8

STATISTICAL HYPOTHESIS TESTING

QUESTION 1.

Convert the following statement into a form which can be used in hypothesis testing. In other words, formulate the null and alternative hypotheses.

A group of students have agreed to take part in an experiment to find out how their exam results improve if they revise. The mean mark for the 50 students involved in previous exams was 45%. They all have agreed to record how much revision they do one week before the exam.

If the mean exam result is greater than 45%, then revising improves the exam result.
If the mean mark is less or the same then revising does not improve exam performance.

The null hypothesis $H_0: \mu_\sigma = \mu_0$
The alternative hypothesis $H_1: \mu_\sigma \neq \mu_0$

QUESTION 2.

Interpret the following hypotheses, where μ_0 represents the hypothesised value:

$H_0: \mu = \mu_0$	sample mean the same as the population mean.
$H_1: \mu \neq \mu_0$	sample mean not equal to the population mean.
$H_1: \mu > \mu_0$	sample mean greater than the population mean.
$H_1: \mu < \mu_0$	sample mean less than the population mean.

QUESTION 3.

Complete the following statement:

Hypotheses are statements about the mean in which either the null hypothesis or the alternative hypothesis can be true but not both.

QUESTION 4.

Complete the following statement:

The golden rule to remember for hypothesis testing is that hypotheses refer to the population and the test refers to the sample.

QUESTION 5.

Decide whether the following statements are true or false:

A result is statistically significant if you are 95% certain your results are genuine.	True
It is possible to be 100% certain your sample results are true in the population.	False
My data can tell me if the correct result was a guess.	True

QUESTION 6.

In the box below, explain the difference between a one-tailed test and a two-tailed test.

The difference between the tests is:

A one-tailed test is used if the analysis refers to a result being either greater than or less than the test statistic but not both.

A two-tailed test is used when the analysis could be either greater than or less than the test statistic.

QUESTION 7.

The mean lifetime of a sample of 100 water valves is found to be 1570 hours with a standard deviation of 120 hours. If α is the mean lifetime of all the water valves, test the hypothesis that the mean lifetime is equal to 1600 hours against the alternative hypothesis that it is not. You should test at (a) 0.05 and (b) 0.01 level of significance.

H_0: $\mu =$	1600 hrs
H_1: $\mu \neq$	1600 hrs

Complete the following sentences:

The most appropriate test is a two-tailed test.

H_0 will be rejected if the Z score of the sample mean is outside the range –1.96 to +1.96.

If the Z score of the sample mean is within the range –1.96 to 1.96, H_0 will be accepted or a decision will be withheld subject to further investigation.

$\mu_{\bar{x}} =$	1600
$\sigma_{\bar{x}} =$	12
$Z =$	–2.5

At the 0.05 significance level H_0 is rejected since the Z score is outside the range –1.96 to + 1.96.

At the 0.01 significance level H_0 is accepted since the Z score is inside the range –2.58 to + 2.58.

QUESTION 8.

Use the data from question 7 to compute the p-value.

The p-value of the two-tailed test is $P(Z \leq -2.5) + P(Z \geq 2.5) = 0.0124$.

This means the probability of a mean lifetime for the water valves of less than 1570 hours or more than 1630 hours would occur by chance if H_0 were true.

QUESTION 9.

You are the teacher of two classes of students who were given the same exam. Class A had 50 students and class B 40. The mean mark for class A was 78 with a standard deviation of 7 and in class B the mean mark was 74 with a standard deviation of 8.

You would like to be 95% confident that the classes performed equally well. However, you would be reasonably happy if you could be 90% confident that the classes performed the same.

Also you would like to know the probability that the observed statistics would occur in the same population.

Complete the following tasks:

a) Construct the null and alternative hypotheses and explain the hypotheses.

		Explanation
H_0:	$\mu_1 = \mu_2$	difference due to chance
H_1:	$\mu_1 \neq \mu_2$	there is a significant difference between classes

Assuming both classes come from the same population, calculate the difference in the means and the standard deviations:

$$\mu_{\bar{X}_1 - \bar{X}_2} = 0$$

$$\sigma_{\bar{X}_1 - \bar{X}_2} = 1.606$$

Calculate the value for Z:

$$Z = \frac{\bar{X}_1 - \bar{X}_2}{\sigma_{\bar{X}_1 - \bar{X}_2}} = -2.49$$

Complete the following:

I would use a two-tailed test with a significance level of 0.05 to be 95% confident and a significance level of 0.01 to be 99% confident that the classes' performances were the same.

The probability that the observed statistics would occur in the same population is $P(Z \leq -1.96), P(Z \geq 1.96)$ at the 0.05 level of significance.

My conclusions are
At the 0.05 level of significance the results are significant but at the 0.01 level they are not.
The conclusion is the results are probably significant.

MINI PROJECT

You have been appointed the Quality Control Manager of a company that manufactures roller bearings. You have suspicions that a certain machine is not producing roller bearings within the expected tolerances.

The machine in question is expected to produce roller bearings with a mean diameter of 0.574 cm and a standard deviation of 0.008 cm. The standard quality control procedure is to take a sample of 6 roller bearings every 2 hours and calculate the mean from this sample.

Tasks:

a) Construct the null and alternative hypotheses.
b) Perform the necessary calculations to determine if your suspicions are justified. You would like to be at least 90% confident that the machine is working within the required tolerances.
c) You have read in the company's Quality Control manual that the Production Director expects all machines to work to a 3σ confidence level, but you have never come across this specification before. You need to know if the testing you are doing conforms to the company's expectations. Write down a brief explanation of the 3σ limits and comment on how your testing compares with this.

And finally...

If there are multiple worlds, how do you know which one is the 'null' and which one is the 'alternative'?