Ch 14 實習 (珊彗)

Agenda

- ANOVA使用條件與情境
- ANOVA檢定步驟
- One way ANOVA
  - 例題1:考計算
  - 例題2: 考表格
- Two way ANOVA
  - Random block ANOVA
  - Two factor ANOVA
- 事後檢定

三、Two way ANOVA 使用情境

	Random block	Two factor
探討變數	只對A變數有興趣，但怕會受 block 影響	對A B都有興趣，而且會想知道有無交互作用
兩變數關係	獨立	相依
變異數組成	SST(total)=SSA+SSB+SSE	SST(total)=SSA+SSB+SSAB+SSE
EXCEL	雙因子變異數分析（無重複實驗）	雙因子變異數分析（重複）

三、Two way ANOVA 資料型態

	Factor A	
Block	1	2
1	6	12
2	9	15

Random Block

Factor A	Factor B	
	1	2
1	6	12
	9	10
	7	11
2	9	15
	10	14
	5	10

Two Factors

3.1 One Random Block ANOVA

Source of Variance	sum of square	df	Mean square	F
Treatment /Between group/ explain variance	SST	k-1	MST=SST/k-1	MST/MSE
Block	SSB	b-1	MST=SSB/b-1	MSB/MSE
Error/within group/ unexplained variance	SSE	n-k-b+1 (k-1)(b-1)	MSE=SSE/n-k-b+1	
Total	SST (toatl)	n-1		

3.1 One Random Block ANOVA

- Formula for the calculation of the sums of

$$SS(Total) = (x_{11} - \overline{\overline{X}})^2 + (x_{21} - \overline{\overline{X}})^2 + \ldots + (x_{12} - \overline{\overline{X}})^2 + (x_{22} - \overline{\overline{X}})^2 + \ldots + (X_{1k} - \overline{\overline{X}})^2 + (x_{2k} - \overline{\overline{X}})^2 + \ldots =$$

	Treatment				
Block	1	2		k	**Block mean**
1	X11	X12	. . .	X1k	$\bar{x}[B]_1$
2	X21	X22		X2k	$\bar{x}[B]_2$
.					
.					
.					
b	Xb1	Xb2		Xbk	
Treatment mean	$\bar{x}[T]_1$	$\bar{x}[T]_2$		$\bar{x}[T]_k$	$\overline{\overline{x}}$

SSB=

$$k\left((\bar{x}[B]_1) - \overline{\overline{X}}\right)^2 +$$

$$k\left((\bar{x}[B]_2) - \overline{\overline{X}}\right)^2 +$$

$$\ldots$$

$$k\left((\bar{x}[B]_k) - \overline{\overline{X}}\right)^2$$

$$SST = b\left((\bar{x}[T]_1) - \overline{\overline{X}}\right)^2 + b\left((\bar{x}[T]_2) - \overline{\overline{X}}\right)^2 + \ldots + b\left((\bar{x}[T]_k) - \overline{\overline{X}}\right)^2$$

3.1 One Random Block ANOVA

$$SSE = (x_{11} - \bar{x}[T]_1 - \bar{x}[B]_1 + \bar{\bar{X}})^2 + (x_{21} - \bar{x}[T]_1 - \bar{x}[B]_2 + \bar{\bar{X}})^2 + \dots$$
$$(x_{12} - \bar{x}[T]_2 - \bar{x}[B]_1 + \bar{\bar{X}})^2 + (x_{22} - \bar{x}[T]_2 - \bar{x}[B]_2 + \bar{\bar{X}})^2 + \dots$$
$$(x_{1k} - \bar{x}[T]_k - \bar{x}[B]_1 + \bar{\bar{X}})^2 + (x_{2k} - \bar{x}[T]_k - \bar{x}[B]_2 + \bar{\bar{X}})^2 + \dots$$

很難算，忽略他吧

	Treatment				
Block	1	2		k	**Block mean**
1	X11	X12	. . .	X1k	$\bar{x}[B]_1$
2	X21	X22		X2k	$\bar{x}[B]_2$
.					
.					
.					
b	Xb1	Xb2		Xbk	
Treatment mean	$\bar{x}[T]_1$	$\bar{x}[T]_2$		$\bar{x}[T]_k$	$\bar{\bar{x}}$

SSB=

$$k\left((\bar{x}[B]_1) - \bar{\bar{X}}\right)^2 +$$
$$k\left((\bar{x}[B]_2) - \bar{\bar{X}}\right)^2 +$$
$$\dots\dots$$
$$k\left((\bar{x}[B]_k) - \bar{\bar{X}}\right)^2$$

$$SST = b\left((\bar{x}[T]_1) - \bar{\bar{X}}\right)^2 + b\left((\bar{x}[T]_2) - \bar{\bar{X}}\right)^2 + \dots + b\left((\bar{x}[T]_k) - \bar{\bar{X}}\right)^2$$

Example 1:類似 1,2,3

How well do diets work. In a preliminary study, 20 people who were more than 50 pounds overweight were recruited to compare four diets. The people were matched by age. The oldest four became block 1, the next oldest four became block 2, and so on. The number of pounds that each person lost is listed in the following table. Can we infer at the 1 % significance level that there are differences between the four diets?

	Diet			
Block	**1**	**2**	**3**	**4**
1	5	2	6	8
2	4	7	8	10
3	6	12	9	2
4	7	11	16	7
5	9	8	15	14

Solution 1

S1:假設檢定：

H_0 : The means of the 2 levels of factor A are equal

H_1 : At least two means differ

H_0 : The means of the 2 levels of Block are equal

H_1 : At least two means differ

Source of Variance	sum of square	df	Mean square	F
Treatment /Between group/ explain variance	SST	k-1	MST=SST/k-1	MST/MSE
Block	SSB	b-1	MST=SSB/b-1	MSB/MSE
Error/within group/ unexplained variance	SSE	n-k-b+1 (k-1)(b-1)	MSE=SSE/n-k-b+1	
Total	SST (toatl)	n-1		

S2:決定 critical point

	F (0.05, v1, v2)
Factor A	F (0.01, k-1, n-k-b+1)=F(0.01,3,12)=5.95
Factor B	F (0.01, b-1, n-k-b+1) =F(0.01,4,12)

S3: 算統計量

SSA/SSB/SSE

BLOCK	factor a 1	2	3	4	平均
1	5	2	6	8	5.25
2	4	7	8	10	7.25
3	6	12	9	2	7.25
4	7	11	16	7	10.25
5	9	8	15	14	11.5
平均	6.2	8	10.8	8.2	8.3

$$SSA = 5*((6.2-8.3)^2+(8-8.3)^2+(10.8-8.3)^2+(8.2-8.3)^2)=53.8$$

$$SSB = 4*((5.25-8.3)^2+(7.25-8.3)^2+(7.25-8.3)^2+(10.25-8.3)^2+(11.5-8.3)^2)=102.2$$

$$SST(total)=[(5-8.3)^2+(4-8.3)^2+(6-8.3)^2+....(14-8.3)^2]=286.2$$

$$SSE = 286.2-53.8-102.2=130.2$$

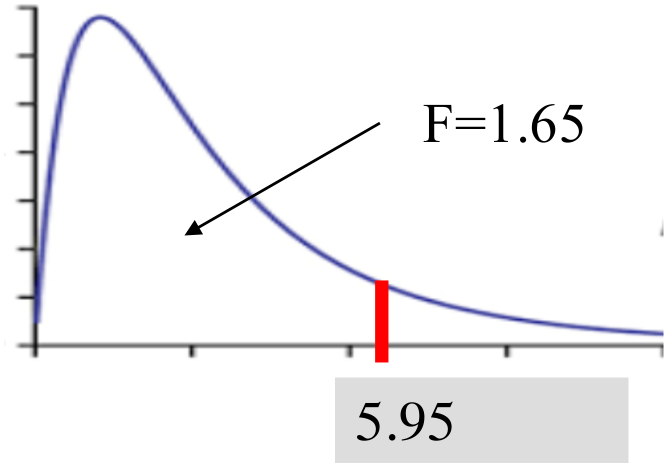
法一、MST/MSE

S4: 下結論

ANOVA Table

Source	Degrees of Freedom	Sum of Squares	Mean Squares	F
Treatments	3	53.8	17.93	1.65
Blocks	4	102.2	25.55	2.35
Error	12	130.2	10.85	
Total	19	286.2		

F=1.65<5.95, There is not enough evidence to conclude there are differences between the four diets.



3.2 Two-way ANOVA (two factors)

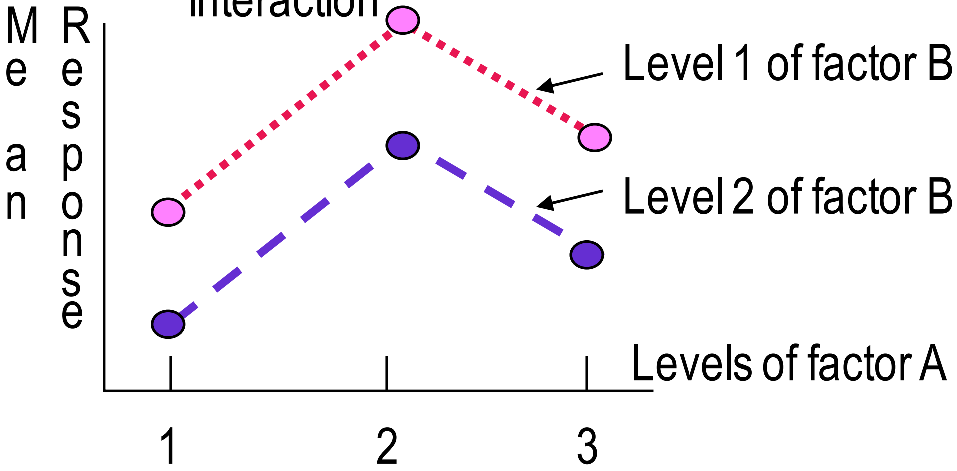
Factor A: Marketing strategy

Factor B: Advertising media

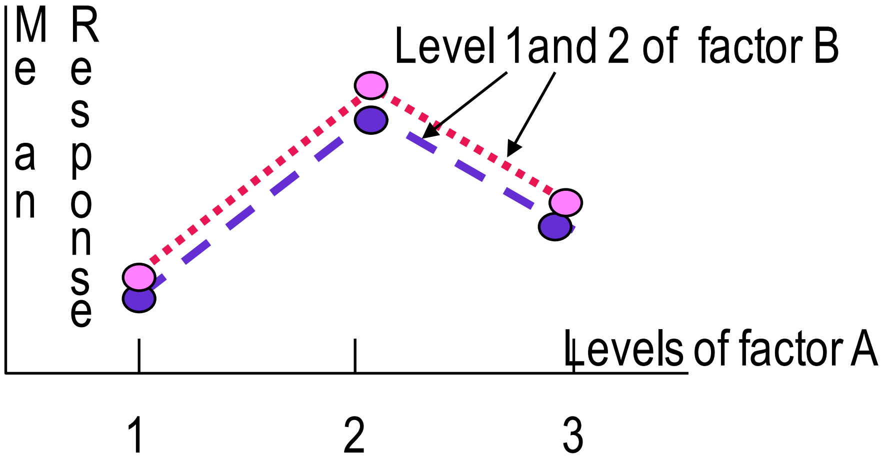
	Convenience	Quality	Price
TV	City 1 sales	City3 sales	City 5 sales
Newspapers	City 2 sales	City 4 sales	City 6 sales

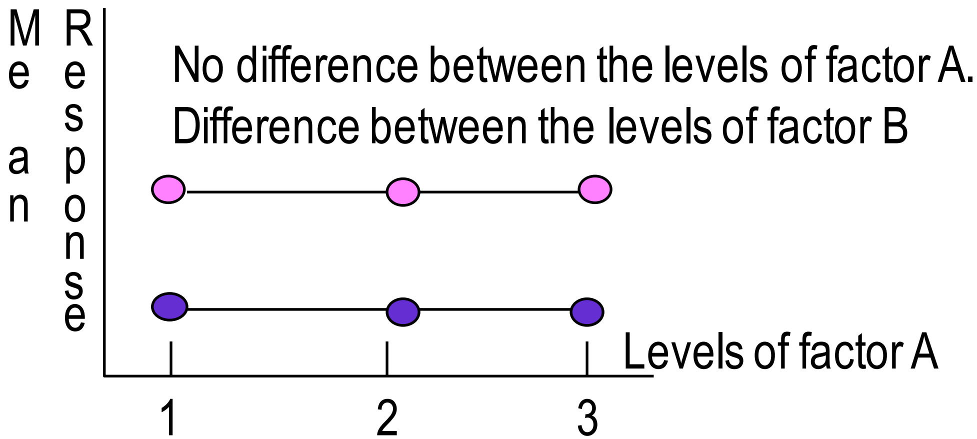
Interaction

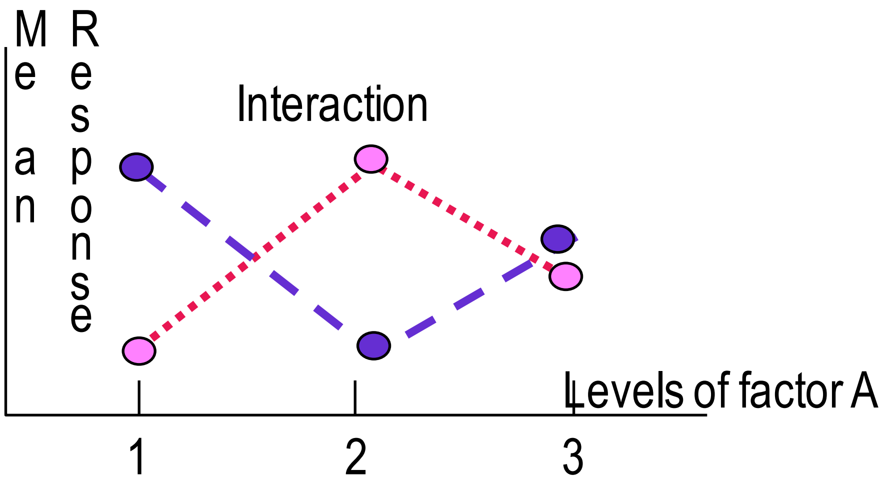
Difference between the levels of factor A, and difference between the levels of factor B; no interaction



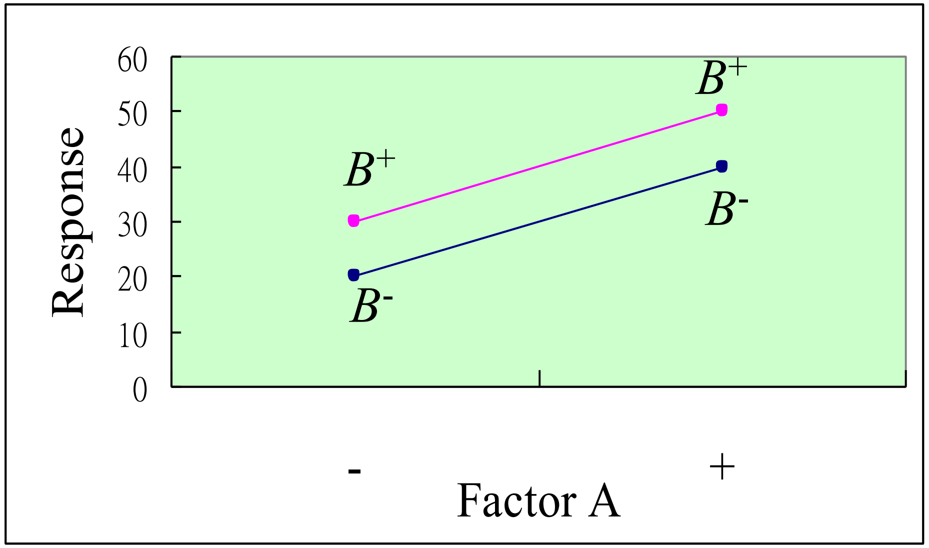
Difference between the levels of factor A
No difference between the levels of factor B



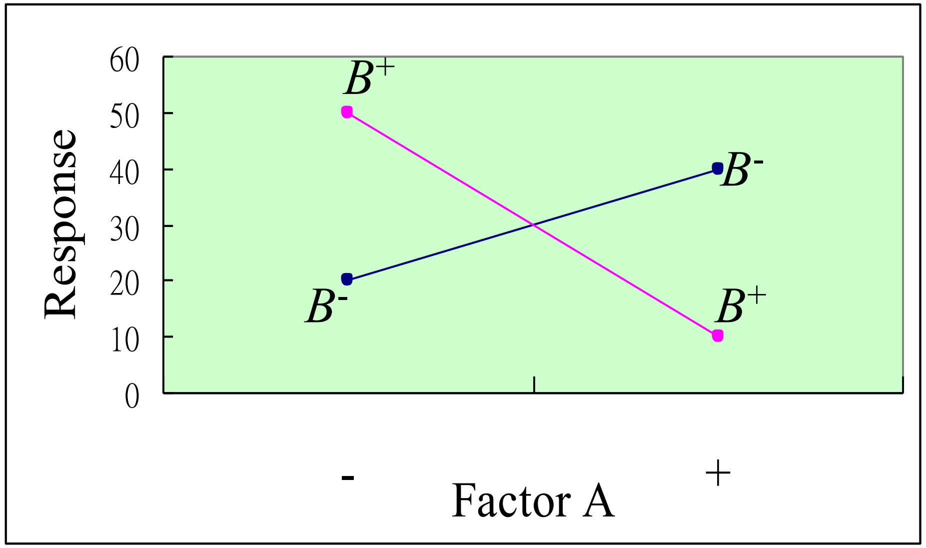




Interaction



Without interaction



With interaction

3.2 Two-way ANOVA (two factors)

Source of Variance	sum of square	df	Mean square	F
Factor A	SST	a-1	MST=SST/a-1	MST/MSE
Factor B	SSB	b-1	MST=SSB/b-1	MSB/MSE
Factor AB	SS(AB)	(a-1)(b-1)	MS(AB)=SS(AB)/[(a-1)(b-1)]	MSAB/MSE
Error/within group/ unexplained variance	SSE	ab(r-1)= n-ab	MSE=SSE/(n-ab)	
Total	SST (toatl)	n-1		

Sums of squares

	Factor A					Sum
Factor B	1	2	$\cdots\cdots\cdots\cdots\cdots$	a		
1	$X_{111}, X_{112}, \cdots X_{11r}$	$X_{121}, X_{122}, \cdots, X_{12r}$	$\cdots\cdots\cdots\cdots\cdots$	$X_{1a1}, X_{1a2}, \cdots, X_{1ar}$	$\overline{X}_{1..}$	$T_{1..}$
2	$X_{211}, X_{212}, \cdots X_{21r}$	$X_{221}, X_{222}, \cdots, X_{22r}$	$\cdots\cdots\cdots\cdots\cdots$	$X_{2a1}, X_{2a2}, \cdots, X_{2ar}$	$\overline{X}_{2..}$	$T_{2..}$
.	r, $\overline{X}[AB]$	.	$\cdots\cdots\cdots\cdots\cdots$	.	a, $\overline{X}[B]$	.
.		.	$\cdots\cdots\cdots\cdots\cdots$	.		.
.		.	$\cdots\cdots\cdots\cdots\cdots$	.		.
$\vdots$	$\vdots$	$\vdots$	$\cdots\cdots\cdots\cdots\cdots$	$\vdots$		$\vdots$
.	.	.	$\cdots\cdots\cdots\cdots\cdots$	.		.
.	.	.	$\cdots\cdots\cdots\cdots\cdots$	.		.
.	.	.	$\cdots\cdots\cdots\cdots\cdots$	.		.
b	$X_{b11}, X_{b12}, \cdots X_{b1r}$	$X_{b21}, X_{b22}, \cdots X_{b2r}$	$\cdots\cdots\cdots\cdots\cdots$	$X_{ba1}, X_{ba2}, \cdots X_{bar}$	$\overline{X}_{b..}$	$T_{b..}$
Sum	$\overline{X}_{.1.}$ $T_{.1.}$	$\overline{X}_{.2.}$ $T_{.2.}$	$\cdots\cdots\cdots\cdots\cdots$ $\cdots\cdots\cdots\cdots\cdots$	$\overline{X}_{.a.}$ $T_{.a.}$	$\overline{\overline{X}}$	$T_{...}$

b, $\overline{X}[A]$

法一、Sums of squares

$$SS(A) = rb\sum_{i=1}^{a}(\bar{x}[A]_i - \bar{\bar{x}})^2 \Rightarrow (10(2)\{(\bar{x}_{conv.} - \bar{\bar{x}})^2 + (\bar{x}_{quality} - \bar{\bar{x}})^2 + (\bar{x}_{price} - \bar{\bar{x}})^2\}$$

$$SS(B) = ra\sum_{j=1}^{b}(\bar{x}[B]_j - \bar{\bar{x}})^2 \Rightarrow (10)(3)\{(\bar{x}_{TV} - \bar{\bar{x}})^2 + (\bar{x}_{Newspaper} - \bar{\bar{x}})^2\}$$

$$SS(AB) = r\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}[AB]_{ij} - \bar{x}[A]_i - \bar{x}[B]_j + \bar{\bar{x}})^2$$

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(x_{ijk} - \bar{x}[AB]_{ij})^2$$

Example 2 (類似4,5,6)

- The following data were generated from a 2 X 2 factorial experiment with 3 replicates.

Factor A	Factor B	
	1	2
1	6	12
	9	10
	7	11
2	9	15
	10	14
	5	10

Ex2 - continued

- a. Test at the 5% significance level to determine whether factors A and B interact.
- b. Test at the 5% significance level to determine whether differences exists between the levels of factor A.
- c. Test at the 5% significance level to determine whether differences exist between the levels of factor B.

Solution 2

S1:假設檢定：

H_0 : Factor A and Factor B do not interact to affect the mean response

H_1 : FactorA and B do interact to affect the mean responses.

H_0 : The means of the 2 levels of factor A are equal

H_1 : At least two means differ

H_0 : The means of the 2 levels of factor B are equal

H_1 : At least two means differ

Source of Variance	sum of square	df	Mean square	F
Factor A	SST	a-1	MST=SST/a-1	MST/MSE
Factor B	SSB	b-1	MST=SSB/b-1	MSB/MSE
Factor AB	SS(AB)	(a-1)(b-1)	MS(AB)=SS(AB)/[(a-1)(b-1)]	MSAB/MSE
Error/within group/ unexplained variance	SSE	n-ab	MSE=SSE/(n-ab)	
Total	SST (toatl)	n-1		

S2:決定 critical point

	F (0.05, v1, v2)
Factor A	F (0.05, a-1, n-ab)＝ F(0.05,1,8)
Factor B	F (0.05, b-1, n-ab) ＝ F(0.05,1,8)
Factor AB	F (0.05, (a-1)(b-1), n-ab)

Solution 2 - continued

F (0.05, a-1, n-ab)

S3: 算統計量

	A		B	C	D	E	F	G
23	ANOVA							
24	*Source of Variation*		*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
25	Sample	SSA	5.33	1	5.33	1.23	0.2995	5.32
26	Columns	SSB	56.33	1	56.33	13.00	0.0069	5.32
27	Interaction	SSAB	1.33	1	1.33	0.31	0.5943	5.32
28	Within	SSE	34.67	8	4.33			
29								
30	Total		97.67	11				

S4: 下結論

- a F = .31, p-value = .5943. There is not enough evidence to conclude that factors A and B interact.
- b F = 1.23, p-value = .2995. There is not enough evidence to conclude that differences exist between the levels of factor A.
- c F = 13.00, p-value = .0069. There is enough evidence to conclude that differences exist between the levels of factor B.

Solution 2 - continued

a= 2, b=2, r=3

Factor A	Factor B	
	1	2
1	6	12
	9	10
	7	11
2	9	15
	10	14
	5	10

	Factor A		
Factor B	1	2	平均
1	6	12	
	9 7.3	10 11	9.2
	7	11	
2	9	15	
	10 8	14 13	10.5
	5	10	
平均	7.7	12.0	9.8

Solution 2 - continued

$$SS(A) = rb\sum_{i=1}^{a}(\bar{x}[A]_i - \bar{\bar{x}})^2 \Rightarrow (3(2)\{(7.7-9.8)^2+(12-9.8)^2\} = 55.5$$

$$SS(B) = ra\sum_{j=1}^{b}(\bar{x}[B]_j - \bar{\bar{x}})^2 \Rightarrow (3)(2)\{(9.2-9.8)^2+(10.5-9.8)^2\} = 5.1$$

	Factor A		
Factor B	1	2	平均
1	6	12	
	9 7.3	10 11	9.2
	7	11	
2	9	15	
	10 8	14 13	10.5
	5	10	
平均	7.7	12.0	9.8

Solution 2 - continued

	Factor A		
Factor B	1	2	平均
1	6	12	
	9 7.3	10 11	9.2
	7	11	
2	9	15	
	10 8	14 13	10.5
	5	10	
平均	7.7	12.0	9.8

(7.3-9.8)=(9.2-9.8)+(7.7-9.8)
=7.3-9.8-9.2-7.7+9.8+9.8
=7.3-9.2-7.7+9.8

$$SSAB = (3(\{(7.3-7.7-9.2+9.8)^2 + (8-7.7-10.5+9.8)^2 + (11-12-9.2+9.8)^2 + (13-12-10.5+9.8)^2\} = 1.35$$

$$SSE = (\{(6-7.3)^2 + (9-7.3)^2 + \ldots\ldots$$

$$SS(AB) = r\sum_{i=1}^{a}\sum_{j=1}^{b}(\bar{x}[AB]_{ij} - \bar{x}[A]_i - \bar{x}[B]_j + \bar{\bar{x}})^2$$

$$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{r}(x_{ijk} - \bar{x}[AB]_{ij})^2 =$$

(a)

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F critical*
Factor A	5.333	1	5.333	1.231	0.2995	5.318
Factor B	56.333	1	56.333	13.00	0.0069	5.318
Interaction	1.333	1	1.333	0.308	0.5943	5.318
Error	34.667	8	4.333			
Total	97.667	11				

(b) H_0: Factors A and B do not interact

H_1: Factors A and B do interact.

Conclusion: $F = 0.308 < 5.318$. do not reject

(c) H_0: No difference among the means of the two levels of factor A

H_1: the two means differ

Conclusion: $F = 1.231 < 5.318$.

Don't reject the null hypothesis. No, differences do not exist among the levels of factor A (injection procedures), according to this data.

Example 3

- Provide an example for a randomized block design with three treatments ($k = 3$) and four blocks ($b = 4$), in which SST is equal to zero and SSB and SSE are not equal to zero.

Treatment

Block	1	2	3
1			
2			
3			
4			

Example 3

- Provide an example for a randomized block design with three treatments (k = 3) and four blocks (b = 4), in which SST is equal to zero and SSB and SSE are not equal to zero.

Treatment

Block	1	2	3
1	2	3	1
2	3	4	3
3	5	4	6
4	5	4	5