

χ^2 -tests

A2

Read online at pastopic.com/cambridge/further-mathematics-9231/notes/chi-squared-tests

Last checked 1 October 2026

Read first: [Continuous random variables](#)

© 2026 Pastopic. Free for personal study. Please share the link, not the file.

What you need to know

By the end of this topic you should be able to:

1. **Fit a distribution to data.** Given a hypothesis about how the data should behave (equally likely outcomes, outcomes in a given ratio, a binomial, Poisson, normal or other named distribution, or a given probability density function), work out the **expected frequencies** it predicts for each class. Sometimes you estimate the distribution's parameter from the data first. The calculations are kept short, and the table is often partly filled in for you.
2. **Carry out a χ^2 goodness of fit test:** combine classes so that every expected frequency is at least 5, work out the test statistic, choose the right number of **degrees of freedom**, compare with the χ^2 table and conclude in context.
3. **Carry out a χ^2 test for independence in a contingency table:** find the expected frequency of every cell, combine rows or columns if any expected frequency is below 5, use $(r - 1)(c - 1)$ degrees of freedom and conclude in context.

Yates' correction is not in the syllabus, so 2×2 tables are tested exactly like bigger ones. You are expected to know from AS & A Level Mathematics (9709): the binomial, Poisson and normal distributions, and the language of hypothesis tests.

Understand it

Observed and expected frequencies

A spinner has four equal sectors, A, B, C and D. If it is fair, each sector should come up a quarter of the time. It is spun 80 times:

Sector	A	B	C	D
Observed, O	14	26	17	23
Expected if fair, E	20	20	20	20

The **observed frequencies** O never match the **expected frequencies** E exactly, even for a perfectly fair spinner, because samples vary. The question is whether the gaps are small enough to be chance, or so big that the model "the spinner is fair" is hard to believe. A χ^2 test measures the gaps with one number.

The test statistic

For each class, square the gap and divide by what you expected:

$$X^2 = \sum \frac{(O - E)^2}{E}.$$

- Squaring stops positive and negative gaps cancelling.
- Dividing by E makes a gap count for more when little was expected: being 6 out when you expected 20 matters more than being 6 out when you expected 500.

For the spinner: $\frac{36}{20} + \frac{36}{20} + \frac{9}{20} + \frac{9}{20} = 1.8 + 1.8 + 0.45 + 0.45 = 4.5$.

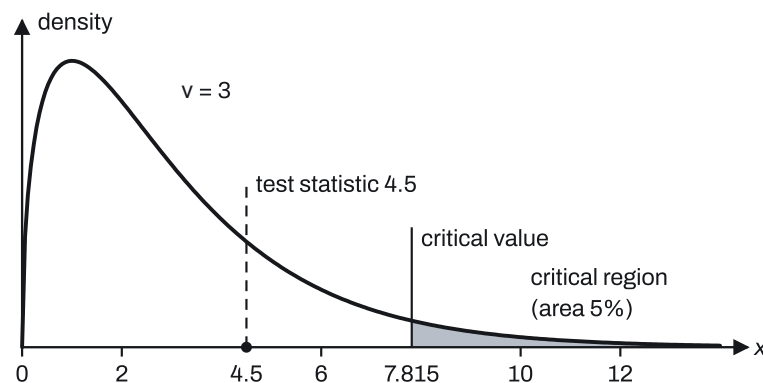
An equivalent form, quicker with a calculator, is $X^2 = \sum \frac{O^2}{E} - N$, where N is the total frequency: $\frac{14^2 + 26^2 + 17^2 + 23^2}{20} - 80 = 84.5 - 80 = 4.5$.

A small X^2 means the data fit the model well; a large one means they don't. Only **large** values count against the model, so the test always uses the **upper tail**.

The χ^2 distribution and degrees of freedom

If the model is true, X^2 roughly follows a **χ^2 distribution**. There is a different χ^2 distribution for each number of **degrees of freedom**, ν , which counts how many of the frequencies are free to vary.

For the spinner there are 4 classes, but the expected frequencies were made to add up to 80, the same total as the observed ones. Once three gaps are known, the fourth is fixed. So $\nu = 4 - 1 = 3$. Every extra fact taken from the data to build the expected frequencies (such as a mean) costs one more degree of freedom.



The χ^2 table in MF19 gives, for each ν , the value that X^2 exceeds with a small probability when the model is true. With $\nu = 3$ at the 5% level it is 7.815 (the table's $p = 0.95$ column). The spinner's 4.5 is below 7.815, so there is **not** enough evidence to say it is unfair.

Why every expected frequency must be at least 5

The χ^2 distribution is only an approximation to how X^2 behaves, and it is poor when some expected frequencies are small: dividing by a tiny E can make one class blow up the total. So before calculating, **combine neighbouring classes** until every expected frequency is at least 5. Combine the observed frequencies in the same way, and count the classes **after** combining when you find ν .

Two kinds of test

- **Goodness of fit:** one row of data and a model that predicts the proportions (equal, a ratio, binomial, Poisson, normal, a given pdf). H_0 : the model fits the data.
- **Independence in a contingency table:** two ways of classifying the same individuals, such as membership type and time of visit. H_0 : the two are independent. If they are, the proportion in each row is the same in every column, which gives

$$E = \frac{\text{row total} \times \text{column total}}{\text{grand total}}.$$

An $r \times c$ table has $\nu = (r - 1)(c - 1)$: once $r - 1$ rows and $c - 1$ columns are filled in, the totals fix the rest.

A test never **proves** anything. Rejecting H_0 means the data give evidence against the model; not rejecting means there isn't enough evidence to doubt it.

Key facts and formulas

"Given" means it is in the MF19 list of formulae and statistical tables in the exam. "Learn it" means you must remember it.

Formula	Status
$X^2 = \sum \frac{(O - E)^2}{E}$, or equivalently $X^2 = \sum \frac{O^2}{E} - N$	Learn it
Goodness of fit: $\nu = (\text{number of classes after combining}) - 1 - (\text{number of parameters estimated from the data})$	Learn it
Contingency table: $E = \frac{\text{row total} \times \text{column total}}{\text{grand total}}$	Learn it
Contingency table: $\nu = (r - 1)(c - 1)$, counting rows and columns after combining	Learn it
Every expected frequency must be at least 5; combine classes until it is	Learn it
Expected frequency = $N \times P(\text{class})$, and all the E add up to N	Learn it
Binomial: $p_r = \binom{n}{r} p^r (1 - p)^{n-r}$, $\mu = np$	Given
Poisson: $p_r = e^{-\lambda} \frac{\lambda^r}{r!}$, $\mu = \lambda$	Given
Normal: $P(a \leq X < b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$ (Φ is in the normal table)	Learn it
Given pdf: $P(a \leq X < b) = \int_a^b f(x) dx$	Learn it
Mean from a frequency table: $\bar{x} = \frac{\sum xf}{\sum f}$	Given
Critical values of the χ^2 distribution (upper tail)	Given

The critical values you use most (all in MF19):

ν	10%	5%	2.5%	1%	0.5%
1	2.706	3.841	5.024	6.635	7.879
2	4.605	5.991	7.378	9.210	10.60
3	6.251	7.815	9.348	11.34	12.84
4	7.779	9.488	11.14	13.28	14.86
5	9.236	11.07	12.83	15.09	16.75
6	10.64	12.59	14.45	16.81	18.55

A test at the 10% level uses the table's $p = 0.9$ column, at 5% the $p = 0.95$ column, and so on.

How to do it

In the Paper 4 questions we have tagged for this topic (32 questions, 2021–2025), the types came up this often. Each question is one type.

Question type	Questions
Test for independence in a contingency table	12
Goodness of fit: Poisson with a given mean	8
Goodness of fit: binomial with a given p	4
Goodness of fit: parameter estimated from the data	4
Goodness of fit: normal distribution or a given pdf	3
Goodness of fit: equally likely outcomes or a given ratio	1

Every test finishes the same way, so learn these five closing steps once:

1. **Hypotheses in context**, naming the model and the data: " H_0 : Po(1.6) fits the data on the number of calls; H_1 : Po(1.6) does not fit the data." For a table: " H_0 : time of visit is independent of membership type; H_1 : time of visit is not independent of membership type."
2. **Test statistic** $X^2 = \sum \frac{(O-E)^2}{E}$, showing the expected frequencies and at least a few of the contributions.
3. **Degrees of freedom**, from the classes left **after** combining.
4. **Compare** X^2 with the critical value: bigger means reject H_0 (significant), smaller means do not reject H_0 .
5. **Conclude in context, with uncertainty**: "There is sufficient evidence to suggest that ..." or "There is insufficient evidence to suggest that ...".

1. Test for independence in a contingency table (12 questions)

1. Write H_0 : "[first variable] is independent of [second variable]", and H_1 : "... is not independent of ...", using the question's words.
2. Find the row totals, column totals and grand total if they aren't given.

- Expected frequency of each cell** = $\frac{\text{row total} \times \text{column total}}{\text{grand total}}$. For example, a cell in a row with total 60 and a column with total 45, out of 200 altogether, has $E = \frac{60 \times 45}{200} = 13.5$. Write every E in the table (to at least 3 s.f.); mark schemes award marks for seeing them.
- If any E is below 5, combine that row or column with a neighbouring one (the one that makes sense, such as two adjacent age bands), then recalculate.
- Work out every contribution $\frac{(O-E)^2}{E}$ and add them.
- $\nu = (r - 1)(c - 1)$ from the table's size after combining: a 3×4 table has $\nu = 2 \times 3 = 6$; a 2×2 table has $\nu = 1$.
- Compare and conclude, mentioning **both** variables.

Variations you will meet:

- A 2×2 table.** Same method with $\nu = 1$. Yates' correction is not needed.
- "Compare ..." or "which group does better?"** after the test: look at the cells with the biggest contributions and say whether O is above or below E there. In Example 2 below, the morning column gives the biggest contributions: pay-as-you-go members visited in the morning more often than independence predicts, and monthly members less often.
- The same data grouped two ways** (for example two grades, then three grades), with the second statistic given. Finish the second test with its own ν , then say which you prefer: usually the one with more categories, because it keeps more detail. It does **not** use more data; the sample is the same.
- Totals not given:** add them up yourself first, and check the row totals and column totals give the same grand total.

2. Goodness of fit: Poisson with a given mean (8 questions)

- H_0 : "Po(λ) fits the data on [what was counted]"; H_1 : "Po(λ) does not fit the data".
- Expected frequencies:** $E = N \times e^{-\lambda} \frac{\lambda^r}{r!}$ for each value r . For example, with $\lambda = 1.6$ and $N = 250$: $E(0) = 250e^{-1.6} = 50.47$ and $E(1) = 250e^{-1.6} \times 1.6 = 80.76$.
- The **last class** is open, such as "5 or more". Its expected frequency is N minus all the others, never just $N \times P(X = 5)$. Check that all the E add up to N .
- Combine** classes at the ends (usually the top end) until every E is at least 5. Add the observed frequencies of the same classes.
- $\nu = (\text{classes after combining}) - 1$, because λ was given, not estimated.
- Sum the contributions, compare and conclude.

Variations you will meet:

- "Find the missing expected frequencies p and q ":** use the formula for the one in the middle and " N minus the rest" for the open-ended last class. Give the accuracy asked for (often 3 d.p.).
- A table of contributions is given with one missing:** find the missing E , then its contribution $\frac{(O-E)^2}{E}$, then add it to the given ones.
- Small expected frequencies at the start as well as the end:** for a larger mean, $E(0)$ can be below 5; combine the first classes too.
- An observed frequency of 0** is still a class: keep it, and give it its expected frequency.

3. Goodness of fit: binomial with a given p (4 questions)

- H_0 : "B(n, p) fits the data on [what was counted]".

- Expected frequencies:** $E = N \times \binom{n}{r} p^r (1-p)^{n-r}$. For example, packs of 4 with $p = 0.2$ and $N = 120$ packs:
 $E(0) = 120 \times 0.8^4 = 49.152$ and $E(4) = 120 \times 0.2^4 = 0.192$.
- Combine classes with $E < 5$. With binomial data this often means combining **several** classes at one end (or at both ends when p is near $\frac{1}{2}$).
- $\nu = (\text{classes after combining}) - 1$.
- Sum the contributions, compare and conclude, for example "there is insufficient evidence to reject the claim that 20% of the items in a pack are faulty".

4. Goodness of fit: parameter estimated from the data (4 questions)

When the question gives no value for p or λ , estimate it from the sample, then test as in types 2 and 3 but with **one fewer degree of freedom**.

- Estimate the parameter** from the frequency table: $\bar{x} = \frac{\sum rf}{\sum f}$.
 - Poisson: $\lambda = \bar{x}$.
 - Binomial $B(n, p)$: $np = \bar{x}$, so $p = \frac{\bar{x}}{n}$. For example, 80 bags of 4 bulbs with $\sum rf = 96$ faulty bulbs give $\bar{x} = 1.2$ and $p = \frac{1.2}{4} = 0.3$. Equivalently, $p = \frac{\text{total successes}}{\text{total trials}} = \frac{96}{320}$.
- Find the expected frequencies and combine classes as before.
- $\nu = (\text{classes after combining}) - 1 - 1$: one for the total and one for the estimated parameter.
- Sum, compare and conclude.

Variations you will meet:

- "Explain how the mean and standard deviation support a Poisson model"**: for a Poisson distribution the mean equals the variance. Square the standard deviation and say it is **close to** the mean (they won't be exactly equal for a sample).
- "Give a reason why a binomial model may not be suitable"**: the trials may not be independent (one event makes another more likely, such as faulty bulbs in one bag all coming from the same bad batch), or the probability may not be the same for every trial. Say it in the question's context.
- "Show that $p = \dots$ "** from the data: show $\sum rf$ and the division.

5. Goodness of fit: normal distribution or a given pdf (3 questions)

The data come in class intervals, such as $54 \leq x < 62$.

- H_0 : " $N(\mu, \sigma^2)$ fits the data" or "the distribution with pdf f fits the data".
- Probability of each interval:**
 - Normal: standardise both ends and subtract. For $N(50, 8^2)$ and $54 \leq x < 62$ with $N = 120$: $P = \Phi(1.5) - \Phi(0.5) = 0.9332 - 0.6915 = 0.2417$, so $E = 120 \times 0.2417 = 29.01$.
 - A given pdf: integrate over the interval. For $f(x) = \frac{3}{64}x^2$ on $0 \leq x \leq 4$ and $N = 160$: $P(2 \leq X < 3) = \left[\frac{x^3}{64} \right]_2^3 = \frac{19}{64}$, so $E = 160 \times \frac{19}{64} = 47.5$.
- The end intervals (such as $x < 40$) run to the end of the distribution, so use Φ of one end only, or subtract from N .
- Combine intervals with $E < 5$, then $\nu = (\text{intervals after combining}) - 1$ when the parameters were given. (Each parameter estimated from the data would take off one more.)
- Sum, compare and conclude.

A "show that" for an expected frequency needs the working written out: the standardised values or the integral with its limits, the probability, then $\times N$. If the answer is given to 4 d.p., work to at least 5 d.p. or exactly.

6. Goodness of fit: equally likely outcomes or a given ratio (1 question)

1. H_0 : "the outcomes are equally likely" (a uniform distribution fits the data), or "the ratio 1 : 2 : 1 fits the data".
2. **Expected frequencies** = $N \times$ the proportion. With 5 equally likely categories and $N = 150$, each $E = 30$. With a ratio 1 : 2 : 1, the proportions are $\frac{1}{4}, \frac{1}{2}, \frac{1}{4}$.
3. $\nu = (\text{number of categories}) - 1$, so 5 categories give $\nu = 4$.
4. Sum, compare and conclude in context.

Worked examples

The bold codes show where marks are earned, as in Cambridge mark schemes. The number is how many marks.

- **M1** method mark: for a correct method, even if the numbers slip.
- **A1** accuracy mark: for a correct answer; it needs the M mark before it.
- **B1** mark for a correct result on its own, with no method needed.

Example 1

A gardener says that seeds of a certain plant produce red, pink and white flowers in the ratio 1 : 2 : 1. In a random sample of 160 plants grown from these seeds, 29 had red flowers, 94 pink and 37 white.

Test, at the 10% significance level, whether the gardener's claim is supported by the data. **[6]**

H_0 : the ratio 1 : 2 : 1 fits the data on flower colour. H_1 : the ratio 1 : 2 : 1 does not fit the data. **B1**

Expected frequencies: $160 \times \frac{1}{4} = 40$, $160 \times \frac{1}{2} = 80$, $160 \times \frac{1}{4} = 40$. **B1**

$$X^2 = \frac{(29 - 40)^2}{40} + \frac{(94 - 80)^2}{80} + \frac{(37 - 40)^2}{40} = 3.025 + 2.45 + 0.225 = 5.7$$

M1 for the expression, **A1** for 5.7.

$\nu = 3 - 1 = 2$, so the critical value is 4.605. $5.7 > 4.605$, so reject H_0 . **M1**

There is sufficient evidence to suggest that the gardener's claim is incorrect: the flower colours do not follow the ratio 1 : 2 : 1. **A1**

Example 2

A gym has two types of membership, monthly and pay-as-you-go. For a random sample of 150 members, the time of day of their last visit was recorded.

	Morning	Afternoon	Evening	Late night
Monthly	21	38	26	3

	Morning	Afternoon	Evening	Late night
Pay-as-you-go	29	17	12	4

Test, at the 5% significance level, whether time of visit is independent of membership type. **[8]**

H_0 : time of visit is independent of membership type. H_1 : time of visit is not independent of membership type. **B1**

Row totals are 88 and 62; column totals are 50, 55, 38, 7; the grand total is 150. The late-night expected frequencies are $\frac{88 \times 7}{150} = 4.107$ and $\frac{62 \times 7}{150} = 2.893$, both below 5, so combine "Evening" and "Late night" into one column, "Evening or later" (total 45). **M1**

Observed (expected) frequencies:

	Morning	Afternoon	Evening or later
Monthly	21 (29.333)	38 (32.267)	29 (26.4)
Pay-as-you-go	29 (20.667)	17 (22.733)	16 (18.6)

M1 for the method $\frac{\text{row total} \times \text{column total}}{150}$, **A1** for all six values.

Contributions:

	Morning	Afternoon	Evening or later
Monthly	2.3674	1.0187	0.2561
Pay-as-you-go	3.3602	1.4459	0.3634

$$X^2 = 8.81. \text{ **M1 A1**}$$

$\nu = (2 - 1)(3 - 1) = 2$, so the critical value is 5.991. $8.81 > 5.991$, so reject H_0 . **M1**

There is sufficient evidence to suggest that time of visit is not independent of membership type. **A1**

(The biggest contributions are in the morning column: pay-as-you-go members came in the morning more often than independence would predict, and monthly members less often.)

Example 3

The numbers of typing errors on 100 randomly chosen pages of a report are shown in the table. No page had more than 5 errors.

Number of errors	0	1	2	3	4	5
Number of pages	31	26	20	12	7	4

(a) Show that the mean number of errors per page is 1.5. **[1]**

(b) Test, at the 5% significance level, whether a Poisson distribution is a suitable model for the number of errors per page. **[8]**

(a) $\sum rf = 0 + 26 + 40 + 36 + 28 + 20 = 150$, so $\bar{x} = \frac{150}{100} = 1.5$. **B1**

(b) H_0 : a Poisson distribution fits the data on errors per page. H_1 : a Poisson distribution does not fit the data. **B1**

Use $Po(1.5)$, with the mean estimated from the data. **M1** for $E = 100e^{-1.5} \frac{1.5^r}{r!}$:

Number of errors	0	1	2	3	4	≥ 5
O	31	26	20	12	7	4
E	22.313	33.470	25.102	12.551	4.707	1.858

A1. The last class is "5 or more", found as $100 -$ the rest, so the E add up to 100.

4.707 and 1.858 are below 5, so combine the last two classes: $O = 11$, $E = 6.564$. **M1**

$$X^2 = \frac{(31 - 22.313)^2}{22.313} + \frac{(26 - 33.470)^2}{33.470} + \frac{(20 - 25.102)^2}{25.102} + \frac{(12 - 12.551)^2}{12.551} + \frac{(11 - 6.564)^2}{6.564}$$

$$= 3.382 + 1.667 + 1.037 + 0.024 + 2.997 = 9.11$$

M1 A1

There are 5 classes, and one parameter was estimated, so $\nu = 5 - 1 - 1 = 3$ and the critical value is 7.815. $9.11 > 7.815$, so reject H_0 . **M1**

There is sufficient evidence to suggest that a Poisson distribution is not a suitable model for the number of errors per page. **A1** (There are more pages with no errors, and more with many, than a Poisson model predicts.)

Example 4

The masses, x grams, of 200 randomly chosen apples from an orchard are summarised below. The grower claims that the masses follow the distribution $N(150, 20^2)$.

Mass	$x < 130$	$130 \leq x < 150$	$150 \leq x < 170$	$170 \leq x < 190$	$x \geq 190$
Observed	22	78	60	29	11
Expected	31.73	68.27	68.27	a	4.55

(a) Show that $a = 27.18$, correct to 2 decimal places. **[2]**

(b) Carry out a goodness of fit test, at the 5% significance level, to test the grower's claim. **[6]**

(a) $P(170 \leq X < 190) = P\left(\frac{170-150}{20} \leq Z < \frac{190-150}{20}\right) = \Phi(2) - \Phi(1)$. **M1**

$= 0.97725 - 0.84134 = 0.13591$, and $200 \times 0.13591 = 27.18$. **A1**

(b) H_0 : $N(150, 20^2)$ fits the data on apple masses. H_1 : $N(150, 20^2)$ does not fit the data. **B1**

The last expected frequency, 4.55, is below 5, so combine the last two classes: $O = 40$, $E = 31.73$. **M1**

$$X^2 = \frac{(22 - 31.73)^2}{31.73} + \frac{(78 - 68.27)^2}{68.27} + \frac{(60 - 68.27)^2}{68.27} + \frac{(40 - 31.73)^2}{31.73} = 2.984 + 1.387 + 1.002 + 2.155 = 7.53$$

M1 A1

4 classes, no parameters estimated: $\nu = 3$, critical value 7.815. $7.53 < 7.815$, so do not reject H_0 . **M1**

There is insufficient evidence to suggest that $N(150, 20^2)$ is not a good model for the masses, so the grower's claim is not contradicted. **A1**

(At the 10% level the critical value is 6.251 and the answer would change: always use the level asked for.)

Common mistakes

- **Hypotheses without context.** "The distribution fits" or "they are independent" is not enough. Name the model **and** the data ("Po(0.9) fits the data on the number of calls"), or **both** variables ("time of visit is independent of membership type").
- **Swapping the hypotheses.** H_0 is always the model or independence. Examiners withhold the final mark when H_1 says "independent".
- **Not combining classes**, or combining too many. Combine only until every E is at least 5, and check again after each combination. Examiners see both errors every year.
- **Getting the last class wrong.** "7 or more" is not "exactly 7", and a class with $O = 0$ must not be dropped. Find its E by subtracting the others from N , and check that all the E add up to N .
- **Wrong degrees of freedom.** Count the classes **after** combining, and take off one more for each parameter estimated from the data. Using the number of classes before combining, or forgetting the estimated mean, gives the wrong critical value.
- **Rounding too early.** Keep expected frequencies to at least 3 d.p. (4 s.f.) so the test statistic is right to 3 s.f.
- **Conclusions that are too sure.** A test can't "prove" a model right or wrong. Write "there is sufficient (or insufficient) evidence to suggest ...".
- **A conclusion that contradicts the test.** If you reject H_0 , the evidence suggests the variables are **not** independent. "There is insufficient evidence to suggest they are independent" says something else and loses the mark.
- **"Mean $\approx$ variance" with the wrong numbers.** The variance is the standard deviation **squared**; state both values and say they are **approximately** equal.
- **"More data"** as the reason for preferring a finer table: the sample is the same; it has more categories, more detail and more degrees of freedom.

How to get the marks

- **Hypotheses (B1):** in words, in context, the right way round. For goodness of fit mention both the distribution and the data.
- **Expected frequencies must be seen.** Write them in a table; mark schemes usually give a method mark for some correct and an accuracy mark for all correct, to 3 s.f. or better.
- **Show contributions.** A method mark needs a few $\frac{(O-E)^2}{E}$ terms (or $\sum \frac{O^2}{E} - N$) with numbers in. A bare total usually earns much less than the full working.
- **Combining:** show which classes you combined and the new O and E , for example " $O = 11, E = 6.564$ ".
- **Comparison (M1):** state the critical value with its ν , the inequality and "reject H_0 " or "do not reject H_0 ". This mark follows through from your own statistic, and a critical value for your own (wrong) number of classes can still earn it.
- **Conclusion (A1):** in context, with words like "sufficient evidence to suggest". This mark needs correct work throughout.
- **"Complete the table" or "find p and q ":** give the accuracy asked for (often 3 d.p.). Using N minus the others for the last class is accepted.

- **"Show that"**: the answer is printed, so show every step, such as $\Phi(2) - \Phi(1)$, the probability and $\times N$, or $\sum r f$ and the division.
- **"Explain" or "give a reason"** (one mark): one clear point in context is enough, such as "bulbs in the same bag may share a fault, so they are not independent".
- **Significance levels**: 10% uses the 0.9 column of the χ^2 table, 5% the 0.95 column, 2.5% the 0.975 column, 0.5% the 0.995 column.

Quick check

1. State the number of degrees of freedom for each test. (a) A Poisson goodness of fit test with the mean estimated from the data, and 7 classes after combining. (b) A test for independence in a 3×5 contingency table with no combining. (c) A goodness of fit test of $B(10, 0.3)$, with 6 classes after combining.
2. The number of successes in 6 trials is modelled by $B(6, 0.25)$. In 200 sets of 6 trials, 2 successes occurred 51 times. Find the expected frequency for 2 successes, and this class's contribution to the test statistic.
3. In a contingency table with grand total 240, a cell lies in a row with total 64 and a column with total 45. Its observed frequency is 18. Find its expected frequency and its contribution to X^2 .
4. A random variable has pdf $f(x) = \frac{x}{8}$ for $0 \leq x \leq 4$. In a goodness of fit test with 80 observations, find the expected frequency for the interval $1 \leq x < 2$.
5. A goodness of fit test with 4 degrees of freedom gives $X^2 = 11.6$. State the result of the test at the 5% level and at the 1% level.

Answers

- (a) $\nu = 7 - 1 - 1 = 5$. (b) $\nu = (3 - 1)(5 - 1) = 8$. (c) $\nu = 6 - 1 = 5$, since p was given.
- $E = 200 \times \binom{6}{2} (0.25)^2 (0.75)^4 = 200 \times 0.29663 = 59.33$. Contribution = $\frac{(51 - 59.33)^2}{59.33} = 1.17$.
- $E = \frac{64 \times 45}{240} = 12$. Contribution = $\frac{(18 - 12)^2}{12} = 3$.
- $P(1 \leq X < 2) = \left[\frac{x^2}{16} \right]_1^2 = \frac{4 - 1}{16} = \frac{3}{16}$, so $E = 80 \times \frac{3}{16} = 15$.
- At 5% the critical value is 9.488: $11.6 > 9.488$, so reject H_0 (there is sufficient evidence that the model does not fit).
At 1% it is 13.28: $11.6 < 13.28$, so do not reject H_0 .

Past questions to try

- 9231/43/M/J/24 Q3**: independence in a 3×3 contingency table.
- 9231/42/O/N/23 Q2**: a Poisson model with a given mean, combining classes at both ends.
- 9231/42/O/N/24 Q3**: a binomial model with p estimated from the data.
- 9231/42/M/J/23 Q3**: expected frequencies from a given pdf, then a goodness of fit test.