Coláiste na Tríonóide, Baile Átha Cliath
Trinity College Dublin
Ollscoil Átha Cliath | The University of Dublin

Faculty of Science, Technology, Engineering and Mathematics

School of Computer Science & Statistics
Statistics and Information Systems

SF MSISS — **Trinity Term 2021**
SF Mathematics
SF and JF Mathematics combinations
SF Computer Science

STU22005: Applied Probability II

17/05/2021 — **Online real-time exam** — **12:00 - 14:00**

Prof. Caroline Brophy

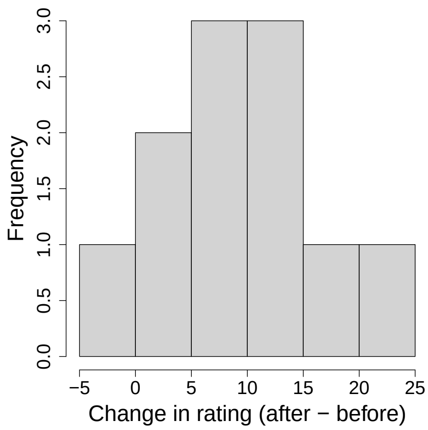
Instructions to Candidates:

Full marks will be awarded for complete solutions to all **three** questions. Each full question is worth either **33 or 34 marks**, and marks for component parts are indicated in brackets. Statistical tables have been provided at the end of the exam paper.

Materials permitted for this examination:

This is an open book examination. This is an individual assessment and collaboration is **not** permitted. You must complete and submit the declaration form included to confirm you understand this.

1. A large company introduced a range of new initiatives to improve employee satisfaction. They selected 11 employees at random from the company and measured their change in satisfaction ratings before and after the initiatives were introduced (change = after - before). The mean change was 9.36, with standard deviation 6.727. The values are displayed in the histogram.



(a) i. Briefly explain the commonly-used notation μ, $\hat{\mu}$, $\bar{x}$, σ, $\hat{\sigma}$ and s in the context of this example. [3 marks]

ii. Construct and interpret a 99% confidence interval. [7 marks]

iii. List any assumptions that are made in the construction of the confidence interval, and comment on how they could be assessed.

[6 marks]

(b) Let $U_1, U_2, ..., U_{40}$ be uniformly distributed random variables from a to b. The probability density function is:

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{otherwise} \end{cases}$$

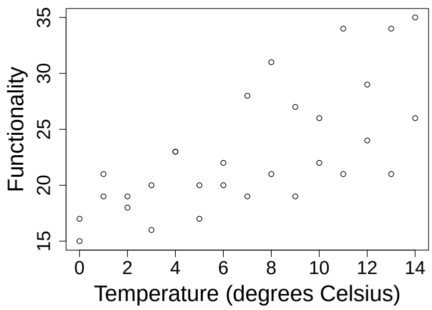
$\mathsf{E}[U_i] = \dfrac{1}{2}(b+a)$ and $\mathsf{Var}(U_i) = \dfrac{1}{12}(b-a)^2$. Let $T = \sum_{i=1}^{40} U_i$.

i. Show $\mathsf{E}[T] = 20(a+b)$ and $\mathsf{Var}(T) = \dfrac{10}{3}(b-a)^2$. [8 marks]

ii. Let $a = 5$ and $b = 15$. Use the central limit theorem (CLT) to find $P(410 \leq T \leq 420)$, approximately. Explain in your own words (2-3 sentences) why the CLT is appropriate to use here. [9 marks]

(33 marks)

2. The functionality of an electrical component of a machine was tested under a range of temperatures (temp = 0, 1, 2, ..., 14 °C) in an experiment. A scatter plot of the data is shown, there were 30 data points in total.



(a) A simple linear regression model was fitted to this dataset using R. Here is some of the output:

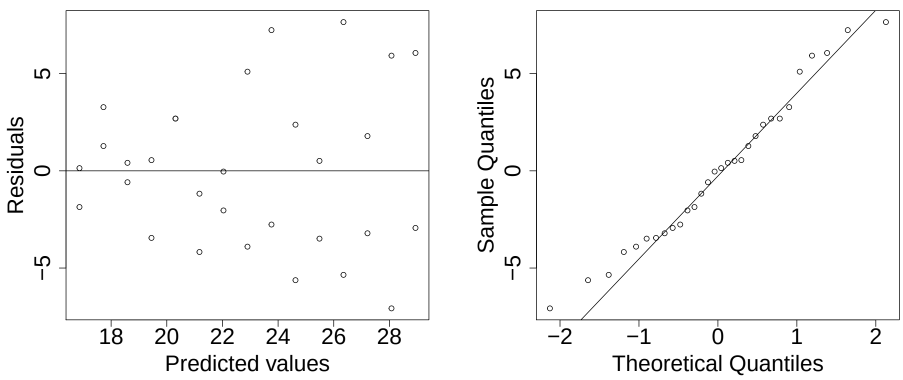
	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	16.8625	1.3906	12.126	1.16e-12
temp	0.8625	0.1691	5.102	2.10e-05

i. Write out the estimated equation of the line and interpret the parameter estimates. [7 marks]

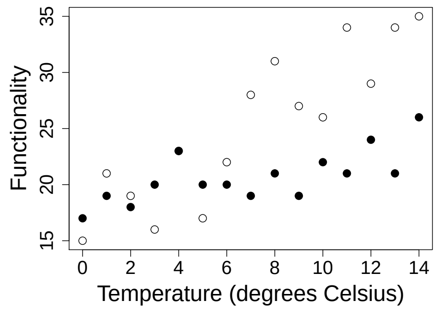
ii. Interpret the hypothesis test for the temp parameter. [7 marks]

iii. The minimum acceptable functionality of the component is 20. Test whether the average functionality for when temperature equals 0°C is lower than 20, using $\alpha = 0.05$. [7 marks]

iv. State the model assumptions and identify if the assumptions are reasonably met using the following residual plots. [6 marks]



(b) Some new information came to light during discussions between the statistician and the person who carried out the experiment: there were two types of components used in the experiment and each was tested once under each temperature 0, 1, 2,...,14 °C. The scatter plot shows the data with the points for component A shown in empty circles and for component B in filled circles.



The dataset was re-analysed by fitting this model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} + \epsilon_i$$

where y_i is the functionality value for the ith experimental unit, x_{i1} is equal to 1 if the ith experimental unit was a type A component and 0 for a type B component, and x_{i2} is the temperature for the ith experimental unit.

Write out this model in matrix notation, clearly showing the structure of each matrix. [7 marks]

(34 marks)

3. (a) Let $X_1, X_2, ..., X_n$ be an independent and identically distributed sample from a distribution with probability density function:

$$f(x) = \lambda^2 x e^{-\lambda x}$$

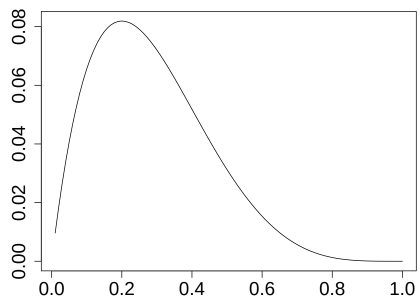
Derive the maximum likelihood estimator (MLE) of λ. [12 marks]

(b) Independent Bernoulli trials were performed until a success was observed. Let X be the number of trials until the first success.

i. What distribution does X follow? [2 marks]

ii. On the 4th trial, the first success was observed. The likelihood function was constructed for this data, with a graph of it shown below. Briefly explain (1-2 sentences) what is on the x and y axes. In your own words (1-2 sentences), explain how the graph can aid finding the MLE.

[8 marks]



(c) A sample of data of size 10 was collected at random from a population and the median was calculated. Sampling from the original data, with replacement, 1000 bootstrap samples were found and the median computed for each.

i. If you had the vector of 1000 bootstrap medians, describe how you would use it to construct a 90% confidence interval. [7 marks]

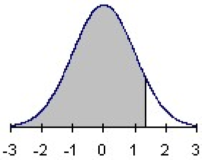
ii. Suppose a histogram was generated of the vector of bootstrap medians. Explain (2-3 sentences) what the histogram is approximating. [4 marks]

(33 marks)

Statistical Tables

Area under the standard Normal curve

The table gives the area left of z. For example, if z = 1.23, the area, shaded in the illustration below, is .8907.



z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

Critical values of the *t* distribution with degrees of freedom df

For example, for df = 3, $P(T_3 > 2.353) = 0.05$.

df \ P	0.3	0.25	0.20	0.15	0.1	0.05	0.025	0.01	0.005
3	0.584	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841
4	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604
5	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032
6	0.553	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707
7	0.549	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499
8	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355
9	0.543	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250
10	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169
11	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106
12	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055
13	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012
14	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977
15	0.536	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947
16	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898
18	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878
19	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861
20	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845
21	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831
22	0.532	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819
23	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807
24	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797
25	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787
26	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779
27	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771
28	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763
29	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756
30	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750
40	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704
50	0.528	0.679	0.849	1.047	1.299	1.676	2.009	2.403	2.678
60	0.527	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660
∞	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576

Critical values of the χ^2 distribution with degrees of freedom df

For example, for df = 1, $P(\chi_1^2 > 3.8414) = 0.05$.

P df	.995	.990	.975	.950	.900	.750	.500	.250	.100	.050	.025	.010	.005
1	0.00004	0.00016	0.00098	0.00393	0.01579	0.10153	0.45494	1.32330	2.70554	3.84146	5.02389	6.63490	7.87944
2	0.01003	0.02010	0.05064	0.10259	0.21072	0.57536	1.38629	2.77259	4.60517	5.99146	7.37776	9.21034	10.59663
3	0.07172	0.11483	0.21580	0.35185	0.58437	1.21253	2.36597	4.10834	6.25139	7.81473	9.34840	11.34487	12.83816
4	0.20699	0.29711	0.48442	0.71072	1.06362	1.92256	3.35669	5.38527	7.77944	9.48773	11.14329	13.27670	14.86026
5	0.41174	0.55430	0.83121	1.14548	1.61031	2.67460	4.35146	6.62568	9.23636	11.07050	12.83250	15.08627	16.74960
6	0.67573	0.87209	1.23734	1.63538	2.20413	3.45460	5.34812	7.84080	10.64464	12.59159	14.44938	16.81189	18.54758
7	0.98926	1.23904	1.68987	2.16735	2.83311	4.25485	6.34581	9.03715	12.01704	14.06714	16.01276	18.47531	20.27774
8	1.34441	1.64650	2.17973	2.73264	3.48954	5.07064	7.34412	10.21885	13.36157	15.50731	17.53455	20.09024	21.95495
9	1.73493	2.08790	2.70039	3.32511	4.16816	5.89883	8.34283	11.38875	14.68366	16.91898	19.02277	21.66599	23.58935
10	2.15586	2.55821	3.24697	3.94030	4.86518	6.73720	9.34182	12.54886	15.98718	18.30704	20.48318	23.20925	25.18818
11	2.60322	3.05348	3.81575	4.57481	5.57778	7.58414	10.34100	13.70069	17.27501	19.67514	21.92005	24.72497	26.75685
12	3.07382	3.57057	4.40379	5.22603	6.30380	8.43842	11.34032	14.84540	18.54935	21.02607	23.33666	26.21697	28.29952
13	3.56503	4.10692	5.00875	5.89186	7.04150	9.29907	12.33976	15.98391	19.81193	22.36203	24.73560	27.68825	29.81947
14	4.07467	4.66043	5.62873	6.57063	7.78953	10.16531	13.33927	17.11693	21.06414	23.68479	26.11895	29.14124	31.31935
15	4.60092	5.22935	6.26214	7.26094	8.54676	11.03654	14.33886	18.24509	22.30713	24.99579	27.48839	30.57791	32.80132
16	5.14221	5.81221	6.90766	7.96165	9.31224	11.91222	15.33850	19.36886	23.54183	26.29623	28.84535	31.99993	34.26719
17	5.69722	6.40776	7.56419	8.67176	10.08519	12.79193	16.33818	20.48868	24.76904	27.58711	30.19101	33.40866	35.71847
18	6.26480	7.01491	8.23075	9.39046	10.86494	13.67529	17.33790	21.60489	25.98942	28.86930	31.52638	34.80531	37.15645
19	6.84397	7.63273	8.90652	10.11701	11.65091	14.56200	18.33765	22.71781	27.20357	30.14353	32.85233	36.19087	38.58226
20	7.43384	8.26040	9.59078	10.85081	12.44261	15.45177	19.33743	23.82769	28.41198	31.41043	34.16961	37.56623	39.99685
21	8.03365	8.89720	10.28290	11.59131	13.23960	16.34438	20.33723	24.93478	29.61509	32.67057	35.47888	38.93217	41.40106
22	8.64272	9.54249	10.98232	12.33801	14.04149	17.23962	21.33704	26.03927	30.81328	33.92444	36.78071	40.28936	42.79565
23	9.26042	10.19572	11.68855	13.09051	14.84796	18.13730	22.33688	27.14134	32.00690	35.17246	38.07563	41.63840	44.18128
24	9.88623	10.85636	12.40115	13.84843	15.65868	19.03725	23.33673	28.24115	33.19624	36.41503	39.36408	42.97982	45.55851
25	10.51965	11.52398	13.11972	14.61141	16.47341	19.93934	24.33659	29.33885	34.38159	37.65248	40.64647	44.31410	46.92789
26	11.16024	12.19815	13.84390	15.37916	17.29188	20.84343	25.33646	30.43457	35.56317	38.88514	41.92317	45.64168	48.28988
27	11.80759	12.87850	14.57338	16.15140	18.11390	21.74940	26.33634	31.52841	36.74122	40.11327	43.19451	46.96294	49.64492
28	12.46134	13.56471	15.30786	16.92788	18.93924	22.65716	27.33623	32.62049	37.91592	41.33714	44.46079	48.27824	50.99338
29	13.12115	14.25645	16.04707	17.70837	19.76774	23.56659	28.33613	33.71091	39.08747	42.55697	45.72229	49.58788	52.33562
30	13.78672	14.95346	16.79077	18.49266	20.59923	24.47761	29.33603	34.79974	40.25602	43.77297	46.97924	50.89218	53.67196