UNIVERSITY OF LONDON
IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2000

BEng Honours Degree in Computing Part II
MEng Honours Degrees in Computing Part II
for Internal Students of the Imperial College of Science, Technology and Medicine

This paper is also taken for the relevant examinations for the Associateship of the City and Guilds of London Institute

PAPER C245

STATISTICS

Wednesday 3 May 2000, 16:00
Duration: 90 minutes
(Reading time 5 minutes)

Answer THREE questions

Paper contains 4 questions

1. (i) The state of a network is characterised by two related measures, A and B, each of which can equal 0 or 1. A equals 1 with probability 2/5. When A equals 0, then, with probability 2/3, B also equals 0. When A equals 1, then, with probability 3/4, B also equals 1.

Find, showing your reasoning, the probability that

(a) both $A = 0$ and $B = 0$;

(b) $B = 1$;

(c) at least one of A or B equals 0;

(d) A equals 0 when it is known that B equals 1.

(ii) The network described in part (i) is monitored each day. If we assume that the values taken by A and B on each day are independent of the values taken on every other day, calculate the following, showing your reasoning:

(e) the probability that A will be in state 0 in only one of four consecutive days;

(f) if, in one week (consisting of seven days), we observe that B is in state 1 every day, what is the probability that A will be in state 1 on less than 3 of those days?

(g) If we restrict our attention to those weeks in which no Bs in state 0 are observed, what is the mean and standard deviation of the number of days per week on which we will find A in state 1?

2. (i) The files in a computer are modified periodically. The times between modification of each file are distributed independently as a random variable, T, which has an exponential probability density function $f(t) = \lambda e^{-\lambda t}, \quad t \geq 0,$ with mean 10 days.

   (a) What is the value of λ and what is the variance of the times between modifications?

   (b) What proportion of the files will still not be modified after 40 days?

   (c) What is the median time between modifications?

   (d) Find forms for the reliability function and the hazard function of T.

   (e) For those files which have still not been modified after 10 days, what is the mean time remaining until they will be modified?

   (ii) The lifetime, T, of an electronic component is known to follow a distribution with survivor function,

$$F(t) = 1 - e^{-t}(1 + t),$$

   where t is the time in minutes.

   (a) Find the hazard function of these components. Comment on its shape when t is large.

   (b) Find the overall system reliability of the following system, which consists of three components, C1, C2, and C3 of the type described in part (ii)(a) and which operate independently.

```
         |------C1-------C2---------|
         |                          |
 --------|                          |-----------
         |                          |
         |-------------C 3----------|
```

   (c) Using the result for (ii)(b), find the probability that the overall system will run for more than 1 minute.

3. The standard type of battery used in a particular make of laptop computer is known to yield a mean lifetime of 4 hours. A manufacturer claims that a new type of battery will yield a mean lifetime of more than 4 hours. To test this, a sample of the new batteries was used, yielding the data (in hours) which are summarised below. It may be assumed that battery lifetimes are normally distributed about their mean. In answering the questions below, you may wish to use the table of the t-distribution given below.

Sample size $=$ 5,

Sample mean $=$ 6.2,

Sample standard deviation $=$ 1.26.

(i) State appropriate null and alternative hypotheses for this problem. State whether the test you have in mind will be a one-tailed or two-tailed test and justify your choice.

(ii) Carry out a test of the null hypothesis at the 1% level, showing clearly the form of the test statistic you use, stating what distribution you expect the test statistic to have under the null hypothesis and describing the critical region. State your conclusion clearly.

(iii) Suppose it was then pointed out that the mean value of 4 hours of the standard battery type was in fact also based on a sample of only 5. Given that the standard deviation of this sample also happened to be 1.26, carry out a test that the mean lifetimes of batteries manufactured by the two processes are the same. Again show the form of the test statistic, state the distribution you expect it to have under the null hypothesis and describe the critical region.

(iv) Explain any difference in the results you obtained to parts (ii) and (iii).

Extract from a table of the t-distribution:

df	99% quantile	df	99% quantile
2	6.97	12	2.68
4	3.75	14	2.62
6	3.14	16	2.58
8	2.90	18	2.55
10	2.76	20	2.53

4. (i) What would you say about the relative size of the mean and median from:

   (a) a symmetric distribution?

   (b) a positively skewed distribution?

   (ii) A random sample of 64 observations from a population produced the following summary statistics:

$$\sum x_i = 5376; \qquad \sum (x_i - \overline{x})^2 = 100800.$$

   (a) Find a 95% confidence interval for the mean μ of the population from which the sample was drawn, assuming that the sample size is large enough for normal approximations to apply

   (b) Interpret this confidence interval.

   (iii) A random sample of observations is drawn from a random variable, X, which is known to follow a Poisson distribution with probability function

$$P(X = x) = p_x = \frac{e^{-\lambda}\lambda^x}{x!}.$$

   Write down the likelihood function for λ and show that the sample mean is the maximum likelihood estimator of λ.

END OF PAPER

DEPARTMENT OF MATHEMATICS

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

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

$$e^x = 1 + x + x^2/2! + \dots + x^n/n! + \dots,$$

$$\cos x = 1 - x^2/2! + x^4/4! - \dots + (-1)^n x^{2n}/(2n)! + \dots,$$

$$\sin x = x - x^3/3! + x^5/5! - \dots + (-1)^n x^{2n+1}/(2n+1)! + \dots,$$

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

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 rule:

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

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

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

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):

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

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

(b) Some indefinite integrals:

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

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

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

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

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

$$x_{n+1} = x_n - [f(x_n)/f'(x_n)], \quad n = 0, 1, 2 \ldots$$

(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

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

where

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

and

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

October 1998