Further Math with Ali Hashir

Chi-Squared Tests

Topical Mark Schemes (CIE 9231)

Mark schemes corresponding to the topical questions booklet for this topic, compiled from official CIE 9231 marking schemes (2015–2025). Question numbers match the questions booklet.

About this resource: I teach CIE A Level Further Mathematics, and over the years I found that there simply weren't many well-organised, topic-wise resources available for this subject – it's a niche A Level with relatively few candidates, so it doesn't get much attention from resource makers. These topical booklets are my attempt to fix that: every past-paper question, sorted by topic, with its official mark scheme alongside it.

Help us improve these documents for future students. If you spot a mistake, email us at **vectorspacepk@gmail.com**.

What kind of mistakes might you find? These booklets are compiled by hand from hundreds of past-paper PDFs. I've checked them as carefully as I can, but I'm only human, and the odd error may have slipped through – so please help me catch what I missed. Things to look out for:

- Spelling or typing mistakes
- Wrong final answers
- A mark scheme that doesn't line up with its question (misaligned pages)
- A mark scheme that belongs to the wrong question entirely
- Pre-2020 questions that are technically out of the current syllabus (the syllabus changed in 2020, but some older questions are still included here for practice)
- Missing, extra, or wrongly-labelled marks (e.g. an M1 or A1 that shouldn't be there, or one that's missing)
- Diagrams or images that are unclear, missing, or don't match the question
- Questions placed under the wrong topic

If you spot any of these (or anything else that looks off), please let me know.

This is **Version 1** of this booklet. As readers point out errors, I'll keep correcting and improving these documents, so I'd recommend checking back on the website from time to time and using the latest version rather than relying on a saved copy.

If you think these documents brought you value, consider following me on social media – it helps more than you'd think.

Contents

Click a question number below to jump straight to its mark scheme. Every box is a live link.

For the best hyperlink experience, open this PDF in a dedicated reader – such as Adobe Acrobat, Apple Preview, or your browser's built-in PDF viewer – rather than a basic in-app or embedded preview, where links sometimes don't work.

Q1

9231/23 • May/June 2015 • Question 6

↑ Back to contents

Mark scheme.

State (at least) null hypothesis (AEF): H_0: reliability is independent of supplier.	**B1**
Find expected values (to 1 d.p.): 62.253, 57.423, 41.323, 53.747, 49.577, 35.677 (lose A1 if rounded to integers).	**M1 A1**
Calculate χ^2: $\chi^2 = 0.1212 + 0.5416 + 1.6764 + 0.1404 + 0.6274 + 1.9417 = 5.05$ (allow 5.03 if 1 d.p. expected values used).	**M1 A1**
State or use correct tabular χ^2 value (to 3 s.f.): $\chi^2_{2,0.95} = 5.99[1]$.	**B1**
Valid method for reaching conclusion: accept H_0 if $\chi^2 <$ tabular value.	**M1**
Correct conclusion, from correct values (AEF): reliability is independent of supplier.	**A1**
	Total: 8

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q2

9231/21, 9231/22, 9231/23 • October/November 2015 • Question 8

↑ Back to contents

Mark scheme.

Mean of sample data $\lambda = 220/100 = 2.2$.	**B1**
State null hypothesis: H_0: Poisson distribution fits the data ($\lambda = 2.2$).	**B1**
Expected values $100\lambda^r e^{-\lambda}/r!$: 11.080, 24.377, 26.814, 19.664, 10.815, 4.759, 2.491.	**M1A1**
Combine last two cells so expected value ≥ 5: $O = 3$, $E = 7.25$.	**M1**
Calculate $\chi^2 = 0.076 + 2.879 + 0.653 + 1.448 + 0.441 + 2.491 = 7.99$.	**M1A1**
Tabular value (6 cells, 4 d.f.): $\chi^2_{4,0.95} = 9.488$.	**B1**
Valid method for conclusion: accept H_0 if $\chi^2 <$ tabular value.	**M1**
Conclusion: distribution fits the data ($\lambda = 2.2$).	**A1**
	Total: 10

Q3

9231/21, 9231/22 • May/June 2016 • Question 9

↑ Back to contents

Mark scheme.

State (at least) null hypothesis in full: H_0: Given distribution fits data	**B1**
Find expected values using $150\,{}^6C_i q^{6-i}p^i$ with $p = 0.6$, $q = 0.4$: 0.6144, 5.5296, 20.736, 41.472, 46.656, 27.9936, 6.9984	**M1 A2**
Combine 0.6144 & 5.5296 since 1st expected value < 5: $O_i : 4\ldots$, $E_i : 6.144\ldots$	**B1**
Calculate χ^2 (to 3 s.f.): $\chi^2 = 0.748+0.877+2.189+1.606+0.144+3.570 = 9.13$	**M1 A1**
State or use consistent tabular value: 6 cells: $\chi^2_{5,0.95} = 11.07$	**B1**
State or imply valid method for conclusion: accept H_0 if $\chi^2 <$ tabular value	**M1**
Conclusion: $9.13 < 11.1$ so distribution fits data	**A1**
	Total: 10

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q4

9231/23 • May/June 2016 • Question 9

↑ Back to contents

Mark scheme.

State (at least) null hypothesis: H_0: No difference in preferences	**B1**
Find expected values (to 1 d.p.): 21.867, 18.133, 29.333, 10.667, 19.133, 15.867, 25.667, 9.333	**M1 A2**
Calculate value of χ^2 (to 3 s.f.): $\chi^2 = 0.449 + 2.075 + 0.015 + 0.510 + 0.513 + 2.371 + 0.017 + 0.583 = 6.53$	**M1 A1**
State or use correct tabular χ^2 value: $\chi^2_{3,0.9} = 6.251$	**B1**
State or imply valid method for conclusion: reject H_0 if $\chi^2 \geq$ tabular value	**M1**
Correct conclusion: difference in preferences	**A1**
	Total: 9

Q5

9231/21, 9231/22, 9231/23 • October/November 2016 • Question 9

↑ Back to contents

Mark scheme.

(i)	
Find mean of sample data: $\bar{x} = 240/60 = 4$	**B1**
Find variance: $\sigma^2 = 1174/60 - 4^2 = 3.57$	**B1**
State valid reason why Poisson distribution suitable (AEF): $4 \approx 3.57$ (no $\surd$ on $\bar{x}, \sigma^2$) (allow "unsuitable since $4 \neq 3.57$")	**B1**
(ii)	
Find expected values $60\lambda^r e^{-\lambda}/r!$ with $\lambda = 4$: 4.40; 11.7 (to 1 d.p.)	**B1; B1**
(iii)	
State (at least) null hypothesis (AEF): H_0: Poisson distribution fits data	**B1**
Combine cells so that all expected values $\geqslant 5$: O_i : 4 12 … 6 8; E_i : 5.50 8.79 … 6.25 6.64	***M1 A1**
Calculate value of χ^2 (to 2 d.p.; A1 dep *M1): $\chi^2 = 0.409 + 1.172 + 1.181 + 0.044 + 1.397 + 0.81 + 0.279 = 5.29$	**M1 DA1**
State or use consistent tabular value (to 2 d.p.): 7 cells: $\chi^2_{5,0.9} = 9.236$ (8 cells: 10.64; 9 cells: 12.02; 10 cells: 13.36)	**B1✓**
State or imply valid method for conclusion: accept H_0 if $\chi^2 <$ tabular value	**M1**
Conclusion (AEF, requires both values correct): $5.29 < 9.24$ so distribution fits	**A1**
	Total: 13

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q6

9231/23 • May/June 2017 • Question 10

↑ Back to contents

Mark scheme.

Verify given mean (B0 for $\bar{x} = 325/100 = 3.25$ without working): $\bar{x} = (1/100)\Sigma x\,\mathrm{f}(x) = 325/100 = 3.25$ (AG). **(ii)**	**B1**
State expression for required expected value E_3 (M1 for E_3 or E_6): $E_3 = 100\lambda^3\mathrm{e}^{-\lambda}/3!$ with $\lambda = 3.25$.	**M1,A1**
Find $E_6 = 100\lambda^6\mathrm{e}^{-\lambda}/6! = 6.35$.	**A1**
Find $E_{\geqslant 7} = 100 - \Sigma_0^6 E_i = 4.77$. **(iii)**	**A1**
State (at least) null hypothesis in full: H_0: distribution fits data (AEF).	**B1**
Combine values consistent with all expected values $\geqslant 5$ (FT on $E_6, E_{\geqslant 7}$): O_i: 13, 18, 26, 20, 16, 7; E_i: 16.48, 20.48, 22.18, 18.02, 11.72, 11.12 (combining first two and last two cells).	**M1,FT**
Find $\chi^2 = 0.735 + 0.300 + 0.658 + 0.218 + 1.563 + 1.527$.	**M1**
$= 5.00$.	**A1**
State or use consistent tabular value $\chi^2_{n-2,0.95}$ (FT on number n of cells): using $n = 6$, $\chi^2_{4,0.95} = 9.488$.	**B1,FT**
State or imply valid method for conclusion: accept H_0 if $\chi^2 <$ tabular value: $5.00\,[\pm 0.1] < 9.49$ so distribution fits [data].	**M1**
or distribution is a suitable model (AEF).	**A1**
	Total: 12

Q7

9231/21, 9231/22, 9231/23 • October/November 2017 • Question 8

↑ Back to contents

Mark scheme.

State (at least) the null hypothesis: H_0: No association (method is independent of gender).	**B1**
Find expected values E_i: 9.0, 19.8, 31.2 (male row); 6.0, 13.2, 20.8 (female row).	**M1, A1**
Find value of χ^2 from $\sum(E_i - O_i)^2/E_i$: $\chi^2 = 0.111 + 3.073 + 1.482 + 0.167 + 4.610 + 2.223 = 11.7$.	**M1, A1**
State or use correct tabular value: $\chi^2_{2,0.99} = 9.21$.	**B1**
Valid method for reaching conclusion: reject H_0 if $\chi^2 >$ tabular value.	**M1**
$11.7 > 9.21$, so there is an association between method of voting and gender.	**A1**
	Total: 8

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q8

9231/21, 9231/22 • May/June 2018 • Question 8

↑ Back to contents

Mark scheme.

State null hypothesis: colour is independent of type (no association)	**B1**
Find expected values E_i: 47.67, 43.33, 39; 36.67, 33.33, 30; 25.67, 23.33, 21 (3 s.f.)	**M1 A1**
Find $\chi^2 = \sum(E_i - O_i)^2/E_i = 0.597 + 1.241 + 0.103 + 1.603 + 1.333 + 0.033 + 0.212 + 0.019 + 0.429 = 5.57$ (allow 5.64 if integer E_i used)	**M1 A1**
State or use correct tabular value: $\chi^2_{4,0.9} = 7.779$	**B1**
Valid method for reaching conclusion: accept H_0 if $X^2 <$ tabular value	**M1**
$5.57 < 7.78$, so independent (no association)	**A1**
	Total: 8

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q9

9231/22 • October/November 2018 • Question 10

↑ Back to contents

Mark scheme.

(i)

$\bar{x} = \Sigma x\,\mathrm{f}(x) = 118/40 = 2.95$ AG (verify mean of sample data, B0 for $\bar{x} = 118/40 = 2.95$)	**B1**
$\sigma_n^2 = (454 - 118^2/40)/40 = 2.65$ or $\sigma_{n-1}^2 = (454 - 118^2/40)/39 = 2.72$ (find variance of sample data, accept either)	**B1**
$2.95 \approx 2.65$ or 2.72 (valid comment on correct values)	**B1**

(ii)

$E_4 = 40\lambda^4 \mathrm{e}^{-\lambda}/4!$ with $\lambda = 2.95$ [$= 40 \times 0.1652$] (allow $(\lambda/4) \times E_3 = 0.7375 \times 8.96$) (state expression for required expected value E_4)	**M1A1**

(iii)

H_0: [Poisson] distribution fits data (AEF) (state at least null hypothesis in full)	**B1**
Combine values consistent with all expected values $\geqslant 5$: O_i : $8, 8, 10, 5, 9$; E_i : $8.27, 9.11, 8.96, 6.61, 7.05$	**M1**
$X^2 = 0.0088 + 0.1352 + 0.1207 + 0.3921 + 0.5394$ (find value of X^2 from $\Sigma(E_i - O_i)^2/E_i$ or $\Sigma O_i^2/E_i - n$)	**M1**
$= 1.20$ (to 3 s.f.)	**A1**
No. n of cells: 5; $\chi^2_{n-2,0.95} = 7.815$ (state or use consistent tabular value $\chi^2_{n-2,0.95}$ to 3 s.f., FT on number n of cells used to find X^2)	**B1**
Accept H_0 if $X^2 <$ tabular value (AEF) (state or imply valid method for conclusion)	**M1**
$1.20[\pm 0.1] < 7.81[5]$ so [Poisson] distn. fits [data] or distn. is a suitable model (AEF) (conclusion requires both values correct)	**A1**

Total: 12

Q10

9231/21, 9231/22 • May/June 2019 • Question 9

↑ Back to contents

Mark scheme.

State or imply expression for required expected value E_3 of X: $E_3 = (3/16)\int_{1.6}^{2.4}(4-x)^{1/2}\,\mathrm{d}x = (3/16)[-(2/3)(4-x)^{3/2}]_{1.6}^{2.4}$	**M1**
Find expected value E_3 (may be implied in finding $50E_3$): $= (2.4^{3/2} - 1.6^{3/2})/8 = 1.694/8 = 0.2118$	**A1**
Hence verify corresponding expected frequency: $50E_3 = 10.59$ (AG) (ii)	**A1**
State (at least) null hypothesis in full: H_0: distribution fits data	**B1**
Combine values consistent with all expected values $\geqslant 5$: O_i : $18, 16, 8, 8$; E_i : $14.22, 12.54, 10.59, 12.65$	**M1**
Find value of X^2 from $\Sigma(E_i - O_i)^2/E_i$: $X^2 = 1.005+0.955+0.633+1.709 = 4.30$	**M1 A1**
FT on number, n, of cells used to find X^2; state or use consistent tabular value: $\chi^2_{4-1,0.95} = 7.815$	**B1**
State or imply valid method for conclusion: accept H_0 if $X^2 <$ tabular value	**M1**
Conclusion (requires both values approx. correct): $4.30 < 7.815$ so distribution fits data	**A1**
	Total: 10

Q11

9231/23 • May/June 2019 • Question 8

↑ Back to contents

Mark scheme.

State (at least) null hypothesis: H_0: holidays are independent of salesman (no association)	**B1**
Find expected values (lose A1 if rounded to integers): E_i : 29.68, 33.04, 21.28, 23.32, 25.96, 16.72	**M1 A1**
Find value of X^2 from $\Sigma(E_i - O_i)^2/E_i$: $X^2 = 0.7380 + 0.7446 + 0.0037 + 0.9392 + 0.9477 + 0.0047$	**M1**
$= 3.38$	**A1**
State or use correct tabular χ^2 value: $\chi^2_{2,0.9} = 4.605$	**B1**
Compare their calculated value with their X^2 value and give appropriate conclusion: accept H_0 if $X^2 <$ tabular value	**M1**
Correct conclusion: type of holidays is independent of salesman	**A1**
	Total: 8

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q12

9231/41, 9231/42 • May/June 2020 • Question 1

↑ Back to contents

Mark scheme.

Expected frequencies: $49 \to 56.8$, $51 \to 47.6$, $40 \to 35.6$, $93 \to 85.2$, $68 \to 71.4$, $49 \to 53.4$.	**M1A1**
$\sum \frac{(O-E)^2}{E} = \frac{(49-56.8)^2}{56.8} + \frac{(51-47.6)^2}{47.6} + \frac{(40-35.6)^2}{35.6} + \frac{(93-85.2)^2}{85.2} + \frac{(68-71.4)^2}{71.4} + \frac{(49-53.4)^2}{53.4}$	**M1**
$= 3.096$ (3.10)	**A1**
Use appropriate tabular value $= 4.605$	**M1**
$3.096 < 4.605$ so the grades awarded are independent of the background	**A1**
	Total: 6

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q13

9231/43 • May/June 2020 • Question 1

↑ Back to contents

Mark scheme.

Expected frequencies: $31 \to 22$, $35 \to 37.4$, $22 \to 28.6$, $19 \to 28$, $50 \to 47.6$, $43 \to 36.4$.	**M1A1**
$\sum \frac{(O-E)^2}{E} = \frac{(31-22)^2}{22} + \frac{(35-37.4)^2}{37.4} + \frac{(22-28.6)^2}{28.6} + \frac{(19-28)^2}{28} + \frac{(50-47.6)^2}{47.6} + \frac{(43-36.4)^2}{36.4}$	**M1**
$= 9.569\ (9.57)$	**A1**
Tabular value $= 5.991$ and comparing $9.569 > 5.991$	**M1**
So standard achieved is dependent on reading scheme used	**A1**
	Total: 6

Q14

9231/41, 9231/43 • October/November 2020 • Question 3

↑ Back to contents

Mark scheme.

Attempt at E values using Bin$(5, 0.35)$: 17.404, 46.858, 50.462, 27.172, 7.316, 0.7878 (at least 4 correct; all correct to 2 dp)	**M1 A1**
Add last two columns: 16, 8.104	**M1**
$\frac{(17.404-12)^2}{17.404} + \frac{(46.858-39)^2}{46.858} + \frac{(50.462-46)^2}{50.462} + \frac{(27.172-37)^2}{27.172} + \frac{(8.104-16)^2}{8.104} = 14.64$ or 14.65 (accept 14.6–14.7)	**M1 A1**
Compare '14.64' > 9.49	**M1**
Freya's claim is not supported / data does not fit the distribution (SC for Poisson M0M1M1M0 max 2)	**A1 FT**
	Total: 7

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q15

9231/42 • October/November 2020 • Question 3

↑ Back to contents

Mark scheme.

M1 $\int_0^{0.5}\left(\frac{4}{9}x-\frac{1}{9}x^2\right)\mathrm{d}x=\frac{1}{9}\left[2\times0.5^2-\frac{1}{3}0.5^3\right]$ $(=0.0509(26)$ or $\frac{11}{216})$ – statement of integration with correct limits	**(a)**
Freq $= p = 0.0509(26)\times200 = 10.19$ AG (requires sight of 0.0509 or $\frac{11}{216}$)	**A1**
By addition to 200, $q = 26.85$ $(\frac{725}{27})$ (or by integration)	**B1**
M1 O: 5, 23, 40, 41, 46, 45; E: 10.19, 26.85, 37.96, 43.52, 43.52, 37.96 (at least 4 correct)	**(b)**
Test statistic $= 2.6433+0.5520+0.1096+0.1459+0.1413+1.3056 = 4.89$ or 4.90	**A1**
'4.90' < 11.07: compare their value with 11.07	**M1**
PDF is a satisfactory model for the data	**A1 FT**
	Total: 7

Q16

9231/41, 9231/42 • May/June 2021 • Question 2

↑ Back to contents

Mark scheme.

Expected values: 72(61.425), 42(44.46), 52(54.405), 68(73.71), 33(43.575), 34(31.54), 41(38.595), 58(52.29) (at least 3 correct to 4sf)	**M1**
All expected values correct to 4sf	**A1**
Test value: $\sum \frac{(O-E)^2}{E} = 1.8206+0.1361+0.1063+0.4423+2.5664+0.1919+0.1499+0.6235$ (correct formula used)	**M1**
$= 6.037$ (correct to 3sf)	**A1**
H_0: (driving) test success is independent of instructor	**B1**
Tabular value: 3 df, 10% = 6.251; $6.04 < 6.251$, accept H_0 (compare with correct tabular value and conclusion)	**M1**
Sufficient evidence to suggest that (driving) test success is independent of instructor (allow ± 1 in third significant figure of test value)	**A1**
	Total: 7

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q17

9231/43 • May/June 2021 • Question 5

↑ Back to contents

Mark scheme.

(a)	
(Mean of distribution = 2.4) Variance = 2.25 (allow 1.5^2)	**M1**
Mean approx. equal to variance (so Poisson might be suitable) (must have 'approximately' OE, not equality; must have 2.25)	**A1**
(b)	
Po(2.5) leads to frequencies: (at least 3 correct to 3sf)	**M1**
14.77[5], 36.938, 46.173, 38.477, 24.048, 12.024, 5.010, 2.554 (all correct to 4sf)	**A1**
Combine last two values: 7.564 (FT their table values if final figure is less than 5)	**M1**
$\sum \frac{(O-E)^2}{E} = 2.6227 + 0.4198 + 0.00065 + 1.4529 + 0.04570 + 1.3148 + 0.7845$ (apply correct formula to their frequencies)	**M1**
= 6.64 (correct to 3sf)	**A1**
H_0: distribution fits the data (must mention data and distribution)	**B1**
Tabular value, 5 degrees of freedom, is 9.236; $6.64 < 9.236$, so accept H_0 (compare with correct tabular value and conclusion)	**M1**
Insufficient evidence to show that data does not follow a Poisson distribution (correct conclusion, in context, following correct work; allow ±1 difference in third significant figure)	**A1**
	Total: 10

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q18

9231/41, 9231/43 • October/November 2021 • Question 3

↑ Back to contents

Mark scheme.

$p = 57.704$ (3 dp or better).	**B1**
$q = 0.035$ or 0.036 (accept numbers which round to this).	**B1**

$p = 57.704$, $q = 0.035$ or 0.036

(b)

Combine frequencies less than 5: last 4 columns give 24/15.7812 (allow last 2 or 3 columns combined).	**M1**
$\sum \dfrac{(O-E)^2}{E} = 0.8444 + 1.6319 + 1.5199 + 4.2803.$	**M1**
8.28 (accept 8.27–8.28).	**A1**
H_0: B(8, 0.15) fits the data. Must mention distribution and data.	**B1**
3 degrees of freedom, tabular value = 7.815: $8.28 > 7.815$, reject H_0.	**M1**
Insufficient evidence to support manager's claim, in context, with level of uncertainty in language.	**A1**

Total: 8

Q19

9231/42 • October/November 2021 • Question 2

↑ Back to contents

Mark scheme.

$\mathrm{P}(130 \leqslant x < 140) = \mathrm{P}\left(\frac{130-125}{12} \leqslant Z < \frac{140-125}{12}\right) = \mathrm{P}(0.4167 < Z < 1.25)$. Must see 0.4167 and 1.25 (or 0.8944 and 0.6616).	**M1**
$0.8944 - 0.6616 = 0.2328$; multiply by 60 leads to 13.97 (AG).	**A1**

Expected frequency = 13.97 (AG)

(b)

Combine first two values.	**M1**
Test stat $= 1.7588 + 0.0761 + 0.6743 + 0.6309 + 0.4352$.	**M1**
3.58 (accept 3.57–3.58).	**A1**
H_0: distribution fits data; $\mathrm{N}(125, 12^2)$ is a good model for the data. Must mention distribution and data.	**B1**
4 degrees of freedom, tabular value = 9.488: $3.58 < 9.488$, accept H_0.	**M1**
Sufficient evidence that the normal distribution fits the data / claim is supported, in context, with level of uncertainty in language.	**A1**

Total: 8

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q20

9231/41, 9231/42 • May/June 2022 • Question 4

↑ Back to contents

Mark scheme.

$p = 16.03$.	**B1**
$q = 1.70$ (condone 1.7). **(b)**	**B1**
Combine last two columns: $O = 1$, $E = 5.04$ (add last 2 columns, or 3 columns: 13.06).	**M1**
Calculate values of $\frac{(O-E)^2}{E}$: 0.4693, 0.8704, 0.8853, 1.5720, 0.5727, 0.5088, 3.2384.	**M1**
Test statistic = 8.12.	**A1**
H_0: data fits a Poisson distribution with mean 2.5; H_1: data does not fit a Poisson distribution with mean 2.5.	**B1**
Critical value of chi-squared = 10.64. Compare: $8.12 < 10.64$, accept H_0 (allow 9.236 if 3 columns combined, or 12.02 if none combined).	**M1**
There is sufficient evidence to support the scientist's claim / data fits a Poisson distribution with mean 2.5 (correct conclusion in context).	**A1**
	Total: 8

Q21

9231/43 • May/June 2022 • Question 2

↑ Back to contents

Mark scheme.

H_0: the size of shell is independent of the location; H_1: the size of shell is not independent of the location.	**B1**
Attempt at calculation of E values: Large row 60.29, 77.87, 100.48, 75.36; Small row 35.71, 46.13, 59.52, 44.64.	**M1, A1**
Calculate values of $\frac{(O-E)^2}{E}$: 0.9859, 1.0104, 0.1997, 0.4221, 1.6646, 1.7055, 0.3372, 0.7126 (accuracy of ± 0.1).	**M1**
Test statistic = 7.04 (correct total to 3sf).	**A1**
Tabular value for 3 df = 6.251. $7.04 > 6.251$, reject H_0 (compare with 6.251, correct FT conclusion).	**M1**
Evidence suggests that size of shell is dependent on (beach) location (correct conclusion in context).	**A1**
	Total: 7

Q22

9231/41, 9231/43 • October/November 2022 • Question 2

↑ Back to contents

Mark scheme.

One of p, q, r correct.	**B1**
Other two correct. ($p = 0.098$, $q = 41.804$, $r = 31.353$) **(b)**	**B1**
H_0: the data fits the binomial distribution B(8, 0.6); H_1: the data does not fit the binomial distribution B(8, 0.6)	**B1**
Combine first 3 columns: $O = 8$, $E = 7.471$; AND last two columns: $O = 11$, $E = 15.956$ (both)	**M1**
Chi-squared sum: $0.0374 + 1.5817 + 2.9655 + 0.8058 + 0.9139 + 1.5394$ (method must be seen)	**M1**
$\chi^2 = 7.84$	**A1**
Critical value, 5 df = 9.236; compare calculated value with 9.236; $7.84 < 9.236$, accept H_0	**M1**
There is insufficient evidence to reject the director's claim.	**A1**
	Total: 8

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q23

9231/42 • October/November 2022 • Question 2

↑ Back to contents

Mark scheme.

H_0: performance is independent of region; H_1: performance is not independent of region	**B1**
Calculate expected frequencies (allow 1 error): 62(53.9), 41(44.1), 44(49), 102(106.7), 94(87.3), 95(97), 56(59.4), 45(48.6), 61(54)	**M1 A1**
Test statistic = 1.217+0.218+0.510+0.207+0.514+0.041+0.195+0.267+0.907	**M1**
$\chi^2 = 4.08$	**A1**
Tabular value, 4 df = 7.779; $4.08 < 7.779$, accept H_0	**M1**
Insufficient evidence that performance is not independent of region.	**A1**
	Total: 7

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q24

9231/41, 9231/42 • May/June 2023 • Question 3

↑ Back to contents

Mark scheme.

$\int_{1.5}^{2} \frac{1}{24}\left(\frac{4}{x^2}+x^2\right)\mathrm{d}x = \frac{1}{24}\left[-\frac{4}{x}+\frac{x^3}{3}\right]$. Integration with correct powers and correct limits seen.	**M1**
$a = 50 \times 0.092014 = 4.6007$ (AG). $\frac{53}{576}$ or $\frac{1325}{288}$ or 0.092014 or 4.60069 seen.	**A1**
$b = 11.4211$. Or 11.421(0). **(b)**	**B1**
Combine first two columns: 6, 9.0278. Must be seen, or implied by 1.0155.	**M1**
Calculate value of chi-squared: $1.0155 + 0.5715 + 0.7661 + 0.2183 + 0.5885$. At least 2 correct values (at least 3sf) or expressions seen. Allow columns not combined or three columns combined.	**M1**
3.16. CWO. Dependent on M1M1 scored. SC B1 for 3.16 with no working.	**A1**
H_0: f is a good fit for the data; H_1: f is not a good fit for the data.	**B1**
Tabular value of chi-squared: 7.78. '3.16' ¡ 7.78 and accept H_0. Correct tabular value: allow correct FT value if columns not combined (9.236) or three columns combined (6.251).	**M1**
Insufficient evidence to suggest that f is not a good fit for the data (condone: sufficient evidence to suggest that f is a good fit for the data). Correct conclusion in context, following correct work, level of uncertainty in language, no contradictions. A0 if hypotheses reversed.	**A1**
	Total: 9

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q25

9231/43 • May/June 2023 • Question 6

↑ Back to contents

Mark scheme.

H_0: grade is independent of age; H_1: grade is not independent of age. Condone 'Ability to remember' instead of 'grade'.	**B1**
Calculate expected values, shown in table in bold: $21.2, 24.4, 14.4$ (row 1), $31.8, 36.6, 21.6$ (row 2). At least 2 correct values or expressions.	**M1**
6 correct values or expressions.	**A1**
Calculate chi-squared values: $0.6811+2.8918+1.4694+0.4541+1.9279+0.9796$. At least 2 correct values (at least 3sf) seen.	**M1**
8.40[4].	**A1**
Tabular value, 2 degrees of freedom = 7.378; '8.404' > 7.378 and reject H_0/significant. Compare their value with 7.378 and conclusion without context. Condone 'accept H_1'.	**M1**
Sufficient evidence to suggest that grade is not independent of age. Correct conclusion in context, level of uncertainty in language, no contradictions. A0 if hypotheses wrong way round. **(b)**	**A1**
Value, 4 degrees of freedom = 11.14; $10.91 < 11.14$, accept H_0/not significant. Condone 'reject H_1'.	**M1**
Insufficient evidence to suggest that grade is not independent of age. CAO. **(c)**	**A1**
For example, result in part (b) because the table contains more information, e.g. more degrees of freedom, more groups, more detail. Any appropriate comment to support part (a) or part (b). Allow 'more specific'. Not 'more data' or 'more accurate' on its own.	**B1**
	Total: 10

Q26

9231/41, 9231/43 • October/November 2023 • Question 2

↑ Back to contents

Mark scheme.

H_0: opinion is independent of street; H_1: opinion is not independent of street.	**B1**
At least 2 correct expected values or expressions.	**M1**
All 6 correct expected values: 32.24, 57.66, 43.4, 21.7, 19.76, 35.34, 26.6, 13.3.	**A1**
Calculate chi-squared values (correct to 3 d.p.): $0.0179 + 0.2323 + 0.0452 + 0.8521 + 0.0292 + 0.3790 + 0.0737 + 1.3902$.	**M1**
$\chi^2 = 3.02$ (SC B1 following M1M0; SC B2 following M0M0).	**A1**
Tabular value, 3 d.f. $= 7.815$: $3.02 < 7.815$, accept H_0 (allow 'not significant').	**M1**
Insufficient evidence to suggest that opinion depends on street (CWO, correct conclusion in context; A0 if hypotheses wrong way round or missing).	**A1**
	Total: 7

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q27

9231/42 • October/November 2023 • Question 2

↑ Back to contents

Mark scheme.

Expected frequency for 3 breakdowns = 19.421; each value (2 marks total).	**B1**
Expected frequency for 6 breakdowns = 6.939. For part (b):	**B1**
H_0: Po(3.5) is a good fit to the data; H_1: Po(3.5) is not a good fit to the data.	**B1**
Combine first 2 columns: $5, 12.23$; combine last 2 columns: $9, 5.876$.	**M1**
Chi-squared values: $4.274 + 0.799 + 0.302 + 0.945 + 1.417 + 0.612 + 1.661$ (allow if columns not combined correctly; at least two correct values/expressions).	**M1**
$\chi^2 = 10.0$ (AWRT; SC B1 for 10.0 if M0 awarded).	**A1**
Tabular value: 10.64. $10.0 < 10.64$, accept H_0/not significant (allow equivalent tabular values 12.02 or 13.36 depending on grouping).	**M1**
Insufficient evidence to suggest that Po(3.5) is not a good fit to the data (correct conclusion in context; A0 if hypotheses wrong way round or missing).	**A1**
	Total: 8

Q28

9231/41, 9231/42 • May/June 2024 • Question 5

↑ Back to contents

Mark scheme.

Expected values: P: (25), (40.625), (59.375); Q: (15), (24.375), (35.625). Calculate E values to at least 2dp, allow one error. These must be seen. For A1, working to at least 2dp.	**M1 A1**
Chi-squared contributions: P: 1.96, 0.1388, 0.3603; Q: 3.2667, 0.2314, 0.6004. At least 2 correct values or expressions seen. Or may be implied by AWRT 6.56.	**M1**
Test statistic = 6.56 (AWRT 6.56).	**A1**
H_0: Quality (of items) is independent of company; H_1: Quality (of items) is not independent of company.	**B1**
$6.56 > 5.991$ Reject H_0. '6.56' from their attempt at chi-squared, 5.991 must be correct. Consistent signs in comparison. FT conclusion from their 6.56.	**M1**
There is sufficient evidence to suggest that the quality of brushes is not independent of company. (b)	**A1**
P produces better quality brushes (than Q) or Q produces worse quality brushes (than P); P has fewer poor quality brushes than expected AND Q has more poor quality brushes than expected.	**B1**
	Total: 8

Q29

9231/43 • May/June 2024 • Question 3

↑ Back to contents

Mark scheme.

Expected values: Early: (17.28), (21.60), (15.12); On time: (39.68), (49.60), (34.72); Late: (23.04), (28.80), (20.16). Calculate expected values, at least 4 correct to 3sf. These must be seen. A1 for all correct to at least 3sf.	**M1 A1**
Contributions: 1.2893, 0.00741, 1.7338, 2.3615, 0.1161, 1.5265, 1.0678, 0.2722, 0.2314. At least 4 correct values (to 2sf) or expressions. May be implied by 8.600–8.615.	**M1**
Test statistic = 8.61 (8.600–8.615). Unsupported value scores M0A0 M1A1.	**A1**
H_0: Reliability of buses is independent of bus company; H_1: Reliability of buses is not independent of/dependent on bus company.	**B1**
'8.61' $<$ 9.488 Accept H_0. '8.61' from an attempt at chi-squared, 9.488 must be correct. FT conclusion from their 8.61.	**M1**
There is insufficient evidence to support/suggest that reliability of buses is dependent/not independent of bus company.	**A1**
	Total: 7

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q30

9231/41, 9231/43 • October/November 2024 • Question 3

↑ Back to contents

Mark scheme.

(a)	
One of 17.098, 8.122, 3.086 correct.	**B1**
All three correct.	**B1**
(b)	
Combine last two or three columns: classes $0, 1, 2, 3, 4$ or more with observed $10, 18, 35, 21, 16$ and expected $14.957, 28.418, 26.997, 17.098, 12.53$.	**M1**
Contributions to test statistic: $1.6428, 3.8192, 2.3724, 0.8905, 0.9609(7)$ (may be implied by awrt 9.69).	**M1**
Test statistic is 9.69 (9.686).	**A1**
H_0: Po(1.9) is a good fit for the data; H_1: Po(1.9) is not a good fit for the data.	**B1**
Critical value is 7.779 (4 degrees of freedom); compare $9.69 > 7.779$, reject H_0.	**M1**
Sufficient evidence to suggest that Po(1.9) is not a good fit for the data / sufficient evidence to reject the statistician's claim.	**A1**
	Total: 8

Q31

9231/42 • October/November 2024 • Question 3

↑ Back to contents

Mark scheme.

(a)	
$\bar{x} = \dfrac{40+86+105+64+20}{150} = \dfrac{315}{150} = 2.1$, $p = \dfrac{2.1}{5} = 0.42$ (AG; must see either 315 or 2.1).	**B1**
(b)	
Expected frequencies: 9.846, 35.6475, 51.627, 37.3845, 13.536, 1.9605 – at least 2 correct to 2 dp.	**B1**
At least 4 correct to 2 dp.	**B1**
Combine last two columns (giving 20, 15.50).	**M1**
Chi-squared contributions: $0.4712, 0.5314, 1.4416, 0.1521, 1.3088$ (at least 2 correct).	**M1**
Test statistic = 3.905 (accept 3.90–3.910).	**A1**
H_0: Binomial B$(5, 0.42)$ fits the data; H_1: Binomial B$(5, 0.42)$ does not fit the data.	**B1**
Critical value is 6.251 (must come from combined columns; allow 7.779).	**B1**
$3.905 < 6.251$, accept H_0.	**M1**
Insufficient evidence to suggest that B$(5, 0.42)$ is not a good fit to the data.	**A1**
	Total: 10

Q32

9231/41, 9231/42 • May/June 2025 • Question 3

↑ Back to contents

Mark scheme.

B1 $\dfrac{0\times 1844+1\times 143+2\times 11+3\times 0+4\times 1+5\times 0+6\times 1}{2000\times 6}=\dfrac{175}{12000}$, AWRT 0.01458. **(a)**

B1 $a=2000-(1831.3+6.016+0.119+0.001)=162.6$ (AG, shown convincingly). [Alt: $a=\binom{6}{1}\times\frac{7}{480}\times(1-\frac{7}{480})^5\times 2000=162.6$] **(b)**

B1 H_0: A binomial distribution is a satisfactory model for the data; H_1: it is not (must mention binomial and data). **(c)**

M1 Correct expression combining cells (e.g. combining last 4 cells): $\dfrac{(1844-1831.3)^2}{1831.3}+\dfrac{(143-162.6)^2}{162.6}+\dfrac{(13-6.136)^2}{6.136}$.

A1 $=10.1$ (AWRT 10.1; allow AWRT 10.3 if using 6.1).

M1 Compare their 10.1 with critical value 7.879 (1 d.f.) and appropriate conclusion.

A1 $10.1>7.879$, reject H_0; sufficient evidence to suggest that the manager's model is not satisfactory.

B1 Not independent, or probability of "success" may not be constant, or plausible context reason (e.g. boxes being dropped causing more multiple breakages). **(d)**

Total: 8

Q33

9231/43 • May/June 2025 • Question 1

↑ Back to contents

Mark scheme.

H_0: uniform distribution fits the data; H_1: it does not (must mention data and uniform distribution).	**B1**
Expected frequencies are 40, 40, 40.	**B1**
$\frac{(38-40)^2}{40} + \frac{(52-40)^2}{40} + \frac{(30-40)^2}{40}$ $[= 0.1 + 3.6 + 2.5]$.	**M1**
$= 6.2$.	**A1**
Compare their 6.2 with critical value 5.991 (2 d.f.) and appropriate conclusion.	**M1**
$6.2 > 5.991$, reject H_0; sufficient evidence to suggest the scientist's claim is incorrect (eye colour is not uniformly distributed).	**A1**
	Total: 6

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q34

9231/44 • May/June 2025 • Question 3

↑ Back to contents

Mark scheme.

H_0: type of item sold is independent of salesperson; H_1: not independent.	**B1**
Calculate expected values (must be seen); M1 for at least 4 correct to 3sf.	**M1**
All expected values correct to 3sf or better: (28.5, 41.8, 24.70; 29.1, 42.68, 25.22; 17.4, 25.52, 15.08).	**A1**
At least 2 terms of correct form seen: $\sum \dfrac{(O-E)^2}{E}$.	**M1**
Test statistic = 3.67 (or 3.68).	**A1**
Compare their 3.67 with critical value 7.779 (4 d.f.) and appropriate conclusion.	**M1**
$3.67 < 7.779$, accept H_0 (not significant); insufficient evidence to suggest that type of item sold and salesperson are not independent.	**A1**
	Total: 7

Q35

9231/41, 9231/43 • October/November 2025 • Question 2

↑ Back to contents

Mark scheme.

$p = \dfrac{e^{-2.8}(2.8)^3}{3!} \times 100 = 22.248$ (or by subtraction; allow 22.2 or 22.3).	**B1**
$q = \dfrac{(29 - 22.25)^2}{22.25} = 2.05$ (AWRT 2.05). (b)	**B1**
H_0: Po(2.8) fits the data; H_1: Po(2.8) does not fit the data.	**B1**
Sum of chi-square contributions using their q: $2.739 + 0.241 + 2.152 + 2.049 + 0.425 + 3.753 + 0.037 = 11.396$.	**M1**
Compare $11.4 < 12.59$, accept H_0 / not significant.	**M1**
Insufficient evidence to suggest that Po(2.8) is not a good fit to the data / that the manager's claim is false.	**A1**
	Total: 6

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Q36

9231/42 • October/November 2025 • Question 3

↑ Back to contents

Mark scheme.

$m = 69.522$.	**B1**
$n = 0.156$ or 0.157. (b)	**B1**
H_0: Po(0.7) fits the data; H_1: Po(0.7) does not fit the data.	**B1**
Combine last 3 columns to give 13 and 6.828 (or 6.829).	**M1**
$\frac{(88-99.317)^2}{99.317} + \frac{(73-69.552)^2}{69.552} + \cdots + \frac{(13-6.829)^2}{6.829}$.	**M1**
7.15 or 7.16.	**A1**
Compare $7.15 < 7.815$, accept H_0 / not significant.	**M1**
Insufficient evidence to reject expert's claim / to suggest that Po(0.7) is not a good fit/model for the data.	**A1**
	Total: 8

Q37

9231/44 • October/November 2025 • Question 5

↑ Back to contents

Mark scheme.

H_0: performance is independent of age; H_1: performance is not independent of age.	**B1**
Expected frequencies: 34(27.93), 41(44.69), 6(8.38), 16(22.07), 39(35.31), 9(6.62) (at least 3 correct to 3 s.f.).	**M1**
All correct to 3 s.f.	**A1**
$\frac{(34-27.93)^2}{27.93}+\frac{(41-44.69)^2}{44.69}+\cdots+\frac{(9-6.62)^2}{6.62}$.	**M1**
AWRT 5.21.	**A1**
Compare $5.21 > 4.605$, reject H_0 / significant.	**M1**
Sufficient evidence to suggest that performance and age are not independent (dependent/associated); supports the instructor's belief.	**A1**
	Total: 7

Spotted a mistake? Email us at **vectorspacepk@gmail.com**.

Help us improve these documents for future students. If you spot a mistake, email us at **vectorspacepk@gmail.com**.

Written by Ali Hashir.