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, $\bar{x}$, variance s^2, 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, $\hat{x}$; estimated standard deviation, $\hat{\sigma}$; estimated variance, $\hat{\sigma}^2$.

Thus if you take a random sample in which the mean is $\bar{x}$, you can use $\bar{x}$ to estimate the parent mean, μ.
So $\hat{\mu} = \bar{x}$. If in a particular sample $\bar{x} = 23.4$, then $\hat{\mu} = 23.4$
The true value of μ will generally be different from $\hat{\mu}$

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

THE AVONFORD STAR

Avonford set to become greenhouse?

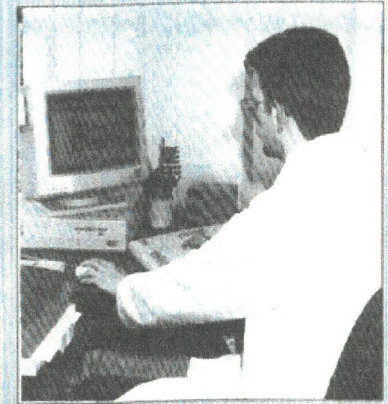
From our Science Correspondent Ama Williams

On a recent visit to the Avonford Community College, I was intrigued to find experiments being conducted to measure the level of carbon dioxide in the air we are all breathing. Readers will of course know that high levels of carbon dioxide are associated with the greenhouse effect.

Lecturer Ray Sharp showed me round his laboratory. "It is delicate work, measuring parts per million but I am trying to establish what is the normal level in this area. Yesterday we took ten readings and you can see the results for yourself: 336, 334, 332, 332, 331, 331 330, 330, 328, 326."

When I commented that there seemed to be a lot of variation between the readings, Ray assured me that that was quite in order.

"I have taken hundreds of these measurements in the past," he said. "There is always a standard deviation of 2.5. That's just natural variation."



Scientist Ray Sharp's tests could spell trouble for Avonford.

I suggested to Ray that his students should test whether these results are significantly above the accepted value of 328 parts per million. Certainly they made me feel uneasy. Is the greenhouse effect starting here in Avonford?

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?

$\frac{9}{10}$ of the values in the sample are greater or equal to than 328, so yes.

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

Yes. If a suitable sampling procedure has not been used, then the resulting data may be worthless, even misleading.

Estimating the parent mean, μ

Ray Sharp's sample data were as follows

336+334+332+332+331+331+330+330+328+326 =

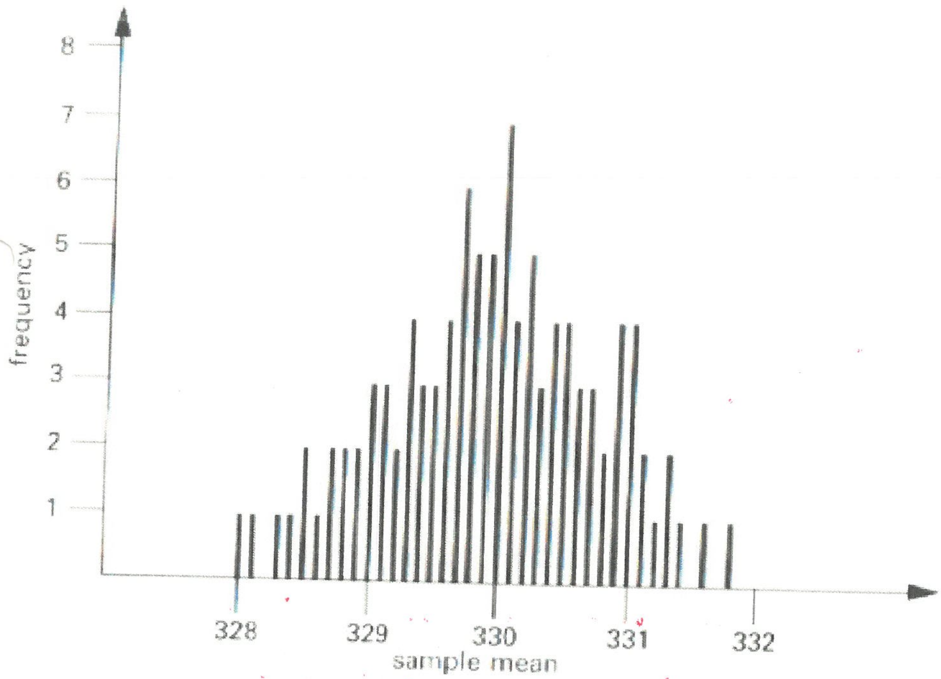
His intention in collecting them was to estimate the mean of the parent population.

The sample mean, $\bar{x}$ = ~~3346~~ 3310 ÷ 10 = 331

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

- That the parent mean ~ 331, but does not tell you that it is definitely exactly 331
- If Ray took another sample, its mean would probably not be 331, but you would be surprised if it were very far away from it.
- If he took lots of samples, all of size 10, you would expect their means to be close together but certainly not all the same.

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 σ^2, the distribution of the sample mean is approximately $N\left(\mu, \frac{\sigma^2}{n}\right)$ 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.

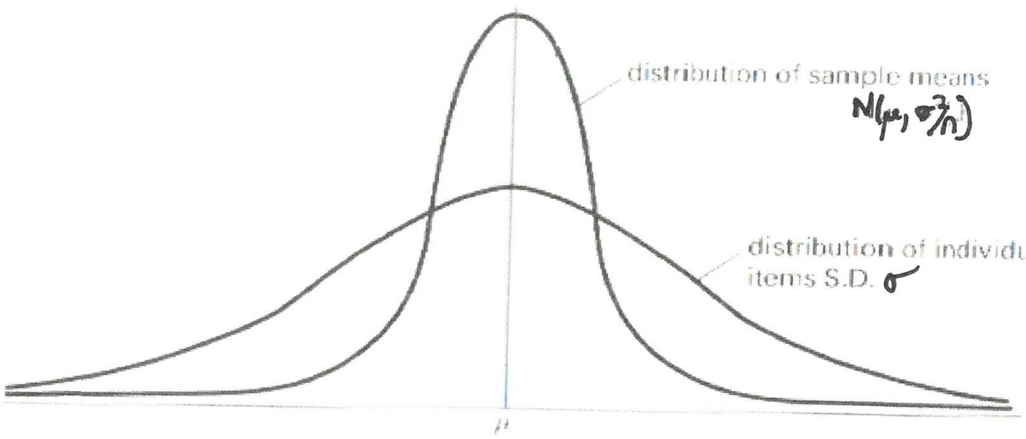
1. The mean of the sample means is μ, the parent mean of the original distribution.
2. The standard deviation of the sample means is $\frac{\sigma^2}{\sqrt{n}}$. This is often called the *standard error.*

   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 $\sqrt{n}$.

3. 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 CO_2 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 CO_2 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 CO_2 level within them is determined. The results are as follows:

336, 334, 332, 332, 331, 331, 330, 330, 328, 326 $\bar{x} = 331$.

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 CO_2 at Avonford is above normal.

Parent Population, X is the random variable parts of CO_2 per million in a specimen $X \sim N(328, 2.5^2)$

Sample Population of Means, $\bar{X} \sim \left(328, \frac{2.5}{10}^2\right) \Rightarrow \sigma = \frac{2.5}{\sqrt{10}}$

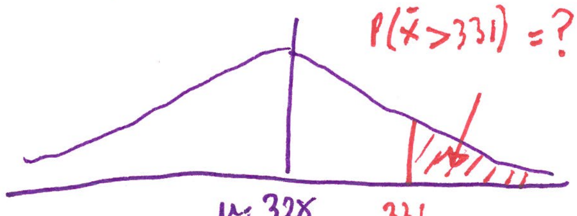
Calculated sample mean $\bar{x} = 331$.

Null Hypothesis: $H_0 : \mu = 328$ The level of CO_2 in Avonford is normal

Alternative Hyp: $H_1 : \mu > 328$ " " " " " is above Normal

1-tail test @ 0.1% significance

METHOD 1 p-value

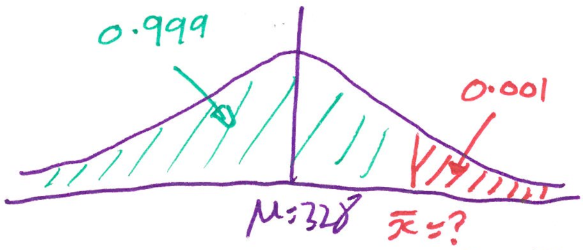


P-value $= P(\bar{X} > 331) = 7.39 \times 10^{-5}$
$= 0.0000739$

Since $p < 0.001$, the null hypothesis is rejected.

There is some evidence to suggest that CO_2 levels in Avonford are above normal.

Method 2 Find critical region



$= 330.44$

$\therefore$ Critical Region $\bar{X} > 330.44$

The sample mean $\bar{x}$ lies within this critical region $\therefore$ the null hypothesis is rejected

Whenever concluding the outcome of your test make sure it is given in terms of the context of the problem. First state whether H_0 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.

<u>Eg2</u> 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. ← hypothesis informed by this sample, so new sample tested

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.

Let X be the random variable "test score"

where $X \sim N(100, 15^2)$ ⇒ Parent Population

and $\bar{X} \sim N\left(100, \frac{15^2}{500}\right)$ ⇒ Sample Mean Population, $\sigma = \frac{15}{\sqrt{500}}$

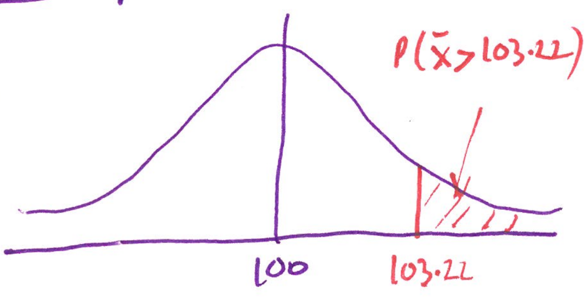
and $\bar{x} = 103.22$

<u>Null Hypothesis</u> $H_0: \mu = 100$ people are no more intelligent than 50 yrs ago

<u>Alt Hypothesis</u> $H_1: \mu > 100$ people are becoming more intelligent than 50 yrs ago

1-tail test @ 1% significance level.

<u>Method 1</u> p-value



p-value $= P(\bar{x} > 103.22) = 7.9296 \times 10^{-7}$

$= 0.00000079$

So $p < 0.01$, so reject H_0

<u>Method 2</u> critical region

∴ critical region $\bar{X} > 101.6$

$\bar{x} = 103.22$, within critical region

∴ reject H_0

There is sufficient evidence to suggest that IQ scores have increased on this test.

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

(i) the probability that an individual paperclip, chosen at random, has a mass greater than 4.04g

(ii) the standard error of the mass for random samples of 25 paperclips

(iii) 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

(iv) 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.

Let M be the random variable "the mass of a randomly selected paperclip"

$M \sim N(4.00, 0.08^2) \Rightarrow$ parent population

(i) $P(M > 4.04) = 0.309$

(ii) $S.E = \frac{\sigma}{\sqrt{n}} = \frac{0.08}{\sqrt{25}} = 0.016$

(iii) Sample Population of Means $\bar{M} \sim N\left(4.00, \frac{0.08^2}{25}\right)$, $\sigma = \frac{0.08}{\sqrt{25}} = 0.016$

$P(\bar{M} > 4.04) = 6.21 \times 10^{-3}$

$= 0.0062$

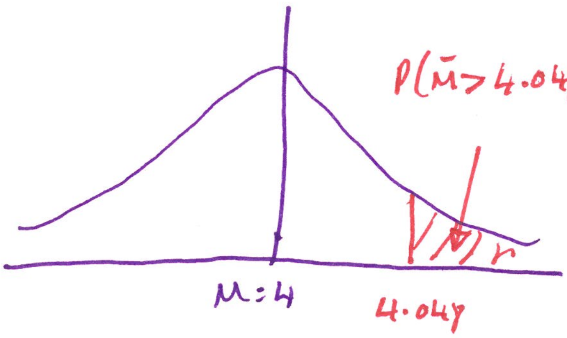
(iv) Calculated sample mean $\bar{m} = 101.2 \div 25 = 4.048g$

$H_0 : \mu = 4.00$ no increase in mean mass of paperclips

$H_1 : \mu > 4.00$ the mean mass has increased

1-tail test @ 5% significance

Method 1 p-value

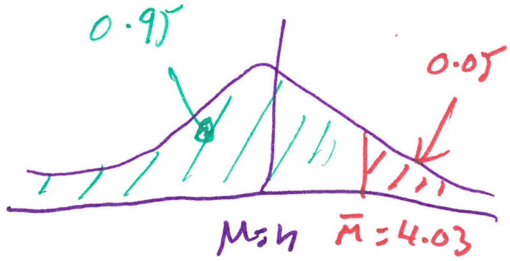


p-value $= P(\bar{M} > 4.048) = 1.35 \times 10^{-3}$

$= 0.00135$

So $p < 0.05$, hence H_0 rejected

Method 2



Critical Region $\bar{M} > 4.03$

Sample mean, 4.048, lies within Critical region ∴ H_0 rejected

Hence there is sufficient evidence to suggest that the mean mass has increased

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.

let W be the random variable "the weight of a randomly selected bag"

parent population $W \sim N(50, 0.96^2)$

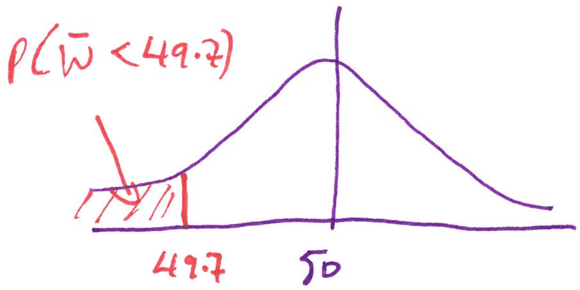
Sample population of means $\bar{W} \sim N\left(50, \frac{0.96^2}{36}\right)$, $\sigma = \frac{0.96}{\sqrt{36}}$

Calculated sample mean: $1789.20 \div 36 = 49.7$ $(\bar{w})$

$H_0: \mu = 50$ no fall in the mean weight of bags.

$H_1: \mu < 50$ mean weight has been reduced.

1-tail test, no significance level given, so can't find critical region



p-Value $p = P(\bar{W} < 49.7) = 0.03$

@ 1% sig level $p > 0.01$ ∴ H_0 accepted

@ 5% " " $p < 0.05$ ∴ H_0 rejected

So $0.01 < p < 0.05$, there is some evidence to suggest the null hypothesis can be rejected. Possibly require an additional sample.

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

(i) Write down, in precise form, the zoologist's null and alternative hypotheses, and state whether a 1-tail or 2-tail test is appropriate.

(ii) Carry out the test at the 5% significance level and write down your conclusion.

(iii) Would your conclusion have been the same at the 10% significance level?

Let M be the random variable "the mass of a randomly selected lizard"

Parent population $M \sim N(72.7, 4.8^2)$

Sample mean population $\bar{M} \sim N\left(72.7, \frac{4.8^2}{12}\right)$, $\sigma = \frac{4.8}{\sqrt{12}}$

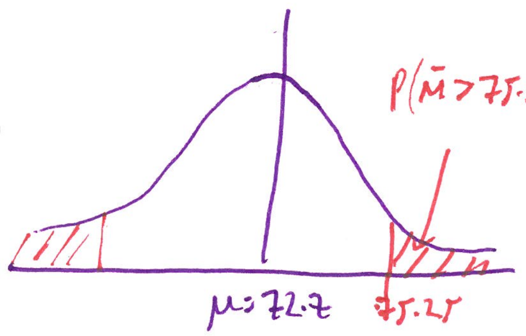
Calculated mean $\bar{M} = 75.25$

(i) $H_0: \mu = 72.7$ lizards are the same type.

$H_1: \mu \neq 72.7$ lizards are different type.

2-tail test @ 5% significance level, ie 2.5% each tail.

(ii) Method 1 p-value



p-value, $p = P(\bar{M} > 75.25)$

$= 0.033$

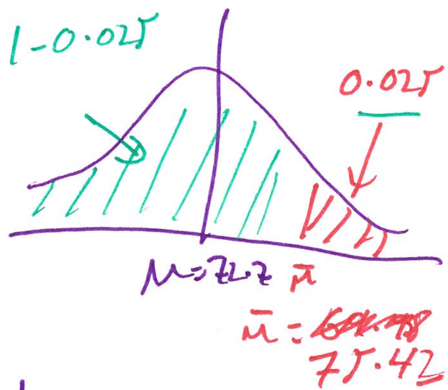
Sig level @ upper tail $= 0.025$

hence $p > 0.025$

∴ not enough evidence to ~~rej~~ reject H_0.

Which suggest the found lizards are the same as the known lizards.

Method 2



So critical region $\bar{M} > 75.42$

Calculated sample mean < 75.42

∴ outside critical region

reject H_1

(iii) @ 10% sig level, 5% each tail

p-value < 0.05

∴ H_0 would have been rejected.

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

(i) 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.

(ii) Would the removal if this outlier alter the conclusion?

(i) let S be the random variable "the speed of a randomly selected green car"

parent population $S \sim N(80, 10^2)$

Sample population of means $\bar{S} \sim N\left(80, \frac{10^2}{20}\right)$ $\sigma = \frac{10}{\sqrt{20}}$

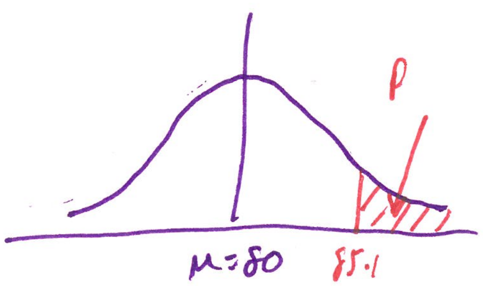
Calculated sample mean $\bar{s} = 85.1$

$H_0: \mu = 80$ no change in motorists mean speeds

$H_1: \mu \neq 80$ motorists mean speed has changed

2-tail test @ 5% so 2.5% each tail.

Method 1 p-value

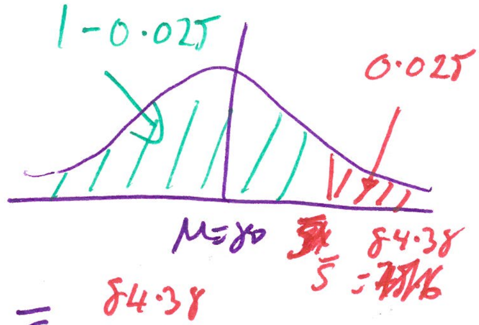


p-value, $p = P(\bar{S} > 85.1) = 0.0113$

$p < 0.025$, so reject H_0

there is sufficient evidence to suggest a change in motorists speeds.

Method 2



Critical Region $\bar{S} >$ 84.38 mph

Calculated mean = 85.1, within crit region

∴ reject H_0.

Remove 124mph $\bar{s} = 83.05$

$\bar{S} \sim N\left(80, \frac{10^2}{19}\right)$ $\sigma = \frac{10}{\sqrt{19}}$

Critical region $\bar{S} > 84.50$ mph

$\bar{s}$ outside critical region hence insufficient evidence to reject H_0

Exercise 3.1

1. For each of the following, a random sample of size n is taken from a Normal population with known standard deviation. The sample mean $\bar{x}$ is calculated.
In each case find the p-value and carry out the hypothesis tests, given H_0 and H_1 at the significance level indicated.

	σ	n	$\bar{x}$	H_0	H_1	Sig. level
(i)	14	20	247	$\mu = 240$	$\mu \neq 240$	2%
(ii)	15	10	172.9	$\mu = 165$	$\mu > 165$	5%
(iii)	24	15	226	$\mu = 240$	$\mu < 240$	1%

2. The masses of adult students are known to be Normally distributed with mean 67.4 kg and standard deviation 3.8 kg. A sample of size 24 is taken and the mean found to be 65.8 kg. Assuming that the standard deviation is unchanged, test, at the 1% significance level, whether the mean mass of adult students has decreased, giving the critical region for the test.

3. Ball bearings produced by a machine should have diameters of 15 mm. A random sample of the diameters of 80 ball bearings gave a sample mean of 14.64 mm and an unbiased estimate for the standard deviation of 1.41 mm. Test, at the 2% significance level, whether the mean diameter of the ball bearings has changed, giving the critical region for the test. Assume that the diameters of the ball bearings are Normally distributed.

4. Packets of cereal are claimed to have a weight of 500 g. The standard deviation is known to be 5 g. An inspector wants to find out if the packets are underweight. She weighs 15 packets of cereal and finds that they have a total weight of 7458 g. Carry out a hypothesis test at the 2% level.

5. A machine is set to produce rods of length L cm, with the distribution $L \sim N(10.6, 0.8)$.
A quality control check on 50 rods showed a mean length of 10.79 cm. Carry out a test at the 10% significance level to find out whether the machine should be recalibrated.

6. The lengths of metal rods produced by a particular machine have been Normally distributed with mean 65 cm and standard deviation 6 cm. A statistical test is conducted on a randomly chosen sample of 25 rods. What would be the critical regions in testing
(i) whether the rods are now shorter, using the 5% significance level?
(ii) whether the lengths of the rods have altered, using the 2% significance level?

7. A particular type of light bulb is supposed to last on average for 1050 hours. The manufacturer claims that he has improved the quality of the light bulbs so that on average they now last for more than 1050 hours. In order to test this claim, at the 5% significance level, the retailer takes a random sample of 50 bulbs, which gave a sample mean lifetime of 1065 hours and an unbiased estimate for the standard deviation of 69 hours. Assume that the lifetimes of the bulbs are Normally distributed.
   (i) State the null and alternative hypotheses.
   (ii) Calculate the *p*-value.
   (iii) State whether the manufacturer's claim is justified.

8. It has been found from experience that a particular type of thread has breaking strengths that are Normally distributed with mean 11.9 N and variance 4.3 N^2. In a random sample of 40 threads, taken from a large batch of recently produced threads, the mean breaking strength was found to be 11.2 N. Test, at the 5% significance level, whether there has been a change in the breaking strength of this type of thread.

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

Ex 3.1

WJEC Hypothesis Testing – Ex Level

1 (i) $H_0: \mu = 240$

$H_1: \mu \neq 240$ 2-tail test

$\bar{X} \sim N\left(240, \frac{14^2}{20}\right)$

where μ is the population mean. For 2-tail test @ 2% sig ~~each tail 0.01~~

$P(\bar{X} > 247) = 0.0127$

2-tail test @ 2% sig level

p-value $= 2 \times 0.0127 = 0.0253 > 0.02$, so accept H_0

ie there is insufficient evidence to suggest that the population mean is different from 240

(ii) $H_0: \mu = 165$

$H_1: \mu > 165$

$\bar{X} \sim N\left(165, \frac{15^2}{10}\right)$

where μ is the population mean.

$p = P(\bar{X} > 172.9) = 0.0479 \Rightarrow$ P-value

1 tail test @ 5% significance

$p < 0.05$, so reject H_1

The evidence suggests that the population mean is greater than 165

(1)(iii) $H_0: \mu = 240$

$H_1: \mu < 240$

$\bar{X} \sim N\left(240, \frac{24^2}{15}\right)$

where μ is the population mean.

$p = P(\bar{X} < 226) = 0.01193$

1 tail test @ 1% sig level

$p > 0.01$, so accept H_0

ie there is insufficient evidence to suggest that the population mean is less than 240

(2) $H_0: \mu = 67.4$

$H_1: \mu < 67.4$

$\bar{X} \sim N\left(67.4, \frac{3.8^2}{24}\right)$

where μ is the mean mass of adults in the population

$p = P(\bar{X} < 65.8) = 0.0196$

1 tail test @ 1% sig level

$p > 0.01$, so accept H_0

ie there is insufficient evidence to suggest that the mean mass of adult students has fallen below 67.4 kg

OR $P(\bar{x} < n) = 0.01$

$n = 65.60$, so critical value $= 65.60$ kg

critical region $\bar{x} < 65.60$

So $\bar{x} = 67.4$ does not lie within critical region $\therefore$ there is insufficient evidence

(3) $H_0: \mu = 15$

$H_1: \mu \neq 15$ 2-tailed test @ 2% sig

$\bar{X} \sim N\left(15, \frac{1.41^2}{80}\right)$

where μ is the mean diameter of population of ball bearings.

$P(\bar{X} \leq n) = 0.01$

$n = 14.633$ ~~so critical value = 14.633, critical region $\bar{X} \leq 14.633$~~

~~if $\bar{x} = 14.64$ this is within critical region $\therefore$ reject H_0.~~

~~ie there is sufficient evidence to suggest that the mean diameter of ball bearings has changed.~~

$P(\bar{X} > n) = 0.99$

$n = 15.37$

So critical region is $\bar{x} < 14.633$ and $\bar{X} > 15.37$

$\bar{x} = 14.64$ does not lie in this region so there is not sufficient evidence to suggest that the mean diameter has changed

or $p(\bar{X} \leq 14.64) = 0.0112$

2-tailed so p-value $= 2 \times 0.0112 = 0.0224$

$p > 0.02$ so reject H_0.

Ex 3.1

Ex Level 2

(4) $H_0 : \mu = 500$

$H_1 : \mu < 500$ @ 2% sig level

Sample mean $\bar{x} = \frac{7458}{15} = 497.2g$

μ is the mean weight of the population

$\bar{X} = N\left(500, \frac{5^2}{15}\right)$

$P(\bar{X} < ?) = 0.02$

critical value $n = 497.35$ so critical region $\bar{x} < 497.35$

Sample mean lies within critical region $\therefore$ reject H_0

ie there is sufficient evidence to suggest that the mean weight of the population is less than 500g

or p-value $P(\bar{X} < 497.2) = 0.015$

@ 2% significance, $p < 0.02 \therefore \ldots$

(5) $H_0 : \mu = 10.6$

$H_1 : \mu \neq 10.6$ 2-tailed test @ 10% sig

where μ is the mean rod length of the population

$\bar{L} \sim N\left(10.6, \frac{\sqrt{0.8}}{50}^2\right)$ $\frac{\sigma}{\sqrt{n}} = \frac{\sqrt{0.8}}{\sqrt{50}}$

p-value $P(\bar{L} > 10.79) = 0.0665$ $\times 2 = 0.1331$

Now $p > 0.1$, so accept H_0

There is insufficient evidence to suggest the mean length has changed so no need to recalibrate the machine.

(3) $L \sim N(65, 6^2)$

$H_0 : \mu = 65$

$H_1 : \mu < 65$ @ 5% sig

where μ is the population mean length.

$\bar{L} \sim N\left(65, \frac{6^2}{25}\right)$

(i) ~~first ...~~ $P(\bar{L} < n) = 0.05$

critical value, $n = 63.03$ cm   critical region $\bar{L} < 63.03$ cm

(ii) $H_0 : \mu = 65$

$H_1 : \mu \neq 65$ 2-tail test @ 2%

$P(\bar{L} < n) = 0.01$ $P(\bar{L} < n) = 0.99$

$n = 62.2$ $n = 67.79$

So critical region $\bar{L} < 62.2$ and $\bar{L} > 67.79$

(4) (i) $H_0 : \mu = 1050$

$H_1 : \mu > 1050$ where μ is the population mean lifetime

$\bar{L} \sim N\left(1050, \frac{69^2}{50}\right)$

Sample Mean $\bar{L} = 1065$

(ii) p value $P(\bar{L} > 1065) = 0.0621$

testing @ 5% sig level, p-value > 0.05 $\therefore$ accept H_0

So there is not sufficient evidence to justify the manufacturers claim.

(8) $X \sim N(11.9, \sqrt{4.3}^2)$

$H_0 : \mu = 11.9$

$H_1 : \mu \neq 11.9$ 2-tail @ 5% sig level

where μ is the mean breaking strength of the population.

$\bar{X} \sim N\left(11.9, \frac{\sqrt{4.3}}{\sqrt{40}}^2\right)$

$P(\bar{X} \leq 11.2) = 0.0164$

p-value $= 2 \times$ $= 0.033$

p-value < 0.05 $\therefore$ reject H_0

ie there is sufficient evidence to suggest that there has been a change in the breaking strength.