Multiple comparisons

The problem: why not just look at all the pairs of means with t-tests?

α: probability of rejecting the null hypothesis, given that it is true

ex. 4 means, 6 tests at $\alpha = 0.05$

test 1	H_0: $\mu_1 - \mu_2 = 0$	(.95)
test 2	H_0: $\mu_1 - \mu_3 = 0$	(.95)
test 3	H_0: $\mu_1 - \mu_4 = 0$	(.95)
⋮	⋮	⋮
test 6	H_0: $\mu_3 - \mu_4 = 0$	(.95)

If the test statistics are independent (they are not), the probability that *all six* tests avoided type I errors (given H_0 is true for all six tests) is

$$(1 - 0.05)^6 = (0.95)^6 = 0.735$$

So, the probability that *one or more* tests commit a type I error (given H_0's true) is

$$\alpha_E = 1 - 0.735 = 0.265$$

(α_E is the **experimentwise type I error probability**)

Contrasts

$$l = a_1\mu_1 + a_2\mu_2 + \cdots + a_t\mu_t$$

is a **linear contrast** of the **population means** μ_1, μ_2, ..., μ_t. Here the a_i's are constants which sum to zero:

$$a_1 + a_2 + \cdots + a_t = 0$$

ex. $a_1 = 0$, $a_2 = 0$, $a_3 = 1$, $a_4 = -1$

$$l = \mu_3 - \mu_4 \qquad \text{(pairwise comparison)}$$

ex. $a_1 = 1$, $a_2 = -\frac{1}{3}$, $a_3 = -\frac{1}{3}$, $a_4 = -\frac{1}{3}$

$$l = \mu_1 - \frac{\mu_2 + \mu_3 + \mu_4}{3} \qquad \text{(compares } \mu_1 \text{ to the others)}$$

(usually redefine as $l = 3\mu_1 - \mu_2 - \mu_3 - \mu_4$, so that all the a_i's are integers)

Assume the usual model for 1-way AOV: within each of t populations, $Y_i \sim \text{normal}(\mu_i, \sigma^2)$; random sample of size n_i from each population. Parameter estimates:

$$\widehat{\mu}_i = \overline{y}_{i\cdot} \ , \qquad \hat{\sigma}^2 = s_{\mathrm{W}}^2 = \frac{\mathrm{SSW}}{(n_{\mathrm{T}} - t)}$$

(Estimate of) a linear contrast:

$$\widehat{l} = a_1\overline{y}_{1\cdot} + a_2\overline{y}_{2\cdot} + \cdots + a_t\overline{y}_{t\cdot} \qquad \text{(unbiased)}$$

$$V\left(\widehat{l}\right) = \sigma^2\left(\frac{a_1^2}{n_1} + \frac{a_2^2}{n_2} + \cdots + \frac{a_t^2}{n_t}\right)$$

$$\widehat{V}\left(\widehat{l}\right) = s_{\mathrm{W}}^2\left(\frac{a_1^2}{n_1} + \frac{a_2^2}{n_2} + \cdots + \frac{a_t^2}{n_t}\right)$$

$$\left(= \frac{s_{\mathrm{W}}^2}{n}\left(a_1^2 + a_2^2 + \cdots + a_t^2\right) \text{ if } n_i\text{'s are the same}\right)$$

100(1 − α)% CI for l:

$$\widehat{l} \pm t_{\alpha/2}\sqrt{\widehat{V}\left(\widehat{l}\right)}$$

where $t_{\alpha/2}$ is the percentile from the Student's t distribution with $n_{\mathrm{T}} - t$ degrees of freedom

Hypothesis test

Hypotheses:

H_0: $l = 0$

$$H_a: \; l \left\{ \begin{array}{c} > \\ < \\ \neq \end{array} \right\} 0$$

Test statistic:

$$t = \frac{\hat{l} - 0}{\sqrt{\hat{V}\left(\hat{l}\right)}}$$

Rejection region uses Student's t percentile:

$$\text{reject } H_0 \text{ if} \left\{ \begin{array}{c} t \geq t_\alpha \\ t \leq -t_\alpha \\ |t| \geq t_{\alpha/2} \end{array} \right\} \text{ where df} = n_T - t$$

Note: book uses an F test for the two-sided test with H_a: $l \neq 0$. Interestingly, if T has a Student's t distribution with v df, then $F = T^2$ has an F distribution with (1, v) df. So,

reject H_0 if $|t| \geq t_{\alpha/2}$ $(n_T - t$ df)

is the same as

reject H_0 if $t^2 = f \geq f_\alpha$ $(1, n_T - t$ df)

SAS: (suppose model has four means)

```
  ⋮
PROC GLM;
CLASS TRT;
MODEL Y=TRT;
CONTRAST '1 VS 2' TRT 1 -1 0 0;
CONTRAST '1 VS 3' TRT 1 0 -1 0;
CONTRAST 'CTRL VS REST' TRT -3 1 1 1;
```

Orthogonal contrasts

Two (estimated) contrasts given by $l_1 = \Sigma a_i \overline{y}_{i\cdot}$ and $l_2 = \Sigma b_i \overline{y}_{i\cdot}$ are **orthogonal** if

$$\frac{a_1 b_1}{n_1} + \frac{a_2 b_2}{n_2} + \cdots + \frac{a_t b_t}{n_t} = 0$$

$(a_1 b_1 + a_2 b_2 + \cdots + a_t b_t = 0$ if the n_i's are equal)

Sum of squares for a contrast:

$$\text{SSC} = \frac{\widehat{l}^{\,2}}{\left(\frac{a_1^2}{n_1} + \frac{a_2^2}{n_2} + \cdots + \frac{a_t^2}{n_t}\right)}$$

Orthogonal contrasts allow partitioning of SSB (treatment sum of squares) into $t-1$ additive components:

Source	df	SS	MS	test stat
trt	$t-1$	SSB	$\left(\frac{\text{SSB}}{t-1}\right)$	f
contrast 1	1	SSC_1	SSC_1	f_1
contrast 2	1	SSC_2	SSC_2	f_2
$\vdots$				$\vdots$
contrast $t-1$	1	SSC_{t-1}	SSC_{t-1}	f_{t-1}
error	$n_\text{T}-t$	SSW	$\left(\frac{\text{SSW}}{n_\text{T}-t}\right)$	

ex. 5 weed control treatments: control, biol control 1, biol control 2, chem control 1, chem control 2

orthogonal contrasts of possible interest:

control	biol 1	biol 2	chem 1	chem 2
4	− 1	− 1	− 1	− 1
0	1	1	− 1	− 1
0	1	− 1	0	0
0	0	0	1	− 1

With orthogonal contrasts, not only does the SSB become partitioned into components (in 1-way AOV), but also, the SSC_1, SSC_2, ..., SSC_{t-1} are independent random variables.

However, the f statistics $\left(f_i = \frac{\text{SSC}_i}{\text{SSW}/(n_\text{T}-t)}\right)$ are *not* independent independent.

Experimentwise error rate (EER) control

Bonferroni approach

A_1, A_2, ..., A_m are events; $P(A_i) = \alpha$

Bonferroni inequality:

$$P(\overline{A}_1 \text{ and } \overline{A}_2 \text{ and } \cdots \text{ and } \overline{A}_m) \geq 1 - m\alpha$$

Choose α_{E} (say, 0.05)

m hypothesis tests: use $\alpha_{\mathrm{I}} = \frac{\alpha_{\mathrm{E}}}{m}$ for each individual test

m confidence intervals: calculate each CI at $100\left(1 - \frac{\alpha_{\mathrm{E}}}{m}\right)\%$ confidence

Bonferroni:

- fine for small number of comparisons
- very conservative (individual tests not powerful, CIs too big)