**ASSIGNMENT AMA 4414: MULTIVRIATE ANALYSIS**

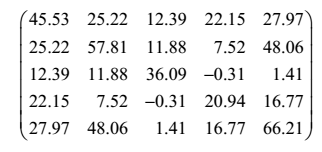
A group of archaeologists are studying remains found at a number of sites. They have measurements on a set of 32 skulls, 17 of which were found at one archaeological site and the other 15 at another site. They believe that each of these two sites was inhabited by a different tribe of people. They are now working on other sites in the same region and wish to decide which tribe the skulls they are finding belong to.

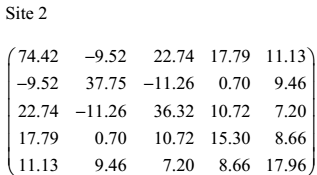
The measurements comprised 5 different dimensions of the skulls, all in mm.

* 1. Explain how linear *discriminant analysis* might be useful in studying these data, and describe the way in which the results may be applied to data from the new sites.
  2. State the assumptions that would need to be made about the data in order for linear discriminant analysis to be valid. Describe the checks that you could do to investigate these assumptions, stating any limitations on the methods.
  3. Summary statistics for each of the two groups of skulls are shown below. Describe the main features in relation to your answers to (i) and (ii).

Means for the 5 variables

|  |  |  |
| --- | --- | --- |
| *Variable* | *Site 1 (n = 17)* | *Site 2 (n = 15)* |
| *x*1 | 174.82 | 185.73 |
| *x*2 | 139.35 | 138.73 |
| *x*3 | 132.00 | 134.77 |
| *x*4 | 69.82 | 76.47 |
| *x*5 | 130.55 | 137.50 |

Variance-covariance matrices: Site 1



* 1. Below is shown a summary of two discriminant analyses on these data. Describe the differences between the two sets of results. Which model do you prefer and why?

