Pearson LCCI

Certificate in Business Statistics (VRQ)
Level 3

Monday 5 September 2016
Time: 3 hours

Paper Reference
ASE20100

Complete the details below in block capitals.

Candidate name

Centre Code

Candidate Number

Candidate ID Number

You must have:
Statistical formulae sheet (enclosed), HB pencil, eraser

Total Marks

Instructions

- Use **black** ink or ball-point pen
  – pencil can only be used for graphs, charts, diagrams, etc.
- **Fill in the boxes** at the top of this page with your name, candidate number, centre code and your candidate ID number.
- Answer **all** questions.
- Answer the questions in the spaces provided
  – there may be more space than you need.
- Answers should be given to an appropriate degree of accuracy.

Information

- The total mark for this paper is 100.
- The marks for **each** question are shown in brackets
  – use this as a guide as to how much time to spend on each question.
- A formulae sheet is provided at the front of the question paper.
- Calculators may be used.

Advice

- Read each question carefully before you start to answer it.
- Try to answer every question.
- You are advised to show your workings.
- Check your answers if you have time at the end.

Turn over ▶

P51677A

1/1/1/1/1/1/1

P 5 1 6 7 7 A 0 1 2 4

PEARSON

Pearson LCCI Level 3 Certificate in Business Statistics formulae sheet

Median for grouped data $l_m + \frac{c_m}{f_m}\left(\frac{n}{2} - F_{m-1}\right)$

Where l_m, c_m and f_m are the lower boundary, width and frequency respectively of the median class, n is the total number of observations and F_{m-1} is the cumulative frequency corresponding to l_m.

Mean for ungrouped data $\bar{x} = \frac{\Sigma x}{n}$

Mean for grouped data $\bar{x} = \frac{\Sigma fx}{\Sigma f}$

Standard deviation for ungrouped data $s = \sqrt{\frac{\Sigma x^2}{n} - (\bar{x})^2}$

Standard deviation for grouped data $s = \sqrt{\frac{\Sigma fx^2}{\Sigma f} - \left(\frac{\Sigma fx}{\Sigma f}\right)^2}$

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

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

Multiplication rule of probability $P(A \cap B) = P(A) \times P(B)$ if A and B independent

Addition rule of probability $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

	Price	Quantity
Laspeyres index	$\frac{\Sigma p_1 q_0}{\Sigma p_0 q_0} \times 100$	$\frac{\Sigma p_0 q_1}{\Sigma p_0 q_0} \times 100$
Paasche index	$\frac{\Sigma p_1 q_1}{\Sigma p_0 q_1} \times 100$	$\frac{\Sigma p_1 q_1}{\Sigma p_1 q_0} \times 100$
Weighted index	$\frac{\Sigma WI}{\Sigma W}$	

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

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

Least squares regression line $\hat{y} = a + bx$

$$a = \frac{\sum y}{n} - \frac{b\sum x}{n}$$

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

One sample z test

Mean $z = \frac{\bar{x}-\mu}{\frac{\sigma}{\sqrt{n}}}$ 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 $p = \frac{n_1p_1+n_2p_2}{n_1+n_2}$

One sample t test

$t = \frac{\bar{x}-\mu}{\frac{s}{\sqrt{n}}}$ 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 squared test $x^2 = \sum\frac{(O-E)^2}{E}$

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

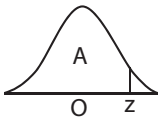
DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

Table 1: Normal distribution

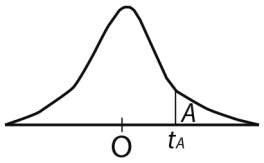
A is the area to the left of the given value of z



z	A	z	A	z	A	z	A	z	A
0.00	0.5000	0.50	0.6915	1.00	0.8413	1.50	0.9332	2.00	0.9772
0.01	0.5040	0.51	0.6950	1.01	0.8438	1.51	0.9345	2.02	0.9783
0.02	0.5080	0.52	0.6985	1.02	0.8461	1.52	0.9357	2.04	0.9793
0.03	0.5120	0.53	0.7019	1.03	0.8485	1.53	0.9370	2.06	0.9803
0.04	0.5160	0.54	0.7054	1.04	0.8508	1.54	0.9382	2.08	0.9812
0.05	0.5199	0.55	0.7088	1.05	0.8531	1.55	0.9394	2.10	0.9821
0.06	0.5239	0.56	0.7123	1.06	0.8554	1.56	0.9406	2.12	0.9830
0.07	0.5279	0.57	0.7157	1.07	0.8577	1.57	0.9418	2.14	0.9838
0.08	0.5319	0.58	0.7190	1.08	0.8599	1.58	0.9429	2.16	0.9846
0.09	0.5359	0.59	0.7224	1.09	0.8621	1.59	0.9441	2.18	0.9854
0.10	0.5398	0.60	0.7257	1.10	0.8643	1.60	0.9452	2.20	0.9861
0.11	0.5438	0.61	0.7291	1.11	0.8665	1.61	0.9463	2.22	0.9868
0.12	0.5478	0.62	0.7324	1.12	0.8686	1.62	0.9474	2.24	0.9875
0.13	0.5517	0.63	0.7357	1.13	0.8708	1.63	0.9484	2.26	0.9881
0.14	0.5557	0.64	0.7389	1.14	0.8729	1.64	0.9495	2.28	0.9887
0.15	0.5596	0.65	0.7422	1.15	0.8749	1.65	0.9505	2.30	0.9893
0.16	0.5636	0.66	0.7454	1.16	0.8770	1.66	0.9515	2.32	0.9898
0.17	0.5675	0.67	0.7486	1.17	0.8790	1.67	0.9525	2.34	0.9904
0.18	0.5714	0.68	0.7517	1.18	0.8810	1.68	0.9535	2.36	0.9909
0.19	0.5753	0.69	0.7549	1.19	0.8830	1.69	0.9545	2.38	0.9913
0.20	0.5793	0.70	0.7580	1.20	0.8849	1.70	0.9554	2.40	0.9918
0.21	0.5832	0.71	0.7611	1.21	0.8869	1.71	0.9564	2.42	0.9922
0.22	0.5871	0.72	0.7642	1.22	0.8888	1.72	0.9573	2.44	0.9927
0.23	0.5910	0.73	0.7673	1.23	0.8907	1.73	0.9582	2.46	0.9931
0.24	0.5948	0.74	0.7704	1.24	0.8925	1.74	0.9591	2.48	0.9934
0.25	0.5987	0.75	0.7734	1.25	0.8944	1.75	0.9599	2.50	0.9938
0.26	0.6026	0.76	0.7764	1.26	0.8962	1.76	0.9608	2.55	0.9946
0.27	0.6064	0.77	0.7794	1.27	0.8980	1.77	0.9616	2.60	0.9953
0.28	0.6103	0.78	0.7823	1.28	0.8997	1.78	0.9625	2.65	0.9960
0.29	0.6141	0.79	0.7852	1.29	0.9015	1.79	0.9633	2.70	0.9965
0.30	0.6179	0.80	0.7881	1.30	0.9032	1.80	0.9641	2.75	0.9970
0.31	0.6217	0.81	0.7910	1.31	0.9049	1.81	0.9649	2.80	0.9974
0.32	0.6255	0.82	0.7939	1.32	0.9066	1.82	0.9656	2.85	0.9978
0.33	0.6293	0.83	0.7967	1.33	0.9082	1.83	0.9664	2.90	0.9981
0.34	0.6331	0.84	0.7995	1.34	0.9099	1.84	0.9671	2.95	0.9984
0.35	0.6368	0.85	0.8023	1.35	0.9115	1.85	0.9678	3.00	0.9987
0.36	0.6406	0.86	0.8051	1.36	0.9131	1.86	0.9686	3.05	0.9989
0.37	0.6443	0.87	0.8078	1.37	0.9147	1.87	0.9693	3.10	0.9990
0.38	0.6480	0.88	0.8106	1.38	0.9162	1.88	0.9699	3.15	0.9992
0.39	0.6517	0.89	0.8133	1.39	0.9177	1.89	0.9706	3.20	0.9993
0.40	0.6554	0.90	0.8159	1.40	0.9192	1.90	0.9713	3.25	0.9994
0.41	0.6591	0.91	0.8186	1.41	0.9207	1.91	0.9719	3.30	0.9995
0.42	0.6628	0.92	0.8212	1.42	0.9222	1.92	0.9726	3.35	0.9996
0.43	0.6664	0.93	0.8238	1.43	0.9236	1.93	0.9732	3.40	0.9997
0.44	0.6700	0.94	0.8264	1.44	0.9251	1.94	0.9738	3.50	0.9998
0.45	0.6736	0.95	0.8289	1.45	0.9265	1.95	0.9744	3.60	0.9998
0.46	0.6772	0.96	0.8315	1.46	0.9279	1.96	0.9750	3.70	0.9999
0.47	0.6808	0.97	0.8340	1.47	0.9292	1.97	0.9756	3.80	0.9999
0.48	0.6844	0.98	0.8365	1.48	0.9306	1.98	0.9761	3.90	1.0000
0.49	0.6879	0.99	0.8389	1.49	0.9319	1.99	0.9767	4.00	1.0000
0.50	0.6915	1.00	0.8413	1.50	0.9332	2.00	0.9772		

Table 2: t distribution

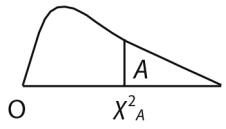
t_A is the value of the t statistic with v degrees of freedom with area A to the right of it



v	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
v	**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
v	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**
$t_{0.05}$	1.74	1.73	1.73	1.73	1.73	1.72	1.71	1.71
$t_{0.025}$	2.11	2.10	2.09	2.09	2.09	2.08	2.07	2.06
$t_{0.01}$	2.57	2.55	2.54	2.54	2.53	2.52	2.50	2.49
$t_{0.005}$	2.90	2.88	2.86	2.86	2.85	2.83	2.81	2.80

Table 3: Chi squared distribution table

χ^2_A is the value of the χ^2 statistic with v degrees of freedom with area A to the right of it



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

Answer ALL questions. Write your answers in the spaces provided.

1 Sport Mag plc publishes programmes for sports matches and has a number of sports teams as customers. Before each print run the production press must first be cleaned and then set up for production. Records show that, in each case, the time taken to clean the production press and the set up time are independently normally distributed.

The following table summarises these times.

	Mean	**Standard deviation**
Cleaning time (minutes)	15	3
Set up time (minutes)	20	4

(a) Calculate, for a randomly selected print run, the probability that

(i) the cleaning time is less than 12 minutes

(3)

(ii) the set up time is more than 18 minutes

(3)

(iii) the total time required for cleaning and set up is greater than 45 minutes.

(5)

To avoid future delays, the production manager requires that no more than 5% of the set up times for a print run will be outside the range 15 to 25 minutes.

(b) Calculate the maximum value of the standard deviation that will meet this requirement, assuming that the value of the mean remains at 20 minutes.

(4)

As the number of customers has fallen recently, the management at Sport Mag plc agrees to conduct a customer satisfaction survey. Records show that 35% of customers are small-sized companies, 51% are medium-sized companies, and the rest are large-sized companies.

(c) (i) Describe, using the breakdown of company size, how you would obtain a stratified sample of 80 customers.

(3)

(ii) Give **one** advantage and **one** disadvantage of using a stratified sample compared to a simple random sample.

(2)

Advantage .....

Disadvantage .....

(Total for Question 1 = 20 marks)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

2 Apollo Bank wishes to determine the average amount of time that cashiers spend serving each customer. It obtains the following data from 100 customers:

- the sample mean time taken is 7.4 minutes
- the sample standard deviation of the time taken is 4.2 minutes.

(a) (i) Calculate the value of the estimated standard error of the sample mean.

(2)

(ii) Calculate a 99% confidence interval for the population mean time that cashiers take to serve a customer.

(3)

(iii) Describe, on the basis of the confidence interval from (a) (ii), what would have been the outcome of a test that the population mean time cashiers spend serving customers is 10 minutes.

(2)

(iv) Give a reason for your answer in (a) (iii).

(1)

The Financial Controller at Apollo Bank is concerned about the number of borrowers failing to make loan repayments on time. She suspects that debtors who owe more to the bank are taking longer to settle. A sample of 500 debts is produced and summarised in the table of observed frequencies below.

Observed frequencies		**Time to settle**		
		Less than 1 month	Between 1 and 6 months	Over 6 months
Amount of debt	Small (below $5000)	120	80	50
	Medium (between $5000 and $10 000)	60	60	30
	Large (over $10 000)	20	40	40

For this contingency table the expected frequencies have been calculated and shown in the following table.

Expected frequencies		**Time to settle**		
		Less than 1 month	Between 1 and 6 months	Over 6 months
Amount of debt	Small (below $5000)	100	90	60
	Medium (between $5000 and $10 000)	60	54	36
	Large (over $10 000)	40	36	24

(b) (i) Test, stating both hypotheses and using the information from both tables, whether there is a significant association between the amount of debt and the time taken to settle.

(7)

(ii) State, in context, what the test in (b) (i) shows.

(1)

(c) Calculate, using the table of **observed frequencies**, the probability that

(i) a randomly chosen borrower with a small debt takes over 6 months to settle

(1)

(ii) a randomly chosen borrower with a large debt takes over 6 months to settle.

(1)

(d) Explain how the two tables, and the probabilities calculated in (c), support the Financial Controller's suspicions that customers with large debts take longer to settle.

(2)

.....

.....

.....

.....

(Total for Question 2 = 20 marks)

3 The charity HELP-US is a direct mail fundraiser that collects funds by sending mailings directly to potential donors. Data is available showing the donation from the latest campaign ($*y*) and the donation from a previous campaign ($*x*) for a sample of 20 donors.

(a) State why 'donation from the latest campaign' is identified as the response variable and the 'donation from the previous campaign' as the explanatory variable.

(1)

An analysis of the data has been carried out using a statistical software package. The output from the analysis is given below.

Slope coefficient	1.25
Intercept coefficient	2.6
Product moment correlation coefficient	0.32

(b) (i) Explain, in the context of the information provided, what the slope and intercept coefficients represent.

(3)

Slope

Intercept

(ii) Estimate the donation in the latest campaign for a donor who gave $25 in the previous campaign.

(2)

(iii) Give a reason why the estimate in (b) (ii) may not be reliable.

(1)

(c) (i) Test, at the 1% significance level, whether the product moment correlation coefficient is significantly different from zero. State your hypotheses clearly.

(5)

(ii) Interpret, in context, what this test shows.

(1)

(iii) Explain why your interpretation in (c) (ii) would not change if the test had been carried out at the 10% level.

(2)

(d) Give **one** advantage of using a 1% level of significance rather than a 10% level in a hypothesis test.

(1)

.....

.....

The data for the 20 pairs of donations for the previous and latest campaigns were ranked. A spreadsheet analysis of the ranked data gave a value for $\sum d^2$ of 995

(e) (i) Calculate Spearman's rank correlation coefficient.

(3)

.....

.....

.....

.....

.....

.....

(ii) Explain why the value of Spearman's rank correlation coefficient differs from the value of the product moment correlation coefficient.

(1)

.....

.....

.....

(Total for Question 3 = 20 marks)

4 Steelite Engineering plc manufactures a range of metal components for the aeronautical industry. The weekly sales (in $000) over the last year are shown in the table below.

Weekly sales ($000)	Frequency
20 and less than 30	8
30 and less than 50	15
50 and less than 70	13
70 and less than 100	10
100 and less than 150	6

(a) Calculate estimated values of the:

(i) mean weekly sales

(3)

(ii) standard deviation of weekly sales.

(3)

An estimate of the median is $54,600

(b) (i) Calculate a measure of skewness using this estimate and the estimates in (a).

(2)

(ii) Describe the shape of the distribution of weekly sales.

(1)

The diameter of an engineered component, manufactured by Steelite Engineering plc, is normally distributed with a mean of 50 cm and standard deviation of 1.5 cm. The quality of the manufacturing process is monitored by taking samples of size 9 at hourly intervals.

(c) Calculate, in cm, the upper and lower (95%) warning limits and (99.8%) action limits for sample means from samples of size 9

(3)

Five samples are taken during a shift. The sample means are given in the table below.

Sample number	1	2	3	4	5
Sample mean	50.8	50.4	49.7	50.6	49.5

(d) (i) Comment, without plotting a graph, on the stability of the manufacturing process.

(1)

(ii) State what action, if any, you would recommend about the manufacturing process during this shift.

(1)

During a different shift a new engineering component was manufactured and strength tested. It was found that, out of a random sample of 75 components, 60 passed a strength test, the remainder did not pass the strength test.

(e) Calculate a 95% confidence interval for the proportion of components that pass the strength test.

(5)

The marketing information for Steelite Engineering plc claims that 90% of components pass the strength test.

(f) Explain, using the confidence interval in (e), whether the manufacturer's claim is likely to be correct for this new component.

(1)

.....

.....

.....

(Total for Question 4 = 20 marks)

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

5 Mactar plc assembles computers for a number of education and government organisations. In order to improve the productivity of its employees, the company has decided to introduce a new payment scheme for its assembly workers.

To assess the effectiveness of this scheme a random sample of 8 assembly workers was taken prior to the introduction of the scheme and their productivity measured in a 40-hour working week. A second, different, random sample of 7 assembly workers was taken after the introduction of the new scheme and their productivity measured over an equivalent working week.

Both sets of results are shown in the following table.

Productivity (units assembled)	**Old scheme**	27	32	23	29	33	30	27	31
	New scheme	29	31	31	34	36	30	33	

(a) (i) Explain why a paired t-test is not appropriate to test if there has been an improvement in productivity using these data.

(1)

.....

.....

(ii) Give **one** disadvantage of using an independent sample t-test, rather than a **paired** t-test, to assess the effectiveness of such a scheme.

(1)

.....

.....

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

The productivity results can be summarised in the following table.

	Mean	Standard deviation
Old scheme	29	3.25
New scheme	32	2.45

(b) (i) Carry out an independent sample t-test to establish whether the new payment scheme has resulted in a significant increase in productivity. State your hypotheses clearly.

(7)

(ii) State, in context, what the test in (b) (i) shows.

(1)

An equivalent payment scheme was offered to Mactar plc's sales team. To monitor the effectiveness of this change, a random sample of 10 sales representatives was chosen and their individual sales figures were recorded one month before and one month after the introduction of the new payment scheme.

The differences between the sales figures before and after the introduction of the new payment scheme were analysed using a paired t-test.

(c) (i) State the **two** hypotheses used for this paired t-test.

(1)

The calculated value for the test statistic for this paired t-test is 2.03

(ii) Test, at the 5% significance level, whether the introduction of the new payment scheme has changed the sales figures.

(3)

(iii) State, in context, what the test in (c) (ii) shows.

(1)

The average salary for all employees at Mactar plc, together with a national index of retail prices (based on the year 2005), for the period 2013 to 2015 is given in the table below.

Year	Average salary ($)	index of retail prices (2005 = 100)
2013	32,150	132.0
2014	32,540	137.2
2015	32,720	141.5

(d) (i) Convert the index of retail prices to a series with 2013 as the base year.

(2)

(ii) Calculate the average salary for the years 2013 to 2015 based on the retail prices in 2013.

(2)

(iii) Describe the purchasing power of the average salary over the period 2013 to 2015.

(1)

(Total for Question 5 = 20 marks)

TOTAL FOR PAPER = 100 MARKS

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

DO NOT WRITE IN THIS AREA

BLANK PAGE