem

$$\frac{1}{L}\int_{-L}^{L} [f(x)]^2\, dx = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}\left(a_n^2 + b_n^2\right).$$

1. Probabilities for events

For events A, B, and C $\quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$

More generally $\quad P(\bigcup A_i) = \sum P(A_i) - \sum P(A_i \cap A_j) + \sum P(A_i \cap A_j \cap A_k) - \cdots$

The odds in favour of A $\quad P(A) \,/\, P(\overline{A})$

Conditional probability $\quad P(A \mid B) = \dfrac{P(A \cap B)}{P(B)}$ provided that $P(B) > 0$

Chain rule $\quad P(A \cap B \cap C) = P(A)\,P(B \mid A)\,P(C \mid A \cap B)$

Bayes' rule $\quad P(A \mid B) = \dfrac{P(A)\,P(B \mid A)}{P(A)\,P(B \mid A) + P(\overline{A})\,P(B \mid \overline{A})}$

A and B are independent if $\quad P(B \mid A) = P(B)$

A, B, and C are independent if $\quad P(A \cap B \cap C) = P(A)P(B)P(C)$, and

$$P(A \cap B) = P(A)P(B), \quad P(B \cap C) = P(B)P(C), \quad P(C \cap A) = P(C)P(A)$$

2. Probability distribution, expectation and variance

The probability distribution for a discrete random variable X is called the probability mass function (pmf) and is the complete set of probabilities $\{p_x\} = \{P(X = x)\}$

Expectation $\quad E(X) = \mu = \sum_x x p_x$

For function $g(x)$ of x, $\quad E\{g(X)\} = \sum_x g(x) p_x$, so $E(X^2) = \sum_x x^2 p_x$

Sample mean $\quad \overline{x} = \dfrac{1}{n} \sum_k x_k$ estimates μ from random sample $x_1, x_2, \ldots, x_n$

Variance $\quad \mathrm{var}\,(X) = \sigma^2 = E\{(X - \mu)^2\} = E(X^2) - \mu^2$

Sample variance $\quad s^2 = \dfrac{1}{n-1} \left\{ \sum_k x_k^2 - \dfrac{1}{n} \left(\sum_j x_j \right)^2 \right\}$ estimates σ^2

Standard deviation $\quad \mathrm{sd}\,(X) = \sigma$

If value y is observed with frequency n_y

$$n = \sum_y n_y, \quad \sum_k x_k = \sum_y y n_y, \quad \sum_k x_k^2 = \sum_y y^2 n_y$$

Skewness $\quad \beta_1 = E\left(\dfrac{X - \mu}{\sigma}\right)^3$ is estimated by $\dfrac{1}{n-1} \sum \left(\dfrac{x_i - \overline{x}}{s}\right)^3$

Kurtosis $\quad \beta_2 = E\left(\dfrac{X - \mu}{\sigma}\right)^4 - 3$ is estimated by $\dfrac{1}{n-1} \sum \left(\dfrac{x_i - \overline{x}}{s}\right)^4 - 3$

Sample median $\quad \tilde{x}$ or x_{med}. Half the sample values are smaller and half larger

If the sample values $x_1, \ldots, x_n$ are ordered as $x_{(1)} \le x_{(2)} \le \cdots \le x_{(n)}$, then $\tilde{x} = x_{(\frac{n+1}{2})}$ if n is odd, and $\tilde{x} = \frac{1}{2}\left(x_{(\frac{n}{2})} + x_{(\frac{n+2}{2})}\right)$ if n is even

α-quantile $Q(\alpha)$ is such that $P(X \leq Q(\alpha)) = \alpha$

Sample α-quantile $\widehat{Q}(\alpha)$ Proportion α of the data values are smaller

Lower quartile Q1 $= \widehat{Q}(0.25)$ one quarter are smaller

Upper quartile Q3 $= \widehat{Q}(0.75)$ three quarters are smaller

Sample median $\widetilde{x} = \widehat{Q}(0.5)$ estimates the population median $Q(0.5)$

3. Probability distribution for a continuous random variable

The cumulative distribution function (cdf) $F(x) = P(X \leq x) = \int_{x_0=-\infty}^{x} f(x_0)\mathrm{d}x_0$

The probability density function (pdf) $f(x) = \dfrac{\mathrm{d}F(x)}{\mathrm{d}x}$

$E(X) = \mu = \int_{-\infty}^{\infty} x\, f(x)\mathrm{d}x$, $\mathrm{var}\,(X) = \sigma^2 = E(X^2) - \mu^2$, where $E(X^2) = \int_{-\infty}^{\infty} x^2\, f(x)\mathrm{d}x$

4. Discrete probability distributions

Discrete Uniform *Uniform* (n)

$$p_x = \frac{1}{n} \quad (x = 1, 2, \ldots, n) \qquad \mu = (n+1)/2,\ \sigma^2 = (n^2-1)/12$$

Binomial distribution *Binomial* (n, θ)

$$p_x = \binom{n}{x} \theta^x (1-\theta)^{n-x} \quad (x = 0, 1, 2, \ldots, n) \qquad \mu = n\theta,\ \sigma^2 = n\theta(1-\theta)$$

Poisson distribution *Poisson* (λ)

$$p_x = \frac{\lambda^x e^{-\lambda}}{x!} \quad (x = 0, 1, 2, \ldots) \quad \text{(with } \lambda > 0) \qquad \mu = \lambda,\ \sigma^2 = \lambda$$

Geometric distribution *Geometric* (θ)

$$p_x = (1-\theta)^{x-1}\theta \quad (x = 1, 2, 3, \ldots) \qquad \mu = \frac{1}{\theta},\ \sigma^2 = \frac{1-\theta}{\theta^2}$$

5. Continuous probability distributions

Uniform distribution *Uniform* (α, β)

$$f(x) = \begin{cases} \dfrac{1}{\beta - \alpha} & (\alpha < x < \beta), \\ 0 & \text{(otherwise)}. \end{cases} \qquad \mu = (\alpha+\beta)/2,\ \sigma^2 = (\beta-\alpha)^2/12$$

Exponential distribution *Exponential* (λ)

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & (0 < x < \infty), \\ 0 & (-\infty < x \leq 0). \end{cases} \qquad \mu = 1/\lambda,\ \sigma^2 = 1/\lambda^2$$

Normal distribution $N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\} \quad (-\infty < x < \infty), \quad E(X) = \mu, \quad \mathrm{var}(X) = \sigma^2$$

Standard normal distribution $N(0,1)$

If X is $N(\mu, \sigma^2)$, then $Y = \dfrac{X-\mu}{\sigma}$ is $N(0,1)$

6. Reliability

For a device in continuous operation with failure time random variable T having pdf $f(t)$ $(t > 0)$

The reliability function at time t	$R(t)$	$=$	$P(T > t)$
The failure rate or hazard function	$h(t)$	$=$	$f(t)/R(t)$
The cumulative hazard function	$H(t)$	$=$	$\int_0^t h(t_0)\,\mathrm{d}t_0 = -\ln\{R(t)\}$
The Weibull(α, β) distribution has	$H(t)$	$=$	βt^α

7. System reliability

For a system of k devices, which operate independently, let

$R_i = P(D_i) = P($"device i operates"$)$

The system reliability, R, is the probability of a path of operating devices

A system of devices in series operates only if every device operates

$$R = P(D_1 \cap D_2 \cap \cdots \cap D_k) = R_1 R_2 \cdots R_k$$

A system of devices in parallel operates if any device operates

$$R = P(D_1 \cup D_2 \cup \cdots \cup D_k) = 1 - (1-R_1)(1-R_2)\cdots(1-R_k)$$

8. Covariance and correlation

The covariance of X and Y $\quad \mathrm{cov}(X,Y) = E(XY) - \{E(X)\}\{E(Y)\}$

From pairs of observations $(x_1, y_1), \ldots, (x_n, y_n)$ $\quad S_{xy} = \sum_k x_k y_k - \frac{1}{n}(\sum_i x_i)(\sum_j y_j)$

$$S_{xx} = \sum_k x_k^2 - \frac{1}{n}(\sum_i x_i)^2, \qquad S_{yy} = \sum_k y_k^2 - \frac{1}{n}(\sum_j y_j)^2$$

Sample covariance	$s_{xy} = \dfrac{1}{n-1} S_{xy}$		estimates $\mathrm{cov}(X,Y)$
Correlation coefficient	$\rho = \mathrm{corr}(X,Y)$	$= \dfrac{\mathrm{cov}(X,Y)}{\mathrm{sd}(X) \cdot \mathrm{sd}(Y)}$	
Sample correlation coefficient	$r = \dfrac{S_{xy}}{\sqrt{S_{xx} S_{yy}}}$		estimates ρ

9. Sums of random variables

$$
\begin{aligned}
E(X+Y) &= E(X)+E(Y)\\
\operatorname{var}(X+Y) &= \operatorname{var}(X)+\operatorname{var}(Y)+2\operatorname{cov}(X,Y)\\
\operatorname{cov}(aX+bY,\ cX+dY) &= (ac)\operatorname{var}(X)+(bd)\operatorname{var}(Y)+(ad+bc)\operatorname{cov}(X,Y)
\end{aligned}
$$

If X is $N(\mu_1, \sigma_1^2)$, Y is $N(\mu_2, \sigma_2^2)$, and $\operatorname{cov}(X,Y)=c$, then $X+Y$ is $N(\mu_1+\mu_2,\ \sigma_1^2+\sigma_2^2+2c)$

10. Bias, standard error, mean square error

If t estimates θ (with random variable T giving t)

Bias of t $\quad \operatorname{bias}(t) = E(T)-\theta$

Standard error of t $\quad \operatorname{se}(t) = \operatorname{sd}(T)$

Mean square error of t $\quad \mathsf{MSE}(t) = E\{(T-\theta)^2\} = \{\operatorname{se}(t)\}^2+\{\operatorname{bias}(t)\}^2$

If $\overline{x}$ estimates μ, then $\operatorname{bias}(\overline{x})=0$, $\operatorname{se}(\overline{x})=\sigma/\sqrt{n}$, $\mathsf{MSE}(\overline{x})=\sigma^2/n$, $\widehat{\operatorname{se}}(\overline{x})=s/\sqrt{n}$

Central limit property $\quad$ If n is fairly large, $\overline{x}$ is from $N(\mu,\ \sigma^2/n)$ approximately

11. Likelihood

The likelihood is the joint probability as a function of the unknown parameter θ.

For a random sample $x_1, x_2, \ldots, x_n$

$$
\begin{aligned}
\ell(\theta;\, x_1,x_2,\ldots,x_n) &= P(X_1=x_1 \,|\, \theta)\cdots P(X_n=x_n \,|\, \theta) && \text{(discrete distribution)}\\
\ell(\theta;\, x_1,x_2,\ldots,x_n) &= f(x_1 \,|\, \theta)\, f(x_2 \,|\, \theta)\cdots f(x_n \,|\, \theta) && \text{(continuous distribution)}
\end{aligned}
$$

The maximum likelihood estimator (MLE) is $\widehat{\theta}$ for which the likelihood is a maximum

12. Confidence intervals

If $x_1, x_2, \ldots, x_n$ are a random sample from $N(\mu, \sigma^2)$ and σ^2 is known, then the 95% confidence interval for μ is $(\overline{x}-1.96\frac{\sigma}{\sqrt{n}},\ \overline{x}+1.96\frac{\sigma}{\sqrt{n}})$

If σ^2 is estimated, then from the Student t table for t_{n-1} we find $t_0 = t_{n-1,0.05}$

The 95% confidence interval for μ is $(\overline{x}-t_0\frac{s}{\sqrt{n}},\ \overline{x}+t_0\frac{s}{\sqrt{n}})$

13. Standard normal table — Values of pdf $\phi(y) = f(y)$ and cdf $\Phi(y) = F(y)$

y	$\phi(y)$	$\Phi(y)$	y	$\phi(y)$	$\Phi(y)$	y	$\phi(y)$	$\Phi(y)$	y	$\Phi(y)$
0	.399	.5	.9	.266	.816	1.8	.079	.964	2.8	.997
.1	.397	.540	1.0	.242	.841	1.9	.066	.971	3.0	.999
.2	.391	.579	1.1	.218	.864	2.0	.054	.977	0.841	.8
.3	.381	.618	1.2	.194	.885	2.1	.044	.982	1.282	.9
.4	.368	.655	1.3	.171	.903	2.2	.035	.986	1.645	.95
.5	.352	.691	1.4	.150	.919	2.3	.028	.989	1.96	.975
.6	.333	.726	1.5	.130	.933	2.4	.022	.992	2.326	.99
.7	.312	.758	1.6	.111	.945	2.5	.018	.994	2.576	.995
.8	.290	.788	1.7	.094	.955	2.6	.014	.995	3.09	.999

14. Student t table — Values $t_{m,p}$ of x for which $P(|X| > x) = p$, when X is t_m

m	p= 0.10	0.05	0.02	0.01	m	p= 0.10	0.05	0.02	0.01
1	6.31	12.71	31.82	63.66	9	1.83	2.26	2.82	3.25
2	2.92	4.30	6.96	9.92	10	1.81	2.23	2.76	3.17
3	2.35	3.18	4.54	5.84	12	1.78	2.18	2.68	3.05
4	2.13	2.78	3.75	4.60	15	1.75	2.13	2.60	2.95
5	2.02	2.57	3.36	4.03	20	1.72	2.09	2.53	2.85
6	1.94	2.45	3.14	3.71	25	1.71	2.06	2.48	2.78
7	1.89	2.36	3.00	3.50	40	1.68	2.02	2.42	2.70
8	1.86	2.31	2.90	3.36	∞	1.645	1.96	2.326	2.576

15. Chi-squared table — Values $\chi^2_{k,p}$ of x for which $P(X > x) = p$, when X is χ^2_k and $p = .995, .975,$ *etc*

k	.995	.975	.05	.025	.01	.005
1	.000	.001	3.84	5.02	6.63	7.88
2	.010	.051	5.99	7.38	9.21	10.60
3	.072	.216	7.81	9.35	11.34	12.84
4	.207	.484	9.49	11.14	13.28	14.86
5	.412	.831	11.07	12.83	15.09	16.75
6	.676	1.24	12.59	14.45	16.81	18.55
7	.990	1.69	14.07	16.01	18.48	20.28
8	1.34	2.18	15.51	17.53	20.09	21.95
9	1.73	2.70	16.92	19.02	21.67	23.59
10	2.16	3.25	13.31	20.48	23.21	25.19
12	3.07	4.40	21.03	23.34	26.22	28.30
14	4.07	5.63	23.68	26.12	29.14	31.32
16	5.14	6.91	26.30	28.85	32.00	34.27

k	.995	.975	.05	.025	.01	.005
18	6.26	8.23	28.87	31.53	34.81	37.16
20	7.43	9.59	31.42	34.17	37.57	40.00
22	8.64	10.98	33.92	36.78	40.29	42.80
24	9.89	12.40	36.42	39.36	42.98	45.56
26	11.16	13.84	38.89	41.92	45.64	48.29
28	12.46	15.31	41.34	44.46	48.28	50.99
30	13.79	16.79	43.77	46.98	50.89	53.67
40	20.71	24.43	55.76	59.34	63.69	66.77
50	27.99	32.36	67.50	71.41	76.15	79.49
60	35.53	40.48	79.08	83.30	88.38	91.95
70	43.28	48.76	90.53	95.02	100.4	104.2
80	51.17	57.15	101.9	106.6	112.3	116.3
100	67.33	74.22	124.3	129.6	135.8	140.2

16. $\boxed{\text{The chi-squared goodness-of-fit test}}$

The frequencies n_y are grouped so that the fitted frequency $\widehat{n}_y$ for every group exceeds about 5.

$X^2 = \sum_y \frac{(n_y - \widehat{n}_y)^2}{\widehat{n}_y}$ is referred to the table of χ^2_k with significance point p,

where k is the number of terms summed, less one for each constraint, *eg* matching total frequency, and matching $\overline{x}$ with μ

17. $\boxed{\text{Joint probability distributions}}$

Discrete distribution $\{p_{xy}\}$, where $p_{xy} = P(\{X = x\} \cap \{Y = y\})$.

Let $p_{x\bullet} = P(X = x)$, and $p_{\bullet y} = P(Y = y)$, then

$$p_{x\bullet} = \sum_y p_{xy} \quad \text{and} \quad P(X = x \mid Y = y) = \frac{p_{xy}}{p_{\bullet y}}$$

$\boxed{\text{Continuous distribution}}$

Joint cdf $\quad F(x,y) = P(\{X \le x\} \cap \{Y \le y\}) = \int_{x_0=-\infty}^{x} \int_{y_0=-\infty}^{y} f(x_0, y_0)\, \mathrm{d}x_0\, \mathrm{d}y_0$

Joint pdf $\quad f(x,y) = \frac{\mathrm{d}^2 F(x,y)}{\mathrm{d}x\, \mathrm{d}y}$

Marginal pdf of X $\quad f_X(x) = \int_{-\infty}^{\infty} f(x, y_0)\, \mathrm{d}y_0$

Conditional pdf of X given $Y = y$ $\quad f_{X|Y}(x|y) = \frac{f(x,y)}{f_Y(y)}$ (provided $f_Y(y) > 0$)

18. $\boxed{\text{Linear regression}}$

To fit the linear regression model $y = \alpha + \beta x$ by $\widehat{y}_x = \widehat{\alpha} + \widehat{\beta} x$ from observations $(x_1, y_1), \ldots, (x_n, y_n)$, the least squares fit is $\quad \widehat{\alpha} = \overline{y} - \overline{x}\widehat{\beta}, \quad \widehat{\beta} = \frac{S_{xy}}{S_{xx}}$

The residual sum of squares $\quad \text{RSS} = S_{yy} - \frac{S_{xy}^2}{S_{xx}}$

$\widehat{\sigma^2} = \frac{\text{RSS}}{n-2} \qquad \frac{n-2}{\sigma^2}\widehat{\sigma^2}$ is from χ^2_{n-2}

$$E(\widehat{\alpha}) = \alpha, \quad E(\widehat{\beta}) = \beta,$$

$$\text{var}(\widehat{\alpha}) = \frac{\sum x_i^2}{n\, S_{xx}} \sigma^2, \quad \text{var}(\widehat{\beta}) = \frac{\sigma^2}{S_{xx}}, \quad \text{cov}(\widehat{\alpha}, \widehat{\beta}) = -\frac{\overline{x}}{S_{xx}} \sigma^2$$

$$\widehat{y}_x = \widehat{\alpha} + \widehat{\beta} x, \qquad E(\widehat{y}_x) = \alpha + \beta x, \qquad \text{var}(\widehat{y}_x) = \left\{ \frac{1}{n} + \frac{(x - \overline{x})^2}{S_{xx}} \right\} \sigma^2$$

$\frac{\widehat{\alpha} - \alpha}{\widehat{\text{se}}(\widehat{\alpha})}, \quad \frac{\widehat{\beta} - \beta}{\widehat{\text{se}}(\widehat{\beta})}, \quad \frac{\widehat{y}_x - \alpha - \beta x}{\widehat{\text{se}}(\widehat{y}_x)}$ are each from t_{n-2}

E2.9 (Maths 4)

	EXAMINATION QUESTIONS/SOLUTIONS 2010-11	Course EE2 (1) Paper 4
Question 4 Q1		Marks & seen/unseen

Parts

If $A\underline{a}_i = \lambda_i \underline{a}_i$, construct $P = (\underline{a}_1\, \underline{a}_2\, \underline{a}_3 \cdots)$ & form. — seen

$$P^T P = \begin{pmatrix} \underline{a}_1^T \\ \underline{a}_2^T \\ \underline{a}_3^T \\ \vdots \end{pmatrix} (\underline{a}_1\, \underline{a}_2\, \underline{a}_3 \cdots) = \begin{pmatrix} \underline{a}_1^T \underline{a}_1 & \underline{a}_1^T \underline{a}_2 & \cdots \\ \vdots & \underline{a}_2^T \underline{a}_2 & \\ \vdots & & \ddots \\ \vdots & & & \underline{a}_n^T \underline{a}_n \end{pmatrix}$$ 3

$$= \{\underline{a}_i^T \underline{a}_j\} = \{\delta_{ij}\} = I.$$ 2

using orthonormality property.

$\therefore P^T = P^{-1}$ 1

unseen

For $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ symm $\lambda_1 = 4$ & $(\lambda - 2)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0$

$\lambda_2 = 3 \quad \lambda_3 = 1$ — 3 for evals

$(0, 0, 1)_{\lambda_1 = 4}$; $\frac{1}{\sqrt{2}}(1, 1, 0)_{\lambda_2 = 3}$ & $\frac{1}{\sqrt{2}}(1, -1, 0)_{\lambda_3 = 1}$ — 6 for evecs

$$\therefore P = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \\ \sqrt{2} & 0 & 0 \end{pmatrix} \Rightarrow P^T = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 & 0 & \sqrt{2} \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$$

Either a) Multiply out $P^T A P$ to get Λ — 5

or b) Use the more general proof:

$$AP = P\Lambda \Rightarrow P^{-1} A P = \Lambda$$
$$\downarrow$$
$$P^T A P = \Lambda$$

where $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_3 \end{pmatrix}$

Setter's initials: JDG. Checker's initials: RLC Page number

	EXAMINATION QUESTIONS/SOLUTIONS 2010-11	Course EE2 Paper 4 (2)
Question 4 Q2		Marks & seen/unseen

Parts

$$Q = 4x_1^2 + 4x_1x_2 + x_2^2 + 4x_3^2 = \underline{x}^T A \underline{x}$$ — 2

where $A = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ symm $\quad \therefore \lambda_1 = 4$ & $(\lambda-4)(\lambda-1)-4=0$

$\therefore \lambda^2 - 5\lambda = 0 \quad$ so $\quad \lambda_2 = 5 \quad \lambda_3 = 0$ — 3 for evs

$(0,0,1)_{\lambda_1=4}$; $\frac{1}{\sqrt{5}}(2,1,0)_{\lambda_2=5}$; $\frac{1}{\sqrt{5}}(1,-2,0)_{\lambda_3=0}$ — 6 for evecs.

Construct $P = (\underline{a}_1, \underline{a}_2\ \underline{a}_3) = \frac{1}{\sqrt{5}}\begin{pmatrix} 0 & 2 & 1 \\ 0 & 1 & -2 \\ \sqrt{5} & 0 & 0 \end{pmatrix}$

Now define $\underline{x} = P\underline{y}$ so

$$Q = \underline{x}^T A \underline{x} = \underline{y}^T (P^T A P)\underline{y}$$

and $AP = P\Lambda \Rightarrow P^{-1}AP = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2 \ldots)$ — 4

so $P^T A P = \Lambda$

$$\therefore Q = \underline{y}^T \Lambda \underline{y} = \sum_{i=1}^{n} \lambda_i y_i^2$$

$$= 4y_1^2 + 5y_2^2 + 0y_3^2$$ — 2

Now $x = P\underline{y} \rightarrow y = P^{-1}\underline{x} = P^T\underline{x}$ — 3

$$\therefore \underline{y} = \frac{1}{\sqrt{5}}\begin{pmatrix} 0 & 0 & \sqrt{5} \\ 2 & 1 & 0 \\ 1 & -2 & 0 \end{pmatrix}\underline{x}$$

Setter's initials	Checker's initials	Page number
JDG	RK	

	EXAMINATION QUESTIONS/SOLUTIONS 2010-11	Course EE2 Paper 4
Question 4 Q1		Marks & seen/unseen

Parts

If $A\underline{a}_i = \lambda_i \underline{a}_i$, construct $P = (\underline{a}_1\, \underline{a}_2\, \underline{a}_3 \cdots)$ & form. — seen

$$P^T P = \begin{pmatrix} \underline{a}_1^T \\ \underline{a}_2^T \\ \underline{a}_3^T \\ \vdots \end{pmatrix} (\underline{a}_1\, \underline{a}_2\, \underline{a}_3 \cdots) = \begin{pmatrix} \underline{a}_1^T\underline{a}_1 & \underline{a}_1^T\underline{a}_2 & \cdots & \\ \vdots & \underline{a}_2^T\underline{a}_2 & & \\ \vdots & & \ddots & \\ \vdots & & & \underline{a}_n^T\underline{a}_n \end{pmatrix}$$ — 3

$$= \{\underline{a}_i^T \underline{a}_j\} = \{\delta_{ij}\} = I.$$ — 2

using orthonormality property.

$\therefore P^T = P^{-1}$ — 1

unseen

For $A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ symm, $\lambda_1 = 4$ & $(\lambda-2)^2 - 1 = \lambda^2 - 4\lambda + 3 = 0$

$\lambda_2 = 3 \quad \lambda_3 = 1$ — 3 for evals

$(0, 0, 1)_{\lambda_1 = 4}$; $\frac{1}{\sqrt{2}}(1, 1, 0)_{\lambda_2 = 3}$ & $\frac{1}{\sqrt{2}}(1, -1, 0)_{\lambda_3 = 1}$ — 6 for evecs

$$\therefore P = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & -1 \\ \sqrt{2} & 0 & 0 \end{pmatrix} \Rightarrow P^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 0 & \sqrt{2} \\ 1 & 1 & 0 \\ 1 & -1 & 0 \end{pmatrix}$$

Either a) Multiply out P^TAP to get Λ — 5

or b) Use the more general proof:

$$AP = P\Lambda \Rightarrow P^{-1}AP = \Lambda$$

$$\downarrow$$

$$P^TAP = \Lambda$$

where $\Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & \lambda_2 & \\ 0 & & \lambda_3 \end{pmatrix}$

Setter's initials	Checker's initials	Page number
JDG.	RLC	

	EXAMINATION QUESTIONS/SOLUTIONS 2010-11	Course EE2 Paper 4
Question 4 Q2		Marks & seen/unseen

Parts

$Q = 4x_1^2 + 4x_1x_2 + x_2^2 + 4x_3^2 = \underline{x}^T A \underline{x}$ — 2

where $A = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 4 \end{pmatrix}$ symm $\therefore \lambda_1 = 4$ & $(\lambda-4)(\lambda-1) - 4 = 0$

$\therefore \lambda^2 - 5\lambda = 0$ so $\lambda_2 = 5 \quad \lambda_3 = 0$ — 3 for evs

$(0,0,1)_{\lambda_1=4}$; $\frac{1}{\sqrt{5}}(2,1,0)_{\lambda_2=5}$; $\frac{1}{\sqrt{5}}(1,-2,0)_{\lambda_3=0}$ — 6 for evecs.

Construct $P = (\underline{a}_1, \underline{a}_2\ \underline{a}_3) = \frac{1}{\sqrt{5}}\begin{pmatrix} 0 & 2 & 1 \\ 0 & 1 & -2 \\ \sqrt{5} & 0 & 0 \end{pmatrix}$

Now define $\underline{x} = P\underline{y}$ so

$$Q = \underline{x}^T A \underline{x} = \underline{y}^T (P^T A P)\underline{y}$$

and $AP = P\Lambda \Rightarrow P^{-1}AP = \Lambda = \mathrm{diag}(\lambda_1, \lambda_2 \ldots)$ — 4

so $P^T A P = \Lambda$

$$\therefore Q = \underline{y}^T \Lambda \underline{y} = \sum_{i=1}^{n} \lambda_i y_i^2$$

$$= 4y_1^2 + 5y_2^2 + 0y_3^2$$ — 2

Now $\underline{x} = P\underline{y} \Rightarrow \underline{y} = P^{-1}\underline{x} = P^T\underline{x}$ — 3

$$\therefore \underline{y} = \frac{1}{\sqrt{5}}\begin{pmatrix} 0 & 0 & \sqrt{5} \\ 2 & 1 & 0 \\ 1 & -2 & 0 \end{pmatrix}\underline{x}$$

Setter's initials	Checker's initials	Page number
JDG	RK	

3. (a) For a patient chosen randomly from the population, we define the events:
   C : Patient has the condition
   P : Patient tests positive for the condition
   N : Patient tests negative for the condition
   I : The test is inconclusive

   The probability that a patient tests positive is

$$\begin{aligned} P(P) &= P(P|C)P(C) + P(P|\bar{C})P(\bar{C}) \\ &= 0.94 \times 0.001 + 0.02 \times 0.999 \\ &= 0.00094 + 0.01998 = 0.02092\,. \end{aligned}$$

Seen similar — 5 MARKS

(b) The probability is

$$P(C|P) = \frac{P(P|C)P(C)}{P(P)} = \frac{0.00094}{0.02092} = 0.045\,.$$

Seen similar — 4 MARKS

(c) The probability that a patient tests negative is

$$\begin{aligned} P(N) &= P(N|C)P(C) + P(N|\bar{C})P(\bar{C}) \\ &= 0.05 \times 0.001 + 0.93 \times 0.999 \\ &= 0.929\,, \end{aligned}$$

so the probability of a conclusive test is

$$P(\bar{I}) = P(P) + P(N) = 0.021 + 0.929 = 0.950\,.$$

Let X be the number of tests required. Each test is an independent Bernoulli trial (with probability of success 0.95), and we repeat the experiment until the first success. Thus, $X \sim \text{Geo}(0.95)$, a geometric distribution.

Unseen — 6 MARKS

(d)

$$P(X > 3) = P(\text{'first three inconclusive'}) = (1 - 0.95)^3 = 0.000125\,,$$

so, in a population of 1 million,

$$1000000 \times 0.000125 = 125$$

people (on average) will require more than 3 tests.

Unseen — 5 MARKS

4. (a) Clearly, $f_Y(y) \geq 0$ for all y. We now need to verify that $f_Y(y)$ integrates to 1 over $(-\infty, \infty)$. We have:

$$\begin{aligned}
\int_{-\infty}^{\infty} f_Y(y)dy &= \int_0^{\infty} \theta^{-2} y e^{-y/\theta} dy \\
&= -\int_0^{\infty} \theta^{-1} y \left(e^{-y/\theta}\right)' dy \\
&= -\left[\theta^{-1} y e^{-y/\theta}\right]_0^{\infty} + \int_0^{\infty} \left(\theta^{-1} y\right)' e^{-y/\theta} dy \\
&= 0 + \int_0^{\infty} \theta^{-1} e^{-y/\theta} dy \\
&= \left[-e^{-y/\theta}\right]_0^{\infty} \\
&= 1\,,
\end{aligned}$$

so $f_Y(y)$ is a valid PDF.

Unseen — 3 MARKS

(b) The MGF is

$$\begin{aligned}
M_Y(t) &= E(e^{tY}) \\
&= \int_{-\infty}^{\infty} e^{ty} f_Y(y) dy \\
&= \int_0^{\infty} e^{ty} \theta^{-2} y e^{-y/\theta} dy \\
&= \theta^{-2} \int_0^{\infty} y e^{-y(1/\theta - t)} dy
\end{aligned}$$

If we set $\lambda = (1/\theta - t)^{-1}$, we can rearrange the expression inside the integral so that it has the same form as the PDF (with parameter λ) and integrates to 1.

$$\begin{aligned}
M_Y(t) &= \theta^{-2} \int_0^{\infty} y e^{-y/\lambda} dy \\
&= \theta^{-2} \lambda^2 \int_0^{\infty} \lambda^{-2} y e^{-y/\lambda} dy \\
&= \theta^{-2} (1/\theta - t)^{-2} \\
&= (1 - \theta t)^{-2}
\end{aligned}$$

Unseen — 4 MARKS

For the integral to converge, we need $\lambda > 0$ or, equivalently, $t < 1/\theta$.

Unseen — 1 MARK

Differentiating the MGF, we find

$$\begin{aligned}
M_Y'(t) &= \frac{d}{dt}(1 - \theta t)^{-2} = 2\theta(1 - \theta t)^{-3} \\
M_Y''(t) &= \frac{d}{dt} 2\theta(1 - \theta t)^{-3} = 6\theta^2 (1 - \theta t)^{-4}\,,
\end{aligned}$$

so the mean and variance are

$$E(Y) = M_Y'(0) = 2\theta$$

$$\mathrm{Var}(Y) = E(Y^2) - E(Y)^2 = M_Y''(0) - (M_Y'(0))^2 = 6\theta^2 - 4\theta^2 = 2\theta^2$$

Unseen — 3 MARKS

(c) Let $\hat{\theta}_{MM}$ be the method of moments estimator. Set the sample mean equal to the population mean to find

$$\bar{Y} = 2\hat{\theta}_{MM} \Rightarrow \hat{\theta}_{MM} = \frac{\bar{Y}}{2}.$$

Its expectation is

$$E(\hat{\theta}_{MM}) = E(\bar{Y}/2) = \frac{2\theta}{2} = \theta,$$

so it is an unbiased estimator of θ.

Unseen — 4 MARKS

(d) The likelihood function is

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} f_Y(y_i) = \prod_{i=1}^{n} \theta^{-2} y_i e^{-y_i/\theta} \\ &= \theta^{-2n} \left(\prod_{i=1}^{n} y_i \right) e^{-\sum_i y_i/\theta} \\ &= \theta^{-2n} \left(\prod_{i=1}^{n} y_i \right) e^{-n\bar{y}/\theta}. \end{aligned}$$

The log-likelihood is

$$\ell(\theta) = \log(L(\theta)) = -2n \log\theta - \frac{n\bar{y}}{\theta} + C$$

We differentiate with respect to θ and set the derivative equal to 0 to obtain the maximum-likelihood estimator $\hat{\theta}_{ML}$.

$$\frac{d}{d\theta}\ell(\theta) = -\frac{2n}{\theta} + \frac{n\bar{y}}{\theta^2} \Rightarrow -\frac{2n}{\hat{\theta}_{ML}} + \frac{n\bar{Y}}{\hat{\theta}_{ML}^2} = 0 \Rightarrow \hat{\theta}_{ML} = \bar{Y}/2,$$

which is the same as the method of moments estimator. We know that it is unbiased, so its MLE is equal to its variance:

$$\mathrm{MSE}(\hat{\theta}_{ML}) = \mathrm{Var}(\hat{\theta}_{ML}) = \mathrm{Var}(\bar{Y}/2) = \frac{1}{4}\mathrm{Var}(\bar{Y}) = \frac{1}{4}\frac{2\theta^2}{n} = \frac{\theta^2}{2n}$$

Unseen — 5 MARKS

5. (a) i.

$$P(\text{`Team X wins'}) = \sum_{x>y} p_{X,Y}(x,y)$$
$$= 0.10 + 0.07 + 0.02 + 0.06 + 0.02 + 0.03 = 0.3$$

$$P(\text{`draw'}) = \sum_{x=y} p_{X,Y}(x,y)$$
$$= 0.25 + 0.09 + 0.04 + 0.02 = 0.4$$

Unseen — 2 MARKS

ii.

$$p_X(0) = \sum_{y} p_{X,Y}(0,y)$$
$$= 0.25 + 0.12 + 0.05 + 0.02 = 0.44$$

and similarly for the other of the values of $p_X(x)$, $p_Y(y)$. The resulting marginals are

x	0	1	2	3
$p_X(x)$	0.44	0.27	0.20	0.09

and

y	0	1	2	3
$p_Y(y)$	0.44	0.29	0.18	0.09

Directly from the marginal:

$$E(Y) = 0.44 \times 0 + 0.29 \times 1 + 0.18 \times 2 + 0.09 \times 3 = 0.92$$

and, similarly, $E(X) = 0.94$.

Unseen — 4 MARKS

iii. From the marginal distribution of Y:

$$E(Y^2) = 0.44 \times 0^2 + 0.29 \times 1^2 + 0.18 \times 2^2 + 0.09 \times 3^2 = 1.82\,,$$

so the variance is

$$\mathrm{Var}(Y) = E(Y^2) - E(Y)^2 = 1.82 - 0.92^2 = 0.9736\,.$$

Similarly, $E(X^2) = 1.88$ and $\mathrm{Var}(X) = 0.9964$.
We have:

$$E(XY) = \sum_{x,y} xy\,p_{X,Y}(x,y)$$
$$= 0 \times 0 \times 0.25 + 1 \times 0 \times 0.10 + \ldots$$
$$= 1.15$$

and so

$$\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y) = 0.2852\,.$$

X,Y are *not* uncorrelated, because $\mathrm{Cov}(X,Y) \neq 0$.

Unseen — 3 MARKS

iv.

$$P(Y=0|X=0) = \frac{p_{X,Y}(0,0)}{p_X(0)} = \frac{0.25}{0.44} = 0.568$$

and similarly for the remaining probabilities. The conditional PMF is given by

y	0	1	2	3
$P(Y=y\|X=0)$	0.568	0.273	0.114	0.045

This is different to the marginal for Y, so the two random variables are not independent. (Or: X,Y are correlated, so they are certainly not independent.)

Unseen — 2 MARKS

(b) A system consists of k components in parallel, i.e. it functions as long as at least one component functions. The components are unreliable: they fail with probability 0.20.

i. The probability is

$$\begin{aligned} P(\text{'system functions'}) &= P(\text{'at least one component functions'}) \\ &= 1 - P(\text{'all components fail'}) = 1 - 0.2^k\,. \end{aligned}$$

Seen similar — 2 MARKS

ii.

$$\begin{aligned} P(\text{'system functions'}) &> 1 - 0.001 \\ 1 - 0.2^k &> 1 - 0.001 \\ 0.2^k &< 0.001 \\ k\log(0.2) &< \log(0.001) \\ -1.609k &< -6.908 \\ k &> \frac{-6.908}{-1.609} = 4.29\,, \end{aligned}$$

so we need $k = 5$ components.

Unseen — 2 MARKS

(c) For $t > 1$, the cumulative hazard is

$$H(t) = \int_0^t h(s)ds = \int_1^t \frac{\alpha}{s}ds = [\alpha \log s]_1^t = \alpha \log t$$

Since

$$\begin{aligned} H(t) &= -\log(R(t)) = -\log(1-F(t)) \Rightarrow \\ e^{-H(t)} &= 1 - F(t) \Rightarrow \\ F(t) &= 1 - e^{-H(t)}\,, \end{aligned}$$

the CDF is

$$F(t) = 1 - e^{-\alpha \log t} = 1 - t^{-\alpha},$$

for $t > 1$ (and $F_X(x) = 0$ otherwise).
The PDF is

$$f(t) = \frac{d}{dt}F(t) = \alpha t^{-\alpha-1}$$

for $t > 1$ (and $f_X(x) = 0$ otherwise).

Unseen — 5 MARKS

6. (a)

$$\bar{x} = \frac{1}{n_x}\sum_i x_i = \frac{1}{12} \times 11.26 = 0.938$$

$$\text{median} = \frac{1}{2}\left(x_{(\frac{n_x}{2})} + x_{(\frac{n_x}{2}+1)}\right) = \frac{1}{2}(0.91 + 0.93) = 0.92$$

$$s_x^2 = \frac{1}{n_x - 1}\left(\sum_i x_i^2 - n_x\bar{x}^2\right) = 0.0083$$

Seen similar — 2 MARKS

(b) The 95% CI is

$$\begin{aligned}\bar{x} \pm t_{0.975,(n_x-1)}\frac{s_x}{\sqrt{n_x}} &= 0.939 \pm 2.20\frac{\sqrt{0.0081}}{\sqrt{12}} \\ &= 0.939 \pm 0.057 \\ &= (0.882, 0.996)\end{aligned}$$

We need to assume normality, because the sample size is small.
Note: The tables provided give $t_{0.975,(10)} = 2.23$ and $t_{0.975,(12)} = 2.18$. Any interpolated value between the two will be accepted.

Seen similar — 6 MARKS

(c) The interval extends a fair way beyond 0.97 — little evidence that the car meets emission standards.

Seen similar — 2 MARKS

(d) The mean and variance of the new dataset are

$$\bar{y} = 0.849, \quad s_y^2 = 0.0100.$$

Assuming equal variances, the pooled variance estimate is

$$s_p^2 = \frac{(n_x - 1)s_x^2 + (n_y - 1)s_y^2}{n_x + n_y - 2} = 0.0091$$

The null and alternative hypotheses are

$$H_0 : \mu_x = \mu_y$$
$$H_1 : \mu_x > \mu_y \,.$$

It is a one-sided test, because we are interesting in detecting a *reduction* in emissions. The appropriate test statistic is

$$T = \frac{(\bar{x} - \bar{y}) - 0}{s_p \sqrt{(1/n_x) + (1/n_y)}} = 2.357 \,.$$

At the 10% level, the critical value is $t_{0.9,(23)} = 2.07$, so we reject H_0. At the 4% level (the tables provided don't give 5% one-tail probabilities), the critical value is $t_{0.96,(23)} = 2.50$, so we do not reject H_0.

We conclude that there is weak evidence of a reduction in emissions.

Seen similar — 10 MARKS