Hale School

Mathematics Specialist

Term 3 2018

Test 6 - Statistical Inference

Name: Sam Plemean

/36

Instructions:

- **TIME ALLOWED: 40 Minutes**
- **CAS calculators are allowed**
- **External notes are not allowed**
- **Show your working clearly**
- **Use the method specified (if any) in the question to show your working (Otherwise, no marks awarded)**
- **This test contributes to 6% of the year (school) mark**

Question 1 **(8 marks)**

A sample of 20 values is taken from a population that is distributed normally with a mean value of 28 and a variance of 16.

(a) Calculate the mean and standard deviation of the distribution of the sample means. (2 marks)

Mean = 28 ✓ mean

$\sigma_{\bar{x}} = \frac{\sqrt{16}}{\sqrt{20}} \doteq \frac{2}{\sqrt{5}} = 0.894$ (3dp) ✓ s.d.

(b) Describe the sampling distribution of the sample means. (2 marks)

The sampling distribution will be normal with a mean of 28 and standard deviation 0.894

✓ normal

✓ parameters

(c) Find the probability that the mean of the sample lies between 27 and 30. (2 marks)

$p(27 \le \bar{x} \le 30) = 0.8556$ (4dp) ✓ 0.8556

normCDf$(27, 30, 4/\sqrt{20}, 28)$ ✓ correct statement and notation

(d) If 4 samples of size 20 are taken, find the probability that all of the samples have a mean between 27 and 30. (2 marks)

$0.8556^4 = \underline{\underline{0.5358}}$ (4dp) ✓ concept

✓ answer

Question 2 **(4 marks)**

A factory puts baked beans into cans which are labelled as containing 300g. The company wants to know about the distribution of the weight of the baked beans in their cans. They weigh the contents of 40 random samples of cans. Each sample contains 36 cans and the distribution of weights was found to be as in the table below.

Mean Weight	Number of samples
$299.5 < m \leq 300$	2
$300 < m \leq 300.5$	5
$300.5 < m \leq 301$	8
$301 < m \leq 301.5$	12
$301.5 < m \leq 302$	7
$302 < m \leq 302.5$	4
$302.5 < m \leq 303$	2

Use this information to,

(a) estimate the mean and standard deviation of the sampling distribution of sample means. (2 marks)

Mean = 301.21 g ✓ mean

$s_{\bar{x}}$ = 0.7458 (or 0.7364) ✓ s.d.

(b) Estimate the population mean and standard deviation. (2 marks)

Mean = 301.21 g ✓ mean

Std dev = $0.7458 \times \sqrt{36}$

= 4.475 (3dp) (or 4.418) ✓ s.d.

Question 3 **(11 marks)**

Customers who pay for a particular National Broadband Network service are supposed to receive a bandwidth of 40 Mb/s. Callum tests a random sample of 80 customers. The sample mean is found to be 42 Mb/s and the sample standard deviation 10 Mb/s.

(a) Based on Callum's sample, obtain a 95% confidence interval for μ, the population mean bandwidth, giving your answers to 1 decimal place. (2 marks)

95% C.I. is $\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}}$ ✓ uses 1.96

$42 \pm 1.96 \times \frac{10}{\sqrt{80}}$

$39.8 \leq \mu \leq 44.2$ ✓ correct statement

(b) State whether each of the following situations is certain, likely or unlikely to occur. Provide reasons for your answer.

i) If another sample of 80 customers is taken, then the sample mean will fall within the confidence interval found at part (a). (2 marks)

This is quite likely. Since we are 95% confident that the mean lies in this range, there is a distinct chance that the sample mean will also fall in the range.

✓ likely

✓ reason

ii) If a single customer is selected at random, then the bandwidth will fall within the confidence interval found at part (a). (2 marks)

This is possible but less likely as the original population values are more spread out. A spread of 4.4 is less than half of 1 std deviation.

✓ unlikely

✓ reason

Tayla, a colleague of Callum, said, 'The bandwidths are not normally distributed, so the calculation for the confidence interval is incorrect'.

(c) How should Callum respond to Tayla's comment? (2 marks)

As the sample size is more than 30, it does not matter if the original population is distributed normally.

✓ $n \geq 30$

✓ clear statement

A different sample of 36 customers is taken and it is found that the standard deviation is 8 Mb/s. A confidence interval for the population mean bandwidth is determined to be $40.58 \leq \mu \leq 45.42$.

(d) Determine the confidence level, to the nearest 0.1%, used to calculate this interval. (3 marks)

Mean $= 43$

and $45.42 - 43 = 2.42$

$\therefore \quad z \times \frac{8}{\sqrt{36}} = 2.42$ ✓ equation

$z = 1.815$ ✓ z value

$P(-1.815 \leq z \leq 1.815) = 0.93047$

$\therefore$ 93% confidence interval used ✓ conclusion

Question 4 **(7 marks)**

A golf academy ran a week long programme for 40 people and measured the improvement in players' golf scores (lower scores indicate better performance). They calculated a confidence interval based on a 90% level of confidence and found it to be 10.9 to 12.1.

(a) What was the average score improvement? (1 mark)

Average score improvement = 11.5

✓ 11.5

(b) What was the standard deviation of the score improvements for the group of 40 people? (2 marks)

$$0.6 = \frac{1.6449 \times \sigma}{\sqrt{40}} \Rightarrow \sigma = \frac{0.6\sqrt{40}}{1.6449}$$

$$\sigma = 2.31$$

✓ $\frac{1.6449\sigma}{\sqrt{n}}$

✓ answer

There were 10 additional players who started the course but did not finish. If these players were credited with having 0 improvement and included in the calculations,

i) determine the new mean for the group (2 marks)

$$\Sigma x = 11.5 \times 40 = 460$$

$$\text{New mean} = \frac{460}{50} = 9.2$$

✓ equation

✓ answer

ii) state with reasons whether the standard deviation for the group will increase or decrease (2 marks)

Original population distributed with mean 11.5 and std dev of 2.31. New 0 values are well below the mean so will increase the spread ∴ std. dev increases

✓ increase

✓ reason

or

$$2.31^2 = \frac{\Sigma x^2}{40} - 11.5^2 \Rightarrow \Sigma x^2 = 5503.444$$

✓ finds Σx^2

$$\therefore \text{New } \sigma = \sqrt{\frac{5503.444}{50} - 9.2^2} = 5.04$$

(which is higher

✓ higher

Question 5 **(6 marks)**

For a continuous uniform distribution in the range $a \le x \le b$ the mean and standard deviation are given by $\dfrac{a+b}{2}$ and $\dfrac{b-a}{\sqrt{12}}$ respectively.

The crowd entering Optus Stadium has to pass through a bag check procedure. For each individual the length of time taken to pass the bag check is thought to follow a uniform distribution between 0 and 30 seconds.

(a) Find the probability that the total time for 40 people to pass through the bag check is greater than 11 minutes. (3 marks)

$$11 \text{ mins} = 660\text{s} \Rightarrow \bar{t} = \frac{660}{40} = 16.5$$

$\checkmark\ \bar{t} = 16.5$

$$\text{and } \bar{t} \sim N\left(15, \left(\frac{30/\sqrt{12}}{\sqrt{40}}\right)^2\right) = N(15, 1.369^2)$$

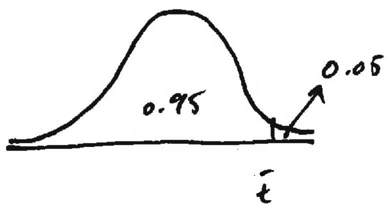
$\checkmark\ \frac{\sigma}{\sqrt{n}} = 1.369$

$$P(\bar{t} \geq 16.5) = 0.137 \ (3\text{dp})$$

$\checkmark$ answer

$$\text{normCDf}\left(16.5, \infty, \left(\frac{30}{\sqrt{12 \times 40}}\right), 15\right)$$

(b) What is the greatest number of people that can be in the queue so there is a probability of at least 0.95 that they all get through the bag check in 5 minutes? (3 marks)



Need 0.05 above $\frac{300}{n}$

$\checkmark\ z = 1.6449$

$$\text{Since } \bar{t} \sim N\left(15, \left(\frac{30}{\sqrt{12n}}\right)^2\right)$$

$\checkmark$ equation

$$\text{Need } \frac{300}{n} = 15 + \frac{1.6449 \times 30}{\sqrt{12n}}$$

$$n = 16.18$$

$\checkmark$ conclusion

$\therefore$ 16 is the most

Or on CAS $\quad P\left(\bar{t} \geq \frac{300}{x}\right) = 0.05$

$$\text{solve}\left(\text{normCDf}\left(\frac{300}{x}, \infty, \frac{30}{\sqrt{12x}}, 15\right) = 0.05, x\right)$$

$x = 16.18 \quad \therefore \quad 16$ is the most

END OF TEST