Statistical Hypothesis testing

Mixed practice 18

1. A market researcher is asked to conduct a survey outside a library. He is asked to sample 100 male library users, 100 female library users, 50 males who have not been into the library and 50 females who have not been into the library.

   What type of sampling method is this?

2. The organisers of the school concert want to find out how many of the students are planning to attend the concert. The school has 48 different tutor groups, and they decide to select a sample of students in the following way:

   - They choose five tutor groups randomly.
   - From each tutor group, they select a random sample of ten students.

   a What name is given to this type of sampling procedure?

   b Explain why this procedure might not give a representative sample in this case.

   The organisers later decide that they should take a simple random sample of 50 students instead.

   c Describe how they might obtain such a sample.

3. Gavin has a six-sided dice that he thinks is biased and shows more 5s than it should. He wants to conduct a hypothesis test to test his belief.

   a State suitable null and alternative hypotheses.

   Gavin rolls the dice 75 times and obtains 18 5s.

   b Conduct the test at the 5% significance level, stating your conclusion clearly.

4. Lisa, who takes the bus to school, is late for school on average once in every eight days. She has recently moved closer to the school and now walks. In the last 30 days she was late only twice.

   Is there evidence, at the 10% significance level, that the probability of Lisa being late for school has decreased? State your hypotheses and your conclusion clearly.

5. A village has a population of 600 people. A sample of 12 people is obtained as follows. A list of all 600 people is obtained and a three-digit number, between 001 and 600 inclusive, is allocated to each name in alphabetical order. Twelve three-digit random numbers, between 001 and 600 inclusive, are obtained and the people whose names correspond to those numbers are chosen.

   a Find the probability that the first number chosen is 500 or less.

   b When the selection has been made, it is found that all of the numbers chosen are 500 or less. One of the people in the village says: 'The sampling method must have been biased.' Comment on this statement.

6. The head teacher of a school asks for volunteers from among the pupils to take part in a survey on political interests.

   i Explain why a sample consisting of all the volunteers is unlikely to give a true picture of the political interests of all pupils in the school.

   ii Describe a better method of obtaining the sample.

7 A doctor knows that 20% of people suffer from side effects when treated with a certain drug. He wants to see if the proportion of people suffering from side effects is lower with a new drug. He looks at a random sample of 30 people treated with the new drug.

What is the largest number of people who could suffer from side effects and still conclude at 5% significance that the new drug has a lower proportion of side effects?

8 In the UK, the proportion of families who own their home (as opposed to renting) is 64%. Sabina wants to find out whether this proportion is different in Germany. She surveys a random sample of 180 families in Germany and finds that 98 of them own their homes.

Conduct a hypothesis test at the 5% significance level to test whether the proportion of families in Germany who own their home is different from that in the UK.

9 **a** Define a simple random sample.

Aneka is investigating attitudes to sport among students at her school. She decides to carry out a survey using a sample of 70 students. There are the same number of boys and girls at the school, so Aneka randomly chooses 35 boys and 35 girls.

b **i** State the name for this type of sample.

ii Explain why in this case, this type of sample is better than a simple random sample.

One of Aneka's questions is about participation in school sports teams. She wants to find out whether more than 40% of students play for a school team. She sets up the following hypotheses: $H_0: p = 0.4$, $H_1: p < 0.4$, where p is the proportion of students who play for a school sports team.

c Find the critical region for the hypothesis test at the 10% significance level, using a sample of 70 students.

d What is the probability of incorrectly rejecting the null hypothesis?

e In Aneka's sample, 32 students play for a school team. State the conclusion of the test.

10 In a rearrangement code, the letters of a message are rearranged so that the frequency with which any particular letter appears is the same as in the original message. In ordinary German the letter *e* appears 19% of the time. A certain encoded message of 20 letters contains one letter *e*.

i Using an exact binomial distribution, test at the 10% significance level whether there is evidence that the proportion of the letter *e* in the language from which this message is a sample is less than in German, i.e., less than 19%.

ii Give a reason why a binomial distribution might not be an appropriate model in this context.

© OCR, AS GCE Mathematics, Paper 4733, June 2007

11 A test is constructed to see if a coin is biased. It is tossed 10 times and if there are 10 heads, 9 heads, 1 head or 0 heads it is declared to be biased.

For each of the following, explain whether or not it could be the significance level for this test:

a 1% **b** 2% **c** 10% **d** 20%

Mixed practice 18

1 Quota sampling

2 **a** Cluster sampling

b Students from the same tutor group are likely to go or not go together

c Obtain a list of all students, number sequentially, use a random number generator to generate 50 numbers, select students with those numbers (ignore repeats and numbers larger than the number of students)

3 **a** $H_0 : p = \frac{1}{6}$ $H_1 : p > \frac{1}{6}$ where p is the underlying probability of rolling a 5

b There is insufficient evidence that the probability of rolling a 5 is greater than $\frac{1}{6}$

4 $H_0 : p = \frac{1}{8}$ $H_1 : p > \frac{1}{8}$ where p is the underlying proportion of days that Lisa is late. There is insufficient evidence that Lisa's probability of being late has decreased from 1 in 8

5 **a** $\frac{5}{6}$

b Not correct; this sample has the same probability as any other sample

6 **a** Pupils who are interested in politic are more likely to volunteer.

b simple random sample – obtain a list of all pupils, number sequentially, use a random number generator to generate 50 numbers, select pupils with those numbers (ignore repeats and numbers larger than the number of pupils)

7 2

8 $p = 0.00523 < 0.025$, sufficient evidence to reject H_0, there is evidence to conclude that home ownership in Germany is lower than 64%

9 **a** Each possible group of a given size has an equal probability to be included within the sample.

b **i** Stratified sample

ii Attitudes to sport may differ between boys and girls.

c $X \geqslant 34$

d 0.0906

e There is insufficient evidence that the proportion of students who play in a school sports team is greater than 40%.

10 **a** H_0: $p = 0.19$, H1: $p < 0.19$. Reject H_0 at 10%, $p = 0.0841$

b Letters are not distributed independently.

11 **a** No **b** No **c** Yes **d** No