Handout #14

Cell Tower Repairs: Comparing Problem Types - R

Problem 6.41 "When cell phone relay towers are not working properly, wireless providers can lose great amounts of money so it is important to be able to fix problems expeditiously." Based on the severity and complexity of the problem and the experience of the engineer assigned to it, a cell tower problem was classified into one of 8 groups. "[For each problem,] two times were observed: the time to assess the problem and plan an attack and the time to implement the solution." Data was collected on two problems within each classification group.

Is there evidence for differences between groups in mean time to assess the problem and implement the solution? Test using an appropriate MANOVA.

Here we have a one-way MANOVA with $p=2$ variables, $g=8$ groups, and $n_1=n_2=\cdots=n_8=2$, for a total sample size of $n=16$. Letting $\underline{\mu}_\ell=\begin{pmatrix}\mu_{\ell 1}\\ \mu_{\ell 2}\end{pmatrix}$ denote the mean vector for the ℓ^{th} group for $\ell=1,\ldots,8$, the null hypothesis is $H_0\colon \underline{\mu}_1=\cdots=\underline{\mu}_8$ (the mean times for both problem assessment and problem solution are the same) vs. $H_a\colon \underline{\mu}_\ell\neq\underline{\mu}_m$, for some $\ell\neq m$. Equivalently, in terms of group effects the hypotheses are $H_0\colon \underline{\tau}_1=\cdots=\underline{\tau}_8=\mathbf{0}$ vs. $H_a\colon \underline{\tau}_\ell\neq\mathbf{0}$, for some ℓ.

Since $p=2$ and $g=8>2$, by line 2 of Table 6.3 we know that the test statistic here is $F^*=\left(\frac{1-\sqrt{\Lambda^*}}{\sqrt{\Lambda^*}}\right)$ which under the null hypothesis will have an *F*-distribution with 14 and 14 degrees of freedom.

The code and output from ***R*** is below. The Wilks' Lambda test statistics is reported to be F^*=17.694 and the associated *p*-value is 1.6×10^{-6}, indicating very strong evidence that there are differences in the mean vectors between the groups.

```
TX<- c(1, 0,  0,  0,  3.0,   6.3,   9.3,
+  1, 0,  0,  0,  2.3,   5.3,   7.6,
+  2, 0,  0,  1,  1.7,   2.1,   3.8,
+  2, 0,  0,  1,  1.2,   1.6,   2.8,
+  3, 0,  1,  0,  6.7,  12.6,  19.3,
+  3, 0,  1,  0,  7.1,  12.8,  19.9,
+  4, 0,  1,  1,  5.6,   8.8,  14.4,
+  4, 0,  1,  1,  4.5,   9.2,  13.7,
+  5, 1,  0,  0,  4.5,   9.5,  14.0,
+  5, 1,  0,  0,  4.7,  10.7,  15.4,
+  6, 1,  0,  1,  3.1,   6.3,   9.4,
+  6, 1,  0,  1,  3.0,   5.6,   8.6,
+  7, 1,  1,  0,  7.9,  15.6,  23.5,
+  7, 1,  1,  0,  6.9,  14.9,  21.8,
+  8, 1,  1,  1,  5.0,  10.4,  15.4,
+  8,1,  1,  1,  5.3,  10.4,  15.7)
> X<-matrix(data=TX,ncol=7,byrow=TRUE)
> CellDat <- data.frame(Group = X[,1],Severity = X[,2],Complexity = X[,3], Experience =
X[,4],Assess = X[,5], Implement = X[,6], Resolve = X[,7])
> fit<-manova(cbind(Assess,Implement) ~ factor(Group),data=CellDat)
> # residuals <- manova(cbind(Assess,Implement) ~ factor(Group),data=CellDat)$residuals
> # Produce Wilks' Lambda and F
> sumfit<-summary(fit,test="Wilks")
> sumfit
              Df     Wilks approx F num Df den Df    Pr(>F)
factor(Group)  7 0.0028615   17.694     14     14 1.614e-06 ***
Residuals      8
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> # Produce Univariate Analyses
> summary.aov(fit)
 Response Assess :
              Df Sum Sq Mean Sq F value    Pr(>F)
factor(Group)  7 59.449  8.4928   41.81 1.114e-05 ***
Residuals      8  1.625  0.2031

 Response Implement :
              Df  Sum Sq Mean Sq F value    Pr(>F)
factor(Group)  7 252.949  36.136   149.4 7.675e-08 ***
Residuals      8   1.935   0.242
```

```
> # Display the "Between" and "Within" Sum of Squares and Cross Products Matrices
> sumfit$SS
$`factor(Group)` (This is the "Between" SSP matrix)
             Assess Implement
Assess     59.44938  120.4094
Implement 120.40937  252.9494

$Residuals (This is the "Within" SSP matrix)
          Assess Implement
Assess     1.625     0.800
Implement  0.800     1.935
> # Estimated Group Means (one variable at a time)
> #
> fitAssess<-aov(Assess ~ 0+factor(Group),data=CellDat)
> model.tables(fitAssess)
Tables of effects
 factor(Group)
factor(Group)
   1    2    3    4    5    6    7    8
2.65 1.45 6.90 5.05 4.60 3.05 7.40 5.15
> fitImplement<-aov(Implement ~ 0+factor(Group),data=CellDat)
> model.tables(fitImplement)
Tables of effects
 factor(Group)
factor(Group)
    1     2     3     4     5     6     7     8
 5.80  1.85 12.70  9.00 10.10  5.95 15.25 10.40
>
```

Follow-Up Analysis: Comparison of Group Means for Each Variable

For each of the $p = 2$ variables we can compare the mean time between all pairs of $g = 8$ groups, for a total of $p\binom{g}{2} = \frac{pg(g-1)}{2} = \frac{2(8)7}{2} = 56$ comparisons. To control the overall error rate at level α, use Bonferroni to make individual comparisons at level $\alpha/56$ which for $\alpha = 0.05$ becomes 0.000893. Here, a confidence interval for the difference between means is:

$(\bar{X}_{\ell i} - \bar{X}_{ki}) \pm t_{n-g}(\frac{\alpha}{pg(g-1)})\sqrt{(\frac{1}{n_\ell} + \frac{1}{n_k})\frac{\mathbf{W}_{ii}}{(n-g)}}$ where $i = 1,\ldots, p$ and $k, \ell = 1,\ldots, g$. For our data this becomes

$(\bar{X}_{\ell 1} - \bar{X}_{k1}) \pm t_8(\frac{0.05}{112})\sqrt{(\frac{1}{2} + \frac{1}{2})\frac{\mathbf{W}_{11}}{8}}$ for comparisons between groups k and ℓ on the mean of variable assessment time and

$(\bar{X}_{\ell 2} - \bar{X}_{k2}) \pm t_8(\frac{0.05}{112})\sqrt{(\frac{1}{2} + \frac{1}{2})\frac{\mathbf{W}_{22}}{8}}$ for comparisons between groups k and ℓ on the mean of variable implementation time.

From the ***R*** output above we get $\mathbf{W} = \begin{pmatrix} 1.625 & 0.8 \\ 0.8 & 1.935 \end{pmatrix}$, and $t_8(\frac{0.05}{112}) = t_8(0.000464) = 5.13$ so that for assessment time, group means must differ by at least $t_8(\frac{0.05}{112})\sqrt{\frac{\mathbf{W}_{11}}{8}} = 5.13\sqrt{\frac{1.625}{8}} = 2.31$ in order to be declared significantly different at (overall) level $\alpha = 0.05$; and for implementation time, group means must differ by at least $t_8(\frac{0.05}{112})\sqrt{\frac{\mathbf{W}_{22}}{8}} = 5.13\sqrt{\frac{1.935}{8}} = 2.52$ in order to be declared significantly different at (overall) level $\alpha = 0.05$. Based on the sample means below, what do you conclude concerning significant differences?

Group Means	2	1	6	5	4	8	3	7
Assessment Time	1.45	2.65	3.05	4.60	5.05	5.15	6.90	7.40

Group Means	2	1	6	4	5	8	3	7
Implementation Time	1.85	5.80	5.95	9.00	10.10	10.40	12.70	15.25

(Courtesy of Dr. Roy St. Laurent)