eClassroom

INTERNATIONAL A LEVEL MATHEMATICS • STATISTICS 2 (WST02)

Statistics 2 Practice Paper 1

Worked Solutions

AIMING FOR B

Written in the style of the Pearson Edexcel International Advanced Level paper (WST02/01) • 75 marks

Every question solved in full on the paper itself, the way your teacher would write it out

How to read these solutions. Each page is the real question page, with a full answer written on the answer lines.

Red pen

M1 A1

boxed

the working a full-marks answer shows, with a short note on why each step is taken

the mark earned on that line: M = method, A = accuracy, B = independent mark

the final answer to each part

Sit the paper first. These solutions are most useful after you have tried every question yourself under timed conditions. Mark your paper with the mark scheme, then use this booklet to see exactly how each lost mark could have been earned.

Q	Topic	Marks	Page
1	A binomial model	8	2
2	A census and a sampling distribution	9	3
3	A Poisson model for calls	10	4
4	A rod broken at random	9	5
5	Poisson and normal approximations	11	6
6	Two one-tailed binomial tests	12	7
7	A continuous random variable	16	8

Question paper & mark scheme | Exam questions by topic | More practice papers

1. On any day, the probability that a particular train arrives late is 0.15. A commuter records the number of days, X, on which the train arrives late over a period of 20 days.

(a) State two assumptions, in context, that are needed for X to be modelled by a binomial distribution.

(2)

(b) Find $P(X = 3)$

(2)

(c) Find $P(2 \leqslant X < 6)$

(2)

(d) Find the mean and the variance of X

(2)

(a) Whether the train is late on one day is independent of the other days. B1

The probability that the train is late is the same, 0.15, on every day. B1

(b) $X \sim B(20, 0.15)$: $P(X = 3) = {}^{20}C_3 \times 0.15^3 \times 0.85^{17} = 1140 \times 0.003375 \times 0.06311\ldots$ M1

$= 0.2428\ldots$ $\boxed{0.243}$ A1

(c) $P(2 \leqslant X < 6) = P(X \leqslant 5) - P(X \leqslant 1) = 0.9327 - 0.1756$ M1

$\boxed{0.7571}$ A1

(d) Mean $= np = 20 \times 0.15$ $\boxed{3}$ B1

Variance $= np(1 - p) = 20 \times 0.15 \times 0.85$ $\boxed{2.55}$ B1

(Total for Question 1 is 8 marks)

DO NOT WRITE IN THIS AREA

Join students already acing their exams.
Book your FREE lesson at www.eclassroom.org.uk

2. The owner of a café wants to find out what her customers think of a new menu.

(a) Give one advantage and one disadvantage of asking every customer, rather than a sample of customers.

(2)

(b) Suggest a suitable sampling frame.

(1)

The café sells coffee in three sizes. A small coffee costs £2, a medium coffee costs £3 and a large coffee costs £5. Of the coffees sold, 50% are small, 30% are medium and 20% are large. A random sample of two coffee orders is taken, and T is the total cost, in pounds, of the two coffees.

(c) Find the sampling distribution of T

(5)

(d) Find P($T > 6$)

(1)

(a) Advantage: every customer's opinion is included, so the result is accurate. B1

Disadvantage: it would take a long time and cost a lot to ask every customer. B1

(b) A list of the café's customers, e.g. everyone on its loyalty card scheme. B1

(c) Totals: 2 + 2 = 4, 2 + 3 = 5, 3 + 3 = 6, 2 + 5 = 7, 3 + 5 = 8, 5 + 5 = 10 B1

Mixed pairs can come in either order: $P(T = 5) = 2 \times 0.5 \times 0.3$ M1

$P(T = 4) = 0.5^2 = 0.25$; $P(T = 5) = 0.3$ A1

$P(T = 6) = 0.3^2 = 0.09$; $P(T = 7) = 2 \times 0.5 \times 0.2 = 0.2$ A1

$P(T = 8) = 2 \times 0.3 \times 0.2 = 0.12$; $P(T = 10) = 0.2^2 = 0.04$ (total 1 ✓) A1

(d) $P(T > 6) = 0.2 + 0.12 + 0.04$ $\boxed{0.36}$ B1

(Total for Question 2 is 9 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

3. Calls to a helpline arrive at random at a mean rate of 3 every 10 minutes.

(a) Give two reasons why a Poisson distribution may be a suitable model for the number of calls in a 10-minute period.

(2)

(b) Find the probability that exactly 4 calls arrive in a 10-minute period.

(2)

(c) Find the probability that at least 8 calls arrive in a 30-minute period.

(3)

(d) Find the probability that more than 4 but fewer than 10 calls arrive in a 20-minute period.

(3)

(a) Calls arrive independently of each other, one at a time. B1

Calls arrive at a constant average rate. B1

(b) $X \sim Po(3)$: $P(X = 4) = e^{-3} \times 3^4 \div 4!$ M1

$= 0.1680...$ $\boxed{0.168}$ A1

(c) 30 minutes: $\lambda = 3 \times 3 = 9$, $Y \sim Po(9)$ B1

$P(Y \geqslant 8) = 1 - P(Y \leqslant 7) = 1 - 0.3239$ M1

$\boxed{0.6761}$ A1

(d) 20 minutes: $\lambda = 6$, $W \sim Po(6)$ B1

$P(4 < W < 10) = P(W \leqslant 9) - P(W \leqslant 4) = 0.9161 - 0.2851$ M1

$\boxed{0.6310}$ A1

(Total for Question 3 is 10 marks)

DO NOT WRITE IN THIS AREA

★★★★★ Join students already acing their exams.
Book your FREE lesson at www.eclassroom.org.uk

4.

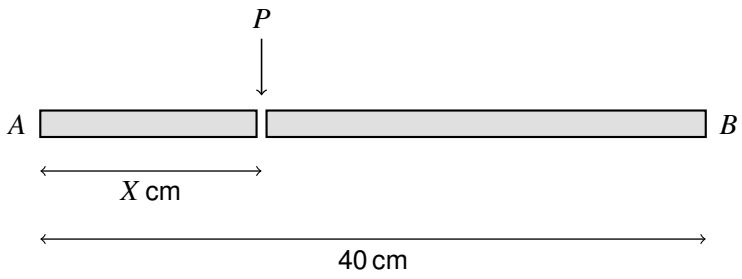


Figure 1

A rod AB of length 40 cm is broken into two pieces at a point P, chosen at random along its length. The random variable X is the length, in cm, of AP, as shown in Figure 1.

(a) Find P$(12 < X < 30)$

(2)

(b) Find E(X) and Var(X)

(3)

(c) Find the probability that the longer piece is more than four times as long as the shorter piece.

(2)

Four such rods are each broken at random, independently.

(d) Find the probability that for exactly one of the four rods the longer piece is more than four times as long as the shorter piece.

(2)

(a) $X \sim U[0, 40]$: $P(12 < X < 30) = (30 - 12) \div 40$ M1

$= 18/40$ 0.45 A1

(b) $E(X) = (0 + 40) \div 2$ 20 B1

$Var(X) = (40 - 0)^2 \div 12 = 1600 \div 12$ M1

$= 133.3...$ 133 A1

(c) Longer > 4 × shorter: the shorter piece is under $40 \div 5 = 8$ cm: $X < 8$ or $X > 32$ M1

$P = (8 + 8) \div 40 = 16/40$ 0.4 A1

(d) Y = number of such rods, $Y \sim B(4, 0.4)$: $P(Y = 1) = 4 \times 0.4 \times 0.6^3$ M1

$= 4 \times 0.4 \times 0.216$ 0.3456 A1

(Total for Question 4 is 9 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

★★★★★
Join students already acing their exams.
Book your FREE lesson at www.eclassroom.org.uk

5. **(i)** A factory makes light bulbs, and 2% of the bulbs are faulty. A box contains 200 bulbs, chosen at random.

(a) Using a suitable approximation, find the probability that the box contains more than 6 faulty bulbs.

(4)

(b) Explain why the approximation you used in part (a) is suitable.

(1)

(ii) In a large town, 40% of the customers of a shop pay by card. A random sample of 150 customers is taken.

(c) Using a suitable approximation, find the probability that more than 66 of the customers in the sample pay by card.

(6)

(a) $X \sim B(200, 0.02)$, n large and p small: use a Poisson approximation. M1

$\lambda = np = 200 \times 0.02 = 4$, so $X \sim Po(4)$ approximately A1

$P(X > 6) = 1 - P(X \leqslant 6) = 1 - 0.8893$ M1

$\boxed{0.1107}$ A1

(b) n = 200 is large and p = 0.02 is small ($np = 4$ is small). B1

(c) $Y \sim B(150, 0.4)$: $np = 60$, $np(1-p) = 36$. Normal approximation $W \sim N(60, 6^2)$ M1 A1

Continuity correction: $P(Y > 66) = P(Y \geqslant 67) \approx P(W > 66.5)$ M1

$z = (66.5 - 60) \div 6 = 1.0833... = 1.08$ (2 d.p.) M1

$P(Z > 1.08) = 1 - 0.8599$ A1

$\boxed{0.1401}$ A1

(Total for Question 5 is 11 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

6. A seed supplier claims that 30% of a type of plant flower in their first year. A gardener believes that the proportion is less than 30%. She plants 20 of the seeds and records the number, X, of plants that flower in their first year.

(a) Write down suitable hypotheses to test the gardener's belief.

(1)

(b) Using a 5% level of significance, find the critical region for a test of these hypotheses.

(3)

(c) Write down the actual significance level of this test.

(1)

In her sample, 3 of the plants flower in their first year.

(d) Comment on the gardener's belief, giving a reason for your answer.

(2)

The supplier now sells a new fertiliser and claims that, when it is used, more than 30% of the plants flower in their first year. The gardener uses the fertiliser on 25 plants and 12 of them flower in their first year.

(e) Test the supplier's claim at the 5% level of significance. State your hypotheses clearly.

(5)

(a) $H_0: p = 0.3$, $H_1: p < 0.3$ B1

(b) Under H_0, $X \sim B(20, 0.3)$. Lower tail, 5%: M1

$P(X \leqslant 2) = 0.0355 < 0.05$, but $P(X \leqslant 3) = 0.1071 > 0.05$ M1

Critical region $X \leqslant 2$ A1

(c) Actual significance level = $P(X \leqslant 2)$ 0.0355 B1

(d) 3 is not in the critical region, so do not reject H_0. M1

There is insufficient evidence to support the gardener's belief that fewer than 30% flower. A1

(e) $H_0: p = 0.3$, $H_1: p > 0.3$ B1

Under H_0, $Y \sim B(25, 0.3)$: $P(Y \geqslant 12) = 1 - P(Y \leqslant 11)$ M1

$= 1 - 0.9558 = 0.0442$ A1

$0.0442 < 0.05$, so the result is significant: reject H_0. M1

There is evidence to support the claim that more than 30% flower with the fertiliser. A1

(Total for Question 6 is 12 marks)

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

★★★★★
Join students already acing their exams.
Book your FREE lesson at www.eclassroom.org.uk

7. The continuous random variable X has probability density function

$$f(x) = \begin{cases} kx^2(3-x) & 0 \leqslant x \leqslant 3 \\ 0 & \text{otherwise} \end{cases}$$

where k is a constant.

(a) Show that $k = \frac{4}{27}$

(3)

(b) Sketch the graph of $f(x)$ for $0 \leqslant x \leqslant 3$

(2)

(c) Find the cumulative distribution function $\mathrm{F}(x)$ for all values of x

(4)

(d) Show that the median, m, of X satisfies $2m^4 - 8m^3 + 27 = 0$

(2)

(e) Show that the median of X lies between 1.8 and 1.9

(2)

(f) Use calculus to find the mode of X

(2)

(g) Describe the skewness of the distribution, giving a reason for your answer.

(1)

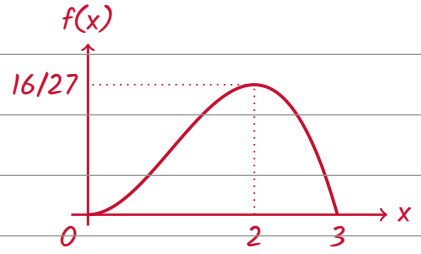
(a) Total probability is 1: $\int$ from 0 to 3 of $k(3x^2 - x^3)\,dx = 1$ M1

$k[x^3 - x^4/4]$ from 0 to 3 $= 1$ A1

$k(27 - 81/4) = 27k/4 = 1 \Rightarrow k = 4/27$ ✓ A1*

(b) $f(0) = 0$, $f(3) = 0$; turning point at $x = 2$, where $f(2) = (4/27) \times 4 \times 1 = 16/27$.

B1 B1



(c) For $0 \leqslant x \leqslant 3$: $F(x) = \int$ from 0 to x of $(4/27)(3t^2 - t^3)\,dt$ M1

$= (4/27)(x^3 - x^4/4) = (4x^3 - x^4)/27$ A1

$F(x) = 0$ for $x < 0$, and $F(x) = 1$ for $x > 3$ B1

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

Question 7 continued

$F(x) = 0\ (x < 0);\ (4x^3 - x^4)/27\ (0 \leqslant x \leqslant 3);\ 1\ (x > 3)$ A1

(d) Median: $F(m) = 0.5 \Rightarrow (4m^3 - m^4)/27 = 1/2$ M1

$8m^3 - 2m^4 = 27 \Rightarrow 2m^4 - 8m^3 + 27 = 0$ ✓ A1*

(e) $g(m) = 2m^4 - 8m^3 + 27$: $g(1.8) = 1.339...$, $g(1.9) = -1.808...$ M1

Change of sign (g is continuous), so the median lies between 1.8 and 1.9 ✓ A1

(f) $f'(x) = (4/27)(6x - 3x^2) = (4/27) \times 3x(2 - x) = 0$ M1

$x = 0$ is the end of the range, so mode = 2 A1

(g) median (between 1.8 and 1.9) < mode (2): negative skew B1

(Total for Question 7 is 16 marks)

TOTAL FOR PAPER IS 75 MARKS

END

DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA
DO NOT WRITE IN THIS AREA

★★★★★
Join students already acing their exams.
Book your FREE lesson at www.eclassroom.org.uk