# GCE Examinations

# Statistics
# Module S1

Advanced Subsidiary / Advanced Level

Paper L

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 shop recorded the number of pairs of gloves, $n$, that it sold and the average daytime temperature, $T$ °C, for each month over a 12-month period.

The data was then summarised as follows:

$$\Sigma T = 124, \quad \Sigma n = 384, \quad \Sigma T^2 = 1802, \quad \Sigma n^2 = 18\,518, \quad \Sigma Tn = 2583.$$

*(a)* Calculate the product moment correlation coefficient for these data. **(5 marks)**

*(b)* Comment on what your value shows and suggest a reason for this. **(2 marks)**

---

**2.** Events $A$ and $B$ are independent.

Given also that

$$\mathrm{P}(A) = \tfrac{3}{4} \quad \text{and} \quad \mathrm{P}(A \cap B') = \tfrac{1}{4},$$

Find

*(a)* $\mathrm{P}(A \cap B)$, **(2 marks)**

*(b)* $\mathrm{P}(B)$, **(3 marks)**

*(c)* $\mathrm{P}(A' \cap B')$. **(3 marks)**

---

**3.** The random variable $X$ is such that

$$\mathrm{E}(X) = a \quad \text{and} \quad \mathrm{Var}(X) = b.$$

Find expressions in terms of $a$ and $b$ for

*(a)* $\mathrm{E}(2X + 3)$, **(1 mark)**

*(b)* $\mathrm{Var}(2X + 3)$, **(2 marks)**

*(c)* $\mathrm{E}(X^2)$. **(3 marks)**

*(d)* Show that

$$\mathrm{E}[(X + 1)^2] = (a + 1)^2 + b.$$ **(4 marks)**

---

© *Solomon Press*

**4.** An engineer tested a new material under extreme conditions in a wind tunnel. He recorded the number of microfractures, $n$, that formed and the wind speed, $v$ metres per second, for 8 different values of $v$ with all other conditions remaining constant. He then coded the data using $x = v - 700$ and $y = n - 20$ and calculated the following summary statistics.

$$\Sigma x = 100, \quad \Sigma y = 23, \quad \Sigma x^2 = 215\,000, \quad \Sigma xy = 11\,600.$$

*(a)* Find an equation of the regression line of $y$ on $x$. **(7 marks)**

*(b)* Hence, find an equation of the regression line of $n$ on $v$. **(2 marks)**

*(c)* Use your regression line to estimate the number of microfractures that would be formed if the material was tested in a wind speed of 900 metres per second with all other conditions remaining constant.

**(2 marks)**

---

**5.** An antiques shop recorded the value of items stolen to the nearest pound during each week for a year giving the data in the table below.

| Value of goods stolen (£) | Number of weeks |
|---|---|
| 0 - 199 | 31 |
| 200 - 399 | 6 |
| 400 - 599 | 3 |
| 600 - 799 | 4 |
| 800 - 999 | 5 |
| 1000 - 1999 | 2 |
| 2000 - 2999 | 1 |

Letting $x$ represent the mid-point of each group and using the coding $y = \dfrac{x - 699.5}{200}$,

*(a)* find $\sum fy$. **(3 marks)**

*(b)* estimate to the nearest pound the mean and standard deviation of the value of the goods stolen each week using your value for $\sum fy$ and $\sum fy^2 = 424$.

**(6 marks)**

The median for these data is £82.

*(c)* Explain why the manager of the shop might be reluctant to use either the mean or the median in summarising these data.

**(3 marks)**

---

***Turn over***

© *Solomon Press*

**6.** At the start of a gameshow there are 10 contestants of which 6 are female. In each round of the game, one contestant is eliminated. All of the contestants have the same chance of progressing to the next round each time.

*(a)* Show that the probability that the first two contestants to be eliminated are both male is $\frac{2}{15}$. **(3 marks)**

*(b)* Find the probability that more females than males are eliminated in the first three rounds of the game. **(6 marks)**

*(c)* Given that the first contestant to be eliminated is male, find the probability that the next two contestants to be eliminated are both female. **(3 marks)**

---

**7.** A cyber-cafe recorded how long each user stayed during one day giving the following results.

| Length of stay (minutes) | 0 – | 30 – | 60 – | 90 – | 120 – | 240 – | 360 – |
|---|---|---|---|---|---|---|---|
| Number of users | 15 | 31 | 32 | 23 | 17 | 2 | 0 |

*(a)* Use linear interpolation to estimate the median and quartiles of these data. **(6 marks)**

The results of a previous study had led to the suggestion that the length of time each user stays can be modelled by a normal distribution with a mean of 72 minutes and a standard deviation of 48 minutes.

*(b)* Find the median and quartiles that this model would predict. **(7 marks)**

*(c)* Comment on the suitability of the suggested model in the light of the new results. **(2 marks)**

---

**END**

© *Solomon Press*