Estimation, Confidence Intervals and Tests Using Normal Distribution

Questions

Q1.

A machine is set to fill pots with yoghurt such that the mean weight of yoghurt in a pot is 505 grams.

To check that the machine is working properly, a random sample of 8 pots is selected. The weight of yoghurt, in grams, in each pot is as follows

508 510 500 500 498 503 508 505

Given that the weights of the yoghurt delivered by the machine follow a normal distribution with standard deviation 5.4 grams,

(a) find a 95% confidence interval for the mean weight, μ grams, of yoghurt in a pot.

Give your answers to 2 decimal places.

(4)

(b) Comment on whether or not the machine is working properly, giving a reason for your answer.

(1)

(c) State the probability that a 95% confidence interval for μ will not contain μ grams.

(1)

(d) Without carrying out any further calculations, explain the changes, if any, that would need to be made in calculating the confidence interval in part (a) if the standard deviation was unknown. Give a reason for your answer.
You may assume that the weights of the yoghurt delivered by the machine still follow a normal distribution.

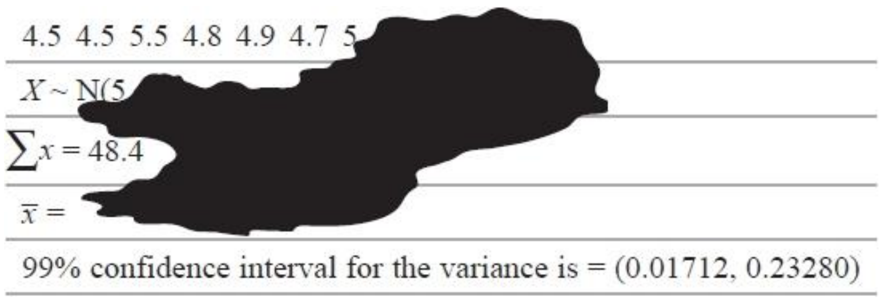
(2)

(Total for question = 8 marks)

Q2.

A company manufactures bolts. The diameter of the bolts follows a normal distribution with a mean diameter of 5 mm.

Stan believes that the mean diameter of the bolts is less than 5 mm. He takes a random sample of 10 bolts and measures their diameters. He calculates some statistics but spills ink on his work before completing them. The only information he has left is as follows



Stating your hypotheses clearly, test, at the 5% level of significance, whether or not Stan's belief is supported.

(9)

(Total for question = 9 marks)

Q3.

Paul takes the company bus to work. According to the bus timetable he should arrive at work at 08 31. Paul believes the bus is not reliable and often arrives late. Paul decides to test the arrival time of the bus and carries out a survey. He records the values of the random variable

X = number of minutes after 08 31 when the bus arrives.

His results are summarised below.

$$n = 15 \qquad \sum x = 60 \qquad \sum x^2 = 1946$$

(a) Calculate unbiased estimates of the mean, μ, and the variance of X.

(3)

Using the mean of Paul's sample and given $X \sim N(\mu, 10^2)$

(b) (i) calculate a 95% confidence interval for the mean arrival time at work for this company bus.

(ii) State an assumption you made about the values in the sample obtained by Paul.

(5)

(c) Comment on Paul's belief. Justify your answer.

(2)

(Total for question = 10 marks)

Q4.

The independent random variables X_1 and X_2 are each distributed B(n, p), where $n > 1$
An unbiased estimator for p is given by

$$\hat{p} = \frac{aX_1 + bX_2}{n}$$

where a and b are constants.

[You may assume that if X_1 and X_2 are independent then $E(X_1X_2) = E(X_1)E(X_2)$]

(a) Show that $a + b = 1$

(2)

(b) Show that $\text{Var}(\hat{p}) = \dfrac{(2a^2 - 2a + 1)p(1 - p)}{n}$

(4)

(c) Hence, justifying your answer, determine the value of a and the value of b for which $\hat{p}$ has minimum variance.

(d) (i) Show that $\hat{p}^2$ is a biased estimator for p^2

(ii) Show that the bias → 0 as $n \to \infty$

(5)

(e) By considering $E[X_1(X_1 - 1)]$ find an unbiased estimator for p^2

(3)

(Total for question = 14 marks)

Mark Scheme – Estimation, Confidence Intervals and Tests Using Normal distribution

Q1.

Question	Scheme	Marks	AOs
(a)	Mean = 504	B1	1.1b
	1.96	B1	3.3
	$504 \pm \frac{5.4}{\sqrt{8}} \times "1.96"$	M1	2.1
	(500.258, 507.742)	A1	1.1b
		(4)	
(b)	**<u>505</u>** is in the confidence **<u>interval</u>** therefore there is evidence that the machine is **<u>working</u>** properly	B1ft	2.2b
		(1)	
(c)	5% oe	B1	1.1b
		(1)	
(d)	s needs to be used instead of σ **and** a t-value instead of the z value	B1	3.3
	since the sample is small therefore you can't use the normal distribution	B1	3.5b
		(2)	
			(8 marks)

Notes:		
(a)	B1	504 may be seen in part(b)
	B1	For realising a normal distribution must be used as a model and finding the correct value 1.96
	M1	For $504 \pm \frac{5.4}{\sqrt{8}} \times "z \text{ value}"$. $\lvert z \rvert > 1$ May be implied by a correct CI
	A1	awrt 500.26 and 507.74 NB using t gives 500.29 and 507.71
(b)	B1ft	Drawing a correct inference (ft) using their answer to part (a) and the 505 from the question. Reason must be given. Ignore incorrect non – contextual
(c)	B1	5%
(d)	B1	create new model by using s and t. Allow if state use CI $\mu \pm \frac{s}{\sqrt{n}} \times "t"$ or use $s = 4.44$ and $t = 2.365$
	B1	For recognising that the sample is small

Q2.

Question	Scheme	Marks	AOs
	99% confidence interval for Var uses χ^2 values of 1.735 or 23.589	B1	3.3
	$\frac{9s^2}{1.735} = 0.2328$ or $\frac{9s^2}{23.589} = 0.01712$	M1	2.1
	$s^2 = \frac{0.2328 \times "1.735"}{9}$ or $\frac{0.01712 \times "23.589"}{9}$ $[= 0.04487...]$	dM1	1.1b
	$\bar{x} = 4.84$	B1	1.1b
	$H_0 : \mu = 5$ $H_1 : \mu < 5$	B1	2.5
	CV $t_9 = -1.833$	B1	1.1b
	$t = \pm \frac{"4.84" - 5}{\sqrt{"0.0449"/10}}$	M1	1.1b
	$=$ awrt -2.39	A1	1.1b
	Stan's belief is supported **or** there is evidence that the **mean diameter** of the bolts is less than **5mm**	A1ft	2.2b
		(9)	
			(9 marks)

Notes:
B1: For realising a χ^2 distribution must be used as a model and finding a correct value
M1: For realising the need to set $\frac{9s^2}{"\text{smallest } \chi^2"} = 0.2328$ or $\frac{9s^2}{"\text{largest } \chi^2"} = 0.01712$
dM1: correct method used to solve equation to find s^2
B1: awrt 4.84
B1: Both hypotheses correct using the notation μ
B1: $\pm$ 1.833
M1: For us of correct formula ie $\pm \frac{"\text{their } 4.84" - 5}{\sqrt{"\text{their } 0.0449"/10}}$ If "4.84" not shown it must be correct here
A1: -2.39
A1ft: Drawing a correct inference following through their CV and test statistic (must have matching signs)
NB if chi squared values not shown $s^2 = 0.045$ or 0.0449 award B0 M1M1 for awrt 0.04487 award B1 M1 A1
Use of $2(2.5758)\frac{\sigma}{\sqrt{10}} = 0.21568$ gives $\sigma = \sqrt{0.0175}$ could get B0M0M0B1B1B1M0A0A0 Unless continue to get $s^2 = \frac{10}{9}0.0175 = 0.0194...$
Use of $2(1.833)\frac{s}{\sqrt{10}} = 0.21568$ gives $s = 0.1860$ could get B0M0M0B1B1B1M1A0A1

Q3.

Question Number	Scheme	Notes	Marks
(a)	$\bar{x} = \frac{60}{15} = 4$	4 cao	B1
	$s^2 = \frac{1}{14}\left(1946 - 15\times 4^2\right) = 121.857...$	M1 Use of complete, correct formula and attempt to substitute. A1 awrt 122 or $\frac{853}{7}$	M1,A1
			(3)
(b)(i)	$\bar{x} \pm 1.96\times\frac{10}{\sqrt{15}} = 4 \pm 5.06$	Accept use of $\bar{x} \pm z\times\frac{10 \text{ or "their } s\text{"}}{\sqrt{15}}$, A1 all correct. Accept $\bar{x} = 0835$.	M1,A1
	$(-1.06, 9.06)$	Can be implied from correct interval below.	A1
	$(0829.56, 0840.04)$	Accept $(0829.94, 0840.06)$ or expressed using words or as an inequality. Accept answers to the nearest minute ie $(0830, 0840)$.	A1
(ii)	Paul samples times of **buses randomly** or **independently** of each other	Context required.	B1
			(5)
(c)	0 / 0831 / 8.31(am) is 'contained in' the confidence interval	Award if comment about their interval is correct. Only accept 'above the lower limit of' etc if the statement taken as a whole clearly means 'contained in'.	M1
	Paul's belief is not supported / 0831 arrival time is reasonable	Must contain some context	A1cao
			(2)
			Total 10

Q4.

Question Number	Scheme	Marks
(a)	$\mathrm{E}\left(\frac{aX_1+bX_2}{n}\right)=\frac{anp+bnp}{n}=ap+bp=(a+b)p$	M1
	$a+b=1$ *	A1* cso (2)
(b)	$\mathrm{Var}\left(\frac{aX_1+bX_2}{n}\right)=\frac{1}{n^2}\left(a^2np(1-p)+b^2np(1-p)\right)$	M1 A1
	$=\frac{p(1-p)\left(a^2+b^2\right)}{n}$	
	$=\frac{p(1-p)\left(a^2+(1-a)^2\right)}{n}$	M1d
	$=\frac{\left(2a^2-2a+1\right)p(1-p)}{n}$ *	A1* cso
		(4)
(c)	Min value when $\frac{(4a-2)p(1-p)}{n}=0$	M1A1
	$\Rightarrow 4a-2=0$	
	$a=\frac{1}{2},\ b=\frac{1}{2}$	A1A1ft
	$\frac{\mathrm{d}^2\mathrm{Var}(\hat{p})}{\mathrm{d}a^2}=\frac{4p(1-p)}{n}>0$ or $\because$ quadratic with positive x^2 $\therefore$ minimum point or sketch	B1 (5)

(d)(i)	$\mathrm{E}\left(\frac{aX_1+bX_2}{n}\right)^2=\mathrm{E}\left(\frac{a^2X_1^{\,2}+b^2X_2^{\,2}+2abX_1X_2}{n^2}\right)$	M1
	$=\frac{1}{n^2}\left(a^2np(1-p)+a^2n^2p^2+b^2np(1-p)+b^2n^2p^2+2abn^2p^2\right)$	M1d
	$=\frac{(a^2+b^2)np(1-p)+(a+b)^2n^2p^2}{n^2}$	
	$=\frac{(a^2+b^2)p(1-p)}{n}+p^2(a+b)^2$	
	$=\frac{(a^2+b^2)p(1-p)}{n}+p^2$; $>p^2$ since $\frac{(a^2+b^2)p(1-p)}{n}>0$ oe $\therefore$ biased	A1;A1
(ii)	As $n\to\infty$ $\mathrm{E}\left(\hat{p}^2\right)\to p^2$ Therefore bias $\to 0$	B1
		(5)
(e)	$\mathrm{E}(X_1(X_1-1))=\mathrm{E}\left(X_1^{\,2}\right)-\mathrm{E}(X_1)$	
	$=np(1-p)+n^2p^2-np$	M1
	$=np-np^2+n^2p^2-np$	
	$=np^2(n-1)$	A1
	Unbiased estimator $=\frac{X_1(X_1-1)}{n(n-1)}$	A1 (3) Total 19

	Notes	
(a)	M1 Using $\frac{a\text{E}(X_1) + b\text{E}(X_2)}{n}$ and subst $\text{E}(X_1) = np$ and $\text{E}(X_2) = np$	
	Acso* Answer given . Need $p(a + b) = p$ and statement $a + b = 1$ and no errors	
(b)	M1 Using $\frac{a^2\text{Var}(X_1) + b^2\text{Var}(X_2)}{n^2}$ and subst $\text{Var}(X_1) = np(1 - p)$ – may be implied by $\frac{1}{n^2}\left(a^2np(1-p)+b^2np(1-p)\right)$ A1 correct answer in any form M1d dep on 1st M1 Subst $b = 1 - a$ A1cso* method must be shown and no errors.	
(c)	M1 $\frac{\text{d}}{\text{da}}(\text{Var})$ (must differentiate with respect a) or attempt to complete the square A1 correct diff = 0 or $2\left(a-\tfrac{1}{2}\right)^2+\tfrac{1}{2}$ A1 $a = 0.5$ A1 ft for $b = 1 - a$ B1 for a reason why minimum	
(d)(i) (ii)	M1 multiplying out and using $\text{E}(aX) = a\,\text{E}(X)$ [may use their values of a and b] M1d dependent on previous M being awarded Using $\text{E}(X^2) = \text{Var}(X) + [\text{E}(X)]^2$ A1 $\frac{(2a^2-2\text{a}+1)p(1-p)}{n}+p^2$ or $\frac{(a^2+b^2)p(1-p)}{n}+p^2$ must be of the form p^2 + a single term A1 for a reason why it is not equal p^2 plus statement to say biased. B1 Follow on from their expression $p^2 + \ldots$ with a and b.	
(e)	M1 multiplying out correctly and subst np for $\text{E}(X)$ or using $\text{E}(\hat{p})^2 = \text{Var}(\hat{p}) + [\text{E}(\hat{p})]^2$ Allow $=\frac{(2a^2-2\text{a}+1)p(1-p)}{n}+p^2$ A1 $np^2(n-1)$ A1 $\frac{X_1(X_1-1)}{n(n-1)}$	
	NB $\frac{X_1(X_1-1)}{n(n-1)}$ gains all 3 marks.	