

1. A total of 228 items are collected from an archaeological site. The distance from the centre of the site is recorded for each item. The results are summarised in the table below.

Distance from the centre of the site (m)	0–1	1–2	2–4	4–6	6–9	9–12
Number of items	22	15	44	37	52	58

Test, at the 5% level of significance, whether or not the data can be modelled by a continuous uniform distribution. State your hypotheses clearly.

(Total 12 marks)

1.

Distance from centre of site (m)	0–1	1–2	2–4	4–6	6–9	9–12	
$b - a$	1	1	2	2	3	3	M1
No of artefacts	22	15	44	37	52	58	
$P(a \leq X < b)$	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{4}$	$\frac{1}{4}$	A1
$228 \times P(a \leq X < b)$	19	19	38	38	57	57	A1

Class	O_i	E_i	$\frac{(O_i - E_i)^2}{E_i}$	$\frac{O_i^2}{E_i}$	
0–1	22	19	$\frac{9}{19} = 0.4736\dots$	25.57...	
1–2	15	19	$\frac{16}{19} = 0.8421\dots$	11.84...	M1
2–4	44	38	$\frac{36}{38} = 0.9473\dots$	50.94...	
4–6	37	38	$\frac{1}{38} = 0.0263\dots$	36.02...	
6–9	52	57	$\frac{25}{57} = 0.4385\dots$	47.43...	A1
9–12	58	57	$\frac{1}{57} = 0.0175\dots$	59.01...	

H_0 : continuous uniform distribution is a good fit B1

H_1 : continuous uniform distribution is not a good fit

$$\sum \frac{(O_i - E_i)^2}{E_i} = \frac{313}{114} = 2.75$$

or

$$\sum \frac{O_i^2}{E_i} - 228 = 230.745... - 228 = ... \quad (\text{awrt } \mathbf{2.75}) \quad \text{dM1 A1}$$

$$\nu = 6 - 1 = 5 \quad \text{B1}$$

(ft their ν i.e. $\chi^2(0.05)$) B1ft

$$\chi^2_5(0.05) = 11.070$$

$2.75 < 11.070$, insufficient evidence to reject H_0 M1

Continuous uniform distribution is a suitable model A1

-

Note

1st M1 for calculation of at least 3 widths and attempting proportions/probs.
or for 1:2:3 ratio seen

1st A1 for correct probabilities

2nd A1 for all correct expected frequencies

2nd M1 for attempting $\frac{(O-E)^2}{E}$ or $\frac{O^2}{E}$, at least 3 correct
expressions or values.

Follow through their E_i provided they are not all = 38

3rd A1 for a correct set of calcs – 3rd or 4th column. (2 dp or better and
allow e.g. 0.94...)

3rd dM1 **dependent on 2nd M1** for attempting a correct sum or calculation
(must see at least 3 terms and +)

The first three Ms and As can be implied by a test statistic of awrt 2.75

4th M1 for a correct statement based on their test statistic (> 1) and
their cv (> 3.8) Contradictory statements score M0 e.g.
“significant” do not reject H_0 .

5th A1 for a correct comment suggesting that continuous uniform model is
suitable. No ft

[12]

1. Some of the weaker candidates assumed that the expected frequencies would all equal 38 and they did not score many marks. Most though handled the unequal class widths correctly and were able to calculate a correct test statistic. Some thought the degrees of freedom should be 4 not 5 but for many candidates this was another good source of marks.