

Lab Assignment No. 8: Canonical Correlation

PSGA7816 Multivariate Statistical Analysis

In completing your assignment you are expected to be specific and analytical in your responses. Simply producing tables or output is **not sufficient** and if not accompanied by any explanation or discussion it will not be given **any** credit.

Please append the full syntax you used to compute your answers. Your typed assignment answers should include the products of your computations (estimates, tables and/or charts) pasted into the body of your paper with appropriate remarks and comments (preferably in APA style).

No e-mail submissions will be accepted. Late assignments will be penalized 2 points per day. The dataset for the assignment is marsh.xls, consisting of 97 samples of Florida marshes and their chemical compositions found in either water or soil: Mercury in water (merc_water), methyl mercury in water (methyl_water), turbidity of water (turbidity_water), total organic carbon in water (carbon_water), mercury in soil (merc_soil), total sulfates found in soil (sulf_soil) and total phosphorus in soil (phos_soil).. Homework 8 is due **April 30, 2009 at 1:30 PM**.

- 1) Review your data and report the appropriate univariate statistics. Report the variable correlations. Test whether the correlation matrices for each group have a significantly different structure from a corresponding identity matrix. What can you infer about this dataset? What are you not able to infer from this correlation matrix that a canonical correlation would provide? **(1pt)**
- 2) Look at the eigenvalues produced from a principal components analysis (use proc princomp). What information can you deduce from it? If you were interested in accounting for at least 90% of the data how many components would you retain? **(1pt)**
- 3) Test the overall null hypothesis that canonical correlations are equal to zero. Give your test statistic, d.f., and p-value. **(1pt)**
- 4) Test the null hypothesis that the first, second and third canonical correlations are equal to zero. Give your test statistic, d.f., and p-value. **(1pt)**
- 5) What are the correlations between each variable and their corresponding canonical variables? Discuss. **(1pt)**
- 6) What are the correlations between each variable and the canonical variables from the opposite group? Discuss. **(1pt)**
- 7) Reproduce in full at least one canonical correlation. **(3pt)**
- 8) What remarks about the canonical correlations can you make? What inferences can you draw based on your analysis? **(1pt)**