# GCE Examinations

# Statistics
# Module S3

Advanced Subsidiary / Advanced Level

Paper E

Time: 1 hour 30 minutes

*Instructions and Information*

Candidates may use any calculator except those with a facility for symbolic algebra and / or calculus.

Full marks may be obtained for answers to ALL questions.

Mathematical and statistical formulae and tables are available.

This paper has 7 questions.

*Advice to Candidates*

You must show sufficient working to make your methods clear to an examiner. Answers without working will gain no credit.




*Written by Shaun Armstrong & Chris Huffer*

*© Solomon Press*

*These sheets may be copied for use solely by the purchaser's institute.*

**1.** A Veterinary Surgeon wishes to survey a stratified sample of size 100 from those people who have pets registered at her surgery. The list below shows the strata to be used and the number in each group.

- people who own just dogs – 165,
- people who own just cats – 140,
- people who own just small mammals – 105,
- others, including those who own more than one type of pet – 90.

*(a)* Find how many members of each group should be included in the sample. **(2 marks)**

*(b)* Give two advantages of using stratified sampling. **(2 marks)**

---

**2.** A psychologist is investigating the numbers people choose when asked to pick a number at random in a given interval. He finds that when asked to pick a number between 0 and 100 people are less likely to pick certain numbers, such as multiples of ten. He believes, however that if people are asked to pick an odd number between 0 and 100 they are equally likely to pick a number ending in any of the digits 1, 3, 5, 7 or 9.

To test this theory he asks 80 people to pick an odd number between 0 and 100 and records the last digit of the numbers chosen. His results are shown in the table below.

| Last Digit | 1 | 3 | 5 | 7 | 9 |
|---|---|---|---|---|---|
| Frequency | 16 | 20 | 14 | 17 | 13 |

Stating your hypotheses clearly and using a 10% level of significance test the psychologist's theory.

**(9 marks)**

---

**3.** A clothes manufacturer wishes to find out if adult females have become taller on average since twenty years ago when their mean height was 5 ft 6 inches.

Studies over time have shown that the standard deviation of the height of adult females has been fairly constant at 2.3 inches. The manager wishes to test if the mean height is now more than 5 ft 6 inches and takes a sample of 150 adult females.

*(a)* Stating your hypotheses clearly, find the critical region for the mean height of the sample for a test at the 5% level of significance.

**(6 marks)**

The total height of the females in the sample is 832 ft.

*(b)* Carry out the test making your conclusion clear. **(4 marks)**

---

© *Solomon Press*

**4.** For a project a student collects data on engine size and sales over a period of time for the models of cars made by one particular manufacturer. Her results are shown in the table below.

| Engine Capacity (litres) | 1.1 | 1.3 | 1.6 | 2.1 | 2.4 | 2.6 | 2.8 | 3.0 |
|---|---|---|---|---|---|---|---|---|
| Sales | 527 | 632 | 840 | 619 | 350 | 425 | 487 | 401 |

*(a)* Calculate Spearman's rank correlation coefficient for these data. **(6 marks)**

*(b)* Stating your hypotheses clearly, test at the 5% level of significance whether or not there is any evidence of correlation.

**(4 marks)**

*(c)* Explain why it is more appropriate to use Spearman's rank correlation coefficient for this test than the product moment correlation coefficient.

**(2 marks)**

---

**5.** A child is playing with a set of red and blue wooden cubes. The side length of the red cubes is normally distributed with a mean of 14.5 cm and a variance of 16.0 $cm^2$. The side length of the blue cubes is normally distributed with a mean of 12.2 cm and a variance of 9.0 $cm^2$.

*(a)* Find the probability that a randomly chosen red cube will have a side length of more than 3 cm greater than a randomly chosen blue cube.

**(5 marks)**

The child makes two towers, one from 4 red cubes and one from 5 blue cubes. Assuming that the cubes for each colour of tower were chosen at random,

*(b)* find the probability that the red tower is taller than the blue tower. **(6 marks)**

*(c)* Explain why the assumption that the cubes for each tower were chosen at random is unlikely to be realistic.

**(1 mark)**

---

***Turn over***

© *Solomon Press*

**6.** A market researcher recorded the number of adverts for vehicles in each of three categories on *ITV*, *Channel 4* and *Channel 5* over a period of time. The results are shown in the table below.

| | *ITV* | *Channel 4* | *Channel 5* |
|---|---|---|---|
| Family Saloon | 69 | 35 | 28 |
| Sports Car | 20 | 28 | 18 |
| Off-road Vehicle | 12 | 22 | 8 |

*(a)* Stating your hypotheses clearly, test at the 5% level of significance whether or not there is evidence of the proportion of adverts for each type of vehicle being dependent on the channel.

**(13 marks)**

*(b)* Suggest a reason for your result in part *(a)*. **(1 mark)**

---

**7.** *(a)* Briefly state the central limit theorem. **(3 marks)**

A student throws ten dice and records the number of sixes showing. The dice are fair, numbered 1 to 6 on the faces.

*(b)* Write down the distribution of the number of sixes obtained when the ten dice are thrown.

**(2 marks)**

*(c)* Find the mean and variance of this distribution. **(3 marks)**

The student throws the ten dice 100 times, recording the number of sixes showing each time.

*(d)* Find the probability that the mean number of sixes obtained is more than 1.8

**(6 marks)**

---

**END**

© *Solomon Press*