



**NAMIBIA UNIVERSITY  
OF SCIENCE AND TECHNOLOGY  
FACULTY OF HEALTH AND APPLIED SCIENCES**

**DEPARTMENT OF MATHEMATICS AND STATISTICS**

<b>QUALIFICATION:</b> Bachelor of Science Honours in Applied Statistics	
<b>QUALIFICATION CODE:</b> O8BSSH	<b>LEVEL:</b> 8
<b>COURSE CODE:</b> MVA802S	<b>COURSE NAME:</b> MULTIVARIATE ANALYSIS
<b>SESSION:</b> JANUARY 2019	<b>PAPER:</b> THEORY
<b>DURATION:</b> 3 HOURS	<b>MARKS:</b> 100

<b>SECOND OPPORTUNITY EXAMINATION QUESTION PAPER</b>	
<b>EXAMINER</b>	Dr D. B. GEMECHU
<b>MODERATOR:</b>	PROF. S. SUSUMAN

<b>INSTRUCTIONS</b>
<ol style="list-style-type: none"><li>1. Answer ALL the questions in the booklet provided.</li><li>2. Show clearly all the steps used in the calculations.</li><li>3. All written work must be done in blue or black ink and sketches must be done in pencil.</li></ol>

**PERMISSIBLE MATERIALS**

1. Non-programmable calculator without a cover.

**ATTACHMENTS**

1. Statistical tables ( $Z$ ,  $\chi^2$  and  $F$ ).

**THIS QUESTION PAPER CONSISTS OF 9 PAGES** (Including this front page)

**Question 1****[8 marks]**

- 1.1. Define multivariate analysis [2]
- 1.2. State two advantages of multivariate approach to hypothesis testing as compared to univariate test. [2]
- 1.3. Briefly explain Principal Components Analysis (PCA) and state two common criteria used in determining the number of principal components to retain. [2+2]

**Question 2****[15 marks]**

2. The following sample data represents the leaf size ( $y_1$ ), colour of flower ( $y_2$ ), and height of plant ( $y_3$ ) of four plants.

$$Y = \begin{pmatrix} 6.1 & 2 & 12 \\ 8.2 & 1 & 8 \\ 5.3 & 0 & 9 \\ 6.4 & 2 & 10 \end{pmatrix}$$

Then calculate:

- 2.1. the sample mean vector  $\bar{y}$ . [3]
- 2.2. the sample variance-covariance matrix  $S$ . [6]
- 2.3. the correlation matrix,  $R$  and interpret your result. [6]

**Question 3****[13 marks]**

3. If  $y \sim N_p(\mu, \Sigma)$  and  $z = (\Sigma^{1/2})^{-1}(y - \mu)$ , then show that  $z \sim N_p(0, I)$ .  
Hint: use the uniqueness properties of moment generating function. [13]

**Question 4****[26 marks]**

4. Let  $x \sim N_3(\mu, \Sigma)$ , where  $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ ,  $\mu = \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$  and  $\Sigma = \begin{pmatrix} 1 & \rho & \rho^2 \\ \rho & 1 & 0 \\ \rho^2 & 0 & 1 \end{pmatrix}$ .

- 4.1. Drive the conditional distribution of  $(x_1, x_2)$  given  $x_3$ . [8]
- 4.2. If we define a new random variable  $y = x_1 + \frac{1}{3}x_2 - \frac{1}{2}x_3$  and the value of  $\rho = 0.5$ , then:
  - 4.2.1. derive the distribution of  $y$  and compute  $P(y < -2)$  [9]
  - 4.2.2. derive the joint distribution of  $x_3$  and  $y$ . Are they independently distributed? Provide explanation for your answer. [9]

**Question 5****[10 marks]**

5. The height and weight for a sample of 20 male college students were measured and the sample mean vector where obtained to be  $\bar{y} = (71.45 \ 164.7)'$ . Assume that the sample were originated form a bivariate normal  $N_2(\mu, \Sigma)$  with  $\Sigma = \begin{pmatrix} 20 & 100 \\ 100 & 1000 \end{pmatrix}$ . Using these result, test the hypothesis  $H_0: \mu = (70, 170)'$  vs  $H_1: \mu \neq (70, 170)'$  at 5% level of significance.

Your solution should include the following:

- 5.1. State the test statistics to be used and its corresponding distribution [2]
- 5.2. State the decision (rejection) rule and compute the tabulated value using an appropriate statistical table [2]
- 5.3. Compute the test statistics and write up your decision and conclusion [6]

**Question 6****[15 marks]**

6. An experiment was conducted to determine whether protein and fiber content for wheat grown with fertilizer 1 is different from that for wheat grown with fertilizer 2. Wheat was grown in 22 plots. On 11 of these plots, fertilizer A was used; on the other 11 plots, fertilizer B was used. The protein and fiber content (in percent) of the wheat from each plot was measured. Assume that the observations are bivariate and follow multivariate normal distributions for  $N(\mu_i, \Sigma)$ ,  $i = 1$  and  $2$ . The sample mean vectors and sample covariance matrices from these measurements are:

$$\bar{y}_1 = (12.1 \ 14.3)', \quad \bar{y}_2 = (10.1 \ 13.3)'$$

$$S_1 = \begin{pmatrix} 2.2 & -1.1 \\ -1.1 & 0.9 \end{pmatrix}, \quad S_2 = \begin{pmatrix} 2.3 & -1.0 \\ -1.0 & 1.1 \end{pmatrix}$$

- 6.1. Find the pooled estimate of the covariance matrix for this data. **[3]**
- 6.2. Test the null hypothesis that the mean protein and fiber content is the same for both fertilizers at 5% level of significance. **[12]**

**Question 7****[13 marks]**

7. The study was conducted to investigate the differences in worker productivity measured by income level (rincom) and hours (hrs) worked for individuals of different age category (agecat). For this study, 685 individuals were randomly selected and the summary statistics of the data are given in the attached SPSS output (Tables 1-4 given below).
- 7.1. Draw conclusion of the Box test for equality of covariance matrix using the 5% significance level. Your conclusion should include the F and  $p$  - value. **[3]**
- 7.2. Are there significant mean differences in worker productivity (as measure by the combinations of income and hours worked) for individuals of different ages? Your answer should include the hypothesis to be tested, test statistics and  $p$  - value and conclusion. **[4]**
- 7.3. Are there significant difference in income levels for individuals of different ages? If so, which age categories differ? **[4]**
- 7.4. Are there significant mean different in hours worked for individuals of different ages? If so, which age categories differ? **[2]**

**Table 1: Box's Test of Equality of Covariance Matrices<sup>a</sup>**

Box's M	6.936
F	.766
df1	9
df2	2886560.794
Sig.	.648

Tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups.

a. Design: Intercept + agecat



Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.957	7507.272 <sup>b</sup>	2.000	680.000	.000	.957
	Wilks' Lambda	.043	7507.272 <sup>b</sup>	2.000	680.000	.000	.957
	Hotelling's Trace	22.080	7507.272 <sup>b</sup>	2.000	680.000	.000	.957
	Roy's Largest Root	22.080	7507.272 <sup>b</sup>	2.000	680.000	.000	.957
agecat4	Pillai's Trace	.091	10.791	6.000	1362.000	.000	.045
	Wilks' Lambda	.909	11.035 <sup>b</sup>	6.000	1360.000	.000	.046
	Hotelling's Trace	.100	11.279	6.000	1358.000	.000	.047
	Roy's Largest Root	.099	22.457 <sup>c</sup>	3.000	681.000	.000	.090

a. Design: Intercept + agecat

b. Exact statistic

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	rincom	1029.016 <sup>a</sup>	3	343.005	20.995	.000	.085
	hrs	64.281 <sup>b</sup>	3	21.427	.167	.919	.001
Intercept	rincom	128493.515	1	128493.515	7864.965	.000	.920
	hrs	1410953.564	1	1410953.564	10972.708	.000	.942
agecat4	rincom	1029.016	3	343.005	20.995	.000	.085
	hrs	64.281	3	21.427	.167	.919	.001
Error	rincom	11125.807	681	16.337			
	hrs	87568.119	681	128.588			
Total	rincom	149966.000	685				
	hrs	1575151.000	685				
Corrected Total	rincom	12154.823	684				
	hrs	87632.400	684				

a. R Squared = .085 (Adjusted R Squared = .081)

b. R Squared = .001 (Adjusted R Squared = -.004)

**Table 4: Pairwise Comparisons**

Dependent Variable	(I) 4 categories of age	(J) 4 categories of age	Mean Difference (I-J)	Std. Error	Sig. <sup>b</sup>	95% Confidence Interval for Difference <sup>b</sup>	
						Lower Bound	Upper Bound
rincom	18-29	30-39	-2.164*	.449	.000	-3.351	-.977
		40-49	-3.458*	.460	.000	-4.675	-2.240
		50+	-3.090*	.493	.000	-4.396	-1.785
	30-39	18-29	2.164*	.449	.000	.977	3.351
		40-49	-1.293*	.397	.007	-2.344	-.242
		50+	-.926	.435	.203	-2.078	.226
	40-49	18-29	3.458*	.460	.000	2.240	4.675
		30-39	1.293*	.397	.007	.242	2.344
		50+	.367	.447	1.000	-.816	1.551
	50+	18-29	3.090*	.493	.000	1.785	4.396
		30-39	.926	.435	.203	-.226	2.078
		40-49	-.367	.447	1.000	-1.551	.816
hrs	18-29	30-39	-.711	1.258	1.000	-4.041	2.619
		40-49	-.169	1.291	1.000	-3.586	3.247
		50+	-.006	1.384	1.000	-3.669	3.657
	30-39	18-29	.711	1.258	1.000	-2.619	4.041
		40-49	.542	1.114	1.000	-2.407	3.491
		50+	.705	1.221	1.000	-2.526	3.936
	40-49	18-29	.169	1.291	1.000	-3.247	3.586
		30-39	-.542	1.114	1.000	-3.491	2.407
		50+	.163	1.255	1.000	-3.157	3.484
	50+	18-29	.006	1.384	1.000	-3.657	3.669
		30-39	-.705	1.221	1.000	-3.936	2.526
		40-49	-.163	1.255	1.000	-3.484	3.157

Based on estimated marginal means

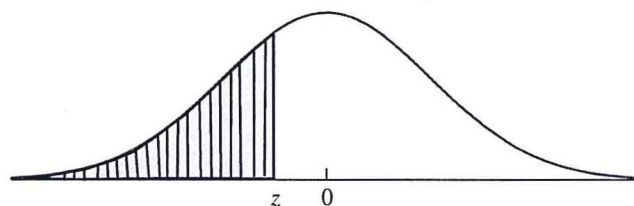
\*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

====END=====



Cumulative normal distribution (z table)

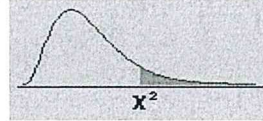


<b>z</b>	<b>0.00</b>	<b>0.01</b>	<b>0.02</b>	<b>0.03</b>	<b>0.04</b>	<b>0.05</b>	<b>0.06</b>	<b>0.07</b>	<b>0.08</b>	<b>0.09</b>
-3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641





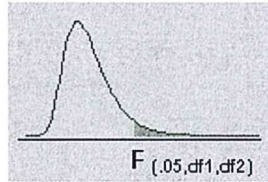
Right tail areas for the Chi-square Distribution



df\area	.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



Table for  $\alpha=.05$



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