Jack Rodoni
STAT 642
HW 04
2/24/22

Read H.O.4 - Chp 4 in Design/Anova Book

See R code H.O.4: • pg 14,

1.) Use the data from the heat loss study to answer the following questions. Contrasts of interest to the researchers were:

- C_1: Control vs Average of the means of the 4 thickness levels
- C_2: Linear trend across the 5 thickness levels
- C_3: Quadratic Trend across the 5 thickness levels
- C_4: Cubic trend across the 5 thickness levels

- See H.O.4 pg 50

$C_1 = 4\mu_1 - (\mu_2 + \mu_3 + \mu_4 + \mu_5)$

$C_2 = -2\mu_1 - \mu_2 + \mu_4 + 2\mu_5$

$C_3 = 2\mu_1 - \mu_2 - 2\mu_3 - \mu_4 + 2\mu_5$

$C_4 = -\mu_1 + 2\mu_2 - 2\mu_4 + \mu_5$

(a) Are the four contrasts mutually orthogonal? If not, select three contrasts from the four contrasts which are mutually orthogonal, — what does she mean by this? Show 3 of the 4 are mutually orthogonal?

NOTE: See H.O.4 pg 8:

Two contrasts $C_1 = \sum_{i=1}^{t} k_i\mu_i$ $C_2 = \sum_{i=1}^{t} d_i\mu_i$ are said to be orthogonal if $\sum_{i=1}^{t} k_i d_i = 0$

- $C_1 \cdot C_2 = [4\ -1\ -1\ -1\ -1] \cdot [-2\ -1\ 0\ 1\ 2]' = (4)(-2) + (-1)(-1) + (-1)(1) + (-1)(2)$
  $= -10 \neq 0 \Rightarrow C_1$ & C_2 are **NOT** orthogonal
- $C_1 \cdot C_3 = (4)(2) + (-1)(-1) + (-1)(-2) + (-1)(-1) + (-1)(2) = 10 \neq 0 \Rightarrow C_1$ & C_2 are **NOT** orthogonal
- $C_1 \cdot C_4 = (4)(-1) + (-1)(2) + (-1)(-2) + (-1)(1) = -5 \neq 0 \Rightarrow C_1$ & C_3 are **NOT** orthogonal

mutually orthogonal

- $C_2 \cdot C_4 = (-2)(-1) + (-1)(2) + (0)(0) + (1)(-2) + (2)(1) = 0 \Rightarrow C_2$ & C_4 are orthogonal
- $C_2 \cdot C_3 = (-2)(2) + (-1)(-1) + (0)(-2) + (1)(-1) + (2)(2) = 0 \Rightarrow C_2$ & C_3 are orthogonal
- $C_3 \cdot C_4 = (2)(-1) + (-1)(2) + (-2)(0) + (-1)(-2) + (2)(1) = 0 \Rightarrow C_3$ & C_4 are orthogonal.

1.) (Cont'd)

(f) Is there a trend in the mean heat losses as a function of coating thickness. Justify your answer.

- Yes, there seems to be a linear trend in the mean heat losses as a function of coating thickness.
  - First I fit the cubic model to the data, the cubic term was not significant, so I removed it. After removing the cubic term, I fit a quadratic model to the data and the quadratic term wasn't significant so I removed it. After the linear model to the data, we see the linear trend is significant.

2.) Use the data from the heat loss study to answer the following questions.

(a) Which thicknesses have the smallest mean heat loss w/ probability of correct selection of 0.95.

[* Hsu's procedure]

- The two largest thicknesses 60 & 80 have the smallest mean heat loss w/ probability of correct selection of 0.95.

~~(b)~~ Do any of the coatings have a mean heat loss less than the mean heat loss for the panes w/ no coating? Use $\alpha = 0.05$ in your answer.

- Yes, using Dunnett's procedure we found the coatings w/ thicknesses 60 & 80 have a mean heat loss less than the mean heat loss for the panes w/ no coating at the $\alpha_2 = 0.05$ level.

(c) Which pairs of the 5 treatment means are different using the Tukey procedure at the $\alpha = 0.05$ significance level.

- The pairs that have significant differences at the $\alpha = 0.05$ significance level are (0,60), (0,80), (20,60), (20,80), (40,60), (40,80)

1.) (Contd)

(b) Provide an estimate of each contrast along w/ the standard error of the estimator

	Estimate	Std Error
C_1:	6.21	1.034
C_2:	-6.12	0.731
C_3	-0.14	0.465
C_4	0.19	0.731

(c) Use the Scheffé test at the $\alpha = 0.05$ level of significance to test the significance of the four contrasts.

Contrast	$\lvert \hat{C}_i \rvert$	S_i	Conclusion
C_1	6.21	3.322	Significant evidence that $C_1 \neq 0$
C_2	6.12	2.349	Significant evidence that $C_2 \neq 0$
C_3	0.14	1.779	Evidence is not significant that $C_3 \neq 0$
C_4	0.19	2.349	Evidence is not significant that $C_4 \neq 0$

(d) Use the Bonferroni test w/ $\alpha_E = 0.05$ level of significance to test the significance of the 4 contrasts.

Contrast	p-value	α_E/m	Conclusion (significant if p-value < α_f)
C_1	3.084×10^{-7}	0.0125	Significant evidence that $C_1 \neq 0$
C_2	1.013×10^{-10}	0.0125	Significant evidence that $C_2 \neq 0$
C_3	0.8722	0.0125	Evidence not significant that $C_3 \neq 0$
C_4	0.796	0.0125	Evidence not significant that $C_4 \neq 0$

~~(e)~~ Test the three trend contrasts simultaneously using the matrix approach from H.O. 4.

$$H = \begin{bmatrix} -2 & -1 & 0 & 1 & 2 \\ 2 & -1 & -2 & -1 & 2 \\ -1 & 2 & 0 & -2 & 1 \end{bmatrix} \qquad \hat{\mu}' = [10.73, 9.92, 9.15, 8.62, 8.32]$$

$$(X'X) = \begin{bmatrix} 10 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 \\ 0 & 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 & 10 \end{bmatrix}$$

• Test: $H_0: C_2 = C_3 = C_4 = 0$ H_A: At least one $C_i \neq 0$. $\Leftrightarrow$ $H_0: H\underline{\mu} = 0$ $H_A: H\underline{\mu} \neq 0$.

$$SS_H = (H\hat{\underline{\mu}} - \underline{0})' \left(H(X'X)^{-1}H'\right)^{-1} (H\hat{\underline{\mu}} - \underline{0}) = 37.5045$$

$$F = \frac{SS_H / k}{MSE} = \frac{37.5045/3}{0.5727891} = 21.82566$$

$$P\left[F_{3,46} \geq 21.82566\right] = 6.15 \times 10^{-9} \Rightarrow$$

There is significant evidence that one of our trend contrasts $C_i \neq 0$.

3) In a DOE textbook, the author states that, "If you are testing M hypotheses: Equality of the relationship $\alpha_F \leq M\alpha_{pc}$ holds when the M tests are independent." Is this statement true? If yes, provide a proof. If no, provide a condition under which equality doesn't hold.
[* see H.O.4 pg 19-21]

No, this statement isn't true. Equality holds if the tests are disjoint.

- Let A_i be the event that we commit a type I error and let
$$P[A_i] = \alpha_{pc} \quad \forall i \in [1, \dots, n]$$
- Then α_F = P[at least 1 type I error occurs]
$$= P\left[\bigcup_{i=1}^{M} A_i\right] = \sum_{i=1}^{M} P(A_i) - \sum_{i<j}^{M} P(A_i \cap A_j) + \sum_{i<j<k}^{M} P(A_i \cap A_j \cap A_k) + \dots + (-1)^{M-1} \sum_{i<\dots<n} P\left(\bigcap_{i=1}^{M} A_i\right)$$
- Say
$$C = -\sum_{i<j}^{M} P(A_i \cap A_j) + \sum_{i<j<k}^{M} P(A_i \cap A_j \cap A_k) + \dots + (-1)^{M-1} \sum_{i<\dots<n} P\left(\bigcap_{i=1}^{M} A_i\right)$$

$$\alpha_F = P\left[\bigcup_{i=1}^{M} A_i\right] = M\alpha_{pc}$$

- note $C = 0 \iff \forall i \neq j \in [1, \dots, M]$
$$P(A_i \cap A_j) = 0.$$
- Thus, equality doesn't hold if $\exists\, i \neq j \in [1, \dots M]$ s.t.
$$P(A_i \cap A_j) > 0.$$

4) A CRD w/ $t=4$ fixed effects treatments and $n_1=5, n_2=4, n_3=5$ & $n_4=6$ reps/treatment was run. The experimenter constructed two contrasts in the treatment means:

$$C_1 = 3\mu_1 + \mu_2 - \mu_3 - 3\mu_4 \qquad C_2 = -\mu_1 + \mu_2 + \mu_3 - \mu_4$$

(a) Are the two contrasts orthogonal

[Recall: Two contrasts $C_1 = \sum k_i\mu_i$, $C_2 = \sum d_i\mu_i$ are orthogonal if $\sum k_i d_i = 0$.]

Let $c_1' = [3\ 1\ -1\ -3]$, $c_2' = [-1\ 1\ 1\ -1]$

$c_1 \cdot c_2 = (-3) + (1) + (-1) + (3) = 0,$

• Yes the two contrasts C_1 & C_2 are orthogonal.

(b) Are the sample estimators of the two contrasts independently distributed?

[* see H.O.4 pg 8]

• $\sum \frac{k_i d_i}{n_i} = \frac{(-3)}{5} + \frac{(1)}{4} + \frac{(-1)}{5} + \frac{3}{6} = -\frac{4}{5} + \frac{3}{4} \neq 0$

$\Rightarrow$ the sample estimators of the two contrasts are not independently distributed.

(c) Construct a contrast, other than C_1 which is orthogonal to C_2.

$c_2' = [-1, 1\ 1\ -1]$ $\qquad c_3' \cdot c_2' = (1) + (-1) = 0.$

$c_3' = [-1\ 0\ -1\ 0]$

5.) A study was conducted to separately analyze the effects of 10 SNPs comparing people w/ type I diabetes vs. controls. The p-values from the 10 separate analyses are given in the following table.

SNP	P-value	SNP	P-value
1	0.0001	6	0.0911
2	0.0058	7	0.2012
3	0.0132	8	0.5718
4	0.0289	9	0.8912
5	0.0498	10	0.9011

(a) w/ no adjustment for multiple testing which SNP's have significant effects. Use PCR = 0.05

SNPs 1-5 are significant using $\alpha_{pc} = 0.05$

(b) Using a Bonferroni adjustment for multiple testing, which SNPs have significant effects? Use FWER

- $\alpha_F = 0.05 \Rightarrow \alpha_{pc} = \frac{0.05}{10} = 0.005$
- SNP 1 is the only SNP w/ a significant effect using the Bonferroni adjustment for multiple testing.

(c) Using the FDR method to adjust for multiple testing, which SNPs have significant effects. Use FDR = 0.05

- SNPs 1-3 have significant effects, using the FDR method to adjust for multiple testing.

6.) (a) False, must also have the same sample sizes, or $\sum k_i d_i / n_i = 0$.

(b) True

(c) Same as (a), must also have same sample size or $\sum k_i d_i / n_i = 0$.

(d) False, The BH FDR testing procedure has higher power than say the Bonferroni procedure. Thus, the BH FDR testing procedure will find at least as many pairs of treatment means to be different

(e) True, see H.0.4, bottom pg 32.

(f) False, see H.0.4, top pg 33

(g) False, see H.0.4, top pg 34

(h) True if equal sample sizes, False if NOT equal sample size

(i) False

(j) True, we can fit a polynomial up to the degree t-1. In this case t-1 = 4