

השוואות מרובות – שיטות נוספות:

נעזר בחומר משקפים של די"ר נירה גלאי

קונטרסטים זה יפה אבל לא מספיק...

- פירוק סכום הריבועים לקונטרסטים מהווה תוצאה יפה. אבל באמצעות פרוק זה לא ניתן לענות על כל השאלות אשר מעניינות אותנו (בעיקר השוואה/דדרוג של הטיפולים).
- בנוסף, לפעמים אנו רוצים להשתמש במידע הנוסף שיש לנו לאחר ביצוע ניתוח שונות (post-hoc) לצורך בחירה של השוואות מעניינות.
- אז נלמד שיטות נוספות...

נלמד הרבה שיטות.....

שיטת LSD

שיטת Bonferroni

שיטת Sidak

שיטת Scheffe

שיטת Tukey

~~שיטת Duncan~~

שיטת Dunnett

~~שיטת Newman-Keuls~~

- עבור כל שיטה יש לשאול:
 - מה השיטה נועדה לעשות?
 - תחת איזה הנחה הניתוח מדויק?
 - מהי רמת המובהקות הכוללת (ניסוי) ומתן רמות המובהקות של כל השוואה והשוואה?
 - האם השיטה מתאימה לניתוח מתוכנן מראש או ל post-hoc?
 - מהי העוצמה (אולי ביחס לשיטות אחרות)?
 - כיצד מבצעים ב SAS?
 - מהו הפלט ב SAS?

נתמקד בהסקה לגבי:

$$\mu_i - \mu_j \quad i, j \in \{1, \dots, a\}$$

חלק מהשיטות אשר נציג כאן מתאימות גם לקונטרסטים כללים יותר, אבל הפרש זוגות התוחלות (דרוג כל התוחלות) הוא המקרה המעניין ביותר וזה אשר מקבלים עבורו פלט ב SAS כברירת מחדל.

LSD שיטת

השוואות בזוגות ברמת מובהקות α

$$T = \frac{\bar{y}_{i\cdot} - \bar{y}_{j\cdot}}{\sqrt{MS_E \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}$$
$$|\bar{y}_{i\cdot} - \bar{y}_{j\cdot}| > t_{1-\alpha/2, N-a} \underbrace{\sqrt{MS_E \left(\frac{1}{n_i} + \frac{1}{n_j} \right)}}_{LSD_{ij}}$$

נחשב את הפרשים של כל זוגות הממוצעים ונשווה ל LSD

פלט עבור דוגמת Tensile Strength

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	2.08596
Least Significant Difference	3.7455

Means with the same letter are not significantly different.			
t Grouping	Mean	N	CottonPercent
A	21.600	5	30
B	17.600	5	25
B	15.400	5	20
C	10.800	5	35
C	9.800	5	15

הערות: הפלט יהיה שונה
בדיוק, יחידות, תדירות

הערות לגבי LSD:

- שיטת Post-Hoc הכי פשוטה שיש.
- השיטה אינה שומרת על רמת המובהקות של הניסוי!!!
- לפעמים ייתכן כי סטטיסטי F בניתוח שונות יהיה מובהק אבל אף השוואה בLSD לא תהייה מובהקת.

שיטת Bonferroni

השוואות מרובות: שיטת בונפרוני

- על-פי מבחן post-hoc זה מתבצעים $\binom{a}{2}$ מבחני t להשוואת תוחלות של

$$\alpha_c = \frac{\alpha_e}{k} = \frac{\alpha_e}{\binom{a}{2}}$$

- אוכלוסיות בלתי תלויות (כמו ב LSD).
■ על-פי תיקון בונפרוני, מתבצע כל מבחן ברמת מובהקות

- סטיסטי המבחן הוא כמו ב LSD.

- בדוגמת הכתנה בשיטת בונפרוני, מתבצע כל מבחן ברמת מובהקות 0.005, עבור רמת מובהקות כוללת השווה ל-0.05.
- את ערכי סטיסטי המבחן (המבוסס על מבחן t) משווים לערך הקריטי של $t_{\alpha_c, N-a}$.

פלט עבור דוגמת Tensile Strength

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of t	3.15340
Minimum Significant Difference	5.6621

Means with the same letter are not significantly different.			
Bon Grouping	Mean	N	CottonPrecent
A	21.600	5	30
A			
B	17.600	5	25
B			
B	15.400	5	20
C			
C	10.800	5	35
C			
C	9.800	5	15

הערות לגבי Bonferroni:

- כאן הצגנו את שיטת Post-Hoc של Bonferroni. חשוב להדגיש כי בניית Post-Hoc צריך לבצע את כל ההשוואות האפשריות. חלופה אחרת הייתה לתכנן מראש מספר השוואות (עם רמת מובהקות המחושבת ע"י Bonferroni).
- בכל מקרה אסור לבצע השוואות בודדות במסגרת Post-Hoc כי כל התיקון של רמת-המובהקות מזדהם.

שיטת Sidak

השוואות מרובות: תיקון סידאק (Sidak)

- תיקון Sidak:
- על-פי מבחן post-hoc זה מתבצעים $\binom{\alpha}{2}$ מבחני t להשוואת תוחלות של

- אוכלוסיות בלתי תלויות.
לפי תיקון סידאק, מתבצע כל מבחן ברמת מובהקות

$$\alpha_c = 1 - (1 - \alpha_e)^{\frac{1}{\binom{\alpha}{2}}} = 1 - (1 - \alpha_e)^{\frac{1}{2}}$$

- סטטיסטי המבחן הוא כמו שראינו קודם.

- בהשוואה לבונפרוני, גישה זו מחמירה פחות ברמת המובהקות של כל השוואה ועדיין שומרת על α_e קבועה.

פלט עבור דוגמת Tensile Strength

CottonPercent	TensileStrength LSMEAN	LSMEAN Number
15	9.8000000	1
20	15.4000000	2
25	17.6000000	3
30	21.6000000	4
35	10.8000000	5

פלט מ Linear Models
(Proc GLM)

Least Squares Means for effect CottonPercent Pr > t for H0: LSMean(i)=LSMean(j)					
Dependent Variable: TensileStrength					
i/j	1	2	3	4	5
1		0.0528	0.0031	<.0001	0.9998
2	0.0528		0.9311	0.0249	0.1711
3	0.0031	0.9311		0.3179	0.0115
4	<.0001	0.0249	0.3179		<.0001
5	0.9998	0.1711	0.0115	<.0001	

אלו הם ערכי p-value מתוקנים לניסוי (ולא להשוואה בודדת)

שיטת Scheffe

Scheffe – השוואה של כל הקונטרסטים במקביל

מ קונטרסטים

$$\Gamma_u = C_{1u}\mu_1 + C_{2u}\mu_2 + \dots + C_{au}\mu_a \quad u = 1, \dots, m$$

מ אמדים לקונטרסטים

$$C_u = C_{1u} \bar{y}_{1\bullet} + C_{2u} \bar{y}_{2\bullet} + \dots + C_{au} \bar{y}_{a\bullet} \quad u = 1, \dots, m$$

מ Standard Errors עבור קונטרסטים

$$S_{C_u} = \sqrt{MS_E \sum_{i=1}^a \frac{C_{iu}^2}{n_i}} \quad u = 1, \dots, m$$

המבחן ורווח הסמך

$$S_{\alpha,u} = S_{C_u} \sqrt{(a-1)F_{\alpha,a-1,N-a}}$$

כלל החלטה:

$$\text{reject if } |C_u| > S_{\alpha,u}$$

$$C_u - S_{\alpha,u} \leq \Gamma_u \leq C_u + S_{\alpha,u}$$

רווח סמך:

פלט עבור דוגמת Tensile Strength

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of F	2.86608
Minimum Significant Difference	6.0796

Means with the same letter are not significantly different.			
Scheffe Grouping	Mean	N	CottonPercent
A	21.600	5	30
A			
B	17.600	5	25
B			
B	15.400	5	20
C			
C	10.800	5	35
C			
C	9.800	5	15

כאן בוצע Scheffe
עבור כל הקונטרסטים
המשווים זוגות של
תוחלות

רווחי סמך של Scheffe

הוספת שורה בקוד של Proc GLM:

means CottonPercent / scheffe cldiff;

Comparisons significant at the 0.05 level are indicated by ***.

CottonPercent Comparison	Difference Between Means	Simultaneous 95% Confidence Limits	
30 - 25	4.000	-2.080 10.080	
30 - 20	6.200	0.120 12.280	***
30 - 35	10.800	4.720 16.880	***
30 - 15	11.800	5.720 17.880	***
25 - 30	-4.000	-10.080 2.080	
25 - 20	2.200	-3.880 8.280	
25 - 35	6.800	0.720 12.880	***
25 - 15	7.800	1.720 13.880	***
20 - 30	-6.200	-12.280 -0.120	***
20 - 25	-2.200	-8.280 3.880	
20 - 35	4.600	-1.480 10.680	
20 - 15	5.600	-0.480 11.680	
35 - 30	-10.800	-16.880 -4.720	***
35 - 25	-6.800	-12.880 -0.720	***
35 - 20	-4.600	-10.680 1.480	
35 - 15	1.000	-5.080 7.080	
15 - 30	-11.800	-17.880 -5.720	***
15 - 25	-7.800	-13.880 -1.720	***
15 - 20	-5.600	-11.680 0.480	
15 - 35	-1.000	-7.080 5.080	

הערות

- העוצמה של שיטת Scheffe נמוכה ביחס לשיטות אחרות (הסיבה היא שהשיטה מאפשרת להשוואת במקביל מספר קונטרסטים כרצוננו).
- השיטה נכונה גם כאשר המדגמים אינם מאוזנים.

שיטת Tukey

שיטת טוקי (Tukey) : Studentized Range

- משתמשים במבחנים אלו בדרך כלל כאשר גודלי המדגם שווים $n_i = n$.
- מניחים כי ההשערות המעניינות הן השוואות בזוגות, כלומר השערות מהצורה

$$H_0 : \mu_i = \mu_j$$

- מבחנים אלו מבוססים על טווח ממוצעי a הקבוצות כאשר הטווח מוגדר כהפרש בין הממוצע הגדול ביותר לנמוך ביותר .

- לטווח הזה , אחרי חלוקה בסטית התקן (כלומר "תיקנון") יש התפלגות שנקראת התפלגות studentized range , המסומנת ב-
- Standard Error :

$$\sqrt{\widehat{Var}(\bar{y}_{i\cdot})} = \sqrt{\frac{MS_E}{n}}$$

שיטת טוקי (Tukey) : studentized range

- בהתפלגות ז', הטווח מגיע מקבוצה של a תצפיות בלתי תלויות מהתפלגות נורמלית עם תוחלת μ ושונות קבועה .
- סטטיסטי המבחן הוא:

$$Q = \frac{\max_i (\bar{y}_i) - \min_i (\bar{y}_i)}{\sqrt{\frac{MSE}{n}}}$$

- נשווה אותו לערך הקריטי מהתפלגות ה- studentized range :
 $q_{1-\alpha, a, N-a}$
- כאשר גודלי המדגם בקבוצות אינו שווה יש תיקונים לשיטת טוקי.
- לשיטת טוקי עוצמה גבוהה עבור השוואות בזוגות

פלט עבור דוגמת Tensile Strength

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of Studentized Range	4.23186
Minimum Significant Difference	5.373

Means with the same letter are not significantly different.			
Tukey Grouping	Mean	N	CottonPercent
A	21.600	5	30
A			
A	17.600	5	25
B			
B	15.400	5	20
C			
C	10.800	5	35
D			
D	9.800	5	15

תיקון כאשר המדגמים אינם מאוזנים

$$n_h = \frac{a}{\sum_{i=1}^a \frac{1}{n_i}}$$

■ הערה:

■ נגדיר ממוצע ההרמוני של a איברים n_1, n_2, \dots, n_a על ידי

■ עבור מבחנים בהם ניתן לבצע את ההשוואות לא רק עבור גודלי מדגם שווים, יש להחליף את n ב- n_h .

שיטת Duncan

שיטת Duncan

- הרעיון במבחן זה הוא לאתר ממוצעים 'מספיק רחוקים' שיעידו על הפרש מובהק בין התוחלות.
- תחילה מושווים שני הממוצעים הרחוקים ביותר.
- אם אין הבדל מובהק ביניהם, אין צורך להשוות ממוצעים קרובים יותר.
- אם יש הבדל מובהק מושווה הממוצע השני בגודלו לממוצע הקטן ביותר ומושווה הממוצע הגדול ביותר לממוצע השני הקטן וכך הלאה.
- כיוון שבכל שלב שכזה, יורד קצה מסוים בהשוואה, הרי שערכי סטיסטי המבחן מושווים בכל שלב לערך אחר של התפלגות Q.

□ בשלב הראשון מושווה ערך סטיסטי המבחן ל-

$$Q(1 - (1 - \alpha)^{k-1}, k, N - k)$$

□ בשלב השני (אם מגיעים אליו) ל- וכך הלאה.

$$Q(1 - (1 - \alpha)^{k-2}, k - 1, N - k)$$

□ הערה: שיטת דאנקן, לא תמיד שומרת על רמת המובהקות הכוללת הרצויה ולכן לא מומלצת לשימוש!!

פלט עבור דוגמת Tensile Strength

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06

Number of Means	2	3	4	5
Critical Range	3.745	3.931	4.050	4.132

Means with the same letter are not significantly different.			
Duncan Grouping	Mean	N	CottonPercent
A	21.600	5	30
B	17.600	5	25
B			
B	15.400	5	20
C	10.800	5	35
C			
C	9.800	5	15

שיטת Dunnett

שיטת Dunnett

- במבחן זה משתמשים כאשר ישנן מספר קבוצות ואחת מהן היא קבוצת ביקורת. רוצים להשוות את הקבוצות רק לקבוצת הביקורת ולא לבין לבין עצמן.
- נניח בדוגמת הכותנה (TensileStrength) שטיפול 15 אחוז כותנה הוא קבוצת ביקורת.

$$|\bar{y}_{1\bullet} - \bar{y}_{j\bullet}| > d_{1-\alpha}(a-1, N-a) \sqrt{MS_E \left(\frac{1}{n_1} + \frac{1}{n_j} \right)}$$

פלט עבור דוגמת Tensile Strength

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06
Critical Value of Dunnett's t	2.65112
Minimum Significant Difference	4.7602

Comparisons significant at the 0.05 level are indicated by ***.		
CottonPrecent Comparison	Difference Between Means	Simultaneous 95% Confidence Limits
30 - 15	11.800	7.040 16.560 ***
25 - 15	7.800	3.040 12.560 ***
20 - 15	5.600	0.840 10.360 ***
35 - 15	1.000	-3.760 5.760

שיטת Newman-Keuls

פלט עבור דוגמת Tensile Strength

Alpha	0.05
Error Degrees of Freedom	20
Error Mean Square	8.06

Number of Means	2	3	4	5
Critical Range	3.7454539	4.5427095	5.0256316	5.3729604

Means with the same letter are not significantly different.

SNK Grouping	Mean	N	CottonPercent
A	21.600	5	30
B	17.600	5	25
B			
B	15.400	5	20
C	10.800	5	35
C			
C	9.800	5	15

סיכום

סיכום

- LSD פוגעת ב רמת המובהקות של הניסוי, שאר השיטות לא (או לא באופן משמעותי). באופן כללי Bonferroni או Sidak עדיפים על LSD, זאת במידה ורוצים לקבע טעות מסוג I ומוכנים לשלם בעוצמה.
- שיטת Scheffe מאפשרת עבודה עם קונטרסטים כללים וגודלי מדגם לא מאוזנים באופן מדויק אבל עוצמת המבחן שלה נמוכה ביחס לשיטות אחרות.
- שיטת Tukey היא הרבה פעמים השיטה המועדפת להשוואה בין זוגות, יש לה עוצמה טובה והיא משמרת את רמת המובהקות, למרות זאת השיטה רגישה למדגמים לא מאוזנים.
- שיטת Dunnett נועדה במיוחד למבחנים בעלי קבוצת ביקורת.