H. Spliid, October 20, 2010

Solution to exercise 38

A reasonable model for the experiment is:

$$Y_{ij} = \mu + \tau_j + E_{ij}$$

that is a one-way analysis of variance model.

Variation	SSQ	f	s^2	F
Sugar type	9.9176	4	2.48	23.9
Residual	2.0760	20	0.104	
Total	11.9936	24		

$$F = \frac{9.9176/4}{2.0760/20} = 23.9 > F(4,20)_{0.001}$$

Thus a significant difference between the yield for different sugar types is seen.

Question 2
Since there is no obvious or natural structure for the 5 types of sugar, it is reasonable to try either Duncans multiple range test or Newman-Keuls test in order to assess whether one or more sugar types are clearly different (smaller or larger) than the others or if some kind of pattern emerges. I prefer Newman-Keuls test, since it is a bit more conservative:

Sugar types	A	B	C	D	E
$\overline{Y}_i$	1.96	1.76	2.42	2.74	3.54
	(2)	(1)	(3)	(4)	(5)

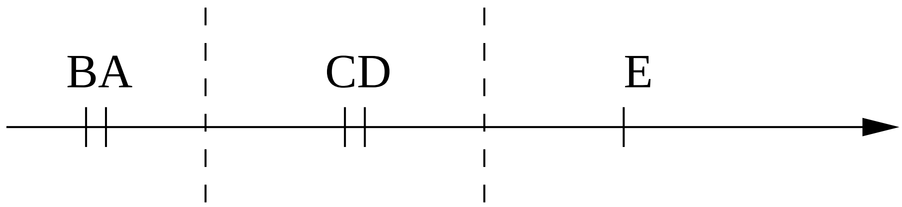
$p =$	2	3	4	5
Studentized ranges ($d.f. = 20$)	2.95	3.58	3.96	4.24
Least sign. range for $\overline{Y}$	0.425	0.516	0.571	0.611

We have used $s_{\overline{Y}} = \sqrt{0.1038/5} = 0.1441$

(5) - (1) = 1.78 > 0.611
(4) - (1) = 0.98 > 0.571
(3) - (1) = 0.66 > 0.516
(2) - (1) = 0.20 < 0.425
(5) - (2) = 1.58 > 0.571

$$
\begin{aligned}
(4) - (2) &= 0.78 > 0.516 \\
(3) - (2) &= 0.46 > 0.425 \\
(5) - (3) &= 1.12 > 0.516 \\
(4) - (3) &= 0.32 < 0.425 \\
(5) - (4) &= 0.80 > 0.425
\end{aligned}
$$

Conclusion:



Question 3
A polynomium of degree 3 is compared with the data be means of orthogonal polynomials.

					$\sum \xi_i^2$	λ
Linear	-3	-1	1	3	20	2
Quadratic	1	-1	-1	1	4	1
Cubic	-1	3	-3	1	20	10/3
Data : $T_{.j}$	9.5	12.8	12.9	11.4		

Using these coefficients we find contrasts and sums of squares for the 1st, 2nd and 3rd order polynomials:

$$
\begin{array}{lcrlcccl}
C_{lin} & = & 5.8 & SSQ_{lin} & = & (5.8^2)/(4 \cdot 20) & = & 0.4205 \\
C_{quad} & = & \text{-}4.8 & SSQ_{quad} & = & (4.8^2)/(4 \cdot 4) & = & 1.4400 \\
C_{cub} & = & 1.6 & SSQ_{cub} & = & (1.6^2)/(4 \cdot 20) & = & 0.0320 \\
\text{Control:} & & & SSQ_{conc} & = & & & 1.8925
\end{array}
$$

ANOVA

Variation	SSQ	f	Test	
Concentration	1.8925	3		
Linear	0.4205	1	5.34	$> F(1,12)_{0.05}$
Quadratic	1.4400	1	18.29	$>> F(1,12)_{0.01}$
Cubic	0.0320	1	0.41	$< F(1,12)_{0.50}$
Residual	0.9450	12		
Total	2.8375	15		

The 3rd order polynomium is not significant, but the quadratic is strongly significant ($\alpha << 0.01$). It is therefore concluded that a polynomium of degree 2 is a reasonable choice for describing the influence of the concentration on the yield.

The polynomial model has the following appearance:

$$Y_{ij} = A_0 \cdot P_0 + A_1 \cdot P_1(z_j) + A_2 \cdot P_2(z_j) + E_{ij}$$

where

$$z_j = \frac{x_j - \overline{x}}{\Delta_x}$$

and $\overline{x} = 5$ and $\Delta_x = 2$.

Estimation of polynomium:

$$\begin{array}{llllll} \widehat{A}_0 & = \overline{Y} & = 2.9125 & P_0 & = 1 \\ \widehat{A}_1 & = 5.8/(4 \cdot 20) & = 0.0725 & P_1(z) & = 2z \\ \widehat{A}_2 & = -4.8/(4 \cdot 4) & = \text{-}0.3000 & P_2(z) & = 1\,(z^2 - (4^2-1)/12) = z^2 - 5/4 \end{array}$$

Note that the coefficient estimates ($\widehat{A}_i$) are computed exactly like the corresponding SSQ's except that the numerators (the contrasts) are not squared.

Polynomium:

$$\begin{aligned} \widehat{Y}(z) &= 2.9125 + 0.0725 \cdot 2z - 0.3(z^2 - 5/4) \\ &= 3.2875 + 0.145z - 0.3z^2 \end{aligned}$$

Since $\overline{x} = 5$, og $\Delta_x = 2$, we have $z = (x-5)/2$. Inserting this the estimated polynomium is found to be:

$$\widehat{Y}(x) = 1.05 + 0.823x - 0.075x^2$$

Finally the variance corresponding to the uncertainty is esimated

$$\widehat{\sigma}_E^2 = (0.032 + 0.945)/(1+12) = 0.0751 \simeq 0.27^2$$

where the sum of squares from the cubic term not used in the model is included.

. o O o .