NAMIBIA UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Faculty of Health, Natural Resources and Applied Sciences

School of Natural and Applied Sciences

Department of Mathematics, Statistics and Actuarial Science

13 Jackson Kaujeua Street
Private Bag 13388
Windhoek
NAMIBIA

T: +264 61 207 2913
E: msas@nust.na
W: www.nust.na

QUALIFICATION: Bachelor of Science Honours in Applied Statistics	
QUALIFICATION CODE: 08BSHS	**LEVEL:** 8
COURSE CODE: MVA802S	**COURSE NAME:** MULTIVARIATE ANALYSIS
SESSION: JANUARY 2024	**PAPER:** THEORY
DURATION: 3 HOURS	**MARKS:** 100

SUPPLEMENTARY / SECOND OPPORTUNITY EXAMINATION QUESTION PAPER	
EXAMINER	Dr D. B. GEMECHU
MODERATOR:	Prof L. PAZVAKAWAMBWA

INSTRUCTIONS
1. There are 6 questions, answer ALL the questions by showing all the necessary steps.
2. Write clearly and neatly.
3. Number the answers clearly.
4. Round your answers to at least four decimal places, if applicable.

PERMISSIBLE MATERIALS

1. Nonprogrammable scientific calculators with no cover.

THIS QUESTION PAPER CONSISTS OF 5 PAGES (Including this front page)

ATTACHMENTS

Two statistical distribution tables (z-and F-distribution tables)

Question 1 [15 Marks]

1.1. Briefly explain Principal components analysis (PCA) and state three assumptions of PCA [5]

1.2. State three reason why multivariate approach to hypothesis testing instead of univariate approach in inference about multivariate mean vectors. [3]

1.3. Briefly discuss a One-Sample Profile Analysis. Your answer should include (Definition of profile analysis, assumptions of the variable, hypothesis to be tested, the contrast matrix, the test statistics and the rejection rule). [7]

Question 2 [10 Marks]

2. The following data represent measurements of blood glucose levels on three occasions (y_1, y_2 and y_3) for 4 women patients who gave consent to participate on the study. The results obtained are listed below:

Individual	y_1	y_2	y_3
1	60	69	62
2	56	53	84
3	62	75	68
4	73	70	64

Then compute

2.1. the sample mean vector $\overline{\boldsymbol{y}}$. [3]

2.2. the sample variance-covariance matrix, $\boldsymbol{S}$. [5]

2.3. the total sample variance. [2]

Question 3 [23 Marks]

3.1. Given that $\boldsymbol{y} \sim N_p(\boldsymbol{\mu}_y, \boldsymbol{\Sigma}_y)$ a random variable z is defined as a linear combination of $\boldsymbol{y} = (y_1, y_2, \ldots, y_p)'$ as $z_i = a_1 y_{i1} + a_2 y_{i2} + \cdot \cdot \cdot + a_p y_{ip}$, for $i = 1, 2, \ldots, n$, then show that $\bar{z} = \boldsymbol{a}'\overline{\boldsymbol{y}}$, where $\boldsymbol{a}' = (a_1\ a_2 \cdots a_p)$ and $\overline{\boldsymbol{y}}$ is the sample mean vector of the p-variables. [5]

3.2. Suppose that test 1 (x_1) and test 2 (x_2) scores of MVA students that follow a bivariate normal distribution with parameters mean $\mu_1 = 70$ and $\mu_2 = 60$, the standard deviations $\sigma_1 = 10$ and $\sigma_2 = 15$, and $\rho = 0.6$.

3.2.1. Express a given information in the form of matrix notation, thus what would be $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$? [2]

3.2.2. If a student is selected randomly, then find the probability that

3.2.2.1. the score of a randomly selected student is above 75 on test 2? [4]

3.2.2.2. the score of a randomly selected student is above 75 on test 2 given that the student scored 80 on Test 1. [6]

3.2.2.3. the sum of the score of a randomly selected student on both tests is above 150. [3]

3.2.2.4. the students performance in test 1 is better than test 2. [3]

Question 4 [20 Marks]

4. A medical researcher is interested in two particular fatty acids (A and B) found in human blood. Measurements (micrograms per gram) were taken on 16 new-born babies with Down's syndrome. The sample means were 70 and 50 for fatty acids A and B, respectively and the corresponding sample covariance matrix was

$$S = \begin{pmatrix} 100 & 80 \\ 80 & 100 \end{pmatrix}.$$

For non-down's syndrome new-born babies the expected fatty acid levels are 80 and 65 for A and B, respectively. Use the multivariate hypothesis test technique to assess whether the observed data for the Down's syndrome babies are consistent with the expected values for non-Down's syndrome babies. Your solution should include the following:

4.1. State the null and alternative hypothesis to be tested **[1]**

4.2. State the test statistics to be used and its corresponding distribution **[2]**

4.3. State the decision (rejection) rule and compute the tabulated value using an appropriate statistical table **[2]**

4.4. Compute the test statistics and write up your decision and conclusion **[6]**

4.5. Construct a 95% T^2 interval for $\mu_1 - \mu_2$ **[5]**

4.6. Assuming that the purpose is to make only two confidence statements (i.e. $m = 2$), construct a 95% Bonferroni confidence interval for μ_1 **[4]**

Question 5 [20 Marks]

5. Observations on two responses are collected for three treatments. The observation vectors $\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$ are

Treatment 1: $\begin{bmatrix} 6 \\ 7 \end{bmatrix}$, $\begin{bmatrix} 5 \\ 3 \end{bmatrix}$

Treatment 2: $\begin{bmatrix} 3 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 6 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$

Treatment 3: $\begin{bmatrix} 2 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 5 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 5 \\ 8 \end{bmatrix}$

5.1. Construct the one-way MANOVA table **[14]**

5.2. Evaluate Wilks' lambda, Λ_{wilks} **[3]**

5.3. Test for vector of treatment effe cts at 5% level of significance. Your answer should include specification of the null and alternative hypothesis. **[3]**

Hint: Use the test statistics: $\left(\frac{N-g-1}{g-1}\right)\left(\frac{1-\sqrt{\Lambda_{wilks}}}{\sqrt{\Lambda_{wilks}}}\right) \sim F_{2(g-1),2(N-g-1)}$

Question 6 [12 Marks]

6. A marketing researcher conducted a study to evaluate the effect of advertisement on two different types of products. For this research, a random sample of eight ex-customers and eight customers were shown advertisements on both products and asked to rate them on the two variables (Ability to gain attention (y1) and Persuasiveness (purchase y2)). The software output of the analysis of

the data is given in Tables 1-3. Your answer to each question below should include **the hypothesis to be tested, test statics** and $p - value$ and **conclusion**.

6.1. Draw conclusion of the Box test for equality of covariance matrix using the 5% significance level. **[3]**

6.2. Is there any significance interaction effects of customer status and type of product on the ratting? **[3]**

6.3. Test if there is a significant of effects of customer status on the rating of the two variables. **[3]**

6.4. Test if there is a significant of effects of product type on the rating of the two variables. **[3]**

Table 1: Box's Test of Equality of Covariance Matrices[a]

Box's M	3.285
F	.253
df1	9
df2	1650.212
Sig.	.986
Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.	
a. Design: Intercept + customersatus + producttype + customersatus * producttype	

Table 2: Multivariate Tests[a]

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.990	565.442[b]	2.000	11.000	<.001	.990
	Wilks' Lambda	.010	565.442[b]	2.000	11.000	<.001	.990
	Hotelling's Trace	102.808	565.442[b]	2.000	11.000	<.001	.990
	Roy's Largest Root	102.808	565.442[b]	2.000	11.000	<.001	.990
customersatus	Pillai's Trace	.925	68.073[b]	2.000	11.000	<.001	.925
	Wilks' Lambda	.075	68.073[b]	2.000	11.000	<.001	.925
	Hotelling's Trace	12.377	68.073[b]	2.000	11.000	<.001	.925
	Roy's Largest Root	12.377	68.073[b]	2.000	11.000	<.001	.925
typeprodct	Pillai's Trace	.746	16.119[b]	2.000	11.000	<.001	.746
	Wilks' Lambda	.254	16.119[b]	2.000	11.000	<.001	.746
	Hotelling's Trace	2.931	16.119[b]	2.000	11.000	<.001	.746
	Roy's Largest Root	2.931	16.119[b]	2.000	11.000	<.001	.746
customersatus * producttype	Pillai's Trace	.274	2.073[b]	2.000	11.000	.172	.274
	Wilks' Lambda	.726	2.073[b]	2.000	11.000	.172	.274
	Hotelling's Trace	.377	2.073[b]	2.000	11.000	.172	.274
	Roy's Largest Root	.377	2.073[b]	2.000	11.000	.172	.274
a. Design: Intercept + customersatus + producttype + customersatus * producttype							
b. Exact statistic							

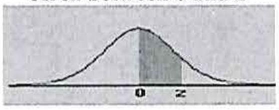
Table 3: Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	Gain attention (y1)	52.000[a]	3	17.333	26.000	<.001	.867
	Purchase (y2)	60.688[b]	3	20.229	29.424	<.001	.880
Intercept	Gain attention (y1)	324.000	1	324.000	486.000	<.001	.976
	Purchase (y2)	410.063	1	410.063	596.455	<.001	.980
customersatus	Gain attention (y1)	36.000	1	36.000	54.000	<.001	.818
	Purchase (y2)	52.563	1	52.563	76.455	<.001	.864
producttype	Gain attention (y1)	16.000	1	16.000	24.000	<.001	.667
	Purchase (y2)	5.063	1	5.063	7.364	.019	.380
customersatus * producttype	Gain attention (y1)	.000	1	.000	.000	1.000	.000
	Purchase (y2)	3.063	1	3.063	4.455	.056	.271
Error	Gain attention (y1)	8.000	12	.667			
	Purchase (y2)	8.250	12	.688			
Total	Gain attention (y1)	384.000	16				
	Purchase (y2)	479.000	16				
Corrected Total	Gain attention (y1)	60.000	15				
	Purchase (y2)	68.938	15				

a. R Squared = .867 (Adjusted R Squared = .833)

b. R Squared = .880 (Adjusted R Squared = .850)

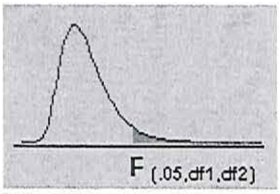
Area between 0 and z



	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990

Table for α=.05

df1



df2

df2\df1	1	2	3	4	5	6	7	8	9	10	12
1	161.448	199.500	215.707	224.583	230.162	233.986	236.768	238.883	240.543	241.882	243.906
2	18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371	19.384	19.396	19.413
3	10.128	9.552	9.277	9.117	9.014	8.941	8.887	8.845	8.812	8.786	8.745
4	7.709	6.944	6.591	6.388	6.256	6.163	6.0942	6.041	5.998	5.964	5.912
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818	4.772	4.735	4.678
6	5.987	5.143	4.757	4.533	4.387	4.284	4.207	4.147	4.099	4.060	3.999
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726	3.676	3.637	3.575
8	5.318	4.459	4.066	3.838	3.688	3.581	3.501	3.438	3.388	3.347	3.284
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.229	3.178	3.137	3.073
10	4.965	4.103	3.708	3.478	3.326	3.217	3.136	3.072	3.020	2.978	2.913
11	4.844	3.982	3.587	3.358	3.204	3.095	3.012	2.948	2.896	2.854	2.788
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849	2.796	2.753	2.687
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767	2.714	2.671	2.604
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699	2.645	2.602	2.534
15	4.543	3.682	3.287	3.056	2.901	2.791	2.707	2.641	2.587	2.544	2.475
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591	2.537	2.494	2.425
17	4.451	3.591	3.197	2.965	2.810	2.699	2.614	2.548	2.494	2.450	2.381
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510	2.456	2.412	2.342
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477	2.423	2.378	2.308
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.441	2.393	2.348	2.278