TRINITY COLLEGE

Year 12 Methods Units 3,4
Test 5 2020

Calculator Assumed
Normal Distribution & Sampling Techniques

STUDENT'S NAME MARKING KEY [KRISZYK]

DATE: Monday 24th August **TIME:** 50 minutes **MARKS:** 49

INSTRUCTIONS:
Standard Items: Pens, pencils, drawing templates, eraser
Special Items: Three calculators, notes on one side of a single A4 page (these notes to be handed in with this assessment)

Questions or parts of questions worth more than 2 marks require working to be shown to receive full marks.

---

1. (2 marks)

Oscar is the quality control expert at the Honey Joys cereal factory where 750 g boxes of Honey Joys are normally distributed with mean weight of 758 grams.

Rafael, the Company Director, wants to reassure customers that boxes of Honey Joys are rarely underweight so instructs Ben, the PR representative, to issue the statement that: "The 0.1 quantile is 752.6 grams and the 0.8 quantile is 761.6 grams".

Ben wants to issue a statement that everyday people will understand. Complete his statement below:

Honey Joys are pleased to reassure you, our valued customers, that our 750 gram boxes of Honey Joys cereal have an average weight of 758 *grams, with at least* 90 *%* *containing* 752.6 *grams or more, and in fact* 20 *% contain at least* 761.6 *grams .*

2. (5 marks)

(a) Jack hosts a podcast and he is curious how much his listeners like his show. He decides to poll the next 100 listeners who send him fan emails.

They don't all respond, but 94 of the 97 listeners who responded said they "loved" his show.

(i) What is the main source of bias in this scenario? [1]

voluntary response.

(ii) Would the bias lead to an underestimate or overestimate for the population parameter? [1]

Overestimate

(b) A high school has a policy that students' phones must be kept away during class. A principal used the school roster to poll a random sample of 50 students, and only 10%, percent said that they ever had their phone out during class. The next day, the principal observed classrooms and noticed that approximately 25%, percent of students had their phone out at some point during class.

Name a major source of bias in the sample. Explain your answer. [2]

response bias (social desiribility)
students do not feel comfortable telling the principal they use their phones in class therefore do not answer truthfully.

(c) A campaign manager is curious what percent of voters will and will not vote in an upcoming election. They carry out a phone poll on a random sample of registered voters that includes the question, "Will you stay home and pass up your opportunity to vote in the upcoming election?" Only 5%, percent of the 250 people reached in the poll say they plan on staying home.

Which of these is the best explanation for why this result is probably biased? [1]

(i) Some people in the sample probably couldn't be reached, so this is likely an overestimate.

(ii) Some people in the sample probably couldn't be reached, so this is likely an underestimate.

(iii) People may not want to admit they plan on staying home rather than voting, so this is likely an overestimate.

(iv) People may not want to admit they plan on staying home rather than voting, so this is likely an underestimate.

3. (4 marks)

Mr Shinkfield wishes to conduct a survey of the year 12 students on the use of the year 12 kitchenette. He requires a sample of 30 students to complete the survey.

(a) Mr Shinkfield creates a survey online and writes the address on a flyer which he leaves in the kitchenette. He then waits for the first 30 responses.

Name a source of bias in the sample providing justification. [2]

- convenience sampling
  Mr S left the flyers in the kitchenette so respondents were self-selected.

- voluntary response bias
  (as above)

Any other with justification

(b) Lucas suggests a more effective way to attain a sample is to use his Classpad to generate 30 random numbers.

Suggest a process he could use to generate a sample of 30 year 12 students and the relevant Classpad inputs required. (Assume there are 120 current year 12 students). [2]

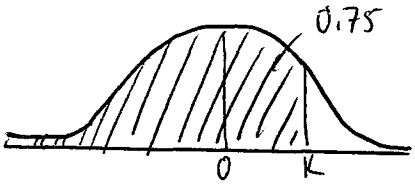
- Number each Yr 12 student #1 → #120 and use CP to generate 30 random numbers between those values.

- randList (30, 1, 120)

4. (7 marks)

For the standard normal distribution Z, determine the following.

(a) The 3rd quartile. [2]



$k = 0.674$

(b) $P(Z = 0.25)$ [1]

0

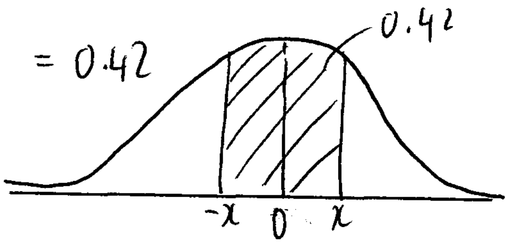
(c) $P(-\sigma < Z < 2\sigma)$ [2]

$= P(-1 < Z < 2)$

$= 0.8186$

(d) x given $P(\mu \pm x) = 0.42$ [2]

$P(-x < \mu < x) = 0.42$



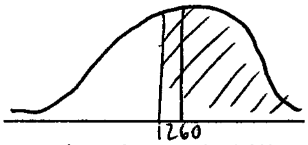
$\therefore P(0 \leq Z \leq x) = 0.21$

$x = 0.553$

5. (13 marks)

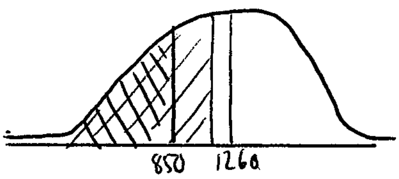
The amount of water used by a single residential WA household each day is normally distributed with a mean of 1260 L and standard deviation of 185 L. The Water Corporation of WA recommends a single residential household uses no more than 1 kL of water per day. Let X represent a WA household that is selected a random:

(a) Determine the probability that on any given day, the amount of water used exceeds 1 kL. [2]



$P(X > 1000) = 0.9200$

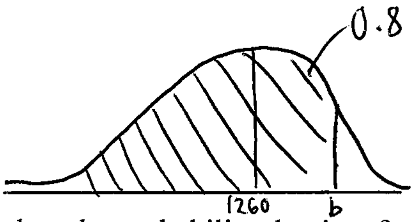
(b) Determine the probability the water usage on any given day is less than 850 L if the household follows the Water Corporation's daily usage recommendations. [3]



$P(X < 850 \mid X < 1000)$

$= \dfrac{0.0133}{0.08} = 0.1663$

(c) The probability that the daily water usage does not exceed b L is 0.8. Calculate the value of b. [2]



$P(X < b) = 0.8$

$b = 1415.7\text{ L}$

(d) Calculate the probability that in a fortnight, the water usage exceeded 1400 L on more than 6 occasions. [3]

$P(X > 1400) = 0.2246$

Y : number of days water usage exceeds 1400L

$Y \sim B(14, 0.2246)$

$P(Y > 6) = 0.0219$

The WA government proposes that water-saving devices should be installed on all taps to reduce the amount of water used.

(e) (i) Determine the mean daily water usage if the probability that the daily water usage exceeds 1200 L is reduced to 0.25 when the water-saving devices are installed. Assume that the standard deviation remains unchanged. [2]

$P = 0.25$

$L = 1200$

$U = \infty$

$S = 185$

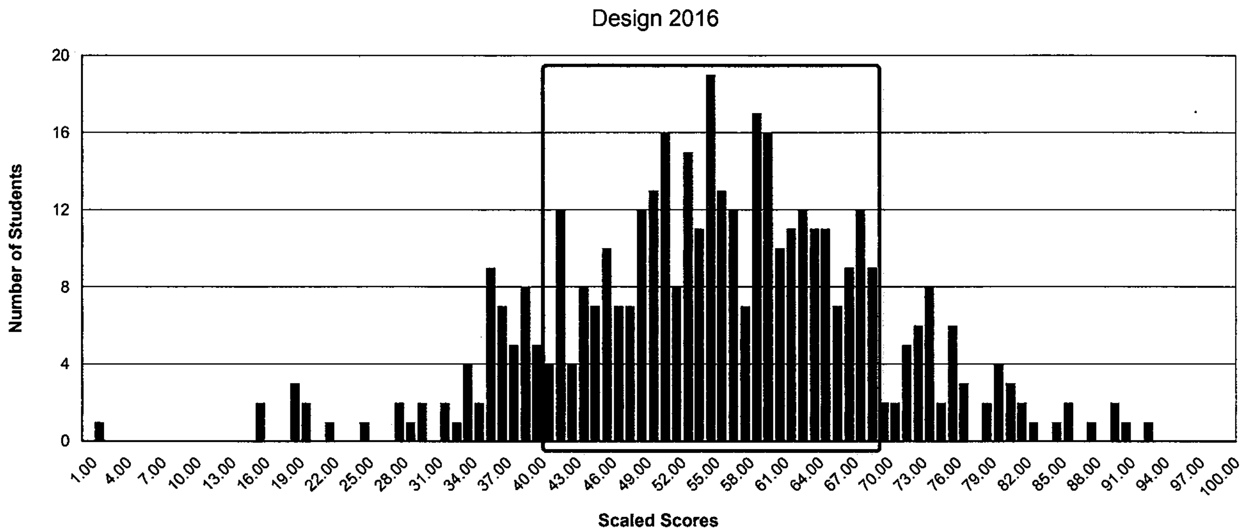
$\mu = 1075.22$

(ii) Comment on your findings. [1]

The water saving devices have been sucessful as μ has dropped almost 200L or one σ

6. (3 marks)

The plot below shows the final scaled scores for the 2016 year 12 WACE students who studied the Design course. The total number of students sitting the final exam was 422. The mean scaled score and standard deviation were 54.9 and 13.7 respectively. The rectangle shows the scaled scores which fit within one standard deviation either side of the mean. There were 310 students who scored within this range.



The use of the Normal Distribution for modelling real-life data is only appropriate if it gives a close approximation to reality.
For this scenario we will assume that if the actual proportion within 1 standard deviation of the mean is within 3% of the theoretical proportion as given by a normal distribution, then the normal distribution can be used to model this data.

Based on this statement, justify (using calculations) whether it is appropriate or not to use a Normal Distribution to model this data.

Actual proportion within $-1 < Z < 1$ $= \frac{310}{422}$

$= 0.7346$

Theoretical proportion within : $= 0.6827$
i.e. $P(-1 < Z < 1)$ if $Z \sim N(0,1)$

Difference in proportions $= 0.7346 - 0.6827$

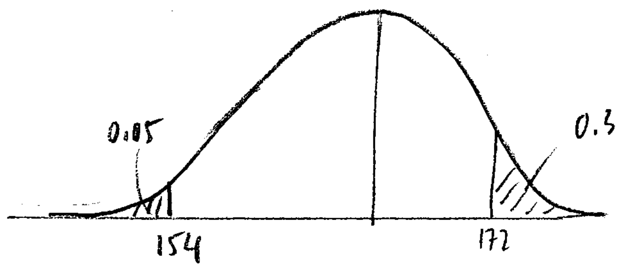
$= 0.052 \sim 5.2\%$

The model is not considered appropriate as $5.2\% > 3\%$ threshold.

7. (5 marks)

The heights of a population of women are normally distributed with mean μ and standard deviation σ cm. It is known that 30% of the women are taller than 172 cm and the 5th percentile is equal to 154 cm.

(a) Sketch a normal distribution showing the above information. [1]



(b) Determine the mean and standard deviation of the heights of the women. [4]

Using $Z \sim N(0,1)$

172 ~ 0.5244σ above μ

154 ~ -1.6449σ below μ

$$Z = \frac{x-\mu}{\sigma}$$

① $0.5244 = \dfrac{172-\mu}{\sigma}$

② $-1.6449 = \dfrac{154-\mu}{\sigma}$

$\mu = 167.65$

$\sigma = 8.298$

8. (10 marks)

Tuna are caught off the coast of Western Australia. The weights of the tuna are normally distributed with a mean of 22 kg. The heaviest 4% of the tuna are processed as fertilizer and the lightest 6% are exported fresh as Class A eating fish. The remainder, which weigh between 17.33 kg and 27.25 kg, are canned.

(a) Determine the probability that a randomly selected tuna is canned. [1]

$P(\text{canned}) = 0.9$

(b) Show the standard deviation of tuna caught off the coast of WA is 3kg. [2]

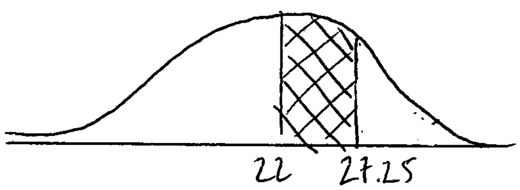
$\mu = 22 \quad P(X \leq 17.33) = 0.06$

$\sigma = 3$

(c) A fishing boat catches a haul of 280 tuna. How many of these would they expect to weigh over 22 kg and be used in cans? [3]

$P(22 < X < 27.25)$

$= \underline{0.4599}$ ✓✓



$= 0.4599 \times 280 \quad \sim \underline{\underline{129}}$ ✓

Kingfish are also caught of the WA coast. Their weights are also normally distributed, with a mean of 20 kg and a standard deviation of 4 kg. At a particular time of the year, the ratio of the numbers of tuna to the numbers of kingfish is 20:1.

(d) Determine the probability that a randomly selected fish which has been caught, known to be either a tuna or a kingfish, weighs less than 22 kg. [4]

$P(KF) = \frac{1}{21} \qquad P(T < 22) = \frac{20}{21} \times 0.5$ ✓

$P(T) = \frac{20}{21} \qquad P(KF < 22) = \frac{1}{21} \times 0.6914$ ✓✓

$P(\text{Fish} < 22) = 0.509.$