MAS/MASOR/MASSM - Statistical Analysis - Autumn Term

Solutions 4

1. (a) > pixels<-c(128,126,135,131,137,134,122,128,119,121,117,112,114,112,117,114)
   > mat<-rep(1:4,rep(4,4))
   > material<-factor(mat)
   > plasma<-data.frame(material,pixels)
   > plasma.aov<-aov(pixels~material,data=plasma)
   > summary(plasma.aov)

   Df  Sum Sq Mean Sq F value    Pr(>F)
   material  3  844.69 281.57 14.3016 0.00029
   Residuals 12  236.25  19.68

   Since $p = 0.00029$ then there is strong evidence that there are differences in the number of faulty pixels due to the material type.

   (b) > mean(pixels)
      [1] 122.9375

   > unlist(by(pixels,material,mean))-mean(pixels)
      1 2 3 4
     7.0625 7.3125 -5.6875 -8.6875

   Estimates of mean and effects of material types 1 to 4 are 122.94, 7.0625, 7.3125, -5.6875, and -8.6875, respectively.

   (c) Assuming that none of the objects used to create the data frame have been removed yet, then the confidence interval can be calculated as follows:

   > SED<-sqrt(19.6875/2)
   > tt<-qt(0.975,12)
   > diff<-mean(pixels[material==1])-mean(pixels[material==4])
   > CI<-c(diff-tt*SED,diff+tt*SED)
   > CI
      [1]  8.914029 22.585971

   (d) > multicomp(plasma.aov,error.type="cwe",method="lsd")

   95 % non-simultaneous confidence intervals for specified linear combinations, by the Fisher LSD method

   critical point: 2.1788
   response variable: pixels

   intervals excluding 0 are flagged by '****'

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std.Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>-0.25</td>
<td>3.14</td>
<td>-7.09</td>
</tr>
<tr>
<td>1-3</td>
<td>12.80</td>
<td>3.14</td>
<td>5.91</td>
</tr>
<tr>
<td>1-4</td>
<td>15.80</td>
<td>3.14</td>
<td>8.91</td>
</tr>
<tr>
<td>2-3</td>
<td>13.00</td>
<td>3.14</td>
<td>6.16</td>
</tr>
</tbody>
</table>

1
The only significant differences at the (default) 5% level are between materials 1 and 3, 1 and 4, 2 and 3, and 2 and 4.

(e) The values of the confidence intervals indicate that no improvement can be made by switching to material 3, and that the number of faulty pixels would increase by switching to either material 1 or 2. So should stick with using material 4.

<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2-4</td>
<td>16.00</td>
<td>3.14</td>
<td>9.16</td>
<td>22.80</td>
<td>****</td>
</tr>
<tr>
<td>3-4</td>
<td>3.00</td>
<td>3.14</td>
<td>-3.84</td>
<td>9.84</td>
<td></td>
</tr>
</tbody>
</table>
2. (a)

```r
> time<-c(10,13,11,9,16,21,22,24,18,31,7,6,9,17,7)
> m<-rep(1:3,rep(5,3))
> circuit<-factor(m)
> acm<-data.frame(time,circuit)
> acm.aov<-aov(time~circuit,data=acm)
> summary(acm.aov)
```

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum of Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>circuit</td>
<td>2</td>
<td>554.5333</td>
<td>277.2667</td>
<td>16.12016</td>
</tr>
<tr>
<td>Residuals</td>
<td>12</td>
<td>206.4000</td>
<td>17.2000</td>
<td></td>
</tr>
</tbody>
</table>

The p value is less than 0.01, thus there are differences in the response time according to the circuit type.

(b) > multicomp(acm.aov,alpha=0.01)

99 % simultaneous confidence intervals for specified linear combinations, by the Tukey method

critical point: 3.5681
response variable: time

intervals excluding 0 are flagged by '****'

<table>
<thead>
<tr>
<th>Estimate</th>
<th>Std.Error</th>
<th>Lower Bound</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>-11.4</td>
<td>2.62</td>
<td>-20.80</td>
</tr>
<tr>
<td>1-3</td>
<td>2.6</td>
<td>2.62</td>
<td>-6.76</td>
</tr>
<tr>
<td>2-3</td>
<td>14.0</td>
<td>2.62</td>
<td>4.64</td>
</tr>
</tbody>
</table>

Significant differences between circuits 1 and 2, and also 2 and 3, but not between 1 and 3.

(c) Since we have equal numbers of observations in each group, then we seek to test whether or not $\mu_2 = (\mu_1 + \mu_3)/2$. This yields the vector of coefficients $(1, -2, 1)'$. Testing $\mu_1 = \mu_3$ yields the vector $(1, 0, -1)$. The two vectors are orthogonal to each other.

```r
> c1<-c(1,-2,1)
> c2<-c(1,0,-1)
> ctr<-matrix(c(c1,c2),nrow=3)
> contrasts(acm$circuit)<-ctr
> acm.aov<-aov(time~circuit,data=acm)
> summary(acm.aov, split=list(circuit=list("2 v others"=1, "1 v 3"=2)))
```

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum of Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>circuit</td>
<td>2</td>
<td>554.5333</td>
<td>277.2667</td>
<td>16.12016</td>
</tr>
<tr>
<td>circuit: 2 v others</td>
<td>1</td>
<td>537.6333</td>
<td>537.6333</td>
<td>31.25775</td>
</tr>
<tr>
<td>circuit: 1 v 3</td>
<td>1</td>
<td>16.9000</td>
<td>16.9000</td>
<td>0.98256</td>
</tr>
<tr>
<td>Residuals</td>
<td>12</td>
<td>206.4000</td>
<td>17.2000</td>
<td></td>
</tr>
</tbody>
</table>

ANOVA table shows that responses for circuit 2 differ from those of circuits 1 and 3 ($p \ll 0.01$), but no significant differences between circuits 1 and 3.