UNIVERSITY OF LONDON

[C245 2001]

B.ENG. AND M.ENG. EXAMINATIONS 2001

For Internal Students of the Imperial College of Science, Technology and Medicine

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

COMPUTING C245

STATISTICS

Date Tuesday 8th May 2001 2.00 - 3.30 pm

Answer THREE questions

[*Before starting, please make sure that the paper is complete. There should be a total of FOUR questions. Ask the invigilator for a replacement if this copy is faulty.*]

1. (i) A software system has six internal states, which can be described in terms of two random variables, A and B. Random variable A can take three values: 1, 2 and 3. Random variable B can take two values: 1 and 2. Various probabilities are known for the system. In particular, it is known that:

$$\begin{aligned} P(B=1|A=1) &= 0.2 \\ P(B=1|A=2) &= 0.25 \\ P(B=1|A=3) &= 0 \\ P(A=2) &= 0.4 \\ P(A=3) &= 0.1 \end{aligned}$$

Find, showing your reasoning, the probability that

(a) $A = 1$,

(b) $B = 1$,

(c) $B = 2$ when it is known that $A = 1$,

(d) $A = 1$ when it is known that $B = 1$.

(ii) The system described in part (i) is observed every hour and is found to be in a random state, with no apparent dependence from one hour to the next.

(a) What is the probability that the system first enters state $B = 1$ at hour 3?

(b) What is the probability that the system is in state $B = 1$ in the first hour, $B = 2$ in the second hour and $B = 1$ in the third hour?

(c) What is the probability that the system first enters state $B = 1$ at the third hour or later?

(d) What is the probability that two consecutive states will be state $(A = 1,\ B = 1)$ followed by state $(A = 2,\ B = 1)$?

(e) If we observe that the system is in state $(A = 1,\ B = 1)$, what is the probability that the next state will be $(A = 2,\ B = 1)$?

(f) Write down the relevant probabilities and say whether you are more likely to observe the two consecutive states $(A = 1,\ B = 2)$, $(A = 3,\ B = 2)$ or the two consecutive states $(A = 1,\ B = 1)$, $(A = 2,\ B = 2)$.

2. (i) A particular type of component does not deteriorate or become more liable to failure as time progresses but is also very fragile. The hazard function, $h(t)$, of such components is thus constant.

   (a) The distribution of the lifetimes of such components has sometimes been called the 'glass figurine distribution'. What is its scientific name?

   (b) Derive the form of the survivor function and the probability density function of the lifetimes of such components.

   (c) If components have a probability of only 0.2 of surviving for more than two years, calculate the mean lifetime of the components.

   (d) What is the probability that a component will survive for more than 4 years?

   (ii) Six components, $C1, C2, \ldots, C6$, are used to construct a system in such a way that, if any of components $C1$ to $C4$ fails, the system fails, and if both of components $C5$ and $C6$ fail, the system fails.

   (a) Draw a series/parallel diagram for this system, labelling the components.

   (b) If the components have exponential lifetime distributions, all with the same parameter, λ, derive an expression for the overall reliability function of the system.

   (iii) A company C makes battery systems using three identical components in a series arrangement. Each component has an exponential lifetime distribution. Two manufacturers produce components of the kind required. With time measured in years, components from manufacturer A have $\lambda = 1/18$, while those from manufacturer B have $\lambda = 1/24$. Those from manufacturer A cost £2 each, while those from B cost twice as much. Any battery system which lasts more than two years earns the company C a profit of £10. Any battery system which fails before this time is replaced free of charge and earns C nothing. Which manufacturer should company C buy its components from, given that it must buy all components from one of the manufacturers?

3. (i) Explain the following terms, illustrating your explanation with diagrams or numerical examples as appropriate.

 (a) A *positively-skewed* distribution;

 (b) the *first quartile* of a sample;

 (c) the *mode* of a sample;

 (d) the *standard deviation* of a sample;

 (e) the *significance level* of a hypothesis test.

(ii) *Sample A*, below, gives the times, in minutes, it took 15 assembly line workers, trained using method A, to assemble a product. Calculate the mean, median and variance of this sample:

Sample A: 9.2, 9.6, 9.1, 9.1, 9.0, 9.1, 9.3, 9.7, 9.2, 8.4, 9.5, 8.9, 8.7, 9.3, 8.8.

(iii) For normally distributed populations with equal variances, the following formula gives the t-statistic in a test to see if two population means are equal:

$$t = \frac{(\overline{X}_1 - \overline{X}_2)}{S\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} .$$

 (a) Explain the meaning of the terms $\overline{X}_1$, $\overline{X}_2$, S, n_1, and n_2;

 (b) Given that you may assume that the underlying populations of times are normally distributed, carry out a test to see if *Sample B*, given below, has been taken from a population with the same mean as that from which *Sample A* was drawn. State clearly your null and alternative hypotheses, the distribution of the test statistic under the null hypothesis, the significance level of the test and your conclusion.

Sample B: 8.5, 8.4, 8.8, 8.6, 8.8, 9.7, 8.7, 9.0, 8.2, 9.5, 8.8, 8.9, 8.5, 8.7.

4. (i) (a) A sample of n observations, $x_1, \ldots, x_n$, is drawn from a normal distribution. Show that the maximum likelihood of the mean of the distribution is given by $\sum x_i/n$.

(b) Using the result in part (a), find a 95% confidence interval for the mean of the population from which the data below are drawn, given that you may assume that the underlying population is normal.

1.31, 1.09, 0.85, 1.05, 1.13, 0.63, 1.39, 1.09, 0.95, 1.04, 1.07, 1.20.

(ii) An intermittent fault on a laptop computer means that users have, on average, problems once in every five times that they try to send an email. Assuming that the success or failure of each attempt is independent of the success or failure of every other attempt, answer the following questions, in each case showing your reasoning:

(a) What is the probability that they will be successful on all 10 of 10 successive attempts?

(b) What is the probability that they will be successful on 8 or more of 10 successive attempts?

(c) What is the expected number of successful attempts in 10 successive attempts?

(d) What is the variance of the number of successful results in 10 successive attempts?

(e) What is the most likely number of successful results in 10 successive attempts?

END OF PAPER

COMP 245: Probability and Statistics for Students of Computing
2000

This sheet contains important formulae you may need in the examination. It does not contain definitions, concepts, or other material, and it does not contain simple formulae you would be expected to be able to derive or remember yourselves.

(Arithmetic) mean $\bar{x} = \frac{1}{n}\sum x_i$

Median: order the sample values $\{x_1, x_2, \ldots, x_n\}$ so that $x_{(1)}$ is the smallest, $x_{(2)}$ is the next smallest, and so on, then the median is the value $x_{(n+1)/2}$.

Quartiles: the first quartile is $x_{([n+1]/4)}$, using the same ideas as in defining the median.

Geometric mean: $x_G = \sqrt[n]{\prod x_i}$ and *Harmonic mean:* $x_H = \left(\frac{1}{n}\sum\frac{1}{x_i}\right)^{-1} = \frac{n}{\sum 1/x_i}$

Variance $s^2 = \frac{1}{(n-1)}\sum(x_i - \bar{x})^2$, *standard deviation*: $\sqrt{\text{variance}}$

Skewness: $\frac{1}{n-1}\sum\left(\frac{x_i - \bar{x}}{s}\right)^3$

S the set of all possible events, ϕ the empty set
Notation: $s \in S$ $\quad A \subset B$
$\phi \subset A \subset S$ for all A
$A \cup B$ (A or B) $\quad A \cap B$ (A and B) $\quad$ both commutative
$A \cap B$ is the *joint event* of A and B
A and B are *disjoint events* if $A \cap B = \phi$
Complement of A denoted by : A' or $\bar{A}$

$P(\phi) = 0 \qquad P(S) = 1 \qquad P(A) = 1 - P(A')$

For two disjoint events A and B (i.e. events for which $A \cap B = \phi$) $P(A \cup B) = P(A) + P(B)$

For *any* two events A and B: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Generalise: $P(\cup_i A_i) = \sum_i P(A_i) - \sum P(A_i \cap A_j) + \sum P(A_i \cap A_j \cap A_k) + \ldots\ldots$

Two events are said to be *independent* if the occurrence or non-occurrence of one is not affected by whether or not the other occurs

If two events A and B are independent, then $P(A \cap B) = P(A) \,.\, P(B)$

The probability that A will occur, *given that B has occurred* is denoted $P(A \mid B)$
If A and B are independent then $P(A \mid B) = P(A)$.

In general, $P(A \cap B) = P(A \mid B) \,.\, P(B)$ and $P(A \cap B) = P(B \mid A) \,.\, P(A)$

From this $P(A \mid B) = P(B \mid A)P(A)/P(B)$ *(Bayes theorem)*

Now $P(B) = P(B \mid A)P(A) + P(B \mid A')P(A')$ *(theorem of total probability)*

So Bayes theorem can also be written as $P(A|B) = \dfrac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A')P(A')}$

The mean or *expected value* of a random variable is $E(x) = \sum_x xP(x)$, often denoted μ.

The variance of a random variable is
$V(x) = \sum_x (x-\mu)^2 P(x) = E(X^2) - E(X)^2 = E[(X - E(X))^2]$, often denoted σ^2.

The skewness of a random variable is $S(x) = \sum\left(\dfrac{x-\mu}{\sigma}\right)^3 P(x) = \dfrac{E[(x-\mu)^3]}{\sigma^3}$

$E(aX + bY) = aE(X) + bE(Y)$
$V(aX + bY) = a^2V(X) + b^2V(Y)$ if X and Y are independent
$V(aX + bY) = a^2V(X) + b^2V(Y) + 2abCov(X,Y)$, always, with $Cov(X,Y)$ the *covariance* of X and Y, defined as $Cov(X,Y) = E[(X - E(X))(Y - E(Y))]$

More generally, the mean of the sum of a weighted combination of $X_1,\ldots,X_n$, $\sum a_iX_i$ is $\mu = \sum a_i\mu_i$ and the variance of the sum of a weighted combination of *independent* $X_1,\ldots,X_n$, $\sum a_iX_i$ is $\sigma^2 = \sum a_i^2\sigma_i^2$. If they are not independent then
$\sigma^2 = \sum a_i^2\sigma_i^2 + \sum_{i \neq j} a_ia_jCov(X_i,X_j)$

The discrete uniform distribution

Let S be the set of integers from 1 to n.

$P(X = x) = 1/n$ with $\mu = \dfrac{(n+1)}{2}$ and $\sigma^2 = \dfrac{1}{12}(n^2 - 1)$

Bernoulli distribution

Let $P(E) = P(X{=}1) = p$ and $P(E') = P(X{=}0) = 1\text{-}p = q$

$P(X = x) = p^xq^{1-x}$ with $\mu = p$ $\sigma^2 = pq$

Binomial distribution

$P(X = x) = \dbinom{n}{x}p^xq^{n-x}$, Notation: $B(n,p)$ $\mu = np$ $\sigma^2 = npq$

Geometric (e.g. prob x failures before first success)

$P(X = x) = q^xp$ $\mu = \dfrac{q}{p}$ $V(Y) = \dfrac{q}{p^2}$ $x = 0,1,2,3,\ldots.$

Poisson

$P(Z = z) = \dfrac{e^{-\mu}\mu^x}{x!}$ Mean = variance $= \mu$

The *probability distribution function* (or *cumulative distribution function*, the cdf) is $F(x) = P(X \le x)$. The *probability density function* or pdf is $f(x) = F'(x)$ (the derivative of F), so that $F(x) = \int_{-\infty}^{x} f(y)dy$

$$\mu = E(X) = \int x f(x)dx \qquad \sigma^2 = E(x^2) - E(x)^2$$

Uniform: Pdf $$f(x) = \begin{cases} 1/(b-a) & a \le x < b \\ 0 & otherwise \end{cases}$$

The Exponential distribution: Pdf $$f(x) = \begin{cases} \lambda \exp(-\lambda x) & x > 0 \\ 0 & x \le 0 \end{cases}$$

and cdf $F(x) = 1 - \exp(-\lambda x)$ when $x > 0$

$\mu = 1/\lambda \qquad \sigma^2 = 1/\lambda^2$

The Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \quad \text{for } -\infty < x < \infty$$

μ is the mean and σ is the standard deviation

The *standard normal distribution* has $\mu = 0$ and $\sigma = 1$: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

If a random variable X follows a $N(\mu, \sigma^2)$ distribution, then the random variable $(x-\mu)/\sigma$ follows a standard normal distribution $N(0,1)$, often denoted $\phi(x)$. The cdf of the standard normal distribution is often denoted $\Phi(x)$.

The area between two points a and b under a normal curve $N(\mu, \sigma^2)$ is the same as the area under a N(0,1) curve between points $(a-\mu)/\sigma$ and $(b-\mu)/\sigma$.

Joint, marginal, and conditional densities

$$f(x,y) \qquad f_X(x) = \int f(x,y)dy \qquad f(y \mid x) = \frac{f(x,y)}{f_X(x)}$$

Given a random sample $x_1, \ldots, x_n$ from a distribution $p(x;\theta)$, the likelihood function for θ is $L(\theta) = \prod_{i=1}^{n} p(x_i;\theta)$. A 95% confidence interval for the mean μ of a distribution is approximately given by $\bar{x} \pm 1.96 \times s/\sqrt{n}$.

If T is a random variable denoting the lifetime of a component, with pdf $f(T)$ and cdf $F(T)$ the *survivor function* or *reliability function* is $R(T) = 1 - F(T)$ and the hazard function is $r(t) = f(t)/R(t)$. $R(t) = \exp\left[-\int_0^t r(s)ds\right]$

The standard normal tables gives values of $\Phi(x) = F(x)$ for a N(0,1) distribution:

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
.0	.5	.9	.816	1.8	.964	2.8	.997
.1	.540	1.0	.841	1.9	.971	3.0	.998
.2	.579	1.1	.864	2.0	.977	3.5	.9998
.3	.618	1.2	.885	2.1	.982	1.282	.9
.4	.655	1.3	.903	2.2	.986	1.645	.95
.5	.691	1.4	.919	2.3	.989	1.96	.975
.6	.726	1.5	.933	2.4	.992	2.326	.99
.7	.758	1.6	.945	2.5	.994	2.576	.995
.8	.788	1.7	.955	2.6	.995	3.09	.999

The chi-squared table gives the values of x for which $\chi^2(k)$ has $P(X > x) = p$, where $\chi^2(k)$ is the chi-squared distribution with k degrees of freedom.

k	**.995**	**.975**	**.05**	**.025**	**.01**	***k***	**.995**	**.975**	**.05**	**.025**	**.01**
1	.000	.001	3.84	5.02	6.63	**18**	6.26	8.23	28.87	31.53	34.81
2	.010	.051	5.99	7.38	9.21	**20**	7.43	9.59	31.42	34.17	37.57
3	.072	.216	7.81	9.35	11.34	**22**	8.64	10.98	33.92	36.78	40.29
4	.207	.484	9.49	11.14	13.28	**24**	9.89	12.40	36.42	39.36	42.98
5	.412	.831	11.07	12.83	15.09	**26**	11.16	13.84	38.89	41.92	45.64
6	.676	1.24	12.59	14.45	16.81	**28**	12.46	15.31	41.34	44.46	48.28
7	.990	1.69	14.07	16.01	18.48	**30**	13.79	16.79	43.77	46.98	50.89
8	1.34	2.18	15.51	17.53	20.09	**40**	20.71	24.43	55.76	59.34	63.69
9	1.73	2.70	16.92	19.02	21.67	**50**	27.99	32.36	67.50	71.41	76.15
10	2.16	3.25	13.31	20.48	23.21	**60**	35.53	40.48	79.08	83.30	88.38
12	3.07	4.40	21.03	23.34	26.22	**70**	43.28	48.76	90.53	95.02	100.4
14	4.07	5.63	23.68	26.12	29.14	**80**	51.17	57.15	101.9	106.6	112.3
16	5.14	6.91	26.30	28.85	32.00	**100**	67.33	74.22	124.3	129.6	135.8

The Student's t table gives the values of x for which $t(v)$ has $P(|X| > x) = p$, where $t(v)$ is the Student t distribution with v degrees of freedom.

v	**.10**	**.05**	**.02**	**.001**	***v***	**.10**	**.05**	**.02**	**.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