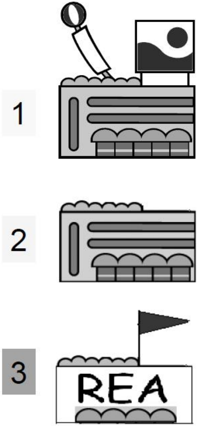
Lec 1: An Introduction to ANOVA

Ying Li
Stockholm University

October 31, 2011

- Three end-aisle displays
- Which is the best?



Design of the Experiment

- Identify the stores of the similar size and type.
- The displays are randomly assigned to use.

Design of the Experiment

- Identify the stores of the similar size and type.
- The displays are randomly assigned to use.

First Principle

RANDOMIZATION

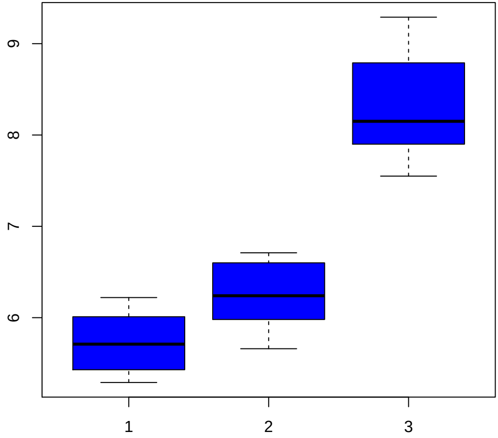
Observations

■ **TABLE 3.12**

The End-Aisle Display Experimental Design

Display Design	Sample Observations, Percent Increase in Sales				
1	5.43	5.71	6.22	6.01	5.29
2	6.24	6.71	5.98	5.66	6.60
3	8.79	9.20	7.90	8.15	7.55

- 3 levels
- 5 replicates



9
8
7
6
1
2
3

The Analysis of Variance

TABLE 3.2
Typical Data for a Single-Factor Experiment

Treatment (Level)	Observations				Totals	Averages
1	y_{11}	y_{12}	...	y_{1n}	$y_{1.}$	$\bar{y}_{1.}$
2	y_{21}	y_{22}	...	y_{2n}	$y_{2.}$	$\bar{y}_{2.}$
⋮	⋮	⋮	⋮ ⋮ ⋮	⋮	⋮	⋮
a	y_{a1}	y_{a2}	...	y_{an}	$y_{a.}$	$\bar{y}_{a.}$
					$y_{..}$	$\bar{y}_{..}$

- a level of the factors (a treatments)
- n replicates
- $N = a \times n$ runs
- Completely randomized design

Statistical Model

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}$$

$$i = 1, \cdots a$$

$$j = 1, \cdots n$$

- μ : overall mean
- τ_i : the effect of ith treatment
- $\varepsilon_{ij} \sim N(0, \sigma^2), i.i.d.$
- $\sum_i^a \tau_i = 0$

The Analysis of Variance

We are interested in testing the equality of treatment means

$$E(\bar{y}_{i.}) = \mu + \tau_i = \mu_i$$

$i = 1, \cdots a$ The appropriate hypothesis are

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_a$$

$$H_1 : \mu_i \neq \mu_j$$

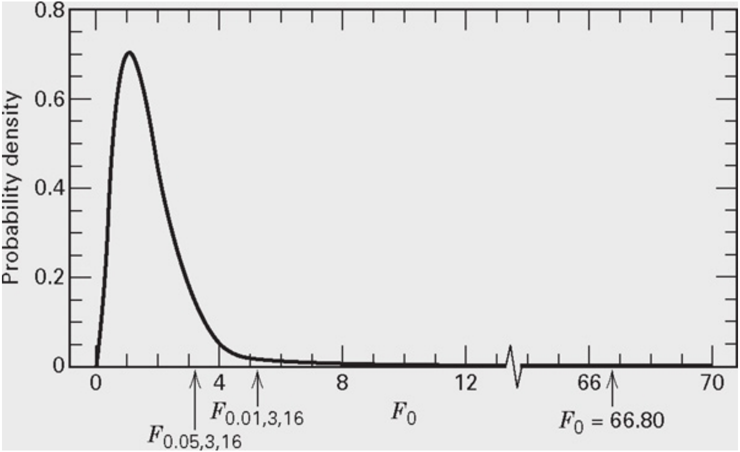
for at least one pair (i, j)

TABLE 3.3

The Analysis of Variance Table for the Single-Factor, Fixed Effects Model

Source of Variation	**Sum of Squares**	**Degrees of Freedom**	**Mean Square**	F_0
Between treatments	$SS_{\text{Treatments}} = n\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2$	$a-1$	$MS_{\text{Treatments}}$	$F_0 = \frac{MS_{\text{Treatments}}}{MS_E}$
Error (within treatments)	$SS_E = SS_T - SS_{\text{Treatments}}$	$N-a$	MS_E	
Total	$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n}(y_{ij} - \bar{y}_{..})^2$	$N-1$		

- If H_0 is true, the distribution of F_0 is $F_{a-1,a(n-1)}$
- Reject H_0 if $F_0 > F_{\alpha,a-1,a(n-1)}$



0.8
0.6
0.4
0.2
0
Probability density
0
4
8
12
66
70
$F_{0.05,3,16}$
$F_{0.01,3,16}$
F_0
F_0 = 66.80

Equation for manual calculation

Balanced Data:

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n} y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$SS_{Treatment} = \frac{1}{n}\sum_{i=1}^{a} y_{i.}^2 - \frac{y_{..}^2}{N}$$

$$SS_E = SS_T - SS_{Treatment}$$

Unbalanced Data

$$SS_T = \sum_{i=1}^{a}\sum_{j=1}^{n_i} y_{ij}^2 - \frac{y_{..}^2}{N}$$

$$SS_{Treatment} = \sum_{i=1}^{a} \frac{y_{i.}^2}{n_i} - \frac{y_{..}^2}{N}$$

$$SS_E = SS_T - SS_{Treatment}$$

Example

■ **TABLE 3.12**

The End-Aisle Display Experimental Design

Display Design	Sample Observations, Percent Increase in Sales				
1	5.43	5.71	6.22	6.01	5.29
2	6.24	6.71	5.98	5.66	6.60
3	8.79	9.20	7.90	8.15	7.55

$\bar{y}_{1.} = 5.73, \bar{y}_{2.} = 6.24, \ \bar{y}_{3.} = 8.32, \bar{y}_{..} = 6.76$

$SS_T = 21.93$

$SS_{Treatment} = 18.78$

$SS_E = SS_T - SS_{Treatment} = 3.15$

	DF	SS	MS	F	P
Treatment	2	18.78	9.39	35.77	$<$0.001
Error	12	3.15	0.26		
Total	14	21.93			

	DF	SS	MS	F	P
Treatment	2	18.78	9.39	35.77	<0.001
Error	12	3.15	0.26		
Total	14	21.93			

Which ones cause the differ?

Multiple Comparison Methods

We know the end-aisle display different than others. We might suspect the first and the second are different. Thus one reasonable test hypothesis would be

$$H_0 : \mu_1 = \mu_2$$

$$H_1 : \mu_1 \neq \mu_2$$

or,

$$H_0 : \mu_1 - \mu_2 = 0$$

$$H_1 : \mu_1 - \mu_2 \neq 0$$

Contrasts

In general, a contrast is a linear combination of the parameters of the form

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i$$

where c_i are the contrast constants, and $\sum_{i=1}^{a} c_i = 0$.
For unbalance design $\sum_{i=1}^{a} c_i n_i = 0$.

The hypothesis can be expressed in the terms of contrasts

$$H_0 : \sum_{i=1}^{a} c_i \mu_i = 0$$

$$H_1 : \sum_{i=1}^{a} c_i \mu_i \neq 0$$

In our example, $c_1 = 1, c_2 = -1,\ c_3 = 0$.

Contrast Tests

Testing hypothesis involving contrast can be done in two basic ways

1. t-test:

$$t_0 = \frac{\sum_{i=1}^{a} c_i \bar{y}_{i.}}{\sqrt{\frac{MS_E}{n} \sum_{i=1}^{a} c_i^2}}$$

2. F-test:

$$F_0 = \frac{MS_c}{MS_E} = \frac{SS_c/1}{MS_E}$$

where

$$SS_c = \frac{(\sum_{i=1}^{a} c_i \bar{y}_{i.})^2}{\frac{1}{n} \sum_{i=1}^{a} c_i^2}$$

Orthogonal Contrasts

Two contrasts $\sum_{i=1}^{a} = c_i \mu_i = 0$ and $\sum_{i=1}^{a} = d_i \mu_i = 0$ are orthogonal if

$$\sum_{i=1}^{a} c_i d_i = 0$$

or in unbalanced case if

$$\sum_{i=1}^{a} n_i c_i d_i = 0$$

We can specify $a - 1$ orthogonal contrasts. Tests performed on orthogonal contrasts are independent.

Example

Treatment	Coefficients for Orthogonal Contrasts	
1 (control)	−2	0
2 (level 1)	1	−1
3 (level 2)	1	1

Pairwise Comparison

The null hypothesis are

$$H_0 : \mu_i = \mu_j, i \neq j$$

Pairwise Comparison

The null hypothesis are

$$H_0 : \mu_i = \mu_j, i \neq j$$

Tukey's test

Make use of the distribution of studentized range statistics

$$q = \frac{\bar{y}_{max} - \bar{y}_{min}}{\sqrt{\frac{MS_E}{2}(\frac{1}{n_i} + \frac{1}{n_j})}}$$

Appendix Table 7 contains upper percentiles for q.
The difference d between two averages is significant if

$$|d| > \frac{q_\alpha(a, f)}{\sqrt{2}} \sqrt{MS_E(\frac{1}{n_i} + \frac{1}{n_j})}$$

The fisher least significant difference method (LSD)

The fisher least significant difference method procedure use T test for testing

$$H_0 : \mu_i = \mu_j, i \neq j$$

The test statistics is:

$$t_0 = \frac{\bar{y}_{i.} - \bar{y}_{j.}}{\sqrt{MS_E(\frac{1}{n_i} + \frac{1}{n_j})}}$$

The quantity: least significant difference (LSD)

$$LSD = t_{\alpha/2,N-a}\sqrt{MS_E(\frac{1}{n_i} + \frac{1}{n_j})}$$

If

$$|\bar{y}_{i.} - \bar{y}_{j.}| > LSD$$

we conclude that the population means μ_i and μ_j differ.

Questions:1

Why in some situation overall F test of ANOVA is significant, but the pairwise comparison fails to reveal the differences?

Questions:2

Many tests available:
Duncan, Student-Newman-Keuls, REGWQ, Bonferroni, Sidak, Scheff, ·
Which pairwise comparison method do I use?

Questions:2

Many tests available:
Duncan, Student-Newman-Keuls, REGWQ, Bonferroni, Sidak, Scheff, ·
Which pairwise comparison method do I use?

- unfortunately, no-clear cut answer.
- Fisher's LSD procedure, only apply if F test is significant
- Tukey's method control overall error rate, many statisticians prefer.

Dunnett's Test

Often one of the treatments is a control treatment, and we want to compare the other treatments with the control treatment. We want to test $a - 1$ hypotheses:

$$H_0 : \mu_j = \mu_1$$

where

$$i = 2, \cdots , a$$

Dunnett's Test

For the ith treatment, H_0 is rejected if

$$|\bar{y}_i - \bar{y}_1| > d_\alpha(a-1, f)\sqrt{MS_E(\frac{1}{n_i} + \frac{1}{n_j})}$$

Appendix 8 gives upper percentiles for $d_\alpha(a-1, f)$

One Example

■ **TABLE 3.1**
Etch Rate Data (in Å/min) from the Plasma Etching Experiment

Power (W)	Observations 1	2	3	4	5	Totals	Averages
160	575	542	530	539	570	2756	551.2
180	565	593	590	579	610	2937	587.4
200	600	651	610	637	629	3127	625.4
220	725	700	715	685	710	3535	707.0

One Example

■ TABLE 3.4
ANOVA for the Plasma Etching Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
RF Power	66,870.55	3	22,290.18	$F_0 = 66.80$	<0.01
Error	5339.20	16	333.70		
Total	72,209.75	19			

One Example

■ **TABLE 3.11**

Analysis of Variance for the Plasma Etching Experiment

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F_0	*P*-Value
Power setting	66,870.55	3	22,290.18	66.80	<0.001
Orthogonal contrasts					
$C_1 : \mu_1 = \mu_2$	(3276.10)	1	3276.10	9.82	<0.01
$C_2 : \mu_1 + \mu_3 = \mu_3 + \mu_4$	(46,948.05)	1	46,948.05	140.69	<0.001
$C_3 : \mu_3 = \mu_4$	(16,646.40)	1	16,646.40	49.88	<0.001
Error	5,339.20	16	333.70		
Total	72,209.75	19			