Data Science and Statistics in Research: unlocking the power of your data

Session 2.4:

Hypothesis testing II

Outline

Independent t-test

EXAMPLE: MOTOR TREND CAR ROAD TESTS

- The `mtcars` dataset in `R` contains fuel consumption and 10 other aspects of 32 cars from 1973-74.
- Suppose a car company wants to see if the true mean fuel consumption (in miles per gallon) differs significantly in automatic and manual cars.
- We set up the following hypothesis
  - null: the true mean fuel consumption is the same between automatic and manual cars
  - alternative: the true mean fuel consumption differs between automatic and manual cars.
- How do we test this hypothesis?

Independent t-test

What is it for?

- An independent *t*-test is used to compare the means of two independent groups.

What does it do?

- It tells you whether the means of the two groups are significantly different.

What is the output?

- A p-value which indicates the probability that the data are consistent with the null hypothesis of no difference between the groups.

Independent t-test

How do you interpret the output?

- If the p-value is small, typically 0.05, then we have enough evidence to reject the null hypothesis.

What restrictions are there on its use?

- If the data are severely skewed (non-normal) or sample sizes are small, then a non-parametric test may be more suitable.

Example: Motor Trend Car Road Tests

- We want to test the following hypotheses
  - null: the true mean fuel consumption is the same between automatic and manual cars
  - alternative: the true mean fuel consumption differs between automatic and manual cars.
- We choose a significance level of 0.05 for our test and construct the following statistical decision rule
  - **IF** the p-value is less than 0.05
  - **THEN** we have enough evidence to reject the null hypothesis
  - **OTHERWISE** there is not enough evidence to reject the null hypothesis.

Example: Motor Trend Car Road Tests

- There are 32 samples in our dataset.
- Automatic cars have a sample mean 17.1 and sample standard deviation 3.8.
- Manual cars have a sample mean 24.4 and sample standard deviation 6.2.
- Performing an independent t-test gives us a p-value of 0.0013.
- This is less than the chosen significance level.
- We therefore have enough evidence to reject the null hypothesis.
- We conclude that the true mean fuel consumption is different between automatic and manual cars.

Paired t-test

Example: Diet Comparison

- A study was performed to assess the effect of different diets on LDL cholesterol in men.
- 12 men went under two diets, with a 'washout' period between them.
- Cholesterol was measured in all men after each diet.
- A dietician wanted to test whether there is a difference in the true mean cholesterol after the two diets
  - null: there is no difference in the true mean cholesterol between the two diets
  - alternative: there is a difference in the true mean cholesterol between the two diets.
- How do we test this hypothesis?

PAIRED T-TEST

What is it for?

- A paired *t*-test is used to compare the means of two paired groups, when the two groups are of equal size and subjects in one sample are paired with one in the other.

What does it do?

- It tells you whether the means of the two groups are significantly different.

What is the output?

- A p-value which indicates the probability that the data are consistent with the null hypothesis of no difference between the groups.

Paired t-test

How do you interpret the output?

- If the p-value is small, typically 0.05, then we have enough evidence to reject the null hypothesis.

What restrictions are there on its use?

- If the data are severely skewed (non-normal) or sample sizes are small, then a non-parametric test may be more suitable.

Example: Diet Comparison

- We want to test the following hypotheses
  - null: there is no difference in the true mean cholesterol between the diets
  - alternative: there is a difference in the true mean cholesterol between the diets
- We choose a significance level of 0.05 for our test and construct the following statistical decision rule
  - **IF** the p-value is less than 0.05
  - **THEN** we have enough evidence to reject the null hypothesis
  - **OTHERWISE** there is not enough evidence to reject the null hypothesis.

Example: Diet Comparison

- The sample mean cholesterol level for diet 1 is 4.73, and for diet 2 is 4.26.
- The sample mean of the differences is 0.46.
- Performing a paired t-test gives us a p-value of 0.0018.
- This is less than the chosen significance level.
- We therefore have enough evidence to reject the null hypothesis.
- We conclude that the true mean level of cholesterol differs between the diets.

Chi-squared test

Testing proportions

- So far we have used hypothesis tests between groups using continuous data.
- There may be a situation where collected data will be categorical.

Example: Low Birth Weights in India

- A study was performed to assess the effects of smokeless tobacco use and low birth weight in India.
- We have frequencies of low birth-weight babies given by categories related to the mothers' tobacco usage.

	Regular Birth Weight	Low Birth Weight	Total
Non-users	646	160	806
1–4 times	85	27	112
5 or more times	35	21	56
Total	766	208	974

Testing proportions

- In this case, using t-tests will not be appropriate to test whether there is a difference between groups.
- However, we can test for association by comparing proportions of the outcomes of two categorical variables.

Example: Low Birth Weights in India

- We can view the data as proportions rather than frequencies

	Regular Birth Weight	Low Birth Weight	Total
Non-users	0.80	0.20	1.00
1–4 times	0.76	0.24	1.00
5 or more times	0.62	0.38	1.00
Total	0.79	0.21	1.00

- We observe that the proportions for low-birth rate increase as smokeless tobacco use increases.
- This indicates that there is some association between the two outcomes.

Example: Low Birth Weights in India

- Suppose a doctor wanted to test whether the association between smokeless tobacco use and low-birth rate is significant.
- We set up the following hypothesis
  - null: there is no difference in the proportions between low-birth rate and smokeless tobacco use (i.e. the outcomes are independent and there is no association)
  - alternative: there is a difference in the proportions between low-birth rate and smokeless tobacco use (i.e. the outcomes are dependent and there is an association)
- How do we test this hypothesis?

Chi-Squared Test

What is it for?

- A Chi-Squared test is used to compare the proportions of outcomes in different groups.

What does it do?

- It tells you whether there is a difference in the proportions between the groups.

What is the output?

- A p-value which indicates the probability of the data arising from the null hypothesis that there is no association between the proportions of each outcome and the group they are in.

Chi-Squared Test

How do you interpret the output?

- If the p-value is small, typically 0.05, then there is enough evidence to reject the null hypothesis. This will represent an association between outcome and group.

What restrictions are there on its use?

- Sample sizes need to be large enough for the normal assumption to the binomial distribution to hold. Fisher's exact test may be more appropriate for small samples.

Example: Low Birth Weights in India

- We want to test the following hypotheses
  - null: there is no difference in the proportions between low-birth rate and smokeless tobacco use (i.e. the outcomes are independent and there is no association)
  - alternative: there is a difference in the proportions between low-birth rate and smokeless tobacco use (i.e. the outcomes are dependent and there is an association).

Example: Low Birth Weights in India

- We choose a significance level of 0.05 for our test and construct the following statistical decision rule
  - **IF** the p-value is less than 0.05
  - **THEN** we have enough evidence to reject the null hypothesis
  - **OTHERWISE** there is not enough evidence to reject the null hypothesis.
- Performing a Chi-Squared test gives us a p-value of 0.0059.
- This is less than the chosen significance level.
- We therefore reject the null hypothesis in favour of the alternative.
- We conclude that the proportions are not the same between the groups and there is an association between low-birth weight and smokeless tobacco use.

Any Questions?