PSYC214: Statistics
Lecture 1 – Measurement, variance and inferential statistics

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

1

1

Lecture 1 – Measurement, variance and inferential statistics

Agenda/Content

- Experimental science
- Variables
- Descriptive statistics
  - Levels of measurement
  - Measures of central tendency
  - Measures of variability
- Distributions
- Inferential statistics and hypotheses
- Within and between participant designs

2

2

Controlled experiment

A scientific investigation in which both the control group and experimental group(s) are kept under similar conditions apart from the factor under study, so that the effect of influence of that factor can be identified or determined.

3

3

Experimental science

Lancaster University

Population versus sample

- Population is every individual you are interested in

4

4

Experimental science

Lancaster University

Population versus sample

- Population is every individual you are interested in
- The sample is a subset of your population of interest. We examine samples because it is typically impossible to sample everyone in the population

5

5

Experimental science

Lancaster University

Population versus sample

- You should always opt for random sampling, where you pick your sample randomly
- However, in reality, we often use opportunity sampling where we recruit who we have access to

6

6

Variables

Lancaster University

Independent Variable

- The variable (FACTOR) the experimenter manipulates or changes, which may be assumed to have a direct effect (i.e., influences change) on the dependent variable.

Dependent Variable

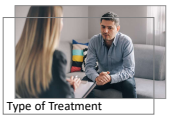
- The outcome of interest. It is the variable being tested and measured in an experiment. It is 'dependent' on the effect (i.e., influence) of the independent variable.

7

7

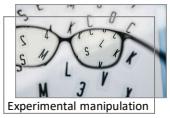
Independent variable

Lancaster University



Type of Treatment

Treatment factor



Experimental manipulation

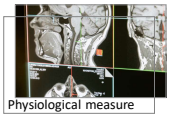
8

8

Dependent variable – i.e., outcome

Lancaster University

Behavioural measure



Physiological measure

Self-report measure

9

9

Statistics

- Use descriptive statistics to describe characteristics and tendencies of your sample
- Use inferential statistics to determine whether the performance and characteristics of your sample generalizes to the population

10

10

PSYC214: Statistics
Lecture 1 – Measurement, variance and inferential statistics – Part 2

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

11

11

Descriptive statistics

1. Levels of measurement
2. Measures of central tendency
3. Measures of variability

12

12

1. Levels of measurement

Lancaster University

Nominal, Ordinal, Interval, Ratio

13

13

1. Levels of measurement - Types

Lancaster University

	Nominal	Ordinal	Interval	Ratio
Categories, Names	●	●	●	●
Rank or order		●	●	●
Known and proportionate intervals			●	●
True zero				●

14

14

1. Levels of measurement - Examples

Lancaster University

	Nominal	Ordinal	Interval	Ratio
Categories, Names				
Rank or order				
Known and proportionate intervals				
True zero				

15

15

1. Levels of measurement - Examples

Lancaster University

Nominal	Numbers assigned to an athlete			
Ordinal	Rank order of winners			
Interval	Performance rating on a scale (0-10)	9.8	9.6	8.6
Ratio	Time to finish in seconds	21.07	21.14	21.79

16

16

1. Levels of measurement - Examples

Lancaster University

Categorical	Nominal	Numbers assigned to an athlete			
	Ordinal	Rank order of winners			
Quantitative	Interval	Performance rating on a scale (0-10)	9.8	9.6	8.6
	Ratio	Time to finish in seconds	21.07	21.14	21.79

17

17

Lancaster University

PSYC214: Statistics
Lecture 1 – Measurement, variance and inferential statistics – Part 3

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

18

18

Descriptive statistics

Lancaster University

1. Levels of measurement
2. Measures of central tendency
3. Measures of variability

19

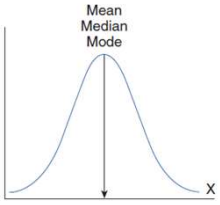
19

2. Measures of central tendency

Lancaster University

A single value that describes the way in which a group of data clusters around a central value, i.e., the centre of the data set

- There are three measures of central tendency
  - Mode
  - Median
  - Mean



20

20

2. Measures of central tendency

Lancaster University

	Nominal	Ordinal	Interval	Ratio
Categories, Names	Mode, % frequencies	Mode, % frequencies	Mode, % frequencies	Mode, % frequencies
Rank or order		Median, percentile	Median, percentile	Median, percentile
Known and proportionate intervals			Mean, standard deviation	Mean, standard deviation
True zero				All above

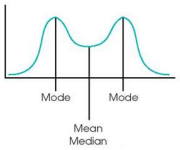
21

21

2. Measures of central tendency - Mode

The most frequent score/data

- Level of measurement: Nominal, ordinal or interval/ratio
- Shape of distribution: Bimodal or multimodal



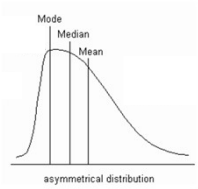
22

22

2. Measures of central tendency - Median

The middle number when data are ordered

- Level of measurement: Ordinal or interval/ratio
- Shape of distribution: Highly skewed



23

23

2. Measures of central tendency - Mean ($\bar{X}$)

The average, is the sum (Σ) of all scores (x) divided by the number of scores (N)

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

24

24

2. Measures of central tendency - Mean ($\bar{X}$)

Lancaster University

The average, i.e., the sum (Σ) of all scores (x) divided by the number of scores (N)

Total set of scores

$$\bar{X} = \frac{\sum X}{N} \qquad \bar{X} = \frac{5+7+7+6+2+3+4+5}{8}$$

Mean of a set of numbers

Number of scores

25

25

2. Measures of central tendency - Mean ($\bar{X}$)

Lancaster University

The average, i.e., the sum (Σ) of all scores (x) divided by the number of scores (N)

Total set of scores

$$\bar{X} = \frac{\sum X}{N} \qquad \bar{X} = 4.875$$

Mean of a set of numbers

Number of scores

26

26

Lancaster University

PSYC214: Statistics
Lecture 1 – Measurement, variance and inferential statistics – Part 4

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

27

27

Descriptive statistics

Lancaster University

1. Levels of measurement
2. Measures of central tendency
3. Measures of variability

28

28

3. Measures of variability

Lancaster University

The spread or dispersion of scores in relation to the midpoint of data.

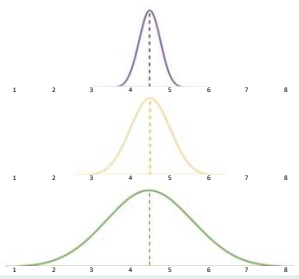
- Range
- Sum of squares
- Variance
- Standard deviation

29

29

3. Measures of variability - why care?

Lancaster University



30

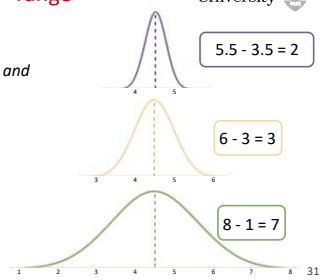
30

3. Measures of variability - range

Lancaster University

The difference between the highest and lowest score

- Subtract the lowest value in the distribution by the highest value



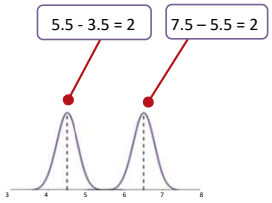
31

31

3. Measures of variability - range

Lancaster University

When is it not useful?



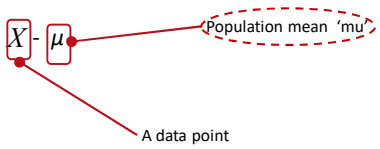
32

32

3. Measures of variability - sum of squares

Lancaster University

1. Calculate difference



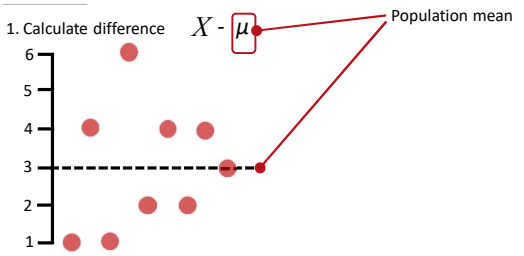
33

33

3. Measures of variability - sum of squares

Lancaster University

1. Calculate difference $X - \mu$



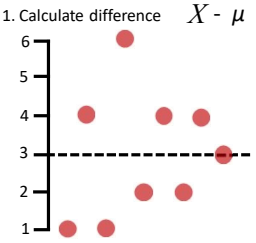
34

34

3. Measures of variability - sum of squares

Lancaster University

1. Calculate difference $X - \mu$



Data point	χ - μ
χ1	-2
χ2	1
χ3	-2
χ4	3
χ5	-1
χ6	1
χ7	-1
χ8	1
χ9	0
Total	0

35

35

3. Measures of variability - sum of squares

Lancaster University

1. Calculate difference $X - \mu$
2. Calculate the sum of squares

$$\text{Sum of squares (SS)} = \sum(\mu - x_i)^2$$

is the sum of all data

is the population mean

is each data point

Data point	χ - μ	(χ - μ)²
χ1	-2	4
χ2	1	1
χ3	-2	4
χ4	3	9
χ5	-1	1
χ6	1	1
χ7	-1	1
χ8	1	1
χ9	0	0
Total	0	22

36

36

3. Measures of variability - variance

- Variance: Average deviation around the mean of a distribution (average of sum of squares)

$$\text{Variance } (\sigma^2) = \frac{\sum(\mu - x_i)^2}{n-1}$$

Sum of squares

Degrees of freedom

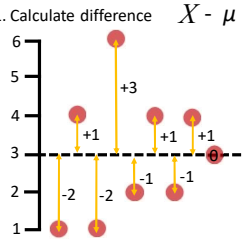
Where μ is the mean
x_i is each data point
n is the number of data points

37

RECAP – find the difference

Lancaster University

1. Calculate difference $X - \mu$



Data point	χ - μ
χ1	-2
χ2	1
χ3	-2
χ4	3
χ5	-1
χ6	1
χ7	-1
χ8	1
χ9	0
Total	0

38

38

RECAP - sum of squares

Lancaster University

1. Calculate difference $X - \mu$
2. Calculate the sum of squares

$$\text{Sum of squares (SS)} = \sum(\mu - x_i)^2$$

is the sum of all data

is the population mean

is each data point

Data point	χ - μ	(χ - μ)²
χ1	-2	4
χ2	1	1
χ3	-2	4
χ4	3	9
χ5	-1	1
χ6	1	1
χ7	-1	1
χ8	1	1
χ9	0	0
Total	0	22

39

39

RECAP – variance (to take into account the number of data points!)

- Variance: Average deviation around the mean of a distribution (average of sum of squares)

$$\text{Variance } (\sigma^2) = \frac{\sum(\mu - x_i)^2}{n-1}$$

Sum of squares

Degrees of freedom

Where μ is the mean
x_i is each data point
n is the number of data points

40

3. Measures of variability – standard deviation

- Standard deviation (σ): Measure of the typical deviation from the mean. It is the squared root of the variance

$$\text{Standard Deviation } (\sigma) = \sqrt{\frac{\sum(\mu - x_i)^2}{n-1}}$$

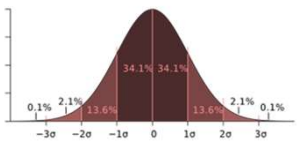
Variance

Where μ is the mean
x_i is each data point
n is the number of data points

41

3. Measures of variability – standard deviation

- Standard deviation (σ): Measure of the typical deviation from the mean. It is the squared root of the variance



42

43

43

Lancaster University

PSYC214: Statistics
Lecture 1 – Measurement, variance and inferential statistics – Part 5

Michaelmas Term
Dr Sam Russell
s.russell1@lancaster.ac.uk

44

44

Inferential statistics

Lancaster University

1. Allow you to draw conclusions based on extrapolations
2. Use data from the sample of participants in the experiment to compare the treatment groups and make generalizations about the larger population of participants
3. Provide a quantitative method to decide if the null hypothesis (H_0) should be rejected

45

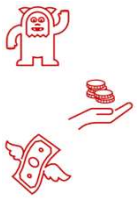
45

Inferential statistics - comparing groups

Often, a researcher is interested in gathering information about different populations in order to compare them

- What is the effect of our treatment/manipulation on an outcome of interest?

- Compare anxiety levels in different age groups
- Compare charitable behaviour before and after Christmas
- Compare Pre and Post consumer behaviour of Covid-19



46

46

Inferential statistics - Hypotheses

H_0 the Null Hypothesis

- H_0: there is no significant difference between the conditions/groups and the null hypothesis is accepted.
- Under H_0, the samples come from the same population.

H_1 the Experimental Hypothesis

- H_1: there is a significant difference between the conditions/groups and the null hypothesis is rejected.
- Under H_1, the samples come from the different populations.

47

47

Inferential statistics - (Non)parametric tests

- Statistical tests can be separated into:
  - Parametric
  - Non-parametric

While parametric tests are the norm in psychology and are generally more powerful than non-parametric tests, they require that the scores be an interval or ratio measure and there needs to be homogeneity of variance

48

48

Example set 1

Lancaster University

George observed a group of patients before, during and after the administration of a drug X to evaluate the effectiveness of the treatment.

Fozia measured participants' scores on a Psychological test of creativity in the morning, noon and afternoon in order to see whether there are any differences throughout the day.

Kenji measured the visual acuity of a single group of observers. He asked each subject to complete a vision test after they wore each of the five different brands of contact lenses under investigation.



49

49

In all cases

Lancaster University

The same participant (used to be called **subject**) is being tested in different conditions:

- Before, during and after treatment
- Morning, noon and afternoon
- Five different brands of contact lenses

As each participant (subject) provides scores on the different conditions, we say that the measures are **related** and **correlated**

50

50

In all cases

Lancaster University

In these three cases, the Independent Factor is said to be a WITHIN-subject factor as it is altered within each subject.

51

51

Example set 2

Lancaster University

Lucy is interested in age differences in mental toughness. She recruits 20 young adults, 20 middle-aged adults and 20 older adults, and asks them to complete a Hardiness Test.

Manuel is studying whether statistics lectures are more effective in the morning or in the afternoon. He administers a pop quiz to the morning and afternoon classes and compares the performance.

Mo wants to examine differences in personality traits between students from different universities. He recruits students from Lancaster, York and Bath and asks them to complete a sociability questionnaire.



52

52

In all cases

Lancaster University

The different subjects are being tested in different conditions

- Young, Middle-Aged and Older Adults
- Morning class and Afternoon class
- Lancaster, York and Bath Universities

Because different observer provides scores on the different conditions, we say that the measures are **unrelated** and **uncorrelated**

53

53

In all cases

Lancaster University

In this case, the Independent factor is said to be a BETWEEN-subject factor as it is altered between each subject.

54

54

Lecture 1 – Measurement, variance and inferential statistics

Lancaster University

Review

- Experimental science
- Variables and levels of measurement
- Descriptive statistics
  - Levels of measurement
  - Measures of central tendency
  - Measures of variability
- Distributions
- Inferential statistics and hypotheses
- Within and between participant designs

55

55

Lancaster University

Thank you for attention!

56

56