One-way ANOVA (univariate)

Tukey’s honest significant differences

07/27/2021

Berkeley
UNIVERSITY OF CALIFORNIA

Simultaneous Bonferroni CIs

Theorem B. Under the one-way ANOVA model assumptions,

$$\frac{\sqrt{n_i}\left(\bar{Y}_{i\cdot} - \mu - \alpha_i\right)}{S_i} \sim t_{n_i - 1}, \text{ for any } i = 1, \ \ldots, \ k.$$

If one establishes m confidence intervals, and wishes to have an overall confidence level of $1 - \alpha$, each individual confidence interval can be adjusted to the level of $1 - \frac{\alpha}{m}$.

$\mu + \alpha_i$, $i = 1, \cdots k$

Simultaneous $(1 - \alpha) \times 100\%$ CIs for population means :

$$\bar{Y}_{i\cdot} \pm t_{n_i - 1}(\alpha'/2)\frac{S_i}{\sqrt{n_i}} \quad \text{for any } i = 1, \ \ldots, \ k,$$

where $\alpha' = \cancel{2\alpha/[k(k-1)]}$.

α / k

```r
get_bound <- function(vec, k, alpha = 0.05){
  center = mean(vec)
  n = length(vec)
  halfwidth = qt(1 - (alpha/2)/k, df = n - 1)*sd(vec)/sqrt(n)

  return(c(center - halfwidth, center + halfwidth))
}
```

Simultaneous inference statements: Bonferroni

Corollary B. We can utilize the duality between the CIs and the HTs to perform testings on

$$H_0^i : \underbrace{\alpha_i = 0}_{\mu+\alpha_i=\mu} \quad \text{versus} \quad H_1^i : \alpha_i \neq 0.$$

The conclusions we draw about each test hold **simultaneously**.

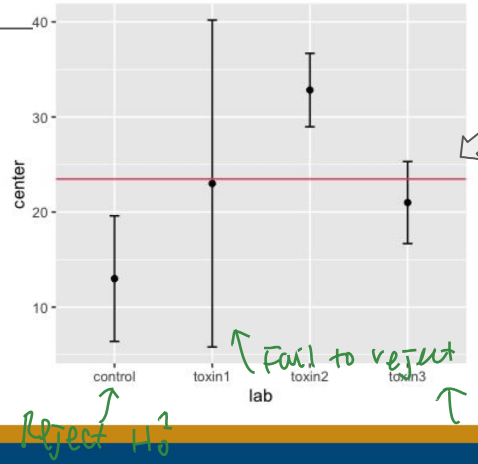
Toxin 1	Toxin 2	Toxin 3	Control
28	33	18	11
23	36	21	14
14	34	20	11
27	29	22	16
	31	24	
	34		

list

```
> CIs <- sapply(Observations, get_bound, k=k, alpha=0.05)
> CIs
          [,1]     [,2]     [,3]     [,4]
[1,]  5.807657 28.97077 16.68534 6.396238
[2,] 40.192343 36.69589 25.31466 19.603762
```

```
library(ggplot2)
centers <- sapply(Observations, mean)
CI_df <- data.frame(lab = group_ind, center = centers,
                    lower = CIs[1,], upper = CIs[2,])
ggplot(CI_df, aes(x = lab, y = center)) + geom_point() +
  geom_errorbar(width = 0.1, aes(ymin = lower, ymax = upper)) +
  geom_hline(yintercept = mean(unlist(Observations)), col = 2)
```

$= \mu$



Simultaneous inference statements: Bonferroni for pairwise differences

$\bar{Y}_{i\cdot} \sim N(\mu+\alpha_i, \frac{\sigma^2}{n_i})$ $\rightarrow \bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}$

$\bar{Y}_{r\cdot} \sim N(\mu+\alpha_r, \frac{\sigma^2}{n_r})$ $\sim N(\alpha_i - \alpha_r, \sigma^2(\frac{1}{n_i}+\frac{1}{n_r}))$

Theorem C. Under the one-way ANOVA model assumptions,

$$\frac{(\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) - (\alpha_i - \alpha_j)}{\sqrt{S_{ir}^2\left(\frac{1}{n_i} + \frac{1}{n_r}\right)}} \sim t_{n_i+n_r-2}, \text{ for any pair } i \neq r.$$

list [[i]]

c(i, r)

Simultaneous $(1-\alpha)\times 100\%$ CIs for pairwise differences :

$$(\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) \pm t_{n_i+n_r-2}(\alpha'/2)\sqrt{S_{ir}^2\left(\frac{1}{n_i} + \frac{1}{n_r}\right)} \text{ for all pairs } i \neq r,$$

where $\alpha' = 2\alpha/[k(k-1)]$.

$\binom{k}{2}$

```r
get_pair_diff_CI <- function(pair, Observations, alpha=0.05){
  vec1 = Observations[[pair[1]]]
  vec2 = Observations[[pair[2]]]

  # bookkeeping
  center1 = mean(vec1)
  center2 = mean(vec2)
  n1 = length(vec1)
  n2 = length(vec2)
  k = length(Observations); m = k*(k-1)/2

  # Get CI bounds
  Sp2 = {(n1-1)*var(vec1) + (n2-1)*var(vec2)}/(n1+n2-2)
  halfwidth = qt(1-(alpha/2)/m, df = n1+n2-2)*sqrt(Sp2*(1/n1+1/n2))
  return(c(center1-center2,
           center1-center2-halfwidth, center1-center2+halfwidth))
}
```

i   r

Simultaneous inference statements: Bonferroni

apply(matrix, 2, FUN)
mean
= colMeans(matrix)

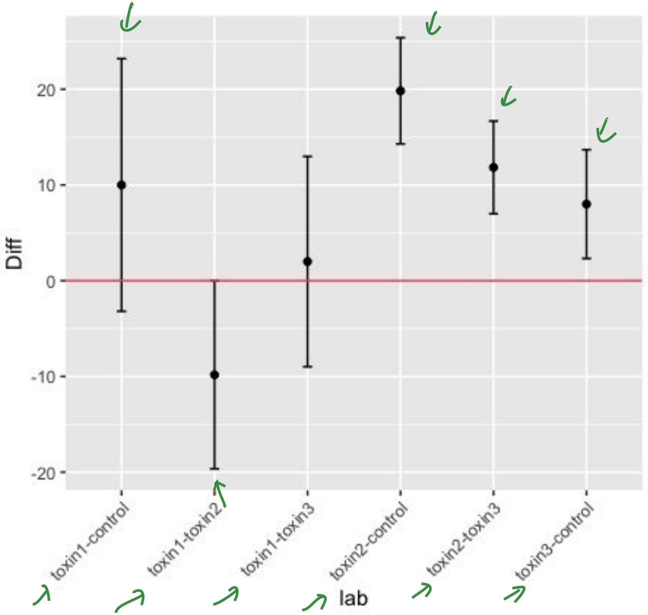
Corollary C. We can utilize the duality between the CIs and the HTs to perform testings on

$$H_0^{ir}: \underbrace{\alpha_i = \alpha_r}_{} \quad \text{versus} \quad H_1^{ir}: \alpha_i \neq \alpha_r.$$

$\alpha_i - \alpha_r = 0$

The conclusions we draw about each test hold **simultaneously**.

apply(matrix, 1, mean)
= rowMeans(matrix)




```
> combn(k, 2)
     [,1] [,2] [,3] [,4] [,5] [,6]
[1,]    1    1    1    2    2    3
[2,]    2    3    4    3    4    4
> CIs = apply(combn(k, 2), 2, get_pair_diff_CI, Observations)
> CIs
             [,1]      [,2]      [,3]      [,4]     [,5]      [,6]
[1,]  -9.83333333  2.000000 10.000000 11.833333 19.83333  8.000000
[2,] -19.64850220 -8.985056 -3.194625  6.991933 14.28520  2.317363
[3,]  -0.01816447 12.985056 23.194625 16.674734 25.38147 13.682637
> pair_names = apply(combn(k, 2), 2,
+       function(pair) paste0(group_ind[pair[1]],'-', group_ind[pair[2]]))
> pair_names
[1] "toxin1-toxin2"  "toxin1-toxin3"  "toxin1-control" "toxin2-toxin3"
[5] "toxin2-control" "toxin3-control"
> CI_df <- data.frame(lab = pair_names, Diff = CIs[1,],
+                    lower = CIs[2,], upper = CIs[3,])
> ggplot(CI_df, aes(x = lab, y = Diff)) + geom_point() +
+   geom_errorbar(width = 0.1, aes(ymin = lower, ymax = upper)) +
+   geom_hline(yintercept = 0, col = 2)
```

$Y_1, \dots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$ $\quad \max \frac{\sqrt{n}\,|Y_i - Y_j|}{S} \sim q_{n,n}$ →

$\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot} \sim N\left(\alpha_i - \alpha_j,\ \sigma^2\left(\frac{1}{n_i} + \frac{1}{n_r}\right)\right)$

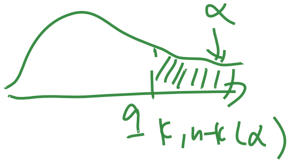
Simultaneous inference statements: Tukey's method

Theorem D. Under the one-way ANOVA model assumptions,

$$\max_{i,\,r} \frac{\left|(\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) - (\alpha_i - \alpha_j)\right|}{\sqrt{MS_W\left(\frac{1}{n_i} + \frac{1}{n_r}\right)}} \sim q_{k,\,n-k}.$$

$MS_W = \frac{SS_W}{n-k}$

studentized range distribution



Simultaneous $(1-\alpha) \times 100\%$ CIs for pairwise differences :

$$(\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) \pm q_{k,\,n-k}(\alpha)\sqrt{MS_W\left(\frac{1}{n_i} + \frac{1}{n_r}\right)} \quad \text{for all pairs } i \neq r.$$

```
qtukey(alpha, nmeans = k, df = n-k, lower.tail = FALSE)
```

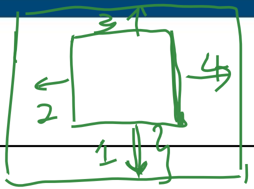
By definition of upper-tail quantile,

$$1-\alpha = P\left(\max_{i,r} \frac{\left|(\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) - (\alpha_i - \alpha_j)\right|}{\sqrt{MS_W\left(\frac{1}{n_i} + \frac{1}{n_r}\right)}} \le q_{k,n-k}(\alpha)\right)$$

$$= P\left(\frac{\left|(\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) - (\alpha_i - \alpha_r)\right|}{\sqrt{MS_W\left(\frac{1}{n_i} + \frac{1}{n_r}\right)}} \le q_{k,n-k}(\alpha) \text{ for any pair } i \neq r\right)$$

$$= P\left((\alpha_i - \alpha_r) \in (\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) \pm \text{Tukey\_halfwidth for any pair } i \neq r\right)$$

Simultaneous inference statements: Tukey's method



Theorem D. Under the one-way ANOVA model assumptions,

$$\max_{i,\,r} \frac{\left|(\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) - (\alpha_i - \alpha_j)\right|}{\sqrt{MS_W\left(\frac{1}{n_i} + \frac{1}{n_r}\right)}} \sim q_{k,\,n-k}.$$

Simultaneous $(1-\alpha)\times 100\%$ CIs for pairwise differences :

$$(\bar{Y}_{i\cdot} - \bar{Y}_{r\cdot}) \pm q_{k,\,n-k}(\alpha)\sqrt{MS_W\left(\frac{1}{n_i} + \frac{1}{n_r}\right)} \quad \text{for all pairs } i \neq r.$$

```
qtukey(alpha, nmeans = k, df = n-k, lower.tail = FALSE)
```

fit = lm(values ~ ind, data = input)
anova(fit)

values ind
toxin1
control


```
> anova_fit <- aov(values ~ ind, data = input)
> summary(anova_fit)
            Df Sum Sq Mean Sq F value   Pr(>F)    
ind          3  995.9   332.0   26.09 3.35e-06 ***
Residuals   15  190.8    12.7                     
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>
> posthoc <- TukeyHSD(x=anova_fit, which = 'ind', conf.level=0.95)
> par(mar=c(5.1, 6.3, 4.1, 2.1))
> plot(posthoc, las = 1)
> par(mar=c(5.1, 4.1, 4.1, 2.1))
```

plot(posthoc)

One-way ANOVA (univariate)

Non-parametric Kruskal-Wallis test

07/27/2021

Kruskal-Wallis test

Example 3. A clinical study is designed to assess differences in albumin levels in adults following diets with different amounts of protein. Is there is a difference in serum albumin levels among subjects on the three different diets?

5% Protein	10% Protein	15% Protein
3.1	3.8	4.0
2.6	4.1	5.5
2.9	2.9	5.0
	3.4	4.8
	4.2	

Hard to say if each column is a Normal variable.

H_0 : The three population medians are equal. $\leftrightarrow$

H_1 : The three population medians are not all equal.

Kruskal-Wallis test

Total Sample (Ordered Smallest to Largest)			Ranks		
5% Protein	10% Protein	15% Protein	5% Protein	10% Protein	15% Protein
2.6			1		
2.9	2.9		2.5	2.5	
3.1			4		
	3.4			5	
	3.8			6	
		4.0			7
	4.1			8	
	4.2			9	
		4.8			10
		5.0			11
		5.5			12

$R_1 = 7.5$ $\quad R_2 = 30.5$ $\quad R_3 = 40$

$\frac{R_1}{n_1} \quad \frac{R_2}{n_2} \quad \frac{R_3}{n_3}$

$\underbrace{0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0 \; 0}_{n}$ $\frac{n+1}{2}$

$R = \{\underbrace{H \geq c}\}$ 5.656

We reject the null hypothesis.

1. Rank all observations from smallest to largest;
2. Record the corresponding ranks;
3. Sum the ranks within each column/treatment, which is denoted by R_i ;
4. The test statistic for the Kruskal Wallis test is denoted H and is defined as follows:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} n_i \left(\frac{R_i}{n_i} - \frac{n+1}{2} \right)^2 .$$

SS_B

$$H = \frac{12}{12 \times 13} \left[3 \times \left(\frac{7.5}{3} - \frac{13}{2}\right)^2 + 5 \times \left(\frac{30.5}{5} - \frac{13}{2}\right)^2 + 4 \times \left(\frac{40}{4} - \frac{13}{2}\right)^2 \right] = 7.523$$

Kruskal-Wallis test

```
> group1 <- c(3.1, 2.6, 2.9)
> group2 <- c(3.8, 4.1, 2.9, 3.4, 4.2)
> group3 <- c(4.0, 5.5, 5.0, 4.8)
>
> #Combine into a vector
> Observations <- c(group1, group2, group3)
> trt_labels <- rep(c('5%', '10%', '15%'), times=c(3,5,4))
> Albumin <- data.frame(Obs = Observations, ind = trt_labels)
> kruskal.test(Obs ~ ind, data = Albumin)

        Kruskal-Wallis rank sum test

data:  Obs by ind
Kruskal-Wallis chi-squared = 7.5495, df = 2, p-value = 0.02294
```

≈ 7.523

```
> fit <- lm(Obs ~ ind, data = Albumin)
> anova(fit)
Analysis of Variance Table

Response: Obs
          Df Sum Sq Mean Sq F value  Pr(>F)
ind        2 6.8470  3.4235  12.617 0.00245 **
Residuals  9 2.4422  0.2714
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

1. Rank all observations from smallest to largest;
2. Record the corresponding ranks;
3. Sum the ranks within each column/treatment, which is denoted by R_i ;
4. The test statistic for the Kruskal Wallis test is denoted H and is defined as follows:

$$H = \frac{12}{n(n+1)} \sum_{i=1}^{k} n_i \left(\frac{R_i}{n_i} - \frac{n+1}{2} \right)^2 .$$

Two-way ANOVA (univariate)

12.3 of Rice

07/27/2021

Berkeley
UNIVERSITY OF CALIFORNIA

Two treatment factors

$9 = 3 \times 3$

Example 4. A fabric plant is studying the effect of several factors on the dyeing of cotton-synthetic cloth. At 300°C, three operators and three cycle times were selected. The finished cloth was compared to a standard, and a numerical score was assigned.

Cycle times	Operators 1	Operators 2	Operators 3
40 mins	23	27	31
	24	28	32
	25	26	29
50 mins	36	34	33
	35	38	34
	36	39	35
60 mins	28	35	26
	24	35	27
	27	34	25

3 x 3 factorial design:

- Main effects of operators;
- Main effects of cycle times;
- Different operators are more effective with different cycle times → Interaction effects.

Two treatment factors

expand.grid (vec1 , vec2)
c(a,b) c(α, β)

a α
b α
a β
b β

Example 3. A fabric plant is studying the effect of several factors on the dyeing of cotton-synthetic cloth. At 300°C, three operators and three cycle times were selected. The finished cloth was compared to a standard, and a numerical score was assigned.

Cycle times	Operators 1	2	3
40 mins	23	27	31
	24	28	32
	25	26	29
50 mins	36	34	33
	35	38	34
	36	39	35
60 mins	28	35	26
	24	35	27
	27	34	25

3 x 3 factorial design:

- Main effects of operators;
- Main effects of cycle times;
- Different operators are more effective with different cycle times → Interaction effects.

```
> head(Obs_w_names, 6)
  cycle_times operator Obs
1          40        1  23
2          40        1  24
3          40        1  25
4          50        1  36
5          50        1  35
6          50        1  36
```

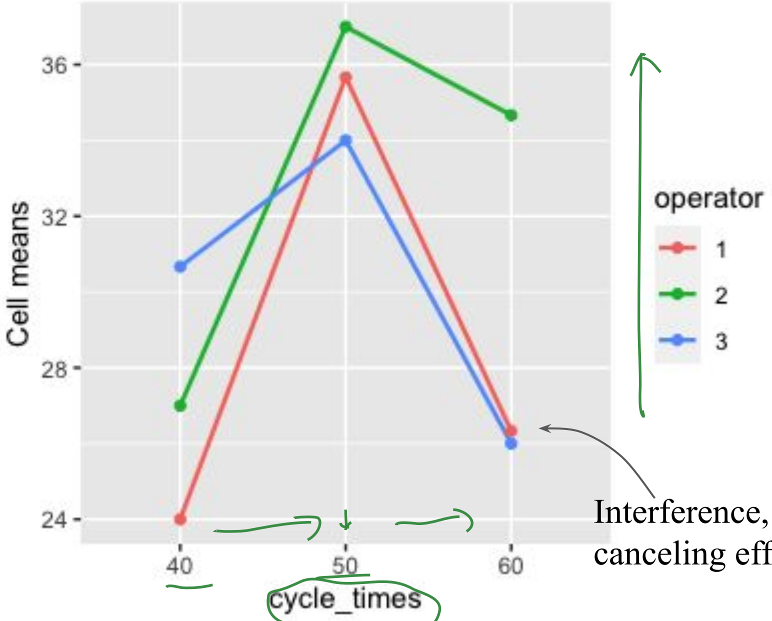
```
p <- 3
cell_names <- expand.grid(cycle_times = c(40, 50, 60), operator = c(1,2,3))
repeats <- rep(1:nrow(cell_names), each = p)
Cell_names <- cell_names[repeats, ]
Obs <- c(23, 24, 25, 36, 35, 36, 28, 24, 27,
         27, 28, 26, 34, 38, 39, 35, 35, 34,
         31, 32, 29, 33, 34, 35, 26, 27, 25)
Obs_w_names <- data.frame(cycle_times = factor(Cell_names$cycle_times),
                          operator = factor(Cell_names$operator), Obs = Obs)
```

Column-major order

Treatment means plot

Connecting **cell means** to examine the main effects and interaction effects.

Synergism, or compounding effect



Interference, or canceling effect

```
> Cell_means <- aggregate(Obs ~ cycle_times + operator,
+                         data = Obs_w_names, FUN = mean)
> head(Cell_means, 6)
  cycle_times operator      Obs
1          40        1 24.00000
2          50        1 35.66667
3          60        1 26.33333
4          40        2 27.00000
5          50        2 37.00000
6          60        2 34.66667
```

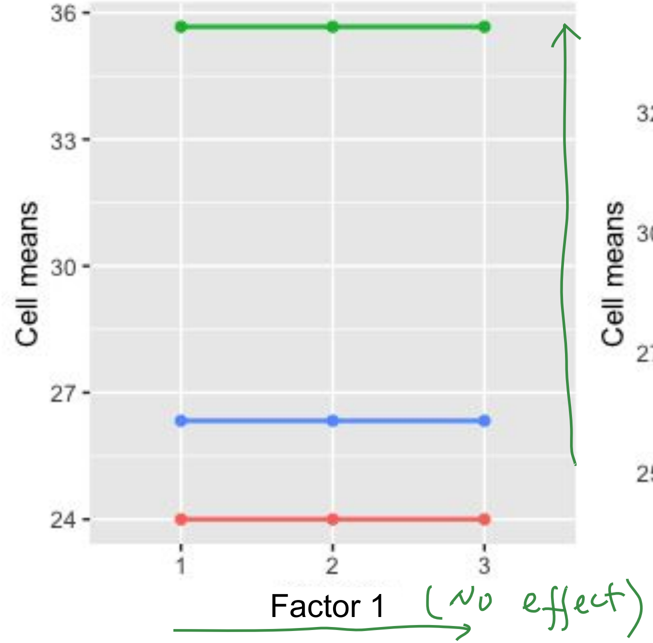
```
ggplot(data=Cell_means,
    aes(x=cycle_times, y=Obs, group=operator, color=operator)) +
  geom_line(size = 0.8) + geom_point() + ylab("Cell means")
```

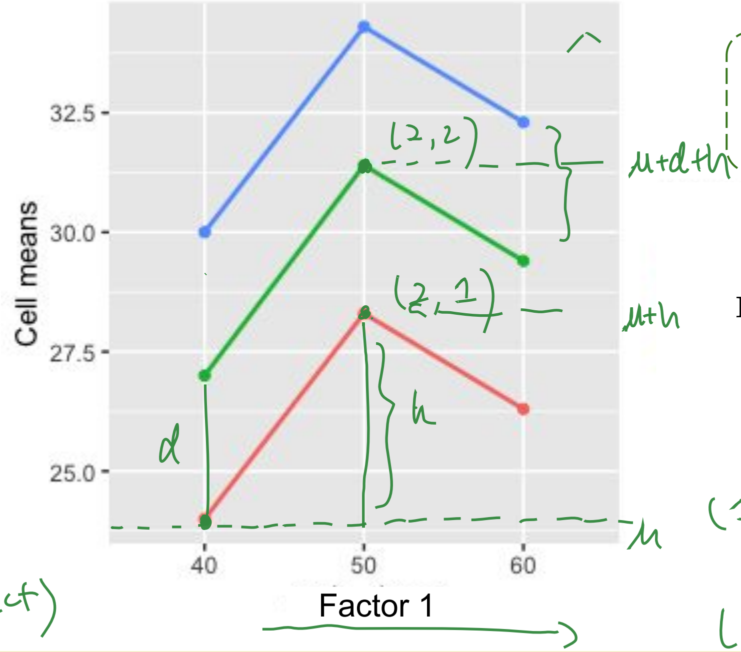
Treatment means plot

Connecting cell means to examine the main effects and interaction effects.

Additive

What if we observe the following plots? Are there any interaction effects?





Will changing operator alter the main effects of cycle times, or vice versa?

Population mean for (i,j)th cell:

$$\mu + \alpha_i + \beta_j$$

$\alpha_1 = 0$, $\beta_1 = 0$

$(1,1) \to \mu$

$\alpha_2 = h$, $\beta_1 = 0$

$(2,1) \to \mu + h$

$(2,2)$

$\alpha_2 = h$

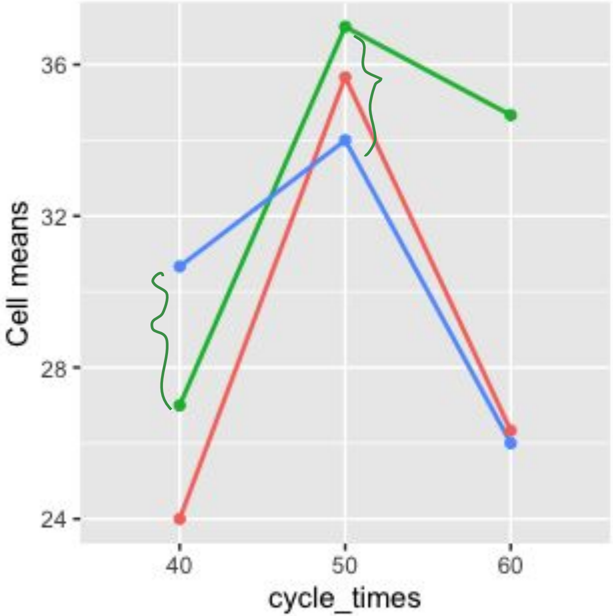
$\beta_2 = d$

$\mu + h + d$

Treatment means plot

Connecting cell means to examine the main effects and interaction effects.

Cycle times	Operators 1	Operators 2	Operators 3
40 mins	23	27	31
	24	28	32
	25	26	29
50 mins	36	34	33
	35	38	34
	36	39	35
60 mins	28	35	26
	24	35	27
	27	34	25



Interaction can be defined as ***failure of the main effects of one factor to be the same at different levels of the other factor****.*

Graphically, non-parallel lines connecting means.

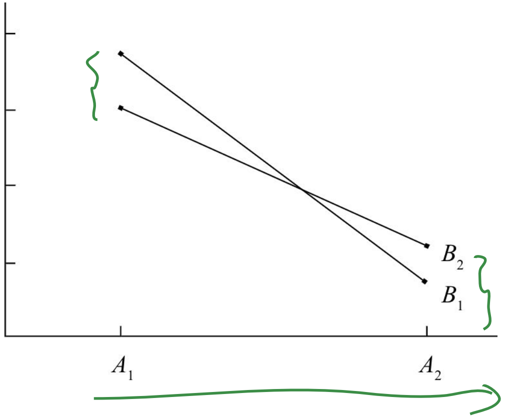
Population mean for (i, j)th cell:

$$\mu + \alpha_i + \beta_j + \delta_{ij}$$

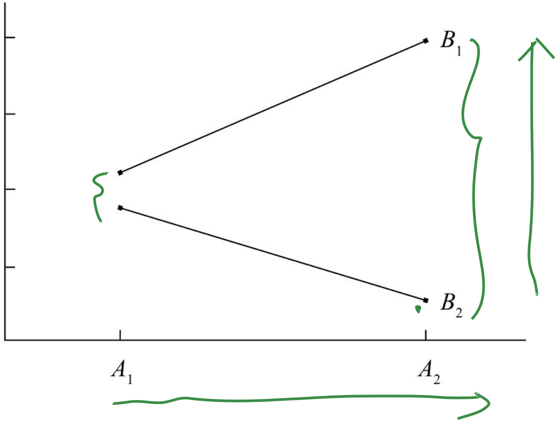
multiplicative

Treatment means plot

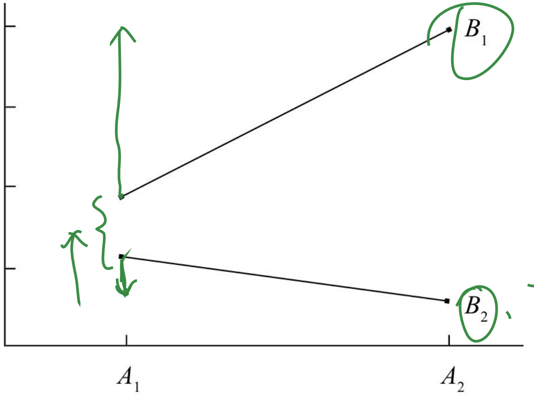
Connecting cell means to examine the main effects and interaction effects.



Large effect of Factor A, <u>no effect</u> of Factor B with a slight interaction.



<u>No effect</u> of Factor A, a large effect of Factor B, with a very large interaction



An effect of Factor A, a large effect of Factor B with a large interaction

Two-way ANOVA: model assumption

$\to (\mu, \bar{\alpha}, \bar{\beta}, \bar{\delta})$

$\to (\mu', \bar{\alpha}', \bar{\beta}', \bar{\delta}')$

$\Sigma \alpha_i \neq 0$, $\Sigma \beta_j \neq 0$, $\delta_{i\cdot} \neq 0$, $\delta_{\cdot j} \neq 0$

$\underbrace{\mu + \bar{\alpha} + \bar{\beta} + \bar{\delta}_{i\cdot} + \bar{\delta}_{\cdot j} - \bar{\delta}_{\cdot\cdot}}_{\mu'} + \underbrace{\alpha_i - \bar{\alpha}}_{\alpha_i'} + \underbrace{\beta_j - \bar{\beta}}_{\beta_j'} + \underbrace{\delta_{ij} - \bar{\delta}_{i\cdot} - \bar{\delta}_{\cdot j} + \bar{\delta}_{\cdot\cdot}}_{\delta_{ij}'}$

Cycle times	Operators 1	Operators 2	Operators 3
40 mins	23	27	31
	24	28	32
	25	26	29
50 mins	36	34	33
	35	38	34
	36	39	35
60 mins	28	35	26
	24	35	27
	27	34	25

Balanced design

May be imbalanced.

Model assumption ($l = 1, \cdots, n_{ij}$, $i = 1, \cdots, I$, $j = 1, \ldots, J$):

$$Y_{ijl} = \underbrace{\mu}_{\text{common mean level}} + \underbrace{\alpha_i}_{\text{main effect of } i} + \underbrace{\beta_i}_{\text{main effect of } j} + \underbrace{\delta_{ij}}_{\text{interaction between } i \text{ and } j} + \underbrace{\epsilon_{ijl}}_{\overset{\text{iid}}{\sim} N(0,\sigma^2)}.$$

$\sum_{i=1}^{I} \alpha_i = 0.$ $\quad \sum_{j=1}^{J} \beta_i = 0.$ $\quad \sum_{i=1}^{I} \delta_{ij} = \sum_{j=1}^{J} \delta_{ij} = 0.$

Indentifiable

$\delta_{\cdot j} = \delta_{i\cdot} = 0$

$H = H_0 \cup H_1$

$$H_0: \alpha_1 = \cdots = \alpha_I = 0 \quad \text{versus} \quad H_1: \text{not all } \alpha_i\text{'s are zero}$$

$$H_0: \beta_1 = \cdots = \beta_J = 0 \quad \text{versus} \quad H_1: \text{not all } \beta_j\text{'s are zero}$$

$$H_0: \text{all } \delta_{ij}\text{'s are zero} \quad \text{versus} \quad H_1: \text{not all } \delta_{ij}\text{'s are zero}$$

Two-way ANOVA: LRT

$n=\sum_{i=1}^{I}\sum_{j=1}^{J} n_{ij}$

$\vec{\alpha}=(\alpha_1,\dots,\alpha_I)$ $\quad$ $\vec{\delta}=(\delta_{ij})$ $\quad i=1,\dots,I$

$\vec{\beta}=(\beta_1,\dots,\beta_J)$ $\quad j=1,\dots,J$

Proposition D. Assume $Y_{ijl} = \mu + \alpha_i + \beta_j + \delta_{ij} + \epsilon_{ijl}$, where $\epsilon_{ijl} \overset{\text{iid}}{\sim} N(0,\sigma^2)$. We can derive the maximum likelihood in $\Theta = \Theta_0 \cup \Theta_1$:

$$\sup_{\Theta} L = \left(\frac{1}{\sqrt{2\pi\hat{\sigma}_n^2}}\right)^n e^{-\frac{n}{2}},$$

where $\hat{\sigma}_n^2 = n^{-1}\sum_i\sum_j\sum_{l=1}^{n_{ij}}\left(Y_{ijl}-\bar{Y}_{ij\cdot}\right)^2$.

Proof*. $L(\vec{\alpha},\vec{\beta},\vec{\delta},\mu,\sigma^2 \mid Y_{ijl}) = \prod_i\prod_j\prod_{l=1}^{n_{ij}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(Y_{ijl}-\overbrace{\mu-\alpha_i-\beta_j-\delta_{ij}})^2}{2\sigma^2}}$

$\ell(\vec{\alpha},\vec{\beta},\vec{\delta},\mu,\sigma^2 \mid Y_{ijl}) = n\log\frac{1}{\sqrt{2\pi\sigma^2}} - \frac{1}{2\sigma^2}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{n_{ij}}(Y_{ijl}-\mu-\alpha_i-\beta_j-\delta_{ij})$

$\Theta = \{\underbrace{\sum_{i=1}^{I}\alpha_i=0}_{I-1}, \underbrace{\sum_{j=1}^{J}\beta_j=0}_{J-1}, \underbrace{\sum_{i=1}^{I}\delta_{ij}=\sum_{j=1}^{J}\delta_{ij}=0}_{(I-1)(J-1)}, \mu\in\mathbb{R}, \sigma^2>0\}$

$\dim\Theta = (I-1)+(J-1)+(I-1)(J-1)+1+1 = IJ+1.$

$\delta_{IJ} = \sum_{i=1}^{I-1}\sum_{j=1}^{J-1}\delta_{ij}$

$-\frac{1}{2\sigma^2}\sum_{l=1}^{n_{IJ}}(Y_{IJl}-\mu-\alpha_I-\beta_J-\delta_{IJ})^2$

$-\frac{1}{2\sigma^2}\sum_{i=1}^{I-1}\sum_{l=1}^{n_{iJ}}(Y_{iJl}-\mu-\alpha_i-\beta_J-\delta_{iJ})^2$

$=\sum_{j=1}^{J-1}\beta_j$ $\quad$ $=\sum_{i=1}^{I-1}\alpha_i$

$$\begin{matrix} \delta_{11} & \delta_{12} & \cdots & \delta_{1J} \\ \delta_{21} & \delta_{22} & \cdots & \delta_{2J} \\ \vdots & \vdots & & \vdots \\ \delta_{I1} & \delta_{I2} & \cdots & \delta_{IJ} \end{matrix}$$

$\delta_{I1} = -\sum_{i=1}^{I-1}\delta_{i1}$

$\ell(\vec{\alpha},\vec{\beta},\vec{\delta},\mu,\sigma^2 \mid Y_{ijl}) = -\frac{n}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{I-1}\sum_{j=1}^{J-1}\sum_{l=1}^{n_{ij}}(Y_{ijl}-\mu-\alpha_i-\beta_j-\delta_{ij})^2 - \frac{1}{2\sigma^2}\sum_{j=1}^{J-1}\sum_{l=1}^{n_{Ij}}(Y_{Ijl}-\mu-\alpha_I-\beta_j-\delta_{Ij})^2$

Two-way ANOVA: LRT

Proof *cont'd*.

$$\vec{\alpha},\ \vec{\beta},\ \vec{\delta},\ \mu,\ \sigma^2$$

$$\frac{\partial l}{\partial \mu} = \frac{1}{\sigma^2}\boxed{\cancel{\sum_i}\sum_j \sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \mu - \alpha_i - \beta_j - \delta_{ij}\right) = 0} \rightarrow \sum_{i,j,l} Y_{ijk} = \sum_{i,j,l}(\mu+\alpha_i+\beta_j+\delta_{ij})$$

$$\left\{\begin{array}{l} \frac{\partial l}{\partial \alpha_1} = \cancel{\frac{1}{\sigma^2}} \sum_{j=1}^{J}\sum_{l=1}^{n_{1j}} \left(Y_{1jl} - \mu - \alpha_1 - \beta_j - \delta_{1j}\right) - \cancel{\frac{1}{\sigma^2}} \sum_{j=1}^{J}\sum_{l=1}^{n_{Ij}} \left(Y_{Ijl} - \mu - \alpha_I - \beta_j - \delta_{Ij}\right) = 0 \\ \vdots \\ \frac{\partial l}{\partial \alpha_{I-1}} = \cancel{\frac{1}{\sigma^2}} \sum_{j=1}^{J}\sum_{l=1}^{n_{(I-1)j}} \left(Y_{(I-1)jl} - \mu - \alpha_{(I-1)} - \beta_j - \delta_{ij}\right) - \cancel{\frac{1}{\sigma^2}} \sum_{j=1}^{J}\sum_{l=1}^{n_{Ij}} \left(Y_{Ijl} - \mu - \alpha_I - \beta_j - \delta_{Ij}\right) = 0 \end{array}\right.$$

Summing all $I-1$ likelihood equations, we get

$$\sum_{i=1}^{I-1}\sum_{j=1}^{J}\sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \mu - \alpha_i - \beta_j - \delta_{ij}\right) - (I-1)\sum_{j=1}^{J}\sum_{l=1}^{n_{Ij}} \left(Y_{Ijl} - \mu - \alpha_I - \beta_j - \delta_{Ij}\right) = 0$$

$$\cancel{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \mu - \alpha_i - \beta_j - \delta_{ij}\right)} - I\sum_{j=1}^{J}\sum_{l=1}^{n_{Ij}} \left(Y_{Ijl} - \mu - \alpha_I - \beta_j - \delta_{Ij}\right) = 0$$

$$\Rightarrow \boxed{\sum_{j=1}^{J}\sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \mu - \alpha_i - \beta_j - \delta_{ij}\right) = 0} \qquad \boxed{\sum_{i=1}^{I}\sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \mu - \alpha_i - \beta_j - \delta_{ij}\right) = 0}$$

Two-way ANOVA: LRT

Proof *cont'd*.

$$\frac{\partial l}{\partial \delta_{ij}} = \frac{1}{\sigma^2}\underbrace{\sum_{l=1}^{n_{ij}}(Y_{ijl}-\mu-\alpha_i-\beta_j-\delta_{ij})}_{=0} + \frac{1}{\sigma^2}\cancel{\sum_{l=1}^{n_{IJ}}(Y_{IJl}-\mu-\alpha_I-\beta_J-\delta_{IJ})}^{=0}$$

$$-\frac{1}{\sigma^2}\cancel{\sum_{l=1}^{n_{Ij}}(Y_{Ijl}-\mu-\alpha_I-\beta_j-\delta_{Ij})}^{=0} - \frac{1}{\sigma^2}\cancel{\sum_{l=1}^{n_{iJ}}(Y_{iJl}-\mu-\alpha_i-\beta_J-\delta_{iJ})}^{=0}$$

$$\sum_{i=1}^{I-1}\sum_{j=1}^{J-1}\frac{\partial l}{\partial \delta_{ij}} = \sum_{i=1}^{I-1}\sum_{j=1}^{J-1}\sum_{l=1}^{n_{ij}}(Y_{ijl}-\mu-\alpha_i-\beta_j-\delta_{ij}) + (I-1)(J-1)\sum_{l=1}^{n_{IJ}}(Y_{IJl}-\mu-\alpha_I-\beta_J-\delta_{IJ})^{=0}$$

$$-(I-1)\sum_{j=1}^{J-1}\sum_{l=1}^{n_{Ij}}(Y_{Ijl}-\mu-\alpha_I-\beta_j-\delta_{Ij}) - (J-1)\sum_{i=1}^{I-1}\sum_{l=1}^{n_{iJ}}(Y_{iJl}-\mu-\alpha_i-\beta_J-\delta_{iJ})$$

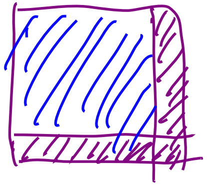
$$= \cancel{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{l=1}^{n_{ij}}(Y_{ijl}-\mu-\alpha_i-\beta_j-\delta_{ij})}^{=0} + IJ\sum_{l=1}^{n_{IJ}}(Y_{IJl}-\mu-\alpha_I-\beta_J-\delta_{IJ})$$

$$-I\cancel{\sum_{j=1}^{J}\sum_{l=1}^{n_{Ij}}(Y_{Ijl}-\mu-\alpha_I-\beta_j-\delta_{Ij})}^{=0} - J\cancel{\sum_{i=1}^{I}\sum_{l=1}^{n_{iJ}}(Y_{iJl}-\mu-\alpha_i-\beta_J-\delta_{iJ})}^{=0}$$

$$\Rightarrow \sum_{l=1}^{n_{IJ}}(Y_{IJl}-\mu-\alpha_I-\beta_J-\delta_{IJ}) = 0$$

$$\sum_{i=1}^{I-1}\frac{\partial l}{\partial \delta_{ij}} = \cancel{\sum_{i=1}^{I-1}\sum_{l=1}^{n_{ij}}(Y_{ijl}-\mu-\alpha_i-\beta_j-\delta_{ij})}^{=0} - (I-1)\overbrace{\boxed{\sum_{l=1}^{n_{Ij}}(Y_{Ijl}-\mu-\alpha_I-\beta_j-\delta_{Ij})}}^{\beta_j} - \cancel{\sum_{i=1}^{I}\sum_{l=1}^{n_{iJ}}(Y_{iJl}-\mu-\alpha_i-\beta_J-\delta_{iJ})}^{=0}$$

$$\sum_{j=1}^{J} \frac{\partial l}{\partial \delta_{ij}} = \cdots = 0 \implies \sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \mu - \alpha_i - \beta_j - \delta_{ij} \right) = 0 .$$



From $\frac{\partial l}{\partial \delta_{ij}} = 0$, we get:

$$\boxed{\sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \mu - \alpha_i - \beta_j - \delta_{ij} \right) = 0, \quad i = 1, \dots I, \quad j = 1 \dots J.}$$

$$\underbrace{\widehat{\mu + \alpha_i + \beta_j + \delta_{ij}}}_{\text{population mean MLE for } (i,j)\text{th cell}} = \frac{1}{n_{ij}} \sum_{l=1}^{n_{ij}} Y_{ijl} = \underbrace{\overline{Y}_{ij\cdot}}_{\text{Sample mean for } (i,j)\text{th cell.}}$$

$$\frac{\partial l}{\partial \sigma^2} = 0 \implies \hat{\sigma}_n^2 = \frac{1}{n} \sum_i \sum_j \sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \mu - \alpha_i - \beta_j - \delta_{ij} \right)^2$$

$$= \frac{1}{n} \sum_i \sum_j \sum_{l=1}^{n_{ij}} \left(Y_{ijl} - \overline{Y}_{ij\cdot} \right)^2 .$$

$$\sup_{(H)} L(\vec{\alpha}, \vec{\beta}, \vec{\delta}, \mu, \sigma^2) = L\left(\widehat{\mu + \alpha_i + \beta_j + \delta_{ij}}, \hat{\sigma}_n^2 \right) = \left(\frac{1}{\sqrt{2\pi}\, \hat{\sigma}_n} \right)^2 e^{-\frac{n}{2}}$$

Tomorrow ...

- Two-way ANOVA cont’d;
- MANOVA demonstration.