ഭാരതീയ സാങ്കേതികവിദ്യാ സ്ഥാപനം പാലക്കാട്
भारतीय प्रौद्योगिकी संस्थान पालक्काड
Indian Institute of Technology Palakkad
Nurturing Minds For a Better World

IIT PALAKKAD

DEPARTMENT OF DATA SCIENCE

DS5610 BUSINESS ANALYTICS

TEST 1 (OPEN BOOK/NOTES)

20 February 2025 **Marks: 50 (Weight: 15 Marks)** **Duration: 60 Mins**

PLEASE READ THE FOLLOWING CAREFULLY

- **Answer all the questions.** There are no choices.
- **Answer all parts of the question together.** For example, if Q. No. 2 contains sub-questions such as 2a, 2b and 2c, all these questions should be answered together in sequence in one place. **Please do not write 2a on page 1 and 2c on the last page. SUCH ANSWERS WILL NOT BE EVALUATED.**
- Draw diagrams neatly and legibly.
- **THIS IS AN OPEN-BOOK EXAMINATION. Textbook, Reading materials, handwritten notes in bound form, and printed slide handouts are allowed. NO OTHER MATERIAL, LOOSE SHEETS are allowed.**
- **The use of a scientific calculator is permitted.** However, **the use of mobile phones, laptops and other electronic devices is NOT PERMITTED during the examination.**
- **No exchange of scientific calculators, books, printed slides, loose sheets, etc. are permitted.**
- **If you find any of the data missing for any of the questions, MAKE YOUR OWN ASSUMPTIONS AND HIGHLIGHT THE SAME BY DRAWING A BOX AROUND IT.**
- **WRITE THE ANSWER IN A BLUE PEN. Writing by Pencil or Red/Green Pen is not allowed.**

Q. No. 1. The operations manager of Sweetjet Airlines had the following data about the passengers who flew from Coimbatore to Colombo during the last 18 weeks.

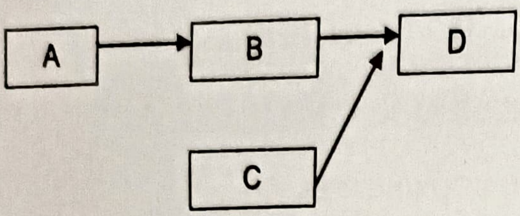
Week	1	2	3	4	5	6	7	8	9
Passengers flown	405	410	420	415	412	420	424	433	438
Week	10	11	12	13	14	15	16	17	18
Passengers flown	440	446	451	455	464	466	474	476	482

a. What graph would you use for visualising this data? Why?
b. Plot the graph and provide your inference.
c. What type of data is this? Why do you think so?
d. Does the data represent a sample or a population? Justify your answer.
e. What statistical measures can you compute to describe the data and make a few decisions?

(2 + 3 + 2 + 2 + 9 = 18 Marks)

P.T.O

Q. No. 2 The diagram shown below represents the operating system of the Hair Oil Division (HOD) of Kerala Cosmetics Limited. Department A (Extraction) receives the dry coconut, crushes it, and extracts the raw coconut oil. This is fed to Department B (Refining), which reduces the fat content and thereby processes it to a suitable thinness for making hair oil. Department C (Herbs) provides the necessary herbal extract to mix with the coconut oil. Department D (Mixing) combines the inputs from both B and C to produce the final aromatic herbal hair oil.



Department	Capacity (Litres Per Hour)
A	100
B	70
C	50
D	120

Instead of time taken by each department, the capacity (in Litres Per Hour or LPH) of each department is shown in the above table. It should be understood here that the ratio for mixing the outputs from departments B and C is two to one – i.e., to get three litres of aromatic herbal oil at department D, it requires mixing two litres of B's output (refined coconut oil) and one litre of C's output (herbal extract). Based on the description of the process, answer the following questions:

a. What is the type of production system followed by HOD? Give reasons.
b. What type of process configuration is exhibited by A, B and D? Give reasons.
c. Comment about the Department C with respect to the process configuration.
d. What is the system's capacity, and which department is the bottleneck? (provide necessary explanation based on the computation)
e. How much slack (unused capacity) is available in the other departments? (provide necessary explanation based on the computation)
f. How much system capacity can be gained by adding capacity to the bottleneck? (provide necessary explanation)

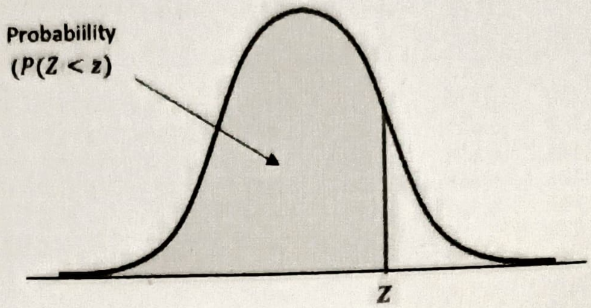
(2 + 2 + 2 + 3 + 3 + 3 = 15 Marks)

Q. No. 3. A pharmaceutical lab is testing a new reagent (Method B) to see if it improves the yield of a specific compound compared to the existing reagent (Method A). The lab runs about 6 trials for each reagent and records the following: Method A had a mean yield of 84.5% with a standard deviation of 2.1%, while Method B recorded a mean yield of 87.2% with a standard deviation of 1.8%. Can we conclude that there is a difference in the mean yield between the two reagents/methods at a significance level of 0.05? In this context, answer the following:

a. What is the parameter of interest for estimation in the above situation?
b. What type/class of hypothesis testing would you be conducting? Why?
c. Write the null and alternative hypotheses, both in the form of a notation as well as in a simple English sentence.
d. What are the underlying assumptions in the chosen test?
e. What type of test would this be, considering the hypothesis? Why?
f. Based on the results obtained, what is your decision?
g. Assuming that the actual p-value turned out to be 0.0649. What is your inference? In this case, what would be the relation between the computed test statistic and the tabulated statistic value?

(1 + 3 + 2 + 2 + 2 + 5 + 2 = 17 Marks)

Standard Normal Table



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	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	0.5754
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	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549
0.7	0.7580	0.7612	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	0.7996	0.8023	0.8051	0.8079	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	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9485	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	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9773	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	0.9865	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	0.9980	0.9980	0.9981
2.9	0.9981	0.9982	0.9983	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	0.9996	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	0.9998	0.9998

Obtained from:
https://d3rw207pwvlq3a.cloudfront.net/attachments/000/125/197/original/z-table_%282%29-crop-1596812289721-crop-1597154408572.png?1597154413

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
						Confidence Level					

Obtained from: https://www.tdistributiontable.com/wp-content/uploads/2020/08/t-table.png

Hypothesis Testing Cheat Sheet

debonomark@gmail.com 99271274 purematematics.mt

Type	Condition	Hypothesis	Statistic	Notes
Population Mean	σ is known n is large ($n>30$)	$H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$ / $\mu>\mu_0$ / $\mu<\mu_0$	$Z=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$	Z follows the the standard normal distribution.
	σ is known n is small ($n\leq 30$)	$H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$ / $\mu>\mu_0$ / $\mu<\mu_0$	$t=\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$	t follows the t distribution with $n-1$ d.o.f.
	σ is unknown n is large ($n>30$)	$H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$ / $\mu>\mu_0$ / $\mu<\mu_0$	$Z=\frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}}$	Z follows the the standard normal distribution.
	σ is unknown n is small ($n\leq 30$)	$H_0: \mu=\mu_0$ $H_1: \mu\neq\mu_0$ / $\mu>\mu_0$ / $\mu<\mu_0$	$t=\frac{\bar{X}-\mu}{\frac{s}{\sqrt{n}}}$	t follows the t distribution with $n-1$ d.o.f.
Population Proportion		$H_0: \pi=\pi_0$ $H_1: \pi\neq\pi_0$ / $\pi>\pi_0$ / $\pi<\pi_0$	$Z=\frac{p-\pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$	Z follows the the standard normal distribution.
Difference between Population Proportions		$H_0: \pi_1-\pi_2=0$ $H_1: \pi_1-\pi_2\neq 0$ / $\pi_1-\pi_2>0$ / $\pi_1-\pi_2<0$	$Z=\frac{p_1-p_2-(\pi_1-\pi_2)}{\sqrt{P(1-P)(\frac{1}{n_1}+\frac{1}{n_2})}}$	Z follows the the standard normal distribution. $P=\frac{x_1+x_2}{n_1+n_2}$ where x_1, x_2 is the number of subjects in group 1, 2 resp. that have the characteristic
Difference between Population Means	σ_1,σ_2 are known	$H_0: \mu_1-\mu_2=\mu_0$ $H_1: \mu_1-\mu_2\neq\mu_0$ / $\mu_1-\mu_2>\mu_0$ / $\mu_1-\mu_2<\mu_0$	$Z=\frac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$	Z follows the the standard normal distribution.
	σ_1,σ_2 are unknown σ_1,σ_2 are unequal n_1,n_2 are large	$H_0: \mu_1-\mu_2=\mu_0$ $H_1: \mu_1-\mu_2\neq\mu_0$ / $\mu_1-\mu_2>\mu_0$ / $\mu_1-\mu_2<\mu_0$	$Z=\frac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}}$	Z follows the the standard normal distribution.
	σ_1,σ_2 are unknown σ_1,σ_2 are equal	$H_0: \mu_1-\mu_2=\mu_0$ $H_1: \mu_1-\mu_2\neq\mu_0$ / $\mu_1-\mu_2>\mu_0$ / $\mu_1-\mu_2<\mu_0$	$t=\frac{\bar{x}_1-\bar{x}_2-(\mu_1-\mu_2)}{\sqrt{s_p^2(\frac{1}{n_1}+\frac{1}{n_2})}}$	t follows the t distribution with n_1+n_2-2 d.o.f.
Population Variance		$H_0: \sigma^2=\sigma_0^2$ $H_1: \sigma^2\neq\sigma_0^2$	$\chi^2=\frac{(n-1)s^2}{\sigma^2}$	χ^2 follows the χ^2 distribution with $n-1$ d.o.f.
Difference between Population Variances		$H_0: \sigma_1^2=\sigma_2^2$ $H_1: \sigma_1^2\neq\sigma_2^2$	$F=\frac{s_1^2}{s_2^2}$	F follows the F distribution with n_1-1 and n_2-1 d.o.f.

Obtained from: https://purematematics.mt/wp-content/uploads/2022/05/Hypothesis-Testing-Cheat-Sheet.jpg

GOOD LUCK