Examination papers and Examiners' reports

Management mathematics

2790**076**

2002, 2003, 2004

Undergraduate study in
Economics, Management, Finance and the Social Sciences

LSE
THE LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

UNIVERSITY OF LONDON EXTERNAL PROGRAMME 1858

Examiner's report 2004

Zone A

General remarks

This year's paper was quite standard in its choice of topics and questions – for many of the questions one could trace a history of similar questions through past papers. At the same time, however, you will also find small new parts to most questions.
As there were only a few Zone A candidates, it is hard to say anything specific (certainly hard to say anything statistically significant) about their performance or about which areas they found particularly difficult. However the performance and areas of strength and weakness seemed to be similar on both zone papers. Candidates are therefore urged to read the reports for both Zone A and Zone B.

The paper proved to be of a suitable length and it was good to see that virtually all candidates tackled five full questions. Equally pleasing from the point of view of the Examiners was that all questions received a reasonably fair share of attempts – although Questions 4 and 5 were a little less popular than the others.

Specific comments on questions

Question 1

(a) Some students failed to recognise the specific title of an Ideal Index and spoke too generally about what an index should do, for example, not to overestimate nor to underestimate. This led some candidates into a worthless discussion about Paasche and Laspeyre type indices. Check page 15 in your subject guide for the Factor Reversal and Time Reversal tests that an Ideal Index should pass. Full marks would be earned by stating the requirement of each of these tests and then defining Fischer's Ideal Index.

(b) There will naturally be some numerical errors with so much data to analyse and with so many indices to calculate. Such errors, however, usually lose few marks. More important are the substantial marks lost by not using the correct index. Furthermore it is worth reiterating the need to read given tables of information very carefully before embarking upon 'battering' the calculator. Note, for example, that the *Earnings* column is for average per person. Thus the total earnings paid by the company in a year will be the *Number* of people in each group multiplied by the average *Earnings* per person in that group – summed over the four Ethnic Groups. Other than that it is really a matter of using great care not to lose one's thought processes part way through the calculations. Remember also that all the indices required will have a base year value of 100 (not 1).
Throughout the following outline answers p_{ij} stands for the average *Earnings* in group i for year j and q_{ij} stands for the *Number* of workers in group i for year j.
i. Total earnings for year 1 =

$$\sum_i p_{i1}q_{i1} = (505)(182) + (245)(103) + (908)(7) + (125)(45) = 129{,}126$$

Similarly earnings for year 2 total =

$$\sum_i p_{i2}q_{i2} = (531)(225) + (268)(66) + (873)(9) + (133)(55) = 152{,}335$$

and for year 3 it is 179,649.

Thus the required index values are 100.0 for year 1, $\frac{152,335}{129,126}x100 = 118.0$ for year 2, and 139.1 for year 3.

(b)ii. For this part we have to first work out the overall average earnings per worker, namely for all workers irrespective of their ethnic group. To do this for each year we should divide the total earnings of that year (as calculated in part i.) and divide by the total number of workers. Thus the average earnings per worker for year 1 is $\frac{129,126}{182+103+7+45} = \frac{129,126}{337} = 383.16.$

Similarly the average earnings for year 2 and year 3 are $\frac{152,335}{355} = 429.11$ and

$\frac{179,649}{381} = 471.52$ respectively.

Hence the required index is 100 for year 1, $\frac{429.11}{383.16}x100 = 111.99$ for year 2 and

$\frac{471.52}{383.16}x100 = 123.06$ for year 3.

(b)iii. With this data we are finally (and probably most easily) required to use the total workers values already calculated in part ii. to determine the indices for workers employed:

Again year 1 = 100, whilst year 2 = $\frac{355}{337}x100 = 105.34$ and for year 3 it is

$\frac{381}{337}x100 = 113.06$

[Notice that each of the three indices requested are using a base year 1 as requested. There is no indication that a chain (year on year) index is required.]

(c) All we are given are prices and hence there can be no possibility of using a weighted index for the data given – see part ii. below.

i. If the basic index definitions are known this should be a simple question.

The Simple Aggregate Index is $\frac{\sum_i p_{it}}{\sum_i p_{i0}}$. This gives an index value of

$\frac{105+85+90+80}{90+65+80+80}x100 = 114.29$ in year 2

In a similar fashion we can calculate the index values of 128.57 for year 3 and 133.33 for year 4.

The Average Price Relative Index is defined as $\frac{100}{n}\sum_i \frac{p_{it}}{p_{i0}}$.

For example, in year 2 we get an index value of $\frac{100}{4}\left\{\frac{105}{90}+\frac{85}{65}+\frac{90}{80}+\frac{80}{80}\right\} = 115.00.$

You should be able to check that we similarly get index values of 128.89 in year 3 and 133.26 in year 4.

ii. All that was required here was the recognition that some form of weighting would be useful, for example, the quantity of each newspaper sold in each year. We could then form a number of different weighted indices.

Overall Question 1 is indeed rather long (and perhaps tedious) but should (and often did) earn high marks for the careful candidate who could remember the basic index definitions. Always remember to state the formula you are going to use, show in detail how it works for one of the years and then give summary results for the other year(s).

Question 2

Part i. was generally well done, part ii. tended to earn half marks for most candidates as they did not consider both alternative answers – this was partly expected, part iii. was found to be surprisingly difficult – with many unable to define the given subset correctly – either in words or diagrammatically.

i. There are several ways of demonstrating some inconsistency in the given data – usually this involves showing that the order of some subset(s) must be negative in an attempt to satisfy the given data. Thus, for example, if we start from $n(W \cap S \cap D) = 50$, we find that $n(W \cap S \cap D^c) = 140$ and $n(W \cap S^c \cap D) = 200$. Hence, in order to make $n(W) = 250$ we must have $n(W \cap S^c \cap D^c) = -140$!! , which is obviously impossible. Alternatively we can show that $n(W^c \cap S^c \cap D)$ is a similarly nonsensical minus 110. Clearly there must therefore be an error in the data.

ii. Given that there is only one error in the given data we must find a piece of given data which, if changed, might allow the data to become entirely feasible. As demonstrated in i. above there are two subsets whose order needs to be negative if the data is to be consistent. Hence the piece of given data that needs to be changed will have to change **both** of these subset orders. With some careful thought you will find that the only possibilities are $n(W \cap S \cap D)$ and $n(W \cap D)$. Note that $n(W \cap S^c \cap D)$ was not a given datum and hence cannot be a candidate for being erroneous!! The candidate should therefore consider $n(W \cap S \cap D)$ and $n(W \cap D)$ in turn. The standard procedure is to let the order of one of these subsets be x, evaluate all the subset orders (there should be 8 of them in a Venn diagram for three overlapping sets) in terms of x and then recognise that the requirement of non-negativity of subset orders will give limits on the value of x. If the error is in $n(W \cap S \cap D)$ we find that its order has to be 190. If the error is in $n(W \cap D)$ we find its order x must be such that $80 \leq x \leq 110$. Many candidates concentrated entirely upon only one of $n(W \cap D)$ or $n(W \cap S \cap D)$.

iii. The brackets **are** important. In words…this subset represents 'Those workers competent with Wordprocessing and those competent with Spreadsheets only'. Other, equivalent, expressions are of course allowable. Little words like 'only' are very important – sometimes omitting them drastically changes the statement's logic. Following on from ii., there will be two possible answers for the order of the given subset. If the original error was in $n(W \cap S \cap D)$ then $n(W \cup (D^c \cap S)) = 440$. However, if the original error was in $n(W \cap D)$ then $330 \leq n(W \cup (D^c \cap S)) \leq 360$.

Question 3

Generally quite a well done question with most students able to obtain good marks. Parts (a) and (b) are very standard type questions testing the student's ability to work with both difference and differential equations. The biggest difficulties arose from deciding upon the appropriate particular solution (especially so for Q3(b))
It seems sensible that this report should merely reproduce the key stages of the solution procedure for each part:

(a) The Auxiliary equation is $m^2 + 4 = 0$ which gives two imaginary solutions of $m = \pm 2i$ and hence a complementary solution of the $y_t = 2^t\left[A\cos\frac{\pi t}{2} + B\sin\frac{\pi t}{2}\right]$.
For a particular solution we try $p = C + Dt$ (there is no need for anything more complicated for this question!) and find that $C = 6$ and $D = 1$. Combining the two solutions and using the initial conditions to solve for the constants we find that $A = -5$ and $B = -2$.
Hence the required solution is $y_t = 6 + t + 2^t\left[-5\cos\frac{\pi t}{2} - 2\sin\frac{\pi t}{2}\right]$
The required graph is oscillating with increasing magnitude and hence is unstable. An easy mark is obtained by simply starting the graph through the (0,1) coordinate !!

(b) The Auxiliary equation is $m^2 - 7m + 12 = 0$ which gives two real solutions of $m = 3$ or 4 and hence a complementary solution of the $y = Ae^{3x} + Be^{4x}$
For a particular solution we try $y = (Cx^2 + Dx)e^{4x}$ and find that $C = 1/2$ and $D = -1$. Learn the rules for creating particular solutions!
Combining the two solutions and using the initial conditions to solve for the constants we find that $A = 0$ and $B = 1$.
Hence the required solution is $y = (\frac{x^2}{2} - x + 1)e^{4x}$.

Question 4

In essence the whole question is quite straightforward (you might say mechanical). Unfortunately the answers often demonstrated that the subject guide is inadequately used. Virtually all the marks could be earned by simply repeating some of the words from the subject guide for parts (a) and (b) and following the standard procedure for creating a Box and Whisker diagram for part (c).

(a) Look at page 86 in the subject guide and you will see that the seven marks for the question correspond precisely to the seven reasons for carrying out an initial simple data analysis listed in the guide.

(b) Again an appropriate answer comes directly from the subject guide. Talk about Box Plots, Scatter plots (of various forms), etc. There are only three marks so nothing too lengthy or profound was required.

(c) Admittedly the data tabulated could have been clearer – however there is nothing at all wrong with it. Candidates should also be aware that there are some small variations in Box & Whisker diagrams produced in certain texts. The definition/methodology expected from the candidates naturally corresponds entirely with the method given in the subject guide. Note: where data spans a class, for example, 2 to 12 then it is usual to treat all values as having a middle value in that class, namely 7 for the class in question.

To finish off the answer candidates were expected to make an appropriate comment about the number of extreme values, outliers and (especially) skewness of the data.

Question 5

Not a particularly elegant question and rather lacking in an obvious application to management. Nonetheless it proved to be quite a good test of candidate's ability to work with some mathematical concepts.

(a)i. A standard Simpson's rule type question with the added complication of needing to be able to handle a *cot* function (and to remember to work with radians for such a question since the integral limits are given in terms of which is presumable in radians and not degrees). Note that there are five ordinates requested which will therefore be for θ = /3, 5 /12, /2, 7 /12, 2 /3 and hence you will need to evaluate *cot*(/6), *tan*(5 /24), …up to *cot*(/3). Many candidates forgot the 2 factor in the denominator. Using Simpson's rule with care should give a value of the integral of 1.0988.

ii. The Examiners predicted that this might be a particularly hard part of the paper for those tackling this question. For those who cannot see the appropriate approach, recognise that *cot* is *cos* divided by *sin* and hence the numerator is the differential of the denominator. Hence the answer is going to involve *log*(*sin* θ/2). In fact the answer is 2 *log*(*sin* θ/2) evaluated between /3 and 2 /3. This gives an answer of 1.0986. Thus the Simpson's rule approach is a good approximation to the precise answer (as you should expect).

(b)i. A few missed this part out – clearly not knowing what an Argand diagram was. Others labelled the axes inappropriately (*x* for the horizontal and *y* for the vertical is not precise enough for this circumstance).
The axes should be Real (horizontally) and Imaginary (vertically). Furthermore the complex number is shown by drawing a line from the origin to the appropriate coordinate point ((-7, 8) in this case). The coordinate point alone is insufficient.

ii. Most candidates recognised the need to work with complex conjugates to solve for *v*. However expansion of the brackets and resulting simplification produced various careless errors. The full answer goes as follows:
If $uv = -29 + 17i$ then

$$v = \frac{-29+17i}{-7+8i} = \frac{(-29+17i)(-7-8i)}{49+64} = \frac{339+113i}{113} = 3+i.$$

iii. First of all one needs to evaluate $v - u$ as a complex number and then determine its magnitude – effectively the length of the line if drawn on an Argand diagram. The answer goes as follows:

$$|v-u| = |3+i+7-8i| = |10-7i| = \sqrt{100+49} = 12.21.$$

Question 6
Once again there is quite a lot of 'bookwork' (or what some students call 'theory'). Indeed the whole of (a) and (b) can be obtained directly from the subject guide (or elsewhere!).

(a)i. This is a stochastic process (either continuous or discrete time and usually discrete state space) which models the way in which a 'particle' changes position. Random walks are highly used in financial market analysis and instrument pricing.

ii. Here you should discuss (very briefly will do) the possible ways in which people might queue, for example, all people queue for one server, or for individual servers, or they queue for a set amount of time, etc. Is there a maximum queue length? Is there a fixed or variable amount of servers, etc. A little network type diagram showing the queuing procedure would be a suitable means of answering the question.

(c)i. For P_1, absorbing states do not exist and hence the Markov chain cannot be absorbing – specifically make this statement.
For P_2, C is an absorbing state and it is an absorbing Markov Chain since all states can lead to the absorbing state.
For P_3, B and C are absorbing states and again the transition matrix does represent an absorbing Markov chain since all states can reach at least one absorbing state.

ii. When drawing the network diagram remember to include an appropriately directed arrow on the arcs, and to show the loops (when they exist) from some states back to themselves.

iii. Perhaps it was lucky for several candidates that more marks were not assigned for this part. One might reach A when t = 1, or when t = 2 or when t = 3. The probabilities for these three events are 1/3, (1/3)(1/3) and (1/3)(1/3)(1/3) which when added gives an answer of 13/27 for reaching A at or before t = 3. Our required answer is therefore 1 – (13/27) = 14/27.

Question 7

A popular question, as always. A few marks were sometimes lost through poor arithmetic or poor (probably careless) choices of companies to cluster. When the methods are entirely confused, however, or when there is no consistency of approach the marks are reduced considerably. Do take care!

i. A similarity matrix counts up the number of 0's or 1's a pair of companies have in common. Whether you divide by 6 (to obtain a proportion of characteristics in common rather than an absolute number) or not is entirely the student's choice – however it is unnecessary and time consuming (and probably hinders accuracy and understanding).
Another unnecessary step is to produce a complete matrix. It is quite acceptable to simply give the upper (or lower if you prefer) triangular matrix. If you do want to include the diagonal elements showing the similarity between company 1 and company 1 (for example) then make certain it gets the correct value i.e. 6 if you are not dividing by 6, or 1 if you are. Candidates must be penalised if they have a mixed approach.

ii. Make certain you have the single and complete methods the right way round – otherwise you are in danger of a heavy penalty indeed. For a similarity matrix where we are trying to join the most similar companies we join companies (or existing clusters) because they have **high** similarities. The way in which we measure the similarity between two clusters is according to the most similar pair (for single linkage) or the most dissimilar pair (for complete linkage). What might be confusing in this example is that the two methods are initially similar in their clustering and, furthermore, there are a lot of arbitrary decisions that the candidate has to make (because of non-unique highest valued clusters). This is one of the things that the candidates were expected to comment on in part iii.

iii. Remember to include a horizontal axis (going from high to low similarity in this case). Also remember to complete the whole dendogram right down to all 6 companies being in the same cluster. Finally remember to avoid lines crossing over each other. Appropriate comments concern the arbitrary choices being made (see ii. above) and the extent (or otherwise) to which the two methods give different clusters.

Question 8

Parts i.–iv. are bookwork (theory) and should have been easy marks. The answers need not be overly long. Something along the following lines will suffice for full marks:

i. The time period between the making of the forecast and the time period which we are forecasting for, namely how far in advance we are forecasting.

ii. Increase the number of past observations used, for example, take a 5-point rather than 3-point moving average.

iii. Use a lower value for the smoothing constant.

iv. The fitted region is where we assume the data is known and we fit the model to this data in an optimal way. The forecast region is where we assume the data is not known and has to be forecast using the model derived from the fitted region.

(b) It was not really intended that candidates should produce forecasts using each smoothing constant – although those that did, and then used RMSE or MAD to decide to use the largest value were given full credit. Candidates were really intended to pick from the smoothing constants by looking at the data and realising that a highly responsive (namely, high smoothing constant) will be best.
Two common errors with exponential smoothing are a) using the data for a period when forecasting the value for the same period! and b) evaluating $F_t = \alpha F_{t-1} + (1-\alpha)X_{t-1}$ rather than $F_t = \alpha X_{t-1} + (1-\alpha)F_{t-1}$. Thus, for example the correct forecast for February is 0.4(580) + 0.6(600) = 592 not 0.4(600) + 0.6(580) = 588. With the right methodology and using the highest smoothing constant one should get a December forecast of 1278.83.

(c) Similar comments should be made about the approach to this part. Again it was really intended that candidates should pre-select the 3 month moving average in order to capture the periodical step changes in the data, namely to make the forecasts more responsive. A further source of confusion to the candidates was whether to use the method for trend determination or for forecasting *per se*. Those who chose the first method placed the moving average in the middle of the months of data they used and then extrapolated (trended) the results out to produce a forecast for December. A better approach (and certainly simpler and more suitable with such a small amount of data) is to simply take the forecast for December as the average of the data for July to November, namely 1283.33.

Examination paper for 2005

There will be no change to the format, style or number of questions in the examination paper for 2005.