Class Notes for Chapter 9

- "Multiple" = several X's; "Multivariate" = several Y's
- Two important distributions here. The first is the multivariate Normal distribution for k variables (stacked into the vector y) given in Equation 9.2. The Bivariate case in 9.1 is for k = 2. The second important distribution is Wishart distribution which generalizes the chi-square distribution to matrices.
- Multivariate tests of a single mean, paired means, and two independent-sample means involve Hotelling's T^2 statistic.
- Example 9.1 concerns testing whether a single mean vector is equal to a constant vector. Details on the 4 commonly used test statistics are given on p.5; the plots confirm Normality. Note that the multivariate test gives p = 0.0649, whereas the three univariate tests give p-values 0.1080, 0.1619, and 0.9354.
- Example 9.2 is a bivariate paired t-test, with one portion of each sample going to lab 1 and the other to lab 2. The first response variable is BOD and the second is SS. The correct p-value is 0.0208; this differs from the NS univariate p-values.
- Example 9.3 concerns MANOVA with factors gender (M, F) and drug (A, B, C) and with two response variables. A log-transformation is first made to render the responses Normal and with constant variance. In Output 9.3a, we start with the gender*drug interaction. Since it is NS, we then look at gender and drug – only drug is significant. The multivariate MCP is then carried out in Outputs 9.3b and 9.3c, and the results are summarized at the top of p.9.
- Example 9.4 is a multivariate COD. Again, we have separate error terms. The factors gender, seq and gender*seq are tested using the *sub(gender*seq)* error term – Output 9.4a. The factors period, dose and co (and partitions of co) are tested

using the *usual error term* – see Outputs 9.4b and 9.4c. There appears to be no carryover effect here – so the seven-day washout period seems ample – but the dose (500, 1000, 2000, and 3000 mg) means are not equal. We partition into the linear, quadratic and cubic orthogonal pieces in Output 9.4e, and conclude that the quadratic and cubic 'terms' (vectors) are zero, but not the linear terms (vectors). This can be seen in the graphs on p.14, but we may have evidence to fit a nonlinear curve here.

- **The model for Multivariate Linear Regression is in Equation 9.5, where the k y-vectors and error-vectors are stacked side-by-side into the Y and U matrices, and we wish to estimate all the parameters in B as well as the variance components in Σ. "Point" estimates of these matrices are given on p.13, and tests regarding B involve F tests.**
- **Example 9.5 has two Y's and one x, and can be fit in SAS' REG procedure. The multivariate test that the 2x1 slope vector is zero is given in Output 9.5b, and the results match our hand calculations for estimates of B and Σ.**
- **Multivariate Nonlinear Regression is more complicated. ML parameter estimates are obtained by minimizing the determinant of $V(\theta) = U^T U$ given in Equation 9.8**
- **Example 9.6 concerns 10 rhino's in S. Africa, x is age, z_1 is the length in mm of the anterior horn and z_2 is the length in mm of the posterior horn. Clarke (1992) proposed the 5-parameter model in Equation 9.9, and since one parameter is shared, it is impossible to fit these curves separately! The log-transformed data are graphed on p.18, and the MLE's minimize the expression in Equation 9.10. These values are given, and used to test for common slopes (LR test in Equation 9.11).**
- **Next: two bivariate binomial models – stay tuned.**