SEMESTER 1
2021-2022

EE304FZ
Probability and Statistics

Dr. Lo Fook Loong, Prof. R. Farrell, Dr Jim Harkin

Time allowed: 2 hours

Question 1 is **compulsory** and is worth 40 marks.

Answer ***three*** questions from the remaining four. Each is worth 20 marks.

Instructions

	Yes	No
Log Books Allowed	ϒ	X
Formula Tables Allowed	X	ϒ
Other Allowed *(enter details)*	X	ϒ

General *(enter details)*

Formula tables are attached to the end of the exam paper.
Non-programmable calculators are allowed.

QUESTION 1

(a) A password can be formed from capital letters 'A' to 'Z', small letters 'a' to 'z' and digits '0' to '9'.

(i) If the password must be 8 characters long and contain at least one capital letter, one small letter and one digit, how many possible passwords are there? (3 marks)

(ii) If the password must be 9 characters long, must contain at least one small letter, one capital letter, one digit and just one of the following special characters '@', '#', '$' and '%', how many possible passwords are there? (2 marks)

(b) When a player in a poker game has the 5 cards Ten, Jack, Queen, King and Ace of any suit, then he is said to be holding a royal flush. What is the probability of getting a royal flush if 5 random cards are taken from a well shuffled pack of 52 cards? (5 marks)

(c) The continuous random variable X has a pdf given below:

$$f(x) = \begin{cases} 0 & x < 0 \\ kx & 0 \leq x < 5 \\ 0 & x > 5 \end{cases}$$

(i) Find the constant *k*. (4 marks)

(ii) Calculate the expected value of X, E(X). (3 marks)

(iii) Find the variance of X, Var(X). (3 marks)

(d) A bottling factory fills each bottle with juice, and it is known that the volume in the bottles has a normal distribution. The mean volume is 200 ml, and standard deviation is 10 ml. What is the probability that a bottle has less than 185 ml of juice? (5 marks)

(The Question is continued on the next page)

QUESTION 1 (continued)

(e) An electricity supply company is interested to know how fast it can respond to reported power outages. The company wants the mean time from report to arrival at the area of power outage to be within 20 minutes. It recorded the response time from 40 outages and found it to have a mean of 18 mins. The standard deviation is known to be 3 mins. Can the company be 95% sure that the response time is acceptable? (5 marks)

(f) The mean weight of a bag of flour is supposed to be $\mu = 500$g. It is known that the standard deviation is $\sigma = 10$g. The null hypothesis is $\mu = 500$, and the alternative hypothesis is $\mu \neq 500$. We will reject the null hypothesis if the sampled means of 50 bags $\bar{x}$ is such that $\bar{x} > 502$ or $\bar{x} < 498$.

(i) What is the test statistic Y that is to be used? (2 marks)

(ii) With the help of the standard normal table, obtain the probability of a Type I error. (3 marks)

(g) The relationship between a dependent variable Y and a variable x is $Y = \beta_0 + \beta_1 x + \epsilon$.

(i) One hypothesis to test for linear regression analysis is whether $\beta_1 = 0$. Briefly explain why. (3 marks)

(ii) Which distributions can be used to test the hypothesis above. (2 marks)

QUESTION 2

Over a long period of time in a town, it was found that the proportion of sunny days is 0.6, cloudy days is 0.3 and rainy days is 0.1. Note that these three types of days are mutually exclusive. It was also found that cloudy and rainy days are twice as likely to be windy compared to sunny days. The record shows that 20% of the days are windy.

(a) Draw the Venn diagram showing the events sunny (S), cloudy (C), rainy (R) and windy (W). (5 marks)

(b) What is the probability that the town has windy conditions given that it is sunny? (5 marks)

(c) What is the probability that the town experience a sunny day given that it is windy? (5 marks)

(d) What is the probability that a day is not rainy but windy, P(~R,W)? (5 marks)

QUESTION 3

Consider the 18-letter string “TO BE OR NOT TO BE”.

(a) If 18 cards each of which has one of the letters above written on it are placed in a bag, what is the probability of picking out the 5 cards with spaces on them without replacement? (5 marks)

(b) How many different arrangements can be made from these 18 characters? (5 marks)

(c) A keyboard with 27 keys, from ‘A’ to ‘Z’, and space (‘ ’) is connected to a monitor screen. What is the number of different 18-letter strings that can be produced from this keyboard? (5 marks)

(d) If a monkey pressed the keys on the keyboard randomly at one keypress per second, what is the expected amount of time before the string above will appear on the monitor screen? (5 marks)

QUESTION 4

A factory making expensive pieces of a delicate sensor needs to check how often each sensor can take a shock before it fails. The results on tests on 10 pieces of sensor is shown below:

10	12	11	13	12	15	13	11	14	13

A Student t-distribution is used to construct a 98% confidence interval for the mean number of shocks before failure.

(a) What is the sample mean for the number of shocks the sensor can take before it fails? (4 marks)

(b) What is the sample standard deviation for the number of shocks? (6 marks)

(c) From the t-distribution table, obtain the critical values for a 98% confidence interval for the mean. (4 marks)

(d) Calculate the range of the estimated mean number of shocks within that 98% confidence interval. (6 marks)

QUESTION 5

The number of customers arriving at a bank is thought to have a Poisson distribution. The number of customers walking into the bank was recorded for a total of 150 randomly selected 1-minute intervals. The result is shown below:

No. of customers	0	1	2	3	>4
Frequencies	27	46	44	33	0

From the table, it can be seen, for example, that there were 27 1-minute intervals where no customer stepped into the bank.

(a) What is the estimated parameter λ? (5 marks)

(b) What is the test statistic to be used in a χ^2 test? (3 marks)

(c) The null hypothesis H_0 is that the arrivals of customers follow a Poisson process. What value must χ_0^2 be greater than to reject H_0 for a 5% significance level? (4 marks)

(d) Calculate the test statistic value. (5 marks)

(e) Explain briefly if we can reject the null hypothesis for significance level of 5%. (3 marks)

Probability

- Basic Probability

$$0 \leq P(E) \leq 1 \qquad \forall E$$

$$P(E \cup F) = P(E) + P(F) - P(E \cap F)$$

$$P(E|F) = \frac{P(E \cap F)}{P(F)}$$

- If $F_1, F_2, ..., F_n$ is a collection of mutually exclusive and exhaustive events, then;

$$P(E) = P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + ... + P(E|F_n)P(F_n)$$

$$P(F_j|E) = \frac{P(E|F_j)P(F_j)}{P(E|F_1)P(F_1) + P(E|F_2)P(F_2) + ... + P(E|F_n)P(F_n)}$$

- Binomial distribution with parameters n and p:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

- Negative binomial distribution with parameters k and r:

$$P(X_r = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

- Poisson distribution with mean λt:

$$P(X = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

- Cumulative distribution of exponential distribution with parameters λ:

$$F(x) = 1 - e^{-\lambda x}$$

- Cumulative distribution of Weibull distribution with shape parameter β and scale parameter δ:

$$F(x) = 1 - e^{-\left(\frac{x}{\delta}\right)^{\beta}}$$

Statistics

$\sim \mathcal{N}(0,1)$ has normal distribution with mean 0 and standard deviation 1.
$\sim t_p$ has t-distribution with p degrees of freedom.
$\sim \chi^2_p$ has chi-squared distribution with p degrees of freedom.

- Estimation of mean (large sample):

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$$

- Estimation of mean (small sample):

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}} \sim t_{n-1}$$

- Estimation of proportion (large sample):

$$Z = \frac{\hat{P} - p}{\sqrt{\frac{p(1-p)}{n}}} \sim \mathcal{N}(0,1)$$

- Chi-squared goodness to fit:

$$\chi^2 = \sum_{i=1}^{n} \frac{(E_i - O_i)^2}{E_i} \sim \chi^2_{n-p-1}$$

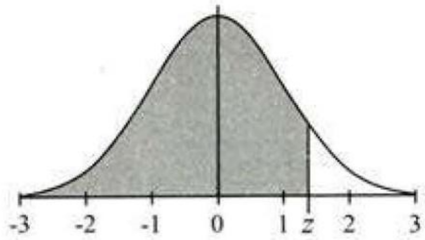
- Chi-squared contingency table:

$$\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(E_{ij} - O_{ij})^2}{E_{ij}} \sim \chi^2_{(r-1)(c-1)}$$

Dóchúlachtaí don dáileadh normalach caighdeánach

I gcás z a thugtar, faightear ón tábla

$$P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2}\, dt$$



Probabilities for the standard normal distribution

For a given z, the table gives

$$P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}t^2}\, dt$$

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	0.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	0.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	0.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	0.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	0.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	0.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	0.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	0.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	0.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	0.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621

Clár / Contents

an dáileadh normalach (ar lean) — normal distribution (continued)

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.1	0.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	0.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	0.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	0.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	0.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	0.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	0.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	0.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	0.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	0.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	0.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	0.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	0.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	0.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	0.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	0.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	0.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	0.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	0.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	0.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

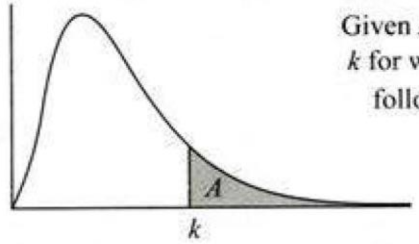
Clár Staitisticí / Statistics Contents

Clár / Contents

Dáileadh chí-chearnaithe

luachanna criticiúla tástála aonfhoircní

Nuair a thugtar A, faightear ón tábla an luach ar k mar a bhfuil $P(X > k) = A$, áit a leanann X dáileadh chí-chearnaithe a bhfuil v céim saoirse aige.



Chi-squared distribution

one-tailed critical values

Given A, the table gives the value of k for which $P(X > k) = A$, where X follows a chi-squared distribution with v degrees of freedom.

v \ A	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
1	0.0000	0.0002	0.0010	0.0039	3.8415	5.0239	6.6349	7.8794
2	0.0100	0.0201	0.0506	0.1026	5.9915	7.3778	9.2103	10.597
3	0.0717	0.1148	0.2158	0.3518	7.8147	9.3484	11.345	12.838
4	0.2070	0.2971	0.4844	0.7107	9.4877	11.143	13.277	14.860
5	0.4117	0.5543	0.8312	1.1455	11.070	12.833	15.086	16.750
6	0.6757	0.8721	1.2373	1.6354	12.592	14.449	16.812	18.548
7	0.9893	1.2390	1.6899	2.1673	14.067	16.013	18.475	20.278
8	1.3444	1.6465	2.1797	2.7326	15.507	17.535	20.090	21.955
9	1.7349	2.0879	2.7004	3.3251	16.919	19.023	21.666	23.589
10	2.1559	2.5582	3.2470	3.9403	18.307	20.483	23.209	25.188
11	2.6032	3.0535	3.8157	4.5748	19.675	21.920	24.725	26.757
12	3.0738	3.5706	4.4038	5.2260	21.026	23.337	26.217	28.300
13	3.5650	4.1069	5.0088	5.8919	22.362	24.736	27.688	29.819
14	4.0747	4.6604	5.6287	6.5706	23.685	26.119	29.141	31.319

Clár / Contents

dáileadh chí-chearnaithe (ar lean) chi-squared distribution (continued)

v \ A	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005
15	4.6009	5.2293	6.2621	7.2609	24.996	27.488	30.578	32.801
16	5.1422	5.8122	6.9077	7.9616	26.296	28.845	32.000	34.267
17	5.6972	6.4078	7.5642	8.6718	27.587	30.191	33.409	35.718
18	6.2648	7.0149	8.2307	9.3905	28.869	31.526	34.805	37.156
19	6.8440	7.6327	8.9065	10.117	30.144	32.852	36.191	38.582
20	7.4338	8.2604	9.5908	10.851	31.410	34.170	37.566	39.997
21	8.0337	8.8972	10.283	11.591	32.671	35.479	38.932	41.401
22	8.6427	9.5425	10.982	12.338	33.924	36.781	40.289	42.796
23	9.2604	10.196	11.689	13.091	35.172	38.076	41.638	44.181
24	9.8862	10.856	12.401	13.848	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	67.505	71.420	76.154	79.490
60	35.534	37.485	40.482	43.188	79.082	83.298	88.379	91.952
70	43.275	45.442	48.758	51.739	90.531	95.023	100.43	104.21
80	51.172	53.540	57.153	60.391	101.88	106.63	112.33	116.32
90	59.196	61.754	65.647	69.126	113.15	118.14	124.12	128.30
100	67.328	70.065	74.222	77.929	124.34	129.56	135.81	140.17

- 39 -

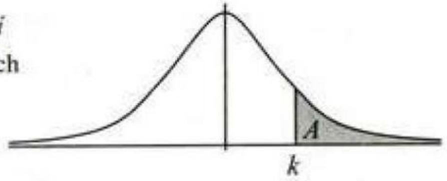
Clár Staitisticí / Statistics Contents

Clár / Contents

t-dháileadh Student

luachanna criticiúla tástála aonfhoirceanna

Nuair a thugtar A, faightear ón tábla an luach ar k mar a bhfuil $P(T > k) = A$, áit a leanann T, t-dháileadh a bhfuil v céim saoirse aige.



Student's t-distribution

one-tailed critical values

Given A, the table gives the value of k for which $P(T > k) = A$, where T follows a t-distribution with v degrees of freedom.

v \ A	0.1	0.05	0.025	0.01	0.005	0.001	0.0005	0.0001	0.00005
1	3.078	6.314	12.71	31.82	63.66	318.3	636.6	3183	6366
2	1.886	2.920	4.303	6.965	9.925	22.33	31.60	70.70	99.99
3	1.638	2.353	3.182	4.541	5.841	10.21	12.92	22.20	28.00
4	1.533	2.132	2.776	3.747	4.604	7.173	8.610	13.03	15.54
5	1.476	2.015	2.571	3.365	4.032	5.893	6.869	9.678	11.18
6	1.440	1.943	2.447	3.143	3.707	5.208	5.959	8.025	9.082
7	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7.063	7.885
8	1.397	1.860	2.306	2.896	3.355	4.501	5.041	6.442	7.120
9	1.383	1.833	2.262	2.821	3.250	4.297	4.781	6.010	6.594
10	1.372	1.812	2.228	2.764	3.169	4.144	4.587	5.694	6.211
11	1.363	1.796	2.201	2.718	3.106	4.025	4.437	5.453	5.921
12	1.356	1.782	2.179	2.681	3.055	3.930	4.318	5.263	5.694
13	1.350	1.771	2.160	2.650	3.012	3.852	4.221	5.111	5.513
14	1.345	1.761	2.145	2.624	2.977	3.787	4.140	4.985	5.363

Clár / Contents

t-dháileadh Student (ar lean) — Student's *t*-distribution (continued)

v \ A	0.1	0.05	0.025	0.01	0.005	0.001	0.0005	0.0001	0.00005
15	1.341	1.753	2.131	2.602	2.947	3.733	4.073	4.880	5.239
16	1.337	1.746	2.120	2.583	2.921	3.686	4.015	4.790	5.134
17	1.333	1.740	2.110	2.567	2.898	3.646	3.965	4.715	5.043
18	1.330	1.734	2.101	2.552	2.878	3.610	3.922	4.648	4.966
19	1.328	1.729	2.093	2.539	2.861	3.579	3.883	4.590	4.899
20	1.325	1.725	2.086	2.528	2.845	3.552	3.850	4.539	4.838
21	1.323	1.721	2.080	2.518	2.831	3.527	3.819	4.492	4.785
22	1.321	1.717	2.074	2.508	2.819	3.505	3.792	4.452	4.736
23	1.319	1.714	2.069	2.500	2.807	3.485	3.768	4.416	4.694
24	1.318	1.711	2.064	2.492	2.797	3.467	3.745	4.382	4.654
25	1.316	1.708	2.060	2.485	2.787	3.450	3.725	4.352	4.619
26	1.315	1.706	2.056	2.479	2.779	3.435	3.707	4.324	4.587
27	1.314	1.703	2.052	2.473	2.771	3.421	3.689	4.299	4.556
28	1.313	1.701	2.048	2.467	2.763	3.408	3.674	4.276	4.531
29	1.311	1.699	2.045	2.462	2.756	3.396	3.660	4.254	4.505
30	1.310	1.697	2.042	2.457	2.750	3.385	3.646	4.234	4.482
40	1.303	1.684	2.021	2.423	2.704	3.307	3.551	4.094	4.321
50	1.299	1.676	2.009	2.403	2.678	3.261	3.496	4.014	4.228
60	1.296	1.671	2.000	2.390	2.660	3.232	3.460	3.962	4.169
80	1.292	1.664	1.990	2.374	2.639	3.195	3.416	3.899	4.095
100	1.290	1.660	1.984	2.364	2.626	3.174	3.390	3.861	4.054
∞	1.282	1.645	1.960	2.326	2.576	3.090	3.290	3.719	3.891

Clár Staitisticí / Statistics Contents

Clár / Contents