

THE ROYAL STATISTICAL SOCIETY

2008 EXAMINATIONS – SOLUTIONS

ORDINARY CERTIFICATE

PAPER I

The Society provides these solutions to assist candidates preparing for the examinations in future years and for the information of any other persons using the examinations.

The solutions should NOT be seen as "model answers". Rather, they have been written out in considerable detail and are intended as learning aids.

Users of the solutions should always be aware that in many cases there are valid alternative methods. Also, in the many cases where discussion is called for, there may be other valid points that could be made.

While every care has been taken with the preparation of these solutions, the Society will not be responsible for any errors or omissions.

The Society will not enter into any correspondence in respect of these solutions.

Ordinary Certificate, Paper I, 2008. Question 1

The letter should be on the organisation's headed notepaper.

We would like to know whether the monthly newsletter meets our members' needs. You have been chosen by a random selection method as one of the members to approach for opinions. We hope that you will complete the short enclosed questionnaire. Your responses will be kept confidential and results of the survey will be published in aggregate form only.

Many questions can be answered by simply ticking the relevant response.

1. In which year did you join the organisation?
2. Are you: male ? female ?
3. Please tick the box showing your **main** employment status.
student
employed part-time
employed full-time
unemployed
retired
4. Do you read everything in the newsletter? Yes No
5. What do you do with the newsletter when you have finished with it?
Keep it for a while
Pass it on to someone else
Throw it away
6. We are particularly interested in knowing members' opinions of the special features in the newsletter. In general, do you think:

There are too many
The number is about right
There are too few

In general do you find them:

- Of no interest
- A few are of interest
- Most or all of them are of interest

Solution continued on next page

7. Do you think that advertisements should be included?

Yes No No opinion

8. Do you agree that publishing a newsletter at intervals of one month is about right?

Yes No

If you answered No, at what interval would you like the newsletter to be published?

.....

9. Please rate your satisfaction with the newsletter on a scale of 1 to 5, where 1 is not at all satisfied and 5 is extremely satisfied.

1	2	3	4	5

Thank you very much for completing this questionnaire. Please return it in the enclosed reply-paid envelope.

Ordinary Certificate, Paper I, 2008. Question 2

Advantages of including a questionnaire as an insert in the newsletter include:

There are no costs of selecting a sample, addressing envelopes etc, or separate postage costs.

The questionnaire is sent to all members so potentially every member could give an opinion of the newsletter.

Disadvantages include:

Could be time consuming/costly to insert a questionnaire in every newsletter.

The extra weight of the questionnaire might increase the postage costs.

Unless the mailing list of members is up to date, some newsletters will not be delivered (and this may not be detected). [The mailing list is likely to be at its most up to date fairly early in the organisation's financial year soon after membership subscriptions have become due and been paid; but it is still very likely that there will be errors and omissions.]

Some members might postpone opening the newsletter if they are busy, or might never open it.

The response rate is likely to be low as people tend to ignore inserts in publications.

Difficult/costly to follow-up non-response and could only be done if names or other identifiers were requested in the questionnaire.

There is likely to be a respondent bias towards, for example, those who have strong views about the newsletter and those who have time on their hands.

Advantages of sending the questionnaire separately to a random sample of members include:

Members might be more likely to open the envelope containing a questionnaire than to open one which they recognise as containing the newsletter.

It is more likely that members will respond to a personal request than to a questionnaire sent as an insert in the newsletter.

It will be easier and cheaper to follow-up non-response even if this means sending reminder letters to everyone if responses are anonymous.

A random sample of members should give a representative sample of members covering all types of members and opinions, provided non-response is minimal.

Disadvantages include:

The financial and time costs of selecting the sample, addressing and stuffing envelopes.

Costs associated with following up non-response.

Ordinary Certificate, Paper I, 2008. Question 3

- (i) Systematic sample from the alphabetical list of members.

This should achieve a sample close to simple random. The proportions in the sample in the three grades of membership should be similar to the actual proportions in the entire membership, and likewise there should be a good spread across the years of joining the organisation. An example of a potential problem is that, if the number in any grade is very small (e.g. there are only a few student members), this method might not select any members at all from that grade.

Systematic sample from the list ordered by year of joining the organisation.

This would achieve a good sample across time, with numbers from different periods of joining represented in much the same proportions as in the entire membership. This is important for picking up any trends over time. Year of joining is likely to be related to some extent to grade of membership, perhaps with those who joined a long time ago more likely to be retired and those who joined recently more likely to be students, and this method would thus also sample the three grades in roughly the same proportions as in the entire membership. However, if very few joined in some years for some reason important to the organisation, such as a sharp rise in subscriptions, then this sampling method might not sample that group at all.

- (ii) As pointed out above, grade of membership and year of joining are likely to be related, so there might not be much advantage to be gained by systematic sampling from the list reordered by grade rather than by year of joining. However, this should ensure that each grade is correctly represented in the sample, and this might therefore be preferable to the other methods, depending on the purpose of the survey. [If the grade information is readily available in this way, it might be desirable to stratify by grade.]

Ordinary Certificate, Paper I, 2008. Question 4

(i) Total of 1162 members.

Need $(141/1162) \times 95 = 0.1213 \times 95 = 11.53$ student members, say 12.

Need $(782/1162) \times 95 = 0.6721 \times 95 = 63.9$ ordinary members, say 64.

Need $(239/1162) \times 95 = 0.2057 \times 95 = 19.54$ retired members, say 20.

Note that $12 + 64 + 20 = 96$ but there is sufficient budget only for 95 altogether. So take one fewer in one of the groups, say 11 student members (as 11.53 is very slightly further from 12 than 19.54 is from 20 – but there is very little in it, and this does depend on the calculations being worked to 2 decimal places).

(ii)

Membership grade	N_i	s_i	$N_i s_i$	$N_i s_i / \sum N_i s_i$	$(N_i s_i / \sum N_i s_i) \times 95$
Student	141	226.21	31895.61	0.0597	5.67
Ordinary	782	550.12	430193.84	0.8046	76.44
Retired	239	303.60	72560.40	0.1357	12.89

$(\sum N_i s_i = 534649.85)$

This suggests 6 student members, 76 ordinary members and 13 retired members.

$6 + 76 + 13 = 95$, so the budget constraint is satisfied.

(iii) The method described in part (ii) assumes that there has been little change in the SDs of expenditure over the two years. If however the SDs now are substantially different, the sample sizes found in part (ii) will not be optimum. In contrast the sampling method in (i) uses only the membership numbers which can assumed to still represent the groups, at least proportionally, and this could seem fairer to members.

Note in particular that in (ii) only 6 student members are selected. If non-response is high in this group, as is likely, then the achieved sample might be very small indeed.

Ordinary Certificate, Paper I, 2008. Question 5

- (i) A pilot survey is a small sample survey carried out at the planning stage of a full scale census or survey. It is used to test the design of the questionnaire, the sample frame and the general administrative procedures. It provides estimates of the costs and response rates, and of the variances of measured variables. It can lead to improvements, technical and/or procedural, in the conduct of the full census or survey.

A sample survey is, as the name suggests, a survey of a sample of a population. The sample would usually be selected by a random procedure and would be considerably larger than the sample used in a pilot survey, but normally considerably smaller than the population. It would be done when it is too costly both financially and in terms of time to survey the whole population. For fixed resources, information can be obtained in greater depth in a sample survey than in a census, and quality control can be more stringent.

A census is strictly a complete count of a population, but is taken more generally to mean a survey of the complete population. It is done when it is important to get information about every member of the population and to provide benchmark figures with which later figures can be compared. A census might also be appropriate when the population is very small so that little is to be gained by sampling it, especially if there is large variability in the variables of interest.

- (ii) Results from a census might differ from the population figures if questions are misunderstood, or if the responses given to questions are incorrect – either because the respondent has answered incorrectly (or not at all) or because responses have been recorded incorrectly or copied incorrectly into a database. Results might also differ because some particular groups in the population are hard to reach and so, in practice, these groups turn out to be under-represented even though it was meant to be a complete census.

Examples will depend on candidates' experience. In the UK Population Census, the homeless and young professionals are found to be hard to reach (the former for obvious reasons, the latter often because they work very long hours and are thus only rarely at home to be interviewed). In rural communities, several groups of workers may be hard to locate, and others may be itinerant.

Ordinary Certificate, Paper I, 2008. Question 6

- (i) A region can be considered to be a cluster of supermarkets. Cluster sampling could consist of selecting one or more (but not all) regions at random. All supermarkets in these regions would then be in the sample of supermarkets (1-stage cluster sampling) or a sample from the supermarkets in these regions could be taken (2-stage cluster sampling).

An advantage of this method is that interviewers would need to be employed in the selected regions only, saving on travel and administrative time and costs.

Disadvantages include: the supermarkets in the chosen region(s) might be atypical; the variances of estimators in this method of sampling tend to be higher than in stratified and simple random sampling; the estimation method is more complex.

- (ii) The supermarkets could be stratified into the three size groups, and a random sample of supermarkets taken from each group.

Advantages of this method include: all three sizes of supermarket will be included; different sampling methods and questionnaires could be used in each group if necessary as the types of customer could be very different in the three types of supermarket; it is easy to produce estimates for each group; variances tend to be lower under stratified sampling than under cluster sampling.

A disadvantage is that the supermarkets could be scattered widely geographically, so that interviewer costs may be high.

Ordinary Certificate, Paper I, 2008. Question 7

- (i) As the interviewers are to select the customers and interview them at the time of selection, quota sampling would be an appropriate sampling method.

Interviewers should go to the stores at varying times during the day and week. They should aim to interview customers of all ages and of both sexes, in about the same proportions as customers using the supermarket. Observing customers at check-outs would give an indication of these proportions, as might asking staff (but staff will not necessarily see the whole range of customers; for example, staff who work only in the evenings are likely to see mainly people who are employed during the day and who do not have children under 16).

Different types of customer should be approached – both those shopping on their own and those who are with others, both those who are doing a big shop and those who have come in for one or two items only. Customers should be selected taking no account of their dress or demeanour, that is interviewers should avoid approaching only those who look respectable, who look as if they might be sympathetic to answering questions, who appear not to be in a hurry, etc. Customers could be selected from those in the check-out queue, from those leaving the store, or by walking round the store. Standing by one display in the store would not be sensible as not all customers will pass it.

Interviewers should dress moderately, approach customers politely, etc. They should not try to persuade customers to respond against their will.

- (ii) Interviewers should treat the customers with respect. They should ask the questions and read any introduction and closing note of thanks exactly as written on the questionnaire, and should use a neutral tone of voice to avoid biasing the customers' answers. They should not comment on or react in any way to the customers' responses. They should record the responses that are given and not make them up or expand on comments. They might (discreetly) record specific observations as regards customers, such as if a customer was hard of hearing leading to communication difficulties.

Ordinary Certificate, Paper I, 2008. Question 8

The observers could stand by the displays of organic fruit and vegetables with a pre-prepared check-list to be completed by ticking boxes indicating characteristics such as sex and estimated age group for customers who put such produce in their baskets or trolleys, and other boxes indicating what organic fruit and vegetables are taken and in what kinds of quantity. However, there might be a problem if fruit and vegetable displays are not very close to each other, in which case individual customers might need to be followed round the displays to collect any useful information.

Alternatively, or as well, the observers could stand by the checkouts to make similar observations.

Observations would need to be made at different times of day and on different days of the week.

Advantages of an observational study in this context over the use of a questionnaire include:

- It does not take up customers' time.

- It does not rely on customers answering questions truthfully.

- Customers do not have to be approached and persuaded to take part.

Disadvantages include:

- It is only possible to observe whether organic fruit and vegetables are bought by a customer at the time of observation. No information can be gained about customers who buy such produce at other times, or who do not buy because they cannot find what they are looking for or think it too expensive.

- No information can be obtained about customer characteristics that cannot be observed, such as household size, place of residence and occupation.

- The results depend very much on the ability of interviewers to decide which customers to observe and their ability to make correct records.

[Note. Possible alternative survey methods include (short and simple) interviews of customers in check-out queues. Excellent information on products bought should be available from the data base holding information from the check-out tills.]