

Applied Multivariate Data Analysis CHE 494/694

Assignment 1

Submission Deadline: October 5th, 5pm

1)

Let X be $N_3(\mu, \Sigma)$ with $\mu' = [2, -3, 4]$ and $\Sigma = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$.

- Find the distribution of AX where $A = \begin{bmatrix} 1 & -2 & 3 \\ 1 & 3 & -1 \end{bmatrix}$
- Find the 2×1 vector a such that x_2 and $x_2 - a' \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ are independent
- Find distribution of $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ given $x_3 = x_3^*$
- Expression for the confidence ellipse in X with 90% confidence

2)

Consider the 2nd and 4th and 6th columns of the *food consumption data set* as a trivariate data set $[X_1 X_2 X_3]$ and compute the following (Use MATLAB wherever necessary)

- Mean, co-variance, correlation matrix and the 2D, 3D scatter plot. Comment on the correlations (+ve or -ve or 0)
- Expression for the confidence ellipse with 90% confidence. The major and minor axis full lengths.
- Can $\mu = [8; 17; 31]$ be the true population mean with 80% confidence? Justify your answer with hypothesis testing.
- Individual confidence intervals for μ_1, μ_2 and μ_3 with a confidence of 95%.
- Simultaneous T^2 and Bonferroni confidence intervals of μ_1, μ_2 and μ_3 with a confidence of 95%. Comment on the results in comparison with part (d).