Paper Reference(s)

6686/01

Edexcel GCE

Statistics S4

Advanced Level

Wednesday 18 June 2008 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Green)	Nil

Candidates may use any calculator allowed by the regulations of the Joint Council for Qualifications. Calculators must not have the facility for symbolic algebra manipulation, differentiation and integration, or have retrievable mathematical formulas stored in them.

Instructions to Candidates

In the boxes on the answer book, write the name of the examining body (Edexcel), your centre number, candidate number, the unit title (Statistics S4), the paper reference (6686), your surname, other name and signature.
Values from the statistical tables should be quoted in full. When a calculator is used, the answer should be given to an appropriate degree of accuracy.

Information for Candidates

A booklet 'Mathematical Formulae and Statistical Tables' is provided.
Full marks may be obtained for answers to ALL questions.
This paper has 7 questions.
The total mark for this paper is 75.

Advice to Candidates

You must ensure that your answers to parts of questions are clearly labelled.
You must show sufficient working to make your methods clear to the Examiner.
Answers without working may gain no credit.

1. A random sample $X_1, X_2, ..., X_{10}$ is taken from a population with mean μ and variance σ^2.

(*a*) Determine the bias, if any, of each of the following estimators of μ.

$$\theta_1 = \frac{X_3 + X_4 + X_5}{3},$$

$$\theta_2 = \frac{X_{10} - X_1}{3},$$

$$\theta_3 = \frac{3X_1 + 2X_2 + X_{10}}{6}.$$

(4)

(*b*) Find the variance of each of these estimators.

(5)

(*c*) State, giving reasons, which of these three estimators for μ is

(i) the best estimator,

(ii) the worst estimator.

(4)

---

2. A large number of students are split into two groups *A* and *B*. The students sit the same test but under different conditions. Group *A* has music playing in the room during the test, and group *B* has no music playing during the test. Small samples are then taken from each group and their marks recorded. The marks are normally distributed.

The marks are as follows:

Sample from Group *A*	42	40	35	37	34	43	42	44	49
Sample from Group *B*	40	44	38	47	38	37	33		

(*a*) Stating your hypotheses clearly, and using a 10% level of significance, test whether or not there is evidence of a difference between the variances of the marks of the two groups.

(8)

(*b*) State clearly an assumption you have made to enable you to carry out the test in part (*a*).

(1)

(*c*) Use a two tailed test, with a 5% level of significance, to determine if the playing of music during the test has made any difference in the mean marks of the two groups. State your hypotheses clearly.

(7)

(*d*) Write down what you can conclude about the effect of music on a student's performance during the test.

(1)

---

3. The weights, in grams, of mice are normally distributed. A biologist takes a random sample of 10 mice. She weighs each mouse and records its weight.

The ten mice are then fed on a special diet. They are weighed again after two weeks.

Their weights in grams are as follows:

Mouse	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*	*I*	*J*
Weight before diet	50.0	48.3	47.5	54.0	38.9	42.7	50.1	46.8	40.3	41.2
Weight after diet	52.1	47.6	50.1	52.3	42.2	44.3	51.8	48.0	41.9	43.6

Stating your hypotheses clearly, and using a 1% level of significance, test whether or not the diet causes an increase in the mean weight of the mice.

(8)

---

4. A town council is concerned that the mean price of renting two bedroom flats in the town has exceeded £650 per month. A random sample of eight two bedroom flats gave the following results, £x, per month.

705, 640, 560, 680, 800, 620, 580, 760

[You may assume $\sum x = 5345$, $\sum x^2 = 3\,621\,025$.]

(*a*) Find a 90% confidence interval for the mean price of renting a two bedroom flat.

(6)

(*b*) State an assumption that is required for the validity of your interval in part (*a*).

(1)

(*c*) Comment on whether or not the town council is justified in being concerned. Give a reason for your answer.

(2)

---

5. A machine is filling bottles of milk. A random sample of 16 bottles was taken and the volume of milk in each bottle was measured and recorded. The volume of milk in a bottle is normally distributed and the unbiased estimate of the variance, s^2, of the volume of milk in a bottle is 0.003.

(*a*) Find a 95% confidence interval for the variance of the population of volumes of milk from which the sample was taken.

(5)

The machine should fill bottles so that the standard deviation of the volumes is equal to 0.07.

(*b*) Comment on this with reference to your 95% confidence interval.

(3)

---

6. A drug is claimed to produce a cure to a certain disease in 35% of people who have the disease. To test this claim a sample of 20 people having this disease is chosen at random and given the drug. If the number of people cured is between 4 and 10 inclusive the claim will be accepted. Otherwise the claim will not be accepted.

(*a*) Write down suitable hypotheses to carry out this test.

(2)

(*b*) Find the probability of making a Type I error.

(3)

The table below gives the value of the probability of the Type II error, to 4 decimal places, for different values of p where p is the probability of the drug curing a person with the disease.

P(cure)	0.2	0.3	0.4	0.5
P(Type II error)	0.5880	r	0.8565	s

(*c*) Calculate the value of r and the value of s.

(3)

(*d*) Calculate the power of the test for $p = 0.2$ and $p = 0.4$

(2)

(*e*) Comment, giving your reasons, on the suitability of this test procedure.

(2)

---

7. An engineering firm buys steel rods. The steel rods from its present supplier are known to have a mean tensile strength of 230 N/mm^2.

A new supplier of steel rods offers to supply rods at a cheaper price than the present supplier. A random sample of ten rods from this new supplier gave tensile strengths, x N/mm^2, which are summarised below.

Sample size	Σx	Σx^2
10	2283	524 079

(*a*) Stating your hypotheses clearly, and using a 5% level of significance, test whether or not the rods from the new supplier have a tensile strength lower than the present supplier. (You may assume that the tensile strength is normally distributed).

(7)

(*b*) In the light of your conclusion to part (*a*) write down what you would recommend the engineering firm to do.

(1)

TOTAL FOR PAPER: 75 MARKS

END

June 2008
6686 Statistics S4
Mark Scheme

Question Number	Scheme	Marks
1 a	$E(\theta_1) = \frac{E(X_3)+E(X_4)+E(X_5)}{3}$ $= \frac{3\mu}{3}$ $= \mu$ Bias = 0 allow unbiased	B1
	$E(\theta_2) = \frac{E(X_{10}) - E(X_1)}{3}$ $= 1/3(\mu - \mu)$ $= 0$ Bias = $-\mu$ allow $\pm\mu$	B1,B1
	$E(\theta_3) = \frac{3E(X_1)+2E(X_2)+E(X_{10})}{6}$ $= \frac{3\mu+2\mu+\mu}{6}$ $= \mu$ Bias = 0 allow unbiased	B1 (4)
b	$Var(\theta_1) = \frac{1}{9}\{(Var\ X_2) + Var(X_3) + Var(X_4)\}$	M1
	$= \frac{1}{9}\{\sigma^2+\sigma^2+\sigma^2\}$	
	$= \frac{1}{3}\sigma^2$	A1
	$Var(\theta_2) = \frac{2}{9}\sigma^2$	B1
	$Var(\theta_3) = \frac{1}{36}\{9\sigma^2+4\sigma^2+\sigma^2\}$	M1
	$= \frac{7}{18}\sigma^2$	A1 (5)
ci)	θ_1 is the better estimator. It has a lower var. and no bias	B1 depB1
ii)	θ_2 is the worst estimator. It is biased	B1 depB1 (4)

Question Number	Scheme	Marks
2 a	$H_1 : \sigma_A^2 = \sigma_B^2$ $H_0 : \sigma_A^2 \neq \sigma_B^2$	B1
	$S_A^2 = 22.5$ $s_B^2 = 21.6$ awrt	M1 A1A1
	$\frac{s_1^2}{s_2^2} = 1.04$	M1 A1
	$F_{(8, 6)} = 4.15$	B1
	1.04 < 4.15 do not reject H_0. The variances are the same.	B1 (8)
b	Assume the samples are selected at random, (independent)	B1 (1)
c	$s^2_p = \frac{8(22.5)+6(21.62)}{14} = 22.12$ awrt 22.1	M1 A1
	$H_0 : \mu_A = \mu_B$ $H_1 : \mu_A \neq \mu_B$	B1
	$t = \frac{40.667 - 39.57}{\sqrt{22.12}\sqrt{\frac{1}{9}+\frac{1}{7}}}$	M1
	= 0.462 0.42 – 0.47	A1
	Critical value = $t_{14}(2.5\%) = 2.145$	B1
	0.462 < 2.145 No evidence to reject H_0. The means are the same	B1 (7)
d	Music has no effect on performance	B1 (1)

Question Number	Scheme	Marks
3	Differences 2.1 -0.7 2.6 -1.7 3.3 1.6 1.7 1.2 1.6 2.4	M1
	$\bar{d}$ = 1.41	M1
	$H_0 : \mu_d = 0$ $H_1 : \mu_d > 0$	B1
	$s = \sqrt{\frac{40.65 - 10 \times 1.41^2}{9}} = 1.5191\ldots$	M1
	$t = \frac{1.41}{\left(\frac{1.519...}{\sqrt{10}}\right)} = 2.935\ldots$ awrt 2.94 /2.93	M1 A1
	t_9 (1%) = 2.821	B1
	2.935.. > 2.821 Evidence to reject H_0. There has been an increase in the mean weight of the mice.	B1ft
		(8)

2 sample test can score
M0 M0
B1 for $H_0 : \mu_A = \mu_B$ $H_1 : \mu_A < \mu_B$

M1 $\frac{9 \times 24.5 + 9 \times 17.16}{18}$
M0 A0
B1 2.552
B1 ft

ie 4/8

Question Number	Scheme	Marks
4a	$\bar{x}$ = 668.125 s = 84.428	M1 M1
	T_7(5%) = 1.895	B1
	Confidence limits = $668.125 \pm \frac{1.895 \times 84.428}{\sqrt{8}}$	M1
	= 611.6 and 724.7	A1A1
	Confidence interval = (612, 725)	(6)
b	Normal distribution	B1 (1)
c	£650 is within the confidence interval. No need to worry.	B1 √ B1 √ (2)

Question Number	Scheme	Marks
5 a	Confidence interval = $\left(\frac{15\times 0.003}{27.488}, \frac{15\times 0.003}{6.262}\right)$	M1 B1B1
	= (0.00164, 0.00719)	A1 A1 (5)
b	0.07^2 =0.0049	M1
	0.0049 is within the 95% confidence interval.	A1
	There is no evidence to reject the idea that the standard deviation of the volumes is not 0.07 or The machine is working well.	A1 (3)

Question Number	Scheme	Marks
6 a	H_0 $p = 0.35$ $H_1 : p \neq 0.35$	B1 B1 (2)
b	Let X = Number cured then $X \sim$ B(20, 0.35)	B1
	α = P(Type I error) = P($x \leq 3$) +P($x \geq 11$) given p= 0.35	M1
	= 0.0444 + 0.0532	
	= 0.0976	A1 (3)
c	β = P(Type II error) = P($4 \leq x \leq 10$)	M1
	p 0.2 0.3 0.4 0.5	
	β 0.5880 **0.8758** 0.8565 **0.5868**	A1A1 (3)
d	Power = 1 – B	M1
	0.4120 0.1435	A1 (2)
e	Not a good procedure.	B1
	Better further away from 0.35 or	B1dep
	This is not a very powerful test (power = 1 - β)	(2)

Question Number	Scheme	Marks
7 a	$H_0 : \mu = 230 \quad H_1 : \mu < 230$	B1
	$\nu = 9$	
	From table critical value $= \pm 1.833$	B1√
	$\bar{x} = 228.3 \quad S = 17.858$	B1 B1
	$t = \frac{\bar{x} - \mu}{\frac{s}{\sqrt{n}}}$	M1
	$= \pm \frac{228.3 - 230}{\frac{17.858}{\sqrt{10}}} = \pm 0.301$	A1
	$\pm 0.301 > \pm 1.833$. No evidence to reject H_0. Mean is 230 N/mm^2	B1 (7)
b	Since the tensile strength is the same and the price is cheaper recommend use new supplier.	B1 (1)