STA 104 Applied Nonparametric Statistics

Chapter 4: One-Way Layout Problems: Nonparametric One-Way Analysis of Variance

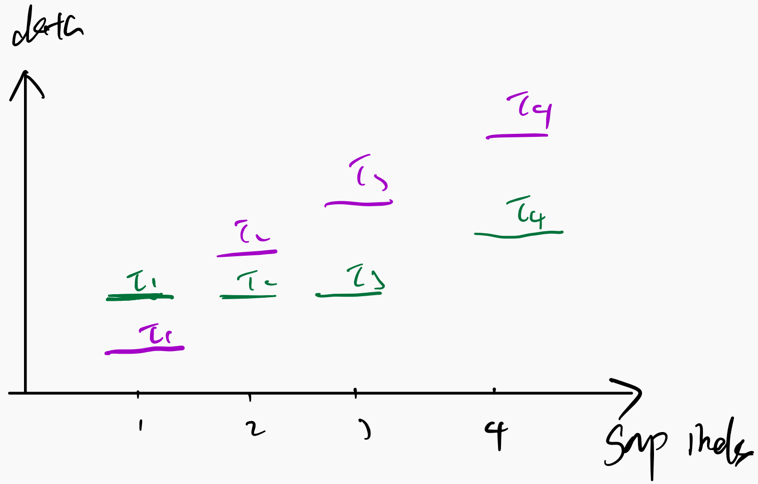
Xiner Zhou

Department of Statistics, University of California, Davis

Table of contents

One-Sided All-Treatments Multiple Comparisons for Ordered Treatment Effects Alternatives

After rejection of



$$H_0 : \underbrace{\tau_1 = \ldots = \tau_k}_{F_1 = F_2 \ldots = F_k \equiv F}$$

$$H_1 : \tau_1 \leq \tau_2 \leq \cdots \leq \tau_k$$ with at least one of $\leq$ is $<$

with the Jonckheere-Terpstra test, it is important to reach conclusions about exactly which $\leq$ is $<$ as opposed to $=$, that is, whether there is strict ordering among all $\binom{k}{2} = k(k-1)/2$ individual differences between pairs of treatment effects (τ_i, τ_j), for $i < j$, and these conclusions are naturally one-sided, in accordance with the ordered alternatives setting.

Hypothesis

$$\left\{\begin{array}{ll} H_0 : \tau_1 = \tau_2 & H_1 : \tau_1 < \tau_2 \\ H_0 : \tau_1 = \tau_3 & H_1 : \tau_1 < \tau_3 \\ \dots & \\ H_0 : \tau_{k-1} = \tau_k & H_1 : \tau_{k-1} < \tau_k \end{array}\right\} \frac{k(k-1)}{2} \text{ simultaneous tests/multiple}$$

comparisons

Motivation

To test for each: $\underbrace{H_0: \tau_i = \tau_j \quad H_1: \tau_i < \tau_j}_{\text{Two-sample problem: one-sided}}$:

$\Rightarrow$ For each pair of treatments (i, j), for $1 \leq i < j \leq k$, let

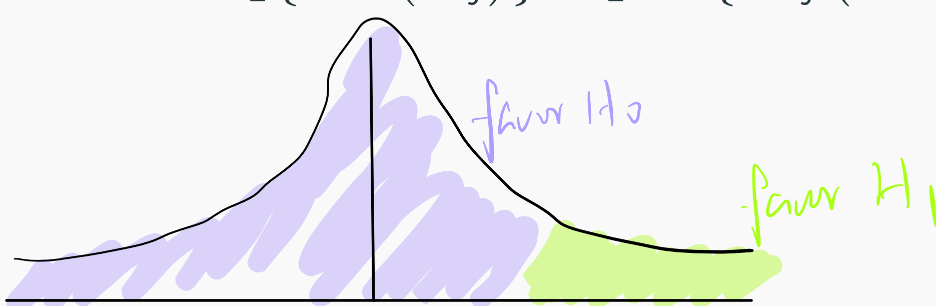
Wilcoxon rank sum test:

$$W_{ij} = \sum_{b=1}^{n_j} R_{jb}$$

where R_{jb} are the ranks of X_{jb} among the combined i th and j th samples; that is, W_{ij} is the Wilcoxon rank sum of the j th sample ranks in the joint two-sample ranking of the i th and j th sample observations.

$\Rightarrow$ standardized (under H_0) version of W_{ij} multiplied by $\sqrt{2}$

$$W_{ij}^{*} = \sqrt{2}\left[\frac{W_{ij} - E_0\left(W_{ij}\right)}{\left\{\operatorname{var}_0\left(W_{ij}\right)\right\}^{1/2}}\right] = \frac{W_{ij} - \frac{n_i\left(n_i+n_j+1\right)}{2}}{\left\{n_i n_j\left(n_i + n_j + 1\right)/24\right\}^{1/2}}, \quad \text{for } 1 \leq i < j \leq k.$$



Motivation

Optional :

$\Rightarrow$

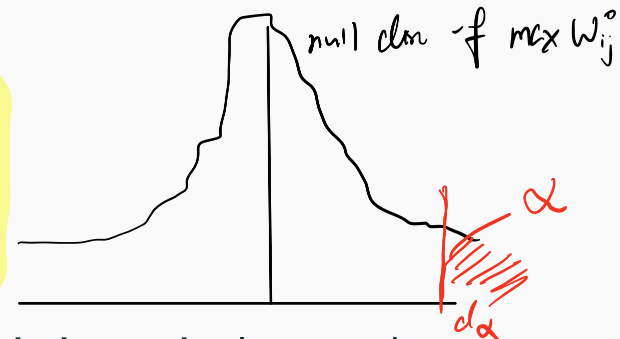
- When H_0 is true, the $[k(k-1)/2]$-component vector $\left(W_{12}^*, W_{13}^*, \ldots, W_{k-1,k}^*\right)$ has, as $\min\left(n_1, \ldots, n_k\right)$ tends to infinity, an asymptotic multivariate normal distribution with mean vector $\mathbf{0}$, and

$$\left(W_{12}^*, W_{13}^*, \ldots, W_{k-1,k}^*\right) \sim \left(\ldots \frac{Z_j - Z_i}{\sqrt{\frac{n_i+n_j}{2n_i n_j}}} \ldots\right)$$

the $[k(k-1)/2]$-component vector of differences where $Z_1, \ldots, Z_k$ are mutually independent and Z_i has an $N\left(0, 1/n_i\right)$ distribution

$\Rightarrow$

$$\max_{1 \leq i < j \leq k} W_{ij}^* \sim \max_{1 \leq i < j \leq k} \frac{Z_j - Z_i}{\sqrt{\frac{n_i+n_j}{2n_i n_j}}}$$



$\Rightarrow$ To get the null distribution for the simultaneous tests, it is equivalent to know the null distribution of $\max_{1 \leq i < j \leq k} W_{ij}^*$

$\Rightarrow$ It is then equivalent to the distribution of the maximum difference when we draw k independent normal random variables with $N(0, 1/n_i)$

Procedure

critical value

For each pair of treatments (i, j), for $1 \leq i < j \leq k$,

Decide $\tau_i < \tau_j$ if $W^*_{ij} \geq d_\alpha$; otherwise decide $\tau_i = \tau_j$.

d_α is the upper α percentile of the maximum range of k normal variates with $N(0, 1/n_i)$.

Example: Motivational Effect of Knowledge of Performance

For Hundal's (1969) study to assess the motivational effects of knowledge of performance, we found using the Jonckheere-Terpstra test that there was sufficient evidence in the sample data to conclude that $\tau_1 \leq \tau_2 \leq \tau_3$ with at least one strict inequality.

To examine which of the types of information (none, rough, or accurate) lead to differences in median numbers of pieces processed.

Control (no information)	Group B (rough information)	Group C (accurate information)
39.5	37.5	48
35	40	40.5
38	47	45
42.5	44	43
44.5	41.5	46
41	42	50

Simultaneous tests :

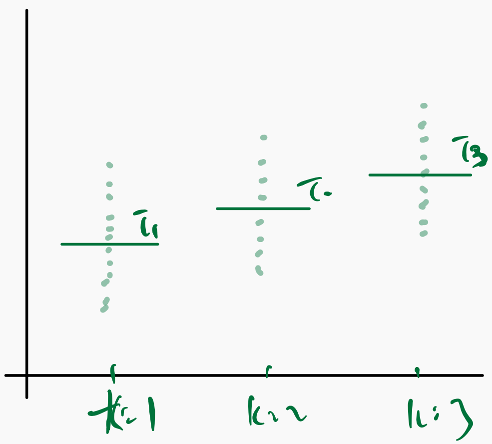
$$\begin{cases} H_0: \tau_1 = \tau_2 & H_1: \tau_1 < \tau_2 \\ H_0: \tau_1 = \tau_3 & H_1: \tau_1 < \tau_3 \\ H_0: \tau_2 = \tau_3 & H_1: \tau_2 < \tau_3 \end{cases}$$

Control experimentwise error rate 5%,

```
> library(NSM3)
> cHayStonLSA(alpha=0.05,k=3)
[1] 2.94
```

To compute critical value d_α

```
> sum(rank(c(39.5,35,38,42.5,44.5,41,37.5,40,47,44,41.5,42))[7:12])
[1] 44
> sum(rank(c(39.5,35,38,42.5,44.5,41,48,40.5,45,43,46,50))[7:12])
[1] 53
> sum(rank(c(37.5,40,47,44,41.5,42,48,40.5,45,43,46,50))[7:12])
[1] 49
```



$$W_{12}^* = \frac{[44 - 6(13)/2]}{\sqrt{6 \times 6 \times 13/24}} = 1.132277$$

$$W_{13}^* = \frac{[53 - 6(13)/2]}{\sqrt{6 \times 6 \times 13/24}} = 3.170376$$

$$W_{23}^* = \frac{[49 - 6(13)/2]}{\sqrt{6 \times 6 \times 13/24}} = 2.264554$$

$$\Rightarrow |W_{12}^*| < 2.94 \quad \Longrightarrow \quad \text{decide } \tau_1 = \tau_2$$

$$|W_{13}^*| > 2.94 \quad \Longrightarrow \quad \text{decide } \tau_1 < \tau_3$$

$$|W_{23}^*| < 2.94 \quad \Longrightarrow \quad \text{decide } \tau_2 = \tau_3$$

Thus, at an experimentwise error rate of .05, we have reached the conclusion that $\tau_1 < \tau_3$ but$\tau_1 = \tau_2$ and $\tau_2 = \tau_3$.