

June 2006
6691 Statistics S3
Mark Scheme

Question Number	Scheme	Marks
1 (a)	<p><u>Advantages:</u></p> <ul style="list-style-type: none"> - does not require the existence of a ^{Sampling frame} population list - <u>field work can be done quickly</u> as representative sample can be achieved with a small sample size - costs kept to a minimum (<u>cheaply</u>) - administration relatively <u>easy</u> - non-response not an issue <p><u>Disadvantages:</u></p> <ul style="list-style-type: none"> - not possible to estimate sampling errors - interviewer choice and may not be able to judge easily / <u>may lead to bias</u> - non-response not recorded - non-random process 	<p style="text-align: right;">any one</p> <p style="text-align: right;">B1</p> <p style="text-align: right;">any one</p> <p style="text-align: right;">B1</p> <p style="text-align: right;">(2)</p>
(b)	<p><u>Advantages:</u></p> <ul style="list-style-type: none"> - <u>random process</u> so possible to <u>estimate sampling errors</u> - free from <u>bias</u> <p><u>Disadvantages:</u></p> <ul style="list-style-type: none"> - not suitable when sample size is large - <u>sampling frame required</u> which <u>may not exist</u> or may be difficult to construct for a large population. 	<p style="text-align: right;">any one</p> <p style="text-align: right;">B1</p> <p style="text-align: right;">any one</p> <p style="text-align: right;">B1 (2)</p> <p style="text-align: right;">TOTAL 4</p>

NO REPETITION / OPPOSITES

Question Number	Scheme	Marks
2 (a)	$\bar{X} \sim N(90, \frac{5^2}{100})$ i.e. $N_9(90, 0.25)$ Application of <u>central limit theorem</u> as (sample large)	M1A1 B1 (3)
(b)	$P(\bar{X} \geq 91) = 1 - P(Z < \frac{91-90}{0.5})$ $= 1 - P(Z < 2)$ $= 1 - 0.9772$ $= 0.0228$ stand. aort 0.0228	M1A1 A1 (3) Total 6
3 (a)	$H_0: \mu_A = \mu_B, H_1: \mu_A \neq \mu_B$ μ_1, μ_2 OK both $se = \sqrt{\frac{47^2}{70} + \frac{23^2}{90}} (= \sqrt{37.43492...})$ Test statistic is $\pm \frac{198-201}{se} = \pm 0.4903$ aort 0.99 M1A1 B1 Probab aort 0.312 B1 Probab cv 0.025 $cv = (\pm) 1.96$ Insufficient evidence to reject H_0 , no significant difference between the mean cholesterol content of the two samples. (require correct comparison for FT) correct required.	B1 M1A1 M1A1 B1 A1 ✓ (7)
(b)	<ul style="list-style-type: none"> - require 1 egg from each of 70 chickens of diet A to ensure <u>independence</u>, similarly for diet B. - no chickens in common between the two samples to ensure <u>independence</u> - not same chickens on diet A and diet B because if it were we need to do a <u>paired analysis</u>. Any 1	B1, B1 (2)
not necessary to have a list of parameters to be used		TOTAL 9

4.

Rank:

Shop	Distance	Price	d	d ²
A	1	9	8	64
B	2	7	5	25
C	3	10	7	49
D	4	6	2	4
E	5	4	1	1
F	6	8	2	4
G	7	2	5	25
H	8	1	7	49
I	9	5	4	16
J	10	3	7	49

Reverse ranking on price, $\sum d^2 = 44$
Hairs

$$(a) \quad r_s = 1 - \frac{6 \times 286}{10(10-1)} = -0.73 \text{ or } \frac{-11}{15} \text{ or } -0.733$$

ranking MI
MI, AI
MI, AI
(5)
or 0.733 for $\sum d^2 = 44$

$$(b) \quad H_0: \rho = 0$$

$$H_1: \rho < 0 \quad (H_1: \rho > 0 \text{ if reverse ranking})$$

$$cv = -0.5636 \quad (0.5636)$$

Reject H_0 , evidence there is a significant negative correlation between the price of an ice cream and the distance from a tourist attraction.

(Ice cream gets cheaper further from the tourist attraction)

(-cv from correct table required) (positive in context)

BI
BI
BI
BI
(4)
TOTAL 9

5.

$M =$ wt of male worker
 $F =$ wt of female worker

$M \sim N(78.5, 12.6^2)$
 $F \sim N(62.0, 9.8^2)$

(a) $W = M_1 + \dots + M_7 + F_1 + \dots + F_8$

$E(W) = 7 \times 78.5 + 8 \times 62.0 = 1045.50$

work 1050 M1A1

$Var(W) = 7 \times 12.6^2 + 8 \times 9.8^2 = 1879.64$

1880 M1A1

(4)

(b) Independent (used in Variance formula)

B1

(1)

(c) $P(W > 1090) = P\left(z > \frac{1090 - 1045.5}{\sqrt{1879.64}}\right)$

M1

$= P(z > 1.03)$

work 1.03

A1

$= 1 - 0.8485$

1 - H1

$= \underline{0.1515}$

A1

(4)

AWAT(0.152)

9

6.	<p>H_0: No association between age and colour (independent)</p> <p>H_1: Association between age and colour (Not independent)</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <thead> <tr> <th>O</th> <th>E</th> <th>$\frac{(O-E)^2}{E}$</th> </tr> </thead> <tbody> <tr> <td>12</td> <td>10.08</td> <td>0.3657...</td> </tr> <tr> <td>6</td> <td>7.92</td> <td>0.4654...</td> </tr> <tr> <td>10</td> <td>9.52</td> <td>0.0242...</td> </tr> <tr> <td>7</td> <td>7.48</td> <td>0.0308...</td> </tr> <tr> <td>6</td> <td>8.4</td> <td>0.6857...</td> </tr> <tr> <td>9</td> <td>6.6</td> <td>0.8727...</td> </tr> </tbody> </table> <p>$\sum \frac{(O-E)^2}{E} = 2.4446...$</p> <p>$\nu = (3-1)(2-1) = 2, \chi^2 = 5.991$</p> <p>Insufficient evidence to reject H_0.</p> <p>No association between age and colour (cv for correct h/c for ft)</p>	O	E	$\frac{(O-E)^2}{E}$	12	10.08	0.3657...	6	7.92	0.4654...	10	9.52	0.0242...	7	7.48	0.0308...	6	8.4	0.6857...	9	6.6	0.8727...	<p>BI</p> <p>BI</p> <p>MIAI</p> <p>MIAI</p> <p>MIAI</p> <p>BI BI√</p> <p>AI√ (ii)</p> <p>TOTAL 11</p>
O	E	$\frac{(O-E)^2}{E}$																					
12	10.08	0.3657...																					
6	7.92	0.4654...																					
10	9.52	0.0242...																					
7	7.48	0.0308...																					
6	8.4	0.6857...																					
9	6.6	0.8727...																					
7.(a)	<p>$\bar{x} = \frac{500}{10} = 50$</p> <p>$s^2 = \frac{1}{9} (25001.74 - \frac{500^2}{10}) = 0.193$</p> <p>limits are $50 \pm 1.96s$ $= (49.02, 50.98)$</p> <p>Confidence interval is $(50 - 2.5758 \times \frac{0.5}{\sqrt{10}}, 50 + 2.5758 \times \frac{0.5}{\sqrt{10}})$ $= (49.59273, 50.40727..)$</p> <p>use of estimate in (a) in (b) and (c) assume MISREAD.</p>	<p>MIAI</p> <p>MIAIAI (5)</p> <p>MIBI</p> <p>AI AI (4)</p> <p>MIBIAI</p> <p>AI AI (5)</p> <p>TOTAL 14</p>																					

8(a)

$B_2(5, 0.5)$

MIAI
(2)

(b) $H_0: B(5, 0.5)$ is a suitable model (good fit)
 $H_1: B(5, 0.5)$ is not a suitable model (not a good fit)
 \checkmark for $\hat{p} = 0.466$.

BIV

No. of heads	0	1	2	3	4	5	100% (100% for bin, 1 correct = AI, All correct = AI, 3st or better)
Expected	3.125	15.625	31.25	31.25	15.625	3.125	
Actual	6	18	29	34	10	3	

MIAIAI

	0	E	$\frac{(O-E)^2}{E}$
0 or 1	24	18.75	1.47
2	29	31.25	0.162
3	34	31.25	0.242
4 or 5	13	18.75	1.763

grouped O and E
All correct 2st or better.

MIAI

$$\sum \frac{(O-E)^2}{E} = 3.6373$$

Σ required, correct 3.64

MIAI

$$\nu = 4 - 1 = 3, \chi^2_{0.10}(3) = 6.251$$

BIBIV

Insufficient evidence to reject H_0
 $B(5, 0.5)$ is a suitable model.

No evidence that coins are biased

AI

(11)

Ungrouped gives a χ^2 of 5.44, $\nu = 5, \chi^2_{0.10}(5) = 9.236$
~~for comparison~~

TOTAL 13