

PSYC214: Statistics Lecture 2 – One factor between-participants ANOVA – Part I

Michaelmas Term,
Dr Richard Philpot
r.philpot@lancaster.ac.uk

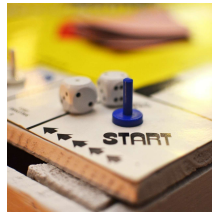
1

1

One factor between-participants ANOVA

Agenda/Content for Lecture 2

- Introduction to analysis of variance (ANOVA)
- Introduction to one factor between-participants design
- Sources of variability in data
- Calculating within-group and between-group variances
- Degrees of Freedom
- Producing the F-statistic



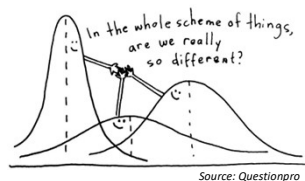
2

2

Introduction to analysis of variance

Why conduct an analysis of variance?

- Compares means and variance
- Allows analysis of group differences for more than two groups
- Several means without inflating Type I error rate



Source: Questionpro

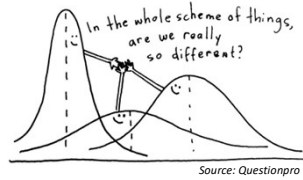
3

3

Introduction to analysis of variance

What do you need for a one factor between participants ANOVA?

- At least one categorical independent variable (i.e., one factor)
- One continuous dependent variable (outcome measure)



4

4

Sources of variability in data

1. Treatment effects
2. Individual differences
3. Random (residual) errors



Within-group variability?



Between-group variability?

5

5

Sources of variability in data

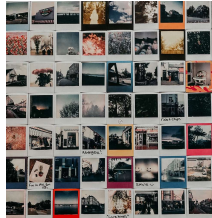
1. **Treatment effects**
2. Individual differences
3. Random (residual) errors

6

6

Treatment effects

- The effects of the independent variable
- This is what we want!
- We want people who are treated differently because of our intervention to behave differently



7

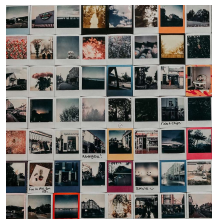
Sources of variability in data

1. Treatment effects
2. Individual differences
3. Random (residual) errors

8

Individual differences

- Some individuals may be more proficient in memory recall
- Maybe some individuals have experience of similar tasks
- Some may have ignored instructions or had lower attention spans / motivation
- A control group can employ their own strategy, increasing the variability



9

Sources of variability in data

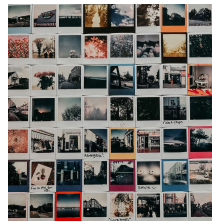
1. Treatment effects
2. Individual differences
3. **Random (residual) errors**

10

10

Random (residual) errors

- Ideally a participant would have a 'true level' at which they perform, which can always be measured accurately
1. Varying external conditions – e.g., temperature, time of day
 2. State of participant (e.g. tired?)
 3. Experimenter's ability to measure accurately...

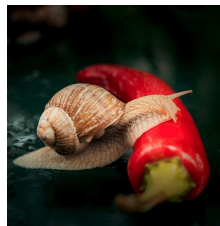


11

11

...Experimenter effects

- Experimenters need to minimise these, so not to obscure the treatment effect
- Spread data away from the true means – i.e., increase variability and standard errors
- Reduce confidence in our estimates and a randomly plucked sample



12

12

Within- and between- group variability

• Within-group variability

- The extent to which participants within a single group or population differ, despite receiving the same treatment



Within-group variability?

• Between-group variability

- The extent to which overall groups differ from one another (hopefully because of our treatment)



Between-group variability?

13

13

Within- and between- group variability

High between-group variability
- no within group-variability

	Group A	Group B	Group C
	10	20	30
	10	20	30
	10	20	30
	10	20	30
	10	20	30
Mean	10	20	30
S	0	0	0

No between-group variability -
high within-group variability

	Group A	Group B	Group C
	10	15	5
	25	20	25
	30	30	25
	35	40	45
	50	45	50
Mean	30	30	30
S	14.6	12.8	18.0

Moderate between-group variability
- moderate within-group variability

	Group A	Group B	Group C
	10	10	20
	10	20	20
	10	20	30
	20	20	30
	20	30	30
Mean	14	20	26
S	5.5	7.1	5.5

Tables adapted from Roberts and Russo (1999)

14

14

PSYC214: Statistics Lecture 2 – One factor between-participants ANOVA – Part II

Michaelmas Term
Dr Richard Philpot
r.philpot@lancaster.ac.uk

15

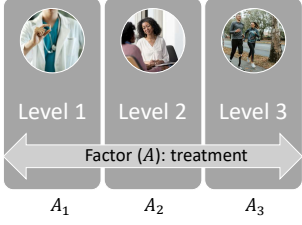
15

Introduction to analysis of variance

Lancaster University

Factors and levels

- Factor: treatment
- 3 levels
 - Medication
 - Counselling
 - Exercise



Level 1 Level 2 Level 3

Factor (A): treatment

A_1 A_2 A_3

16


16

Introduction to analysis of variance

Lancaster University

Factors and levels

- Factor: population
- 3 levels:
 - A_1 Meat eater
 - A_2 Pescatarian
 - A_3 Vegetarian



Level 1 Level 2 Level 3

Factor (A): population

A_1 A_2 A_3

17

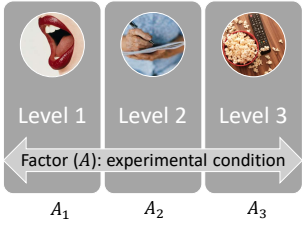
17

Introduction to analysis of variance

Lancaster University

Factors and levels

- Factor: experimental condition
- 3 levels:
 - A_1 Verbal negative feedback
 - A_2 Written negative feedback
 - A_3 Control (no feedback)



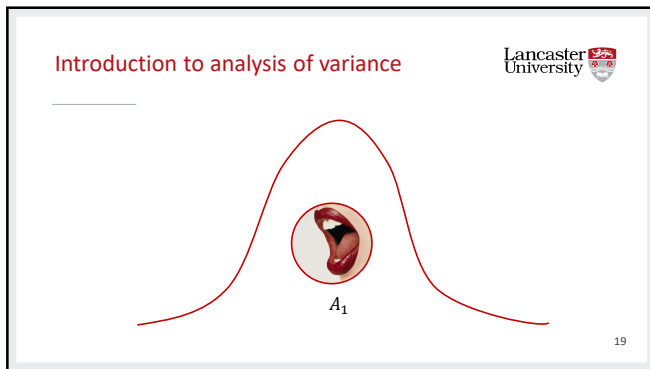
Level 1 Level 2 Level 3

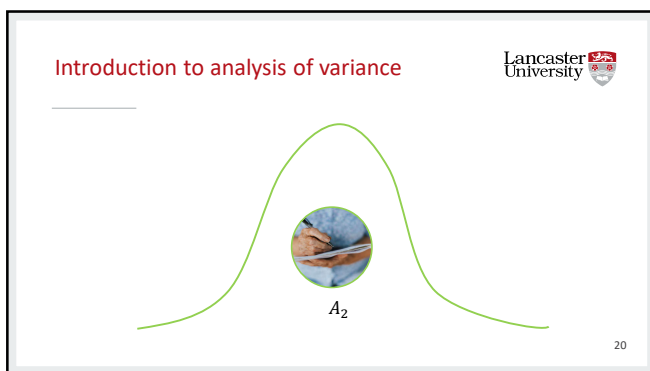
Factor (A): experimental condition

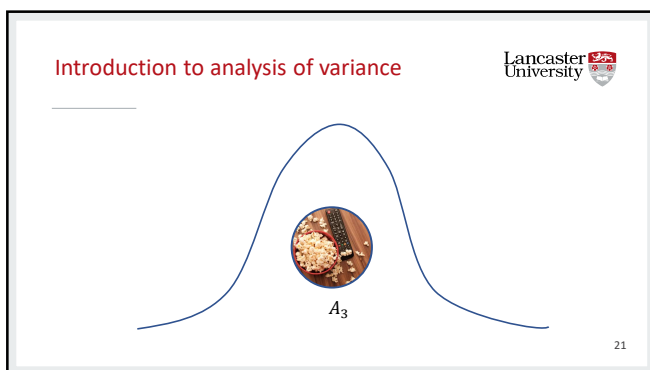
A_1 A_2 A_3

18

18



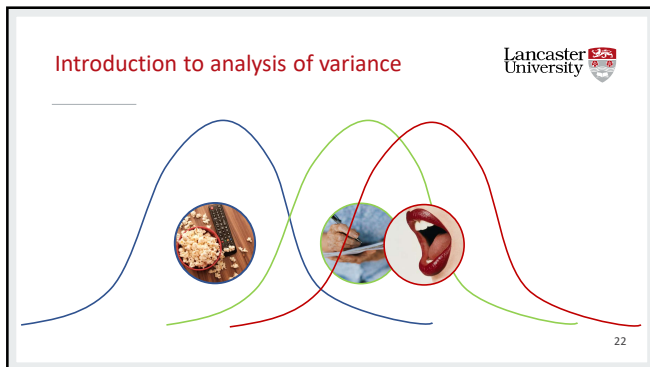




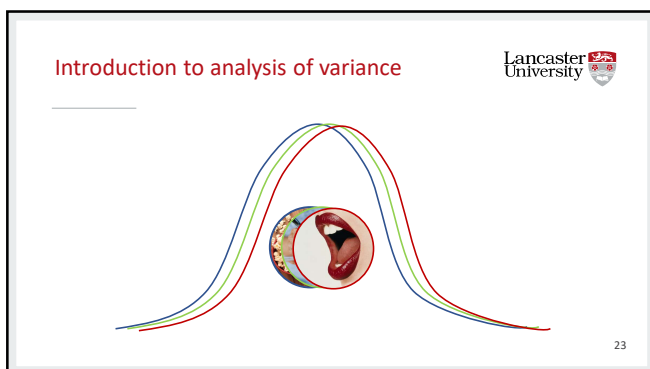
19

20

21



22



23

Testing for differences

Lancaster University

<ul style="list-style-type: none"> • H₀ the Null Hypothesis • Under H₀, the samples come from the <u>same</u> population • $\mu_1 = \mu_2 = \mu_3$ [No difference in the population means] • Experimental effect = 0 • All differences are due to individual differences + random (residual) errors 	<ul style="list-style-type: none"> • H₁ the Experimental Hypothesis • Under H₁, the samples come from the <u>different</u> populations. • $\mu_1 \neq \mu_2 \neq \mu_3$ [Population means are different] • Experimental effect $\neq 0$ • All differences are due to individual differences, random (residual) errors <u>AND</u> the experimental effect
--	--

24

24

Introduction to analysis of variance



$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$

$$F = \frac{\text{Signal}}{\text{Noise}}$$

$$F = \frac{\text{Signal}}{\text{Noise}}$$

25

25

The F ratio

$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$



26

26

The F ratio

$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$



$$F = \frac{\text{treatment effects} + \text{individual differences} + \text{random (residual) errors}}{\text{individual differences} + \text{random (residual) errors}}$$

experimental error


experimental error

27

27

The F ratio

Lancaster University



$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$

$$F = \frac{\text{treatment effects + individual differences + random (residual) errors}}{\text{individual differences + random (residual) errors}}$$


$$F = \frac{\text{treatment effects + experimental error}}{\text{experimental error}}$$

28

28

Introduction to analysis of variance

Lancaster University



$$F = \frac{\text{Signal}}{\text{Noise}}$$

$$F = \frac{\text{Signal}}{\text{Noise}}$$

The larger in magnitude the F value, the more treatment effects are standing out away from experimental error – i.e., the larger the signal is from the noise. The larger the F, the less likely that differences in scores are caused by chance.

29

29

Lancaster University

PSYC214: Statistics
Lecture 2 – One factor between-participants ANOVA – Part III

Michaelmas Term
 Dr Richard Philpot
 r.philpot@lancaster.ac.uk

30

30

Calculating between-group variance

$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$



31

31

Mean (\bar{A})

A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6

Total set of scores

$$\bar{X} = \frac{\sum X}{N}$$

Mean

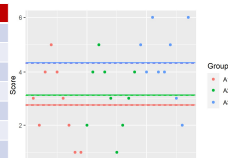
Number of scores

32

32

Mean (\bar{A})

A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$




33

33

Grand Mean (\bar{Y})

Lancaster University



A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

$$\bar{Y} = \frac{\bar{A}_1 + \bar{A}_2 + \bar{A}_3 + \dots \bar{A}_k}{k}$$

\bar{Y} = The grand mean of averages
k = number of levels

$$\bar{Y} = \frac{2.78 + 3.11 + 4.33}{3}$$


$$\bar{Y} = 3.41$$

34

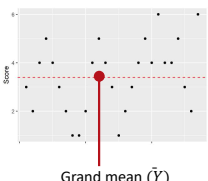
34

Grand Mean (\bar{Y})

Lancaster University



A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$




Grand mean (\bar{Y})

35

35

Total between-group variance

Lancaster University


$$\text{total between group variance} = \frac{N_{A1}(\bar{A}_1 - \bar{Y})^2 + N_{A2}(\bar{A}_2 - \bar{Y})^2 + N_{A3}(\bar{A}_3 - \bar{Y})^2}{\text{total between group degrees of freedom}} \text{ (and so on)}$$



A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

$\bar{Y} = 3.41$

36

36

Total between-group variance 


$$\text{total between group variance} = \frac{N_{A1}(\bar{A}_1 - \bar{Y})^2 + N_{A2}(\bar{A}_2 - \bar{Y})^2 + N_{A3}(\bar{A}_3 - \bar{Y})^2 \text{ (and so on)}}{\text{total between group degrees of freedom}}$$


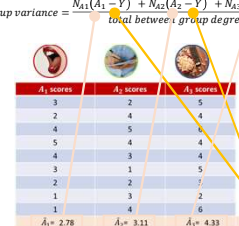
A ₁ scores	A ₂ scores	A ₃ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	2
1	3	2
1	4	6

$\bar{A}_1 = 2.78$ $\bar{A}_2 = 3.11$ $\bar{A}_3 = 4.33$ $\bar{Y} = 3.41$

37

37

Total between-group variance 


$$\text{total between group variance} = \frac{N_{A1}(\bar{A}_1 - \bar{Y})^2 + N_{A2}(\bar{A}_2 - \bar{Y})^2 + N_{A3}(\bar{A}_3 - \bar{Y})^2 \text{ (and so on)}}{\text{total between group degree of freedom}}$$


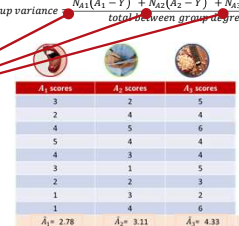
A ₁ scores	A ₂ scores	A ₃ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	2
1	3	2
1	4	6

$\bar{A}_1 = 2.78$ $\bar{A}_2 = 3.11$ $\bar{A}_3 = 4.33$ $\bar{Y} = 3.41$

38

38

Total between-group variance 

$$\text{total between group variance} = \frac{N_{A1}(\bar{A}_1 - \bar{Y})^2 + N_{A2}(\bar{A}_2 - \bar{Y})^2 + N_{A3}(\bar{A}_3 - \bar{Y})^2 \text{ (and so on)}}{\text{total between group degrees of freedom}}$$


N_{A1} = Number of scores for A₁
= 9

N_{A2} = Number of scores for A₂
= 9

N_{A3} = Number of scores for A₃
= 9

A ₁ scores	A ₂ scores	A ₃ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6

$\bar{A}_1 = 2.78$ $\bar{A}_2 = 3.11$ $\bar{A}_3 = 4.33$ $\bar{Y} = 3.41$

39

39

Degrees of freedom

Between-groups degrees of freedom

- The total number of levels minus one
- For example, in our experiment we have three levels [verbal feedback, written feedback, control]
- The between-groups degree of freedom is there 3 levels - 1 = 2
- Between-groups df = 2



40

40

Total between-group variance

$$\text{total between group variance} = \frac{9(2.78 - 3.41)^2 + 9(3.11 - 3.41)^2 + 9(4.33 - 3.41)^2}{2}$$



A ₁ scores	A ₂ scores	A ₃ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

$$\bar{Y} = 3.41$$

41

41

Total between-group variance

$$\text{total between group variance} = \frac{9(-0.63)^2 + 9(-0.30)^2 + 9(0.92)^2}{2}$$



A ₁ scores	A ₂ scores	A ₃ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$




$$\bar{Y} = 3.41$$

42

42

Total between-group variance

total between group variance = $\frac{9(0.40) + 9(0.09) + 9(0.85)}{2}$

A ₁ scores	A ₂ scores	A ₃ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

N_{k1} = Number of scores for A₁

= 9

N_{k2} = Number of scores for A₂

= 9

N_{k3} = Number of scores for A₃

= 9




$F = 3.41$

43

43

Total between-group variance

total between group variance = $\frac{3.60 + 0.81 + 7.65}{2} = 6.037$ (with rounding)

A ₁ scores	A ₂ scores	A ₃ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

$\bar{Y} = 3.41$


44

44

Calculating between-group variance

$F = \frac{\text{between-group variance}}{\text{within-group variance}}$

$F = \frac{6.037}{\text{within-group variance}}$



45

45

PSYC214: Statistics
Lecture 2 – One factor between-participants
ANOVA – Part IV

Michaelmas Term
Dr Richard Philpot
r.philpot@lancaster.ac.uk

46

46

Up to now...

$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$

$$F = \frac{6.037}{\text{within-group variance}}$$



47

47

Calculating within-group variance

$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$



48

48

Total within-group variance



$$\text{total within group variance} = \frac{\text{SS level } A_1 + \text{SS level } A_2 + \text{SS level } A_3 (\text{and so on})}{\text{total within group degrees of freedom}}$$

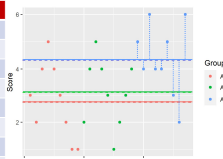
49

49

Mean



A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$



50

50

Total within-group variance



$$\text{total within group variance} = \frac{\text{SS level } A_1 + \text{SS level } A_2 + \text{SS level } A_3 (\text{and so on})}{\text{total within group degrees of freedom}}$$

SS level A_1
= Sums of squares for level 1

SS level A_2
= Sums of squares for level 2

SS level A_3
= Sums of squares for level 3

A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

 $F = 3.41$

51

51

Total within-group variance

$$\text{total within group variance} = \frac{\sum (A_1 - \bar{A}_1)^2 + (A_2 - \bar{A}_2)^2 + (A_3 - \bar{A}_3)^2 + (\text{and so on})}{\text{total within group degrees of freedom}}$$



A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

52

52

Total within-group variance

$$\text{total within group variance} = \frac{\sum (A_1 - 2.78)^2 + (A_2 - 3.11)^2 + (A_3 - 4.33)^2 + (\text{and so on})}{\text{total within group degrees of freedom}}$$



A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

53

53

Degrees of freedom

Within-groups degrees of freedom

- For within-groups degrees of freedom, we add up the number of participants for each level - 1
- Mathematically this is expressed as:

$$= (N_{A1} - 1) + (N_{A2} - 1) + (N_{A3} - 1)$$

$$= (9 - 1) + (9 - 1) + (9 - 1)$$

$$= 24$$



54

54

Total within-group variance

$$\text{total within group variance} = \frac{\sum (A_1 - 2.75)^2 + (A_2 - 3.11)^2 + (A_3 - 4.33)^2}{24}$$



A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

55

55

Total within-group variance

$$\text{total within group variance} = \frac{42.444}{24} = 1.769 \text{ (with rounding)}$$




A_1 scores	A_2 scores	A_3 scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

56

56

The F ratio



$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$

$$F = \frac{6.037}{1.769}$$

$$F = 3.414$$

57

57


$\frac{1}{\sqrt{N}}$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	100	200	216	222	223	224	225	227	230	241	257	284	320	360	400	450	512	583	635
2	85.5	160	172	174	175	176	177	179	182	194	210	232	260	292	328	376	432	500	545
3	70.1	135	145	147	148	149	150	151	153	164	180	200	224	252	284	332	392	456	495
4	57.1	104	110	112	113	114	115	116	117	126	140	156	176	200	228	272	328	392	435
5	46.0	84.0	89.0	90.0	91.0	92.0	93.0	94.0	95.0	104	116	132	152	176	204	248	304	368	411
6	36.0	67.5	71.0	72.0	73.0	74.0	75.0	76.0	77.0	84.0	96.0	112	132	156	184	224	280	344	387
7	28.0	54.0	57.0	58.0	59.0	60.0	61.0	62.0	63.0	68.0	78.0	92.0	108	128	152	188	236	296	341
8	22.0	44.0	46.0	47.0	48.0	49.0	50.0	51.0	52.0	56.0	64.0	76.0	88.0	104	124	156	196	248	293
9	18.0	36.0	38.0	39.0	40.0	41.0	42.0	43.0	44.0	48.0	56.0	68.0	80.0	96.0	112	136	172	216	261
10	15.0	30.0	32.0	33.0	34.0	35.0	36.0	37.0	38.0	42.0	50.0	60.0	72.0	84.0	100	120	152	192	237
12	12.0	24.0	26.0	27.0	28.0	29.0	30.0	31.0	32.0	36.0	44.0	52.0	64.0	76.0	92.0	112	144	184	229
15	9.0	18.0	20.0	21.0	22.0	23.0	24.0	25.0	26.0	30.0	38.0	46.0	56.0	68.0	84.0	104	136	176	221
20	6.8	13.6	15.0	16.0	17.0	18.0	19.0	20.0	21.0	24.0	32.0	40.0	48.0	58.0	72.0	92.0	120	160	205
24	5.8	11.6	12.8	13.6	14.4	15.2	16.0	16.8	17.6	20.0	28.0	36.0	44.0	52.0	64.0	84.0	112	152	197
30	4.8	9.6	10.6	11.4	12.2	13.0	13.8	14.6	15.4	17.6	24.0	32.0	40.0	48.0	60.0	80.0	108	148	193
40	3.6	7.2	7.8	8.4	8.8	9.4	9.8	10.4	10.8	12.4	16.0	20.0	24.0	28.0	36.0	52.0	72.0	104	141
60	2.4	4.8	5.2	5.6	5.8	6.2	6.4	6.6	6.8	7.8	10.4	12.8	15.2	17.6	22.4	32.0	48.0	72.0	111
80	1.8	3.6	3.8	4.0	4.2	4.4	4.6	4.8	5.0	5.6	7.2	8.8	10.4	12.0	15.2	22.4	36.0	56.0	107
100	1.5	3.0	3.2	3.4	3.6	3.8	4.0	4.2	4.4	5.0	6.4	7.6	9.2	10.8	13.6	20.0	32.0	52.0	105
120	1.3	2.6	2.8	3.0	3.2	3.4	3.6	3.8	4.0	4.6	5.6	6.8	8.0	9.6	12.0	18.0	32.0	52.0	103
150	1.0	2.0	2.2	2.4	2.6	2.8	3.0	3.2	3.4	4.0	4.8	5.6	6.8	8.0	10.0	16.0	32.0	52.0	101
200	0.75	1.5	1.6	1.8	1.9	2.0	2.1	2.2	2.3	2.6	3.2	3.8	4.4	5.2	6.4	12.0	32.0	52.0	100
240	0.63	1.26	1.36	1.44	1.52	1.60	1.68	1.76	1.84	2.16	2.80	3.44	4						


Source: E. S. Pearson and H. O. Hartley, *Biometrical Tables for Statisticians*, Vol. 2 (1972), Table 3, page 178, by permission.

58

58


The F ratio





$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$
$$F = \frac{6.037}{1.769}$$

$F = 3.414, p = 0.05$, A statistically significant test result ($P \leq 0.05$)




59

59

Lecture 2 – One factor between-participants ANOVA


Review of lecture 2

- What is Analysis of Variance
- What is a one-factor between-participants design
- Sources of variability in data
- Calculated within-group and between-group variances
- Degrees of Freedom
- Produced the F-statistic




60

60

Lancaster University 

Thank you for attention! Questions?



61
