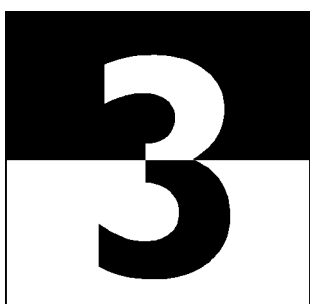


Business Statistics



Level 3

Series 2 2004

(Code 3009)

Model Answers

Business Statistics Level 3

Series 2 2004

How to use this booklet

Model Answers have been developed by LCCIEB to offer additional information and guidance to Centres, teachers and candidates as they prepare for LCCIEB examinations. The contents of this booklet are divided into 3 elements:

- (1) Questions – reproduced from the printed examination paper
- (2) Model Answers – summary of the main points that the Chief Examiner expected to see in the answers to each question in the examination paper, plus a fully worked example or sample answer (where applicable)
- (3) Helpful Hints – where appropriate, additional guidance relating to individual questions or to examination technique

Teachers and candidates should find this booklet an invaluable teaching tool and an aid to success.

The London Chamber of Commerce and Industry Examinations Board provides Model Answers to help candidates gain a general understanding of the standard required. The Board accepts that candidates may offer other answers that could be equally valid.

© Education Development International plc 2004

All rights reserved; no part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise without prior written permission of the Publisher. The book may not be lent, resold, hired out or otherwise disposed of by way of trade in any form of binding or cover, other than that in which it is published, without the prior consent of the Publisher.

Business Statistics Level 3
Series 2 2004

QUESTION 1

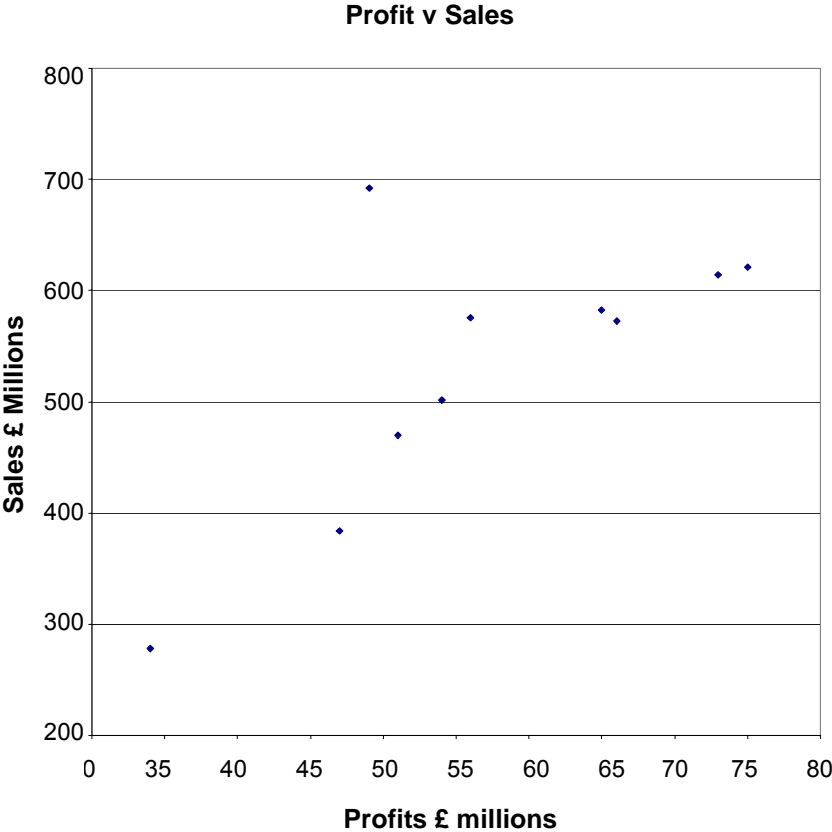
The sales and profits of 10 companies were:

Profits £ millions	47	56	49	51	75	66	34	54	73	65
Sales £ millions	384	576	692	470	621	573	278	501	614	583

- (a) Plot the data on a scatter diagram. (3 marks)
 - (b) Calculate the product-moment correlation coefficient and comment on your answer. (11 marks)
 - (c) Test whether the correlation coefficient differs significantly from zero. (6 marks)
- (Total 20 marks)**

Model Answer to Question 1

(a)



Model Answer to Question 1 continued

(b)

Profits £ millions	Sales £ millions				
x	y	x^2	y^2	xy	$n = 10$
570	5292	33954	2938136	311798	

$$r = \frac{n\sum xy - \sum x \cdot \sum y}{\sqrt{(n(\sum x^2) - (\sum x)^2)(n(\sum y^2) - (\sum y)^2)}}$$

$$\frac{10 \times 311,798 - 570 \times 5,292}{\sqrt{10 \times 33,954 - (570)^2 \times (10 \times 2,938,136 - (5,292)^2)}}$$

$$\frac{101,540}{\sqrt{14640 \times 1376096}}$$

$$\frac{101540.0}{141936.77}$$

$r = 0.72 \quad (0.715389)$

The correlation is positive and fairly strong

(c) Null hypothesis $\rho=0$
Alt hypothesis $\rho \neq 0$

$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$	Degrees Of freedom	$n-2 = 8$
	Crit 't' =	2.31/3.36
$\frac{2.03647}{0.6939}$		
2.93		

There is no reason to reject the null hypothesis at 1% but can reject at 5%.
The correlation coefficient does not differ significantly from 0 (at the 1% significance level).

QUESTION 2

A company classifies its customers as small, medium and large. Data on the length of time a company takes to pay its invoices were collected from a random sample of 320 customers in 2003.

		Company size			Totals
		Small	Medium	Large	
Payment time (days)	Less than 30	30	54	16	100
	30 and less than 50	30	102	28	160
	More than 50	20	24	16	60
		80	180	60	320

- (a) Examine whether a relationship exists between company size and payment time. (12 marks)
- (b) In 2000 36% of payment times were below 30 days, 46% of payment times were between 30 days and 50 days, and 18% of payment times were more than 50 days.

Test whether there is any evidence that payment times in 2003 are significantly different from those in 2000.

(8 marks)

(Total 20 marks)

Model Answer to Question 2

- (a) Null hypothesis: There is no association between size and payment time.
 Alt hypothesis: There is association between size and payment time.

Degrees of freedom (R-1)(C-1) = (3-1) x (3-1) = 2 x 2 = 4
 $\chi^2 = 9.49/13.28$

		Company size			Totals
		Small	Medium	Large	
Payment time (days)	Less than 30	30	54	16	100
	30 and less than 50	30	102	28	160
	More than 50	20	24	16	60
		80	180	60	320

Expected	25	56.25	18.75
Freq	40	90	30
	15	33.75	11.25

Contribution to chi	1	0.09	0.4033
	2.5	1.6	0.1333
	1.6666	2.8166	2.0055

12.216

Conclusions: There is no reason to reject the null hypothesis at 1%. There is no association between company size and payment time. There is reason to reject the null hypothesis at 5%. There is some evidence of association between size and payment time.

- (b) Null hypothesis: There is no change in the payment time distribution.
 Alternative hypothesis: There is a change in the payment time distribution.

Degrees of freedom (n-1) = (3-1) = 2
 Crit $\chi^2 = 5.99/9.21$

	Less than 30	30 and less than 50	50 and over
Expected	36%	46%	18%
	115.2	147.2	57.6
Differences	100	160	60
	15.2	12.8	2.4
Contribution	2.01	1.11	0.10
$\chi^2 =$	3.22		

Conclusion: Chi-sq is less than the critical value at the 5% level. The payment times have not changed.

QUESTION 3

- (a) Explain the factors which should be taken into account when constructing a general price index such as the Retail Price Index.

(10 marks)

In December 2003 the retail price index for a country was as follows (Jan 1987 = 100).

Year	Index
1995	151.9
1996	159.7
1997	165.6
1998	170.1
1999	175.8
2000	185.4
2001	199.8
2002	205.9
2003	212.3

- (b) Re-base the index with 1995 =100.

(4 marks)

- (c) An asset was bought in 1998 for £17,800. If its value has risen in line with inflation, by how much will its value have increased over the 5-year period 1998 to 2003?

(6 marks)

(Total 20 marks)

Model Answer to Question 3

(a) Five factors need to be considered: choice of:

ITEMS – which items should be chosen as typical eg foodstuffs, consumer durables, services etc

PRICES – how should the prices be collected: the most popular price or an average price based on purchases at different prices

WEIGHTS – the relative importance of the goods and services in the total spending pattern based on a survey of consumers spending.

TYPE OF INDEX – is this to be a base year, current year price index or changing cost of a given set of goods and services

BASE YEAR – the base year should be one that is typical ie not one that would exaggerate or underestimate changes in the index

(b)

Year	Retail price index 1987 =100	Retail index re-based 1995 =100
1995	151.9	100.0
1996	159.7	105.1
1997	165.6	109.0
1998	170.1	112.0
1999	175.8	115.7
2000	185.4	122.1
2001	199.8	131.5
2002	205.9	135.5
2003	212.3	139.8

(c) Change in index = $\frac{139.8 - 112}{112} = \frac{27.8}{112} \times 100\%$

$$= 24.8\% \times \text{£}17,800 = \text{£}4,418.2$$

or $\frac{212.3 - 170.1}{170.1} = \frac{42.2}{170.1} \times 100\% = 24.809\% \times \text{£}17,800 = \text{£}4,416$

QUESTION 4

(a) State four uses of the normal distribution?

(4 marks)

(b) A production process takes place in three stages: forming, manufacturing and testing.

The initial forming has a mean time of 7 minutes with a standard deviation of 3 minutes. Manufacturing has a mean time of 15 minutes with a standard deviation of 5 minutes and testing has a mean time of 8 minutes with a standard deviation of 3 minutes.

The times for each process are independent and normally distributed.

(i) Find the probability that an item takes less than 11 minutes to be formed.

(ii) Find the probability that manufacturing an item takes more than 18 minutes.

(iii) Find the probability that testing an item takes between 7 minutes and 13 minutes.

(9 marks)

(c) What is the probability that the total production time for the three processes (forming, manufacturing and testing) takes less than 36 minutes?

(7 marks)

(Total 20 marks)

Model Answer to Question 4

(a) Quality control charts, hypothesis testing, confidence interval estimation, probability estimation

$$(b) (i) z = \frac{\bar{x} - \mu}{\sigma} = \frac{11 - 7}{3} = \frac{4}{3} = 1.33$$

$$p = 0.908$$

$$(ii) z = \frac{18 - 15}{5} = \frac{3}{5} = 0.6$$

$$p = 0.726 \text{ greater than } 18 \text{ mins } 1 - 0.726 = 0.274$$

$$(iii) z = \frac{7 - 8}{3} = \frac{-1}{3} = -0.33$$

$$p = 0.616, 7 \text{ to } 8 \text{ mins} = 0.616 - 0.5 = 0.116$$

$$\frac{13 - 8}{3} = \frac{-5}{3} = 1.67$$

$$p = 0.955 \text{ 8 to } 13 \text{ mins} = 0.955 - 0.5 = 0.455 (z = 1.7)$$

$$P \text{ 7 to } 13 \text{ mins} = 0.116 + 0.455 = 0.571 (0.561, z = 1.6)$$

$$\text{joint mean} = 7 + 15 + 8 = 30$$

$$\begin{aligned} \text{joint sd} &= \sqrt{sd_1^2 + sd_2^2 + sd_3^2} \\ &= \sqrt{3^2 + 5^2 + 3^2} = \sqrt{43} = 6.56 (6.6) \end{aligned}$$

$$(c) \frac{36 - 30}{6.6} = 6/6.6 = 0.91 \quad \text{accept } 0.9$$

$$p = 0.816, 1 - 0.964 = 0.036 \quad (0.971/0.029)$$

QUESTION 5

(a) Explain the difference between a one tail and a two tail test. (4 marks)

(b) Ten male members of staff and ten female members were chosen at random and their output was recorded.

Males	145	168	239	362	222	165	108	96	113	129
Females	164	156	196	173	256	179	198	118	219	247

Test whether there is there any evidence to suggest that the output of males and females is different?

(12 marks)

(c) What is meant by a Type II error? From your result in (b) above could a Type II error have occurred?

(4 marks)

(Total 20 marks)

Model Answer to Question 5

- (a) A one tailed test is when the alternative hypothesis states that the difference is smaller, greater ie there is one direction to the difference. A two tailed test tests to see if there is a change which may be in either direction. Can be both greater or smaller.
- (b) Null hypothesis: The output of male and female workers does not differ.

Alt hypothesis: The output of males and females does differ.

$$\Sigma x_1 \ 1747 \qquad \Sigma x_2 \ 1906$$

$$\bar{x}_1 \ 174.7 \qquad \bar{x}_2 \ 190.6$$

$$SD_1 = 80.97469 \qquad SD_2 = 2.12204$$

$$\text{Pooled SD} = \sqrt{((n_1-1) \times SD_1^2 + (n_1-1) \times SD_2^2) / (n_1 + n_2 - 2)}$$

$$= 64.54$$

$$\begin{array}{ll} \text{df} & 18 \\ \text{t Critical} & 2.10 \end{array}$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$t = \frac{174.7 - 190.6}{64.54 \sqrt{(1/10 + 1/10)}}$$

$$\frac{-15.9}{64.54 \sqrt{0.2}}$$

$$\frac{-15.9}{65.54 \times 0.4472}$$

$$\frac{-15.9}{29.31} = -0.54$$

Conclusion: Accept the null hypothesis. There is no difference in the output of males and females.

- (c) Type II error is accepting a null hypothesis that should be rejected.

The result may be a Type II error as the null hypothesis was accepted.

QUESTION 6

- (a) (i) Explain how a stratified random sample would be taken and give one advantage and one disadvantage of the method.
- (ii) Explain how a quota sample would be taken and give one advantage and one disadvantage of the method. (8 marks)
- (b) Explain the term **sampling distribution of the mean**. (4 marks)
- (c) The data below were obtained from 2 random samples of sales on two different days:

	Monday	Wednesday
Mean daily sales	4,860	4,720
Standard deviation	312	290
Sample size	210	170

Test whether there is a significant difference between the mean daily sales on the two days. (8 marks)

(Total 20 marks)

Model Answer to Question 6

- (a) Explain how a stratified random sample would be taken by allotting all the relevant members of the sampling frame to different strata depending on characteristics. Selection of a random sample in each strata.

Advantages: The sampling error can be calculated, the result can be used in further statistical calculation and it reduces the sampling error. Disadvantages: A sampling frame is necessary, the relevant characteristics need to be identified, the sample may be drawn from a widely dispersed geographical area.

Quota sampling involves taking a sample with a given number of people. Often this is done by stopping people in the street. Interviewers are given quota controls ie the characteristics that the respondents should have. These often relate to age, gender and income. The advantages are that the sample is relatively cheap, there is no need for a sampling frame and no need for call-backs. The disadvantages are that there may be bias in the choice of respondents. Theoretically the standard error cannot be calculated so significance tests are not valid.

- (b) The sampling distribution of the means refers to the variability which occurs in the value of the means as different samples of the same size are taken. The standard error of the means equals σ/\sqrt{n} .
- (c) Null hypothesis: There is no difference in the mean daily sales between Monday and Wednesday. Alternative hypothesis: There is a difference in the mean daily sales between Monday and Wednesday.

Critical z value 1.96/2.58

$$z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$
$$\frac{4,860 - 4,720}{\sqrt{\frac{312^2}{210} + \frac{290^2}{170}}}$$
$$\frac{140}{30.96} = 4.52$$

Reject the null hypothesis. There is evidence to suggest there is a difference between the mean daily output on a Monday and Wednesday.

QUESTION 7

(a) Describe the four components in a time series model. (4 marks)

(b) A company is investigating changes in the pattern, over time, of days lost from work through absenteeism. Using an appropriate model find the centred trend and quarterly seasonal variations for the number of days of absenteeism from work. (10 marks)

	Days lost from absenteeism			
Year	Quarters			
	1	2	3	4
2001	149	136	123	176
2002	153	140	128	177
2003	145	135	131	183

(c) Predict the number of days lost from absenteeism in the first two quarters of 2004 and comment on the reliability of your forecast. (6 marks)

(Total 20 marks)

Model Answer to Question 7

(a) The 4 components are Trend, Cyclical, Seasonal and Random variation. When a time series follows a regular pattern the data can be decomposed into these 4 elements.

(b)

Year	qtr	absenteeism	Total	m avg1	Trend	Additive differences	Multiplicative differences
2001	1	149					
	2	136					
	3	123	584	146	146.5	-23.5	0.83959
	4	176	588	147	147.5	28.5	1.19322
2002	1	153	592	148	148.625	4.375	1.02944
	2	140	597	149.25	149.375	-9.375	0.93724
	3	128	598	149.5	148.5	-20.5	0.86195
	4	177	590	147.5	146.875	30.125	1.20511
2003	1	145	585	146.25	146.625	-1.625	0.98892
	2	135	588	147	147.75	-12.75	0.91371
	3	131	594	148.5			
	4	183					

	qtr1	2	3	4
Seasonal differences	4.375	-9.375	-23.5	28.5
	-1.625	-12.75	-20.5	30.125
Seas. var	1.375	-11.063	-22	29.3125

	qtr1	2	3	4
Seasonal Differences	1.02944	0.93724	0.86195	1.19322
	0.98892	0.91371		
Seas. var	1.00918	0.92547	0.85077	1.19916

(c)

PREDICTIONS

change in trend $\frac{147.75 - 146.5}{7}$

prediction change	0.178571		
	Est trend.	add var	Multi var
next year	148.2857	150	150
	148.4643	137	137

The prediction may be considered to be reasonably accurate as the period of prediction is only for 2 quarters of the next year.

If total/ma1/ma2 and differences misaligned lose 3 method marks (but if the differences calculated correctly as above give full credit).

QUESTION 8

- (a) State two reasons why a business might use statistical quality control.

(4 marks)

A company produces bags of maize set to a mean of 50 kilograms with a standard deviation of 2.5 kilograms. Random samples of 8 bags are taken at regular intervals to ensure the process is under control. Quality control procedures are used which set the warning limits at the 0.025 probability point and action limits at the 0.001 probability point. This means, for example, that the upper action limit is set so that the probability of the means exceeding the limit is 0.001. You may assume that the weight of maize in bags is normally distributed.

- (b) Calculate the warning and action limits and construct the quality control chart.

(8 marks)

Sample	1	2	3	4	5	6
Weight kg	52.5	52.4	49.4	51.6	47.1	50.5

- (c) Plot the results for the last **six** samples and comment on the results.

(4 marks)

- (d) If the mean weight was wrongly set at 50.6 kilograms (the standard deviation remaining at 2.5 kilograms) what would be the probability of the mean of a random sample of eight items exceeding the upper action limit?

(4 marks)

(Total 20 marks)

Mean $\bar{x} = \frac{\sum fx}{\sum f}$

Standard deviation $s = \sqrt{\frac{\sum fx^2}{\sum f} - \left(\frac{\sum fx}{\sum f}\right)^2}$

Pearson measure of skewness $\frac{3(\bar{x} - \text{Median})}{s}$

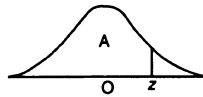
Coefficient of variation $\frac{s}{\bar{x}} \times 100$

Product moment correlation coefficient $r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$

Spearman's rank correlation coefficient $r_s = 1 - \frac{6\sum d^2}{n(n^2 - 1)}$

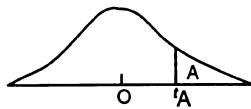
Least squares regression line $\hat{y} = a + bx$
 $b = \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2}$
 $a = \frac{\sum y}{n} - \frac{b\sum x}{n}$

TABLE 1 – NORMAL DISTRIBUTION
A is the area to the left of the given value of z



z	0.0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	
A	.500	.540	.580	.618	.655	.692	.726	.758	.788	.816	
z	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	
A	.841	.864	.885	.903	.919	.933	.945	.955	.964	.971	
z	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0
A	.977	.982	.986	.989	.992	.994	.995	.996	.997	.998	.999

TABLE 2 – t DISTRIBUTION
t_A is the value of the t statistic with ν degrees of freedom with area A to the right of it



ν	1	2	3	4	5	6	7	8
t _{0.05}	6.31	2.92	2.35	2.13	2.02	1.94	1.90	1.86
t _{0.025}	12.71	4.30	3.18	2.78	2.57	2.45	2.37	2.31
t _{0.01}	31.82	6.97	4.54	3.75	3.37	3.14	3.00	2.90
t _{0.005}	63.66	9.93	5.84	4.60	4.03	3.71	3.50	3.36
ν	9	10	11	12	13	14	15	16
t _{0.05}	1.83	1.81	1.80	1.78	1.77	1.76	1.75	1.75
t _{0.025}	2.26	2.23	2.20	2.18	2.16	2.15	2.13	2.12
t _{0.01}	2.82	2.76	2.72	2.68	2.65	2.62	2.60	2.58
t _{0.005}	3.25	3.17	3.11	3.05	3.01	2.98	2.95	2.92
ν	17	18	19	20	21	22	23	24
t _{0.05}	1.74	1.73	1.73	1.73	1.72	1.72	1.71	1.71
t _{0.025}	2.11	2.10	2.09	2.09	2.08	2.07	2.07	2.06
t _{0.01}	2.57	2.55	2.54	2.53	2.52	2.51	2.50	2.49
t _{0.005}	2.90	2.88	2.86	2.85	2.83	2.82	2.81	2.80

One sample z test

Mean $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ Proportion $z = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$

Two sample z test

Mean $z = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ Proportion $z = \frac{p_1 - p_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$

where $p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2}$
Price Quantity

Laspeyres index $\frac{\sum p_1 q_0}{\sum p_0 q_0} \times 100$ $\frac{\sum p_0 q_1}{\sum p_0 q_0} \times 100$

Paasche index $\frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$ $\frac{\sum p_1 q_1}{\sum p_0 q_1} \times 100$

Weighted index $\frac{\sum WI}{\sum W}$

One sample t test

$t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ where $s = \sqrt{\frac{\sum(x - \bar{x})^2}{n - 1}}$

Independent samples t test

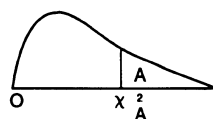
$t = \frac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{n} + \frac{1}{m}}}$ where $s = \sqrt{\frac{\sum(x - \bar{x})^2 + \sum(y - \bar{y})^2}{n + m - 2}}$

Chi-square test $\chi^2 = \sum \frac{(O - E)^2}{E}$

Test for $\rho = 0$ $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$

TABLE 3 – CHI SQUARED DISTRIBUTION

χ²_A is the value of the χ² statistic with ν degrees of freedom with area A to the right of it



ν	1	2	3	4	5	6
χ ² _{0.05}	3.84	5.99	7.81	9.49	11.07	12.59
χ ² _{0.01}	6.63	9.21	11.34	13.28	15.09	16.81
ν	7	8	9	10	11	12
χ ² _{0.05}	14.07	15.51	16.92	18.31	19.68	21.03
χ ² _{0.01}	18.48	20.09	21.67	23.21	24.73	26.22



**EXAMINATIONS
BOARD**

London Chamber of Commerce and Industry Examinations Board

Athena House 112 Station Road

Sidcup Kent DA15 7BJ United Kingdom

Facsimile: +44 (0) 20 8302 4169/+44 (0) 20 8309 5169

Website: www.lccieb.com

Customer Service Team Tel: +44 (0) 8707 202 909 email: custserv@lccieb.org.uk

© Education Development International plc 2004