IV-3-2. F test

F test is used in judgment whether fluctuation of data can be explained by factors. When there is as sub samples in a dataset divided by level of factors, we use F test for judgement whether there are differences among the levels in each factor. More specifically, when we rear small fish in different aquarium of which water temperatures are controlled at different level for comparison of growth, we compare the variance of fluctuation among different aquarium with variance of fluctuations which not include influence in temperature. For this, we have to separate variances caused by difference of temperature and unexplainable variance from total variance, and compare the variance caused by difference of temperature with unexplainable variance using F ratio. When F ration is larger than a value of threshold, we can conclude that the growth is affected by temperature.

Theoretically we can increase factors abundantly, though when we increase factors we have to consider interaction among factors and interpretation of the result becomes complicated. Actual dataset includes replications, nested structure and hierarchical structure. We should consider the model of analysis depending on the structures of dataset. It is difficult to generalize analytical model for the author. Here the author explains the method to separate the variances in the case of one factor with replications (One way ANOVA, analysis of variance), the case of two factors without replications (two way ANOVA)、and the case of two factors with replications.

The author already explained the concept and theory of separation of variances in chapter IV-2-3. Here the author explain only procedure.

IV-3-2-1. One-way ANOVA

Table 22 is an example of dataset which includes sub population. Sub population 1 includes 6 data, sub population 2 includes 7 data, sub population 3 includes 5 data, and sub population 4 includes 6 data. We discuss existence of significant differences among sub populations.

Table 22. Example of dataset which includes sub population.

Sub population	1	2	3	4
	2	10	8	9
	5	2	7	15
	3	4	3	8
	8	9	4	12
	9	13	5	13

	4	14		4
		15		

Procedure

1. Calculation of total sum of square (SS_{total})
2. Calculation of average in each sub population ($\bar{x}_i$)
3. Calculation of sum of square of residual ($SS_{residual}$)
4. Calculation of sum of square among sub populations by deducting $SS_{residual}$ from SS_{total}. ($SS_{factor} = SS_{total} - SS_{residual}$)
5. Calculation of degree of freedom of residual by deducting degree of freedom among factors from total degree of freedom ($df_{residual} = df_{total} - df_{factor}$)。
6. Calculation of variance of residual by dividing sum of square of residual by degree of freedom of residual ($\sigma^2{}_{residula} = \frac{SS_{residual}}{df_{residual}}$)
7. Calculation of variance of factor by dividing sum of square of factor by degree of freedom of factor ($\sigma^2{}_{factor} = \frac{SS_{factor}}{df_{factor}}$)
8. Calculation of F ration by dividing variance of factor by variance of residual ($F = \frac{\sigma^2{}_{factor}}{\sigma^2{}_{residula}}$)
9. Select risk rate and check the threshold value at $\frac{df_{factor}}{df_{residual}}$.
10. Compare the value of threshold and observed F.

Calculation

Sub population	1	2	3	4	sum	
	2	10	8	9		
	5	2	7	15		
	3	4	3	8		
	8	9	4	12		
	9	13	5	13		
	4	14		4		
		15				
n_i	6	7	5	6	24	N
T_i	31	67	27	61	186	T
$\bar{x}_i$	5.16666	9.25	5.4	10.16667		

	7					
S_i	199	791	163	699	1852	S
$\frac{T_i^2}{n_i}$	160.1667	641.2857	145.8	620.1667	1567.419	$\sum \frac{T_i^2}{n_i}$

n_i: numbe or data i sub population i

$T_i: sum\ of\ data\ in\ sab\ population\ i$

$\bar{x}_i$: average of data in subpopulation i

S_i: sum of square of $data\ in\ sub\ population\ i$: $\sum_{j=1}^{n_j} {x_j}^2$

For calculation of SS, it is convenient and easy to use following formula which learned in chapter IV-2-3

$$SS = \sum_{i=1}^{n} (x_i - \bar{x})^2 = S - \frac{T^2}{n}$$

Proof of the formula

$$\sum_{i=1}^{n} (x_i - \bar{x})^2 = \sum_{i=1}^{n} {x_i}^2 - 2\sum_{i=1}^{n} x_i\bar{x} + \sum_{i=1}^{n} \bar{x}^2 = \sum_{i=1}^{n} {x_i}^2 - 2\bar{x}\sum_{i=1}^{n} x_i + n\bar{x}^2$$

$$= \sum_{i=1}^{n} {x_i}^2 - \frac{(\sum_{i=1}^{n} x_i)^2}{n}$$

$$\because \quad \bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$$

Here,

$$S_i = \sum_{j=1}^{n_i} {x_{ij}}^2$$

$$SS_i = S_i - \frac{{T_i}^2}{n_i}$$

$SS_{residual}$ is sum of SS_i

$$SS_{residual} = SS_1 + SS_2 + \cdots + SS_1$$

$$= S_1 - \frac{{T_1}^2}{n_1} + S_2 - \frac{{T_2}^2}{n_2} + \cdots + S_i - \frac{{T_i}^2}{n_i} + \cdots + S_n - \frac{{T_n}^2}{n_n}$$

$$= S_1 + S_2 + \cdots + S_i + \cdots + S_n - \left(\frac{{T_1}^2}{n_1} + \frac{{T_2}^2}{n_2} + \cdots + \frac{{T_i}^2}{n_i} + \cdots + \frac{{T_n}^2}{n_n}\right)$$

$$S = S_1 + S_2 + \cdots + S_i + \cdots + S_n$$

$$\therefore SS_{residual} = S - \left(\frac{{T_1}^2}{n_1} + \frac{{T_2}^2}{n_2} + \cdots + \frac{{T_i}^2}{n_i} + \cdots + \frac{{T_n}^2}{n_n}\right)$$

On the other hand, total SS_{total} is as follow.

$$\mathrm{SS}_{total} = \mathrm{S} - \frac{T^2}{n_{total}}$$

And,

$$\mathrm{SS}_{total} = \mathrm{SS}_{factor} + \mathrm{SS}_{residual}$$

$$\mathrm{SS}_{factor} = \mathrm{SS}_{total} - \mathrm{SS}_{residual}$$

$$= \mathrm{S} - \frac{T^2}{n} - \left\{S - \left(\frac{{T_1}^2}{n_1} + \frac{{T_2}^2}{n_2} + \cdots + \frac{{T_i}^2}{n_i} + \cdots + \frac{{T_n}^2}{n_n}\right)\right\}$$

$$= \left(\frac{{T_1}^2}{n_1} + \frac{{T_2}^2}{n_2} + \cdots + \frac{{T_i}^2}{n_i} + \cdots + \frac{{T_n}^2}{n_n}\right) - \frac{T^2}{n}$$

$$= \sum_{i=1}^{N} \frac{{T_i}^2}{n_i} - \frac{T^2}{n_{total}}$$

Specific calculation

$n_{total} = 24$

$\mathrm{T} = 186$

$\mathrm{S} = 1852$

$$\sum_{i=1}^{4} \frac{{T_i}^2}{n_i} = 1567.419$$

$$SS_{total} = 1852 - \frac{186^2}{24} = 1852 - 1441.5 = 410.5$$

$$SS_{residual} = 1852 - 1567.419 = 284.581$$

$$SS_{factor} = 1567.419 - \frac{186^2}{24} = 1567.419 - 1441.5 = 125.919$$

$\mathrm{df}_{total} = 24 - 1 = 23$

$\mathrm{df}_{factor} = 4 - 1 = 3$

$\mathrm{df}_{factor} = 23 - 3 = 20$

${\sigma_{factor}}^2 = \frac{125.919}{3} = 41.973$

${\sigma_{residual}}^2 = \frac{284.581}{20} = 14.22905$

F ratio$=\frac{{\sigma_{factor}}^2}{{\sigma_{residual}}^2} = \frac{41.973}{14.22905} \approx 2.9498$

Threshold of F in F distribution table (p=0.05) at $\mathrm{df}_{numerator} = 3,\ \mathrm{df}_{denominator} = 20$

is 3.0984

Observed $F = 2.9498 < 3.0983$

We cannot conclude that there is different among sub populations. The result of F test is expressed by following analysis variance table

Table 23. An example of one-way analysis variance table

Source	Sum of square (SS)	degree of freedom (df)	mean square ($MS = \sigma^2$)	$F_{observed}$	p	$F_{threshold}$
among population	125.919	3	41.973	2.2498	0.05	3.0983
residual	284.581	20	14.20995			

IV-3-2-2. Two-way ANOVA without replication

We expand the procedure to the case when a data includes two factors. More specifically, we consider a rearing experiment of fish. There are 12 combinations of two rearing conditions. One condition is type of feed and the other is shape of aquarium. There are 3 types of feed, and 4 shapes of aquarium. One aquarium is used for one combination and 3 fish are reared in a aquarium. We compare the growth rate among 12 combinations of rearing conditions, and we will discuss the impact of type of feeding rate and shape of aquarium. In this example, the levels of conditions are qualitative variables, though the levels is not necessary to be qualitative. They can be feeding rate, water temperature, light condition, oxygen supply and so on. At first, we simplify the data by averaging the growth rate of fish in each aquarium. This is two way ANOVA without replication.

Table 24. Example of dataset for two-way ANOVA without replication.

	A_1	A_2	A_3
B_1	11	11	8
B_2	10	13	19
B_3	9	18	18
B_4	14	18	19

A: type of feed

B: shape of aquarium

We already know the method of separation of variance (See IV-2-3). The author shows process of the calculation

Calculation process

	A_1	A_2	A_3	n	T_i	S_i	$\sum_{i=1}^{n} \frac{{T_i}^2}{n}$	$S_i - \sum_{i=1}^{n} \frac{{T_i}^2}{n}$
B_1	11	11	8	3	30	306	300	6
B_2	10	13	19	3	42	630	588	42
B_3	9	18	18	3	45	729	675	54
B_4	14	18	19	3	51	881	867	14
n	4	4	4	12	168	2546	2352	116
T_j	44	60	64					
S_j	498	938	1110					
$\sum_{j=1}^{n} \frac{{T_j}^2}{n}$	484	900	1024					
$S_j - \sum_{j=1}^{n} \frac{{T_j}^2}{n}$	14	38	86	138		194		

There are 3 origin of data fluctuation, namely, A, B and interaction

$$SS_{total} = 194$$

$$4SS_A = 194 - 138 = 56$$

$$3SS_B = 194 - 116 = 78$$

$$SS_{interaction} = 194 - 56 - 78 = 60$$

$$SS_A = \frac{56}{4} = 14$$

$$SS_B = \frac{78}{3} = 26$$

Table 25. An example of Two wah analysis of variance table without replication.

Source	SS	df	MS	Observed F	p	threshold F
A	16	2	7	0.7	0.05	5.1433
B	26	3	8.6667	0.867	0.05	4.7571
Interaction	60	6	10			
Sum	194	11				

We cannot conclude that there are differences among both condition A and B.

Then we consider the impact of replication.

IV-3-2-3. Two-way ANOVA with replication

Table 26. An example of dataset include replication

	A_1	A_2	A_3
B_1	10	8	6
	11	12	8
	12	13	10
B_2	9	12	18
	9	12	19
	12	15	20
B_3	8	15	17
	9	19	18
	10	20	19
B_4	13	14	18
	13	19	19
	16	21	20

This example dataset includes 3 replications in each combination of two conditions. We can discuss the impacts of 2 conditions by two-way ANOVA without replication after calculation of average in each cell of combined conditions. However, sensitivity of statistical analysis increases with increase of replication. For this reason, it is better to implement statistical analysis remaining each original data without averaging.

Here, we consider degree of freedom when there are m levels in condition A, n levels in condition B and l replications in each combinations of condition A and B.

Total degree of freedom: $\mathrm{df}_{total} = mnl - 1$

Degree of freedom among A: $\mathrm{df}_A = m - 1$

Degree of freedom among B: $\mathrm{df}_B = n - 1$

Degree of freedom of interaction: $\mathrm{df}_{interaction} = (m-1)(n-1)$

Degree of freedom of residual: $\mathrm{df}_{residual} = mn(l-1)$

$$\mathrm{df}_A + \mathrm{df}_B + \mathrm{df}_{interaction} + \mathrm{df}_{residual} = m - 1 + n - 1 + (m-1)(n-1) + mn(l-1)$$
$$= mnl - 1 = \mathrm{df}_{total}$$

There various calculation procedures. An example is as follow.

	A_1	A_2	A_3	Sum	SS
B_1	10	8	6		
	11	12	8		
	12	13	10		
n	3	3	3	9	42
T	33	33	24	90	
S	365	377	200	942	
T^2/n	363	363	192		
SS_{ij}	2	14	8	24	
B_2	9	12	18		
	9	12	19		
	12	15	20		
n	3	3	3	9	140
T	30	39	57	126	
S	306	513	1085	1904	
T^2/n	300	507	1083		
SS_{ij}	6	6	2	14	
B_3	8	15	17		
	9	19	18		
	10	20	19		
n	3	3	3	9	180
T	27	54	54	135	
S	245	986	974	2205	
T^2/n	243	972	972		
SS_{ij}	2	14	2	18	
B_4	13	14	18		
	13	19	19		
	16	21	20		
n	3	3	3	9	76
T	42	54	57	153	
S	594	998	1085	2677	
T^2/n	588	972	1083		
SS_{ij}	6	26	2	34	
T	132	180	192	T_{total}	504
n	12	12	12	n_{total}	36
S	1510	2874	3344	S_{total}	7728
SS	58	174	272	SS_{total}	672

$$\mathrm{SS}_{total} = 4 \times 3 \times \mathrm{SS}_A + 3 \times 3 \times \mathrm{SS}_B + 3 \times \mathrm{SS}_{interaction} + \mathrm{SS}_{residual}$$

$$SS_{residual}$$

$$= \text{sum of sum of squares of replications in each cell (combnation of cndtions)}$$

$$= \sum_{i=1}^{4} \sum_{j=1}^{3} ss_{ij} = 24 + 14 + 18 + 34 = 90$$

Sum of SS in the Row: $\mathrm{SS}_{tota} \quad = 42 + 140 + 180 + 76 = 438$

$3 \times 3 \times \mathrm{SS}_B = \mathrm{SS}_{total} - \mathrm{SS}_{total-B} = 672 - 438 = 234$

Sum of SS in the columns: $\mathrm{SS}_{total-A} = 58 + 174 + 272 = 504$

$4 \times 3 \times \mathrm{SS}_A = \mathrm{SS}_{total} - \mathrm{SS}_{total-A} = 672 - 504 = 168$

$$3 \times \mathrm{SS}_{interaction} = \mathrm{SS}_{total} - 4 \times 3 \times \mathrm{SS}_A - 3 \times 3 \times \mathrm{SS}_B - \mathrm{SS}_{residual}$$
$$= 672 - 168 - 234 - 90 = 180$$

$$\mathrm{SS}_B = \frac{234}{9} = 26$$

$$\mathrm{SS}_A = \frac{168}{12} = 14$$

$$\mathrm{SS}_{interaction} = \frac{180}{3} = 60$$

Table 27, Summary of analysis of variance with replication.

Source	SS	df	MS	$F_{0bserved}{}^1$	$F_{threshold}{}^1$	$F_{0bserved}{}^2$	$F_{threshold}{}^2$
A	14	2	7	1.40625	3.4	0.7	4.1
B	26	3	8.6667	2.31112	3.0	0.8667	3.7
Interaction	60	6	10	2.6667	2.5		
residual	90	24	3.75				
Sum		35					

$p \leq 0.05$

$$F^1: \frac{\sigma_{factor}{}^2}{\sigma_{residual}{}^2},\ F^2: \frac{\sigma_{factor}{}^2}{\sigma_{interaction}{}^2}$$

There two F ratios. F^1 is comparison with residual. F^2 is comparison with interaction., we cannot conclude that A or B has significant impact to the values both from F^1 and F^2. In this case, the conclusions from F^1 and F^2 are similar, though conclusions from F^1 and F^2 are contradictory in several cases. We have to consider which conclusion we should accept. Before this discussion, the author explains the function and meaning of interaction and residual in next paragraph.

IV-3-2-3. Interaction and residual

We implement unrealistic thought experiment for clarification of function and meaning of interaction and residual.

Table 28 is non-randomly made dataset for thought experiment. Differences between A1 and A2 are the same among B. Similarly, differences between A2 and A3 are the same among B. The relation of A among B is similarly as the relation of B among A.

Table 28. Dataset without fluctuation sourcing residual and interaction

	A1	A2	A3
B1	2	6	7
B2	6	10	11
B3	7	11	12
B4	9	13	14

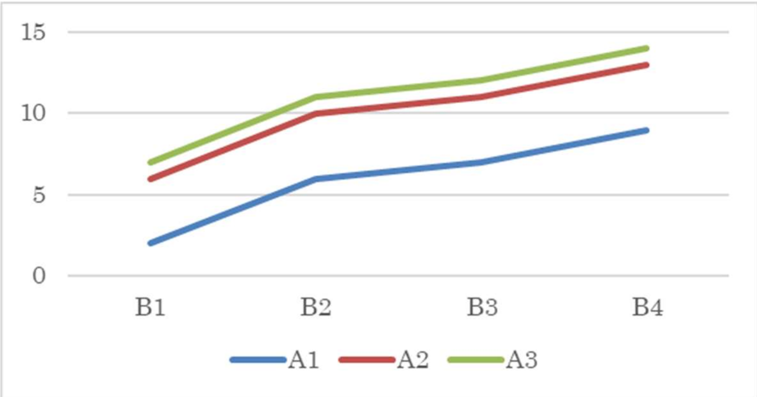

Fig.33. 2-dimensional plots of table 28. We can understand the data changes in parallel relations from Fig.33.

Calculation of F ratio

	A_1	A_2	A_3	n	T_i	S_i	$\sum_{i=1}^{n}\frac{T_i^2}{n}$	$S_i-\sum_{i=1}^{n}\frac{T_i^2}{n}$
B_1	2	6	7	3	15	89	75	14
B_2	6	10	11	3	27	257	243	14
B_3	7	11	12	3	30	314	300	14
B_4	9	13	14	3	36	446	432	14
n	4	4	4	12	108		1050	56
T_j	24	40	44			total S		
S_j	170	426	510			1106		
$\sum_{j=1}^{n}\frac{T_j^2}{n}$	144	400	484	1028	$S_{total}-\frac{T_{total}^2}{n_{total}}$			
$S_j-\sum_{j=1}^{n}\frac{T_j^2}{n}$	26	26	26	78		134		

$$\mathrm{SS}_{total} = 134$$

$$\mathrm{SS}_{total} = 4\mathrm{SS}_A + 3\mathrm{SS}_B + \mathrm{SS}_{interaction}$$

$$4\mathrm{SS}_A = \mathrm{SS}_{total} - (3\mathrm{SS}_B + \mathrm{SS}_{interaction} + \mathrm{SS}_{residual}) = 134 - 78 = 56$$

$$3\mathrm{SS}_B = \mathrm{SS}_{total} - (4\mathrm{SS}_A + \mathrm{SS}_{interaction} + \mathrm{SS}_{residual}) = 134 - 56 = 78$$

$$\mathrm{SS}_A = \frac{56}{4} = 14$$

$$\mathrm{SS}_B = \frac{78}{3} = 26$$

$$\mathrm{SS}_{interaction} = 0$$

Result of ANOVA

Source	SS	df	MS	Observed F	p	threshold F
A	14	2	7	*		
B	26	3	8.6667	*		
Interaction	0	6	0			
Sum	134	11				

In this case, we could not obtain F ratio, because we cannot perform division when dominator is 0. This is thought experiment and the dataset is not realistic. One possible handling is to consider F is infinite. The author considers that this thought experiment is originally unrealistic and absolutely no fluctuation data are significant from the start.

We should make fluctuation of values in the dataset of Table 28 by adding small amount of value to data in each cell to make interaction without influence to variance in A and B. We add small values in Table 29. Sum of adding amounts in B,B,B,*and* B are the same (1.5), and sum of adding amounts in A,A,*and* A are the same (2.0) in order not to make changes in SS and SS.

Table 29. Additional dataset to make interaction

	A_1	A_2	A_3
B_1	0.9	0.5	0.1
B_2	0.4	0.3	0.8
B_3	0.2	0.7	0.6
B_4	0.5	0.5	0.5

Table 30. Revised data by adding additional dataset to dataset in Table 28.

	A1	A2	A3
B1	2.9	6.5	7.1
B2	6.4	10.3	11.8
B3	7.2	11.7	12.6
B4	9.5	13.5	14.5

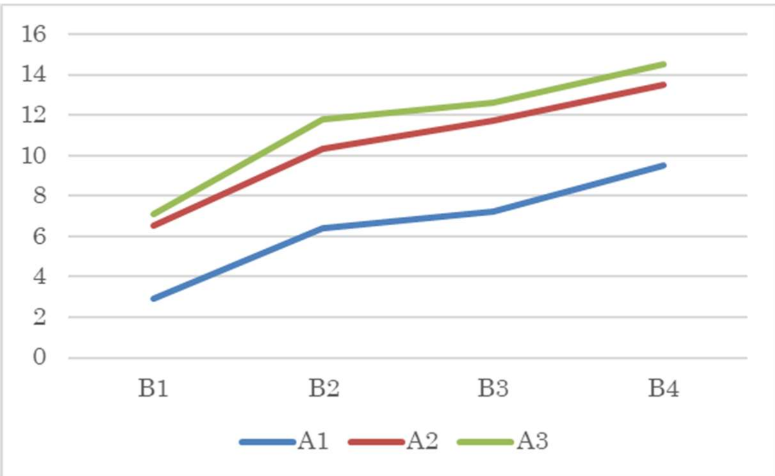

Fig. 34. 2-dimensional plots of table 30.

Calculation

	A_1	A_2	A_3	n	T_i	S_i	$\sum_{i=1}^{n}\frac{{T_i}^2}{n}$	$S_i-\sum_{i=1}^{n}\frac{{T_i}^2}{n}$
B_1	2.9	6.5	7.1	3	16.5	101.07	90.75	10.32
B_2	6.4	10.3	11.8	3	28.5	286.29	270.75	15.54
B_3	7.2	11.7	12.6	3	31.5	347.49	330.75	16.74
B_4	9.5	13.5	14.5	3	37.5	482.75	468.75	14
n	4	4	4	12	114	1217.6	1161	56.6
T_j	26	42	46					
S_j	191.46	467.48	558.66	1217.6				
$\sum_{j=1}^{n}\frac{{T_j}^2}{n}$	169	441	529	1139		134.6		
	22.46	26.48	29.66	78.6				

$$SS_{total} = 134.6$$

$$SS_{total} = 4SS_A + 3SS_B + SS_{interaction}$$

$$4SS_A = SS_{total} - (3SS_B + SS_{interaction} + SS_{residual}) = 134.6 - 78.6 = 56$$

$$3SS_B = SS_{total} - (4SS_A + SS_{interaction} + SS_{residual}) = 134.6 - 56.6 = 78$$

$$SS_A = \frac{56}{4} = 14$$

$$SS_B = \frac{78}{3} = 26$$

$$SS_{interaction} = SS_{total} - (4SS_A + 3SS_B) = 134.6 - 56 - 78 = 0.6$$

Source	SS	df	MS	Observed F	p	threshold F (P=0.005)
A	14	2	7	70		14.554
B	26	3	8.6667	86.667		12.917
Interaction	0.6	6	0.1			
Sum	134.6	11				

From this we can conclude the differences both A and B are significant. We increase the adding values 10 times as in Table 31. Fig.33.

Table 31. Adding value to make 10 times interaction

	A_1	A_2	A_3
B_1	9	5	1
B_2	4	3	8
B_3	2	7	6
B_4	5	5	5

Table 24. Example of dataset for two-way ANOVA without replication.

	A_1	A_2	A_3
B_1	11	11	8
B_2	10	13	19
B_3	9	18	18
B_4	14	18	19

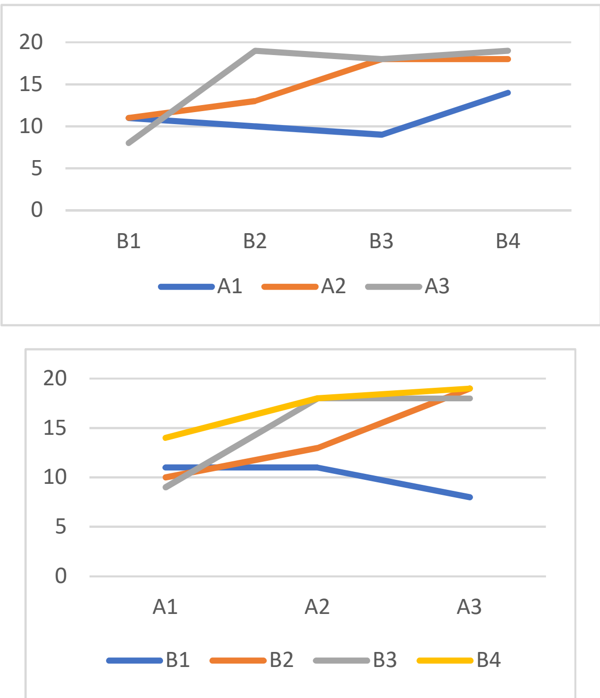

Fig. 35. 2-dimensional plots of table 24.

Calculation

	A_1	A_2	A_3	n	T_i	S_i	$\sum_{i=1}^{n}\frac{T_i^2}{n}$	$S_i-\sum_{i=1}^{n}\frac{T_i^2}{n}$
B_1	11	11	8	3	30	306	300	6
B_2	10	13	19	3	42	630	588	42
B_3	9	18	18	3	45	729	675	54
B_4	14	18	19	3	51	881	867	14
n	4	4	4	12	168	2546	2430	116
T_j	44	60	64					
S_j	498	938	1110	2546				
$\sum_{j=1}^{n}\frac{T_j^2}{n}$	484	900	1024	2408		194		
	14	38	86	138				

$$SS_{total} = 194$$

$$SS_{total} = 4SS_A + 3SS_B + SS_{interaction}$$

$$4SS_A = SS_{total} - (3SS_B + SS_{interaction} + SS_{residual}) = 194 - 138 = 56$$

$$3SS_B = SS_{total} - (4SS_A + SS_{interaction} + SS_{residual}) = 194 - 116 = 78$$

$$SS_A = \frac{56}{4} = 14$$

$$SS_B = \frac{78}{3} = 26$$

$$SS_{interaction} = SS_{total} - (4SS_A + 3SS_B) = 194 - 56 - 78 = 60$$

Result of ANOVA

Source	SS	df	MS	Observed F	p	threshold F (P=0.05)
A	16	2	7	0.7		5.1433
B	26	3	8.6667	0.86667		4.7571
Interaction	60	6	10			
Sum	194	11				

We can conclude that the differences in A and B is not significant, though we cannot say that the differences are insignificant because of large variance of interaction, because variance of interaction in the table of result of ANOVA possibly include random fluctuation. Thus, we should separate effect of interaction and random fluctuation expressed as residual by making replication.

We add dataset shown in table 32 to dataset of table 30. The sums of adding values in a cell are similar among the cells not to make changes in variance of A and B.

Table 32. Dataset to make residual

	A_1	A_2	A_3
B_1	0.9	0.4	0.2
	0.5	0.3	0.7
	0.1	0.8	0.6
B_2	0.9	0.5	0.1
	0.4	0.3	0.8
	0.2	0.7	0.6
B_3	0.5	0.5	0.6
	0.4	0.5	0.3
	0.6	0.5	0.6
B_4	0.6	0.4	0.4
	0.4	0.4	0.2
	0.5	0.7	0.9

Table 33. Dataset in table 30 + dataset in table 32

	A_1	A_2	A_3
B_1	3.8	6.9	7.3
	3.4	6.8	7.8
	3.0	7.3	7.7
B_2	7.3	10.8	11.9
	6.8	10.6	12.6
	6.6	11.0	12.4
B_3	7.7	12.2	13.2
	7.6	12.2	12.9
	7.8	12.2	13.2
B_4	10.1	13.9	14.9
	9.9	13.9	14.7
	10.0	14.2	15.4

Calculation

	A_1	A_2	A_3	Sum	SS
B_1	3.8	6.9	7.3		31.56
	3.4	6.8	7.8		
	3	7.3	7.7		
n	3	3	3	9	
T	10.2	21	22.8	54	
S	35	147.14	173.42	355.56	
T^2/n	34.68	147	173.28		
SS_{ij}	0.32	0.14	0.14	0.6	
B_2	7.3	10.8	11.9		47.22
	6.8	10.6	12.6		
	6.6	11	12.4		
n	3	3	3	9	
T	20.7	32.4	36.9	90	
S	143.09	350	454.13	947.22	
T^2/n	142.83	349.92	453.87		
SS_{ij}	0.26	0.08	0.26	0.6	
B_3	7.7	12.2	13.2		50.3
	7.6	12.2	12.9		
	7.8	12.2	13.2		
n	3	3	3	9	
T	23.1	36.6	39.3	99	
S	177.89	446.52	514.89	1139.3	
T^2/n	177.87	446.52	514.83		
SS_{ij}	0.02	0	0.06	0.08	
B_4	10.1	13.9	14.9		42.34
	9.9	13.9	14.7		
	10	14.2	15.4		
n	3	3	3	9	
T	30	42	45	117	
S	300.02	588.06	675.26	1563.34	
T^2/n	300	588	675		
SS_{ij}	0.02	0.06	0.26	0.34	
T	84	132	144	T_{total}	360
n	12	12	12	n_{total}	36
S	656	1531.72	1817.7	S_{total}	4005.42
SS	68	79.72	89.7	SS_{total}	405.42

$$SS_{total} = 4 \times 3 \times SS_A + 3 \times 3 \times SS_B + 3 \times SS_{interaction} + SS_{residual} = 405.42$$

$$SS_{residual}$$

$$= \text{sum of sum of squares of replications in each cell (combnation of cndtions)}$$

$$= \sum_{i=1}^{4}\sum_{j=1}^{3} ss_{ij} = 0.6 + 0.6 + 0.08 + 0.34 = 1.62$$

Sum of SS in the Row: $SS_{total-B} = 31.56 + 47.22 + 50.3 + 42.34 = 171.42$

$3 \times 3 \times SS_B = SS_{total} - SS_{total} \quad = 405.42 - 171.42 = 234$

Sum of SS in the columns: $SS_{tot} \quad = 68 + 79.72 + 89.7 = 237.42$

$4 \times 3 \times SS_A = SS_{total} - SS_{total-A} = 405.42 - 237.42 = 168$

$3 \times SS_{interaction} = SS_{total} - 4 \times 3 \times SS_A - 3 \times 3 \times SS_B - SS_{residual}$

$= 405.42 - 168 - 234 - 1.62 = 1.8$

$$SS_B = \frac{234}{9} = 26$$

$$SS_A = \frac{168}{12} = 14$$

$$SS_{interaction} = \frac{1.8}{3} = 0.6$$

Result of analysis of variance table

Source	SS	df	MS	$F_{observed}{}^1$	$F_{treshold}{}^1$	$F_{observed}{}^2$	$F_{treshold}{}^2$
A	14	2	7	103.7037	6.98***	70	14.544***
B	26	3	8.6667	128.3956	5.82***	86.667	12.192***
Interaction	0.6	6	0.1	1.481481	2.51*		
Residual	1.62	24	0.0675				
Sum		35					

$F^1: \frac{\sigma}{\sigma_{residual}} \qquad F^2: \frac{\sigma}{\sigma_{interaction}}$

*** $p \leq 0.005$ *: $p \leq 0.05$

In this case, we can deny null hypothesis that the difference in A and B is significant. When we look the SS. Df, and MS of A and B, we notice that the values are similar those values in Table 27.

Table 27, Summary of analysis of variance with replication.

Source	SS	df	MS	$F_{0bserved}{}^{1}$	$F_{threshold}{}^{1}$	$F_{0bserved}{}^{2}$	$F_{threshold}{}^{2}$
A	14	2	7	1.40625	3.4	0.7	4.1
B	26	3	8.6667	2.31112	3.0	0.8667	3.7
Interaction	60	6	10	2.6667	2.5		
residual	90	24	3.75				
Sum		35					

$p \leq 0.05$

$$F^1: \frac{\sigma_{factor}{}^2}{\sigma_{residual}{}^2},\ F^2: \frac{\sigma_{factor}{}^2}{\sigma_{interaction}{}^2}$$

Table 27 is result of ANOVA of dataset in Table 26. From this, the readers understand that the dataset in Table 26 was purposely made dataset. Table 26 was made by combining Table 24 and Table 34 to make residual. The values are larger than the values in Table 32.

Table 34. Adding value to make variance in each cell of Table 24.

B_1	-1	-3	-2
	0	1	0
	1	2	2
B_2	-1	-1	-1
	-1	-1	0
	2	2	1
B_3	-1	-3	-1
	0	1	0
	1	2	1
B_4	-1	-4	-1
	-1	1	0
	2	3	1

As the result, the variances of interaction and residual are larger in table 27 than the variances of interaction and residual in result of analysis of variance table or dataset in Table 33, and we cannot deny the null hypothesis as the result.

From this, we can understand that there are two cases. One is the case when interaction has significance. We should implement one-way ANOVA in each level of one factor to detect specific combination of factors which make prominent increase or decrease in data other than the general trend caused by single factor. Then we judge significance of factors using F^2. We can deny the null hypothesis when F^2 is larger than threshold. In the other case when the interaction is not significant, it is meaningless to consider F^2. We cannot deny the null hypothesis even if value of F^2 is larger than the value of threshold, because we do not need to consider effect of interaction. In this case, we judge only by F^1.