Solutions

Stage 2 Mathematical Methods

Statistics Test

Subtopics 5.1, 5.2, 5.3, 6.1, 6.2, 6.3

Total Marks – 36

(Calculator and one A4 page of hand written notes permitted)

QUESTION 1 (8 marks)

Based on historical breed data, the weight of adult male Persian cats is distributed normally with mean $\mu = 6.0$ kg and standard deviation $\sigma = 0.55$ kg. Adult male Persian cats that weigh more than 7 kg are considered to be overweight.

(a) Based on this distribution, what is the proportion of adult male Persian cats that are overweight?

$X \sim N(6.0, 0.55^2)$

$P(X > 7) = 0.0345$ ie. 3.45% ✓ ✓

[normalcdf(7, E99, 6, 0.55)]

(2 marks)

In 2017, researchers undertook a study into the weights of adult male Persian cats. In this study, 70 adult male Persian cats were randomly selected and weighed. This sample of 70 cats had a mean weight of $\bar{x} = 6.7$ kg.

(b) Assuming that the population standard deviation is $\sigma = 0.55$ kg, use the sample data to calculate a 95% confidence interval for μ, the mean weight of adult male Persian cats in 2017.

Using technology, $6.57 \leq \mu \leq 6.83$ kg ✓ ✓

[ZInterval(0.55, 6.7, 70, 0.95)]

(2 marks)

The researchers wrote a press release entitled 'Increase in mean weight of adult male Persian cats'.

(c) Does the confidence interval that you calculated in part (b) support the title of this press release? Justify your answer.

Yes. The entire confidence interval calculated in part (b) lies above the historical mean of 6·0kg. ✓
This suggests that the mean weight of adult male Persian cats has indeed increased. ✓

(2 marks)

(d) The press release states that between 22% and 38% of adult male Persian cats are overweight.

Provide mathematical calculations that support this statement.

Suppose $X \sim N(6{\cdot}57,\ 0{\cdot}55^2)$

Then $P(X > 7) = 0{\cdot}217$ ie. about 22% ✓ [normalcdf(7, E99, 6·57, 0·55)]

Suppose $X \sim N(6{\cdot}83,\ 0{\cdot}55^2)$

Then $P(X > 7) = 0{\cdot}379$ ie. about 38% ✓ [normalcdf(7, E99, 6·83, 0·55)]

(2 marks)

QUESTION 2 (13 marks)

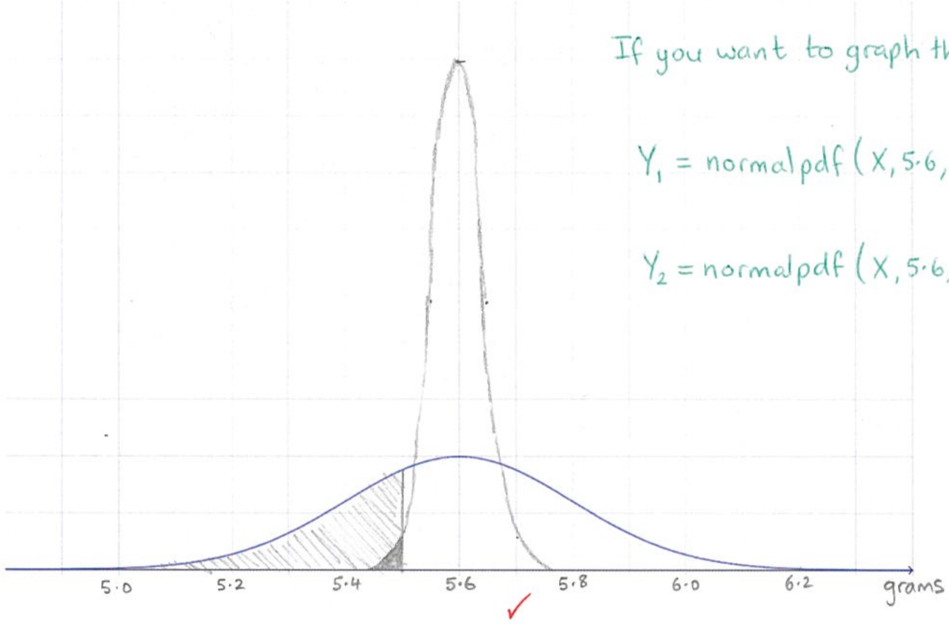
Premium Instant Coffee is produced in sachets. The net weight of the coffee in a randomly chosen sachet can be modelled by W, a normally distributed random variable with a mean of $\mu = 5.6$ grams and a standard deviation of $\sigma = 0.2$ grams.

The distribution of W is graphed below:

If you want to graph them on calculator...

$Y_1 = \text{normalpdf}(X, 5.6, 0.2)$

$Y_2 = \text{normalpdf}(X, 5.6, \frac{0.2}{\sqrt{20}})$



(a) On the horizontal axis of the graph of the normal density curve above, write numbers to illustrate the distribution of W.

(1 mark)

These coffee sachets are sold in packs of twenty. Let $\overline{W}_{20}$ be the average of the net weights of the sachets in a randomly chosen pack.

(b) (i) Write down the mean and standard deviation of the distribution of $\overline{W}_{20}$.

$\mu_{\overline{W}_{20}} = 5.6 \text{ g}$ ✓ $\qquad \mu_{\overline{X}_n} = \mu$

$\sigma_{\overline{W}_{20}} = \frac{0.2}{\sqrt{20}} = 0.0447 \text{g}$ ✓ $\qquad \sigma_{\overline{X}_n} = \frac{\sigma_x}{\sqrt{n}}$

(2 marks)

(ii) On the axes of the graph above, sketch the distribution of $\overline{W}_{20}$.

(A much taller, narrower normal distribution, centred on 5.6g)

(2 marks)

(c) On the graph above, illustrate the fact that $P(W \leq 5.5) > P(\overline{W}_{20} \leq 5.5)$.

lightly shaded area > darkly shaded area ✓

(1 mark)

The packs of twenty coffee sachets are labelled as containing 110 grams net.

(d) Find the probability that a randomly chosen pack of twenty sachets will contain less than its labelled weight.

$P(S_{20} < 110) = P\left(\bar{W}_{20} < \frac{110}{20}\right)$ ✓

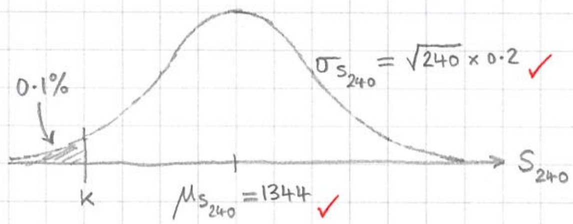
$= 0.0127$ i.e. 1.27% ✓

$\left[\text{normalcdf}\left(-E99, 5.5, 5.6, \frac{0.2}{\sqrt{20}}\right)\right]$

(2 marks)

The coffee sachets are also sold in bulk in catering boxes. These boxes contain 240 sachets.

(e) (i) If 0.1% of catering boxes contain less than k grams, find k to the nearest whole gram.



$\mu_{S_n} = n\mu_X$

$\sigma_{S_n} = \sqrt{n}\,\sigma_X$

$P(S_{240} < k) = 0.001 \implies k = 1334$ g ✓

$\left[\text{invNorm}\left(0.001, 1344, \sqrt{240} \times 0.2\right)\right]$

(3 marks)

(ii) Would it be appropriate to label the catering boxes as containing 1.3 kilograms net?

Give a reason for your answer.

Yes, it would be appropriate. From (e) (i) above, $P(S_{240} < 1334\text{g}) = 0.1\%$

$\therefore P(S_{240} < 1300\text{g}) \ll 0.1\%$ ✓

It is virtually certain that the catering boxes will contain more than 1.3kg net. ✓

(<< means much less than) (2 marks)

QUESTION 3 (7 marks)

The committee of a large sporting association has proposed the major redevelopment of a stadium. For the redevelopment to be approved, all members of the association must vote on the proposal and at least three-quarters must vote yes.

The committee wants to gain information about the likely outcome of the vote. A random sample of 150 members are asked whether they will vote yes or no. Of this sample, 119 plan to vote yes.

(a) What proportion of the sample members is planning to vote yes?

$\hat{p} = \frac{119}{150} = 0.793$ ✓

(1 mark)

(b) Calculate a 95% confidence interval for p, the true proportion of members who are planning to vote yes.

Using technology, $0.729 \leq p \leq 0.858$ ✓ ✓ [1-PropZInt(119, 150, 0.95)]

(2 marks)

(c) *Tick the appropriate box below to indicate your answer.*

This confidence interval suggests that the redevelopment:

Will be approved by the vote ☐

Will not be approved by the vote ☐

May or may not be approved by the vote. ☑ ✓

(1 mark)

(d) Justify the answer you indicated in part (c).

Part of the confidence interval lies below the critical value of 0.75, and part above.
Therefore p could be below or above the $\frac{3}{4}$ mark necessary to approve the redevelopment. ✓

(1 mark)

(e) The committee plans to gain information from a larger sample of members.

Determine the number of members who need to be sampled in order to obtain a 95% confidence interval with a width no greater than 0.065.

$n \geqslant \left(\frac{2 \times 1.96}{0.065}\right)^2 \times 0.793 \times 0.207$ ✓ $n = \left(\frac{2 \times 1.96}{w}\right)^2 p^*(1-p^*)$

$= 596.31$

$\therefore n \geqslant 597$ members ✓

(2 marks)

QUESTION 4 (8 marks)

Heartworm disease – a parasitic condition that affects dogs – is spread by mosquitos. The proportion of dogs affected by the disease varies from place to place. As part of a study, a random sample of 5400 dogs in the Adelaide area were tested for heartworm disease. The study found that 162 of those dogs had heartworm disease.

(a) Calculate $\hat{p}$, the proportion of the sample of dogs that had heartworm disease.

$\hat{p} = \frac{162}{5400} = 0.03$ ✓

(1 mark)

(b) Calculate a 95% confidence interval for p, the proportion of all dogs in the Adelaide area the have heartworm disease.

Using technology, $0.0255 \leqslant p \leqslant 0.0346$ ✓ ✓ [1-PropZInt(162, 5400, 0.95)]

(2 marks)

A smaller, follow-up study is planned

(c) What sample size would be necessary to obtain a confidence interval with a width of 0.032 or less for p, the proportion of all dogs in the Adelaide area that have heartworm disease?

$n \geqslant \left(\frac{2 \times 1.96}{0.032}\right)^2 \times 0.03 \times 0.97$ ✓

$n = \left(\frac{2 \times 1.96}{w}\right)^2 p^*(1-p^*)$

$= 436.68$

$\therefore n \geqslant 437$ ✓

(2 marks)

(d) As a cost saving measure, a sample with $n = 150$ is proposed.

(i) Calculate $n\hat{p}$.

$150 \times 0.03 = 4.5$ ✓

(1 mark)

(ii) Comment on the validity of the confidence interval that would be obtained using this sample size.

The validity of the confidence interval relies on the distribution of $\hat{p}$ being approximately normal. ✓

For this to be the case, we require both $n\hat{p} \geqslant 5$ and $n(1-\hat{p}) \geqslant 5$.

In this case $n\hat{p} < 5$, therefore the confidence interval may not be valid. ✓

(2 marks)