

Introduction to Hypothesis Testing

These questions are for you to complete in your own time. Please use them for extra revision of the concepts discussed in lectures and practised in tutorials. For help with these questions, you can ask any of your tutorial leaders or visit Maths-Aid. Maths-Aid can be found in Room 1.16 on Level 1 of the Marjorie Robinson Library and can be contacted at mathsaid@ncl.ac.uk.

Population vs. Sample Statistics

1. Students in the UK have an average IQ of 110 with a variance of 64. The IQs of two randomly selected students are measured at five different universities across the UK. The data is shown in the table below.

Student	IQ Score
1	115
2	128
3	150
4	146
5	132
6	99
7	121
8	125
9	111
10	107

- a) If possible, identify the population mean (μ), the sample mean (\bar{x}), the population standard deviation (σ) and the sample standard deviation (s) for this data set.
- b) You find out that the five universities that students were sampled from make up the top 5 universities in the UK. How does this change your interpretation of these statistics?

- A teacher is interested in the performance of children in year 8 on working memory task. For this the teacher assesses the digit span (the number of consecutive digits each student can recite) children in her year 8 class. She obtains the following data.

Student	Digit Span	Student	Digit Span
1	5	11	5
2	8	12	5
3	5	13	3
4	6	14	4
5	3	15	5
6	8	16	10
7	7	17	5
8	6	18	6
9	4	19	7
10	4	20	7

If possible, identify the population mean (μ), the sample mean (\bar{x}), the population standard deviation (σ) and the sample standard deviation (s) for this data set.

- A researcher sends out a questionnaire to all houses within a certain postcode area. They would like to know the mean and standard deviation of the number of years spent in full time education for all the people who responded to the questionnaire. Should they use the formula for the population or sample standard deviation?
- For their dissertation, a psychology student is investigating how number of years in education impacts upon performance in the Stroop task. They gather data from a selection of 18 to 22 year olds in Universities who have never left education and from a group of age and gender matched individuals who left school at 16. Should they use the formula for the sample standard deviation or the population standard deviation when calculating their descriptive statistics?

Formulating Statistical Hypotheses

State the null and alternative hypotheses for the following experiments.

- Colour blindness is known to be more common in men than women. A researcher is interested in whether colour blindness in men is associated with career choice (specifically in artists and accountants). The researcher tests all male artists and accountants that she can find in the local area for colour blindness.
- A group of psychology researchers at Newcastle University suspect that having access to a smart phone from a young age influences social skills. They expect that delaying exposure to a smart phone until college (16 years old) will lead to improved social skills in comparison to being given a smart phone much earlier in life. The researchers recruit two groups of subjects, all aged 10, who have not yet been given access to a smart phone for a longitudinal

study. The first group are given a phone at age 11, the second group are never given a phone. The subjects' social skills are monitored until their 17th birthdays.

One or Two Tailed Testing and Significance Levels

1. Are any of these p -values significant at the 1%, 5% or 10% level?
 - a) $p = 0.333$
 - b) $p = 0.082$
 - c) $p = 0.049$
 - d) $p = 0.113$
 - e) $p = 0.077$
 - f) $p = 0.002$

2. As a third year project, a student tests the following hypothesis and obtains a p -value of $p = 0.003$.

H_0 : There is no association between an individual's income and their criminal record. H_1 : Individuals with lower income are more likely to have a criminal record.

- a) Is this a one or two tailed test?
- b) Fill in the blanks in the following statement to interpret this information.

The result is significant at the ___% level. This means there is significant evidence to _____ the _____ hypothesis. Hence, there is significant evidence to suggest that _____.

3. You have tested the statistical hypotheses below and obtained a p -value of $p = 0.025$.

H_0 : There is no association between IQ and susceptibility to mental illness.

H_1 : There is some association between IQ and susceptibility to mental illness.

- a) Construct a statement like the one in the question above to interpret this information.

Solutions

Population vs. Sample Statistics

1. a) $\mu = 110$, $\bar{x} = 123.4$, $\sigma = 8$, $s = 16.37$
2. μ not given, $\bar{x} = 5.65$, σ not given, $s = 1.79$
3. Sample
4. Sample

Formulating Statistical Hypotheses

1. H_0 : There is no association between colour blindness and career choice
 H_1 : There is *some* association between colour blindness and career choice
2. H_0 : There is no association between having access to a smart phone and social skills
 H_1 : Those not exposed to a smart phone will have better social skills at age 16

One or Two Tailed Testing and Significance Levels

1. a) None
b) 10%
c) 5%
d) None
e) 10%
f) 1%
2. a) One tailed
b) The result is significant at the *1% level*. This means there is significant evidence to *reject* the *null* hypothesis. Hence, there is significant evidence to suggest that *individuals with lower income are more likely to have a criminal record*.