NANYANG TECHNOLOGICAL UNIVERSITY
SEMESTER 1 EXAMINATION 2023–2024
MH2814 – PROBABILITY AND STATISTICS

Nov/Dec 2023 TIME ALLOWED: 2 HOURS

Matriculation Number:

Seat Number:

INSTRUCTIONS TO CANDIDATES

1. This examination paper contains **SIX (6)** questions and comprises **SEVENTEEN (17)** pages, including appendix from page 14 to page 17.
2. Tables of some probability distributions are provided in the Appendix.
3. Answer **ALL** questions. The marks for each question are indicated at the beginning of each question.
4. This is a **RESTRICTED OPEN BOOK** exam. You are only allowed to bring into the examination hall **ONE DOUBLE-SIDED A4-SIZE REFERENCE SHEET WITH TEXTS HANDWRITTEN OR TYPED ON THE A4 PAPER WITHOUT ANY ATTACHMENTS** (e.g. sticky notes, post-it notes, gluing or stapling of additional papers).
5. Candidates may use calculators. However, they should write down systematically the steps in the workings.
6. All your solutions should be written in this booklet within the space provided after each question. However, if you write your solutions on other pages, please indicate them clearly.
7. This examination paper is **NOT ALLOWED** to be removed from the examination hall.

For examiners only

Question	Marks
1 (15)	
2 (15)	
3 (20)	

Question	Marks
4 (15)	
5 (15)	
6 (20)	

TOTAL (100)

QUESTION 1. **(15 Marks)**

(a) A bag contains green and red balls, each labelled A or B (but not both). The proportion of green balls is 40%, and the proportion of balls labeled A is 24%. Given that a randomly selected ball is green, the probability that it is also labeled B is 0.8. Find the probability that a randomly selected ball is red, given that it is labeled A.

(b) 500 electronic devices were donated to a recycling company. These devices had an average usage time of 120 hours with a standard deviation of 15 hours. The company refurbishes and sells devices with 60 hours or less of usage time, and disposes of the rest. If a particular device ranks 40th lowest in terms of usage time, does it meet the criteria for sale according to Chebyshev's Theorem's?

Question 1 continues on page 3.

(c) A company manufactures consumer products with unknown mean shelf life and standard deviation, following a normal distribution. It is observed that:

(i) The top 10% of products have a shelf life exceeding 36 months.

(ii) The bottom 33% of products have a shelf life less than 24 months.

Find the standard deviation.

End of Question 1.

QUESTION 2. **(15 Marks)**

In a busy harbor, ships arrive at a rate of 3 ships per day on average, following a Poisson distribution.

(a) Find the probability that on a particular day, at least 2 ships arrive at the harbour.

(b) Find the probability that at least 6 ships arrive at the harbour in a particular 5-day period.

Question 2 continues on page 5.

(c) Find the probability that during a particular 12-day period, there are at least 6 days with at least 2 ships arriving at the harbour per day.

End of Question 2.

QUESTION 3. **(20 Marks)**

The joint density function of X and Y is given as:

$$f(x,y) = \begin{cases} x, & \text{if } 0 < 3x < y \leq 3 \quad \text{or} \quad 0 < 3x < 6 - y < 3; \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Find $P(Y < 1)$.

(b) Find the marginal distributions of X and Y, respectively.

Question 3 continues on page 7.

(c) Find $P(Y > 2|X = 0.5)$.

(d) Find $E(X)$ and $E(Y)$, the expected value of X and Y, respectively.

(e) Find σ_{XY}, the covariance of X and Y.

End of Question 3.

QUESTION 4. **(15 Marks)**

(a) A statistician is studying the distribution of scores for a national exam taken by 10,000 students based on scores from a sample of 35 students. The sample mean and sample variance of the 35 scores are computed, and they are found to be 76 and 5, respectively. The statistician makes the following claims:

(i) By applying the Central Limit Theorem, we may assume that scores of the 10,000 students follows a normal distribution.

(ii) A good estimate for the population variance is $5/\sqrt{34}$.

(iii) Approximately 9,500 of the students have scores lying within the interval (74.3, 77.7).

For each of these claims (i), (ii), and (iii), determine whether it is valid or invalid and provide an explanation for your answer.

Question 4 continues on page 9.

(b) Concrete cylinders are frequently subjected to tests to determine their compressive strength, a critical factor in structural design. A quality control engineer aims to estimate the mean compressive strength of concrete cylinders produced by a manufacturing process. A random sample of 20 concrete cylinders is selected, and their compressive strengths, x_i, are recorded in megapascals (MPa). The following summary statistics are computed:

$$\sum_{i=1}^{20} x_i = 464, \qquad \sum_{i=1}^{20} x_i^2 = 10818,$$

Calculate a 90% confidence interval for the mean compressive strength of concrete cylinders produced by this manufacturing process, assuming that the compressive strengths follow an approximately normal distribution.

End of Question 4.

QUESTION 5. **(15 Marks)**

A study is conducted to determine whether the consumption of avocado affects cholesterol levels. To investigate this, a sample of 15 participants who consume avocados regularly and a sample of 12 participants who do not consume avocados were selected. Their cholesterol levels (in mg/dL) were measured.

Group A (consumes avocados): $n_A = 15,\ x_A = 195,\ s_A = 16,$

Group B (does not consume avocados): $n_B = 12,\ x_B = 207,\ s_B = 21,$

where n_i, x_i, and s_i are the sample size, sample mean, and sample standard deviation of group i, respectively.

Test at a 5% level of significance to determine if there is evidence that consuming avocados is associated with lower mean cholesterol levels.

(a) State the null hypothesis, H_0.

(b) State the alternative hypothesis, H_1.

(c) State the probability of wrongly rejecting your null hypothesis.

Question 5 continues on page 11.

(d) Assume the two sample groups came from normal populations with different variance, determine the test statistics and an approximate p-value, based on the sample data.

(e) State your decision and conclusion.

End of Question 5.

QUESTION 6. **(20 Marks)**

A team of environmental scientists is conducting a study to understand the relationship between the amount of rainfall (in millimeters) and the quantity of air pollution removed (in kilograms) in a specific urban area. They collected data from 20 different days and recorded the amount of rainfall x_i (in millimeters) and the amount of particulate removed y_i (in μ g/m^3) and recorded the following summary statistics:

$$n = 20, \qquad \bar{x} = 6.4, \qquad \bar{y} = 137, \qquad S_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = -127.8,$$

$$S_{xx} = \sum_{i=1}^{n}(x_i - \bar{x})^2 = 18.3, \qquad S_{yy} = \sum_{i=1}^{n}(y_i - \bar{y})^2 = 934.5.$$

(a) Find the estimated simple linear regression line $E(Y) = \beta_0 + \beta_1 \cdot X$, with this dataset.

(b) With this dataset, is there evidence to reject the claim that the slope parameter of the regression model is -6.1, at 0.05 level of significance? Justify your answer with an appropriate statistical test.

Question 6 continues on page 13.

(c) Based on the simple linear model in Part (a), what would be the average amount of particulate removed when the amount of rainfall is 8 millimeters? Give an estimate of this score with a 95% confidence interval.

END OF PAPER

• Z-distribution.
Commonly used z_α where $P(Z > z_\alpha) = \alpha$:

$z_{0.05} = 1.645$	$z_{0.10} = 1.283$
$z_{0.025} = 1.960$	$z_{0.01} = 2.328$
$z_{0.005} = 2.575$	$z_{0.02} = 2.054$

• t-distribution.
Commonly used t_α where $P(T > t_\alpha) = \alpha$, for $\nu = 15$ to $\nu = 30$:

	α						
ν	0.20	0.15	0.10	0.05	0.025	0.02	0.015
15	0.866	1.074	1.341	1.753	2.131	2.249	2.397
16	0.865	1.071	1.337	1.746	2.120	2.235	2.382
17	0.863	1.069	1.333	1.740	2.110	2.224	2.368
18	0.862	1.067	1.330	1.734	2.101	2.214	2.356
19	0.861	1.066	1.328	1.729	2.093	2.205	2.346
20	0.860	1.064	1.325	1.725	2.086	2.197	2.336
21	0.859	1.063	1.323	1.721	2.080	2.189	2.328
22	0.858	1.061	1.321	1.717	2.074	2.183	2.320
23	0.858	1.060	1.319	1.714	2.069	2.177	2.313
24	0.857	1.059	1.318	1.711	2.064	2.172	2.307
25	0.856	1.058	1.316	1.708	2.060	2.167	2.301
26	0.856	1.058	1.315	1.706	2.056	2.162	2.296
27	0.855	1.057	1.314	1.703	2.052	2.158	2.291
28	0.855	1.056	1.314	1.701	2.048	2.154	2.286
29	0.854	1.055	1.313	1.699	2.045	2.150	2.278
30	0.854	1.055	1.310	1.697	2.042	2.147	2.250

- Binomial Probability Sums $\sum_{x=0}^{r} b(r;n,p)$ for $n = 12, 13, 14$

		p									
n	r	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**
12	**0**	0.2824	0.0687	0.0317	0.0138	0.0022	0.0002	0.0000			
	1	0.6590	0.2749	0.1584	0.0850	0.0196	0.0032	0.0003	0.0000		
	2	0.8891	0.5583	0.3907	0.2528	0.0834	0.0193	0.0028	0.0002	0.0000	
	3	0.9744	0.7946	0.6488	0.4925	0.2253	0.0730	0.0153	0.0017	0.0001	
	4	0.9957	0.9274	0.8424	0.7237	0.4382	0.1938	0.0573	0.0095	0.0006	0.0000
	5	0.9995	0.9806	0.9456	0.8822	0.6652	0.3872	0.1582	0.0386	0.0039	0.0001
	6	0.9999	0.9961	0.9857	0.9614	0.8418	0.6128	0.3348	0.1178	0.0194	0.0005
	7	1.0000	0.9994	0.9972	0.9905	0.9427	0.8062	0.5618	0.2763	0.0726	0.0043
	8		0.9999	0.9996	0.9983	0.9847	0.9270	0.7747	0.5075	0.2054	0.0256
	9		1.0000	1.0000	0.9998	0.9972	0.9807	0.9166	0.7472	0.4417	0.1109
	10				1.0000	0.9997	0.9968	0.9804	0.9150	0.7251	0.3410
	11					1.0000	0.9998	0.9978	0.9862	0.9313	0.7176
	12						1.0000	1.0000	1.0000	1.0000	1.0000
13	**0**	0.2542	0.0550	0.0238	0.0097	0.0013	0.0001	0.0000			
	1	0.6213	0.2336	0.1267	0.0637	0.0126	0.0017	0.0001	0.0000		
	2	0.8661	0.5017	0.3326	0.2025	0.0579	0.0112	0.0013	0.0001		
	3	0.9658	0.7473	0.5843	0.4206	0.1686	0.0461	0.0078	0.0007	0.0000	
	4	0.9935	0.9009	0.7940	0.6543	0.3530	0.1334	0.0321	0.0040	0.0002	
	5	0.9991	0.9700	0.9198	0.8346	0.5744	0.2905	0.0977	0.0182	0.0012	0.0000
	6	0.9999	0.9930	0.9757	0.9376	0.7712	0.5000	0.2288	0.0624	0.0070	0.0001
	7	1.0000	0.9988	0.9944	0.9818	0.9023	0.7095	0.4256	0.1654	0.0300	0.0009
	8		0.9998	0.9990	0.9960	0.9679	0.8666	0.6470	0.3457	0.0991	0.0065
	9		1.0000	0.9999	0.9993	0.9922	0.9539	0.8314	0.5794	0.2527	0.0342
	10			1.0000	0.9999	0.9987	0.9888	0.9421	0.7975	0.4983	0.1339
	11				1.0000	0.9999	0.9983	0.9874	0.9363	0.7664	0.3787
	12					1.0000	0.9999	0.9987	0.9903	0.9450	0.7458
	13						1.0000	1.0000	1.0000	1.0000	1.0000
14	**0**	0.2288	0.0440	0.0178	0.0068	0.0008	0.0001	0.0000			
	1	0.5846	0.1979	0.1010	0.0475	0.0081	0.0009	0.0001			
	2	0.8416	0.4481	0.2811	0.1608	0.0398	0.0065	0.0006	0.0000		
	3	0.9559	0.6982	0.5213	0.3552	0.1243	0.0287	0.0039	0.0002		
	4	0.9908	0.8702	0.7415	0.5842	0.2793	0.0898	0.0175	0.0017	0.0000	
	5	0.9985	0.9561	0.8883	0.7805	0.4859	0.2120	0.0583	0.0083	0.0004	
	6	0.9998	0.9884	0.9617	0.9067	0.6925	0.3953	0.1501	0.0315	0.0024	0.0000
	7	1.0000	0.9976	0.9897	0.9685	0.8499	0.6047	0.3075	0.0933	0.0116	0.0002
	8		0.9996	0.9978	0.9917	0.9417	0.7880	0.5141	0.2195	0.0439	0.0015
	9		1.0000	0.9997	0.9983	0.9825	0.9102	0.7207	0.4158	0.1298	0.0092
	10			1.0000	0.9998	0.9961	0.9713	0.8757	0.6448	0.3018	0.0441
	11				1.0000	0.9994	0.9935	0.9602	0.8392	0.5519	0.1584
	12					0.9999	0.9991	0.9919	0.9525	0.8021	0.4154
	13					1.0000	0.9999	0.9992	0.9932	0.9560	0.7712
	14						1.0000	1.0000	1.0000	1.0000	1.0000

- Poisson Probability Sums $\sum_{x=0}^{r} p(x;\mu)$

	μ								
r	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863

	μ								
r	**10.0**	**11.0**	**12.0**	**13.0**	**14.0**	**15.0**	**16.0**	**17.0**	**18.0**
0	0.0000	0.0000	0.0000						
1	0.0005	0.0002	0.0001	0.0000	0.0000				
2	0.0028	0.0012	0.0005	0.0002	0.0001	0.0000	0.0000		
3	0.0103	0.0049	0.0023	0.0011	0.0005	0.0002	0.0001	0.0000	0.0000
4	0.0293	0.0151	0.0076	0.0037	0.0018	0.0009	0.0004	0.0002	0.0001
5	0.0671	0.0375	0.0203	0.0107	0.0055	0.0028	0.0014	0.0007	0.0003
6	0.1301	0.0786	0.0458	0.0259	0.0142	0.0076	0.0040	0.0021	0.0010
7	0.2202	0.1432	0.0895	0.0540	0.0316	0.0180	0.0100	0.0054	0.0029
8	0.3328	0.2320	0.1550	0.0998	0.0621	0.0374	0.0220	0.0126	0.0071
9	0.4579	0.3405	0.2424	0.1658	0.1094	0.0699	0.0433	0.0261	0.0154
10	0.5830	0.4599	0.3472	0.2517	0.1757	0.1185	0.0774	0.0491	0.0304
11	0.6968	0.5793	0.4616	0.3532	0.2600	0.1848	0.1270	0.0847	0.0549
12	0.7916	0.6887	0.5760	0.4631	0.3585	0.2676	0.1931	0.1350	0.0917
13	0.8645	0.7813	0.6815	0.5730	0.4644	0.3632	0.2745	0.2009	0.1426
14	0.9165	0.8540	0.7720	0.6751	0.5704	0.4657	0.3675	0.2808	0.2081
15	0.9513	0.9074	0.8444	0.7636	0.6694	0.5681	0.4667	0.3715	0.2867
16	0.9730	0.9441	0.8987	0.8355	0.7559	0.6641	0.5660	0.4677	0.3751
17	0.9857	0.9678	0.9370	0.8905	0.8272	0.7489	0.6593	0.5640	0.4686
18	0.9928	0.9823	0.9626	0.9302	0.8826	0.8195	0.7423	0.6550	0.5622
19	0.9965	0.9907	0.9787	0.9573	0.9235	0.8752	0.8122	0.7363	0.6509
20	0.9984	0.9953	0.9884	0.9750	0.9521	0.9170	0.8682	0.8055	0.7307

- Standard Normal Cumulative Probability Distribution

$$P(Z < z) \text{ for } z > 0.$$

z	**.00**	**.01**	**.02**	**.03**	**.04**	**.05**	**.06**	**.07**	**.08**	**.09**
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
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

- Note that for $z > 0$, we can calculate $P(Z < -z)$ by

$$P(Z < -z) = P(Z > z) = 1 - P(Z < z).$$

MH2814 PROBABILITY & STATISTICS

Please read the following instructions carefully:

1. **Please do not turn over the question paper until you are told to do so. Disciplinary action may be taken against you if you do so.**

2. You are not allowed to leave the examination hall unless accompanied by an invigilator. You may raise your hand if you need to communicate with the invigilator.

3. Please write your Matriculation Number on the front of the answer book.

4. Please indicate clearly in the answer book (at the appropriate place) if you are continuing the answer to a question elsewhere in the book.