Instructions:
• Write the answers and solutions in the space provided
• Explain your answers. Full credit is only given for a correct answer AND proper justification
• Clearly formulate all theoretical results used in your solution
• Due date: Tuesday, April 24, 5PM

Good Luck!
Problem 1 (Bivariate Normal distribution) [20 points]

Consider a bivariate normal population \( \mathbf{x} = [x_1, x_2]' \) with mean \( \mu = [\mu_1, \mu_2]' \) and variance

\[
\Sigma = \begin{bmatrix}
\sigma_{11} & \sigma_{12} \\
\sigma_{21} & \sigma_{22}
\end{bmatrix}.
\]

Assume \( \mu_1 = 5, \mu_2 = 0, \sigma_{11} = 1, \sigma_{22} = 4, \) and Corr\( (x_1, x_2) = \rho_{12} = -0.5 \).

(a) [10 pts] Write out the squared statistical distance expression \( (\mathbf{x} - \mu)'\Sigma^{-1}(\mathbf{x} - \mu) \) as a function of \( x_1 \) and \( x_2 \). Simplify.

(b) [5 pts] Find \( \Sigma \)
(c) [5 pts] Sketch a constant-density contour (grading is done based on proper orientation and axes ordering, you do not have to report eigenvalues/eigenvectors).
Problem 2 (Multivariate Normal Distribution) [20 points]

Let $\mathbf{X}$ be $N_3(0, \Sigma)$ with

$$\Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 1 \\ 0 & 1 & 1 \end{bmatrix}.$$

(a) [10 pts] Find the distribution of $X_1 + X_2 + X_3$.

(b) [5 pts] For which values of $a$ the variables $a(X_2 + X_3)$ and $2aX_1$ are independent?

(c) [5 pts] For which values of $a, b$ the variables $X_2 + aX_3$ and $X_2 + bX_3$ are independent?
Problem 3 (Multivariate Normal Distribution) [20 points]

Let $X_1, X_2, \ldots, X_{100}$ be a random sample from a three-dimensional Normal distribution with mean $\mu$ and covariance $\Sigma$. Let $S$ be an unbiased estimator for $\Sigma$. Specify the following

(a) [10 pts] The distribution of $(\bar{X} - \mu)$

(b) [5 pts] The distribution of $\Sigma^{-1/2} \bar{X}$

(c) [5 pts] The approximate distribution of $(X_3 - \mu)^T S^{-1} (X_3 - \mu)$
Problem 4 (Principal Component Analysis) [20 points]

(a) [10 pts] Determine the population principal components $Y_1$, $Y_2$, and $Y_3$ for the correlation matrix $\rho$ that corresponds to the covariance

$$
\Sigma = \begin{bmatrix}
9 & 0 & 1 \\
0 & 4 & 0 \\
1 & 0 & 1
\end{bmatrix}
$$

(b) [10 pts] Find the proportion of variance explained by each principal component.
Problem 5 (Factor Analysis) [20 points] The correlation matrix

\[ \rho = \begin{bmatrix}
1 & 0.05 & 0.09 \\
0.05 & 1 & 0.45 \\
0.09 & 0.45 & 1
\end{bmatrix} \]

for a three dimensional variable \( Z = [Z_1, Z_2, Z_3]' \) is approximated by a factor model with \( m = 1 \):

\[ Z_i = l_i F_1 + \epsilon_i, \quad i = 1, 2, 3 \]

and

\[ L = [l_1, l_2, l_3]' = [0.1, 0.5, 0.9]' . \]

(a) [5 pts] Find the communalities \( h_i^2, i = 1, 2, 3 \).

(b) [10 pts] Find the specific variances \( \psi_i, i = 1, 2, 3 \).
(c) [5 pts] Find Corr(Z_i, F_1), i = 1, 2, 3. Which variable might carry the greatest weight in naming the common factor? Why?