ALL SAINTS' COLLEGE

Year 12 Mathematics Specialist 2019
Test Number 6:
Statistical Inference
Resource Rich

Name: ________________________________ Teacher: Mrs Da Cruz

Marks: 46

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 marks]**

A Company is asked to make bolts with length accurate to within ± 0.002 with a 95% confidence level. If $\sigma = 0.004$, determine how many bolt lengths must be averaged to satisfy the request.

From Formula Sheet: $n = \left(\frac{z \times s}{d}\right)^2$

$z = 1.960, \quad n = \left(\frac{1.960 \times 0.004}{0.002}\right)^2 = 15.3664$

$\therefore$ *Require 16 bolt lengths to be averaged.*

Question 2 **[3 marks]**

A valid 90% confidence interval, for a sample of 10 observations based on a population standard deviation of 0.05 with a $\bar{x} = 0.927$, is 0.901 to 0.953.

a) What must be true about the population?

Since n<30, the population must have been normally distributed.

b) Describe the effect on the width of the confidence <u>interval</u> when the confidence <u>level</u> is increased. Justify your answer with some calculations.

As the level of confidence increases the confidence interval **<u>increases in width</u>**.

95% CI: 0.896 – 0.958

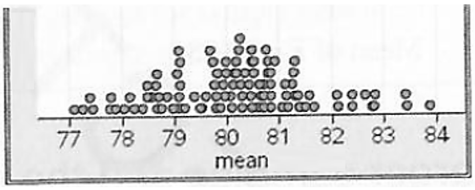
98% CI: 0.890 – 0.964

99% CI: 0.886 – 0.968

At least one new CI shown.

Question 3 **[8 marks]**

Using simulation, 100 random samples, each of size 49, is generated and the sample means are shown in the dot plot.



a) Indicate whether the following statements are true or false.
b) For any of the statements which are false, clearly explain why they are false.

i. As the dot plot appears to show that the sampling distribution of the sample means is normally distributed, the population must have been normal. True / False

Since n>30, the population distribution did not have to be normal.

ii. The dot plot appears to show that the sampling distribution of sample means is normal with a mean of approximately 80 and standard deviation of approximately 1. True / False

iii. The information in the dot plot implies that the population mean was approximately 80 with a standard deviation of approximately 1. True / False

The population mean would be approximately 80 but the standard deviation of the population will always be larger than the standard deviation of the sampling distribution.

iv. The information in the dot plot implies that the population mean was approximately 80 with a standard deviation of approximately 7. True / False

v. The information in the dot plot shows a sample which is normally distributed, mimicking the population distribution. True / False

The information in the dot plots shows the sampling distribution of sample means, not the distribution of a sample. If it was 'a sample' it would mimic the population.

Question 4 **[12 marks]**

The mass of crayfish caught near the Abrolhos Islands is observed to be normally distributed with a mean of $\mu = 1{\cdot}2$ kg and standard deviation of $\sigma = 0{\cdot}25$ kg.

Joe the fisherman catches 65 crayfish.

(a) Determine the probability that:

(i) the mean crayfish mass will be less than 1.15 kg. (3 marks)

(i) $\dfrac{\sigma}{\sqrt{65}} = 0{\cdot}031$ ✓

$\therefore$ $P(\bar{x} \le 1{\cdot}15) = 0{\cdot}053$ using $N(1{\cdot}2, 0{\cdot}031^2)$ ✓✓

(ii) the total mass will be between 75 kg and 80 kg. (3 marks)

(ii) Mean of Joe's catch ~$N(1{\cdot}2, 0{\cdot}031^2)$ ✓

$\therefore$ $P\left(\dfrac{75}{65} \le \text{Mean} \le \dfrac{80}{65}\right) = 0{\cdot}771$ ✓✓

On another fishing trip, we are required to be 98% confident that the mean crayfish mass differs from the population mean by less than 0.05 kg.

(b) Find the number of crayfish that need to be caught. (2 marks)

$\dfrac{\sigma z}{\sqrt{n}} < 0{\cdot}05$ with $\sigma = 0{\cdot}25$ and $z = 2{\cdot}326$

$\therefore$ $n > 135{\cdot}3$ ✓

$\therefore$ Need to catch 136 fish. ✓

A rival crayfisherman, Jamie, has started catching crayfish further out to sea than Joe. Jamie states that the crayfish caught are significantly bigger than in the area that Joe fishes in.

Over a month Jamie catches 220 crayfish with total mass of 270 kg. Assume $\sigma = 0.25$ kg.

(c) Determine whether Jamie's claim is supported at the 95% level of confidence. (4 marks)

(c) $\bar{x} = \frac{270}{220} = 1{\cdot}227$ kg and $\sigma = 0{\cdot}25$ and $n = 220$

$z = 1{\cdot}96$

$\bar{x} \pm z\frac{\sigma}{\sqrt{n}} = 1{\cdot}1976 \rightarrow 1{\cdot}2564$ ✓✓

∴ $\mu = 1{\cdot}2$ is within 95% CI so James' claim is not accepted. ✓

∴ Cannot conclude that crayfish in James' area are significantly bigger ✓

Question 5 **[16 marks]**

The volume of water used by the SavaDaWater company to top up an ornamental pool has been observed to be normally distributed with mean $\mu =$ 175 litres and standard deviation $\sigma =$ 15 litres.

The ornamental pool is topped up 50 times. Determine the probability that the:

(a) sample mean volume will be between 173 and 177 litres. (3 marks)

Solution
Let $\overline{W} =$ the sample mean from 50 times the pool is topped up (litres) $= N\left(175,\ \sigma_{\overline{W}}^2\right)$ where $\sigma_{\overline{W}} = \dfrac{15}{\sqrt{50}} = 2.1213....$ Require $P\left(173 < \overline{W} < 177\right) = 0.6542$
Specific behaviours
✓ states that the sample mean is a normal random variable ✓ states the correct parameters for the normal random variable ✓ determines the correct probability

(b) total volume of water used is less than 8.96 kilolitres. (3 marks)

Solution
For a total of 8.96 kL, the sample mean $\overline{W} = \dfrac{8960}{50} = 179.2$ litres Require $P\left(\overline{W} < 179.2\right) = 0.9761$
Specific behaviours
✓ calculates the sample mean correctly for the total 8.96 kL ✓ writes the correct event in terms of the required sample mean ✓ determines the correct probability

Water is a scarce commodity and accuracy is required. The pool is topped up 50 times and the sample mean obtained is denoted by $\overline{W}$.

(c) If it is required that $P\left(a \le \overline{W} \le b\right) = 0.99$, then determine the values of a and b, each correct to 0.1 litres. (3 marks)

Solution
As $\overline{W} = N\left(175,\ \sigma_{\overline{W}}^2\right)$ where $\sigma_{\overline{W}} = \frac{15}{\sqrt{50}} = 2.1213....$
Given $P(-k < z < k) = 0.99$, $k = 2.5758$
Interval : $175 - 2.5758\left(\sigma_{\overline{W}}\right) < \overline{W} < 175 + 2.5758\left(\sigma_{\overline{W}}\right)$
i.e. $169.536 < \overline{W} < 180.464$
i.e. the sample mean 99% confidence interval is 169.5 litres to 180.5 litres
Specific behaviours
✓ uses the correct parameters for the distribution of $\overline{W}$
✓ determines the value for k for the confidence interval
✓ states the interval correct to 0.1 litres

(d) If the probability for the mean amount of water used differs from μ by less than five litres is 96%, find n, the number of waterings that need to be measured. (3 marks)

Solution
$\sigma_{\overline{W}} = \frac{15}{\sqrt{n}}$ Require $P(-k < z < k) = 0.96$ $\therefore\ k = 2.0537$
Hence $2.0537\left(\frac{15}{\sqrt{n}}\right) < 5$
Solving gives $n > 37.96$
i.e. we require at least 38 waterings to have the mean differ by less than five litres
Specific behaviours
✓ determines the standard z score that represents 96 % confidence
✓ forms the correct inequality to solve for n
✓ states the correct minimum integer value for n

A rival company called WolliWorks takes over the watering of the ornamental pool. Over 36 consecutive days, it was observed that the WolliWorks company used a total of 6.57 kilolitres. The standard deviation for the 36 days was also 15 litres.

A representative from the SavaDaWater company states that 'WolliWorks are using significantly more water than we did when we were watering this pool. They are wasting water'.

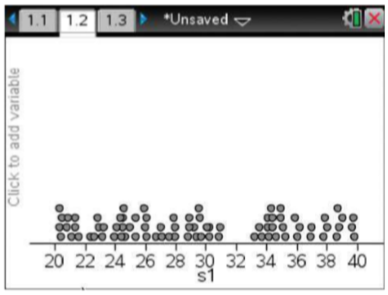
(e) Perform the calculations necessary to comment on this claim. (4 marks)

Solution
Let $\mu_w =$ the population mean for the Waterworks company (litres) For the WolliWorks total of 6 570 litres, this gives $\overline{W} = 182.5$ litres We will estimate μ_w as $N\left(182.5,\ \sigma_{\overline{W}}^2\right)$ where $\sigma_{\overline{W}} = \dfrac{15}{\sqrt{36}} = 2.5$ Confidence Interval for μ_w 95% level $182.5 - 1.96\left(\sigma_{\overline{W}}\right) < \mu_w < 182.5 + 1.96\left(\sigma_{\overline{W}}\right)$ i.e. $177.6 < \mu_w < 187.4$ Confidence Interval for μ_w 99% level $182.5 - 2.58\left(\sigma_{\overline{W}}\right) < \mu_w < 182.5 + 2.58\left(\sigma_{\overline{W}}\right)$ i.e. $176.0 < \mu_w < 189.0$ The SavaDaMoney population mean $\mu =$ 175 is outside the confidence interval using $\overline{W} = 182.5$ and $\sigma = 15$. i.e. the claim is vindicated. i.e. the WolliWorks company IS using significantly more water. Whether they are wasting water cannot be determined from the given data.
Specific behaviours
✓ determines the expected variation using $n = 36$ ✓ determines an appropriate confidence interval for the WolliWorks population mean ✓ states that the SavaDaMoney population mean 175 is outside the confidence interval ✓ concludes correctly by writing a comment about the claim

Question 6 **[5 marks]**

a. Simulate a sample of 80 items from a uniform distribution on the interval [20,40] and sketch the graph of **your** sample. (1 mark)

4 a Similar to:



Scale needs to be visible

b. Simulate 200 sample of 80 items from a uniform distribution on the interval [20,40] and find the mean of the samples. (1 mark)

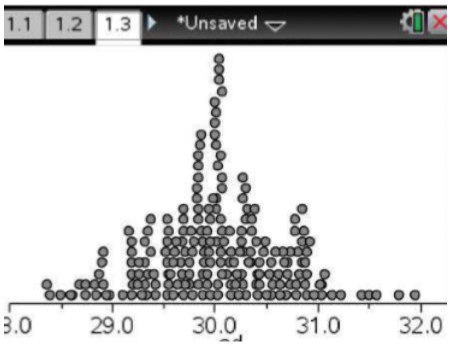
Answer should be **very close to 30** $\bar{X} \sim N\left(30, \left(\frac{5\sqrt{3}}{\sqrt{80}}\right)^2\right) \approx N(30, 0.968^2)$

Only accept answers between 29.6 and 30.4 (which is being very generous)

j

c. Draw a graph of the sampling distribution. (1 mark)

c Similar to:



Scale needs to be visible

d. Comment on the graphs. (2 marks)

- Shapes and location and spread: The single sample is uniformly distributed from 20 to 40, mimicking the population, whilst the sampling distribution is normally distributed from only about 28.3 to 32 with a peak at about 30.

Or

- Location and Spread: They have similar means but the sampling distribution has a significantly smaller standard deviation.

- Shape: The sample is uniformly distributed whilst the sampling distribution is normal.