

MATH 204 - ASSIGNMENT 1

Please Hand in Assignment in the Lecture on Friday 23rd January.

In an observational follow-up study of kidney transplant patients, the level of gene expression (the amount of mRNA measured in a processed tissue sample) for each of three genes (TCL1A, IRF5, DAPK1) was measured. The patients fell into five categories, labelled AB, C, D, E, F:

- AB Health Maintained under Drug Therapy 1
- C Chronic Rejectors
- D Healthy Controls
- E Drug-Free Tolerant Patients
- F Health Maintained under Drug Therapy 2

Summary data for the study are recorded below in terms of group-specific sample means, variances and sums of squares statistics for each gene.

	n_i	TCL1A				IRF5				DAPK1			
		$\bar{x}_i$	s_i^2	SST	SSE	$\bar{x}_i$	s_i^2	SST	SSE	$\bar{x}_i$	s_i^2	SST	SSE
AB	52	-0.413	0.669			0.141	0.999			0.909	0.225		
C	25	0.043	1.660			0.358	1.294			1.259	0.470		
D	20	1.161	0.659			-0.119	1.626			1.094	0.170		
E	31	1.335	1.007			0.061	1.320			0.860	0.191		
F	14	-0.901	1.044			0.031	1.763			1.224	0.361		
				95.446	130.225			2.794	175.393			3.554	36.427

The data may be downloaded in SPSS format from

<http://www.math.mcgill.ca/~dstephens/204/Data/Kidneys.sav>

- (a) Using the ANOVA F-test, assess whether, for any of the three genes, there is any evidence that the level of gene expression is different across the five patient groups. Report the ANOVA table for each gene.

15 Marks

- (b) Report the pooled estimate of variance, s_p^2 , for each gene, and comment (without formal assessment) on the validity of the assumptions necessary for the ANOVA F-test to be valid.

5 Marks

You may use SPSS, and submit extracts from the output generated, provided that you write comments pointing out the key results.