Chi-squared goodness of fit

Worked solutions

Further Maths topic practice

This topic online

1 A spinner has five equal sections coloured red, blue, green, yellow and white. Amira spins it 150 times and records the colour each time.

Colour	Red	Blue	Green	Yellow	White
Frequency	36	24	33	21	36

She wants to test whether the spinner is fair.

a. Write down suitable hypotheses for the test.

b. Find the expected frequency for each colour if the spinner is fair.

c. State the number of degrees of freedom and the critical value for a test at the 5% level of significance.

(3 marks)

a) H_0 is always the model

H_0: the spinner is fair (a discrete uniform model is suitable)

H_1: the spinner is not fair

b) Each colour has probability $\frac{1}{5}$

$E = 150 \times \frac{1}{5}$

$E = 30$

30 for each colour

c) 5 classes, 1 constraint (the total must be 150)

$\nu = 5 - 1$

$\nu = 4$

From the χ^2 table, 5% column

critical value = 9.488

$\nu = 4$, critical value 9.488

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

2 The number of people visiting a village library on each weekday was recorded over a period of several weeks.

Day	Monday	Tuesday	Wednesday	Thursday	Friday
Visits	52	38	45	41	64

Test, at the 5% level of significance, whether visits to the library are equally likely on each of the five days. State your hypotheses clearly. (6 marks)

H_0: visits are equally likely on each day (discrete uniform)

H_1: visits are not equally likely on each day

Total number of visits

$52 + 38 + 45 + 41 + 64 = 240$

Expected frequency = total ÷ number of days

$E = 240 \div 5$

$E = 48$

Each contribution = $(O - E)^2 \div E$

Mon: $\frac{(52 - 48)^2}{48} = 0.333$

Tue: $\frac{(38 - 48)^2}{48} = 2.083$

Wed: $\frac{(45 - 48)^2}{48} = 0.188$

Thu: $\frac{(41 - 48)^2}{48} = 1.021$

Fri: $\frac{(64 - 48)^2}{48} = 5.333$

$\chi^2 = 0.333 + 2.083 + 0.188 + 1.021 + 5.333$

$\chi^2 = 8.958$

5 classes, 1 constraint

$\nu = 5 - 1 = 4$

Critical value, 5% level, $\nu = 4$

critical value $= 9.488$

$8.958 < 9.488$, so do not reject H_0

No evidence that visits are not equally likely: the uniform model is suitable

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

3 Oliver rolls an ordinary six-sided dice 300 times. The frequencies of the scores 1 to 5 are shown in the table, but the frequency of a 6 has been smudged.

Score	1	2	3	4	5	6
Frequency	41	58	47	62	39	f

a. Find the value of f.

b. Explain why a test of whether the dice is fair has 5 degrees of freedom.

c. Test, at the 5% level of significance, whether the dice is fair. State your hypotheses and the critical value clearly.

(7 marks)

a) Frequencies add up to 300

$$41 + 58 + 47 + 62 + 39 = 247$$

$$f = 300 - 247$$

$$f = 53$$

$$\boxed{f = 53}$$

b) Count the classes and the constraints

6 scores, so 6 classes

Only constraint: the total is 300 (nothing estimated)

$$\nu = 6 - 1 = 5$$

$$\boxed{\text{6 classes} - \text{1 constraint} = 5}$$

c)

H_0: the dice is fair (discrete uniform on 1 to 6)

H_1: the dice is not fair

Expected frequency for each score

$$E = 300 \times \frac{1}{6} = 50$$

Each contribution $= (O - E)^2 \div E$

$$1: \frac{(41 - 50)^2}{50} = 1.62$$

$$2: \frac{(58 - 50)^2}{50} = 1.28$$

$$3: \frac{(47 - 50)^2}{50} = 0.18$$

$$4: \frac{(62 - 50)^2}{50} = 2.88$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

5: $\frac{(39-50)^2}{50} = 2.42$

6: $\frac{(53-50)^2}{50} = 0.18$

$\chi^2 = 1.62 + 1.28 + 0.18 + 2.88 + 2.42 + 0.18$

$\chi^2 = 8.560$

Critical value, 5% level, $\nu = 5$

critical value = 11.070

$8.560 < 11.070$, so do not reject H_0

No evidence that the dice is biased

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

4 A quiz has 5 true or false questions. A teacher believes the 96 students in Year 12 simply guessed every answer, so that the number of correct answers, X, for a student can be modelled by B(5, 0.5). The results were:

Number correct	0	1	2	3	4	5
Number of students	1	7	26	35	21	6

a. Find the expected frequencies under the teacher's model.

b. Test, at the 5% level of significance, whether B(5, 0.5) is a suitable model for X. State your hypotheses and the critical value clearly.

(9 marks)

a) Expected frequency = $96 \times P(X = r)$, $X \sim B(5, 0.5)$

$E(0) = 96 \times 0.5^5 = 3$

$E(1) = 96 \times 5 \times 0.5^5 = 15$

$E(2) = 96 \times 10 \times 0.5^5 = 30$

By symmetry

$E(3) = 30$

$E(4) = 15$

$E(5) = 3$

$\boxed{3, 15, 30, 30, 15, 3}$

b)

H_0: B(5, 0.5) is a suitable model for X

H_1: B(5, 0.5) is not a suitable model

E(0) and E(5) are less than 5: combine each with the next class

0 or 1: O = 1 + 7 = 8

E = 3 + 15 = 18

4 or 5: O = 21 + 6 = 27

E = 15 + 3 = 18

Each contribution = $(O - E)^2 \div E$

0 or 1: $\frac{(8 - 18)^2}{18} = 5.556$

2: $\frac{(26 - 30)^2}{30} = 0.533$

3: $\frac{(35 - 30)^2}{30} = 0.833$

4 or 5: $\frac{(27 - 18)^2}{18} = 4.500$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$$\chi^2 = 5.556 + 0.533 + 0.833 + 4.500$$

$$\chi^2 = 11.422$$

4 classes, 1 constraint (p is given, not estimated)

$\nu = 4 - 1 = 3$

Critical value, 5% level, $\nu = 3$

critical value = 7.815

$11.422 > 7.815$, so reject H_0

B(5, 0.5) is not a suitable model: evidence the students were not just guessing

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

5 A farm shop sells eggs in boxes of 6. Priya checks 100 boxes and records the number of cracked eggs in each box.

Number of cracked eggs	0	1	2	3	4	5	6
Number of boxes	38	37	18	5	2	0	0

She wants to test whether a binomial distribution is a suitable model for the number of cracked eggs in a box.

a. Show that a suitable estimate for the probability that an egg is cracked is 0.16

b. Using this estimate, test at the 5% level of significance whether a binomial distribution is a suitable model. State your hypotheses and the critical value clearly.

(9 marks)

a) Mean number of cracked eggs per box

$\Sigma fx = 0 + 37 + 36 + 15 + 8 = 96$

$\text{mean} = \frac{96}{100} = 0.96$

For $B(6, p)$ the mean is $6p$

$6p = 0.96$

$p = 0.16$

$\boxed{p = 0.16}$

b)

H_0: a binomial distribution is a suitable model

H_1: a binomial distribution is not a suitable model

$X \sim B(6, 0.16)$, expected frequency $= 100 \times P(X = r)$

$E(0) = 100 \times 0.84^6 = 35.13$

$E(1) = 100 \times 6 \times 0.16 \times 0.84^5 = 40.15$

$E(2) = 100 \times 15 \times 0.16^2 \times 0.84^4 = 19.12$

$E(3) = 100 \times 20 \times 0.16^3 \times 0.84^3 = 4.86$

$E(3) < 5$, so combine 3, 4, 5 and 6

$E(3 \text{ or more}) = 100 - 35.13 - 40.15 - 19.12 = 5.60$

$O(3 \text{ or more}) = 5 + 2 + 0 + 0 = 7$

Each contribution $= (O - E)^2 \div E$

$0: \frac{(38 - 35.13)^2}{35.13} = 0.235$

$1: \frac{(37 - 40.15)^2}{40.15} = 0.247$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$$2: \frac{(18 - 19.12)^2}{19.12} = 0.065$$

$$3 \text{ or more: } \frac{(7 - 5.60)^2}{5.60} = 0.348$$

$$\chi^2 = 0.235 + 0.247 + 0.065 + 0.348$$

$$\chi^2 = 0.895$$

4 classes, 2 constraints (total and p estimated)

$$\nu = 4 - 2 = 2$$

Critical value, 5% level, $\nu = 2$

critical value = 5.991

$0.895 < 5.991$, so do not reject H_0

A binomial model is suitable for the number of cracked eggs

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

6 A council helpline manager claims that the number of calls arriving in a minute can be modelled by a Poisson distribution with mean 2. The number of calls in each of 120 randomly chosen minutes is recorded.

Number of calls	0	1	2	3	4	5 or more
Number of minutes	14	30	33	22	12	9

Test, at the 5% level of significance, whether the manager's claim is supported by the data. State your hypotheses and the critical value clearly.

(8 marks)

H_0: Po(2) is a suitable model for the number of calls per minute

H_1: Po(2) is not a suitable model

Expected frequency = $120 \times P(X = r)$, $X \sim Po(2)$

$E(0) = 120 \times e^{-2} = 16.24$

$E(1) = 120 \times 2e^{-2} = 32.48$

$E(2) = 120 \times \frac{2^2 e^{-2}}{2!} = 32.48$

$E(3) = 120 \times \frac{2^3 e^{-2}}{3!} = 21.65$

$E(4) = 120 \times \frac{2^4 e^{-2}}{4!} = 10.83$

Last class makes the total 120

$E(5 \text{ or more}) = 120 - 113.68 = 6.32$

Every $E \geq 5$, so no classes are combined

Each contribution = $(O - E)^2 \div E$

$0: \frac{(14 - 16.24)^2}{16.24} = 0.309$

$1: \frac{(30 - 32.48)^2}{32.48} = 0.189$

$2: \frac{(33 - 32.48)^2}{32.48} = 0.008$

$3: \frac{(22 - 21.65)^2}{21.65} = 0.006$

$4: \frac{(12 - 10.83)^2}{10.83} = 0.127$

$5 \text{ or more}: \frac{(9 - 6.32)^2}{6.32} = 1.138$

$\chi^2 = 0.309 + 0.189 + 0.008 + 0.006 + 0.127 + 1.138$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$$\chi^2 = 1.778$$

6 classes, 1 constraint ($\lambda = 2$ is given)

$\nu = 6 - 1 = 5$

Critical value, 5% level, $\nu = 5$

critical value = 11.070

$1.778 < 11.070$, so do not reject H_0

Po(2) is a suitable model for the number of calls per minute

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

7 Kiran is studying daisies on a school field. She places a 1 m by 1 m square frame at 80 random positions and counts the daisies inside it each time.

Number of daisies	0	1	2	3	4	5	6
Frequency	6	15	22	18	11	5	3

a. Show that the mean number of daisies per square is 2.5

b. Test, at the 5% level of significance, whether a Poisson distribution is a suitable model for the number of daisies in a square. State your hypotheses and the critical value clearly.

c. Give one reason, in context, why a Poisson model might be expected to fit. (10 marks)

a) $\Sigma fx \div \Sigma f$

$\Sigma fx = 0 + 15 + 44 + 54 + 44 + 25 + 18 = 200$

$\text{mean} = \frac{200}{80} = 2.5$

mean = 2.5

b)

H_0: a Poisson distribution is a suitable model

H_1: a Poisson distribution is not a suitable model

$X \sim Po(2.5)$, expected frequency $= 80 \times P(X = r)$

$E(0) = 80 \times e^{-2.5} = 6.57$

$E(1) = 80 \times 2.5e^{-2.5} = 16.42$

$E(2) = 80 \times \frac{2.5^2 e^{-2.5}}{2!} = 20.52$

$E(3) = 80 \times \frac{2.5^3 e^{-2.5}}{3!} = 17.10$

$E(4) = 80 \times \frac{2.5^4 e^{-2.5}}{4!} = 10.69$

Use "5 or more" for the last class so the total is 80

$E(5 \text{ or more}) = 80 - 71.29 = 8.71$

$O(5 \text{ or more}) = 5 + 3 = 8$

Each contribution $= (O - E)^2 \div E$

$0: \frac{(6 - 6.57)^2}{6.57} = 0.049$

$1: \frac{(15 - 16.42)^2}{16.42} = 0.122$

$2: \frac{(22 - 20.52)^2}{20.52} = 0.107$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$$3: \frac{(18 - 17.10)^2}{17.10} = 0.047$$

$$4: \frac{(11 - 10.69)^2}{10.69} = 0.009$$

$$5 \text{ or more: } \frac{(8 - 8.71)^2}{8.71} = 0.057$$

$$\chi^2 = 0.049 + 0.122 + 0.107 + 0.047 + 0.009 + 0.057$$

$$\chi^2 = 0.391$$

6 classes, 2 constraints (total and λ estimated)

$\nu = 6 - 2 = 4$

Critical value, 5% level, $\nu = 4$

critical value = 9.488

$0.391 < 9.488$, so do not reject H_0

A Poisson distribution is a suitable model for the number of daisies

c) Poisson conditions in context

Daisies grow singly, randomly and independently at a constant rate

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

8 An editor counts the typing errors on each of 200 pages of a draft novel. The mean number of errors per page is 1.5 and she uses Po(1.5) to model the number of errors on a page. Her results and some of her expected frequencies are shown below.

Number of errors	0	1	2	3	4 or more
Observed	48	70	46	24	12
Expected	r	66.94	s	25.10	13.13

a. Find the value of r and the value of s, giving your answers to 2 decimal places.

b. Explain why the test has 3 degrees of freedom.

c. The test statistic is 0.89. Test, at the 5% level of significance, whether Po(1.5) is a suitable model. State your hypotheses and the critical value clearly.

(7 marks)

a) Expected frequency = $200 \times P(X = r)$, $X \sim Po(1.5)$

$r = 200 \times e^{-1.5}$

$r = 44.63$

$s = 200 \times \frac{1.5^2 e^{-1.5}}{2!}$

$s = 50.20$

$r = 44.63$, $s = 50.20$

b) Count classes, subtract constraints

5 classes

Constraint 1: the expected frequencies total 200

Constraint 2: $\lambda = 1.5$ was estimated from the data

$\nu = 5 - 2 = 3$

c)

H_0: Po(1.5) is a suitable model for the number of errors

H_1: Po(1.5) is not a suitable model

Critical value, 5% level, $\nu = 3$

critical value = 7.815

$0.89 < 7.815$, so do not reject H_0

Po(1.5) is a suitable model for errors per page

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

9 A garden centre sows 10 seeds in each of 150 trays and counts how many seeds germinate in each tray.

Seeds germinating	3	4	5	6	7	8	9	10
Number of trays	2	10	20	24	30	32	26	6

No tray had fewer than 3 seeds germinate. The manager wants to test whether B(10, p) is a suitable model.

a. Show that the estimate of p from these data is 0.7

Using p = 0.7 the manager finds these expected frequencies.

Seeds germinating	4 or fewer	5	6	7	8	9	10
Expected	7.10	15.44	30.02	40.02	35.02	18.16	4.24

b. Explain why the classes 9 and 10 must be combined, and state the number of degrees of freedom for the test.

c. The test statistic is 12.82. Test, at the 1% level of significance, whether B(10, p) is a suitable model. State your hypotheses and the critical value.

d. State, with a reason, whether your conclusion would change at the 5% level.

(8 marks)

a) Mean number germinating per tray

$\Sigma fx = 3\times2 + 4\times10 + 5\times20 + 6\times24 + 7\times30 + 8\times32 + 9\times26 + 10\times6$

$\Sigma fx = 1050$

$\text{mean} = \frac{1050}{150} = 7$

Mean of B(10, p) is 10p

$10p = 7$

$p = 0.7$

b) Expected frequencies must be at least 5

E(10) = 4.24 < 5, so combine 9 and 10

E(9 or 10) = 18.16 + 4.24 = 22.40

Classes: 4 or fewer, 5, 6, 7, 8, 9 or 10: 6 classes

2 constraints: total and p estimated

$\nu = 6 - 2 = 4$

$\boxed{\nu = 4}$

c)

H_0: B(10, p) is a suitable model

H_1: B(10, p) is not a suitable model

1% level, $\nu = 4$

critical value = 13.277

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$12.82 < 13.277$, so do not reject H_0

Binomial model is suitable at the 1% level

d) 5% level, $\nu = 4$

critical value = 9.488

$12.82 > 9.488$, so H_0 would be rejected

Yes: at 5% the binomial model would be rejected

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

10 At a school fair, each of 200 visitors throws 4 hoops at a peg. The number of hoops that land on the peg is recorded for each visitor.

Hoops on the peg	0	1	2	3	4
Number of visitors	18	52	70	46	14

The stall holder, Megan, says the probability that a hoop lands on the peg is 0.4 and that the throws are independent.

a. Test, at the 5% level of significance, whether B(4, 0.4) is a suitable model for the number of hoops on the peg. State your hypotheses and the critical value clearly.

b. Use the data to estimate the probability that a hoop lands on the peg.

c. Using this estimate, the test statistic for a binomial model is 2.21. Stating the degrees of freedom and the critical value, test at the 5% level whether a binomial model with this estimate is suitable.

d. Comment on Megan's claim in the light of both tests.

(10 marks)

a)

H_0: B(4, 0.4) is a suitable model

H_1: B(4, 0.4) is not a suitable model

Expected frequency = $200 \times P(X = r)$, $X \sim B(4, 0.4)$

$E(0) = 200 \times 0.6^4 = 25.92$

$E(1) = 200 \times 4 \times 0.4 \times 0.6^3 = 69.12$

$E(2) = 200 \times 6 \times 0.4^2 \times 0.6^2 = 69.12$

$E(3) = 200 \times 4 \times 0.4^3 \times 0.6 = 30.72$

$E(4) = 200 \times 0.4^4 = 5.12$

All $E \geq 5$, no combining

Each contribution = $(O - E)^2 \div E$

$0: \frac{(18 - 25.92)^2}{25.92} = 2.420$

$1: \frac{(52 - 69.12)^2}{69.12} = 4.240$

$2: \frac{(70 - 69.12)^2}{69.12} = 0.011$

$3: \frac{(46 - 30.72)^2}{30.72} = 7.600$

$4: \frac{(14 - 5.12)^2}{5.12} = 15.401$

$\chi^2 = 2.420 + 4.240 + 0.011 + 7.600 + 15.401$

$\chi^2 = 29.673$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

5 classes, 1 constraint (p given)

$\nu = 5 - 1 = 4$

Critical value, 5% level, $\nu = 4$

critical value = 9.488

$29.673 > 9.488$, so reject H_0

B(4, 0.4) is not a suitable model

b) Mean number on the peg

$\Sigma fx = 52 + 140 + 138 + 56 = 386$

$\text{mean} = \frac{386}{200} = 1.93$

$4p = 1.93$

$p = 0.4825$

$p = 0.4825$

c)

H_0: a binomial distribution is a suitable model

H_1: a binomial distribution is not a suitable model

5 classes, 2 constraints (total and p estimated)

$\nu = 5 - 2 = 3$

critical value = 7.815

$2.21 < 7.815$, so do not reject H_0

A binomial model is suitable

d) Compare the two tests

Binomial shape fits: throws independent, constant probability

But p is about 0.48, not 0.4

Megan underestimates the probability

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

11 The numbers of goals scored in the 380 matches of a football league season are shown below.

Goals in a match	0	1	2	3	4	5	6	7
Number of matches	30	72	98	78	52	26	20	4

a. Show that the mean number of goals per match is 2.6

b. Test, at the 5% level of significance, whether a Poisson distribution is a suitable model for the number of goals in a match. State your hypotheses and the critical value clearly.

c. Use your model to estimate the number of matches in the season with more than 4 goals, and compare this with the data.

(10 marks)

a)

$$\Sigma fx = 0 + 72 + 196 + 234 + 208 + 130 + 120 + 28$$

$$\Sigma fx = 988$$

$$\text{mean} = \frac{988}{380} = 2.6$$

$$\boxed{\text{mean} = 2.6}$$

b)

H_0: a Poisson distribution is a suitable model

H_1: a Poisson distribution is not a suitable model

$X \sim Po(2.6)$, expected frequency $= 380 \times P(X = r)$

$$E(0) = 380 \times \frac{2.6^0 e^{-2.6}}{0!} = 28.22$$

$$E(1) = 380 \times \frac{2.6^1 e^{-2.6}}{1!} = 73.38$$

$$E(2) = 380 \times \frac{2.6^2 e^{-2.6}}{2!} = 95.40$$

$$E(3) = 380 \times \frac{2.6^3 e^{-2.6}}{3!} = 82.68$$

$$E(4) = 380 \times \frac{2.6^4 e^{-2.6}}{4!} = 53.74$$

$$E(5) = 380 \times \frac{2.6^5 e^{-2.6}}{5!} = 27.94$$

$$E(6) = 380 \times \frac{2.6^6 e^{-2.6}}{6!} = 12.11$$

Last class "7 or more" makes the total 380

$$E(7 \text{ or more}) = 380 - 373.48 = 6.52$$

Every $E \geq 5$, so no classes are combined

Each contribution $= (O - E)^2 \div E$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$$0: \frac{(30-28.22)^2}{28.22} = 0.112$$

$$1: \frac{(72-73.38)^2}{73.38} = 0.026$$

$$2: \frac{(98-95.40)^2}{95.40} = 0.071$$

$$3: \frac{(78-82.68)^2}{82.68} = 0.265$$

$$4: \frac{(52-53.74)^2}{53.74} = 0.056$$

$$5: \frac{(26-27.94)^2}{27.94} = 0.135$$

$$6: \frac{(20-12.11)^2}{12.11} = 5.141$$

$$\text{7 or more: } \frac{(4-6.52)^2}{6.52} = 0.977$$

$$\chi^2 = 0.112 + 0.026 + 0.071 + 0.265 + 0.056 + 0.135 + 5.141 + 0.977$$

$$\chi^2 = 6.783$$

8 classes, 2 constraints (total and λ estimated)

$\nu = 8 - 2 = 6$

Critical value, 5% level, $\nu = 6$

critical value = 12.592

$6.783 < 12.592$, so do not reject H_0

A Poisson model is suitable for goals per match

c) More than 4 goals

$P(X \leq 4) = 0.8774$

$P(X > 4) = 1 - 0.8774 = 0.1226$

expected = $380 \times 0.1226 = 46.6$

observed = $26 + 20 + 4 = 50$

About 47 expected, 50 observed: close

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

12 The heights of 300 Year 7 pupils were measured by different members of staff and recorded to the nearest millimetre. The last digit of each recorded height was noted.

Last digit	0	1	2	3	4	5	6	7	8	9
Frequency	58	19	24	22	21	49	23	25	31	28

a. Test, at the 1% level of significance, whether the last digit is equally likely to be any of 0 to 9. State your hypotheses, the degrees of freedom and the critical value clearly.

b. Identify which digits contribute most to the test statistic and suggest a reason in context.

c. The 107 heights ending in 0 or 5 are removed. Test, at the 5% level of significance, whether the remaining digits are equally likely.

(10 marks)

a)

H_0: each last digit 0 to 9 is equally likely

H_1: the last digits are not equally likely

$E = 300 \div 10 = 30$

Each contribution $= (O - E)^2 \div E$

0: $\frac{(58 - 30)^2}{30} = 26.133$

1: $\frac{(19 - 30)^2}{30} = 4.033$

2: $\frac{(24 - 30)^2}{30} = 1.200$

3: $\frac{(22 - 30)^2}{30} = 2.133$

4: $\frac{(21 - 30)^2}{30} = 2.700$

5: $\frac{(49 - 30)^2}{30} = 12.033$

6: $\frac{(23 - 30)^2}{30} = 1.633$

7: $\frac{(25 - 30)^2}{30} = 0.833$

8: $\frac{(31 - 30)^2}{30} = 0.033$

9: $\frac{(28 - 30)^2}{30} = 0.133$

$\chi^2 = 26.133 + 4.033 + 1.200 + 2.133 + 2.700 + 12.033 + 1.633 + 0.833 + 0.033 + 0.133$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$$\chi^2 = 50.867$$

10 classes, 1 constraint

$\nu = 10 - 1 = 9$

Critical value, 1% level, $\nu = 9$

critical value = 21.666

$50.867 > 21.666$, so reject H_0

The last digits are not uniformly distributed

b) Biggest contributions

digit 0: 26.133

digit 5: 12.033

Staff rounded some heights to the nearest 5 mm

c) Remaining heights

$300 - 58 - 49 = 193$

$$E = 193 \div 8 = 24.125$$

H_0: digits 1, 2, 3, 4, 6, 7, 8, 9 are equally likely

H_1: they are not equally likely

Each contribution = $(O - E)^2 \div E$

$$1: \frac{(19 - 24.125)^2}{24.125} = 1.089$$

$$2: \frac{(24 - 24.125)^2}{24.125} = 0.001$$

$$3: \frac{(22 - 24.125)^2}{24.125} = 0.187$$

$$4: \frac{(21 - 24.125)^2}{24.125} = 0.405$$

$$6: \frac{(23 - 24.125)^2}{24.125} = 0.052$$

$$7: \frac{(25 - 24.125)^2}{24.125} = 0.032$$

$$8: \frac{(31 - 24.125)^2}{24.125} = 1.959$$

$$9: \frac{(28 - 24.125)^2}{24.125} = 0.622$$

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.

$$\chi^2 = 1.089 + 0.001 + 0.187 + 0.405 + 0.052 + 0.032 + 1.959 + 0.622$$

$$\chi^2 = 4.347$$

8 classes, 1 constraint

$\nu = 8 - 1 = 7$

Critical value, 5% level, $\nu = 7$

critical value = 14.067

$4.347 < 14.067$, so do not reject H_0

No evidence the other eight digits are not equally likely

Free for classroom and home use: print, photocopy and share. Not for resale or re-hosting.