§10.2: Multiple Comparisons in ANOVA

After doing an ANOVA F-test on MS_F / MS_E and finding a significant p-value, you still need to decide which factor distributions $X^{(f)}$ are different.

Idea: Do t-Test on each pair of factors

Problems: (1) k factors requires $\binom{k}{2} = \frac{k^2-k}{2}$ tests...

(2) "α-Inflation":

More t-Tests ⟹ Higher P(Type I Error)

$$P\left(\text{at least one Type I Error in } n \text{ tests}\right) = 1-(1-\alpha)^n$$

(Assuming all tests are independent)

Example: Do t-Tests with $\alpha = 0.01$ for k factors

k=3: 3 pairs to test, P(Type I Error) ≈ 0.03

k=4: 6 pairs to test, P(Type I Error) ≈ 0.06

k=5: 10 pairs to test, P(Type I Error) ≈ 0.10

k=6: 15 pairs to test, P(Type I Error) ≈ 0.14

k=7: 21 pairs to test, P(Type I Error) ≈ 0.19

Fix: (1) Instead of testing each pair, compute a single "composite" Confidence Interval radius.

$$q_{\alpha, k, N-k} \cdot s/\sqrt{n}$$

→ If one factor's mean is outside the CI of another, conclude they are different.

(2) Use modified t-Test which accounts for #factors

"Studentized Range Distribution" ("Tukey Distribution")

- X sampled N times, split into k factors $X^{(f)}$
- n = #samples of each $X^{(f)}$
- s^2 = pooled sample variance of factors $X^{(f)}$

Then the difference between the largest & smallest factor sample mean ("range of factor means") is

$$\frac{\bar{x}^{(max)} - \bar{x}^{(min)}}{s/\sqrt{n}} \sim Q(k, N-k)$$

Note: k NOT (k-1) !!

In R the Q distribution is "tukey"

Ex ptukey(3, 5, 40) ← "$P(\bar{x}^{(1)} - \bar{x}^{(2)} \leq 3)$"

qtukey(.05, 5, 40) ← "Difference between means w/ Prob < .05"

This can be used to decide if two factor sample means come from different distributions:

Given a factor sample mean $\bar{x}^{(f)}$, the $(1-\alpha)$ CI for other factor sample means in the same distribution is

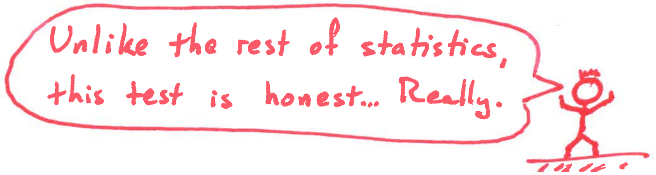
$$\bar{x}^{(f)} \pm \text{qtukey}(1-\alpha,\ k,\ N-k) \cdot \sqrt{MS_E / n}$$

Recall: "Residual Mean Square", MS_E, is the pooled factor sample variance.
So $\sqrt{MS_E}$ is "pooled sample std. dev."

This method is known by various names:

- "The Tukey Method"
- "Tukey's Range Test"
- "Tukey's Honest Significant Difference (HSD)"
- "Tukey's Honest Significance Test"

The emphasis on "honesty" here is worrisome...



Example: Suppose X has 5 factors with means

$\bar{x}^{(1)} = 14.5$ $\quad \bar{x}^{(2)} = 13.8$ $\quad \bar{x}^{(3)} = 13.3$

$\bar{x}^{(4)} = 14.3$ $\quad \bar{x}^{(5)} = 13.1$

ANOVA yields table

	df	SS	MS	F-value	p-value
Factors	4	13.32	3.33	37.84	0.00
Error	40	3.53	.088		
Total	44	16.85			

Apply Tukey's Range Test with $\alpha = 0.05$

There are $44+1 = 45$ samples with $4+1 = 5$ factors.

Assume $n = \frac{45}{5} = 9$ samples per factor.

The cutoff for significant difference between factor means is

$$\text{qtukey}(1-.05,\ 5,\ 40)\ \sqrt{.088/9} \approx 0.40$$

→ Factors with means more than 0.40 apart are "significantly different".

Order factor means by size & group them by CI.

$\bar{x}^{(5)}$	$\bar{x}^{(3)}$	$\bar{x}^{(2)}$	$\bar{x}^{(4)}$	$\bar{x}^{(1)}$
13.1	13.3	13.8	14.3	14.5
(12.7, 13.5)	(12.9, 13.7)	(13.4, 14.2)	(13.9, 14.7)	(14.1, 14.9)

Groups: {13.1, 13.3} (each inside the other's CI), {13.8}, {14.3, 14.5} (each inside the other's CI).

Note: Two factors having means grouped together by Tukey does not imply that they are from same distribution... it only implies that they are "not significantly different" — maybe the difference would become significant if n was larger...

[Tukey's Range Test identifies factors which are "significantly different" — it does not identify factors which are identical.]

Short note from §10.3 "non-uniform sampling"

If $X^{(f)}$ is sampled n_f times
(and not all n_f are the same)
then we cannot use "n" when computing $(1-\alpha)$ CI:

$$\bar{x}^{(f)} \pm qtukey(1-\alpha, k, N-k) \cdot \sqrt{MSE/n}$$

Instead you must do individual confidence intervals for each pair of factor means $\bar{x}^{(f_1)}, \bar{x}^{(f_2)}$

If $X^{(f_1)}$ is sampled n_1 times and $X^{(f_2)}$ is sampled n_2 times then the $(1-\alpha)$ CI for $\bar{x}^{(f_2)}$ around $\bar{x}^{(f_1)}$ is

$$\bar{x}^{(f_2)} = \bar{x}^{(f_1)} \pm qtukey(1-\alpha, k, N-k) \cdot \sqrt{MSE/n}$$

where $\frac{1}{n} = \frac{1}{2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$

(like a simplified version of how degrees of freedom combined in two-sample t-Test.)