Displays for Statistics 5303

Lecture 21

October 23, 2002

Christopher Bingham, Instructor

612-625-7023 (St. Paul)
612-625-1024 (Minneapolis)

Class Web Page
http://www.stat.umn.edu/~kb/classes/5303

Polynomial contrasts with carbon wire data (continued)

```
Cmd> anova("y=(temp+time+degas)^3",pvals:T)
Model used is y=(temp+time+degas)^3
                   DF          SS          MS     P-value
CONSTANT            1       36154       36154     < 1e-08
temp                2      410.69      205.35     < 1e-08
time                2      80.541       40.27     < 1e-08
degas               1     0.46722     0.46722     0.21249
temp.time           4      14.814      3.7035  2.5659e-07
temp.degas          2     0.31361     0.15681     0.58919
time.degas          2     0.70194     0.35097     0.31036
temp.time.degas     4     0.87472     0.21868     0.56559
ERROR1             54       15.85     0.29352

Cmd> c_lin <- vector(-1,0,1); c_quad <- vector(-1,2,-1)

Cmd> contrast(temp,c_lin) #temp main effect line
component: estimate
(1)        5.85      t = 5.85/.1564 = 37.4
component: ss
(1)      410.67
component: se
(1)      0.1564

Cmd> contrast(temp,c_quad) #temp main effect quadratic
component: estimate
(1)       0.075      t = 0.075/0.27089 = -0.277
component: ss
(1)      0.0225
component: se
(1)     0.27089
```

There is a strong linear effect of temperature (t = 37.4) but no quadratic effect (t = 0.277).

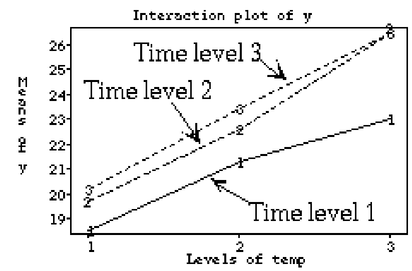
Note $SS_{lin} + SS_{quad} = 410.67 + 0.0225 = 410.69 = SS_{temp}$, in the ANOVA table.

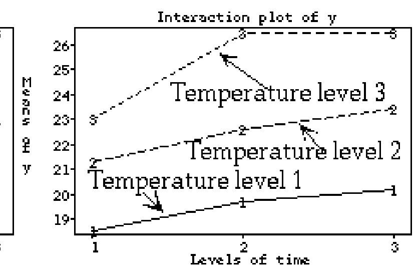
```
Cmd> contrast(time,c_lin) #time main effect linear
component: estimate
(1)      2.4417      t = 2.4417/.1564 = 15.61
component: ss
(1)      71.541
component: se
(1)      0.1564

Cmd> contrast(time,c_quad) #time main effect quadratic
component: estimate
(1)         1.5      t = 1.5/.27089 = 5.537
component: ss
(1)           9
component: se
(1)     0.27089
```

Both linear and quadratic main effect contrasts in time are significant, but the quadratic effect is much smaller. Again $SS_{lin} + SS_{quad} = 71.541 + 9 = 80.541 = SS_{time}$

Look at time by temperature interaction effects.

```
Cmd> interactplot(y,temp,time)

Cmd> interactplot(y,time,temp)
```





The lines are not very parallel, suggesting interaction of **time** and **temp**.

The main effect linear contrasts are isolating the **average** of linear contrasts for each level of the other factor (along each line) separately. That can be near zero when slopes are of opposite signs or curvatures in opposite directions.

It should be fairly clear why we found significant linear main effect contrasts, and why there was some quadratic dependence on time and none on temperature.

Interaction contrasts allow you to extract various features of the interaction. First I calculate the $\overline{y}_{ij\bullet}$.

```
Cmd> ybar_ijdot <- tabs(y,temp,time,means:T); ybar_ijdot
(1,1)      18.538      19.675        20.2
(2,1)      21.275      22.575       23.45
(3,1)          23      26.475      26.488
```

Now I compute linear contrasts comparing **temp** levels in each column (level of **time**) and comparing **time** levels in each row (level of **temp**)

```
Cmd> lins_temp <- vector(sum(c_lin*ybar_ijdot));lins_temp
(1)      4.4625         6.8      6.2875

Cmd> lins_time <- vector(sum(c_lin*ybar_ijdot'));lins_time
(1)      1.6625       2.175      3.4875
```

Note the use of the transpose operator **'** to swap rows and columns so **sum()** would sum accross a row rather than a column.

The averages of these are the same as the main effect linear contrasts computed previously.

```
Cmd> vector(sum(lins_temp)/3, sum(lins_time)/3)
(1)        5.85      2.4417
```


A more "black box" way to find these separate contrasts is using **contrast()** with a third argument.

```
Cmd> contrast(temp,c_lin,time)
component: estimate
(1)      4.4625         6.8      6.2875
component: ss
(1)      79.656      184.96      158.13
component: se
(1)     0.27089     0.27089     0.27089
```

The **estimate** component contains the separate contrasts for each level of **time**.

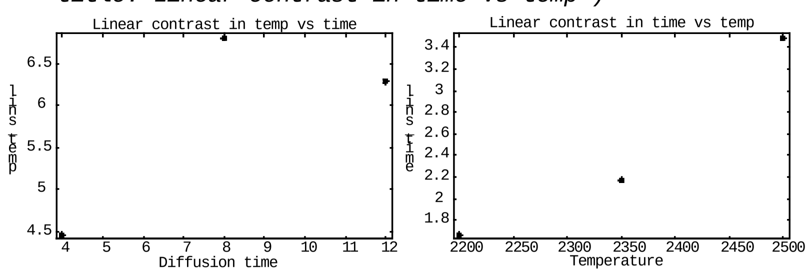
```
Cmd> contrast(time,c_lin,temp)
component: estimate
(1)      1.6625       2.175      3.4875
component: ss
(1)      11.056      18.923      48.651
component: se
(1)     0.27089     0.27089     0.27089
```

Now the **estimate** component contains the separate contrasts for each level of **temp**.

This also provides standard errors and SS.

Let's see how these contrasts vary over levels of the other factor.

```
Cmd> actualtemp <- run(2200,2500,150)

Cmd> actualtime <- run(4,12,4)

Cmd> plot(actualtime,lins_temp,xlab:"Diffusion time",\
    title:"Linear contrast in temp vs time")

Cmd> plot(actualtemp,lins_time,xlab:"Temperature",\
    title:"Linear contrast in time vs temp")
```



The pattern of linear contrasts on the left looks like a curved dependence on time, while the pattern on the right is fairly close to linear on temperature, with only a hint of curvature.

Since a linear contrast is proportional to a least squares slope, it appears the slope on **temp** may depend quadratically on **time**, while the slope on **time** depends linearly on temperature.

Linear contrast coefficients are good for extracting information about a linear trend, so we can apply them to the separate linear contrast values to extract information about their linear dependence on the othre factor.

```
Cmd> sum(c_lin*lins_temp)
(1)       1.825
```

You can do this in one step using an interaction contrast created from the separate contrasts using **outer()**.

Linear by Linear

```
Cmd> outer(c_lin,c_lin) # same thing
(1,1)           1           0          -1
(2,1)           0           0           0
(3,1)          -1           0           1

Cmd> contrast("temp.time",outer(c_lin,c_lin))
component: estimate
(1)       1.825
component: ss
(1)      6.6612
component: se
(1)     0.38309
```

This is the linear by linear interaction contrast computed above. It is highly significant (t = 1.8254/.38309 = 4.76).

The SS = 6.6612 is much less than the overall interaction $SS_{time.temp}$ = 14.814, so there is a lot more interaction to "explain".

Linear in `temp` by quadratic in `time`

```
Cmd> outer(c_lin,c_quad) # linear by quadratic
(1,1)          1          -2          1
(2,1)          0           0          0
(3,1)         -1           2         -1

Cmd> contrast("temp.time",outer(c_lin,c_quad))
component: estimate
(1)        2.85
component: ss
(1)       5.415
component: se
(1)     0.66353
```

t = 2.85/0.66353 = 4.30 is significant

Quadratic in `temp` by linear in `time`

```
Cmd> contrast("temp.time",outer(c_quad,c_lin))
component: estimate
(1)        -0.8
component: ss
(1)     0.42667
component: se
(1)     0.66353
```

t = 0.8/0.66353 = 1.21 is not significant

Quadratic by quadratic

```
Cmd> contrast("temp.time",outer(c_quad,c_quad))
component: estimate
(1)      -3.225
component: ss
(1)      2.3112
component: se
(1)      1.1493
```

t = -3.225/1.1493 = -2.81, significant at the 1% level.

Conclusion:

For each level of **time**, the response on **temp** is curved, but the curvature varies from level to level.

For each level of **temp**, the response on **time** is curved, but the curvature varies from level to level.

The 4 interaction contrasts SS add up to the overall 4 degree of freedom $SS_{temp.time}$

6.6612 + 5.415 + 0.42667 + 2.3112 = 14.814.

What sort of a model underlies the use of these contrasts?

Suppose you have a two factor model where the two factors are defined by values x_{Ai}, i = 1,..., a and x_{Bj}, j = 1,...,b of quantitative variables x_A and x_B so that $\mu_{ij} = f(x_{Ai},x_{Bj})$ for some function $f(x_A,x_B)$.

Suppose for each fixed value x_B, the dependence of $f(x_A,x_B)$ on x_A is quadratic.

$$f(x_A,x_B) = \beta_0(x_B) + \beta_1(x_B)x_A + \beta_2(x_B)x_A^2,$$

where $\beta_k(x_B)$ are the coefficients for that level of x_B.

Now suppose the dependence of each *coefficient* $\beta_k(x_B)$ on x_B is also quadratic

$$\beta_0(x_B) = \delta_{00} + \delta_{01}x_B + \delta_{02}x_B^2$$
$$\beta_1(x_B) = \delta_{10} + \delta_{11}x_B + \delta_{12}x_B^2$$
$$\beta_2(x_B) = \delta_{20} + \delta_{21}x_B + \delta_{22}x_B^2$$

Then

$$f(x_A,x_B) = \delta_{00} + \delta_{10}x_A + \delta_{01}x_B + \delta_{20}x_A^2 + \delta_{11}x_Ax_B + \delta_{02}x_B^2 + \delta_{21}x_A^2x_B + \delta_{12}x_Ax_B^2 + \delta_{22}x_A^2x_B^2$$

In the context of this model, when $\delta_{11} = \delta_{12} = \delta_{21} = \delta_{22} = 0$, all $\alpha\beta_{ij} = 0$ and

- the A_{quad} by B_{quad} F or t tests H_0: $\delta_{22} = 0$.
- Provided $\delta_{22} = 0$, the A_{lin} by B_{quad} F or t tests H_0: $\delta_{12} = 0$
- Provided $\delta_{22} = 0$, the A_{quad} by B_{lin} F or t tests H_0: $\delta_{21} = 0$
- Provided $\delta_{12} = \delta_{21} = \delta_{22} = 0$, the A_{lin} by B_{lin} F or t tests H_0: $\delta_{11} = 0$

One degree of freedom for non-additivity

Part of the purpose of using interaction contrasts is the hope they will help you describe the pattern of interaction.

Another approach, which is particularly useful in experiments with no replication, goes under the name of **Tukey's one degree of freedom for non-additivity**.

Tukey was looking for the simplest sort of model beyond an additive model. He asked the question,

> What sort of a model would I get if the data were derived by some transformation of an additive model?

That is, in the two factor case, suppose there is an **additive model** for a response $\tilde{y}_{ij}$ which is related to y by $\tilde{y}_{ij} = f(y_{ij})$ for some function f. For example, you might have $\tilde{y}_{ij} = \log_{10} y_{ij}$ or $\tilde{y}_{ij} = y_{ij}^{p}$.

To say that $\tilde{y}_{ij}$ has an additive model means that the mean $\tilde{\mu}_{ij}$ of $\tilde{y}_{ij}$ for levels i and j of factors A and B has the form

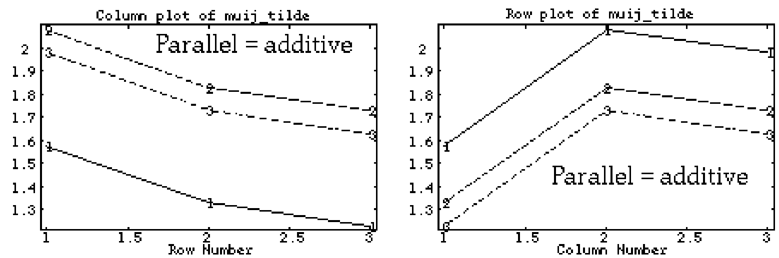
$$\tilde{\mu}_{ij} = \tilde{\mu} + \tilde{\alpha}_i + \tilde{\beta}_j.$$

Suppose $y = g(\tilde{y})$ is the inverse transformation to f. For example, when $\tilde{y}_{ij} = \log_{10} y_{ij}$, $y_{ij} = g(\tilde{y}_{ij}) = 10^{\tilde{y}_{ij}}$ and when $\tilde{y}_{ij} = y_{ij}^{p}$, $y_{ij} = g(\tilde{y}_{ij}) = \tilde{y}_{ij}^{1/p}$.

Although it won't be exact, in many cases the means will be similarly related.
That is $\mu_{ij} = E(y_{ij}) = g(\tilde{\mu}_{ij})$

Here is a 3 by 3 additive table of $\tilde{\mu}_{ij}$

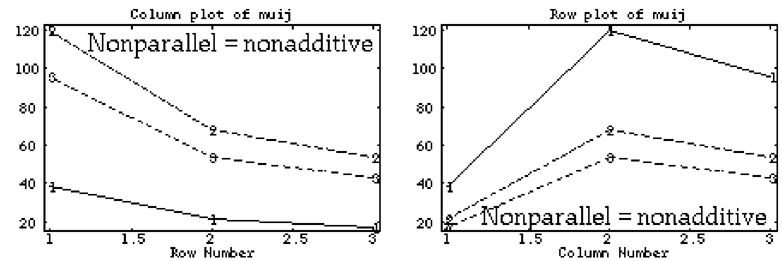
```
Cmd> muij_tilde
(1,1)        1.58         2.08         1.98
(2,1)        1.33         1.83         1.73
(3,1)        1.23         1.73         1.63

Cmd> colplot(muij_tilde,title:"Column plot of muij_tilde")

Cmd> rowplot(muij_tilde,title:"Row plot of muij_tilde")
```



In this scale the factors act additively, with no interaction.

Now transform to $\mu = 10^{\tilde{\mu}}$ and make interaction plots.

```
Cmd> muij <- 10^muij_tilde  # muji <- g(muij_tilde)

Cmd> colplot(muij,title:"Column plot of muij")

Cmd> rowplot(muij,title:"Row plot of muij")
```



The lines are no longer parallel. The model is no longer additive.

When $\mu_{ij} = g(\tilde{\mu} + \tilde{\alpha}_i + \tilde{\beta}_j)$, and the $(\tilde{\alpha}_i + \tilde{\beta}_j)/\tilde{\mu}$ are not too big, approximately,

$$\mu_{ij} \approx g(\tilde{\mu}_{ij}) = g(\tilde{\mu} + \tilde{\alpha}_i + \tilde{\beta}_j)$$
$$\approx g(\tilde{\mu}) + c(\tilde{\alpha}_i + \tilde{\beta}_j) + d(\tilde{\alpha}_i + \tilde{\beta}_j)^2$$

where $c = g'(\tilde{\mu})$ and $d = g''(\tilde{\mu})/2$ (derivatives of $g(\tilde{\mu})$)

After some simplification, leaving out terms of 3rd or 4th degree in $\tilde{\alpha}_i$ and $\tilde{\beta}_j$, you can find effects $\{\alpha_i\}$ and $\{\beta_j\}$ such that

$$\mu_{ij} \cong \mu + \alpha_i + \beta_j + \gamma\alpha_i\beta_j$$

where $\gamma = 2d/c^2 = g''(\tilde{\mu})/g'(\tilde{\mu})^2$. This is a model with interaction with a simple model for the interaction effects:

$$\alpha\beta_{ij} = \gamma\alpha_i\beta_j,\ i = 1,\ldots,a,\ j = 1,\ldots,b$$

requiring just 1 additional parameter γ.

In the particular case when $f(\mu) = \mu^p$ and $g(\tilde{\mu}) = \tilde{\mu}^{1/p}$,

$\gamma = g''(\tilde{\mu})/g'(\tilde{\mu})^2 = (1 - p)/\tilde{\mu}^{1/p} \cong (1 - p)/\mu$

Then $p = 1 - \mu\gamma$. From estimates $\hat{\mu}$ and $\hat{\gamma}$ pf γ and μ you can get a handle on a possible transformation, $y \to y^{1-\hat{\mu}\hat{\gamma}}$.

The model

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma\alpha_i\beta_j + \varepsilon_{ij}$$

might be called the **1-dofna two-factor model**. If it is appropriate, then the null hypothesis H_0: $\alpha\beta_{ij} = 0$ all i and j is equivalent to H_0: $\gamma = 0$

How can you fit the 1-dofna model?

You can do it in two stages:

1. Fit the additive model $\mu_{ij} = \mu + \alpha_i + \beta_j$ and from it find the fitted values $\hat{\mu}_{ij} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$.
2. Compute $z_{ij} = (\hat{\mu}_{ij} - \hat{\mu})^2/2 = (\hat{\alpha}_i + \hat{\beta}_j)^2/2$
3. Fit the model with an additional term

$$y_{ij} = \mu + \alpha_i + \beta_j + \gamma z_{ij} + \varepsilon_{ij}$$

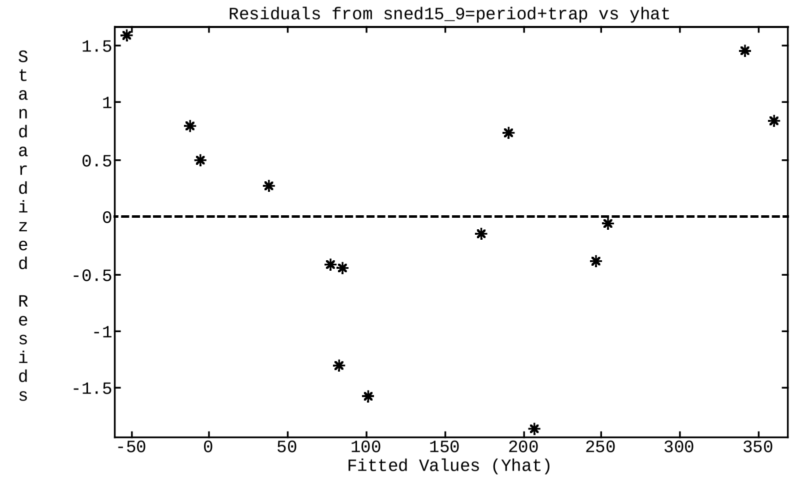
The F- or t-statistic for z is a test of H_0: $\gamma = 0$, that is H_0: model is additive.

Here is an example from Snedecor and Cochran Table 15.9.1. The response is the number of insects caught in a trap over night. There were 5 nights and three traps.

```
Cmd> sned15_9 <- vector(19.1,23.4,29.5,23.4,16.6,\
 50.1,166.1,223.9,58.9,64.6, 123,407.4,398.1,229.1,251.2)

Cmd> print(matrix(sned15_9,5,\
     labels:structure("Trap ","Night ")),format:"11.1f")
MATRIX:
           Night 1     Night 2     Night 3
Trap 1        19.1        50.1       123.0
Trap 2        23.4       166.1       407.4
Trap 3        29.5       223.9       398.1
Trap 4        23.4        58.9       229.1
Trap 5        16.6        64.6       251.2

Cmd> period <- factor(rep(run(5),3))

Cmd> trap <- factor(rep(run(3),rep(5,3)))

Cmd> anova("sned15_9=period + trap", fstat:T)
Model used is sned15_9=period + trap
             DF          SS          MS           F     P-value
CONSTANT      1  2.8965e+05  2.8965e+05    75.70780  2.3739e-05
period        4       52066       13016     3.40223     0.06611
trap          2  1.7333e+05       86667    22.65276  0.00050731
ERROR1        8       30607      3825.9

Cmd> fitted <- sned15_9 - RESIDUALS  # muij_hat
```

`fitted` contains the additive fit because `RESIDUALS = sned15_9 - fitted`.


```
Cmd> resvsyhat(title:"Residuals from sned15_9=period+trap vs
yhat")
```



There seems to be some sort of quadratic dependence of residuals on fitted values.

```
Cmd> grandmean <- describe(sned15_9,mean:T)

Cmd> z <-  (fitted - grandmean)^2/2

Cmd> anova("sned15_9=period + trap + z", fstat:T)
Model used is sned15_9=period + trap + z
WARNING: summaries are sequential
             DF          SS          MS           F     P-value
CONSTANT      1  2.8965e+05  2.8965e+05   322.45184  4.1032e-07
period        4       52066       13016    14.49064   0.0016932
trap          2  1.7333e+05       86667    96.48180  8.0263e-06
z             1       24319       24319    27.07330   0.0012486
ERROR1        7      6287.9      898.27
```

The original SS_{error} = 30607 was effectively the period by trap interaction SS. The 1 degree of freedom SS = 24319 associated with z - **1 degree of freedom for non-additivity** or **1-dofna** - has "explained" about 80% of that. $F_{1,7}$ = 27.07 is highly significant.

The error MS has been reduced from 3825.9 to 898.27 and both main effects are significant.