STAT 5572
Exam I
Due: Monday, Oct.9th by 11:59 pm

(1) Consider a bivariate normal population with $\mu_1 = 0$, $\mu_2 = 2$, $\sigma_{11} = 2$, $\sigma_{22} = 1$, and $\rho_{12} = 0.5$.

(a) (5 pts) Write out the bivariate normal density.
(b) (5 pts) Write out the squared statistical distance expression $(x-\mu)'\Sigma^{-1}(x-\mu)$ as a function of $x_1$and x_2.
(c) (10 pts) Determine (and sketch) the constant-density contour that contains 50% of the probability.

(2) Let X be distributed $N_4(\mu\,,\Sigma)$ such that,

$$\mu = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 0 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 2 & 0 & 1 & 0 \\ 0 & 3 & 0 & 2 \\ 1 & 0 & 5 & 0 \\ 0 & 2 & 0 & 3 \end{bmatrix}$$

(a) (5 pts) Find the marginal distribution of $\chi_1 = [X_1, X_3]$
(b) (10 pts) Find the conditional distribution of $(X_1, X_2 | X_3 = x_3, X_4 = x_4)$.

(3) Suppose X is $N_3(\mu\,,\Sigma)$ with

$$\mu = \begin{bmatrix} -4 \\ 2 \\ 5 \end{bmatrix} \quad \Sigma = \begin{pmatrix} 8 & 0 & -1 \\ 0 & 3 & 0 \\ -1 & 0 & 5 \end{pmatrix}$$

Which of the following are independent? Justify.

(a) (5 pts) X_1 and X_2
(b) (5 pts) X_1 and X_3
(c) (5 pts) (X_1, X_2) and X_3

(4) The datafile '*National_Track_records.dat*' contains the national track records for women in 54 countries.

Let $X_1 = 100m\ (s)$
$X_2 = 200m\ (s)$
$X_3 = 400m\ (s)$
$X_4 = 800m\ (min)$
$X_5 = 1500m\ (min)$

$$X_6 = 3000m\ (min)$$
$$X_7 = Marathon\ (min)$$

Define the following linear combinations,

$$V_1 = \frac{1}{3}X_1 + \frac{1}{6}X_2 + \frac{1}{12}X_3$$

$$V_2 = \frac{15}{8}X_4 + X_5 + \frac{1}{2}X_6 + \frac{1}{28.13}X_7$$

Where V_1 is the average of the short distance times scaled to seconds per 100 meters and V_2 is the average of the long-distance times scaled to seconds per 100 meters.

(a) (8 pts) Calculate the observed values of V_1 and V_2.
(b) (12 pts) Calculate sample means, sample variances and sample covariances of V_1 and V_2.

(5) Consider the '*sweat.dat'* data file. For each of 20 healthy females, three numerical variables that measure aspects of perspiration: $X_1 =$ Sweat (Sweat rate), $X_2 =$ Sodium (Sodium content), and $X_3 =$ Potassium (Potassium content) are included.

(a) (10 pts) Construct univariate QQ-plots for each of the three variables. Also, make the three pairwise scatterplots. Does the multivariate normal assumption seem reasonable?
(b) (15 pts) Test the null hypothesis $\mathrm{H}_0: \mu' = [4.0, 45.0, 10.0]$ at $\alpha = 0.05$ using the Hotelling's T^2 test.
What is the test statistic, critical value, and the p-value? What is your conclusion regarding H_0 ?
(c) (5 pts) Determine the 95% confidence ellipsoid for μ. Where is it centered? What are the corresponding half-lengths of its axes?