$$1-\alpha = P\left(\max_{1\le i_1<i_2\le I} \frac{|(\overline{Y}_{i_1\cdot}-\overline{Y}_{i_2\cdot})-\Delta_{i_1,i_2}|}{s_p/\sqrt{J}} \le q_{I,\,I(J-1)}(\alpha)\right)$$

can test $H_0: \Delta_{i_1,i_2}=$ a constant

if set $\Delta_{i_1,i_2}=0 \Rightarrow$ Tukey's test

if data is fixed

$$= P\left(|(\overline{Y}_{i_1\cdot}-\overline{Y}_{i_2\cdot})-\Delta_{i_1,i_2}| < q_{I,I(J-1)}(\alpha)\,(s_p/\sqrt{J}),\ \text{for any } i_1<i_2\right)$$

– A set of $100(1-\alpha)\%$ simultaneous confidence intervals for all differences Δ_{i_1,i_2}'s is

$$(\overline{Y}_{i_1\cdot}-\overline{Y}_{i_2\cdot}) \pm q_{I,I(J-1)}(\alpha)\,(s_p/\sqrt{J})$$

12/11

Example 4 (Chlorpheniramine maleate in tablet, Tukey's method, cont. Ex.1, LNp.3)

- $s_p = 0.06$ and

$\sqrt{0.0037}$ (LNp.18)

order statistics of $\overline{Y}_{1\cdot},\cdots,\overline{Y}_{I\cdot}$

Lab	4	6	5	2	7	3	1
Mean	3.920	3.955	3.957	3.997	3.998	4.003	4.062
Tukey's test							
t-test using $t_{I(J-1)}(\alpha/2)$							

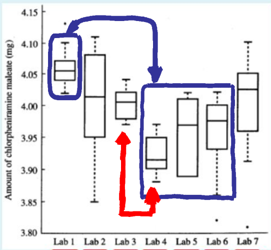
$-0.082 = 3.98$

$\alpha=0.05$

reject if $|\overline{Y}_{i_1\cdot}-\overline{Y}_{i_2\cdot}|>0.053$

- Tukey's test
  - two parameters are $I=7$ and $I(J-1)=63$
    - $q_{7,63}(0.05) = 4.34$ (Table 6, Appendix B, textbook)
    - $q_{7,63}(0.05)(s_p/\sqrt{J}) = 0.082$
  - Conclusions:

irrelevant to i_1, i_2



- Mean from Lab 1 is significantly different from those of Labs 4, 5, 6
- Mean from Lab 3 is significantly different from that of Lab 4
- No other comparisons are significant at 0.05 level

$H_0^{(1,2)}: \mu_1=\mu_2$ not rejected

$H_0^{(2,5)}: \mu_2=\mu_5$ not rejected

Note. These conclusions are not deductive (e.g., Labs 1 and 2 are not significantly different, Labs 2 and 5 are not significantly differnet $\Rightarrow$ cannot deduct that Labs 1 and 5 are not significantly different,

Check LN. CH11, p.14

Q: Why cannot? How to interpret "H_0 not rejected"? Really accept H_0 is true?

Thm10 (LNp.29)

- For t-tests using the critical value $t_{I(J-1)}(\alpha/2)$, two labs would have been declared significantly different if their means has differed by more than

$$t_{63}(0.025)(s_p\sqrt{2/J}) = 0.053 \le 0.082.$$

easier to reject if significance level not adjusted.

no need to assume equal sample size

Theorem 12 (multiple comparisons, Bonferroni method, I-sample normal model)

Consider the model $(*)$ in LNp.5.

Note. do not need any assumption when derive this result.

- **Recall.** $\alpha' \le \alpha^* \le K\alpha'$ (LNp.30)

individual type I error rate

overall type I error rate

- If we set $K\alpha' = \alpha \Leftrightarrow \alpha' = \alpha/K$, then $\alpha^* \le \alpha$.

desired type I error rate

- A desired overall type-I error rate at most α can be guaranteed by testing each null $H_0^{(i_1,i_2)}$ at significance level $\alpha' = \alpha/K = \alpha/\binom{I}{2}$., i.e.,

irrelevant to i_1, i_2

$\because$ 2-sided

$$RR_{i_1,i_2} = \left\{\, |T_{i_1,i_2}| > t_{N-I}\left(\alpha/(2\times K) = \alpha/[2\times\tbinom{I}{2}] = \alpha/[I(I-1)]\right) \right\}$$

- Equivalently, a set of simultaneous $100(1-\alpha)\%$ confidence intervals for all differences Δ_{i_1,i_2}'s is $\leftarrow \binom{I}{2}$ parameters

Tukey's CI_{i_1,i_2}'s in LNp.34 $\overset{cf.}{\longleftrightarrow}$
$$CI_{i_1,i_2} = (\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}) \pm \left(s_p\sqrt{\frac{1}{J_{i_1}} + \frac{1}{J_{i_2}}}\right) \times t_{N-I}\,(\alpha/[I(I-1)])$$

Example 5 (Maleate in tablet, Bonferroni method, cont. Ex.1, LNp.3)

- $K = \binom{7}{2} = 21$ pairwise comparisons among the 7 labs
- A set of 95% simultaneous confidence intervals for the 21 pairwise comparison is $\alpha = 0.05$ Δ_{i_1,i_2}'s
$$(\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}) \pm t_{63}\left(\frac{0.025}{21}\right)(s_p\sqrt{2/J}) = CI_{i_1,i_2}$$
- Two labs would have been declared significantly different if their means has differed by more than $\leftarrow (\Delta_{i_1,i_2} =)\, 0 \notin CI_{i_1,i_2}$
Bonferroni tends not to reject than Tukey
$$t_{63}\left(\frac{0.025}{21}\right)(s_p\sqrt{2/J}) = 0.085 \geq 0.082 \quad (\approx \text{Tukey's method in this case!})$$
- We only conclude that Lab 1 produced significantly higher measurements than Labs 4, 5, 6. $\overset{cf.}{\longleftrightarrow}$ Tukey's results in Ex.4 (LNp.34)

Note 7 (Some notes about multiple comparisons)

- A comparison of Tukey and Bonferroni methods studentized range dist.
  - Tukey's method requires the independent assumption (check (3) in LNp.31) while Bonferroni method does not.

  - Bonferroni method is simple and versatile and, although crude, gives surprisingly good results if K is not too large. ∵ require no assumption (check LNp.30)
  - When K is large, Bonferroni method is conservative, i.e., tends not to declare the treatment means are significantly different. It happens when
$$\alpha' \leqslant \alpha^* \ll K\alpha' = \alpha.$$
e.g. $\alpha = 0.05$, $K = 100$ $\Rightarrow \alpha' = 5\times 10^{-4}$
- In general, for the critical values of the three methods, we have
$$\text{not adjusted } T < \text{Tukey} < \text{Bonferroni (adjusted } T)$$
too sensitive ← not adjusted T; adjusted T → conservative
- When the sample sizes J_i's are different, the Tukey's procedure in Thm 11 (LNp.31-34) can be similarly applied. A modified procedure is called Tukey-Kramer procedure.
- Under the nonparametric model in LNp.23, Bonferroni method can be applied to all $K = \binom{I}{2}$ pairwise comparisons tested by Mann-Whitney tests. cf. not adjusted T (under normality)

If Kruskal-Wallis test is rejected

- See Muller (1981) for other parametric and nonparametric multiple comparison procedures.

❖ **Reading**: textbook, 12.1, 12.2 (12.2.1, 12.2.2, 12.2.3)

- Two-way layout → two factors $\overset{cf.}{\longleftrightarrow}$ one factor ← one-way layout

- A **two-way layout** is an experimental design involving two factors, denoted by A and B, each at two or more levels (e,g, the factor A might be various drugs and the other factor B might be gender). ← cf. → one-way layout
- Suppose that A has I levels and B has J levels
  $\Rightarrow$ total $I \times J$ level combinations, i.e.,

$$(A, B) = (i, j), \quad 1 \leqslant i \leqslant I, 1 \leqslant j \leqslant J$$

  $\Rightarrow$ $I \times J$ samples, each of sample size K_{ij}, $1 \leqslant i \leqslant I$, $1 \leqslant j \leqslant J$.

  └ $I \times J$ populations cf. → J_i in one-way layout

- **Observed data**:
  – let y_{ijk} be the kth observed value of the (i, j)th sample, where

$$i \leftarrow A, \quad j \leftarrow B, \quad k: \text{ replicates}$$

one-way layout (LNp.1) cf. → ⊖ y_{ijk}'s are *continuous* measurements and $y_{ijk} - y_{i'j'k'}$ is *meaningful*

Data

A	B	V
1	1	Y_{111}
⋮	⋮	⋮
1	1	$Y_{11K_{11}}$
⋮	⋮	⋮
1	J	Y_{1J1}
⋮	⋮	⋮
1	J	$Y_{1JK_{1J}}$
2	1	Y_{211}
⋮	⋮	⋮
2	1	$Y_{21K_{21}}$
⋮	⋮	⋮
2	J	Y_{2J1}
⋮	⋮	⋮
2	J	$Y_{2JK_{2J}}$
⋮	⋮	⋮
I	J	Y_{IJ1}
⋮	⋮	⋮
I	J	$Y_{IJK_{IJ}}$

- Statistical modeling: regard y_{ijk} as a realization of a random variable Y_{ijk}

$\boxed{(IJ)\text{-sample model}}$ $\Leftarrow$ (i, j)th sample: $Y_{ij1}, Y_{ij2}, \ldots, Y_{ijK_{ij}} \sim$ i.i.d. F_{ij} $\quad (\triangledown)$

for $i = 1, \ldots, I$, $j = 1, \ldots, J$, and all Y_{ijk}'s are independent. ← ∵ simple random sampling

– Denote the means (or medians) of $F_{11}, \ldots, F_{IJ}$ by $\mu_{11}, \ldots, \mu_{IJ}$.

unknown ─ $F_{11}, \ldots, F_{IJ}$; $\mu_{11}, \ldots, \mu_{IJ}$ ─ parameters

– Data table

crossing structure → cf. ↕ nesting

A	B: 1	2	⋯	J
1	$(A, B) = (1, 1)$, $Y_{111}, \ldots, Y_{11K_{11}}$ $\overset{\text{i.i.d.}}{\sim} F_{11}\ (\mu_{11})$	$(A, B) = (1, 2)$, $Y_{121}, \ldots, Y_{12K_{12}}$ $\overset{\text{i.i.d.}}{\sim} F_{12}\ (\mu_{12})$	⋯	$(A, B) = (1, J)$, $Y_{1J1}, \ldots, Y_{1JK_{1J}}$ $\overset{\text{i.i.d.}}{\sim} F_{1J}\ (\mu_{1J})$
⋮	⋮	⋮	⋮	⋮
I	$(A, B) = (I, 1)$, $Y_{I11}, \ldots, Y_{I1K_{I1}}$ $\overset{\text{i.i.d.}}{\sim} F_{I1}\ (\mu_{I1})$	$(A, B) = (I, 2)$, $Y_{I21}, \ldots, Y_{I2K_{I2}}$ $\overset{\text{i.i.d.}}{\sim} F_{I2}\ (\mu_{I2})$	⋯	$(A, B) = (I, J)$, $Y_{IJ1}, \ldots, Y_{IJK_{IJ}}$ $\overset{\text{i.i.d.}}{\sim} F_{IJ}\ (\mu_{IJ})$

a population distribution F_{ij} ↑ a population → a cell, a sample

Total: $I \times J$ cells ↑

Question 8.

- In one-way ANOVA of I samples, we are primarily concerned with whether the means (or medians) of the I samples are equal or not.
- In two-way case, is it enough to only examine whether the means (or medians) of the IJ samples are equal or not? → i.e., treat it as a one-way layout of IJ samples

In two-way case, we might be intereted in knowing

different types of "systematic difference"

- whether the level changes of factor A (or factor B) cause a systematic difference in μ_{ij}'s? $\Rightarrow$ main effects of A (or B)
- whether the systematic difference caused by the level changes of factor A depends on the level changes of factor B? $\Rightarrow$ Interactions of A and B

Q: How to define these effects?

B=male A=new drug A=old drug ↕ B=female A=new drug A=old drug ↕