Year 12 Mathematics Specialist 2018
Test Number 6:
Statistical Inference
Resource Rich

Name: SOLUTIONS　　　　　　　　　　Teacher: Mrs Da Cruz

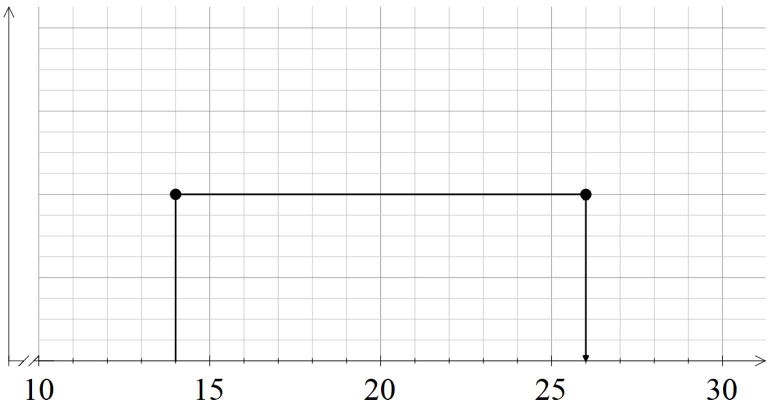
Marks: 44

Time Allowed: 45 minutes

Instructions: You are permitted 1 A4 page of notes and your calculator. Show your working where appropriate remembering you must show working for questions worth more than 2 marks.

Question 1 **[2, 3 = 5 marks]**

Simone takes anywhere from 14 to 26 minutes to travel from home to work each day dependent upon road conditions.
The continuous random variable, X, for her travel times is uniformly distributed: $X \sim U(14, 26)$.



a) If Simone needs to be at work by 8:30am and she leaves home at 8:12 am, what is the probability that she will not be late.

Simone will be late if her journey is longer than 18 minutes.

$$P(X \le 18) = \frac{4}{12} = \frac{1}{3}$$

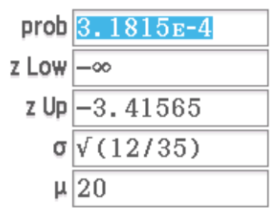
b) Over the next twelve months Simone's travel times for a random sample of 35 work days are recorded. What is the probability that the average of those travel times is less than 18 minutes?

Given: n=35

$$\mu = 20,\ \sigma = \sqrt{\frac{12^2}{12}} = \sqrt{12}$$

$$\bar{X} \sim N\left(20, \sqrt{\frac{12}{35}}^{\,2}\right)$$

$$P(\bar{X} < 18) = 0.00032$$



Question 2 **[2, 4 = 6 marks]**

The confidence level is $\bar{x} \pm z\frac{\sigma}{\sqrt{n}}$. The margin of error being $z\frac{\sigma}{\sqrt{n}}$. A small margin of error is usually preferred. A sample of 2000 male scores gave a mean of 150 and a standard deviation of 40.

a) Give the 95% confidence interval for the population mean, μ.

Given: n=2000, $\bar{x} = 150$, s=40

Lower	148.24695
Upper	151.75305
$\bar{x}$	150
n	2000

C-Level	.95
σ	40
$\bar{x}$	150
n	2000

$148.25 \leq \mu \leq 151.75$

b) How would altering the sample size affect the margin of error? Give examples to demonstrate any statements you make.

For $n = 2000$ the MOE is 1.753

If n increases, the MOE decreases.
(Note: This is because larger samples more closely reflect the population.)

Example if n =3000, the MOE decreases to 1.431 (or similar example)

If n decreases, the MOE increases.
(Note: This is because smaller samples less closely reflect the population.)

Example if n =1000, the MOE increases to 2.479 (or similar example)

Question 3 **[3 marks]**

A statistician working for a telemarketing company informs the bosses that data collected from the population has a standard deviation of 19.95. He assured them that with a confidence interval of 99% and to be within 5 units of the mean, the required sample would have to be 106. Was he correct? Just justify your answer.

$$d = z\frac{\sigma}{\sqrt{n}}$$

$$5 = 2.576\times\frac{19.95}{\sqrt{n}}$$

$$n = 105.64\ldots$$

Yes he is correct.

Question 4 **[2, 1, 3 = 6 marks]**

Dhufish are caught off the coast of North-West Australia. Their weights are normally distributed with a mean of 20 kg and a standard deviation of 2.7 kg.

a) Matt selects a sample of 32 dhufish caught off the North-West coast of Australia and their mean weight is 22.2 kg.

i) Using Matt's sample, determine a 99% confidence interval for the mean weight of dhufish.

$$22.2 - 2.576\times\frac{2.7}{\sqrt{32}} \le \mu \le 22.2 + 2.576\times\frac{2.7}{\sqrt{32}}$$

$$20.97 \le \mu \le 23.43$$

ii) You should observe that the confidence interval from i) does not contain μ. We would expect 99% of such formed intervals to contain μ. Therefore, it is statistically highly improbably that this would happen. In this particular scenario what is the likely reason?

Students are expected to notice that Matt selects the fish and that it doesn't state that the sample is random.

Matt's sample most probably was **<u>not random</u>**. For example - He may have deliberately selected larger fish.

(One alternate answer to be accepted relates to seasonal fish sizes: For example: Matt's sample was chosen at a time of the year when, due to breeding patterns, fish sizes are smaller.
Answers not accepted: mentioning Matt catching the fish as this indicates a misinterpretation of the scenario.)

b) A random selection of 10 dhufish are sent to a restaurant in Perth. Determine the probability that the total weight of the 10 dhufish lies between 175 kg and 202 kg.

$$\bar{X} \sim N\left(20, \left(\frac{2.7}{\sqrt{10}}\right)^2\right)$$

$$P(17.5 < \bar{X} < 20.2) = 0.591$$

Question 5 **[3, 1, 3 = 7 marks]**

Crunchy Crisps are sold in packets labelled as 500 g. It is known that the weight of crisps in each packet is normally distributed with mean 499 g and standard deviation 2 g.

a) If multiple samples of 20 packets of Crunchy Crisps were taken,

i) Describe the sampling distribution of the sample means.

Normal with mean 499 and standard deviation $\frac{2}{\sqrt{20}}$

ii) Calculate the probability of one of these sample having a sample mean of at least 500 g.

$$P(\bar{X} \geq 500) = 0.0127$$

b) Determine the size n of a random sample if the standard deviation of the sampling distribution is to be less than 0.25 g.

$$\sigma_{\bar{X}} < 0.25g$$

$$\frac{\sigma}{\sqrt{n}} < 0.25$$

$$\frac{2}{0.25} < \sqrt{n}$$

$$8 < \sqrt{n}$$

$$64 < n$$

Sample size must be at least 65.

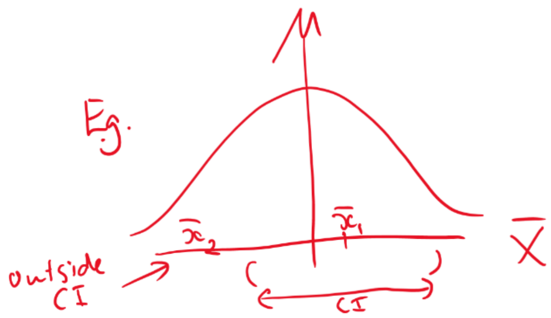
Question 6 **[2, 2, 2, 3 = 9 marks]**

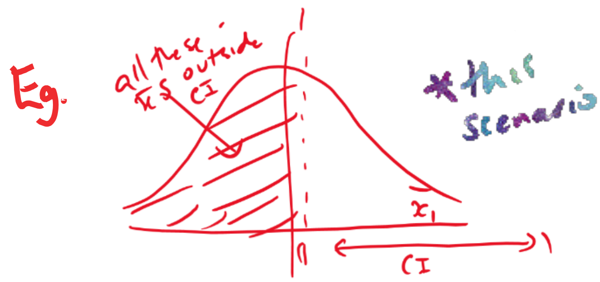
A cable in a bridge is required to support a weight of 10 000 Newtons. Tina tests a random sample of 100 cables from a supplier. The sample mean is found to be 10 300 Newtons and the sample standard deviation 400 Newtons.

Based on Tina's sample, a 95% confidence interval for μ, the population mean cable strength is calculated. State whether each of the following statements is true or false. Provide reasons for your answer and state any assumptions.

(i) If another sample of 100 cables is taken, then the sample mean will fall within the confidence interval produced by Tina's sample.

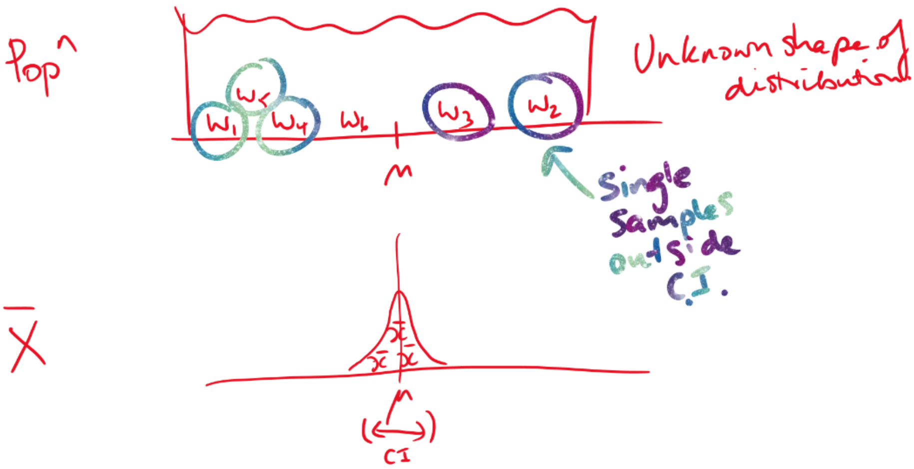
False.
The confidence interval formed only gives an interval within which we have 95% confidence the population mean lies. There is no certainty that the other sample's mean will lie in this bound.





(ii) If a single cable is selected at random, then the strength of the cable will fall within the confidence interval produced by Tina's sample.

False.
Only the population average is likely to be in the interval. A single sample could be from anywhere on the population distribution and not necessarily close to the population average.



Jon, a colleague of Tina, said, ‘The cable strengths are not normally distributed, so the calculation for the confidence interval is incorrect’.

(c) How should Tina respond to Jon’s comment?

The calculation is correct.

The sample size is sufficiently large (>30) such that the distribution of the population is irrelevant for the central limit theorem to be valid.

A different sample of 36 cables is taken and it is found that the standard deviation is 500 Newtons. A confidence interval for the population mean cable strength is determined to be $9900 \leq \mu \leq 10\,200$.

(d) Determine the confidence level, to the nearest 0.1%, used to calculate this interval.

The Margin of Error is 150. (Since the width of the interval is 300.)

$$150 = z \times \frac{500}{\sqrt{36}}$$

$$z = 1.8$$

$$P = 0.928 \;\; or\; 92.8\% \; confidence\ level.$$

prob	0.9281394
z Low	−1.8
z Up	1.8
σ	1
μ	0

Question 7 **[1, 3, 2 = 6 marks]**

The time taken for a Year 8 extension student to complete a particular maths puzzle is normally distributed with mean 3 minutes and standard deviation 20 seconds.
A sample of fifty Year 8 extension students each completed the puzzle.

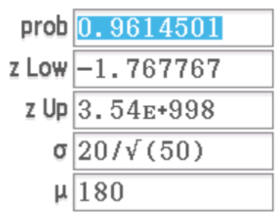
a) Describe how you would expect the 50 sample times to be distributed.

Normal with mean of 3 minutes and standard deviation of 20 seconds.
(The same as the population distribution.)

b) Estimate the probability that the mean time for the sample is greater than 2 minutes and 55 seconds.

$$\bar{X} \sim N\left(180, \left(\frac{20}{\sqrt{50}}\right)^2\right)$$

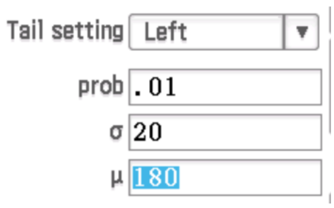
$$P(\bar{X} \geq 175) = 0.96$$

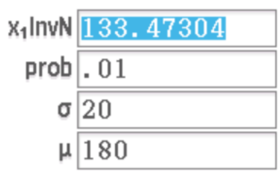


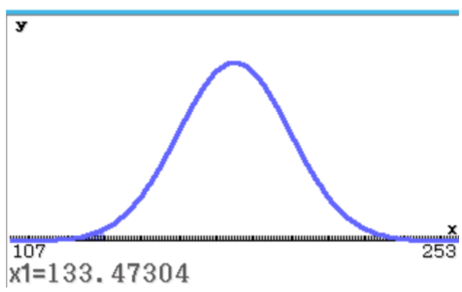
c) Students who complete the puzzle in under k seconds are classified "highly gifted". If 1% of all extension students are classified highly gifted, find k.

$P(X < k) = 0.01$ (from population distribution)

$k \approx 133.5$







Question 8 **[2 marks]**

What will be produced by classpad upon the entry randNorm(2.5, 55, 200)?

200 randomly generated numbers ✓

from the normal distribution with mean 55 and standard deviation 2.5 ✓