

Assignment Sheet No. 5

I apologize for posting this so late. I have included reading for the rest of the semester.

Midterm Examination

I have rescheduled the same for Wednesday, October 26. It will be open book, open notes. Any needed tables other than the normal, χ^2 , Student's t and F distributions in the text or Bonferroni probability points from the handout will be provided. It will cover everything up to and including single sample profile analysis, one-way MANOVA and MANOVA canonical variables.

Reading (tentative)

Weeks of October 17-28: J&W, Sec. 6.8, 7.7, 7.8, 7.10; review Sec. 7.3 - 7.6 on univariate multiple regression

Week of October 31 - November 4: J&W, Chapter 8

Week of November 7 - 11: J&W Sec 9.1 - 9.6

Week of November 14 - 18: J&W, Chapter 10

Week of November 21 - 23: J&W, Sec. 11.1 - 11.4

Week of November 28 - December 2; remainder of Chapter 11, Sec. 12.1-12.3

Week of December 5 - 9 J&W, Chapter 12

Written Assignment (due in class Monday, November 8)

1. Do J&W Ex. 6.24 (JWData5.txt data set T06_13). Use the shortest simultaneous confidence limits that make sense. Also test in several ways the hypothesis that all periods have the same mean vector. Specifically, use Bonferroni F-tests, Roy's maximum relative eigenvalue test, Hotelling's generalized T_0^2 and Pillai's V test. Use $\alpha = .05$. What do the coefficients of the canonical variables tell you about how skull measurements differed between periods?

2. Do J&W Ex. 6.25 (JWData5.txt data set T11_07). Use the shortest simultaneous confidence limits that make sense.

3. In a study of the effects of whole-body radiation, 24 rats were subjected to radiation, 12 at level 500 r. and 12 at level 600 r. In each treatment group there were 3 rats in each of the following 4 classifications: young male, adult male, young female and adult female.

Thus the experiment was a 2^3 factorial experiment with factors A = radiation level, B = age, and C = sex. Weight changes y_1, y_2, y_3 , and y_4 were recorded 1, 3, 6, and 7 days after the irradiation. In addition, the initial weight u was recorded as a covariate but *should be ignored in this assignment*. On the next page is a listing of file radfact.txt which contains a MacAnova array made up of five 5 by 5 hypothesis or error matrices, $H^A, H^B, H^C, H^{(4)}$, and E . There is a link to radfact.txt on <http://www.stat.umn.edu/~kb/classes/5401/datafiles.html>.

H^A, H^B , and H^C are *main effect* hypothesis matrices.

STAT 5401/8401 Assignment Sheet No. 5

Matrix $H^{(4)} = H^{AB} + H^{AC} + H^{BC} + H^{ABC}$ is the sum of hypothesis matrices that might be used to test the indicated interactions but which are not available separately, can be used to test the hypothesis that there are no two-way or three-way *interactions* among the 3 factors. The degrees of freedom associated with $H^{(4)}$ is the sum of the degrees of freedom associated with its component matrices.

Matrix E is the error matrix from the complete factorial model.

```
Cmd> sscp <- read("", "sscp") # read from radfact.txt
sscp  5 5 5 format labels
) Morrison data p. 225
) Five 5 by 5 MANOVA hypothesis and error matrices are
) combined into a 3-dimensional array sscp.
)
) You extract matrix 4, say, by h4 <- matrix(sscp[4,,]) To
) extract the part of matrix 4 having to do only with the
) responses and not U, use h4 <- matrix(sscp[4,-1,-1])
)
) There are 4 hypothesis matrices for a 2x2x2 completely
) randomized factorial experiment with 1 covariate U and 4
) responses Y1 - Y4 in that order.
)
) Var. 1: U = initial wt of rats
) Var. 2: Y1 = weight loss 1 day after irradiation
) Var. 3: Y2 = weight loss 3 day after irradiation
) Var. 4: Y3 = weight loss 6 day after irradiation
) Var. 5: Y4 = weight loss 7 day after irradiation
)
) Factor A: Radiation level -- 500 r. or 600 r.
) Factor B: Age -- young or adult
) Factor C: Sex -- female or male
)
) There were 3 rats for each combination of factors
)
) The hypothesis matrices in the order they appear are
)   HA = sscp[1,,]
)   HB = sscp[2,,]
)   HC = sscp[3,,]
)   H4 = sscp[4,,] = HAB + HAC + HBC + HABC (4 df)
) The error matrix is
)   E = sscp[5,,]
)
Read from file "TP1:Stat5401:Data:radfact.txt"
```

STAT 5401/8401 Assignment Sheet No. 5

```
Cmd> print(sscp,format:"9.2f") # array has labels
```

```
sscp:
```

		InitialWt	Day 1	Day 3	Day 6	Day 7
HA	InitialWt	70.04	10.25	59.79	-52.96	-71.75
	Day 1	10.25	1.50	8.75	-7.75	-10.50
	Day 3	59.79	8.75	51.04	-45.21	-61.25
	Day 6	-52.96	-7.75	-45.21	40.04	54.25
	Day 7	-71.75	-10.50	-61.25	54.25	73.50
HB	InitialWt	11926.04	847.08	66.88	1493.54	1605.00
	Day 1	847.08	60.17	4.75	106.08	114.00
	Day 3	66.88	4.75	0.38	8.38	9.00
	Day 6	1493.54	106.08	8.38	187.04	201.00
	Day 7	1605.00	114.00	9.00	201.00	216.00
HC	InitialWt	51987.04	1210.08	4421.46	6469.29	3816.42
	Day 1	1210.08	28.17	102.92	150.58	88.83
	Day 3	4421.46	102.92	376.04	550.21	324.58
	Day 6	6469.29	150.58	550.21	805.04	474.92
	Day 7	3816.42	88.83	324.58	474.92	280.17
H4	InitialWt	3669.50	-204.66	124.84	32.34	172.75
	Day 1	-204.66	86.01	-52.33	-54.33	-61.16
	Day 3	124.84	-52.33	161.16	98.16	71.76
	Day 6	32.34	-54.33	98.16	128.16	89.76
	Day 7	172.75	-61.16	71.76	89.76	73.51
E	InitialWt	1326.00	36.00	284.65	66.67	-90.67
	Day 1	36.00	95.98	59.33	53.34	49.33
	Day 3	284.65	59.33	411.34	117.67	57.66
	Day 6	66.67	53.34	117.67	342.68	276.32
	Day 7	-90.67	49.33	57.66	276.32	385.32

Let H_{yy}^A , H_{yy}^B , H_{yy}^C , $H_{yy}^{(4)}$, and E_{yy} be the lower right hand 4 by 4 submatrices of these matrices, that is omitting the row and column associated with u .

The parts of the matrices having to do with y can be extracted as follows. In every sub array (matrix) we omit the first row and column (initial weight).

```
Cmd> ha <- matrix(sscp[1,-1,-1])
```

```
Cmd> hb <- matrix(sscp[2,-1,-1])
```

```
Cmd> hc <- matrix(sscp[3,-1,-1])
```

```
Cmd> h4 <- matrix(sscp[4,-1,-1])
```

```
Cmd> e <- matrix(sscp[5,-1,-1])
```

The use of `matrix()` without a dimension, turns its argument into a pure matrix with only two dimensions. Without it, `matrices[i,-1,-1]` would be 1 by 4 by 4 arrays. These would be correctly handled by functions like `solve()` or `releigen()` but the extra dimension might be confusing.

Use H_{yy}^A , H_{yy}^B , H_{yy}^C , $H_{yy}^{(4)}$, and E_{yy} in the following:

(a) Test for the presence of interaction by comparing $H_{yy}^{(4)}$ with E_{yy} using both the Wilks test statistic and the largest root test.

(b) Test for age and sex main effects using the Wilks statistic and in both cases obtain the *coefficients* used to compute the most important canonical variate.