

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

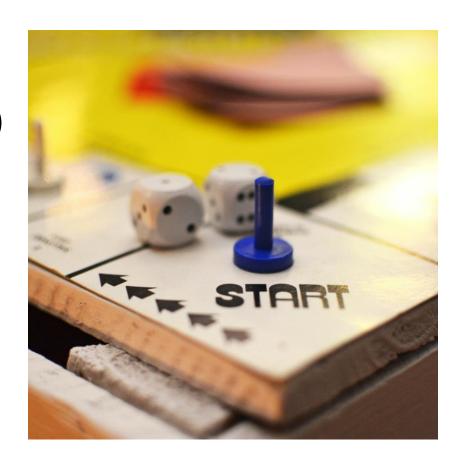
Michaelmas Term,
Dr Richard Philpot
r.philpot@lancaster.ac.uk

### One factor between-participants ANOVA



#### Agenda/Content for Lecture 2

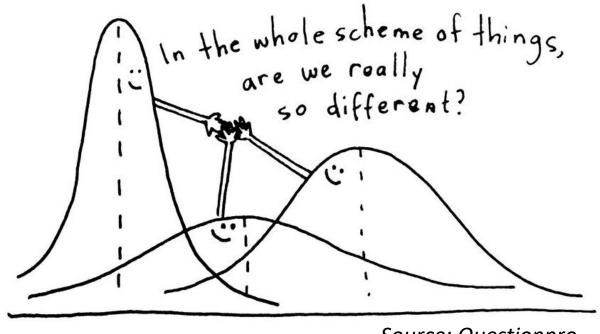
- Introduction to analysis of variance (ANOVA)
- Introduction to one factor betweenparticipants design
- Sources of variability in data
- Calculating within-group and betweengroup variances
- Degrees of Freedom
- Producing the F-statistic





# 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

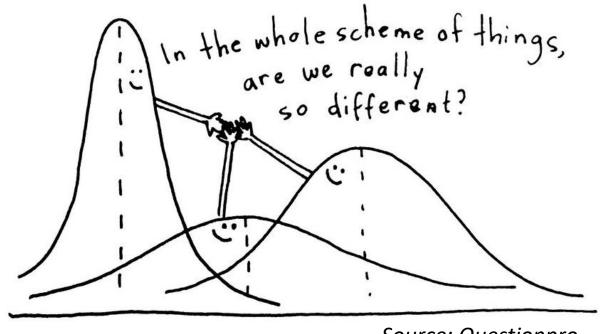


Source: Questionpro



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

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



Source: Questionpro

## Sources of variability in data



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



Within-group variability?



Between-group variability?

## Sources of variability in data



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

### 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



## Sources of variability in data



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

### 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



## Sources of variability in data



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

## 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?)
- Experimenter's ability to measure accurately...



## ...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



## 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?

## Within- and between- group variability



High between-group variability - no within group-variability

No between-group variability - high within-group variability

Moderate between-group variability - moderate 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

	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

	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



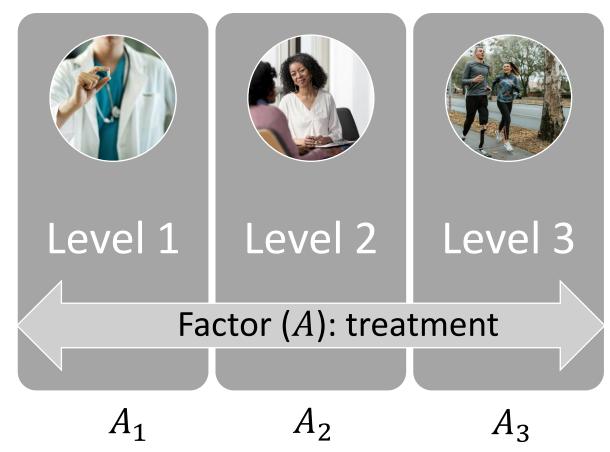
PSYC214: Statistics Lecture 2 – One factor between-participants ANOVA – Part II

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



#### Factors and levels

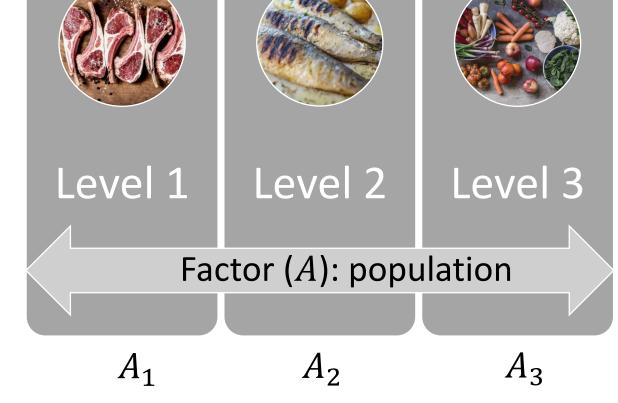
- Factor: treatment
- 3 levels
  - Medication
  - Counselling
  - Exercise





#### Factors and levels

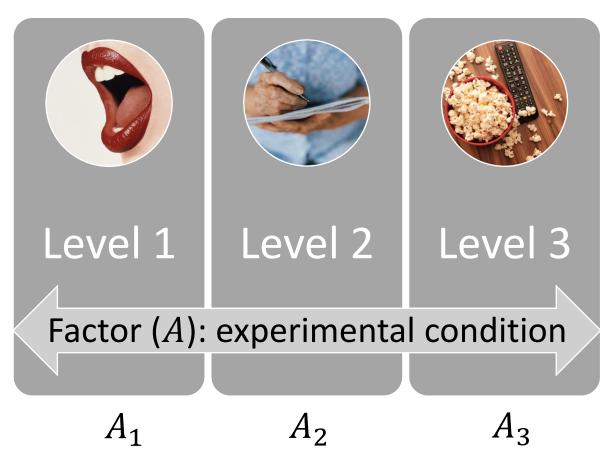
- Factor: population
- 3 levels:
  - $A_1$  Meat eater
  - A<sub>2</sub> Pescatarian
  - A<sub>3</sub> Vegetarian



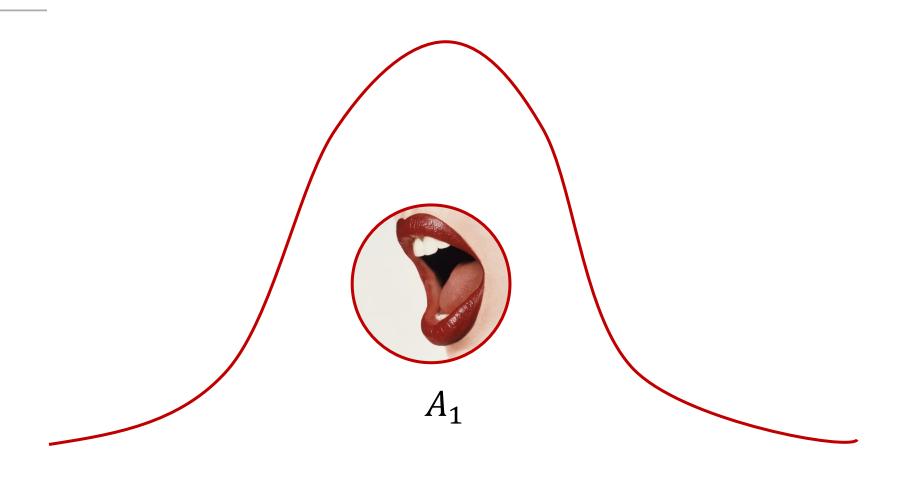


#### Factors and levels

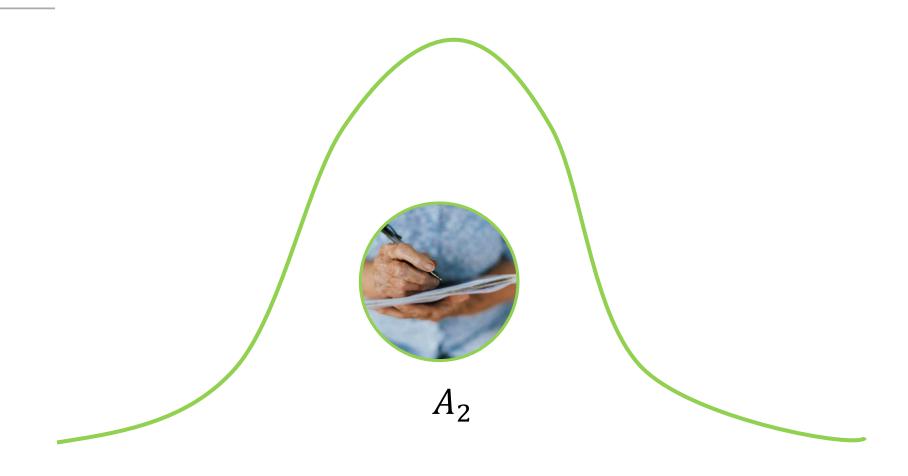
- Factor: experimental condition
- 3 levels:
  - $A_1$  Verbal negative feedback
  - A<sub>2</sub> Written negative feedback
  - A<sub>3</sub> Control (no feedback)



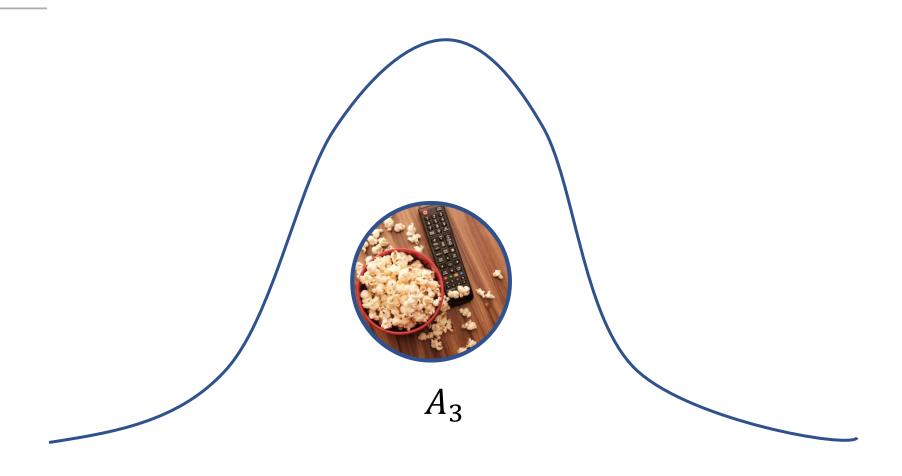




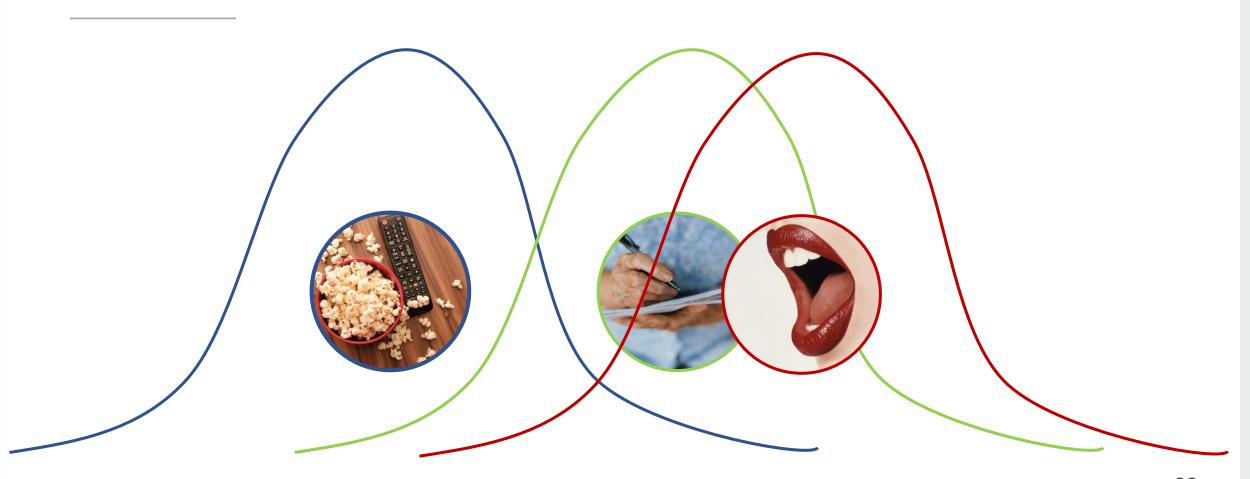




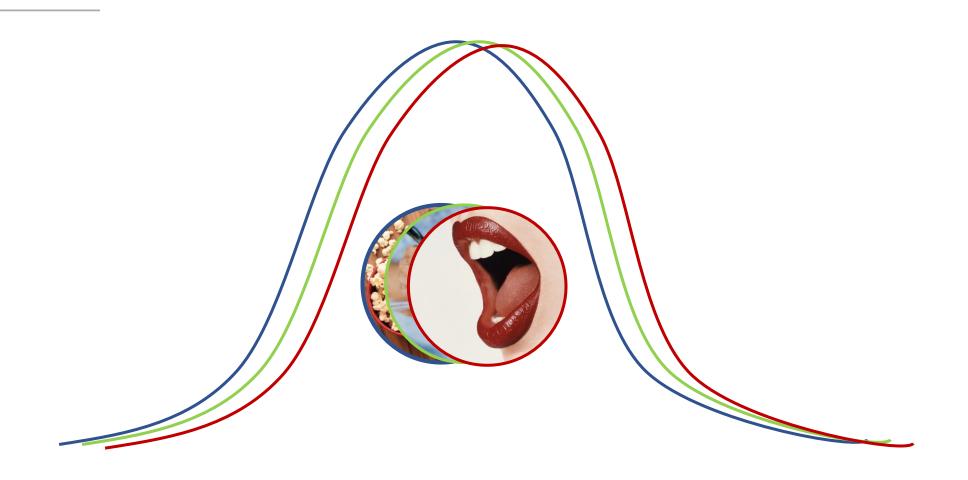












## Testing for differences



- Ho the Null Hypothesis
- Under H0, the samples come from the same 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

- H<sub>1</sub> the Experimental Hypothesis
- Under H1, the samples come from the different populations.
- $\mu_1 \neq \mu_2 \neq \mu_3$  [Population means are different]
- Experimental effect ≠ 0
- All differences are due to individual differences, random (residual) errors AND the experimental effect





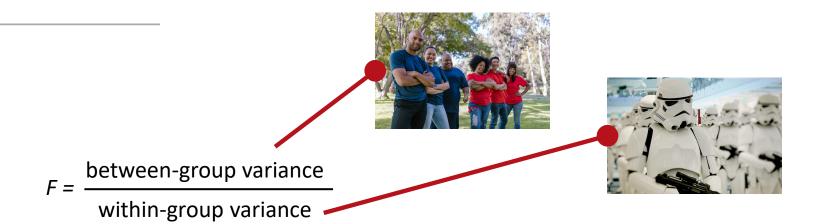
$$F = \frac{\text{between-group variance}}{\text{within-group variance}}$$

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

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

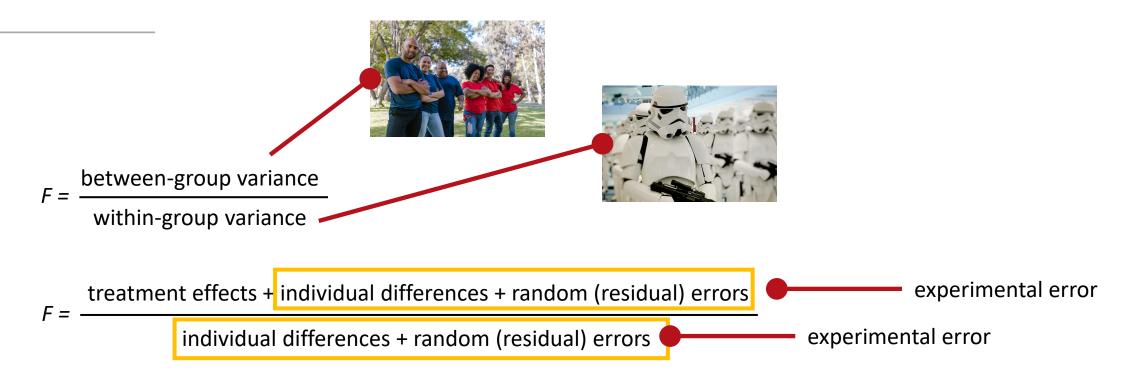
### The F ratio





### The F ratio





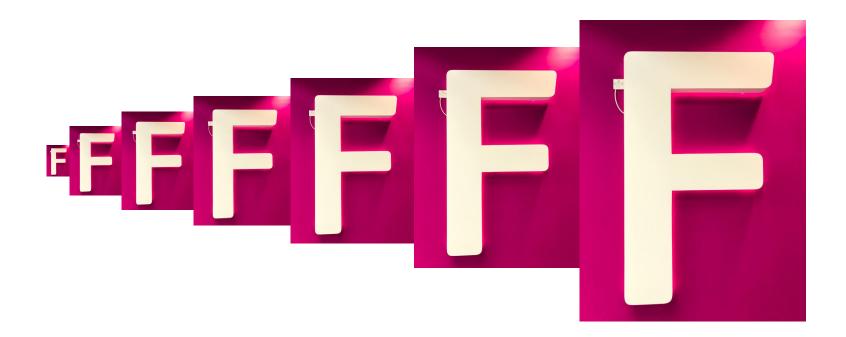
### The F ratio





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





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

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

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.



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

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

## Calculating between-group variance



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



## Mean (Ā)

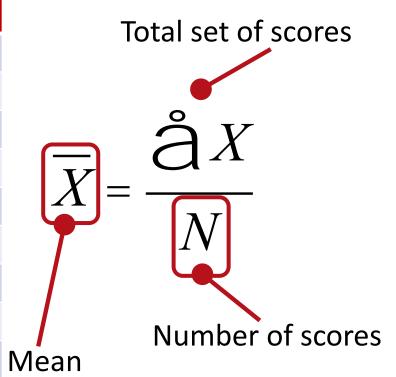








A <sub>1</sub> scores	$A_2$ scores	$A_3$ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6



## Mean (Ā)

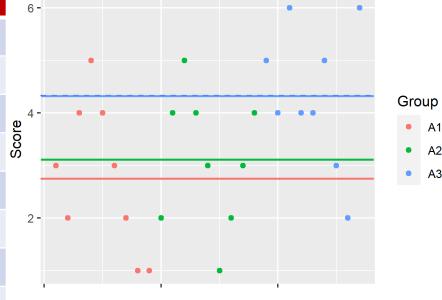








A <sub>2</sub> scores	$A_3$ scores
2	5
4	4
5	6
4	4
3	4
1	5
2	3
3	2
4	6
$\bar{A}_2$ = 3.11	$\bar{A}_3$ = 4.33
	2 4 5 4 3 1 2 3 4



## Grand Mean $(\bar{Y})$









$A_1$ 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
$\bar{A}_1 = 2.78$	$\bar{A}_2 = 3.11$	$\bar{A}_3 = 4.33$

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

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

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

$$\bar{Y} = 3.41$$

## Grand Mean $(\bar{Y})$

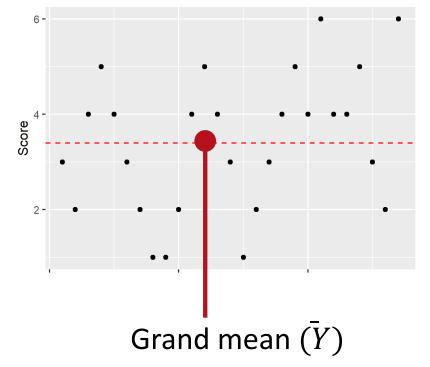








A <sub>1</sub> scores	$A_2$ scores	$A_3$ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1$ = 2.78	$\bar{A}_2$ = 3.11	$\bar{A}_3$ = 4.33



 $\bar{Y} = 3.41$ 

## Total between-group variance



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







 $\bar{Y} = 3.41$ 

$A_1$ 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



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 <sub>2</sub> scores	A <sub>3</sub> scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	1
1	3	2
1	4	6
$\bar{A}_1$ = 2.78	$\bar{A}_2$ = 3.11	$\bar{A_3}$ = 4.33



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

$A_1$ 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	Ş	
1	3	2	
1	4	6	
$\bar{A}_1$ = 2.78	$\bar{A}_2$ = 3.11	$\bar{A}_3$ = 4.33	$\bar{Y} =$



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







$N_{A1}$	= Number of scores for $A_1$
	= 9

$$N_{A2}$$
 = Number of scores for  $A_2$   
= 9

$$N_{A3}$$
 = Number of scores for  $A_3$   
= 9

$A_1$ scores	A <sub>2</sub> 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

#### 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





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







$N_{A1}$	= Number of scores for $A_1$
	= 9

$$N_{A2}$$
 = Number of scores for  $A_2$   
= 9

$$N_{A3}$$
 = Number of scores for  $A_3$   
= 9

A <sub>1</sub> scores	A <sub>2</sub> 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



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







$N_{A1}$ = Number of scores for $A_1$	
= 9	
$N_{A2}$ = Number of scores for $A_2$	
= 9	
·	
$N_{A3}$ = Number of scores for $A_3$	
= 9	

A <sub>1</sub> 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



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







$N_{A1}$	= Nu	mber	of sco	res for	$A_1$
	= 9				
		-			

$$N_{A2}$$
 = Number of scores for  $A_2$   
= 9

$$N_{A3}$$
 = Number of scores for  $A_3$   
= 9

A <sub>1</sub> scores	A <sub>2</sub> 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



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

# Calculating between-group variance



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

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





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

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

#### Up to now...



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

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



# Calculating within-group variance



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





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

#### Mean

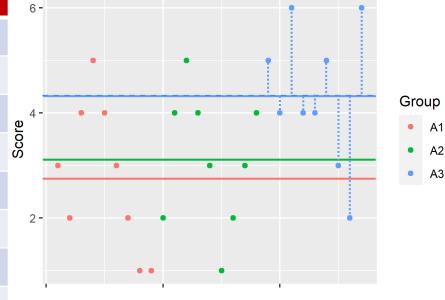








$A_1$ scores	$A_2$ scores	$A_3$ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1$ = 2.78	$\bar{A}_2$ = 3.11	$\bar{A}_3$ = 4.33





total within group variance =  $\frac{SS \ level \ A_1}{total \ within \ group \ degrees \ of \ freedom} = \frac{SS \ level \ A_2}{total \ within \ group \ degrees \ of \ freedom}$ 







 $\bar{Y} = 3.41$ 

$SS$ 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				



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

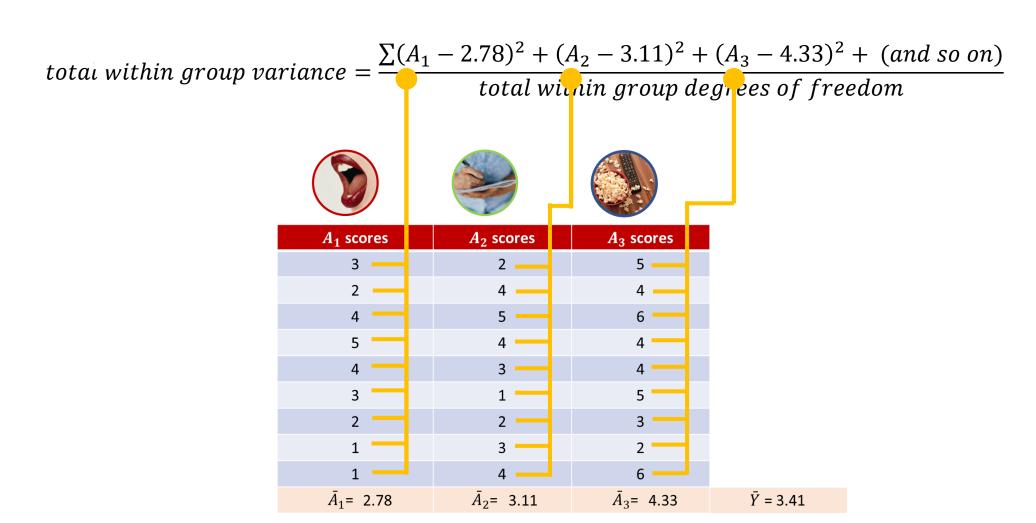






$A_1$ scores	$A_2$ scores	$A_3$ scores
3	2	5
2	4	4
4	5	6
5	4	4
4	3	4
3	1	5
2	2	3
1	3	2
1	4	6
$\bar{A}_1$ = 2.78	$\bar{A}_2$ = 3.11	$\bar{A}_3$ = 4.33





#### 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:

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

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







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



total within group variance = 
$$\frac{42.444}{24}$$
 = 1.769 (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

#### The F ratio





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

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

$$F = 3.414$$

$\nu_1$	1	2	3	4	5	6	7	8	9	10	12	15	20	24	30	40	60	120	∞
1	161	200	216	225	230	234	237	239	241	242	244	246	248	249	250	251	252	253	254
2	18.5	19.0	19.2	19.2	19.3	19.3	19.4	19.4	19.4	19.4	19.4	19.4	19.4	19.5	19.5	19.5	19.5	19.5	19.5
3	10.1	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79	8.74	8.70	8.66	8.64	8.62	8.59	8.57	8.55	8.53
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96	5.91	5.86	5.80	5.77	5.75	5.72	5.69	5.66	5.63
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74	4.68	4.62	4.56	4.53	4.50	4.46	4.43	4.40	4.37
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06	4.00	3.94	3.87	3.84	3.81	3.77	3.74	3.70	3.67
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64	3.57	3.51	3.44	3.41	3.38	3.34	3.30	3.27	3.23
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35	3.28	3.22	3.15	3.12	3.08	3.04	3.01	2.97	2.93
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14	3.07	3.01	2.94	2.90	2.86	2.83	2.79	2.75	2.71
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98	2.91	2.85	2.77	2.74	2.70	2.66	2.62	2.58	2.54
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85	2.79	2.72	2.65	2.61	2.57	2.53	2.49	2.45	2.40
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75	2.69	2.62	2.54	2.51	2.47	2.43	2.38	2.34	2.30
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67	2.60	2.53	2.46	2.42	2.38	2.34	2.30	2.25	2.21
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60	2.53	2.46	2.39	2.35	2.31	2.27	2.22	2.18	2.13
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54	2.48	2.40	2.33	2.29	2.25	2.20	2.16	2.11	2.07
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49	2.42	2.35	2.28	2.24	2.19	2.15	2.11	2.06	2.01
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45	2.38	2.31	2.23	2.19	2.15	2.10	2.06	2.01	1.96
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41	2.34	2.27	2.19	2.15	2.11	2.06	2.02	1.97	1.92
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38	2.31	2.23	2.16	2.11	2.07	2.03	1.98	1.93	1.88
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35	2.28	2.20	2.12	2.08	2.04	1.99	1.95	1.90	1.84
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32	2.25	2.18	2.10	2.05	2.01	1.96	1.92	1.87	1.81
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.23	2.15	2.07	2.03	1.98	1.94	1.89	1.84	1.78
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27	2.20	2.13	2.05	2.01	1.96	1.91	1.86	1.81	1.76
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.18	2.11	2.03	1.98	1.94	1.89	1.84	1.79	1.73
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24	2.16	2.09	2.01	1.96	1.92	1.87	1.82	1.77	1.71
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.15	2.07	1.99	1.95	1.90	1.85	1.80	1.75	1.69
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20	2.13	2.06	1.97	1.93	1.88	1.84	1.79	1.73	1.67
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.12	2.04	1.96	1.91	1.87	1.82	1.77	1.71	1.65
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18	2.10	2.03	1.94	1.90	1.85	1.81	1.75	1.70	1.64
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.09	2.01	1.93	1.89	1.84	1.79	1.74	1.68	1.62
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.00	1.92	1.84	1.79	1.74	1.69	1.64	1.58	1.51
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.92	1.84	1.75	1.70	1.65	1.59	1.53	1.47	1.39
120	3.92	3.07	2.68	2.45	2.29	2.18	2.09	2.02	1.96	1.91	1.83	1.75	1.66	1.61	1.55	1.50	1.43	1.35	1.25
00	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.75	1.67	1.57	1.52	1.46	1.39	1.32	1.22	1.00
			l	l	l							l			l				( I



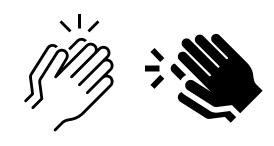
#### The F ratio





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

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



### Lecture 2 – One factor betweenparticipants 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 betweengroup variances
- Degrees of Freedom
- Produced the F-statistic





# Thank you for attention! Questions?

