Spring 2025

Multivariate Analysis
Assignments

- **Assignment 1 due by March 4, 2025**

  1. (20 pt.) Let $\mathbf{X}_1, \ldots, \mathbf{X}_{36}$ be a random sample of size 36 from a three-variate normal distribution having mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$. Specify each of the following completely.
     (a) (10 pt.) The distribution of $\bar{\mathbf{X}}$.
     (b) (10 pt.) The distribution of $n(\bar{\mathbf{X}} - \boldsymbol{\mu})'\boldsymbol{\Sigma}^{-1}(\bar{\mathbf{X}} - \boldsymbol{\mu})$.
  2. (80 pt.) Check whether the following data satisfy the normality assumption.

Company	X_1 =Sales	X_2 = Profits	X_3 = Assets
General Motors	126,974	4,224	173,297
Ford	96,933	3,835	160,893
Exxon	86,656	3,510	83,219
IBM	63,438	3,758	77,734
General Electric	55,264	3,939	128,344
Mobil	50,976	1,809	39,080
Philip Morris	39,069	2,946	38,528
Chrysler	36,156	359	51,038
Du Pont	35,209	2,480	34,715
Texaco	32,416	2,413	25,636

  § Extra 10 points for creating a plot of a bivariate normal distribution with $\mu_1 = \mu_2 = 2, \sigma_1 = \sigma_2 = 1$ and $\rho = 0.5$ using SAS or R.

- **Assignment 2 due by March 11, 2025**

  1. (40 pt.) The scores obtained by 87 college students on the exam are given in the following for X_1= history, X_2= science and X_3=mathematics. Test $H_0 : \boldsymbol{\mu}' = [500, 50, 30]$ versus $H_1 : \boldsymbol{\mu}' \neq [500, 50, 30]$. Here $[500, 50, 30]$ represent average scores for thousands of college students over the last 10 years.
     (a) (12 pt.) Find the value of T^2.
     (b) (12 pt.) Specify the distribution of T^2 in part (a).
     (c) (16 pt.) Make your conclusion for the test.

Individual	X_1	X_2	X_3	Individual	X_1	X_2	X_3
1	468	41	26	45	494	41	24
2	428	39	26	46	541	47	25
3	514	53	21	47	362	36	17
4	547	67	33	48	408	28	17
5	614	61	27	49	594	68	23
6	501	67	29	50	501	25	26
7	421	46	22	51	687	75	33
8	527	50	23	52	633	52	31
9	527	55	19	53	647	67	29
10	620	72	32	54	647	65	34
11	587	63	31	55	614	59	25
12	541	59	19	56	633	65	28
13	561	53	26	57	448	55	24
14	468	62	20	58	408	51	19
15	614	65	28	59	441	35	22
16	527	48	21	60	435	60	20
17	507	32	27	61	501	54	21
18	580	64	21	62	507	42	24
19	507	59	21	63	620	71	36
20	521	54	23	64	415	52	20
21	574	52	25	65	554	69	30
22	587	64	31	66	348	28	18
23	488	51	27	67	468	49	25
24	468	62	18	68	507	54	26
25	587	56	26	69	527	47	31
26	421	38	16	70	527	47	26
27	481	52	26	71	435	50	28
28	428	40	19	72	660	70	25
29	640	65	25	73	733	73	33
30	574	61	28	74	507	45	28
31	547	64	27	75	527	62	29
32	580	64	28	76	428	37	19
33	494	53	26	77	481	48	23
34	554	51	21	78	507	61	19
35	647	58	23	79	527	66	23
36	507	65	23	80	488	41	28
37	454	52	28	81	607	69	28
38	427	57	21	82	561	59	34
39	521	66	26	83	614	70	23
40	468	57	14	84	527	49	30
41	587	55	30	85	474	41	16
42	507	61	31	86	441	47	26
43	574	54	31	87	607	67	32
44	507	53	23				

2. (60 pt.) In a study of the cost of transporting milk from farms to dairy plants, a survey was taken of firms engaged in milk transportation. Cost data on Y_1 =fuel, Y_2 =repair, and Y_3 =capital, all measured on a per-mile basis, are presented as follows. Assume that the normality assumption is satisfied and the covariance matrices are equal. Conduct the one-way MANOVA and make your conclusion at $\alpha = 0.01$.

Gasoline Trucks			Diesel Trucks		
Y_1	Y_2	Y_3	Y_1	Y_2	Y_3
16.44	12.43	11.23	8.50	12.26	9.11
7.19	2.70	3.92	7.42	5.13	17.15
9.92	1.35	9.75	10.28	3.32	11.23
4.24	5.78	7.78	10.16	14.72	5.99
11.20	5.05	10.67	12.79	4.17	29.28
14.25	5.78	9.88	9.60	12.72	11.00
13.50	10.98	10.60	6.47	8.89	19.00
13.32	14.27	9.45	11.35	9.95	14.53
12.68	7.61	10.23	9.70	5.06	20.84
7.51	5.80	8.13	9.77	17.86	35.18
9.90	3.63	9.13	11.61	11.75	17.00
10.25	5.07	10.17	9.09	13.25	20.66
11.11	6.15	7.61	8.53	10.14	17.45
12.17	14.26	14.39	8.29	6.22	16.38
10.24	2.59	6.09	15.90	12.90	19.09
10.18	6.05	12.14	11.94	5.69	14.77
8.88	2.70	12.23	9.54	16.77	22.66
12.34	7.73	11.68	10.43	17.65	10.66
8.51	14.02	12.01	10.87	21.52	28.47
12.95	8.24	7.18	11.88	12.18	21.20
16.93	13.37	17.59	12.03	9.22	23.09
14.70	10.78	14.58			
10.32	5.16	17.00			
8.98	4.49	4.26			

- **Assignment 3 due by March 18, 2025**

1 (30 pt.) Using R or SAS, find the determinant, trace, inverse, eigenvalues and eigenvectors of the matrix A, where

$$A = \begin{bmatrix} 13 & -4 & 2 \\ -4 & 13 & -2 \\ 2 & -2 & 10 \end{bmatrix}.$$

2. (70 pt.) In an effort to develop improved peanuts, crop scientists routinely compare varieties with respect to several variables. A two-factor experiment with two replications was considered. Three varieties were grown at two geographical locations, and the three variables representing yield, sound mature kernels (SMK, weight in grams) and seed size (weight in grams) were measured. The data are shown as follows. Test the effects at $\alpha = 0.05$ and make your conclusion.

Factor 1 Location	Factor 2 Variety	Y_1 Yield	Y_2 SMK	Y_3 Size
1	5	195.3	153.1	51.4
1	5	194.3	167.7	53.7
2	5	189.7	139.5	55.5
2	5	180.4	121.1	44.4
1	6	203.0	156.8	49.8
1	6	195.9	166.0	45.8
2	6	202.7	166.1	60.4
2	6	197.6	161.8	54.1
1	8	193.5	164.5	57.8
1	8	187.0	165.1	58.6
2	8	201.5	166.8	65.0
2	8	200.0	173.8	67.2

- **Assignment 4 due by March 25, 2025**

1. (100 pt.) A study consists of 42 measurements on air-pollution variables recoded at 12:00 noon in the Los Angeles area on different days. The data are shown in the following. Conduct principal component analysis using the covariance and correlation matrices, respectively.

   (a) (20 pt.) Obtain the eigenvalues and eigenvectors for the covariance and the correlation matrices.

   (b) (25 pt.) How many principal components do you choose? Explain.

   (c) (25 pt.) Find the correlations of the original variables with the principle components, and the cumulative percentage of total sample variance explained by the principal component(s) in (b).

   (d) (30 pt.) Interpret the components.

Wind (X_1)	Solar Radiation (X_2)	CO (X_3)	NO (X_4)	NO_2 (X_5)	O_3 (X_6)	HC (X_7)
8	98	7	2	12	8	2
7	107	4	3	9	5	3
7	103	4	3	5	6	3
10	88	5	2	8	15	4
6	91	4	2	8	10	3
8	90	5	2	12	12	4
9	84	7	4	12	15	5
5	72	6	4	21	14	4
7	82	5	1	11	11	3
8	64	5	2	13	9	4
6	71	5	4	10	3	3
6	91	4	2	12	7	3
7	72	7	4	18	10	3
10	70	4	2	11	7	3
10	72	4	1	8	10	3
9	77	4	1	9	10	3
8	76	4	1	7	7	3
8	71	5	3	16	4	4
9	67	4	2	13	2	3
9	69	3	3	9	5	3
10	62	5	3	14	4	4
9	88	4	2	7	6	3
8	80	4	2	13	11	4
5	30	3	3	5	2	3
6	83	5	1	10	23	4
8	84	3	2	7	6	3
6	78	4	2	11	11	3
8	79	2	1	7	10	3
6	62	4	3	9	8	3
10	37	3	1	7	2	3
8	71	4	1	10	7	3
7	52	4	1	12	8	4
5	48	6	5	8	4	3
6	75	4	1	10	24	3
10	35	4	1	6	9	2
8	85	4	1	9	10	2
5	86	3	1	6	12	2
5	86	7	2	13	18	2
7	79	7	4	9	25	3
7	79	5	2	8	6	2
6	68	6	2	11	14	3
8	40	4	3	6	5	2

- **Assignment 5 due by May 6, 2025**

1. (100 pt.) Conduct cluster analysis for the data on national men's track records given below.

Country	100m (s)	200m (s)	400m (s)	800m (min)	1500m (min)	5000m (min)	10000m (min)	Marathon (min)
Argentina	10.23	20.37	46.18	1.77	3.68	13.33	27.65	129.57
Australia	9.93	20.06	44.38	1.74	3.53	12.93	27.53	127.51
Austria	10.15	20.45	45.80	1.77	3.58	13.26	27.72	132.22
Belgium	10.14	20.19	45.02	1.73	3.57	12.83	26.87	127.20
Bermuda	10.27	20.30	45.26	1.79	3.70	14.64	30.49	146.37
Brazil	10.00	19.89	44.29	1.70	3.57	13.48	28.13	126.05
Canada	9.84	20.17	44.72	1.75	3.53	13.23	27.60	130.09
Chile	10.10	20.15	45.92	1.76	3.65	13.39	28.09	132.19
China	10.17	20.42	45.25	1.77	3.61	13.42	28.17	129.18
Columbia	10.29	20.85	45.84	1.80	3.72	13.49	27.88	131.17
Cook Islands	10.97	22.46	51.40	1.94	4.24	16.70	35.38	171.26
Costa Rica	10.32	20.96	46.42	1.87	3.84	13.75	28.81	133.23
Czech Republic	10.24	20.61	45.77	1.75	3.58	13.42	27.80	131.57
Denmark	10.29	20.52	45.89	1.69	3.52	13.42	27.91	129.43
Dominican Republic	10.16	20.65	44.90	1.81	3.73	14.31	30.43	146.00
Finland	10.21	20.47	45.49	1.74	3.61	13.27	27.52	131.15
France	10.02	20.16	44.64	1.72	3.48	12.98	27.38	126.36
Germany	10.06	20.23	44.33	1.73	3.53	12.91	27.36	128.47
Great Britain	9.87	19.94	44.36	1.70	3.49	13.01	27.30	127.13
Greece	10.11	19.85	45.57	1.75	3.61	13.48	28.12	132.04
Guatemala	10.32	21.09	48.44	1.82	3.74	13.98	29.34	132.53
Hungary	10.08	20.11	45.43	1.76	3.59	13.45	28.03	132.10
India	10.33	20.73	45.48	1.76	3.63	13.50	28.81	132.00
Indonesia	10.20	20.93	46.37	1.83	3.77	14.21	29.65	139.18
Ireland	10.35	20.54	45.58	1.75	3.56	13.07	27.78	129.15
Israel	10.20	20.89	46.59	1.80	3.70	13.66	28.72	134.21
Italy	10.01	19.72	45.26	1.73	3.35	13.09	27.28	127.29
Japan	10.00	20.03	44.78	1.77	3.62	13.22	27.58	126.16
Kenya	10.28	20.43	44.18	1.70	3.44	12.66	26.46	124.55
Korea, South	10.34	20.41	45.37	1.74	3.64	13.84	28.51	127.20
Korea, North	10.60	21.23	46.95	1.82	3.77	13.90	28.45	129.26
Luxembourg	10.41	20.77	47.90	1.76	3.67	13.64	28.77	134.03
Malaysia	10.30	20.92	46.41	1.79	3.76	14.11	29.50	149.27
Mauritius	10.13	20.06	44.69	1.80	3.83	14.15	29.84	143.07
Mexico	10.21	20.40	44.31	1.78	3.63	13.13	27.14	127.19
Myanmar(Burma)	10.64	21.52	48.63	1.80	3.80	14.19	29.62	139.57
Netherlands	10.19	20.19	45.68	1.73	3.55	13.22	27.44	128.31
New Zealand	10.11	20.42	46.09	1.74	3.54	13.21	27.70	128.59
Norway	10.08	20.17	46.11	1.71	3.62	13.11	27.54	130.17
Papua New Guinea	10.40	21.18	46.77	1.80	4.00	14.72	31.36	148.13
Philippines	10.57	21.43	45.57	1.80	3.82	13.97	29.04	138.44
Poland	10.00	19.98	44.62	1.72	3.59	13.29	27.89	129.23
Portugal	9.86	20.12	46.11	1.75	3.50	13.05	27.21	126.36
Romania	10.21	20.75	45.77	1.76	3.57	13.25	27.67	132.30
Russia	10.11	20.23	44.60	1.71	3.54	13.20	27.90	129.16
Samoa	10.78	21.86	49.98	1.94	4.01	16.28	34.71	161.50
Singapore	10.37	21.14	47.60	1.84	3.86	14.96	31.32	144.22
Spain	10.17	20.59	44.96	1.73	3.48	13.04	27.24	127.23
Sweden	10.18	20.43	45.54	1.76	3.61	13.29	27.93	130.38
Switzerland	10.16	20.41	44.99	1.71	3.53	13.13	27.90	129.56
Taiwan	10.36	20.81	46.72	1.79	3.77	13.91	29.20	134.35
Thailand	10.23	20.69	46.05	1.81	3.77	14.25	29.67	139.33
Turkey	10.38	21.04	46.63	1.78	3.59	13.45	28.33	130.25
USA	9.78	19.32	43.18	1.71	3.46	12.97	27.23	125.38