MATH 4630.03F

Instructor: Yuehua Wu (Amy)  Phone: 736-2100 ext 33497
Office: N518 ROSS  Office hours: Th 1:00-2:00pm

Course times and places: Tuesdays and Thursdays 10:00 - 11:30am in 117 FC

by Johnson, Richard Arnold & Wichern, Dean W.

References:
1. Multivariate statistical methods, 3rd edition by Morrison, Donald F.

Grading:
Assignments 20% (due on Sept. 27, Oct. 20, Nov. 10, Nov. 29)
Mid-term exam: 30% (Oct. 25)
Final exam: 50% (to be announced)

Tentative course outline:

I. Review and introduction
Matrix algebra, data matrix, summary statistics, linear combinations, graphical representations, some multivariate problems and techniques, SAS IML PROC.

II. Basic properties of random vectors
Random vector, cumulative distribution, probability density function, probability function, marginal cumulative distribution function, independence, mean vector, covariance matrix, correlation matrix, generalized variance, conditional mean vector, conditional variance, regression, characteristic function.

III. Multivariate normal distribution theory
Multivariate normal distribution, conditional and marginal distributions of multinormal variates, central limit theorem, principal axes of the multinormal density, linear forms, Wishart distribution, Cochran theorem, partitioned Wishart matrices, sampling distribution, Hotelling $T^2$ distribution, Mahalanobis distance.

IV. Estimation
Likelihood function, maximum likelihood estimation, multinormal cases with $\Sigma > 0$ (1. Unconstrained; 2. Constraints on the mean vector; 3. Constraints on $\Sigma$; 4. Samples with linked parameters).

V. Tests of hypotheses on means
Likelihood ratio test, union intersection test, test of the hypothesis $H_0 : \mu = \mu_0$, $\Sigma$ known; Hotelling one-sample $T^2$ test, test of the hypothesis $H_0 : \mu_1 = \mu_2$, the analysis of repeated measurements, the paired $T^2$ test, profile analysis, simultaneous confidence intervals, testing the normality assumption, transformation.
VI. Multivariate analysis of variance and multivariate linear regression models

Wilks’ Lambda distribution, one-way classification, two-way classification, extension to higher
designs, multivariate linear regression models, estimability, testability.

VII. Tests of hypotheses on covariance

Test of the hypothesis $H_0 : \Sigma = \Sigma_0$, test of the hypothesis $H_0 : \Sigma_1 = \ldots = \Sigma_k$, test of the hypothesis $H_0 : \Sigma_{12} = 0$, test of the hypothesis $H_0 : \Sigma_{ij} = 0$ for all $i \neq j$

VIII. Canonical correlation analysis

Canonical correlations, canonical correlation variables, population canonical correlation analysis,
sample canonical correlation analysis.

IX. Principal component analysis

Principal components, geometrical meaning of principal components, interpretation of principal
components, sampling properties of principal components.

X. Factor analysis

Mathematical model for factor structure, principal factor analysis, maximum likelihood factor
analysis, testing the goodness of fit of the factor model, rotation of factors, factor scores, relationships between factor analysis and principal component analysis.

XI. Discrimination and classification

Classification for two populations, evaluating classification functions, Fisher’s discrimination
function, classification with several populations.

XII. Cluster analysis

Similarity measures, hierarchical methods, nonhierarchial methods.

Note:

1. If you miss the mid-term exam, no make-up exam will be given. In case that you have a
valid medical (or other) reason for being absent for the exam, the associated fraction of
your final mark will then be transferred to the final exam.

2. The last day to add a Fall Term (F) course with the permission of the instructor is Oct. 7,
2005.

3. The last day to drop a Fall Term (F) course without receiving a grade is Nov. 11, 2005.