

**BSc, MSci and MSc EXAMINATIONS (MATHEMATICS)
May 2024**

This paper is also taken for the relevant examination for the
Associateship of the Royal College of Science

Probability and Statistics

Date: Thursday, May 2, 2024

Time: 14:00 – 17:00 (BST)

Time Allowed: 3 hours

This paper has 6 Questions.

Please Answer Each Question in a Separate Answer Booklet

Candidates should start their solutions to each question on a new sheet of paper.

Supplementary books may only be used after the relevant main book(s) are full.

Any required additional material(s) will be provided.

Allow margins for marking.

Credit will be given for all questions attempted.

Each question carries equal weight.

DO NOT OPEN THIS PAPER UNTIL THE INVIGILATOR TELLS YOU TO

Throughout the exam, we assume that $(\Omega, \mathcal{F}, P)$ denotes a probability space.

Please remember to justify all your answers and state carefully which results from the lectures you apply in your proofs.

Statistical tables are provided on the last pages of the exam paper.

1. (a) A football team consisting of 20 players goes on a training camp. They take 50 identical footballs with them and each player has to carry at least one football in their bags. How many ways are there to distribute the 50 identical footballs amongst the 20 players so that each gets at least one? (4 marks)
Hint: You may leave your answer in terms of factorials or binomial coefficients that do not need to be simplified further.
- (b) If the letters B, O, O, K, K, E, E, P, E, R are arranged at random, what is the probability that
 - (i) the arrangements spell the word BOOKKEEPER? (4 marks)
 - (ii) the arrangements have three adjacent E's? (4 marks)
- (c) Consider a mammography test on a female population with a 0.5% prevalence of breast cancer. The test is positive for around 95% of women with cancer, but it is also positive for around 15% of women without cancer. What is the probability that a woman whose mammography test is positive has breast cancer? (4 marks)
- (d) Let $F_{X,Y}$ denote the joint cumulative distribution function of the random vector (X, Y) . For $x, y \in \mathbb{R}$, find an expression for $P(X > x, Y > y)$ in terms of $F_{X,Y}$. (4 marks)

(Total: 20 marks)

2. (a) Consider an experiment where a (possibly unfair) coin is repeatedly tossed to see how many times in a row it will land with the same side facing upwards. Let X denote the length of the initial run (i.e. if you toss heads first, how many heads do you toss before tossing tails, and vice versa if you toss tails first), and let the probability of heads appearing in one toss be given by $p \in (0, 1)$. Find the probability generating function of X .

(8 marks)

- (b) Consider jointly continuous random variables X and Y with joint probability density function given by

$$f_{X,Y}(x, y) = \begin{cases} cx, & \text{for } -1 \leq x < y \leq 0, \\ 0, & \text{otherwise.} \end{cases}$$

- (i) Show that, for $c = -3$, $f_{X,Y}$ is a valid probability density function. (2 marks)
- (ii) Find the marginal density of X . (2 marks)
- (iii) Find the marginal density of Y . (2 marks)
- (iv) Find the expectation of X . (1 mark)
- (v) Find the expectation of Y . (1 mark)
- (vi) Find the covariance of X and Y . (4 marks)

(Total: 20 marks)

3. (a) Consider two independent and Poisson distributed random variables $X \sim \text{Poi}(\lambda)$ and $Y \sim \text{Poi}(\mu)$ with $\lambda, \mu > 0$. Define the random variable $N := X + Y$. Let $n \in \mathbb{N}$. Clearly stating any results used, show that, conditional on $\{N = n\}$, X follows the $\text{Bin}\left(n, \frac{\lambda}{\lambda + \mu}\right)$ distribution. (6 marks)
- (b) Let $N \sim \text{Poi}(\lambda)$ denote a Poisson distributed random variable with $\lambda > 0$, and let $Y_1, Y_2, \dots$ be independent and identically distributed random variables with finite first and second moment, independent of N . Define a new random variable as $X := \sum_{i=1}^N Y_i$.
- (i) Derive the first and second moment of N . (4 marks)
 - (ii) Find the expectation of X . (5 marks)
 - (iii) Find the variance of X . (5 marks)

Hint: You may state your results in (ii) and (iii) as expressions of suitable moments of Y_1 .

(Total: 20 marks)

4. (a) The variance of the random variable X is defined as $\text{Var}(X) = E[(X - E[X])^2]$. Prove that

$$\text{Var}(X) = \min_{a \in \mathbb{R}} E[(X - a)^2].$$

(4 marks)

- (b) Compute the median and interquartile range of the sample of observations

$$\{18.1, 23.7, 18.2, 24.8, 20.3, 24.9, 20.4, 27.0, 20.5, 28.1, 23.6\}.$$

(2 marks)

- (c) The heights of a group of students are assumed to follow a normal distribution with mean μ and variance σ^2 , where μ and σ^2 are both unknown. Ten students are randomly selected from the group and their heights are measured as $x_1, \dots, x_{10}$, and the sample mean and variance of these measurements are computed as $\bar{x} = 172$ cm and $s^2 = 16$ cm², respectively.

- (i) Compute a 99% confidence interval for the unknown mean, μ , with full justification. Tables 1-4 on pages 10-13 may be helpful. (4 marks)

- (ii) Compute a 95% confidence interval for the unknown variance, σ^2 , with full justification. Tables 1-4 on pages 10-13 may be helpful. (4 marks)

- (d) Suppose that $X_1, \dots, X_n$ are independent and identically distributed random variables each following a uniform distribution on the interval $[0, \theta]$, where $\theta > 0$ is unknown. Let $Y_n = \max\{X_1, \dots, X_n\}$. Compute the cumulative distribution function of Y_n and use this to prove that

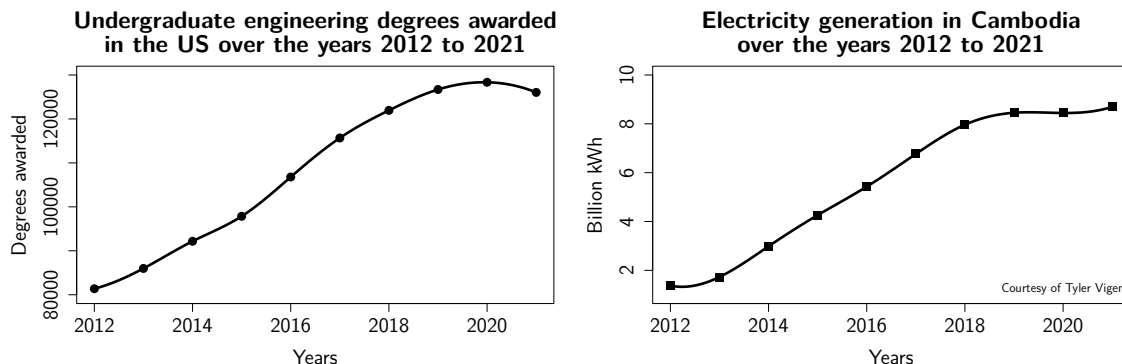
$$\hat{\Theta} = \frac{n+1}{n} Y_n$$

is an unbiased estimator of θ .

(6 marks)

(Total: 20 marks)

5. (a) (i) Prove that if, under the null hypothesis, a test statistic T follows a distribution whose inverse cumulative distribution function exists, then a p -value based on T will follow a $U(0, 1)$ distribution. (4 marks)
- (ii) In hypothesis testing, explain what a significance threshold α is in relation to the p -value and the Type I or Type II error of an experiment. (3 marks)
- (b) Suppose a person at a party claims to be able to taste the difference between Coca-Cola and Pepsi, two popular soft drinks. Their friends decide to put this to the test and conduct an experiment. Twelve glasses are found, and 6 glasses are filled with Coca-Cola and 6 glasses are filled with Pepsi, and while the person is not looking, the glasses are then arranged in a random order. The person is then challenged to taste the drinks in each glass, and identify the 6 glasses that contain Coca-Cola, with the remaining 6 then being identified as containing Pepsi.
- (i) Write down the null and alternative hypotheses for this experiment. (2 marks)
- (ii) Suppose the person correctly identifies 5 out of 6 glasses of Coca-Cola. Is this sufficient evidence to conclude that they can identify the difference between the two soft drinks? Provide justification. (4 marks)
- (c) Suppose that X and Y are two random variables that are dependent and have equal, finite variances, i.e. $\text{Var}(X) = \text{Var}(Y) < \infty$. If $\text{Var}(X + Y)$ and $\text{Var}(X - Y)$ are also both finite, prove that the random variables $X + Y$ and $X - Y$ are uncorrelated. (5 marks)
- (d) The figures below show X , the number of undergraduate engineering degrees awarded in the US for the years 2012 to 2021, and Y , amount of electricity generated in Cambodia for the same years, in billions of kilowatt hours (kWh). The graphs appear to be very similar, and if one computes the sample correlation between the two data sets for X and Y , one obtains a value of $r = 0.997$. Taking this sample correlation value into account, comment on whether or not we can conclude that the two quantities X and Y are related and have any influence on each other. (2 marks)



(Total: 20 marks)

6. (a) Suppose an experiment is performed and the recorded data is used to compute a 95% confidence interval (a, b) for a parameter θ . Is θ in (a, b) with 95% probability? If so, provide justification, if not, provide another interpretation for the confidence interval.

(3 marks)

- (b) Suppose that the data $x_1, x_2, \dots, x_n$ are measurements of the time between phone calls received by a call centre on a weekday morning. A statistician decides to assume independence of the times between phone calls and model this data using an exponential distribution with unknown parameter λ . Let θ be the median of this distribution. Determine the maximum likelihood estimate of θ , given the observed data.

(7 marks)

- (c) Given the data $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ and the model

$$y_i = \beta x_i^2 + e_i, \quad i \in \{1, 2, \dots, n\},$$

derive the estimate $\hat{\beta}$ of the parameter β , such that when the residuals are defined by $\hat{e}_i = y_i - \hat{\beta}x_i^2$, the residual sum of squares

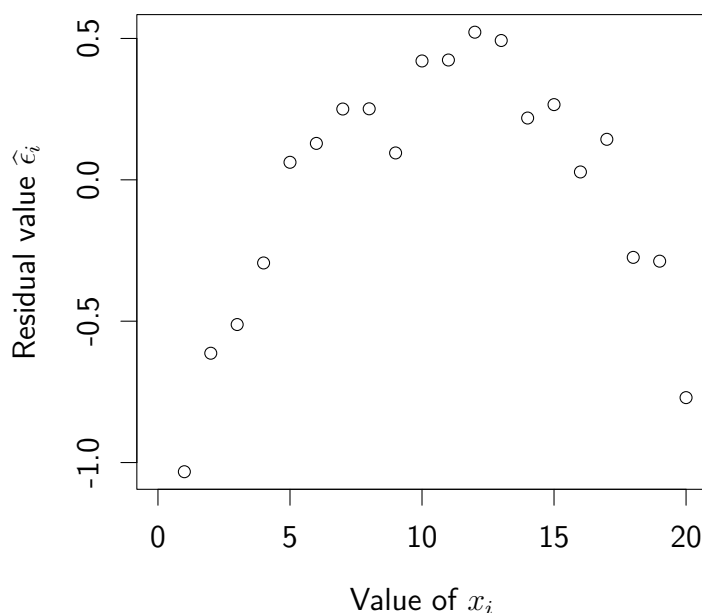
$$\sum_{i=1}^n (\hat{e}_i)^2,$$

is minimised. Express $\hat{\beta}$ in terms of the data.

(6 marks)

- (d) A physicist records data $(x_1, y_1), (x_2, y_2), \dots, (x_{20}, y_{20})$ for two quantities X and Y which are believed to be linearly related. She decides to fit a simple linear regression model to this data, and then obtains the residuals plot below. Does this plot indicate that the model fits the data well or not? Provide justification.

(2 marks)



IMPORTANT: Q6(e) is on the next page

- (e) Suppose that the following data are recorded from a small observational study:

$$\{31, 24, 15, 45, 57\}.$$

For each of the following samples, state whether or not they are bootstrap samples of this data, and provide justification.

- (i) $\{57, 24, 15, 31\}$ (1 mark)
(ii) $\{24, 45, 45, 24, 15\}$ (1 mark)

(Total: 20 marks)

Name of distribution of X	$\text{Im}X$	Parameters	p.m.f. $P(X = x)/\text{p.d.f } f_X(x)$
Discrete distributions			
Bernoulli: $X \sim \text{Bern}(p)$	$\{0, 1\}$	$p \in (0, 1)$	$p^x(1-p)^{1-x}$
Binomial: $X \sim \text{Bin}(n, p)$	$\{0, 1, \dots, n\}$	$n \in \mathbb{N}, p \in (0, 1)$	$\binom{n}{x} p^x (1-p)^{n-x}$
Hypergeometric: $X \sim \text{HGeom}(N, K, n)$	$\{0, 1, \dots, \min(n, K)\}$	$N \in \mathbb{N} \cup \{0\},$ $K, n \in \{0, 1, \dots, N\}$	$\frac{\binom{K}{x} \binom{N-K}{n-x}}{\binom{N}{n}},$ for $x \in \{0, 1, \dots, K\},$ $n-x \in \{0, 1, \dots, N-K\}$
Poisson: $X \sim \text{Poi}(\lambda)$	$\mathbb{N}_0 = \{0, 1, 2, \dots\}$	$\lambda > 0$	$\frac{\lambda^x}{x!} e^{-\lambda}$
Geometric (1): $X \sim \text{Geom}(p)$	$\mathbb{N} = \{1, 2, \dots\}$	$p \in (0, 1)$	$(1-p)^{x-1} p$
Geometric (2): $X \sim \text{Geom}(p)$	$\mathbb{N}_0 = \{0, 1, 2, \dots\}$	$p \in (0, 1)$	$(1-p)^x p$
Negative binomial: $X \sim \text{NBin}(r, p)$	$\mathbb{N}_0 = \{0, 1, 2, \dots\}$	$r \in \mathbb{N}, p \in (0, 1)$	$\binom{x+r-1}{r-1} p^r (1-p)^x$
Continuous distributions			
Uniform: $X \sim \text{U}(a, b)$	(a, b)	$a, b \in \mathbb{R}, a < b$	$\frac{1}{b-a}$
Exponential: $X \sim \text{Exp}(\lambda)$	$(0, \infty)$	$\lambda > 0$	$\lambda e^{-\lambda x}$
Gamma: $X \sim \text{Gamma}(\alpha, \beta)$	$(0, \infty)$	$\alpha, \beta > 0$	$\frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x},$ where $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$
Chi-squared: $X \sim \chi^2(n)$ $(X \sim \chi_n^2)$	$(0, \infty)$	$n \in \mathbb{N}$	$\frac{1}{2\Gamma(n/2)} \left(\frac{x}{2}\right)^{n/2-1} e^{-x/2}$
Beta: $X \sim \text{Beta}(\alpha, \beta)$	$(0, 1)$	$\alpha, \beta > 0$	$\frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1},$ where $B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} dx$ $= \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$
Normal: $X \sim \text{N}(\mu, \sigma^2)$	$\mathbb{R}$	$\mu \in \mathbb{R}, \sigma > 0$	$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$
Student-t: $X \sim \text{Student}(\nu)$	$\mathbb{R}$	$\nu > 0$	$\frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi}\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$

Table 1: $P(Z < z)$ where $Z \sim N(0, 1)$ for values of z between 1.00 and 3.99. One reads the table as $P(Z < z) = p_{r,c}$, where $z = z_r + z_c$ and z_r are the first two digits of z shown in row r , z_c is the third digit of z shown in column c , and $p_{r,c}$ is the (r, c) th entry. For example, $P(Z < 1.02) = 0.8461$, looking in row labelled 1.0 and column labelled 0.02.

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table 2: Values of t for $P(T < t)$, where T has Student's t -distribution with ν degrees of freedom, for $\nu \in \{1, 2, \dots, 30\}$. In this table, $P(T < t_{\nu,p}) = p$, where ν is the ν th row, p is the column with heading p , and $t_{\nu,p}$ is the corresponding entry in the table. For example, $t_{3,0.60} = 0.277$, looking in row labelled 3 and column labelled 0.60.

ν	0.60	0.667	0.75	0.80	0.875	0.90	0.95	0.975	0.99	0.995	0.999
1	0.325	0.577	1.000	1.376	2.414	3.078	6.314	12.706	31.821	63.657	318.31
2	0.289	0.500	0.816	1.061	1.604	1.886	2.920	4.303	6.965	9.925	22.327
3	0.277	0.476	0.765	0.978	1.423	1.638	2.353	3.182	4.541	5.841	10.215
4	0.271	0.464	0.741	0.941	1.344	1.533	2.132	2.776	3.747	4.604	7.173
5	0.267	0.457	0.727	0.920	1.301	1.476	2.015	2.571	3.365	4.032	5.893
6	0.265	0.453	0.718	0.906	1.273	1.440	1.943	2.447	3.143	3.707	5.208
7	0.263	0.449	0.711	0.896	1.254	1.415	1.895	2.365	2.998	3.499	4.785
8	0.262	0.447	0.706	0.889	1.240	1.397	1.860	2.306	2.896	3.355	4.501
9	0.261	0.445	0.703	0.883	1.230	1.383	1.833	2.262	2.821	3.250	4.297
10	0.260	0.444	0.700	0.879	1.221	1.372	1.812	2.228	2.764	3.169	4.144
11	0.260	0.443	0.697	0.876	1.214	1.363	1.796	2.201	2.718	3.106	4.025
12	0.259	0.442	0.695	0.873	1.209	1.356	1.782	2.179	2.681	3.055	3.930
13	0.259	0.441	0.694	0.870	1.204	1.350	1.771	2.160	2.650	3.012	3.852
14	0.258	0.440	0.692	0.868	1.200	1.345	1.761	2.145	2.624	2.977	3.787
15	0.258	0.439	0.691	0.866	1.197	1.341	1.753	2.131	2.602	2.947	3.733
16	0.258	0.439	0.690	0.865	1.194	1.337	1.746	2.120	2.583	2.921	3.686
17	0.257	0.438	0.689	0.863	1.191	1.333	1.740	2.110	2.567	2.898	3.646
18	0.257	0.438	0.688	0.862	1.189	1.330	1.734	2.101	2.552	2.878	3.610
19	0.257	0.438	0.688	0.861	1.187	1.328	1.729	2.093	2.539	2.861	3.579
20	0.257	0.437	0.687	0.860	1.185	1.325	1.725	2.086	2.528	2.845	3.552
21	0.257	0.437	0.686	0.859	1.183	1.323	1.721	2.080	2.518	2.831	3.527
22	0.256	0.437	0.686	0.858	1.182	1.321	1.717	2.074	2.508	2.819	3.505
23	0.256	0.436	0.685	0.858	1.180	1.319	1.714	2.069	2.500	2.807	3.485
24	0.256	0.436	0.685	0.857	1.179	1.318	1.711	2.064	2.492	2.797	3.467
25	0.256	0.436	0.684	0.856	1.178	1.316	1.708	2.060	2.485	2.787	3.450
26	0.256	0.436	0.684	0.856	1.177	1.315	1.706	2.056	2.479	2.779	3.435
27	0.256	0.435	0.684	0.855	1.176	1.314	1.703	2.052	2.473	2.771	3.421
28	0.256	0.435	0.683	0.855	1.175	1.313	1.701	2.048	2.467	2.763	3.408
29	0.256	0.435	0.683	0.854	1.174	1.311	1.699	2.045	2.462	2.756	3.396
30	0.256	0.435	0.683	0.854	1.173	1.310	1.697	2.042	2.457	2.750	3.385

Table 3: Values of x for $P(X < x)$, where $X \sim \chi^2(\nu)$, a chi-squared distribution with ν degrees of freedom. In this table, $P(T < \chi^2_{\nu,p}) = p$, where ν is the ν th row, p is the column with heading p , and $\chi^2_{\nu,p}$ is the corresponding entry in the table. For example, $\chi^2_{3,0.005} = 0.072$, looking in row labelled 3 and column labelled 0.005.

ν	0.001	0.005	0.01	0.025	0.05	0.1	0.125	0.2	0.25	0.333	0.5
1	0.000	0.000	0.000	0.001	0.004	0.016	0.025	0.064	0.102	0.186	0.455
2	0.002	0.010	0.020	0.051	0.103	0.211	0.267	0.446	0.575	0.811	1.386
3	0.024	0.072	0.115	0.216	0.352	0.584	0.692	1.005	1.213	1.568	2.366
4	0.091	0.207	0.297	0.484	0.711	1.064	1.219	1.649	1.923	2.378	3.357
5	0.210	0.412	0.554	0.831	1.145	1.610	1.808	2.343	2.675	3.216	4.351
6	0.381	0.676	0.872	1.237	1.635	2.204	2.441	3.070	3.455	4.074	5.348
7	0.598	0.989	1.239	1.690	2.167	2.833	3.106	3.822	4.255	4.945	6.346
8	0.857	1.344	1.646	2.180	2.733	3.490	3.797	4.594	5.071	5.826	7.344
9	1.152	1.735	2.088	2.700	3.325	4.168	4.507	5.380	5.899	6.716	8.343
10	1.479	2.156	2.558	3.247	3.940	4.865	5.234	6.179	6.737	7.612	9.342
11	1.834	2.603	3.053	3.816	4.575	5.578	5.975	6.989	7.584	8.514	10.341
12	2.214	3.074	3.571	4.404	5.226	6.304	6.729	7.807	8.438	9.420	11.340
13	2.617	3.565	4.107	5.009	5.892	7.042	7.493	8.634	9.299	10.331	12.340
14	3.041	4.075	4.660	5.629	6.571	7.790	8.266	9.467	10.165	11.245	13.339
15	3.483	4.601	5.229	6.262	7.261	8.547	9.048	10.307	11.037	12.163	14.339
16	3.942	5.142	5.812	6.908	7.962	9.312	9.837	11.152	11.912	13.083	15.338
17	4.416	5.697	6.408	7.564	8.672	10.085	10.633	12.002	12.792	14.006	16.338
18	4.905	6.265	7.015	8.231	9.390	10.865	11.435	12.857	13.675	14.931	17.338
19	5.407	6.844	7.633	8.907	10.117	11.651	12.242	13.716	14.562	15.859	18.338
20	5.921	7.434	8.260	9.591	10.851	12.443	13.055	14.578	15.452	16.788	19.337
21	6.447	8.034	8.897	10.283	11.591	13.240	13.873	15.445	16.344	17.720	20.337
22	6.983	8.643	9.542	10.982	12.338	14.041	14.695	16.314	17.240	18.653	21.337
23	7.529	9.260	10.196	11.689	13.091	14.848	15.521	17.187	18.137	19.587	22.337
24	8.085	9.886	10.856	12.401	13.848	15.659	16.351	18.062	19.037	20.523	23.337
25	8.649	10.520	11.524	13.120	14.611	16.473	17.184	18.940	19.939	21.461	24.337
26	9.222	11.160	12.198	13.844	15.379	17.292	18.021	19.820	20.843	22.399	25.336
27	9.803	11.808	12.879	14.573	16.151	18.114	18.861	20.703	21.749	23.339	26.336
28	10.391	12.461	13.565	15.308	16.928	18.939	19.704	21.588	22.657	24.280	27.336
29	10.986	13.121	14.256	16.047	17.708	19.768	20.550	22.475	23.567	25.222	28.336
30	11.588	13.787	14.953	16.791	18.493	20.599	21.399	23.364	24.478	26.165	29.336
35	14.688	17.192	18.509	20.569	22.465	24.797	25.678	27.836	29.054	30.894	34.336
40	17.916	20.707	22.164	24.433	26.509	29.051	30.008	32.345	33.660	35.643	39.335
45	21.251	24.311	25.901	28.366	30.612	33.350	34.379	36.884	38.291	40.407	44.335
50	24.674	27.991	29.707	32.357	34.764	37.689	38.785	41.449	42.942	45.184	49.335

Table 4: Values of x for $P(X < x)$, where $X \sim \chi^2(\nu)$, a chi-squared distribution with ν degrees of freedom. In this table, $P(T < \chi^2_{\nu,p}) = p$, where ν is the ν th row, p is the column with heading p , and $\chi^2_{\nu,p}$ is the corresponding entry in the table. For example, $\chi^2_{3,0.6} = 2.946$, looking in row labelled 3 and column labelled 0.6.

ν	0.6	0.66	0.75	0.80	0.87	0.9	0.95	0.975	0.99	0.995	0.999
1	0.708	0.936	1.323	1.642	2.354	2.706	3.841	5.024	6.635	7.879	10.828
2	1.833	2.197	2.773	3.219	4.159	4.605	5.991	7.378	9.210	10.597	13.816
3	2.946	3.405	4.108	4.642	5.739	6.251	7.815	9.348	11.345	12.838	16.266
4	4.045	4.579	5.385	5.989	7.214	7.779	9.488	11.143	13.277	14.860	18.467
5	5.132	5.730	6.626	7.289	8.625	9.236	11.070	12.833	15.086	16.750	20.515
6	6.211	6.867	7.841	8.558	9.992	10.645	12.592	14.449	16.812	18.548	22.458
7	7.283	7.992	9.037	9.803	11.326	12.017	14.067	16.013	18.475	20.278	24.322
8	8.351	9.107	10.219	11.030	12.636	13.362	15.507	17.535	20.090	21.955	26.125
9	9.414	10.215	11.389	12.242	13.926	14.684	16.919	19.023	21.666	23.589	27.877
10	10.473	11.317	12.549	13.442	15.198	15.987	18.307	20.483	23.209	25.188	29.588
11	11.530	12.414	13.701	14.631	16.457	17.275	19.675	21.920	24.725	26.757	31.264
12	12.584	13.506	14.845	15.812	17.703	18.549	21.026	23.337	26.217	28.300	32.910
13	13.636	14.595	15.984	16.985	18.939	19.812	22.362	24.736	27.688	29.819	34.528
14	14.685	15.680	17.117	18.151	20.166	21.064	23.685	26.119	29.141	31.319	36.123
15	15.733	16.761	18.245	19.311	21.384	22.307	24.996	27.488	30.578	32.801	37.697
16	16.780	17.840	19.369	20.465	22.595	23.542	26.296	28.845	32.000	34.267	39.252
17	17.824	18.917	20.489	21.615	23.799	24.769	27.587	30.191	33.409	35.718	40.790
18	18.868	19.991	21.605	22.760	24.997	25.989	28.869	31.526	34.805	37.156	42.312
19	19.910	21.063	22.718	23.900	26.189	27.204	30.144	32.852	36.191	38.582	43.820
20	20.951	22.133	23.828	25.038	27.376	28.412	31.410	34.170	37.566	39.997	45.315
21	21.991	23.201	24.935	26.171	28.559	29.615	32.671	35.479	38.932	41.401	46.797
22	23.031	24.268	26.039	27.301	29.737	30.813	33.924	36.781	40.289	42.796	48.268
23	24.069	25.333	27.141	28.429	30.911	32.007	35.172	38.076	41.638	44.181	49.728
24	25.106	26.397	28.241	29.553	32.081	33.196	36.415	39.364	42.980	45.559	51.179
25	26.143	27.459	29.339	30.675	33.247	34.382	37.652	40.646	44.314	46.928	52.620
26	27.179	28.520	30.435	31.795	34.410	35.563	38.885	41.923	45.642	48.290	54.052
27	28.214	29.580	31.528	32.912	35.570	36.741	40.113	43.195	46.963	49.645	55.476
28	29.249	30.639	32.620	34.027	36.727	37.916	41.337	44.461	48.278	50.993	56.892
29	30.283	31.697	33.711	35.139	37.881	39.087	42.557	45.722	49.588	52.336	58.301
30	31.316	32.754	34.800	36.250	39.033	40.256	43.773	46.979	50.892	53.672	59.703
35	36.475	38.024	40.223	41.778	44.753	46.059	49.802	53.203	57.342	60.275	66.619
40	41.622	43.275	45.616	47.269	50.424	51.805	55.758	59.342	63.691	66.766	73.402
45	46.761	48.510	50.985	52.729	56.052	57.505	61.656	65.410	69.957	73.166	80.077
50	51.892	53.733	56.334	58.164	61.647	63.167	67.505	71.420	76.154	79.490	86.661

1. (a) We can use the stars and bars argument to solve this problem. We represent the balls as stars and the dividers that separate the players as bars. Since each player must receive at least one ball, we distribute 20 balls initially. This leaves 30 balls to be distributed among the 20 players. Using the stars and bars argument, we have 30 stars (representing the 30 remaining balls) and 19 bars (dividers between the 20 players). The total number of ways to distribute the balls can be calculated as the number of ways to arrange these 30 stars and 19 bars, which is

meth seen ↓

2, B

$$\binom{(30+19)}{19} = \binom{49}{19} = \binom{49}{30}.$$

2, B

Therefore, there are $\binom{49}{19} = \binom{49}{30}$ ways to distribute the 50 identical balls among the 20 players, ensuring that each player receives at least one ball.

[2 marks for the correct calculations and result and 2 marks for the justification.]

meth seen ↓

- (b) (i) Let Ω denote the sample space consisting of all possible arrangements of the letters B, O, O, K, K, E, E, P, E, R. Using the multiplication principles, there are $10!$ possibilities of arranging the 10 letters. Note that E appears three times, we need to adjust for overcounting, these letters can be arranged in $3!$ possible ways; also, O and K appears twice, we need to adjust for overcounting, each of these letters can be arranged in $2!$ possible orders. Altogether we have

$$\text{card}(\Omega) = \frac{10!}{3!2!2!} = 10 \times \cdots \times 5 = \frac{10!}{4!} = (10)_6.$$

[The factorials do not need to be expanded for full marks.]

We define the event A = the arrangements spell the word BOOKKEEPER. Since only one arrangement spells the word BOOKKEEPER, the corresponding probability is given by

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{1}{10!/4!} = \frac{4!}{10!}.$$

[2 marks for the justification and 2 marks for the correct probability]

4, A

- (ii) Consider removing the E's: _ B _ O _ K _ K _ P _ R_. We can arrange the remaining letters in $7!/(2!)^2$ possible ways, where we adjust for the fact that the letters O and K appear twice. In order to ensure that we have three adjacent E's, we must choose one of the blank spaces _ to put the three E's. We can do this in $\binom{8}{1} = 8$ possible ways. We define the event A = the arrangements have three adjacent E's. Then

$$\text{card}(A) = \frac{7!}{(2!)^2} \cdot 8 = \frac{8!}{(2!)^2}.$$

Hence,

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)} = \frac{8!}{(2!)^2} \frac{4!}{10!} = \frac{1}{15}.$$

[2 marks for the justification and 2 marks for the correct probability]

4, A

- (c) Consider a mammography test on a population with a 0.5% prevalence of breast cancer. The test is positive for around 95% of women with cancer, but it is also positive for around 15% of women without cancer. What is the probability that a woman whose mammography test is positive has breast cancer?

meth seen ↓

Define the events: $B :=$ breast cancer present; $TP =$ test is positive. We know that $P(B) = 0.005$ and $P(TP|B) = 0.95$ and $P(TP|B^c) = 0.15$.

We can derive that $P(B^c) = 1 - P(B) = 0.995$.

1, A

Further, by the law of total probability,

$$\begin{aligned} P(TP) &= P(TP|B)P(B) + P(TP|B^c)P(B^c) \\ &= 0.95 \cdot 0.005 + 0.15 \cdot 0.995 = 0.00475 + 0.14925 = 0.154. \end{aligned}$$

Using Bayes' Theorem, we get

$$\begin{aligned} P(B|TP) &= \frac{P(B \cap TP)}{P(TP)} = \frac{P(TP|B)P(B)}{P(TP)} = \frac{0.95 \cdot 0.005}{0.154} \\ &= \frac{0.00475}{0.154} (= 0.03084416 \approx 3\%). \end{aligned}$$

[1 mark for introducing the notation and reading off the probabilities, 1 mark for mentioning both the law of total probability (LTP) and Bayes theorem, 2 marks for the correct calculations.]

3, A

- (d) We note that, for $x, y \in \mathbb{R}$,

meth seen ↓

$$\begin{aligned} P(X > x, Y > y) + P(X > x, Y \leq y) &\stackrel{\text{Law of total prob.}}{=} P(X > x) \\ &= 1 - P(X \leq x) = 1 - F_X(x) = 1 - F_{X,Y}(x, \infty), \end{aligned}$$

which implies that

$$P(X > x, Y > y) = -P(X > x, Y \leq y) + 1 - F_{X,Y}(x, \infty).$$

Also,

$$P(X > x, Y \leq y) + \underbrace{P(X \leq x, Y \leq y)}_{=F_{X,Y}(x,y)} \stackrel{\text{Law of total prob.}}{=} P(Y \leq y) = F_Y(y) = F_{X,Y}(\infty, y),$$

which implies that

$$P(X > x, Y \leq y) = -F_{X,Y}(x, y) + F_{X,Y}(\infty, y).$$

Hence,

$$P(X > x, Y > y) = 1 - F_{X,Y}(x, \infty) - F_{X,Y}(\infty, y) + F_{X,Y}(x, y).$$

[3 marks for the correct calculations, 1 mark for the justification]

4, A

2. (a) Let H denote the event that the first coin toss gives heads. Then H^c is the event that the first coin toss gives tail. Then, for $x \in \mathbb{N}$, we have

seen ↓

$$P(X = x|H) = p^{x-1}(1 - p),$$

since after the initial toss of heads we need to toss heads $x - 1$ times and then toss tail once to obtain a run of heads of length x .

Similarly, we get for $x \in \mathbb{N}$,

$$P(X = x|H^c) = (1 - p)^{x-1}p.$$

By the law of total expectation, we get the following expression for the probability generating function of X :

2, A

unseen ↓

$$G_X(s) := E(s^X) = pE(s^X|H) + (1 - p)E(s^X|H^c),$$

for $|s| < 1$.

2, C

Note that, by the law of the unconscious statistician (LOTUS), we have

$$\begin{aligned} E(s^X|H) &= \sum_{x=1}^{\infty} s^x P(X = x|H) = \sum_{x=1}^{\infty} s^x p^{x-1}(1 - p) \\ &= s(1 - p) \sum_{x=1}^{\infty} (sp)^{x-1} = s(1 - p) \sum_{x=0}^{\infty} (sp)^x = \frac{s(1 - p)}{1 - sp}, \end{aligned}$$

using the geometric series and the assumption that $p < 1, |s| < 1$. Similarly,

$$\begin{aligned} E(s^X|H^c) &= \sum_{x=1}^{\infty} s^x P(X = x|H^c) = \sum_{x=1}^{\infty} s^x (1 - p)^{x-1}p \\ &= sp \sum_{x=1}^{\infty} (s(1 - p))^{x-1} = sp \sum_{x=0}^{\infty} (s(1 - p))^x = \frac{sp}{1 - s(1 - p)}, \end{aligned}$$

using the geometric series and the assumption that $1 - p < 1, |s| < 1$.

3, C

Altogether, we have, for $|s| < 1$,

$$\begin{aligned} G_X(s) &= E(s^X) = pE(s^X|H) + (1 - p)E(s^X|H^c) = p \frac{s(1 - p)}{1 - sp} + (1 - p) \frac{sp}{1 - s(1 - p)} \\ &= \frac{sp(1 - p)}{1 - sp} + \frac{sp(1 - p)}{1 - s(1 - p)}. \end{aligned}$$

1, C

[2 marks for finding the conditional pmf of X (seen), 2 marks for stating and applying the law of total expectation correctly, 2 marks for computing the conditional pgfs correctly, 1 mark for stating the condition that $|s| < 1$, 1 mark for the correct final result.]

(b) (i) We note that

$$\begin{aligned}\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x,y) dy dx &= \int_{-1}^0 \int_x^0 c x dy dx = c \int_{-1}^0 xy \Big|_{y=x}^0 dx = -c \int_{-1}^0 x^2 dx \\ &= -c \frac{x^3}{3} \Big|_{x=-1}^0 = -\frac{c}{3} \stackrel{!}{=} 1,\end{aligned}$$

if and only if $c = -3$. We note that, for $c = -3$, the density is nonnegative for all x, y and integrates to 1.

2, A

(ii)

$$f_X(x) = \int_x^0 -3x dy = 3x^2,$$

for $-1 \leq x \leq 0$ and 0 otherwise.

2, A

(iii)

$$f_Y(y) = \int_{-1}^x -3x dx = \frac{3}{2}(1 - y^2),$$

for $-1 \leq y \leq 0$ and 0 otherwise.

2, A

(iv)

$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx = \int_{-1}^0 3x^3 dx = -\frac{3}{4}.$$

1, B

(v)

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y) dy = \int_{-1}^0 \frac{3}{2}(1 - y^2)y dy = -\frac{3}{8}.$$

1, B

(vi)

$$\begin{aligned}E(XY) &= \int_{-\infty}^{\infty} xy f_{X,Y}(x,y) dy dx = \int_{-1}^0 \int_x^0 xy(-3) dy dx = -3 \int_{-1}^0 \int_x^0 x^2 y dy dx \\ &= -3 \int_{-1}^0 x^2 y^2/2 \Big|_{y=x}^0 dx = \frac{3}{2} \int_{-1}^0 x^4 dx = -\frac{3}{10}.\end{aligned}$$

Hence,

3, B

$$\begin{aligned}\text{Cov}(X, Y) &= E(XY) - E(X)E(Y) = -\frac{3}{10} - \left(-\frac{3}{4}\right) \times \left(-\frac{3}{8}\right) = -\frac{3}{10} - \frac{9}{32} \\ &= -\frac{96 + 90}{320} = -\frac{186}{320} = -\frac{93}{160} [= -0.58125].\end{aligned}$$

1, B

3. (a) We recall from lectures (see Ex 6-12) that $N \sim \text{Poi}(\lambda + \mu)$.

unseen ↓

2, B

Let $n \in \mathbb{N}$ and $k \in \{0, 1, \dots, n\}$, then we have

$$\begin{aligned} P(X = k|N = n) &\stackrel{\text{Bayes}}{=} \frac{P(N = n|X = k)P(X = k)}{P(N = n)} = \frac{P(X + Y = n|X = k)P(X = k)}{P(N = n)} \\ &= \frac{P(Y = n - X|X = k)P(X = k)}{P(N = n)} = \frac{P(Y = n - k|X = k)P(X = k)}{P(X + Y = n)} \\ &\stackrel{X, Y \text{ independent}}{=} \frac{P(Y = n - k)P(X = k)}{P(X + Y = n)} \\ &= \frac{\mu^{n-k}}{(n-k)!} e^{-\mu} \frac{\lambda^k}{k!} e^{-\lambda} \left(\frac{(\lambda + \mu)^n}{n!} e^{-(\lambda + \mu)} \right)^{-1} = \binom{n}{k} \left(\frac{\lambda}{\lambda + \mu} \right)^k \left(\frac{\mu}{\lambda + \mu} \right)^{n-k}, \end{aligned}$$

which is indeed the probability mass function of a $\text{Bin}\left(n, \frac{\lambda}{\lambda + \mu}\right)$ random variable.

[3 marks for the correct computations, 1 mark for the justifications]

4, D

- (b) (i) For $N \sim \text{Poi}(\lambda)$ and $t \in \mathbb{R}$, we derive the moment generating function as

seen ↓

$$\begin{aligned} M_N(t) &= E(e^{tN}) = \sum_{n=0}^{\infty} e^{tn} \frac{\lambda^n}{n!} e^{-\lambda} = \sum_{n=0}^{\infty} \frac{(e^t \lambda)^n}{n!} e^{-\lambda} \\ &= \exp(-\lambda + \lambda e^t) = \exp(\lambda(e^t - 1)). \end{aligned}$$

Then

$$M'_N(t) = \exp(\lambda(e^t - 1))\lambda e^t, \quad M''_N(t) = \exp(\lambda(e^t - 1))\lambda e^t + \exp(\lambda(e^t - 1))(\lambda e^t)^2,$$

and hence $E(N) = M'_N(0) = \lambda$, $E(N^2) = M''_N(0) = \lambda + \lambda^2$.

1, C

[Any correct method can be used for computing the two first moments on N . 2 marks for the mean, 2 marks for the second moment.]

3, B

- (ii) We can use the law of total expectation, to find that

unseen ↓

$$\begin{aligned} E(X) &= \sum_{n=0}^{\infty} E(X|N = n)P(N = n) = \sum_{n=0}^{\infty} E\left(\sum_{i=1}^N Y_i \middle| N = n\right)P(N = n) \\ &= \sum_{n=0}^{\infty} E\left(\sum_{i=1}^n Y_i \middle| N = n\right)P(N = n) \end{aligned}$$

2, C

Next, we use that (Y_i) and N are independent, that the (Y_i) are i.i.d. and $E(N) = \lambda$. Hence,

$$\begin{aligned} E(X) &\stackrel{N, (Y_i) \text{ independent}}{=} \sum_{n=0}^{\infty} E\left(\sum_{i=1}^n Y_i\right)P(N = n) \\ &= \sum_{n=0}^{\infty} nE(Y_1)P(N = n) = E(Y_1) \sum_{n=0}^{\infty} nP(N = n) = E(Y_1)E(N) = E(Y_1)\lambda. \end{aligned}$$

3, D

- (iii) For the second moment, we use the law of total expectation and the LOTUS to conclude that

unseen ↓

$$\begin{aligned}
 E(X^2) &= \sum_{n=0}^{\infty} E(X^2|N=n)P(N=n) = \sum_{n=0}^{\infty} E\left(\left(\sum_{i=1}^N Y_i\right)^2 \middle| N=n\right) P(N=n) \\
 &= \sum_{n=0}^{\infty} E\left(\left(\sum_{i=1}^n Y_i\right)^2 \middle| N=n\right) P(N=n) \\
 &\stackrel{N, (Y_i)_i \text{ independent}}{=} \sum_{n=0}^{\infty} E\left(\left(\sum_{i=1}^n Y_i\right)^2\right) P(N=n),
 \end{aligned}$$

where we used that (Y_i) and N are independent. Using the fact that the (Y_i) are i.i.d., we get

2, D

$$\begin{aligned}
 E\left(\left(\sum_{i=1}^n Y_i\right)^2\right) &= E\left(\sum_{i=1}^n \sum_{j=1}^n Y_i Y_j\right) \\
 &\stackrel{\text{linearity of } E}{=} \sum_{i=1}^n \sum_{j=1}^n E(Y_i Y_j) = \sum_{i=1}^n E(Y_i^2) + \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(Y_i Y_j) \\
 &\stackrel{\text{independence of } (Y_i)_i}{=} \sum_{i=1}^n E(Y_i^2) + \sum_{i=1}^n \sum_{j=1, j \neq i}^n E(Y_i)E(Y_j) \\
 &\stackrel{\text{identical distr. of } (Y_i)_i}{=} nE(Y_1^2) + (n^2 - n)(E(Y_1))^2.
 \end{aligned}$$

Then

$$\begin{aligned}
 E(X^2) &= \sum_{n=0}^{\infty} [nE(Y_1^2) + (n^2 - n)(E(Y_1))^2] P(N=n) \\
 &= E(Y_1^2) \sum_{n=0}^{\infty} nP(N=n) + (E(Y_1))^2 \sum_{n=0}^{\infty} n^2 P(N=n) - (E(Y_1))^2 \sum_{n=0}^{\infty} nP(N=n) \\
 &= E(Y_1^2) E(N) + (E(Y_1))^2 E(N^2) - (E(Y_1))^2 E(N) \\
 &= E(Y_1^2)\lambda + (E(Y_1))^2[\lambda^2 + \lambda - \lambda] \\
 &= E(Y_1^2)\lambda + (E(Y_1))^2\lambda^2
 \end{aligned}$$

Altogether, we have

$$\begin{aligned}
 \text{Var}(X) &= E(X^2) - (E(X))^2 \\
 &= E(Y_1^2)\lambda + (E(Y_1))^2\lambda^2 - (E(Y_1))^2\lambda^2 = E(Y_1^2)\lambda.
 \end{aligned}$$

3, D

4. (a)

seen ↓

$$\begin{aligned}
 E[(X - a)^2] &= E[\{X - E[X] + E[X] - a\}^2] \\
 &= E[\{(X - E[X]) + (E[X] - a)\}^2] \\
 &= E[(X - E[X])^2 + 2(X - E[X])(E[X] - a) + (E[X] - a)^2] \\
 &= E[(X - E[X])^2] + 2E[(X - E[X])(E[X] - a)] + E[(E[X] - a)^2] \quad (\text{linearity of } E).
 \end{aligned}$$

Now, since $(E[X] - a)$ is a constant,

$$\begin{aligned}
 2E[(X - E[X])(E[X] - a)] &= 2(E[X] - a)E[(X - E[X])] \quad (\text{linearity of } E) \\
 &= 2(E[X] - a)(E[X] - E[E[X]]) \quad (\text{linearity of } E) \\
 &= 2(E[X] - a)(E[X] - E[X]) \quad (\text{expected value of a constant } c \text{ is } c) \\
 &= 2(E[X] - a)(0) \\
 &= 0.
 \end{aligned}$$

Therefore,

$$\begin{aligned}
 E[(X - a)^2] &= E[(X - E[X])^2] + 0 + E[(E[X] - a)^2] \\
 &= E[(X - E[X])^2] + (E[X] - a)^2 \quad (\text{expected value of a constant } c \text{ is } c) \\
 &\geq E[(X - E[X])^2],
 \end{aligned}$$

since $(E[X] - a)^2 \geq 0$. Therefore,

$$\text{Var}(X) \leq E[(X - a)^2],$$

with equality when $a = E[X]$.

4, A

(b) We first sort the data:

meth seen ↓

$$\{18.1, 18.2, 20.3, 20.4, 20.5, 23.6, 23.7, 24.8, 24.9, 27.0, 28.1\}$$

Since there are $n = 11$ values, the median is the $i_m = (n+1)/2 = (11+1)/2 = 6$ th smallest value, i.e. $m = 23.6$.

1, A

Next, we compute the lower quartile as the $i_{0.25} = (i_m + 1)/2 = (6 + 1)/2 = 3.5$ th smallest value, i.e. a value between 20.3 and 20.4; we take the average $q_{0.25} = 20.35$.

The upper quartile is the $i_{0.75} = n - i_{0.25} + 1 = 11 - 3.5 + 1 = 8.5$, i.e. a value between 24.8 and 24.9; we take the average $q_{0.75} = 24.85$.

Therefore, the interquartile range is $\text{IQR} = q_{0.75} - q_{0.25} = 24.85 - 20.35 = 4.5$.

1, A

- (c) (i) Since the variance σ^2 is unknown, but the sample variance s^2 is available, we can use Student's t -distribution to construct a confidence interval for μ :

meth seen ↓

$$T = \frac{\mu - \bar{X}}{S/\sqrt{n}},$$

where the sample mean $\bar{X}$ and the sample variance S^2 are defined by

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i,$$

$$S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2.$$

For a $1 - \alpha$ confidence intervals, we find $t_{n-1, 1-\alpha/2} = -t_{n-1, \alpha/2}$ from Table 2 and then

1, A

$$\begin{aligned} P(t_{n-1, \alpha/2} < T < t_{n-1, 1-\alpha/2}) &= 1 - \alpha \\ P\left(-t_{n-1, 1-\alpha/2} < \frac{\mu - \bar{X}}{S/\sqrt{n}} < t_{n-1, 1-\alpha/2}\right) &= 1 - \alpha \\ P\left(-t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}} < \mu - \bar{X} < t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}}\right) &= 1 - \alpha \\ P\left(\bar{X} - t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}} < \mu < \bar{X} + t_{n-1, 1-\alpha/2} \frac{S}{\sqrt{n}}\right) &= 1 - \alpha. \end{aligned}$$

1, A

Since $1 - \alpha = 0.99$, this implies $\alpha = 0.01 \Rightarrow \alpha/2 = 0.005 \Rightarrow 1 - \alpha/2 = 0.995$, and since $n = 10 \Rightarrow n - 1 = 9$, and so $t_{n-1, 1-\alpha/2} = t_{9, 0.995} = 3.250$ from Table 2.

1, A

Therefore, the realised confidence interval for the mean μ is

$$\begin{aligned} \left(\bar{x} - t_{n-1, 1-\alpha/2} \frac{s}{\sqrt{n}}, \bar{x} + t_{n-1, 1-\alpha/2} \frac{s}{\sqrt{n}}\right) &= \left(172 - (3.250) \frac{4}{\sqrt{9}}, 172 + (3.250) \frac{4}{\sqrt{9}}\right) \\ &= \left(172 - \frac{13}{3}, 172 + \frac{13}{3}\right) \\ &\approx (167.67, 176.33), \end{aligned}$$

where any of the numerical expressions on the right hand side of the above equation are acceptable.

1, A

- (ii) There is a theorem in the notes which proves, for $X_1, \dots, X_n \stackrel{\text{i.i.d.}}{\sim} N(\mu, \sigma^2)$, that

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2.$$

1, A

We can construct a $1 - \alpha$ confidence interval as follows:

$$\begin{aligned} P\left(\chi_{n-1, \alpha/2}^2 < \frac{(n-1)S^2}{\sigma^2} < \chi_{n-1, 1-\alpha/2}^2\right) &= 1 - \alpha \\ P\left(\left\{\chi_{n-1, \alpha/2}^2 < \frac{(n-1)S^2}{\sigma^2}\right\} \text{ and } \left\{\frac{(n-1)S^2}{\sigma^2} < \chi_{n-1, 1-\alpha/2}^2\right\}\right) &= 1 - \alpha \\ P\left(\left\{\frac{\sigma^2}{(n-1)S^2} < \frac{1}{\chi_{n-1, \alpha/2}^2}\right\} \text{ and } \left\{\frac{1}{\chi_{n-1, 1-\alpha/2}^2} < \frac{\sigma^2}{(n-1)S^2}\right\}\right) &= 1 - \alpha \\ P\left(\frac{1}{\chi_{n-1, 1-\alpha/2}^2} < \frac{\sigma^2}{(n-1)S^2} < \frac{1}{\chi_{n-1, \alpha/2}^2}\right) &= 1 - \alpha \\ P\left(\frac{(n-1)S^2}{\chi_{n-1, 1-\alpha/2}^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_{n-1, \alpha/2}^2}\right) &= 1 - \alpha. \end{aligned}$$

1, B

Since $1 - \alpha = 0.95 \Rightarrow \alpha = 0.05 \Rightarrow \alpha/2 = 0.025$, using Table 3, we compute $n - 1 = 9$ and $\alpha/2 = 0.025$, then $\chi_{n-1, \alpha/2}^2 = \chi_{9, 0.025}^2 = 2.700$, and $1 - \alpha/2 = 0.975$, so from Table 4 $\chi_{n-1, 1-\alpha/2}^2 = \chi_{9, 0.975}^2 = 19.023$. Then the realised confidence interval is

$$\begin{aligned} \left(\frac{(n-1)s^2}{\chi_{n-1, 1-\alpha/2}^2}, \frac{(n-1)s^2}{\chi_{n-1, \alpha/2}^2}\right) &= \left(\frac{(9)(16)}{(19.023)}, \frac{(9)(16)}{(2.700)}\right) \\ &\approx (7.75, 53.33), \end{aligned}$$

where the final expression can be left as a pair of fractions.

2, A

- (d) It is given in the table of distributions that the probability density function of a $U(a, b)$ random variable is $\frac{1}{b-a}$, and so its cumulative distribution function evaluated at y is

unseen ↓

$$F(y) = \int_a^y \frac{1}{b-a} dt = \frac{y-a}{b-a},$$

and so the cumulative distribution function for each $X_i \sim U(0, \theta)$ is $\frac{y}{\theta}$.

1, D

We now compute the cumulative distribution function $F(y|\theta)$ for Y_n : for $0 < y < \theta$, since $X_1, \dots, X_n$ are independent,

$$\begin{aligned} F(y|\theta) &= P(Y_n \leq y|\theta) = P(\max\{X_1, \dots, X_n\} \leq y|\theta) \\ &= P(X_1 \leq y, \dots, X_n \leq y|\theta) \\ &= \prod_{i=1}^n P(X_i \leq y|\theta) \quad (\text{by independence}) \\ &= \prod_{i=1}^n \left(\frac{y}{\theta}\right) \quad (\text{c.d.f. of } U(0, \theta)) \\ &= \left(\frac{y}{\theta}\right)^n. \end{aligned}$$

2, D

Then we compute the probability density function of Y_n by taking the derivative of its cumulative distribution function:

$$\frac{d}{dy} \left(\frac{y}{\theta}\right)^n = \frac{d}{dy} \left(\frac{y^n}{\theta^n}\right) = \frac{ny^{n-1}}{\theta^n}.$$

1, D

We can then compute the expected value of Y_n as

$$E[Y_n] = \int_0^\theta y \left(\frac{ny^{n-1}}{\theta^n}\right) dy = \frac{n}{\theta^n} \int_0^\theta y^n dy = \frac{n}{\theta^n} \left[\frac{y^{n+1}}{n+1}\right]_{y=0}^{y=\theta} = \frac{n}{\theta^n} \left[\frac{\theta^{n+1}}{n+1}\right] = \frac{n}{n+1}\theta,$$

and so, using linearity of expectation,

1, C

$$E\left[\frac{n+1}{n}Y_n\right] = \frac{n+1}{n}E[Y_n] = \frac{n+1}{n} \left(\frac{n}{n+1}\theta\right) = \theta,$$

showing that $\frac{n+1}{n}Y_n$ is an unbiased estimator of θ , as required.

1, D

5. (a) (i) Let T be a random variable with cumulative distribution function F_T and inverse F_T^{-1} , which is assumed to exist. Let $Z = F_T(T)$.
 Since F_T is a cumulative distribution function, for any $t \in \mathbb{R}$, $0 \leq F_T(t) \leq 1$.
 Then $P(Z \leq 0) = P(F_T(T) \leq 0) = 0$, and $P(Z \leq 1) = P(F_T(T) \leq 1) = 1$.

1, B

Now, for any $z \in (0, 1)$,

$$\begin{aligned} P(Z \leq z) &= P(F_T(T) \leq z) \\ &= P(T \leq F_T^{-1}(z)) \\ &= F_T(F_T^{-1}(z)) \\ &= z. \end{aligned}$$

1, B

Therefore, the cumulative distribution function of the random variable T is

$$F_Z(Z) = P(Z \leq z) = \begin{cases} 0, & \text{if } z \leq 0, \\ z, & \text{if } 0 < z < 1, \\ 1, & \text{if } z \geq 1. \end{cases}$$

which is the cumulative distribution function of a $U(0, 1)$ distribution.

1, B

A p -value can be defined as $p = F_T(t)$ or $p = 1 - F_T(t)$, or something similar; either way, the above calculation shows that the p -value follows a $U(0, 1)$ distribution.

1, B

- (ii) A significance threshold α is a small value chosen in advance of conducting a hypothesis test. If the computed p -value is less than α , then one would reject the null hypothesis.

1, A

The Type I error is the probability of falsely rejecting the null hypothesis when the null hypothesis is actually true.

1, B

Since under the null hypothesis a p -value is uniformly distributed, the Type I error is equal to the significance threshold α .

1, C

- (b) (i) The null hypothesis H_0 and H_1 are:

H_0 : The person CANNOT taste the difference between Coca-Cola and Pepsi.

H_1 : The person CAN taste the difference between Coca-Cola and Pepsi.

(uppercase for emphasis).

2, A

- (ii) We compute the probability of the person getting **at least** 5 out of 6 glasses correct. First, the number of ways of choosing 6 glasses out of 12 is:

$$\binom{12}{6} = \frac{12!}{6!6!} = \frac{12 \times 11 \times \cdots \times 1}{(6 \times 5 \times 4 \times 3 \times 2 \times 1)^2}$$

$$= \frac{12 \times 11 \times 10 \times 9 \times 8 \times 7}{6 \times 5 \times 4 \times 3 \times 2 \times 1}$$

$$= \frac{2 \times 11 \times 2 \times 9 \times 2 \times 7}{1 \times 1 \times 1 \times 3 \times 2 \times 1}$$

$$= \frac{2 \times 11 \times 2 \times 3 \times 1 \times 7}{1 \times 1 \times 1 \times 1 \times 1 \times 1}$$

$$= 11 \times 12 \times 7$$

$$= 11 \times 84 = 924.$$

1, A

Then, the number of ways the person can choose all 6 Coca-Cola glasses correctly is the number of ways of choosing 6 out of 6 Coca-Cola glasses multiplied by choosing 0 out of 6 Pepsi glasses: $\binom{6}{6} \binom{6}{0} = (1)(1) = 1$.

Therefore, the probability of choosing 6 out of 12 glasses such that all 6 glasses are Coca-Cola is

$$\frac{\binom{6}{6} \binom{6}{0}}{\binom{12}{6}} = \frac{1}{924}.$$

1, A

The number of ways the person can choose 5 out of 6 Coca-Cola glasses correctly is the number of ways of choosing 5 out of 6 Coca-Cola glasses multiplied by choosing 1 out of 6 Pepsi glasses: $\binom{5}{6} \binom{6}{1} = (6)(6) = 36$.

Therefore, the probability of choosing 5 out of 12 glasses such that all 6 glasses are Coca-Cola is

$$\frac{\binom{5}{6}\binom{6}{1}}{\binom{12}{6}} = \frac{36}{924}.$$

1, B

Therefore the probability of choosing at least 5 out of 6 glasses of Coca-Cola correctly is

$$\frac{36}{924} + \frac{1}{924} = \frac{37}{924} < \frac{37}{900} < \frac{45}{900} = \frac{5}{100} = 0.05.$$

Therefore, there is evidence to suggest the person can taste the difference between Coca-Cola and Pepsi. A commonly-chosen significance threshold is $\alpha = 0.05$, which means they would only have obtained this result by chance with probability less than 5%.

1, C

(c) We recall the identity from lectures, $\text{Cov}(U, V) = E[UV] - E[U]E[V]$:

$$\begin{aligned}\text{Cov}(U, V) &= E[(U - \mu_U)(V - \mu_V)] \\ &= E[UV - \mu_U V - \mu_V U + \mu_U \mu_V] \\ &= E[UV] - \mu_U E[V] - \mu_V E[U] + E[\mu_U \mu_V] \quad (\text{linearity of } E) \\ &= E[UV] - \mu_U \mu_V - \mu_V \mu_U + \mu_U \mu_V \quad (\text{linearity of a constant } c \text{ is } c) \\ &= E[UV] - \mu_U \mu_V\end{aligned}$$

$$\Rightarrow \text{Cov}(U, V) = E[UV] - E[U]E[V].$$

1, D

Let $U = X + Y$ and $V = X - Y$. Then

$$E[UV] = E[(X + Y)(X - Y)] = E[X^2 - Y^2] = E[X^2] - E[Y^2],$$

and, using the linearity of expectation,

$$\begin{aligned}E[U]E[V] &= E[X + Y]E[X - Y] \\ &= (E[X] + E[Y])(E[X] - E[Y]) = (\mu_X + \mu_Y)(\mu_X - \mu_Y) = \mu_X^2 - \mu_Y^2.\end{aligned}$$

2, D

Therefore,

$$\begin{aligned}\text{Cov}(U, V) &= E[UV] - E[U]E[V] \\ &= \{E[X^2] - E[Y^2]\} - \{\mu_X^2 - \mu_Y^2\} \\ &= \{E[X^2] - \mu_X^2\} - \{E[Y^2] - \mu_Y^2\} \\ &= \{E[X^2] - (E[X])^2\} - \{E[Y^2] - (E[Y])^2\} \\ &= \text{Var}[X] - \text{Var}[Y],\end{aligned}$$

since we have from an exercise in the notes that $\text{Var}[Z] = E[Z^2] - (E[Z])^2$, for any random variable Z .

1, C

Then, since we are given that $\text{Var}[X] = \text{Var}[Y]$,

$$\text{Cov}(U, V) = \text{Var}[X] - \text{Var}[Y] = 0,$$

and therefore the correlation is $\rho(U, V) = 0$.

1, B

- (d) A high correlation does not imply causation, and so one cannot conclude that the number of undergraduate engineering degrees awarded in the US is related to or influenced by the amount of electricity generated in Cambodia, or vice versa.

2, C

6. (a) No, θ is not in (a, b) with 95% probability. In the frequentist view of statistics, the parameter θ is assumed to have a true, but unknown value, and so the true value is either in the realised interval or not.

1, B

If the experiment could be repeated a large number of times, and a confidence interval (a_i, b_i) could be computed for trial i , then, under the null hypothesis, the true parameter value θ would be in approximately 95% of the computed confidence intervals.

2, C

- (b) The pdf of the exponential distribution with parameter λ is

$$f(x) = \lambda \exp(-\lambda x), \quad x \in (0, \infty), \lambda > 0.$$

One first computes the median θ from the equation

$$\begin{aligned} \int_0^\theta \lambda \exp(-\lambda x) &= \frac{1}{2} \\ \Rightarrow [-\exp(-\lambda x)]_0^\theta &= \frac{1}{2} \\ \Rightarrow [-\exp(-\lambda \theta) - (-\exp(0))] &= \frac{1}{2} \\ \Rightarrow [-\exp(-\lambda \theta) + 1] &= \frac{1}{2} \\ \Rightarrow -\exp(-\lambda \theta) &= -\frac{1}{2} \\ \Rightarrow \exp(-\lambda \theta) &= \frac{1}{2} \\ \Rightarrow \log(\exp(-\lambda \theta)) &= \log \frac{1}{2} \\ \Rightarrow -\lambda \theta &= -\log 2 \\ \Rightarrow \theta &= \log 2 / \lambda. \end{aligned}$$

2, D

To find the maximum likelihood estimate of λ , we therefore need to find the maximum likelihood estimate of λ given the data $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$:

$$\begin{aligned} L(\lambda|\mathbf{x}) &= f(\mathbf{x}|\lambda) = \prod_{i=1}^n f(x_i|\lambda) \quad (\text{by independence of } x_1, \dots, x_n) \\ &= \prod_{i=1}^n \lambda \exp(-\lambda x_i) \\ &= \lambda^n \prod_{i=1}^n \exp(-\lambda x_i) \\ &= \lambda^n \exp(-\lambda \sum_{i=1}^n x_i) \\ \Rightarrow L(\lambda|\mathbf{x}) &= \lambda^n \exp(-\lambda n\bar{x}). \end{aligned}$$

1, A

For this function, we rather use the log-likelihood:

$$\begin{aligned}\log L(\lambda|\mathbf{x}) &= \log[\lambda^n] + \log[\exp(-\lambda n\bar{x})] \\ \Rightarrow \log L(\lambda|\mathbf{x}) &= n \log \lambda - \lambda n\bar{x},\end{aligned}$$

And then compute the maximum by taking the first derivative:

$$\frac{d}{d\lambda} \log L(\lambda|\mathbf{x}) = \frac{n}{\lambda} - n\bar{x}.$$

Setting the derivative equal to zero, we find

$$\frac{d}{d\lambda} \log L(\lambda|\mathbf{x}) = 0 \Leftrightarrow \frac{n}{\lambda} - n\bar{x} = 0 \Leftrightarrow \frac{1}{\lambda} = \bar{x} \Leftrightarrow \lambda = (\bar{x})^{-1}.$$

2, A

To check it is a maximum, we compute the second derivative

$$\frac{d^2}{d\lambda^2} \log L(\lambda|\mathbf{x}) = -\frac{n}{\lambda^2},$$

which is negative for all values of $\lambda > 0$, including $\lambda = (\bar{x})^{-1}$.

1, B

To check it is a global maximum, we check the boundary values of λ in its domain $(0, \infty)$:

$$\begin{aligned}\lim_{\lambda \rightarrow 0^+} L(\lambda|\mathbf{x}) &= \lim_{\lambda \rightarrow 0^+} \lambda^n \exp(-\lambda n\bar{x}) = 0 \\ \lim_{\lambda \rightarrow \infty} L(\lambda|\mathbf{x}) &= \lim_{\lambda \rightarrow \infty} \lambda^n \exp(-\lambda n\bar{x}) \\ &= \lim_{\lambda \rightarrow \infty} \frac{\lambda^n}{\exp(\lambda n\bar{x})} \\ &= \lim_{\lambda \rightarrow \infty} \frac{n!}{(n\bar{x})^n \exp(\lambda n\bar{x})} \quad (\text{L'Hospital's rule}) \\ &= 0,\end{aligned}$$

which shows that $L(\lambda|\bar{x})|_{\lambda=(\bar{x})^{-1}} = (\bar{x})^{-n} \exp(-n) > 0$ is a global maximum.

1, D

So $\hat{\lambda} = (\bar{x})^{-1}$ is the maximum likelihood estimate of λ , and so

$$\hat{\theta} = \log 2 / (\bar{x})^{-1} = \bar{x} \log 2$$

is the maximum likelihood estimate of the median θ .

(c) First, instead of trying to fit data to the model

$$y_i = \beta x_i^2 + e_i, \quad i \in \{1, 2, \dots, n\},$$

we define $w_i = x_i^2$ for $i \in \{1, 2, \dots, n\}$, and then consider the model

$$y_i = \beta w_i + e_i, \quad i \in \{1, 2, \dots, n\},$$

which appears more obviously linear.

1, D

We define the residual sum of squares as

$$\begin{aligned} \text{RSS}(\beta) &= \sum_{i=1}^n (\hat{e}_i)^2 = \sum_{i=1}^n [y_i - \beta w_i]^2 \\ &= \sum_{i=1}^n [y_i^2 - 2\beta y_i w_i + \beta^2 w_i^2] \\ &= \sum_{i=1}^n y_i^2 - 2\beta \sum_{i=1}^n y_i w_i + \beta^2 \sum_{i=1}^n w_i^2. \end{aligned}$$

To simplify this expression, let's define

$$Q_y = \sum_{i=1}^n y_i^2, \quad Q_{wy} = \sum_{i=1}^n w_i y_i, \quad Q_w = \sum_{i=1}^n w_i^2.$$

Then

$$\text{RSS}(\beta) = Q_y - 2\beta Q_{wy} + \beta^2 Q_w.$$

1, A

To maximise this function, taking the first derivative,

$$\frac{d\text{RSS}(\beta)}{d\beta} = -2Q_{wy} + 2\beta Q_w$$

and setting it equal to zero,

$$\frac{d\text{RSS}(\beta)}{d\beta} = 0 \Leftrightarrow \beta = \frac{Q_{wy}}{Q_w}.$$

We would still need to check the second derivative, which is

2, B

$$\frac{d\text{RSS}(\beta)}{d\beta} = 2Q_w = 2 \sum_{i=1}^n w_i^2 > 0,$$

since it is the sum of squares, and so $\beta = \frac{Q_{wy}}{Q_w}$ is a local minimum. It is a global minimum, because the expression is a quadratic with positive leading coefficient, and $\lim_{\beta \rightarrow \pm\infty} \text{RSS}(\beta) = \infty$.

2, B

An alternative approach is to completing the square,

$$\begin{aligned}
 \text{RSS}(\beta) &= Q_y - 2\beta Q_{wy} + \beta^2 Q_w \\
 &= Q_w \left(\beta^2 - 2\beta \frac{Q_{wy}}{Q_w} \right) + Q_y \\
 &= Q_w \left(\beta^2 - 2\beta \frac{Q_{wy}}{Q_w} + \frac{Q_{wy}^2}{Q_w^2} \right) + Q_y - \frac{Q_{wy}^2}{Q_w} \\
 &= Q_w \left(\beta - \frac{Q_{wy}}{Q_w} \right)^2 + Q_y - \frac{Q_{wy}^2}{Q_w},
 \end{aligned}$$

which means this expression is minimised when the square is zero and

$$\beta = \frac{Q_{wy}}{Q_w} = \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i^2} = \frac{\sum_{i=1}^n x_i^2 y_i}{\sum_{i=1}^n x_i^4}.$$

Either way the estimate is calculated,

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i^2 y_i}{\sum_{i=1}^n x_i^4}.$$

- (d) The residual plot indicates that the model does not fit the data well. 1, B
 The “U-shape” of the residuals indicates that they are correlated, but under the assumptions of the model they should be independent and identically distributed. 1, C
- (e) (i) This is not a bootstrap sample, since a bootstrap sample needs to be the same size as the original sample (this sample has 4 values while the original sample has 5 values) 1, A
- (ii) This is a bootstrap sample, since it has the same number of values as the original sample and all of the values in this new sample are in the original sample. 1, A

Review of mark distribution:

Total A marks: 48 of 48 marks

Total B marks: 30 of 30 marks

Total C marks: 18 of 18 marks

Total D marks: 24 of 24 marks

Total marks: 120 of 120 marks

Question Marker's comment

- 1 Q1 Overall, the students performed very well on Q1. Most of them correctly solved parts (a) and (b). Although the vast majority correctly solved part (c), many did not justify their steps, such as mentioning the use of Bayes' theorem or the law of total probability. Part (d) was where students performed the worst overall.
- 2 Q2(a) many calculation mistakes (e.g. sum starting at 0 instead of 1), need to explain rigorously how the PDF of X is found (using LTP), many students forgot to check for which values the PGF is defined.(b, i) students often forgot to check if the density is positive everywhere for $c=-3$.(b, ii and iii) many students forgot to indicate the support of the distribution.(b, iv, v, and vi) very often answered correctly, maybe a few calculation mistakes here and there.
- 3 Q3 Only a few students got full marks on this question. The majority did not do very well.3(a): Many students forgot that the sum of independent Poisson is still a Poisson distribution and therefore could not proceed. However, some students succeeded in proving this property from scratch. The question also requires the students to use conditional probability properly, but many struggled with this.3(b)(i): Students performed in general quite good on this one. Some quoted the result directly from the lecture, others used different methods to compute the moments. For students who used MGF as in the solution, many got stuck on converting the infinite sum to exp. A few students forgot that Poisson is discrete.3(b)(ii): Many students did not realize that N is a random variable, which means the attempts are wrong from the start. The proof is straightforward if the student uses law of total expectation. Most students did use the formula of LoTE properly.3(b)(iii): Only a few got this one correct. Similar to (ii), N has to be recognized as a random variable. The question requires the students to use the LoTE to find the $E(X^2)$ first. A few students computed the variance directly without the $E(X^2)$ which requires the use of covariance. Either way, the derivation involves a double summation which many students who attempted failed to compute correctly.
- 4 Q4 Most students did parts (a) and (b) correctly. For part (c), part (i) was answered well, but some students did not recognise part (ii) required a different approach. Part (d) was the question most students struggled with, especially the step deriving the cumulative distribution function from the distribution of the uniform.
- 5 Q5 - Q5(a)(i) was generally not attempted. - Q5(a)(ii) was completed well, with most students achieving full marks. - Q5(b)(i) was generally well done, although not all students got it right. - Q5(b)(ii) was not completed as well as expected. A common mistake was omitting the case where a person can choose 6 out of 6 Coca-Cola glasses correctly. It is worth noting that Q5(b) was very similar to a question on the problem sheet. - Q5(c) was very good, with most students achieving full marks and using different approaches to the formal solutions. - Q5(d) was also very good, with most students achieving full marks.

Question Marker's comment

- 6 Q6(a) Most students correctly identified that the given interpretation of the confidence interval was false and provided the correct interpretation.(b) A significant number of students mistakenly identified the median as the parameter of the exponential distribution. Additionally, many proofs for the MLE missed second-order condition or boundary condition checks.(c) The majority of students who attempted this question incorrectly identified β as the ratio of the mean of y to the mean of x . They mistakenly assumed that the term $\beta * x_i$ could be treated as a constant and applied the known fact that the value minimizing the mean squared error is the mean.(d) Most students correctly recognized that the model did not fit the data and provided accurate and complete answers to this question.(e) The vast majority of students correctly answered this question, demonstrating a solid understanding of the definition of a bootstrap sample.