STA 104 Applied Nonparametric Statistics

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

Xiner Zhou

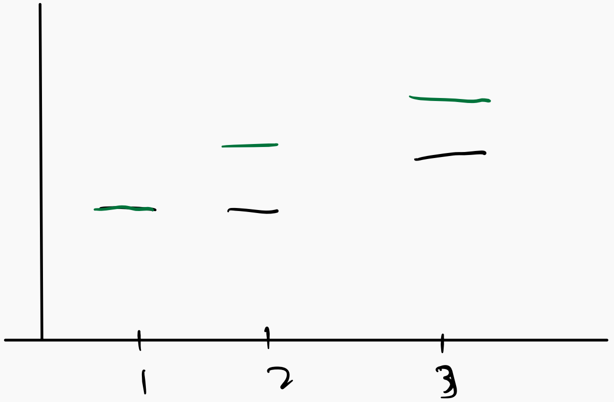
Department of Statistics, University of California, Davis

Table of contents

One-Sided Treatments-versus-Control Multiple Comparisons for Treatment-versus-Control Alternatives

When the main interest is on treatment-versus-control comparisons, we do not compare all treatments, but only each noncontrol treatment with the control on a directional bias. This situation arises, for example, in drug screening in the examination of many new treatments in hopes of improving on a standard, and there is no initial reason to perform between treatment comparisons. Of course, comparisons could be carried out later between treatments that were selected as being better, if there is intention to pick the optimal one.

After rejection of



One-Sided Upper-Tail Test:

$$H_0 : [\tau_i = \tau_1, \text{ for } i = 2, \ldots, k]$$

$$H_1 : [\tau_i \geq \tau_1\ ,, \text{ for } i = 2, \ldots, k, \text{ with at least one strict inequality}]$$

One-Sided Lower-Tail Test:

$$H_0 : [\tau_i = \tau_1, \text{ for } i = 2, \ldots, k]$$

$$H_1 : [\tau_i \leq \tau_1\ ,, \text{ for } i = 2, \ldots, k, \text{ with at least one strict inequality}]$$

with the Fligner-Wolf test, it is important to reach conclusions about exactly which treatment is better than control, and these conclusions are naturally one-sided, in accordance with the directional alternatives setting.

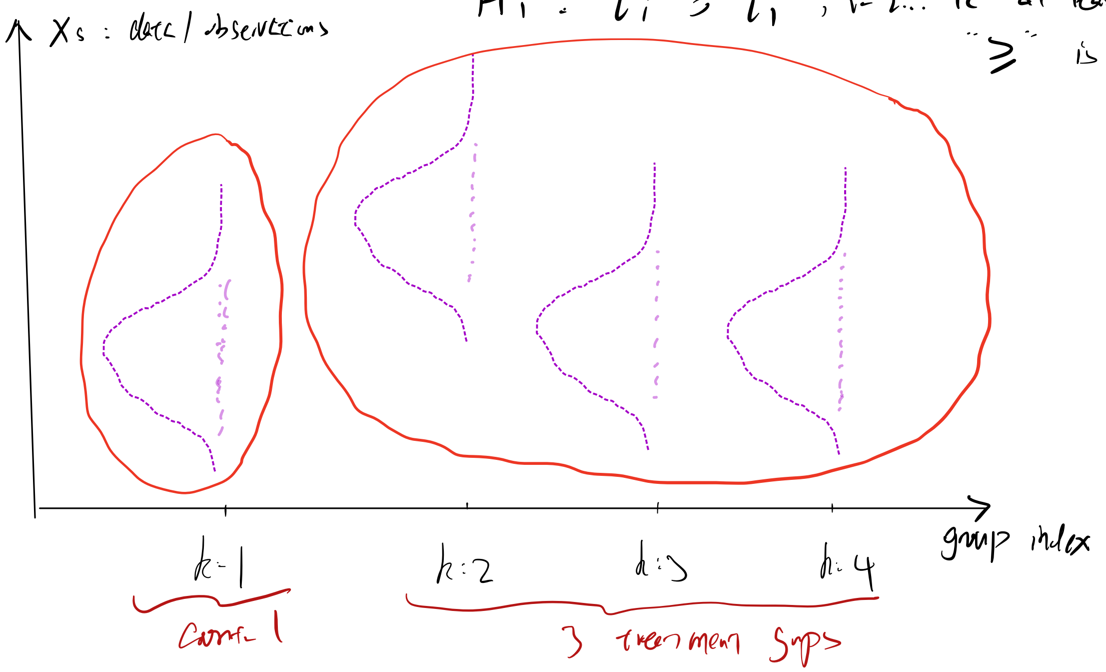
Hypothesis

One-Sided ~~Lower~~ Upper-Tail Test:

$$\left.\begin{cases} H_0 : \tau_1 = \tau_2 & H_1 : \tau_2 > \tau_1 \\ H_0 : \tau_1 = \tau_3 & H_1 : \tau_3 > \tau_1 \\ \dots & \\ H_0 : \tau_{k-1} = \tau_k & H_1 : \tau_k > \tau_1 \end{cases}\right\} k-1 \text{ simultaneous tests/multiple comparisons}$$

Review of FW test:

$H_1: \tau_i \geq \tau_1$, $i = 2 \ldots k$ at least one of "$\geq$" is "$>$"



Step 1: jointly rank X_{ij}'s $\to R_{ij}$

Step 2: Sum up ranks associated with all treatment grps

$$FW = \sum_{i=2}^{K} \sum_{j=1}^{n_i} R_{ij} \quad \text{large}$$

$\Rightarrow$ For each $H_0: \tau_i = \tau_1$ $\quad H_1: \tau_i > \tau_1$:

$$\text{Compare } \underbrace{R_{i\cdot}}_{\text{ave. ranks of ith treatment grp}} - \underbrace{R_{1\cdot}}_{\text{control grp}}$$

Optional :

$\Rightarrow$ Jointly rank all N of the sample observations and let $R_{1.}, \ldots, R_{k.}$ be the averages of these joint ranks associated with treatments $1, \ldots, k$, respectively. ~~(as in Kruskal Wallis statistic.)~~

$\Rightarrow$ For each of the $k-1$ noncontrol treatments, calculate the difference $R_{i.} - R_{1.}$, $i = 2, \ldots, k$. test statistic

$\Rightarrow$ When H_0 is true, and $n_1 = b$ and $n_2 = \cdots = n_k = n$, with both n and b large: the $k-1$-component vector

$$(R_{2.} - R_{1.}, R_{3.} - R_{1.}, \ldots, R_{k.} - R_{1.}) \sim \cancel{(Z_1, \ldots, Z_{k-1})}$$

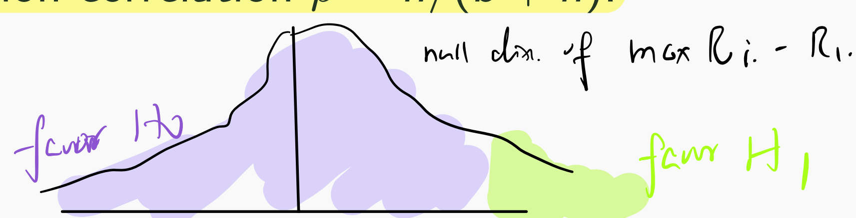
Multivariate Normal Distribution

~~$(k-1)N(0,1)$ variables with common correlation $\rho = n/(b+n)$~~

$\Rightarrow$

$$\max_{2 \le i \le k} R_{i.} - R_{1.} \sim \max_{1 \le i \le k-1} Z_i \times \sqrt{\frac{N(N+1)}{12}} \sqrt{\frac{1}{b} + \frac{1}{n}}$$

$\Rightarrow$ To get the null distribution for the simultaneous tests, it is equivalent to know the null distribution of $\max_{1 \le i \le k-1} Z_i$, the maximum when we draw $k-1$ normal random variables with common correlation $\rho = n/(b+n)$.



Procedure

critical val

When $n_1 = b$ and $n_2 = \cdots = n_k = n$

For each treatments i,

$$\text{Decide } \tau_i > \tau_1 \text{ if } (R_{i.} - R_{1.}) \geq \boxed{m_\alpha^* \left[\frac{N(N+1)}{12}\right]^{1/2} \left(\frac{1}{b} + \frac{1}{n}\right)^{1/2}}$$

otherwise decide $\cancel{\tau_u = \tau_1, u = 2, \ldots, k.}$ $\tau_i = \tau_1$

- m_α^* is the α upper percentile of the $\max_{1 \leq i \leq k-1} Z_i$, the maximum when we draw $k-1$ normal random variables with common correlation $\rho = n/(b+n)$.

General setting: arbitrary sample sizes (Bonferroni's Inequality) [4]

For each treatments i,

$$\text{Decide } \tau_i > \tau_1 \text{ if } (R_{i.} - R_{1.}) \geq z_{\alpha^*} \left[\frac{N(N+1)}{12}\right]^{1/2} \left(\frac{1}{n_1} + \frac{1}{n_u}\right)^{1/2}$$

otherwise decide $\cancel{\tau_u = \tau_1, u = 2, \ldots, k.}$ $\tau_i = \tau_1$

- $\alpha^* = \alpha/(k-1)$

[4] Bonferroni's general approximate procedure can often be quite conservative in practice, as a direct result of the conservative nature of the Bonferroni Inequality.

Example: Motivational Effect of Knowledge of Performance

To further investigate which (if either) of the two types of additional information (rough or accurate) lead to improvement or increase in median numbers of pieces processed relative to the no information control (treatment 1).

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

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

```
# large-sample approximation
> cMaxCorrNor(alpha=0.05,k=2,rho=6/12)
[1] 1.91
> sqrt(18*19/12)*sqrt(1/6+1/6)*1.91
[1] 5.887015
```

$\rightarrow m_i^2$

$\leftarrow$ critical val

Decide $\cancel{\tau_u > \tau_i}$ if $\cancel{(R_u - R_1)} \geq 5.88$.

$\tau_i > \tau_1 \quad R_{i.} - R_{1.}$

```
> ranks=rank(c(39.5,35,38,42.5,44.5,41,37.5,40,47,44,41.5,42,48,40.5,45,43,46,5
> R1=mean(ranks[1:6])
> R2=mean(ranks[7:12])
> R3=mean(ranks[13:18])
> R2-R1
[1] 2.333333
> R3-R1
[1] 7.166667
```

$R_{2.} - R_{1.}$

$R_{3.} - R_{1.}$

$$(R_{2.} - R_{1.}) = 2.3 < 5.88 \Rightarrow \text{ decide } \tau_2 = \tau_1,$$
$$(R_{3.} - R_{1.}) = 7.1 \geq 5.88 \Rightarrow \text{ decide } \tau_3 > \tau_1.$$

Thus at an experimentwise error rate of .05, we have reached the conclusion that accurate information leads to significantly more pieces processed than the no information control, while rough information do not lead to significant improvement compared to no information control.