Paper Reference(s)

6691/01

Edexcel GCE

Statistics S3

Advanced Level

Thursday 20 June 2011 – Morning

Time: 1 hour 30 minutes

Materials required for examination	Items included with question papers
Mathematical Formulae (Pink)	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 S3), the paper reference (6691), 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 not gain full credit.

1. Explain what you understand by the Central Limit Theorem.

(3)

---

2. A county councillor is investigating the level of hardship, h, of a town and the number of calls per 100 people to the emergency services, c. He collects data for 7 randomly selected towns in the county. The results are shown in the table below.

Town	A	B	C	D	E	F	G
h	14	20	16	18	37	19	24
c	52	45	43	42	61	82	55

(*a*) Calculate the Spearman's rank correlation coefficient between h and c.

(6)

After collecting the data, the councillor thinks there is no correlation between hardship and the number of calls to the emergency services.

(*b*) Test, at the 5% level of significance, the councillor's claim. State your hypotheses clearly.

(4)

---

3. A factory manufactures batches of an electronic component. Each component is manufactured in one of three shifts. A component may have one of two types of defect, D_1 or D_2, at the end of the manufacturing process. A production manager believes that the type of defect is dependent upon the shift that manufactured the component. He examines 200 randomly selected defective components and classifies them by defect type and shift.

The results are shown in the table below.

Shift \ Defect type	D_1	D_2
First shift	45	18
Second shift	55	20
Third shift	50	12

Stating your hypotheses, test, at the 10% level of significance, whether or not there is evidence to support the manager's belief. Show your working clearly.

(10)

---

4. A shop manager wants to find out if customers spend more money when music is playing in the shop. The amount of money spent by a customer in the shop is £x. A random sample of 80 customers, who were shopping without music playing, and an independent random sample of 60 customers, who were shopping with music playing, were surveyed. The results of both samples are summarised in the table below.

	$\sum x$	$\sum x^2$	Unbiased estimate of mean	Unbiased estimate of variance
Customers shopping **without** music	5 320	392 000	x	s^2
Customers shopping **with** music	4 140	312 000	69.0	446.44

(*a*) Find the values of $\bar{x}$ and s^2.

(5)

(*b*) Test, at the 5% level of significance, whether or not the mean money spent is greater when music is playing in the shop. State your hypotheses clearly.

(8)

5. The number of hurricanes per year in a particular region was recorded over 80 years. The results are summarised in Table 1 below.

No of hurricanes, h	0	1	2	3	4	5	6	7
Frequency	0	2	5	17	20	12	12	12

Table 1

(*a*) Write down two assumptions that will support modelling the number of hurricanes per year by a Poisson distribution.

(2)

(*b*) Show that the mean number of hurricanes per year from Table 1 is 4.4875.

(2)

(*c*) Use the answer in part (*b*) to calculate the expected frequencies r and s given in Table 2 below to 2 decimal places.

(3)

h	0	1	2	3	4	5	6	7 or more
Expected frequency	0.90	4.04	r	13.55	s	13.65	10.21	13.39

Table 2

(*d*) Test, at the 5% level of significance, whether or not the data can be modelled by a Poisson distribution. State your hypotheses clearly.

(6)

6. The lifetimes of batteries from manufacturer A are normally distributed with mean 20 hours and standard deviation 5 hours when used in a camera.

(*a*) Find the mean and standard deviation of the total lifetime of a pack of 6 batteries from manufacturer A.

(2)

Judy uses a camera that takes one battery at a time. She takes a pack of 6 batteries from manufacturer A to use in her camera on holiday.

(*b*) Find the probability that the batteries will last for more than 110 hours on her holiday.

(2)

The lifetimes of batteries from manufacturer B are normally distributed with mean 35 hours and standard deviation 8 hours when used in a camera.

(*c*) Find the probability that the total lifetime of a pack of 6 batteries from manufacturer A is more than 4 times the lifetime of a single battery from manufacturer B when used in a camera.

(6)

---

7. Roastie's Coffee is sold in packets with a stated weight of 250 g. A supermarket manager claims that the mean weight of the packets is less than the stated weight. She weighs a random sample of 90 packets from their stock and finds that their weights have a mean of 248 g and a standard deviation of 5.4 g.

(*a*) Using a 5% level of significance, test whether or not the manager's claim is justified. State your hypotheses clearly.

(5)

(*b*) Find the 98% confidence interval for the mean weight of a packet of coffee in the supermarket's stock.

(4)

(*c*) State, with a reason, the action you would recommend the manager to take over the weight of a packet of Roastie's Coffee.

(2)

Roastie's Coffee company increase the mean weight of their packets to μ g and reduce the standard deviation to 3 g. The manager takes a sample of size n from these new packets. She uses the sample mean $\overline{X}$ as an estimator of μ.

(*d*) Find the minimum value of n such that $P(|\overline{X} - \mu| < 1) \geq 0.98$.

(5)

TOTAL FOR PAPER: 75 MARKS

END

Question Number	Scheme	Marks
1.	$X_1, X_2, ...X_n$ is a random sample of size n, for large n,	B1
	drawn from a population of any distribution with mean μ and variance σ^2	B1
	then $\bar{X}$ is (approximately) $\mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$	B1
		(3) **3**
2. **(a)**		

Town	A	B	C	D	E	F	G
h rank	1	5	2	3	7	4	6
c rank	4	3	2	1	6	7	5
$\lvert d \rvert$	3	2	0	2	1	3	1
d^2	9	4	0	4	1	9	1

Question Number	Scheme	Marks
	($\lvert d \rvert$ row)	M1
	(d^2 row)	M1
	$\sum d^2 = 28$	M1A1
	$r_s = 1 - \frac{6 \times 28}{7 \times 48}$	M1
	$= 0.5$	A1
		(6)
(b)	$\mathrm{H}_0 : \rho = 0, \mathrm{H}_1 : \rho \neq 0$	B1
	Critical values are $r_s = \pm 0.7857$	B1ft
	0.5<0.7857 insufficient evidence to reject H_0	M1
	Councillor's claim is supported.	A1ft
		(4) **10**

Question Number	Scheme	Marks
3.		

Defect Type / Shift	D_1	D_2	
First Shift	47.25	15.75	63
Second Shift	56.25	18.75	75
Third Shift	46.5	15.5	62
	150	50	200

M1A1

H_0 : Type of defect is independent of Shift (no association)
H_1 : Type of defect is not independent of Shift (association)

B1

O	E	$\frac{(O-E)^2}{E}$	$\frac{O_i^2}{E_i}$
45	47.25	0.1071…	42.857...
18	15.75	0.3214…	20.571..
55	56.25	0.02777…	53.777...
20	18.75	0.0833…	21.333...
50	46.5	0.2634…	53.763...
12	15.5	0.7903…	9.290...

M1A1

$\frac{(O-E)^2}{E}$ =1.5934.. or $\frac{O_i^2}{E_i}$ -200=201.5934-200=1.5934.. awrt1.59 — A1

$\nu = (3-1)(2-1) = 2$ — B1

$\chi_2^2(0.10) = 4.605$ — B1ft

1.59<4.605 so insufficient evidence to reject H_0 — M1

Insufficient evidence to support manager's belief /claim. — A1

10

Question Number	Scheme	Marks
4.		
(a)	$\bar{x} = \frac{5320}{80} = 66.5$	M1,A1
	$s^2 = \frac{392000 - 80 \times (66.5)^2}{79}$	M1A1ft
	$= 483.797\ldots$ awrt 484	A1
		(5)
(b)	$H_0: \mu_m = \mu_{nm}$, $H_1: \mu_m > \mu_{nm}$ (accept μ_1, μ_2 with definition)	B1B1
	$z = \frac{69.0 - 66.5}{\sqrt{\frac{483.797}{80} + \frac{446.44}{60}}}$	M1dM1
	$= 0.6807$ awrt 0.681	A1
	One tailed cv 1.6449 (Probability is awrt 0.752)	B1
	0.6807<1.6449 (or 0.248>0.05) insufficient evidence to reject H_0	dM1
	Mean money spent is not greater with music playing.	A1ft
		(8) **13**

Question Number	Scheme	Marks
5. **(a)**	Hurricanes: occur singly / are independent or occur at random /are a rare event / at a constant rate	B1B1 (2)
(b)	From data $\dfrac{1\times 2+2\times 5+3\times 17+..+7\times 12}{80}=4.4875$	M1A1 (2)
(c)	(see table below)	M1A1A1 (3)
(d)	(see table below)	M1

No of hurricanes, h	0	1	2	3	4	5	6	7+
$80\mathrm{P}(X=h)$	0.9	4..038	***r*=9.06 …**	13.55	***s*=15.205**	13.647…	10.206…	13.388…
Combine to give expected frequencies >5	13.9991			13.55	15.205…	13.647…	10.206…	13.388…
Observed	7			17	20	12	12	12
$\dfrac{(O-E)^2}{E}$	3.499…			0.876…	1.511…	0.198…	0.315…	0.143…
$\dfrac{O_i^2}{E_i}$	3.500...			21.322...	26.306...	10.551...	14.108...	10.755..

Scheme	Marks
H_0: Poisson distribution is a good fit o.e. H_1: Poisson distribution is not a good fit o.e.	B1
$\sum \dfrac{(O_i-E_i)^2}{E_i}$ =6.545.. <u>or</u> $\dfrac{O_i^2}{E_i}$ =86.545-80=6.545.. (awrt **6.55 or 6.54**)	A1
$\nu = 6-2=4$	B1
cv is 9.488 (ft their ν i.e. $\chi_\nu^{\,2}(0.05)$)	B1ft
6.545<9.488 so insufficient evidence to reject H_0	
(Hurricanes) can be modelled by a Poisson distribution	A1 (6)
	13

Question Number	Scheme	Marks
6. (a)	$L = A_1 + A_2 + \ldots + A_6$ Mean is $\text{E}(L) = 6 \times 20 = 120$ Standard deviation is $\sqrt{\text{Var}(W)} = \sqrt{6 \times 5^2} = 5\sqrt{6} = 12.247...$ awrt 12.2	 B1 B1 (2)
(b)	$P(L > 110) \quad = P(Z > \left(\frac{110-120}{12.247...}\right))$ $= P(Z < 0.8164...)$ $= 0.7939$ (or 0.7929 using interpolation or 0.79289 by calc)	M1 A1 (2)
(c)	Let $X = 4B - \sum_1^6 A_i$ $\text{E}(X) = 140 - 120 = 20$ $\text{Var}(X) = 16 \times 8^2 + 6 \times 5^2 = 1174$ $\text{P}(X < 0) = \text{P}\left(Z < \frac{-20}{\sqrt{1174}}\right) = \text{P}(Z < -0.583\ldots)$ $= 0.2797$ (or 0.2810 if no interpolation) or 0.27971 by calc.	 B1 M1M1A1 M1 A1 (6) **10**

Question Number	Scheme	Marks
7. **(a)**	$H_0: \mu = 250$, $H_1: \mu < 250$,	B1
	$z = \dfrac{248 - 250}{\frac{5.4}{\sqrt{90}}}$	M1
	$= -3.513...$ awrt -3.51	A1
	Critical value -1.6449	B1
	-3.513..<-1.6449 so sufficient evidence to reject H_0	
	Manager's claim is justified.	A1
		(5)
(b)	98% CI for μ is	
	$248 \pm 2.3263 \times \dfrac{5.4}{\sqrt{90}}$	M1B1
	= awrt (247,249) dependent upon z value awrt 2.33	A1A1
		(4)
(c)	Hypothesis test is significant or CI does not contain stated weight.	B1
	(Manager should ask the company to investigate if their) stated weight is too high o.e.	B1
		(2)
(d)	$P(\lvert \bar{x} - \mu \rvert < 1) = 0.98$	
	$\dfrac{1}{\frac{3}{\sqrt{n}}} = 2.3263$	M1 A1
	$n = (3 \times 2.3263)^2 = 48.7...$	dM1A1
	Sample size 49 required.	A1
		(5)
		16