Example 3 (Chlorpheniramine maleate, Kruskal-Wallis test, cont. Ex.1, LNp.3)

Ex.2 (LNp.18) cf.

- Statistical modeling of data: assume the nonparametric model in LNp.23.
- $I = 7$ and $J_i = 10 \Rightarrow$ can use χ^2_{I-1} to approximate null distribution of K
- Because $K = 29.51$ and $\chi^2_6(0.005) = 18.55 \Rightarrow p$-value < 0.005, the nonparametric analysis, too, indicates a systematic difference among the labs.

Note 6 (Some notes about the Kruskal-Wallis Test)

- Connection between the Mann-Whitney test (2 samples) and the Kruskal-Wallis test (I samples):
  - test statistic (**recall**: Thm 9, LN, CH11, p.31-33)

$$1+2+\cdots+(m+n) = \frac{(m+n)(m+n+1)}{2}$$

J_1, Y_{1j}, J_2, Y_{2j}, R_{1j}, R_{2j}

$$n\ X_i\text{'s},\ m\ Y_j\text{'s} \Rightarrow R_1,\ldots,R_n,\ R_{n+1},\ldots,R_{m+n} \Rightarrow \underbrace{R_1+\cdots+R_n}_{=W_X} \text{ and } \underbrace{R_{n+1}+\cdots+R_{m+n}}_{=W_Y}$$

$N=m+n$

$$* \ \overline{R}_{1.} = W_X/n,\ \overline{R}_{2.} = W_Y/m, \text{ and } \overline{R}_{..} = \frac{n}{m+n}(W_X/n) + \frac{m}{m+n}(W_X/m)$$

$$* \ SS_B$$
$$= n\left[\frac{W_X}{n} - \left(\frac{n}{m+n}\frac{W_X}{n} + \frac{m}{m+n}\frac{W_Y}{m}\right)\right]^2 + m\left[\frac{W_Y}{m} - \left(\frac{n}{m+n}\frac{W_X}{n} + \frac{m}{m+n}\frac{W_Y}{m}\right)\right]^2$$
$$= n\frac{m^2}{(m+n)^2}\left(\frac{W_X}{n} - \frac{W_Y}{m}\right)^2 + m\frac{n^2}{(m+n)^2}\left(\frac{W_X}{n} - \frac{W_Y}{m}\right)^2 = \frac{mn}{m+n}\left(\frac{W_X}{n} - \frac{W_Y}{m}\right)^2$$
$$= \frac{mn}{m+n}\left[\frac{W_X}{n} - \frac{1}{m}\left(\frac{(m+n)(m+n+1)}{2} - W_X\right)\right]^2$$
$$= \frac{mn}{m+n}\left(\frac{m+n}{mn}W_X - \frac{m+n}{mn} \times \frac{n(m+n+1)}{2}\right)^2 = \frac{m+n}{mn}\left(W_X - \frac{n(m+n+1)}{2}\right)^2$$

symmetric about (LN, CH11, p.33) → $E(W_X)$ under H_0 (Thm12, LN, CH11, p.40)

Reject H_0 if W_X large or small $\Leftrightarrow K > w$

$$* \ K = \frac{12}{\underbrace{(m+n)(m+n+1)}_{N}} SS_B = \left(\frac{W_X - \frac{n(m+n+1)}{2}}{\sqrt{\frac{mn(m+n+1)}{12}}}\right)^2$$

$\overset{D}{\approx} \chi^2_1$ (check Thm 12, LN, CH11, p.40)

$E(W_X)$ under H_0; $Var(W_X)$ under H_0; $\overset{D}{\approx} N(0,1)$

- K and W_X have equivalent exact null distribution; for asymptotic null distribution:

The null distribution of W_X is symmetric about $E(W_X)$

TBp.192 → if $Z \sim N(0,1)$, then $Z^2 \sim \chi^2_1$ (check Thm 13, LN, CH11, p.41)

- the Mann-Whitney test is a special case of the Kruskal-Wallis test where only two groups are being compared ($I = 2$)

cf. Note8 in LN, CH11, p.39

- Advantages of Kruskal-Wallis test
  - it makes no assumption of normality, thus has a wider range of applicability than does the F-test
  - it is especially useful in small-sample situations
  - outliers have less influence on this nonparametric test than on F-test
  - in some applications, the data consist of ranks, e.g., wine tasting

- The problem of multiple comparisons ← This question does not arise in 2-sample case

Question 6.

What is the next question we should ask if the null $H_0: \mu_1 = \cdots = \mu_I$ of ANOVA analysis was rejected? ⇒ μ_i's significantly different → How differ?

- When H_0 was rejected, we concluded that the means are not all equal. But, the ANOVA test gives no information about **how they differ**, e.g., we do not know whether all treatment means are different from each other, or just a few of them are.
- In many applications, real interest may be focused on comparing **pairs** or groups of treatments, and estimating the treatment means and differences.

compare minimum vs. maximum $\Uparrow$

Definition 3 (Multiple comparisons based on pairwise comparison)

Consider the model ($\otimes$) in LNp.1, and assume that μ_1 $(= \overline{\mu} + \alpha_1)$, $\ldots$, μ_I $(= \overline{\mu} + \alpha_I)$ are the means (or medians) of $F_1, \ldots, F_I$, respectively.

$(i_1, i_2) = (3, 4)$

$\overline{Y}_{4\cdot}\ \overline{Y}_{1\cdot}\ \overline{Y}_{2\cdot}\ \overline{Y}_{3\cdot}$

- Comparison of two treatments: For a specific pair (i_1, i_2), $1 \leqslant i_1 < i_2 \leqslant I$, where the choice of this pair does not depend on the data (i.e., i_1 and i_2 are not random variables, priori planned $\overset{\text{cf.}}{\longleftrightarrow}$ post-hoc), test

$\overline{Y}_{3\cdot}\ \overline{Y}_{1\cdot}\ \overline{Y}_{4\cdot}\ \overline{Y}_{2\cdot}$

$(i_1, i_2) = (2, 3)$

$$H_0^{(i_1,i_2)} : \mu_{i_1} = \mu_{i_2} \ (\text{or } \alpha_{i_1} = \alpha_{i_2}) \quad \text{vs.} \quad H_A^{(i_1,i_2)} : \mu_{i_1} \neq \mu_{i_2} \ (\text{or } \alpha_{i_1} \neq \alpha_{i_2})$$

- Multiple comparisons: Simutaneously test these pairwise comparisons

$\binom{I}{2}$ tests

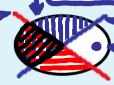
$$H_0^{(i_1,i_2)} : \mu_{i_1} = \mu_{i_2} \ (\text{or } \alpha_{i_1} = \alpha_{i_2}) \quad \text{vs.} \quad H_A^{(i_1,i_2)} : \mu_{i_1} \neq \mu_{i_2} \ (\text{or } \alpha_{i_1} \neq \alpha_{i_2})$$

for *all* $(= \binom{I}{2})$ pairs (i_1, i_2)'s of treatment means. all models $\Omega = H_0 \cup H_A$ $H_0^{(i_1,i_2)}$

- Note that

$$\underbrace{H_0}_{H_0 \text{ of ANOVA}} = \bigcap_{1 \leqslant i_1 < i_2 \leqslant I} H_0^{(i_1,i_2)} \quad \text{and} \quad \underbrace{H_A}_{H_A \text{ of ANOVA}} = \bigcup_{1 \leqslant i_1 < i_2 \leqslant I} H_A^{(i_1,i_2)}$$

$H_A = H_0^c$ $\quad (H_0^{(i_1,i_2)})^c$



$H_0^{(i_1,i_2)}$ a model

Theorem 10 (t-test for comparing two priori-planned treatments, normal model)

Consider the model ($*$) in LNp.5, and define $\rightarrow$ (i_1, i_2) fixed in advance

assume normality

$$\Delta_{i_1,i_2} = \mu_{i_1} - \mu_{i_2} = \alpha_{i_1} - \alpha_{i_2}. \quad \Rightarrow H_0^{(i_1,i_2)}: \Delta_{i_1,i_2} = 0$$

- Since $\overline{Y}_{i_1\cdot} \sim N(\mu_{i_1}, \sigma^2/J_{i_1})$, $\overline{Y}_{i_2\cdot} \sim N(\mu_{i_2}, \sigma^2/J_{i_2})$, and $\overline{Y}_{i_1\cdot}$ and $\overline{Y}_{i_2\cdot}$ are independent, we have $\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot} \sim N(\Delta_{i_1,i_2}, \sigma^2(1/J_{i_1} + 1/J_{i_2}))$. LNp.16

LNp.20

- Since $(N-I)\, s_p^2 = SS_W \sim \sigma^2 \chi^2_{N-I}$, and $\overline{Y}_{i_1\cdot}$, $\overline{Y}_{i_2\cdot}$, and s_p^2 are independent, we have $\sigma^2 \leftarrow e$ Under $\Omega = H_0 \cup H_A$

$\rightarrow \sim N(0,1)$

$$Q_{i_1,i_2,\Delta} \equiv \frac{(\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}) - \Delta_{i_1,i_2}}{s_p\sqrt{\frac{1}{J_{i_1}} + \frac{1}{J_{i_2}}}} = \frac{\left[(\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}) - \Delta_{i_1,i_2}\right] \Big/ \sqrt{\sigma^2\left(\frac{1}{J_{i_1}} + \frac{1}{J_{i_2}}\right)}}{\sqrt{\frac{(N-I)\, s_p^2}{\sigma^2} \times \frac{1}{N-I}}}$$

indep.

$$\sim t_{N-I} \qquad \chi^2_{N-I} \sim$$

pivotal quantity of Δ_{i_1,i_2}

($\overset{\text{cf.}}{\longleftrightarrow}$ pivotal quantity for 2-independent samples, LN, CH11, p.11)

- Test statistic for $H_0^{(i_1,i_2)}$: $\mu_{i_1} = \mu_{i_2}$: $T_{i_1,i_2} \equiv \dfrac{\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}}{s_p\sqrt{\frac{1}{J_{i_1}} + \frac{1}{J_{i_2}}}}$

$\Delta_{i_1,i_2} = 0$

- Null distribution of T_{i_1,i_2}: Under $H_0^{(i_1,i_2)}$, $T_{i_1,i_2} \sim t_{N-I}$

2-sided test

cf.

- Rejection region at significance level α: Reject $H_0^{(i_1,i_2)}$ if $|T_{i_1,i_2}| > t_{N-I}(\alpha/2)$

Question 7. → multiple testing

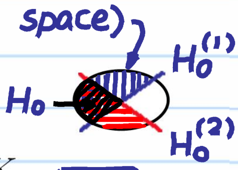
Can we use these $\binom{I}{2}$ level-α rejection regions

$$RR_{i_1,i_2} \equiv \{\text{ data } \mid \underbrace{|T_{i_1,i_2}|}_{Y_{i,j}\text{'s}} > t_{N-I}(\alpha/2)\},$$

$1 \leqslant i_1 < i_2 \leqslant I$, to simutaneously test all pairs of treatment means?

- Difficulty with this approach: If there are multiple comparisons, each at significance level α, the probability of the *overall* type-I error across all comparisons will be inflated, i.e., larger than α.
  - Suppose that there are K null and alternative hypotheses:

Ω (parameter space)



$$H_0^{(1)} \text{ vs. } H_A^{(1)}, \quad \ldots, \quad H_0^{(K)} \text{ vs. } H_A^{(K)}$$

$H_0^{(1)} \cup H_A^{(1)} = \Omega$
$\vdots$
$H_0^{(K)} \cup H_A^{(K)} = \Omega$

Let RR_k be a level-α' rejection region of $H_0^{(k)}$, $k = 1, \ldots, K$.

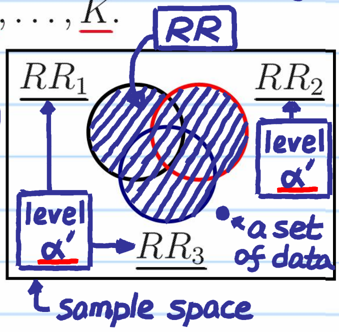
$P(RR_k|H_0^{(k)}) = \alpha'$

  - Consider the null and alternative hypotheses:

Check Def 3 (LNp.28)

$$\underbrace{H_0 : \cap_{k=1}^K H_0^{(k)}}_{\text{all } H_0^{(k)} \text{ are true}} \quad \text{vs.} \quad H_A : \cup_{k=1}^K H_A^{(k)} \leftarrow \text{at least one } H_A^{(k)} \text{ is true}$$

and the rejection region $RR = \cup_{k=1}^K RR_k$.



  - Let α^* be the probability of overall type-I error. Then,

$$\alpha' \leqslant \alpha^* = P(RR|H_0) = P(\cup_{k=1}^K RR_k|H_0) \leqslant \sum_{k=1}^K P(RR_k|H_0^{(k)}) = K\alpha'$$

$\leqslant \sum_{k=1}^K P(RR_k|H_0) \leqslant$

if $\alpha' = 0.05$, $\alpha^* \geq 0.05$

  - For example,
    * in ANOVA the null $H_0 : \mu_1 = \cdots = \mu_I$ is equivalent to the $\cap_{1 \leqslant i_1 < i_2 \leqslant I} H_0^{(i_1,i_2)}$ in multiple comparisons. $\mu_{i_1} = \mu_{i_2}$
    * if the significance levels for ANOVA and for each T_{i_1,i_2} in multiple comparisons are chosen to be the same α, it will be more often to claim "at least one of μ_i's not equal" in multiple comparisons than in ANOVA $\Rightarrow$ need to adjust the significance levels of T_{i_1,i_2}'s (**Q**: How?)

$\because \alpha^* > \alpha$

- By the duality of confidence intervals and hypothesis tests, at significance level α, finding rejection regions RR_{i_1,i_2}'s such that

Δ_{i_1,i_2}'s
$\cap$
$AR = RR^c \leftarrow RR$

$$P(\cup_{1 \leqslant i_1 < i_2 \leqslant I} RR_{i_1,i_2} | H_0) \leqslant \alpha \quad \text{— } (*)$$

is equivalent to finding confidence intervals CI_{i_1,i_2}'s for Δ_{i_1,i_2}'s such that

$\|$

$\cap_{i_1,i_2} RR_{i_1,i_2}^c = \cap_{i_1,i_2} AR_{i_1,i_2}$ → $P(\Delta_{i_1,i_2} \in CI_{i_1,i_2}$ for any (i_1, i_2)'s$) \geqslant 1 - \alpha$.

Theorem 11 (multiple comparisons, Tukey's method, *I*-sample normal model)

Consider the model $(*)$ in LNp.5. For simplicity, assume

assume normality

$$J_1 = \cdots = J_I = J \quad \Rightarrow \quad N = IJ. \leftarrow \text{\# of all obs.}$$

- **Recall.** (1) $\overline{Y}_{i\cdot} \sim N(\mu_i, \sigma^2/J)$, $i = 1, \ldots, I$, (2) $[I(J-1)]\, s_p^2 \sim \sigma^2 \chi^2_{I(J-1)}$, (3) $\overline{Y}_{1\cdot}, \ldots, \overline{Y}_{I\cdot}, s_p^2$ are independent, (4) $T_{i_1,i_2} \overset{\text{LNp.29}}{=} (\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot})/(\sqrt{2}\, s_p/\sqrt{J})$

- Simultaneous rejection regions RR_{i_1,i_2}'s ($RR_{i_1,i_2} = \{\text{data} \mid |T_{i_1,i_2}| > c^*\}$)
  - **Note.** Let $\overline{Y}_{(1)}, \ldots, \overline{Y}_{(I)}$ be the order statistic of $\overline{Y}_{1\cdot}, \ldots, \overline{Y}_{I\cdot}$. Then,

$$|\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}| \leqslant \overline{Y}_{(I)} - \overline{Y}_{(1)} = \max_{1\leqslant i_1<i_2\leqslant I} |\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}|$$

for any $1 \leqslant i_1 < i_2 \leqslant I$.

$\overline{Y}_{4\cdot}\ \overline{Y}_{3\cdot}\ \overline{Y}_{1\cdot}\ \overline{Y}_{5\cdot}\ \overline{Y}_{2\cdot}$

$\overline{Y}_{(1)}\ \overline{Y}_{(2)}\ \overline{Y}_{(3)}\ \overline{Y}_{(4)}\ \overline{Y}_{(5)}$

range $\overline{Y}_{(5)} - \overline{Y}_{(1)}$

  - Under $H_0: \mu_1 = \cdots = \mu_I = \overline{\mu} = \bigcap_{1\leqslant i_1<i_2\leqslant I} H_0^{(i_1,i_2)}$, $\mu_{i_1} = \mu_{i_2}$

Check LNp.31 (*)

$$\alpha = P\Big(\bigcup_{1\leqslant i_1<i_2\leqslant I} \underbrace{\{|T_{i_1,i_2}| > c^*\}}_{RR_{i_1,i_2}}\Big) = P(\text{at least one of } |T_{i_1,i_2}|\text{'s} > c^*)$$

$Z_1 = \dfrac{\overline{Y}_{1\cdot} - \overline{\mu}}{\sigma/\sqrt{J}} \sim N(0,1)$, $\vdots$, $Z_I = \dfrac{\overline{Y}_{I\cdot} - \overline{\mu}}{\sigma/\sqrt{J}} \sim N(0,1)$

$$= P(\max_{1\leqslant i_1<i_2\leqslant I} |T_{i_1,i_2}| > c^*) = P\left(\max_{1\leqslant i_1<i_2\leqslant I} \frac{|\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}|}{s_p/\sqrt{J}} > \sqrt{2}\,c^*\right)$$

$$= P\left(\frac{\overline{Y}_{(I)} - \overline{Y}_{(1)}}{s_p/\sqrt{J}} > \sqrt{2}\,c^*\right) = P\left(\frac{\overbrace{\frac{\overline{Y}_{(I)}-\overline{\mu}}{\sigma/\sqrt{J}} - \frac{\overline{Y}_{(1)}-\overline{\mu}}{\sigma/\sqrt{J}}}^{\boxed{Z_{(I)} - Z_{(1)}}}}{\sqrt{\underbrace{\frac{[I(J-1)]\,s_p^2}{\sigma^2}}_{\sim \chi^2_{I(J-1)}} \times \frac{1}{I(J-1)}}} > \sqrt{2}\,c^*\right) \quad \boxed{\text{indep.}}$$

$$= P\big(SR_{I,I(J-1)} > q_{I,I(J-1)}(\alpha)\big) \quad \Rightarrow \quad c^* = q_{I,I(J-1)}(\alpha)/\sqrt{2},$$

where $q_{I,I(J-1)}(\alpha)$ is the $(1-\alpha)$-quantile of $SR_{I,I(J-1)}$. I $I(J-1)$

  - Studentized range (SR) distribution with parameters n and ν:
    * $Z_1, \ldots, Z_n \sim$ i.i.d. $N(0,1)$, and $\nu\,\hat{\sigma}^2 \sim \chi^2_\nu$ is independent of $Z_1, \ldots, Z_n$.

    * Let $Z_{(1)}, \ldots, Z_{(n)}$ be the order statistics of $Z_1, \ldots, Z_n$.
    * Then,

$$0 < \max_{1\leqslant i_1<i_2\leqslant n} \frac{|Z_{i_1} - Z_{i_2}|}{\sqrt{\underbrace{\nu\,\hat{\sigma}^2}_{\chi^2_\nu}/\nu}} = \frac{\overbrace{Z_{(n)} - Z_{(1)}}^{\text{range}}}{\hat{\sigma}} \sim SR_{n,\nu}$$

studen-dized   indep.

of Z_i's   the d.f. of $\hat{\sigma}^2$

How to derive its pdf or cdf?

  - Simultaneous rejection regions RR_{i_1,i_2}'s: Reject $H_0^{(i_1,i_2)}$ if

$$|T_{i_1,i_2}| > q_{I,I(J-1)}(\alpha)/\sqrt{2} \quad \Leftrightarrow \quad |\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}| > q_{I,I(J-1)}(\alpha)(s_p/\sqrt{J})$$

for any $1 \leqslant i_1 < i_2 \leqslant I$.

- irrelevant to i_1, i_2
- it is related to accuracy. it ↓ if S_p↓ or J↑

- Simultaneous confidence intervals CI_{i_1,i_2}'s
  - Because under $H_0 \cup H_A$,

pivotal quantity of Δ_{i_1,i_2} in LNp.29

$$\max_{1\leqslant i_1<i_2\leqslant I} \sqrt{2}\,|Q_{i_1,i_2,\Delta}| = \max_{1\leqslant i_1<i_2\leqslant I} \frac{|(\overline{Y}_{i_1\cdot} - \overline{Y}_{i_2\cdot}) - \Delta_{i_1,i_2}|}{s_p/\sqrt{J}} \quad = \mu_{i_1} - \mu_{i_2}$$

$$= \max_{1\leqslant i_1<i_2\leqslant I} \frac{\left|\underbrace{\frac{\overline{Y}_{i_1\cdot} - \mu_{i_1}}{\sigma/\sqrt{J}}}_{N(0,1)} - \underbrace{\frac{\overline{Y}_{i_2\cdot} - \mu_{i_2}}{\sigma/\sqrt{J}}}_{N(0,1)}\right|}{\sqrt{\underbrace{\frac{[I(J-1)]\,s_p^2}{\sigma^2}}_{\sim\chi^2_{I(J-1)}} \times \frac{1}{I(J-1)}}} \quad \boxed{\text{indep}} \quad \sim SR_{I,I(J-1)}$$

we have

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

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

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

– 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	$\boxed{4 \quad 6 \quad 5}$			2	7	3	$\boxed{1}$
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