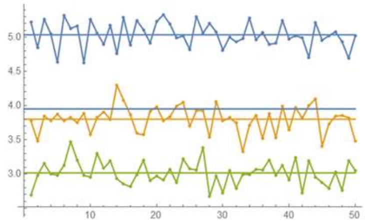


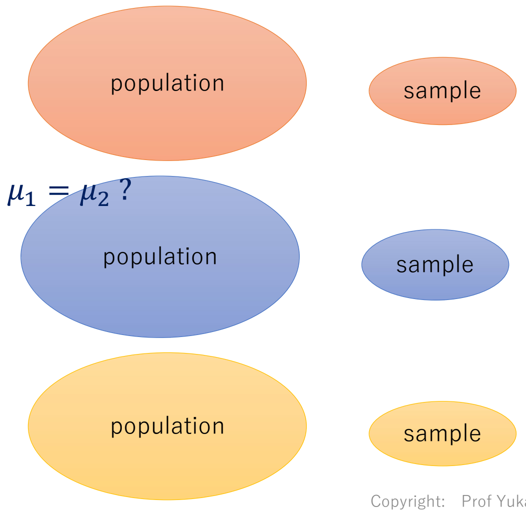
Analysis of Variance Multiple Comparison

2020/07/18

Prof Yukari Shirota (Gakushuin University)
Prof Basabi Chakraborty (Iwate-Prefectural University)

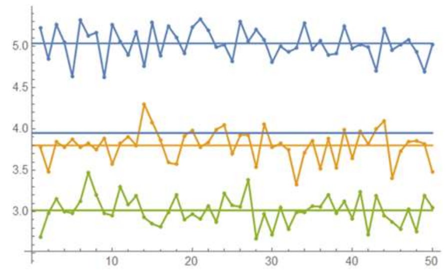
Today I would like to talk about the Analysis of Variance (ANOVA) and multiple comparison. As shown here, the point of ANOVA is the comparison of variances.

Calculate the between-group variance and the within-group variance (error term)



- This is inferential statistics.
- Variance is defined as

$$\frac{\{sum\ of\ squares\ of\ deviations\}}{\{degrees\ of\ freedom\}}$$



In ANOVA, there are two kinds of variances. They are (1) variance between the samples, and (2) variance within each sample. We would like to calculate the two kind of variances. Please do not forget that this is inferential statistics. Our target is the population behavior. A variance can be calculated as sum of squares of deviations divided by degrees of freedom. So we have to count the degrees of freedom.

ANOVA: Analysis of Variance

- One-way ANOVA: to test for differences among means for a single factor when there are three or more groups

  $\boldsymbol{X_{ij} = \mu + \alpha_i + \varepsilon_{ij}}$

  μ **is the mean of the popualtion for all treatments**

  ASSUMPTION: ε_{ij} are error terms, normally distribution $N(0, \sigma^2)$ so that X_{ij} are also normally distribution $N(\mu, \sigma^2)$

  $$Null\ Hypothesis: \mu_1 = \mu_2 = \cdots = \mu$$

  (All treatment means are equal)

  $(Or \quad \boldsymbol{\alpha}_i = 0)$

- Two-way ANOVA: to test the effects of two factors and the possible interaction between them
  - Two-way ANOVA without replication
  - Two-way ANOVA with replication

There are three kinds of ANOVA which are One-way ANOVA, Two-way ANOVA without replication and with replication.
First let's talk about the one-way ANOVA. The null hypothesis is the population average A equals to the population average B and all group population averages are equal. One-way ANOVA has just one factor such as treatment methods and medicines. In two-way ANOVA, there are two factors and an interaction between the two factors may exist.

ANOVA: Analysis of Variance
Two-way ANOVA without replication

- Two-way ANOVA: to test the effects of two factors and the possible interaction between them
  - Two-way ANOVA without replication

$$X_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij}$$

μ is the mean of the popualtion for all treatments

α_i is the treatment effects

β_i is the block effects

ASSUMPTION: ε_{ij} are normally distribution $N(0, \sigma^2)$

so that X_{ij} are also normally distribution $N(\mu, \sigma^2)$

ASSUMPTION: There is no interaction between treatments and blocks

Null Hypothesis (1): *All treatment means are equal* ($\boldsymbol{\alpha}_i = 0$)

Null Hypothesis (2): *All block means are equal* ($\boldsymbol{\beta}_j = 0$)

Let me start with the Two-way ANOVA without replication.
We shall investigate two kinds of effects that are called treatment effects And block effects. The assumption to use the two-way ANOVA is that there is no interaction between two factors.

ANOVA: Analysis of Variance

Two-way ANOVA with replication

- Two-way ANOVA: to test the effects of two factors and the possible interaction between them
  - Two-way ANOVA with replication

$$\boldsymbol{X_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}}$$

μ **is the mean of the popualtion for all treatments**

α_i is the treatment effects, β_i is the block effects

$\boldsymbol{\alpha\beta_{ij}}$ is the treatment-block interaction effects

ASSUMPTION: ε_{ijk} are normally distribution $N(0, \sigma^2)$

so that X_{ijk} are also normally distribution $N(\mu, \sigma^2)$

$$\textit{Null Hypothesis (1): All treatment means are equal } (\boldsymbol{\alpha}_{\mathrm{i}} = 0)$$

$$\textit{Null Hypothesis (2): All block means are equal } (\boldsymbol{\beta}_{\mathrm{j}} = 0)$$

$$\textit{Null Hypothesis (3): There is no interactions between treatments and blocaks } (\boldsymbol{\alpha\beta_{ij}} = 0)$$

Let me move on to the two-way ANOVA with replication. Then the model expression, this alpha beta term is the interaction term. In a two-way ANOVA with replication, there are three null hypotheses.

One-way ANOVA
Effects of Treatments

$$X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

- Null Hypothesis

$$\mu_1 = \mu_2 = \mu_3$$

- Alternative Hypothesis

$$NOT\ [\mu_1 = \mu_2 = \mu_3]$$

Given data of 3 groups

- Treatment 1 2 3

$$\begin{pmatrix} 2.8 & 3.5 & 3.1 \\ 3. & 3.2 & 2.9 \\ 2.7 & 3.1 & 3.4 \\ 2.9 & 3.4 & 2.6 \\ 3.1 & 3.3 & 3. \end{pmatrix}$$

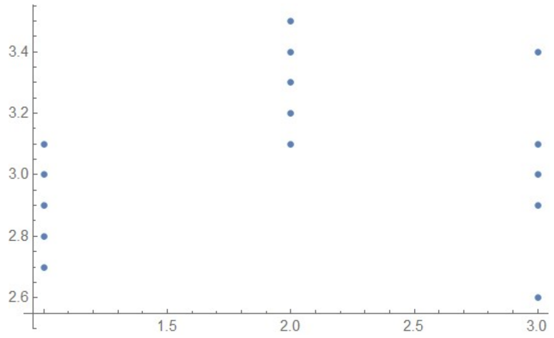
Sample size 5

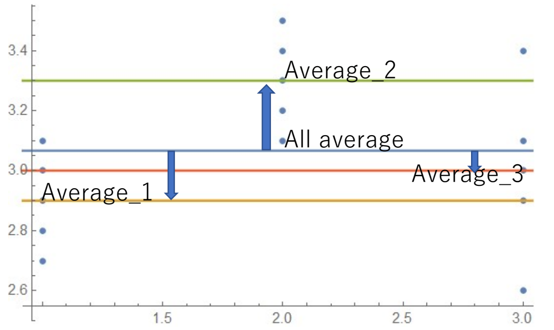
Then Let's start an concrete example. The first one is a one-way ANOVA. The given data of 3 groups is as shown here and the sample size is 5. The null hypothesis is all population averages are equal.

One-way ANOVA
Effects of Treatments

- $X_{ij} = \mu + \alpha_i + \varepsilon_{ij}$

All	3.06667
Model[1]	2.9
Model[2]	3.3
Model[3]	3.





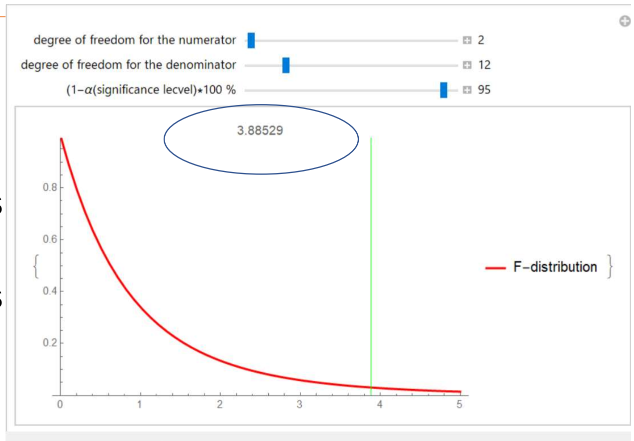
If between-group variance is much greater than within-group variance,
Some treatment effect exists → Rejection of the null hypothesis

Let's plot the data. The all average is 3.0 and each group averages are 2.9, 3.3, and 3.0. The averages are drawn in the plot. We are interested in a between-group variance and a within-group variance. The between-group variance's base line is the all average. The ratio of the two variances is called F value.

One-way ANOVA
Effects of Treatments

		DF	SumOfSq	MeanSq	FRatio	PValue
ANOVA →	Model	2	0.433333	0.216667	4.81481	0.0291604
	Error	12	0.54	0.045		
	Total	14	0.973333			

- F ratio 4.8 falls in the rejection region.
  → Reject the null hypothesis
- P-value 0.029 < 0.05
  → Reject the null hypothesis



We use the F distribution with 2 and 12 which are the degrees of freedom. This data's F value was 4.8 which falls in the rejection region. So we reject the null hypothesis.
This is the end of this ANOVA. But we would like to know which is greater than others. So after ANOVA, we will conduct a multiple comparison.
You are now allowed to repeat t-test repeatedly. We must do a multiple comparison concerning three group comparison.

Multiple Comparison after ANOVA
Tukey Test

- When the null hypothesis of ANOVA is rejected, we want to know where the significant differences in the means lie
- Tukey test: a test used to compare all possible pairs of means; samples in each group must have the same size

Tukey test value

$$q = \frac{\bar{X}_i - \bar{X}_j}{\sqrt{\frac{s_w^2}{n}}}$$

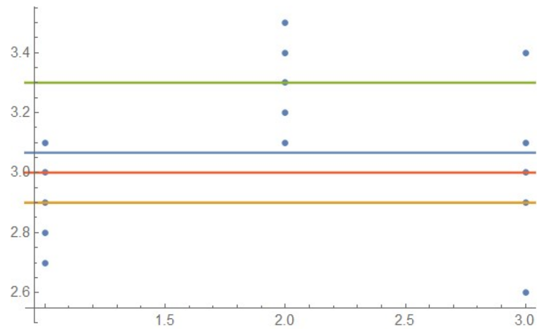
s_w^2 : within − group variance
n : sample size

A multiple comparison method is a test to compare all possible pairs of averages. It tests whether the difference between two group average is significantly large.
There are many multiple comparison methods. Among them Tukey test is a widely used method. The within-group variance includes the all groups' variance information.
Let's use the Tukey method.

Tukey Test Result with 5%

```
PostTests → {Model → Tukey {{1, 2}, {2, 3}}}
```

- Given data of 3 groups



- μ_1 versus μ_2 different
- μ_1 versus μ_3 not different
- μ_2 versus μ_3 different

The Tukey test make the all possible pairs. In this case, we have three pairs. The Tukey test result told us that the difference between group 1 and 2 and the difference between group 2 and 3.

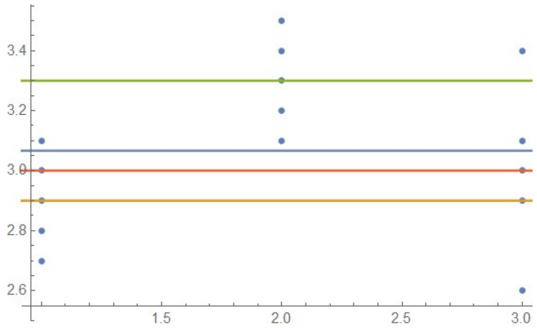
Tukey Test Result with significant level 1%

reducing the significance level by 4%
reducing the rejection region by 4%

`PostTests → {Model → Tukey  {1, 2}`

- Every pair does not fall in the rejection region.
- $\mu_1$versus μ_2 different
- $\mu_1$versus μ_3 not different
- $\mu_2$versus μ_3 not different

- Given data of 3 groups



Let's change the significant level from 5 % to 1 %. The rejection region became smaller. Then the selected significantly different pair was just 1 and 2. We would like to write these kinds of result in our paper. So we shall use the multiple comparison when we handle three or more groups.

Multiple Comparison after ANOVA
Dunnett's Test

- When the null hypothesis of ANOVA is rejected, we want to know where the significant differences in the means lie
- Dunnett's test: a test used to compare one control group mean to the other (n – 1) groups means
- Control group: a group which is not given any special treatment

	Control Group	Treatment 1	Treatment 2	Treatment 3
	24	31	34	29
	23	29	35	28
	25	32	33	31
	26	33	32	30
	22	30	30	29
average	24	31	32.8	29.4

This is another multiple comparison method named Dunnett's test. If you made the control group, then we use the Dunnett's test. The comparison is conducted between the control group average and each group average.

Dunnett's Test Example

- The measurement of the percentage of fat content in cows under three types of diet treatments

	Control Group	Treatment 1	Treatment 2	Treatment 3
	24	31	34	29
	23	29	35	28
	25	32	33	31
	26	33	32	30
	22	30	30	29
average	24	31	32.8	29.4

Significance level	Siginificantly different pairs compared to the control
0.05	Treatment 1, 2, 3

ANOVA →

	DF	SumOfSq	MeanSq	FRatio	PValue
Model	3	216.2	72.0667	28.8267	1.09379×10^{-6}
Error	16	40.	2.5		
Total	19	256.2			

, PostTests → {Model → {Dunnett {1, 2, 3}}}

This is an example of Dunnett's test. For the given data, every group 1, 2, and 3 is significantly different from the control group average.

Two-way ANOVA with Replication

$$X_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$$

μ is the mean of the popualtion for all treatments

α_i is the treatment effects, β_i is the block effects

$\alpha\beta_{ij}$ is the treatment-block interaction effects

ASSUMPTION: ε_{ijk} are normally distribution $N(0, \sigma^2)$

so that X_{ijk} are also normally distribution $N(\mu, \sigma^2)$

Null Hypothesis (1): *All treatment means are equal* ($\alpha_i = 0$)

Null Hypothesis (2): *All block means are equal* ($\beta_j = 0$)

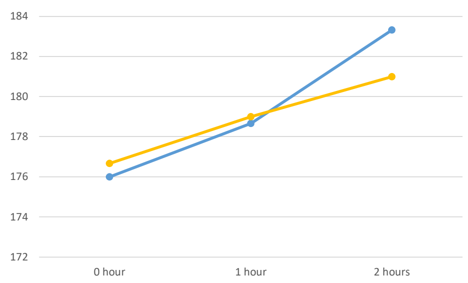
Null Hypothesis (3): *There is no interactions between treatments and blocaks* ($\alpha\beta_{ij} = 0$)

Finally, let me talk about the two-way ANOVA with replication. The model expression included the interaction term. Because the interaction between two factors may exist, we add the third null hypothesis which concerns the interaction existence.

Two-way ANOVA with Replication
No interaction -> Multiple comparison by the factor

- The changes by the medicine XXX and YYY in **systolic** blood pressure of elephants are indicated in the table.

medicine	0 hour	1 hour	2 hours
XXX	176	179	180
	173	176	184
	179	181	186
YYY	177	180	181
	174	176	179
	179	181	183



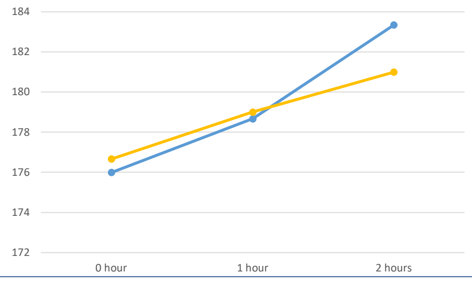
Let's conduct the two-way ANOVA. The data is six elephants' blood pressure data. The first factor is medicine XXX and YYY. The second factor is the periods, a start, after one hour and after 2 hours. The repetition number is three per group. The plot shows the average of the three data. The blue line shows the effect of medicine XXX.

Two-way ANOVA with Replication

No interaction -> Multiple comparison by the factor

- The changes by the medicine XXX and YYY in **systolic** blood pressure of elephants are indicated in the table.

medicine	0 hour	1 hour	2 hours
XXX	176	179	180
	173	176	184
	179	181	186
YYY	177	180	181
	174	176	179
	179	181	183



Null Hypothesis (1): There is no difference in mean blood pressure with respect to medicine
Null Hypothesis (2): There is no difference in mean blood pressure with respect to time
Null Hypothesis (3): There is no interaction between medicines and time

First let's set the three null hypotheses.

Two-way ANOVA with Replication
No interaction -> Multiple comparison by the factor

- The changes by the medicine XXX and YYY in **systolic** blood pressure of elephants are indicated in the table.

		DF	SumOfSq	MeanSq	FRatio	PValue
	factor1	1	0.888889	0.888889	0.126984	0.727767
ANOVA →	factor2	2	102.778	51.3889	7.34127	0.00827417
	factor1 factor2	2	8.11111	4.05556	0.579365	0.57518
	Error	12	84.	7.		
	Total	17	195.778			

No medicine effect
No interaction

Null Hypothesis (1): There is no difference in mean blood pressure with respect to medicine

~~*Null Hypothesis* (2): There is no difference in mean blood pressure with respect to time~~

Null Hypothesis (3): There is no interaction between medicines and time

No interaction -> Multiple comparison by the factor “time”

This is the result of the ANOVA. First we should check the third interaction hypothesis. Because the p-value is 0.57 then we do not reject the null hypothesis. So the interaction does not exist. After we check no interaction, we shall check the two kinds of main effects. As a result, concerning the factor 2, the F ratio falls in the rejection region.

Two-way ANOVA with Replication

No interaction -> Multiple comparison by the factor "time"

Tukey test with 5% level

- 0hour versus 1hour
- 0hour versus 2hours
- 1hour versus 2hours

medicine	0 hour	1 hour	2 hours
XXX	176	179	180
	173	176	184
	179	181	186
YYY	177	180	181
	174	176	179
	179	181	183
average	176.3333	178.8333	182.1667

Model → Tukey {0, 2}

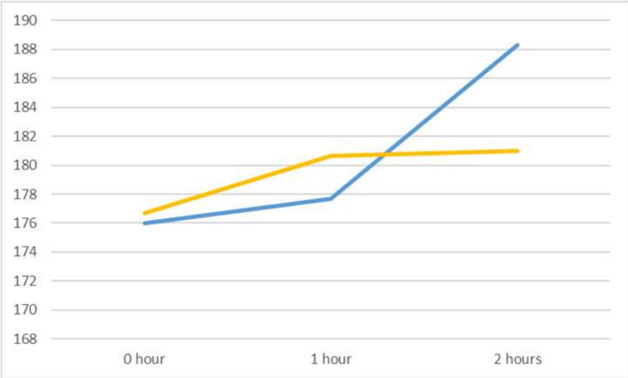
- There is a significant difference between 0 hour and 2 hours.

The factor 2 is the time factor. We would like to know the effect of time in more detail. So we shall do the Tukey test checking all possible pairs. The averages are 176 for 0 hour, 178 for 1 hour, and 182 for 2 hour. The average is the average of 6 data.
The Tukey test told the difference between 0 hours and 2 hours.

Two-way ANOVA with Replication
The case of a significant interaction

- The changes by the medicine AAA and BBB in systolic blood pressure of elephants are indicated in the table.

medicine	0 hour	1 hour	2 hours
AAA	176	179	189
	173	173	189
	179	181	187
BBB	177	181	181
	174	178	179
	179	183	183



We showed the no interaction case. Then let me move on to the interaction existence case. Now the medicine AAA and BBB are investigated. In the plot, the AAA line and the BBB line are crossed. The interaction must exist, we think.

Two-way ANOVA with Replication
The case of a significant interaction

- The changes by the medicine AAA and BBB in **systolic** blood pressure of elephants are indicated in the table.

ANOVA →

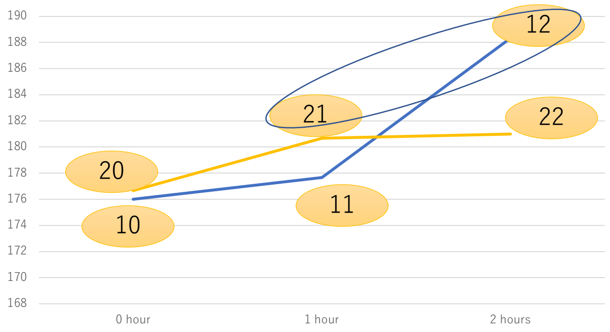
	DF	SumOfSq	MeanSq	FRatio	PValue
factor1	1	6.72222	6.72222	0.909774	0.358994
factor2	2	215.444	107.722	14.5789	0.000614283
factor1 factor2	2	88.1111	44.0556	5.96241	0.015922
Error	12	88.6667	7.38889		
Total	17	398.944			

There is an interaction

- There is a significant interaction between the medicine and the time. $\boldsymbol{X_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}}$
- So no main effects should be interpreted without further study

Then as we expected, there is a significant interaction. The third F ratio's P-value is 0.015 which falls in the rejection region.
If there is an interaction, we should not continue the factor 1 and 2 effect checks. This is because we cannot make the correct decision about alpha and beta owing to the term alpha-beta.

Two-way ANOVA with Replication
Because a significant interaction exists,
we conduct Multiple Comparison of all 6 groups



Tukey {{10, 12}, {11, 12}, {12, 20}, {12, 21}}

Bonferroni {{10, 12}, {11, 12}, {12, 20}}

Then we conduct all 6 groups multiple comparison. We make the six groups and conduct the multiple comparison on them. The result of Tukey test was 4 differences and the result by the Bonferroni method was 3 differences. When we write a paper, we describe "Because a significant interaction exists, we conduct the multiple comparison of all 6 groups, The results are …".

Summary

- Two-way ANOVA without replication
  ASSUMPTION: There is no interaction between treatments and blocks
- Two-way ANOVA with replication
  - When a significant interaction exists, conduct a multiple comparison among all groups.
  - When a significant interaction does not exist, conduct a multiple comparison concerning the factor which has an effect

Welcome your request for teaching materials

- I am Prof Shirota, waiting for your requests and questions about statistics analyses.
- Because I have no idea on your specific research field, I am afraid, I might not be able to answer the question like “What kind of analysis method should I apply for this data (your data) ?” I would be grateful if you could ask a question like “The excellent paper in our field uses XXXX methods. Could you please explain me the XXXX methods ?” In addition, could you please send me the paper ?
- Analysis methods are specific to the research field. So I think I may not be able to answer your question easily. But I would like to look for the approach together with you.