**3. Statistical Hypothesis Testing**

AS Sampling Recap

**Terms & Notation**

A ***sample*** provides a set of data values of a random variable, drawn from all such possible values, the ***parent population***.

A representation of the items available to be sampled is called the ***sampling frame***. This could for example be a list of sheep in a flock, a map marked with a grid or an electoral register. In many situations no sampling frame exists, nor is it possible to devise one, for example for the cod in the North Atlantic. The proportion of the available items that are actually sampled is called the ***sampling fraction***.

A ***parent population***, often just called the ***population***, is described in terms of its ***parameters***, such as its mean, μ and variance σ. By convention Greek letters are used to denote these parameters.

A value derived from a sample is written in Roman letters: mean, , variance *s2*, etc. Such a number is a value of a ***sample statistic***. When sample statistics are used to estimate the parent population parameters they are called ***estimators***. If a value suggested for a population parameter is an estimate obtained from a sample, rather than its true value, this is denoted by a circumflex accent, ^: estimated mean, ; estimated standard deviation, ; estimated variance, .

Thus if you take a random sample in which the mean is , you can use to estimate the parent mean, μ.

So . If in a particular sample , then

The true value of μ will generally be different from

Upper case letters, *X*, *Y*, etc., are used to describe random variables, and lower case letters, *x*, *y*, etc. to denote particular values of them.

For example if *X* is the random variable ‘the score when a fair six-sided dice is rolled’, then

P(*X* = *x*) = 1/6 where *x* = 1, 2, 3, 4, 5, 6

There are essentially two reasons why you might wish to take a sample

* To estimate the values of the parameters of the parent population
* to conduct a hypothesis test

In the AS unit you looked at sampling techniques such as random sampling, systematic sampling, stratified sampling and opportunity sampling. You were also required to be able select appropriate sampling methods for different parent populations and examine critically sampling methods selected by others.

At A2, we will be concentrating on hypothesis testing the mean of a sample in comparison to that of the parent population.

**Interpreting sample data using the Normal distribution**



How do you interpret these figures? Do you think Ama has a point when she states she is worried that the greenhouse effect is already happening in Avonford?

No mention is made in the article about his sampling method, is this something we need to consider?

**Estimating the parent mean, μ**

Ray Sharp’s sample data were as follows

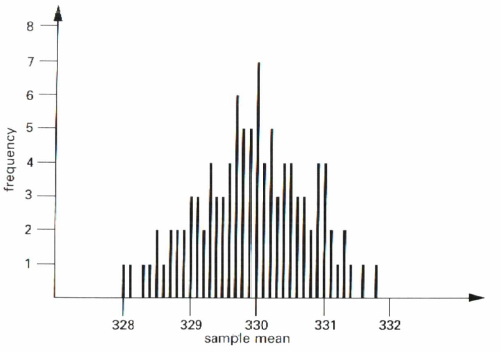
336, 334, 332, 332, 331, 331, 330, 330, 328, 326

His intention in collecting them was to estimate the mean of the parent population.

The sample mean, =

What does this tell us about the parent mean, μ?

If you took 100 such samples, each of size 10, the distributions of their means would look similar to the chart below:



You will notice that this distribution looks rather like the Normal distribution and so may well wonder if this is indeed the case. The answer, provided by the ***central limit theorem***, is yes.

**The Central Limit Theorem**

*For samples of size n drawn from a distribution with mean μ and finite variance σ2, the distribution of the sample mean is approximately for sufficiently large n.*

This theorem is fundamental to much of statistics. It deals with the distribution of sample means. This is called the *sampling distribution* and there are three aspects to it.

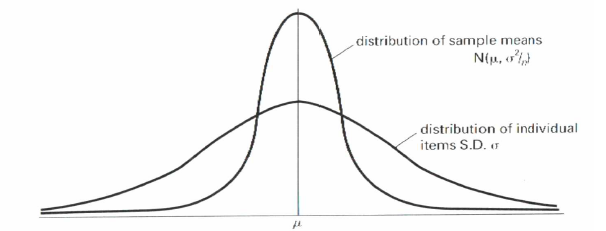
1. The mean of the sample means is μ, the parent mean of the original distribution.
2. The standard deviation of the sample means is . This is often called the *standard error*.

Within a sample you would expect some values above the parent mean, others below it, so that overall the deviations would tend to cancel each other out, and the larger the sample the more this is the case. Consequently the standard deviation of the sample means is smaller than that of the individual items (parent population), by a factor of .

1. The distribution of sample means is approximately Normal.

This is the most surprising part of the theorem. Even if the underlying parent distribution is not Normal, the distribution of the means of the samples of a particular size drawn from it is approximately Normal. The larger the sample size, n, the closer is this distribution to the Normal. For any given value of n the sampling distribution will be closest to Normal where the parent distribution is not unlike the Normal.

In many cases the value of n does not have to be particularly large. For most parent distributions you can rely on the distribution of sample means being Normal if n exceeds about 20 or 25.



So how does all this relate to the air quality in Avonford?

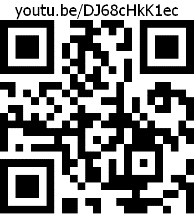
Ray Sharp was mainly interested in establishing data on CO2 levels for Avonford. The newspaper reporter however wanted to know whether levels were above the normal, and so she could have set up and conducted an hypothesis test.

**Hypothesis test for the mean using the Normal distribution**

Eg1 Ama Williams believes that the CO2 level in Avonford has risen above the usual level of 328 parts per million. A sample of 10 specimens of Avonford air are collected and the CO2 level within them is determined. The results are as follows:

336, 334, 332, 332, 331, 331, 330, 330, 328, 326

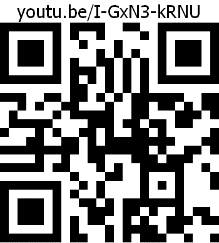
Extensive previous research has shown that the standard deviation of the levels within such samples is 2.5.

Use these data to test, at the 0.1% significance level, Ama’s belief that the level of CO2 at Avonford is above normal.

**Whenever concluding the outcome of your test make sure it is given in terms of the context of the problem. First state whether H0 is to be accepted or rejected, then make a statement beginning “there is evidence to suggest that…” or “there is insufficient evidence to suggest that…”. You should NOT write “this proves that…” or “so the claim is right”. You are not proving anything, only considering evidence.**

**Note that a hypothesis test should be formulated before the data are collected and not after. If the sample data lead you to form a hypothesis, then you should plan a suitable test and collect further data on which to conduct it.**

Eg2 A researcher is interested in establishing the mean IQ of the population and uses a test which is known to give scores with a standard deviation of 15. Initially the researcher takes a sample of 100 people and finds the mean of their scores to be 105.

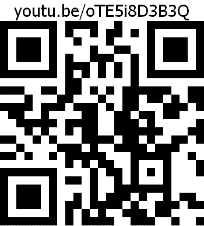
Knowing that 50 years earlier the mean score on this test was 100, the researcher puts forward the theory that people are becoming more intelligent (as measured by this particular test). She selects a random sample of 500 people all of whom take the test. Their mean score is 103.22.

Carry out a suitable hypothesis test on the researchers theory, at the 1% significance level.

Eg3 A machine is designed to make paperclips with mean mass 4.00g and standard deviation 0.08g. The distribution of the masses of the paperclips is Normal. Find

1. the probability that an individual paperclip, chosen at random, has a mass greater than 4.04g
2. the standard error of the mass for random samples of 25 paperclips
3. the probability that the mean mass of a random sample of 25 paperclips is greater than 4.04

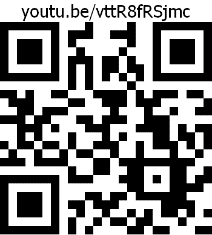
A quality control officer weighs a random sample of 25 paperclips and finds their total mass to be 101.2g

1. Conduct a hypothesis test at the 5% significance level of whether this provides evidence of an increase in the mean mass of the paperclips. State your null and alternative hypotheses clearly.

Eg4 A chemical is packed into bags by a machine. The mean weight of the bags is controlled by the machine operator, but the standard deviation is fixed at 0.96kg. The mean weight should be 50kg, but it is suspected that the machine has been set to give underweight bags. If a random sample of 36 bags has a total weight of 1789.20kg, is there evidence to support the suspicion? (You must state the null and alternative hypotheses and you may assume that the weights of the bags are Normally distributed.

Eg5 A certain type of lizard is known to have mean mass 72.7g with standard deviation 4.8g. A zoologist finds a colony of lizards in a remote place and is not sure whether they are of the same type. In order to test this, she collects a sample of 12 lizards and weighs them with the following results:

80.4 67.2 74.9 78.8 76.5 75.5 80.2 81.9 79.3 70.0 69.2 69.1

1. Write down, in precise form, the zoologist’s null and alternative hypotheses, and state whether a 1-tail or 2-tail test is appropriate.
2. Carry out the test at the 5% significance level and write down your conclusion.
3. Would your conclusion have been the same at the 10% significance level?

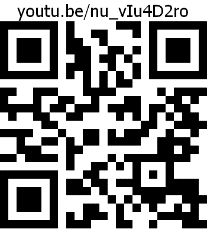
Eg6 Some years ago the police did a large survey of the speeds of motorists along a stretch of motorway, timing cars between two bridges. They concluded that their mean speed was 80mph with standard deviation of 10mph.

Recently the police wanted to investigate whether there had been any change in the motorists’ mean speed. They timed the first 20 green cars between the same two bridges and calculated their speeds in mph to be as follows:

85 75 80 102 78 96 124 70 68 92

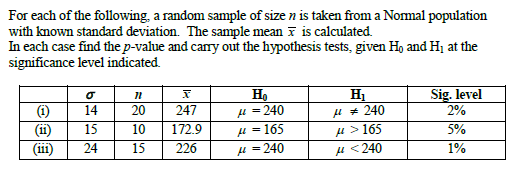
84 69 73 78 86 92 108 78 80 84

1. State suitable null and alternative hypotheses and use the sample data to carry out a hypothesis test at the 5% significance level. State the conclusion.

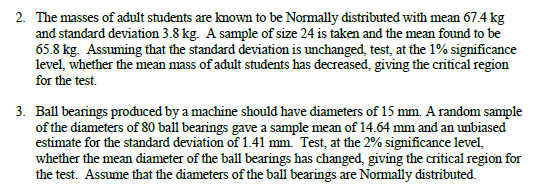
One of the police officers involved in the investigation says that one of the cars in the sample was being driven exceptionally fast and that its speed should not be included within the sample data.

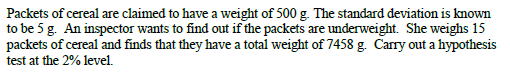
1. Would the removal if this outlier alter the conclusion?

**Exercise 3.1**



1.

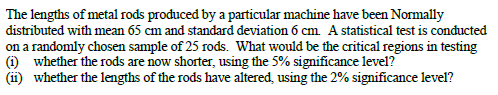




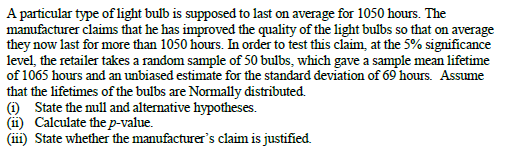
4.



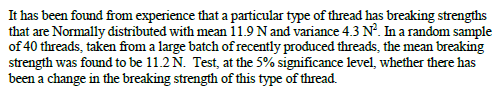
5.



6.



7.



8.

**Answers**

(1i) 0.0127, accept null (ii) 0.0479, reject null (iii) 0.0119, accept null

(2) CR ≤ 65.6, accept null

(3) CR < 14.6, accept null

(4) either p-value = 0.015 or CR < 497.35, reject null

(5) either p-value = 0.067 or CR > 10.81, accept null

(6i) CR < 63.03 (ii) CR < 62.21 and CR > 67.79

(7) p-value = 0.062, accept null

(8) either p-value = 0.016 or CR < 11.26, reject null