IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

EXAMINATIONS 2023

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 COMP50008

PROBABILITY AND STATISTICS

Friday 19th May 2023, 10:00
Duration: 90 minutes

Answer ALL TWO questions

Paper contains 2 questions
Calculators required

1 a A mobile device uploads pictures and audio recordings through a noisy wireless channel. A picture is sent with probability p and lost with probability q due to a transmission error. Audio recordings are sent with probability $1-p$ and lost with probability v. Assume transmission errors to be independent.

i) Given that an upload was successful, what is the probability that it was a transmission of an audio recording?

ii) A new protocol is tested where each upload is performed three times in a row to increase reliability. If at least two out of three uploads for the same picture (or audio recording) are successful, then the upload has succeeded. What would be the success probability for an upload under this protocol?

b Consider a random variable X and the following function $f(x) = C(k)x^{-k-1}$, for $x \in [a,b]$ with $a > 0$ and $k > 0$, and where $C(k)$ is a normalizing constant.

i) Find $C(k)$ so that $f(x)$ is a valid probability density function.

ii) Assume in this part $k = 1$. Derive an expression for the squared coefficient of variation $c^2 = Var(X)/E[X]^2$ of X in terms of a, b and $C(1)$.

iii) Assume in this part $k = 1$ and $b = 4a$. Give an expression for the conditional expectation $E[X|X > 2a]$.

c Consider a discrete-time Markov chain (DTMC) with initial probability vector $\pi_0 = (1,0,0,0)$ and transition probability matrix

$$R = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ a & 0 & 0 & 1-a \end{bmatrix}$$

i) Assume in this part $a = 1/2$. Let X_n the DTMC state at time $n \geq 0$. Calculate the steady-state probability vector π_∞.

ii) Without calculations, state which value of a would you choose to make π_∞ uniformly distributed across the four states. Justify your answer.

The three parts carry, respectively, 30%, 45%, and 25% of the marks.

2a A random sample consists of 13 participants in the UK who rated the Government's performance in the light of the cost of living crisis. They scored the performance in terms of marks out of 100, which are given below:

37, 44, 21, 18, 46, 43, 22, 5, 36, 41, 34, 52, 30

i) Estimate the mean marks the Government might get if the entire UK population were to do this exercise.

ii) Show that sample variance is a biased estimator of the population variance.

iii) Estimate the variance of the marks the Government might get if the entire UK population were to do this exercise.

iv) What is the 98% confidence interval for the mean marks?

You may use the following table of Student's t distribution:

Degrees of freedom (ν)	Amount of area in one tail (α)							
	0.0005	0.001	0.005	0.010	0.025	0.050	0.100	0.200
1	636.6192	318.3088	63.65674	31.82052	12.70620	6.313752	3.077684	1.376382
2	31.59905	22.32712	9.924843	6.964557	4.302653	2.919986	1.885618	1.060660
3	12.92398	10.21453	5.840909	4.540703	3.182446	2.353363	1.637744	0.978472
4	8.610302	7.173182	4.604095	3.746947	2.776445	2.131847	1.533206	0.940965
5	6.868827	5.893430	4.032143	3.364930	2.570582	2.015048	1.475884	0.919544
6	5.958816	5.207626	3.707428	3.142668	2.446912	1.943180	1.439756	0.905703
7	5.407883	4.785290	3.499483	2.997952	2.364624	1.894579	1.414924	0.896030
8	5.041305	4.500791	3.355387	2.896459	2.306004	1.859548	1.396815	0.888890
9	4.780913	4.296806	3.249836	2.821438	2.262157	1.833113	1.383029	0.883404
10	4.586894	4.143700	3.169273	2.763769	2.228139	1.812461	1.372184	0.879058
11	4.436979	4.024701	3.105807	2.718079	2.200985	1.795885	1.363430	0.875530
12	4.317791	3.929633	3.054540	2.680998	2.178813	1.782288	1.356217	0.872609
13	4.220832	3.851982	3.012276	2.650309	2.160369	1.770933	1.350171	0.870152
14	4.140454	3.787390	2.976843	2.624494	2.144787	1.761310	1.345030	0.868055
15	4.072765	3.732834	2.946713	2.602480	2.131450	1.753050	1.340606	0.866245
16	4.014996	3.686155	2.920782	2.583487	2.119905	1.745884	1.336757	0.864667
17	3.965126	3.645767	2.898231	2.566934	2.109816	1.739607	1.333379	0.863279

b Alice believes she owns a biased coin which tends to fall heads more often than tails. However, she does not know the probability (p) of the coin falling heads. She thinks p is $Beta$-distributed as $p \sim Beta(8, 2)$.

Assume the Normal approximation of Beta distribution to be as follows:

$$Beta(a, b) \approx N\left(\mu = \frac{a}{a+b}, \sigma = \sqrt{\frac{ab}{(a+b)^2(a+b+1)}}\right)$$

Also, assume the following critical z-values for Normal distribution:
$z_{0.9} = 1.282$; $z_{0.95} = 1.645$; $z_{0.99} = 2.326$; $z_{0.995} = 2.576$

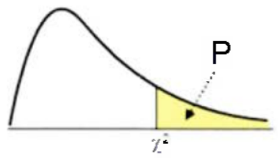
i) Compute the 90% confidence interval of p.

ii) Alice hands the coin to Bob who tosses it fifteen times and observes the coin falling heads in nine of those tosses. Compute the posterior distribution of p and then update the 90% confidence interval of p.

iii) Explain if you can reject the Null Hypothesis that the coin does not fall heads more often than tails at 5% significance level.

c In a survey of 200 independent stock market traders, 40 are Accounting graduates (A), 80 are Business graduates (B), 60 are Computing graduates (C), and the remaining 20 are from other backgrounds (R). In the last financial year: 12 As, 28 Bs, 23 Cs, and 7 Rs made 30% or more profits on their investments; 19 As, 35 Bs, 34 Cs, and 12 Rs made profits greater than 0% but less than 30% on their investments; the remaining incurred losses. The Chi-Squared statistic for the above data is found to be $\chi^2 = 11.001$. Using this, conduct a goodness of fit test at the 20%, 10%, and 5% significance level to check if there is any correlation between the graduating background and performance of the traders.

You may use the following Chi-squared distribution table:



	P										
DF	**0.995**	**0.975**	**0.20**	**0.10**	**0.05**	**0.025**	**0.02**	**0.01**	**0.005**	**0.002**	**0.001**
1	0.0000393	0.000982	1.642	2.706	3.841	5.024	5.412	6.635	7.879	9.550	10.828
2	0.0100	0.0506	3.219	4.605	5.991	7.378	7.824	9.210	10.597	12.429	13.816
3	0.0717	0.216	4.642	6.251	7.815	9.348	9.837	11.345	12.838	14.796	16.266
4	0.207	0.484	5.989	7.779	9.488	11.143	11.668	13.277	14.860	16.924	18.467
5	0.412	0.831	7.289	9.236	11.070	12.833	13.388	15.086	16.750	18.907	20.515
6	0.676	1.237	8.558	10.645	12.592	14.449	15.033	16.812	18.548	20.791	22.458
7	0.989	1.690	9.803	12.017	14.067	16.013	16.622	18.475	20.278	22.601	24.322
8	1.344	2.180	11.030	13.362	15.507	17.535	18.168	20.090	21.955	24.352	26.124
9	1.735	2.700	12.242	14.684	16.919	19.023	19.679	21.666	23.589	26.056	27.877
10	2.156	3.247	13.442	15.987	18.307	20.483	21.161	23.209	25.188	27.722	29.588

The three parts carry, respectively, 35%, 45%, and 20% of the marks.