MEI Structured Mathematics

Practice Comprehension task - 1

Randomised Response Technique

It is often the case that decision makers need accurate information about activities which are criminal or anti-social. In such cases the usual research technique of selecting a random sample of the population under investigation and interviewing them is of no value. People will not answer accurately for fear of recriminating themselves. A procedure known as Randomised Response Technique can result in a much clearer picture of the true state of affairs.

The following case study is based on this procedure.

The Social Services department in a town are concerned about the level of domestic violence, and in particular the incidence of women suffering violence from their husbands They want to establish a shelter for victims of such treatment but need information about the extent of the problem if they are to secure the necessary funding.

Consequently they decide to carry out a survey among married couples. In this context they use the words ***husbands*** and ***wives*** to mean partners of the opposite sex with whom they are living, and describe such couples as ***married***.

They decide to adopt a working description of a violent husband as one who has hit his wife during the last 6 months. However, then are aware that if they were to ask a random sample of husbands if they were violent, few would admit to it.

Instead, having selected a random sample of 100 husbands, they ask them to take part in the following exercise.

Toss a coin. If it comes down Heads, answer question A; if it comes down Tails, answer question B.

Question A: Have you behaved violently towards your wife during the last six months.

Question B: Are you male?

It is explained to those taking part that if they answer ***Yes*** there is no way that anyone can know which question they are answering and so they are not giving away any damaging information about themselves.

The result of the council's survey was:

No 40 Yes 60

From this they argued that since the probability of the coin coming Tails is 0.5, it can be expected to have come down Tails 100×0.5 or 50 times. All of these would result in a Yes response.

Since the total of yes responses was in fact 60, 10 of these could be taken to have come from those whose coin landed Heads, in other words to be positive responses to question A: *Have you behaved violently towards your wife during the last 6 months?*

Since 50 people can be expected to have answered question A, following their coins coming Heads, it follows that the proportion of violent husbands is 10 out of 50, or 20%

This is shown in the following table.

Coin	No	Yes	Total
Heads - A	40	10	50
Tails - B	0	50	50
Totals	40	60	100

Table 1

However, an official from the Council's treasury department raises an objection. He says that he tossed a coin 100 times and it did not come down 50 heads and 50 tails. It actually came down 44 heads and 56 tails. He points out that if you consider both this possibility and that of 56 heads and 44 tails you get wildly different results.

Case 1: 44 heads and 56 tails

Coin	No	Yes	Total
Heads - A	40	4	44
Tails - B	0	56	56
Totals	40	60	100

Table 2

The conclusion here is that the proportion of husbands who are violent is 4 out of 44, about 9%

Case 2: 56 heads and 44 tails

A table similar to table 2 above will indicate that the proportion of husbands who are violent is about 29%.

The difference between 9% and 29% is too great to proceed with any decision making with any confidence so the Social Services Department called in a Statistician for advice.

His report was as follows:

1. Your procedure involved 100 people tossing coins, and your calculations depended on the assumption that the 100 coins came down exactly 50 Heads and 50 Tails.
   This is unlikely to be the case. If you toss 100 coins many times the results will be spread out around the 50-50 result, and there is nothing unusual in the Council official's 44-56 split. It is possible to work out a range of likely results, and doing this shows that in 95% of cases the Head-Tail split should lie between 40-60 and 60-40.

If you repeat your calculations with these extreme values you come out with the proportion of husbands who are violent somewhere in the range 0% to 33%. Clearly this is too wide a range of outcomes to be helpful to you in presenting your case.

2. A very much larger sample will be needed to reduce the possible percentage error to an acceptable level. *60*

3. The design of your questions is flawed. Anyone who answers *No* must be answering question A and so is giving away information that he might not wish to. You have assumed that all the husbands questioned regard violence as a bad thing which they will not admit to. However, the possibility must be considered that there are husbands who for whatever reason would be reluctant to admit to not being violent. *65*

   Consequently your question B should be altered so that *No* is a possible answer to that question as well. One possible design is the following.

   Toss a coin and throw a die together. If the coin comes Heads, answer question B. If it comes Tails then answer question B. *70*

   Question A: Have you behaved violently towards your wife in the past 6 months?
   Question B: Was the score on the die 6?

There is now of course a further source of error, due to the variation involved in throwing the die in question B.

The technique illustrated in this paper is called ***Randomised Response Technique***. The method assumes that people will tell the truth when there is no danger of incriminating themselves and available evidence indicates that this is indeed the case. In practice it is used effectively but only when it is possible to take large samples. Its use to determine the extent of domestic violence might be more appropriate to a national survey carried out by government that to one conducted by the Social Services Department in one town. *75* *80*

Questions.

1 **(i)** Give an example of information (other than concerning domestic violence) where Randomised Response Technique might be an appropriate means of collecting data.

[1]

(ii) Describe briefly the advantages of the technique and state any problems involved.

[2]

2 In Case 1 (line 42 onwards) the table shows that the proportion of husbands who are violent is 4 out of 44, about 9%

(i) Complete the table below for Case 2, where the split Heads-Tails is 56 to 44.

[1]

Coin	No	Yes	Total
Heads - A			56
Tails - B	0	44	44
Totals	40	60	100

(ii) Show from the entries in your table how the figure of about 29% is achieved.

[2]

3 In line 58 the report of the Statistician states that the upper limit for the proportion of violent husbands could vary between 0% and 33% .

(i) Complete the table below, giving figures that could achieve a proportion of 33%.

Coin	No	Yes	Total
Heads - A			
Tails - B	0		
Totals	40	60	100

[3]

(ii) Complete the table below, giving figures that could achieve a proportion of 0%.

Coin	No	Yes	Total
Heads - A			
Tails - B	0		
Totals	40	60	100

[3]

4 The Statistician recommended a much larger sample and also a changed question B.

3600 husbands were included in another survey and were asked the following:

Toss a coin and a die together. If the coin comes down Heads answer question A and if it comes down Tails answer question B.

Question A: Have you behaved violently towards your wife in the last 6 months?
Question B: Was the score of the die 6?

The result was: *No* 2894 *Yes* 706

(i) In order to process these results in the same way as in the procedure described in the paper, what assumptions need to be made? [2]

(ii) With these assumptions, estimate the proportion of violent husbands. [4]

Answers.

<table>
<tr><td>1</td><td>(i)</td><td>Anything that is considered criminal but which may be widespread
e.g. dishonest tax returns
speeding on motorways
Drug abuse</td><td>B1

1</td><td></td></tr>
<tr><td></td><td>(ii)</td><td>It makes it possible to collect information that people might not want to admit to for fear of incrimination
Because of the variation ion the alternative response, a large sample would be needed to ensure any kind of validity.</td><td>B1

B1

2</td><td></td></tr>
<tr><td>2</td><td>(i)</td><td><table><tr><th>Coin</th><th>No</th><th>Yes</th><th>Total</th></tr><tr><td>Heads - A</td><td>40</td><td>16</td><td>56</td></tr><tr><td>Tails - B</td><td>0</td><td>44</td><td>44</td></tr><tr><td>Totals</td><td>40</td><td>60</td><td>100</td></tr></table></td><td>B1

1</td><td>all correct</td></tr>
<tr><td></td><td>(ii)</td><td>Proportion = Heads-Yes entry divided by total heads (56)
i.e. $\frac{16}{56} = 0.2857 \Rightarrow 29\%$</td><td>M1
A1
2</td><td></td></tr>
<tr><td>3</td><td>(i)</td><td><table><tr><th>Coin</th><th>No</th><th>Yes</th><th>Total</th></tr><tr><td>Heads - A</td><td>40</td><td>20</td><td>60</td></tr><tr><td>Tails - B</td><td>0</td><td>40</td><td>40</td></tr><tr><td>Totals</td><td>40</td><td>60</td><td>100</td></tr></table>
Proportion = Heads-Yes entry divided by total heads (60)
i.e. $\frac{20}{60} = 0.3333 \Rightarrow 33\%$</td><td>B1

M1
A1
3</td><td></td></tr>
<tr><td></td><td>(ii)</td><td><table><tr><th>Coin</th><th>No</th><th>Yes</th><th>Total</th></tr><tr><td>Heads - A</td><td>40</td><td>0</td><td>40</td></tr><tr><td>Tails - B</td><td>0</td><td>60</td><td>60</td></tr><tr><td>Totals</td><td>40</td><td>60</td><td>100</td></tr></table>
Proportion = Heads-Yes entry divided by total heads (60)
i.e. $\frac{0}{60} = 0 \Rightarrow 0\%$</td><td>B1

M1
A1
3</td><td></td></tr>
</table>

<table>
<tr><td>4</td><td>(i)</td><td>That the coin comes down 50-50 Heads and tails
That the die shows different numbers an equal number of times. i.e. P(6) = 1/6</td><td>B1
B1
2</td><td></td></tr>
<tr><td></td><td>(ii)</td><td>
<table>
<tr><th>Coin</th><th>No</th><th>Yes</th><th>Total</th></tr>
<tr><td>Heads - A</td><td>1394</td><td>406</td><td>1800</td></tr>
<tr><td>Tails - B</td><td>1500</td><td>300</td><td>1800</td></tr>
<tr><td>Totals</td><td>2894</td><td>706</td><td>3600</td></tr>
</table>
Proportion = Heads-Yes entry divided by total heads (1800)
i.e. $\frac{406}{1800} = 0.2255 \Rightarrow 23\%$
</td><td>B1
B1
M1
A1
4</td><td>Entries of 1800
Other entries</td></tr>
</table>