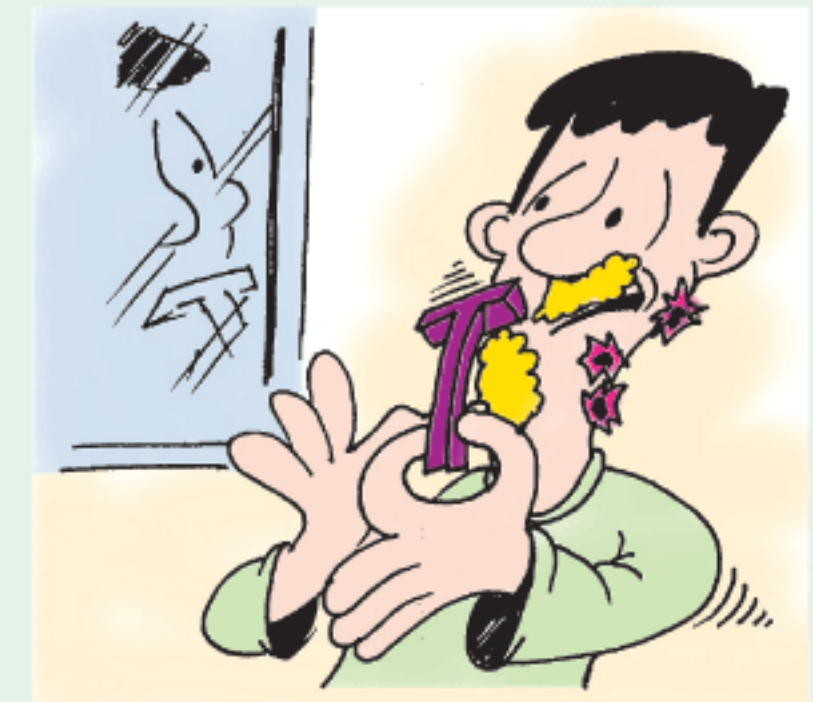


REVIEW SET 16A

- 1 Buses serving route 033F are scheduled to arrive at stop 2 at 9:45 am each day. The bus company wants to determine whether the buses serving this route have been arriving late to stop 2. Write down the set of hypotheses that the bus company should consider.
- 2 A null hypothesis H_0 and an alternative hypothesis H_1 are tested at a 10% significance level. A p -value ≈ 0.0794 was obtained for this particular hypothesis test.
 - a Interpret the p -value.
 - b What does the significance level represent?
 - c Is there enough evidence to reject H_0 in this case?

- 3 Quickshave produces disposable razorblades. They claim that the mean number of shaves before a blade has to be thrown away is 13. A researcher wishing to test the claim asks 30 men to supply data on how many shaves they get from one Quickshave blade. The researcher found the sample mean was 12.8 and the sample standard deviation was 1.6.



Perform a t -test with a 5% level of significance to test the manufacturer's claim.

- 4 Rosario owns an apricot orchard. Last year, the mean weight of his apricots was 90 grams. Rosario fears that severe droughts this year may have reduced the weight of his apricots. To address his concerns, Rosario randomly selected 20 apricots from his current harvest and recorded their weights in grams:

88 72 93 71 86 94 70 99 86 80
92 93 88 78 83 72 79 75 78 84

- a Write down the hypotheses that Rosario should test.
 - b Are Rosario's concerns justified at the 1% level?
- 5 Joe and Ruben are friends who like to fish in their free time. Joe is very competitive and claims that he is a better fisherman than Ruben.

The table below summarises the number of fish that Joe and Ruben have caught during their last 12 fishing trips:

	Mean	Standard deviation
Joe	$\bar{x}_1 = 10.9$	$s_1 \approx 3.34$
Ruben	$\bar{x}_2 = 10.25$	$s_2 \approx 2.26$

- a Write down the hypotheses that Joe should use to test his claim.
 - b Use a two-sample t -test with a 5% level of significance to test Joe's claim.
- 6 Kelly recorded the time spent shopping, in minutes, by a sample of customers in two neighbouring supermarkets.

Supermarket A: 12 28 13 7 22 19 4 13 6 11

Supermarket B: 14 35 32 21 14 8 2 16 24 27 19 42

Conduct a two-sample t -test to determine whether there is a significant difference between the time spent shopping by customers at the supermarkets. Use a 10% level of significance.

- 7 When a menswear store opens, the sizes of its shirts are distributed according to the percentages below.

Size	Small	Medium	Large	X-Large	XX-Large
Percentage (%)	10	20	35	25	10

In the first week, the store sells 70 shirts. The number sold of each size is shown in this table:

Size	Small	Medium	Large	X-Large	XX-Large
Number sold	4	7	22	24	13

To determine whether the sales are consistent with the distribution of the sizes stocked by the store, a χ^2 goodness of fit test is performed at a 5% level of significance.

- a Write down the:
- null hypothesis
 - number of degrees of freedom.
- b Calculate the p -value for the test.
- c Decide whether the store should change the distribution of shirt sizes that it stocks.
- 8 A guide for a particular role playing game lists the following percentages for the rarity of items obtained when a “loot box” is opened.

Item rarity	Percentage chance
super rare	5%
rare	10%
uncommon	25%
common	60%

Emmanuel thinks that the guide is incorrect. To test his suspicions, he opens 250 loot boxes. His results are shown in the table below.

Item rarity	Frequency
super rare	5
rare	17
uncommon	76
common	152
Total	250



- a Calculate the expected frequency for each item rarity.
- b Conduct a χ^2 goodness of fit test with a 1% level of significance to determine whether Emmanuel’s suspicions are justified. Use a critical value to make your decision.
- 9 The table alongside shows the responses to a survey about whether the city speed limit should be increased. Conduct a χ^2 test for independence at a 10% level to determine whether there is any association between the *age of a driver* and their *opinion* on the speed limit.

	Age of driver		
	18 to 30	31 to 54	55+
Increase	234	169	134
No increase	156	191	233

REVIEW SET 16B

1 Quickchick grow chickens to sell to a supermarket chain. However, the buyers believe that the supplied chickens are lighter than the minimum advertised weight of 1.2 kg.
What set of hypotheses should the buyers consider to investigate their concerns?

2 A particular χ^2 test has 6 degrees of freedom (df) and significance level $\alpha = 0.05$.
a State the critical value χ_{crit}^2 for this test.
b The test statistic obtained for this test was $\chi_{calc}^2 \approx 5.71$. Is there sufficient evidence to reject the null hypothesis?



3 As part of a routine health check of its employees, a company wants to check whether their systolic blood pressure is too high. High systolic blood pressure is generally diagnosed when a blood pressure test is more than 140 mm Hg.

A doctor measures the systolic blood pressure of a random sample of 35 company employees. She finds that the sample mean is 143.7 mm Hg with standard deviation 11.2 mm Hg.
Are the company's concerns justified on a 5% level of significance?

4 The average distance Arthur can hit a golf ball is 115 metres.
After spending time with a professional, Arthur measured the distance of 30 drives. The results in metres were as follows:

100	126	93	171	131	94	136	144	138	110
168	132	100	49	156	119	119	150	146	139
149	145	122	56	140	118	115	73	105	133

Is there sufficient evidence at the 5% level to claim that Arthur has improved?

5 To decide the credit limit of a prospective credit card holder, a bank gives points based on factors such as employment, income, home and car ownership, and general credit history.
The points totals of randomly selected people living in the suburbs Maple Grove and Berkton are shown below:

<i>Maple Grove:</i>	14	11	13	13	15	12	12	12	10	11	11	11	12	13
	14	13	11	12	14	14	14	13	15	14	11	10	11	16
	11	12	12	10	11	10	10	12	13	13	13	12		
<i>Berkton:</i>	11	10	12	14	11	12	14	11	11	14	13	13	14	13
	12	12	14	11	12	11	12	12	11	12	13	11	12	13
	12	12	10	10	12	11	11	9	14	10	13	13	10	12

Conduct a two-sample t -test with a 10% level of significance to determine whether there is a difference between the average points total of the two suburbs.

6 The following scores are the final IB examination results for students who took a revision course and students who did not:

<i>Revision course:</i>	32	39	31	35	40	33	34	33	34	35	39	33	32	35
<i>No revision course:</i>	28	31	30	23	33	31	36	36	38	35				

- a** Find the mean and standard deviation of each sample.
- b** Conduct a two-sample t -test with a 10% level of significance to determine whether the revision course was effective.

- 7** A toy company claims to manufacture glass, agate, alabaster, and onyx marbles in the ratio 4 : 2 : 2 : 1. The marbles are sold in bags of 50.

Aggie bought a bag of marbles and counted the number of each type. Her results are shown in the table.

Test the manufacturer's claim using a goodness of fit test with a 5% significance level.

Type	Frequency
glass	19
agate	16
alabaster	13
onyx	2

- 8** Consider the contingency table alongside. Test whether the variables P and Q are independent using a χ^2 test:

- a** at a 5% level of significance
b at a 1% level of significance.

	Q_1	Q_2	Q_3	Q_4
P_1	19	23	27	39
P_2	11	20	27	35
P_3	26	39	21	30

- 9** The following table shows the results from an investigation considering *level of education* and *business success*.

		Education level			
		High school	Graduate certificate	Undergraduate degree	Postgraduate
Business success	No success	35	30	41	25
	Low success	28	41	26	29
	Success	35	24	41	56
	High success	52	38	63	72

At a 1% level with $df = 9$, the critical value is 21.67. Test at a 1% level whether there is a link between *education level* and *business success*.

EXERCISE 16E.1

1 a

	Drove to work	Cycled to work	Public transport
Male	25.3	7.7	11
Female	20.7	6.3	9

b

	Junior school	Middle school	High school
Plays sport	38.28	56.76	69.96
Does not play sport	19.72	29.24	36.04

c

	Wore hat and sunscreen	Wore hat or sunscreen	Wore neither
Sunburnt	10.92	6.16	3.92
Not sunburnt	28.08	15.84	10.08

2 a

	Pass Maths test	Fail Maths test
Male	30	20
Female	30	20

b In a sample of 100 students, we would expect 30 to be male and pass the Maths test.

c

f_{obs}	f_{exp}	$f_{\text{obs}} - f_{\text{exp}}$	$(f_{\text{obs}} - f_{\text{exp}})^2$	$\frac{(f_{\text{obs}} - f_{\text{exp}})^2}{f_{\text{exp}}}$
24	30	-6	36	1.2
26	20	6	36	1.8
36	30	6	36	1.2
14	20	-6	36	1.8
Total				6

$\chi^2_{\text{calc}} = 6$

3 $\chi^2_{\text{calc}} \approx 6.61$, $df = 2$, $p \approx 0.0368$
As $p < 0.05$, we reject H_0 , and conclude that the variables *weight* and *suffering diabetes* are dependent.

4 a $\chi^2_{\text{crit}} = 4.61$
b $\chi^2_{\text{calc}} \approx 8.58$, $df = 2$, $p \approx 0.0137$
As $\chi^2_{\text{calc}} > 4.61$, we reject H_0 . So at a 10% level, we conclude that *age* and the *party they wish to vote for* are dependent.

5 a $\chi^2_{\text{calc}} \approx 23.6$, $df = 3$, $p \approx 0.000\,029\,9$
As $p < 0.05$, we reject H_0 . So at a 5% level, *reason for travelling* and *rating* are dependent.

b Guests travelling for a holiday are more likely to give a higher rating.

6 $\chi^2_{\text{calc}} \approx 7.94$, $df = 6$, $p \approx 0.242$
As $p > 0.1$, we do not reject H_0 . So at a 10% level, *position* and *injury type* are independent.

7 a

	Own a pet?	
	Yes	No
0 - 19	≈ 4.02	≈ 3.98
20 - 29	≈ 27.1	≈ 26.9
30 - 49	≈ 50.2	≈ 49.8
50+	≈ 36.7	≈ 36.3

b Yes, 4.02 and 3.98.

c

	Own a pet?	
	Yes	No
0 - 29	37	25
30 - 49	42	58
50+	39	34

d $\chi^2_{\text{calc}} \approx 5.22$, $df = 2$, $p \approx 0.0735$
As $p > 0.05$, we do not reject H_0 . So at a 5% level, *age* and *owning a pet* are independent.

8 a

	Intelligence level			
	Low	Average	High	Very high
Non smoker	≈ 262	≈ 383	≈ 114	≈ 4.69
Medium level smoker	≈ 133	≈ 194	≈ 57.7	≈ 2.38
Heavy smoker	≈ 107	≈ 157	≈ 46.6	≈ 1.93

b $\chi^2_{\text{calc}} \approx 16.9$, $df = 6$, $p = 0.009\,59$
As $\chi^2_{\text{calc}} > 16.81$, we reject H_0 . So at a 1% level, we conclude that *intelligence level* and *cigarette smoking* are not independent.

c

	Intelligence level			Sum
	Low	Average	High/Very high	
Non smoker	279	386	98	763
Medium level smoker	123	201	63	387
Heavy smoker	100	147	66	313
Sum	502	734	227	1463

d $\chi^2_{\text{calc}} \approx 13.2$, $df = 4$, $p = 0.0104$
As $p > 0.01$, we do not reject H_0 . So at a 1% level, *intelligence level* and *cigarette smoking* are independent. This is a different conclusion from the one in **b**.

EXERCISE 16E.2

1 a

	Result	
	Heads	Tails
Guess	49.4	54.6
Tails	45.6	50.4

b $\chi^2_{\text{calc}} \approx 1.35$

c As $\chi^2_{\text{calc}} < 3.84$, we do not reject H_0 . So at a 5% level, Horace's *guess* and *result* are independent.

d According to this test, Horace's claim is not valid.

2 a

	Result	
	Pass	Fail
France	63.8	21.2
Germany	168.2	55.8

b $\chi^2_{\text{crit}} = 2.71$

c $\chi^2_{\text{calc}} \approx 4.62$

d As $\chi^2_{\text{calc}} > 2.71$, we reject H_0 . So at a 10% level, *motorbike test result* and *country* are dependent.

REVIEW SET 16A

- $H_0: \mu = 0$ {the bus arrives on time}
 $H_1: \mu > 0$ {the bus is late}
- a** There is 7.94% chance of observing this result if the null hypothesis is true.
b For a 10% significance level, we reject H_0 if there is less than 10% chance of observing this result.
c As the p -value $< 0.1 = \alpha$, we reject H_0 in favour of H_1 .
- There is insufficient evidence to reject the manufacturer's claim at the 5% level of significance.
- a** $H_0: \mu = 90$ and $H_1: \mu < 90$
b Rosario's concerns are justified.
- a** Let μ_1 be the population mean number of fish caught by Joe per fishing trip, and μ_2 be the population mean number of fish caught by Ruben per fishing trip.
 $H_0: \mu_1 = \mu_2$ and $H_1: \mu_1 > \mu_2$
b Joe's claim is not justified at a 5% level of significance.
- There is sufficient evidence to support the claim that there is a significant difference between the time spent shopping by customers at the supermarkets at a 10% level.

- 7 a Let $p_1, p_2, p_3, p_4,$ and p_5 be the population proportions of shirts which are small, medium, large, X-large, and XX-large respectively.
- i $H_0: p_1 = 0.1, p_2 = 0.2, p_3 = 0.35, p_4 = 0.25, p_5 = 0.1$
- ii $df = 4$
- b ≈ 0.0134
- c Yes, the p -value < 0.05 , so we reject H_0 in favour of H_1 on a 5% level of significance. Since we accept H_1 , we conclude that the store should change its distribution.

8 a

Item rarity	super rare	rare	uncommon	common
Expected frequency	12.5	25	62.5	150

- b $\chi_{\text{crit}}^2 = 11.34, df = 3, \chi_{\text{calc}}^2 \approx 10.0$
As $\chi_{\text{calc}}^2 < \chi_{\text{crit}}^2$, there is insufficient evidence to justify Emmanuel's suspicions at a 1% significance level.
- 9 $\chi_{\text{calc}}^2 \approx 42.1, df = 2, p = 7.37 \times 10^{-10}$
As $\chi_{\text{calc}}^2 > 4.61$, we reject H_0 at a 10% significance level.
The variables *age of a driver* and their *opinion* are dependent.

REVIEW SET 16B

- 1 $H_0: \mu = 1.2$ {the mean minimum weight of Quickchick chickens is 1.2 kg}
 $H_1: \mu < 1.2$ {the mean minimum weight of Quickchick chickens is less than 1.2 kg}
- 2 a 12.59
b As $\chi_{\text{calc}}^2 < \chi_{\text{crit}}^2$, there is insufficient evidence to reject H_0 .
- 3 There is sufficient evidence to justify the company's concerns that the systolic blood pressure of its employees is too high on a 5% level of significance.
- 4 There is insufficient evidence at the 5% significance level to support the claim that Arthur has improved.
- 5 There is insufficient evidence to support the claim that the average points total of the two suburbs is significantly different at a 10% significance level.
- 6 a
- | | Mean | Standard deviation |
|--------------------|--------------------------|--------------------|
| Revision course | $\bar{x}_1 \approx 34.6$ | $s_1 \approx 2.82$ |
| No revision course | $\bar{x}_2 = 32.1$ | $s_2 \approx 4.48$ |
- b There is sufficient evidence to support the claim that the revision course improved results at a 10% level of significance.
- 7 $\chi_{\text{calc}}^2 = 5.215, df = 3, p \approx 0.157$
As $p > 0.05$, there is insufficient evidence to reject the manufacturer's claim at a 5% significance level.
- 8 $\chi_{\text{calc}}^2 \approx 13.0, df = 6, p \approx 0.0433$
- a As $p < 0.05$, we reject H_0 . So, at a 5% level, P and Q are dependent.
- b As $p > 0.01$, we do not reject H_0 . So, at a 1% level, P and Q are independent.
- 9 $\chi_{\text{calc}}^2 \approx 25.6, df = 9, p \approx 0.00241$
As $\chi_{\text{calc}}^2 > 21.67$, we reject H_0 . So at a 1% level, *education level* and *business success* are dependent.

EXERCISE 17A

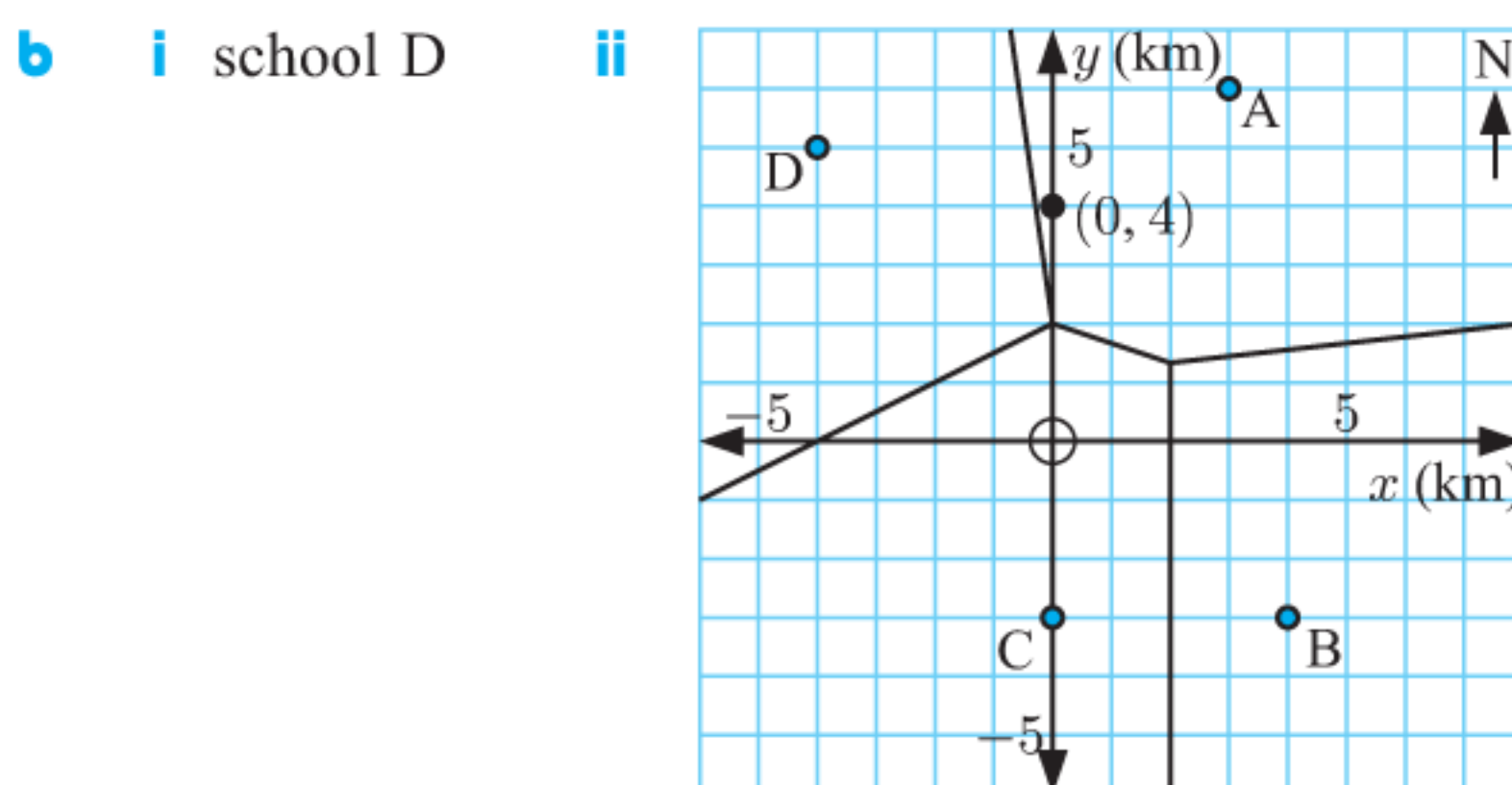
- 1 a i 3 ii 3 iii 1
b i Site B as P lies in cell B.
ii $PB = \sqrt{10} \approx 3.16$ units, $PA = 3\sqrt{2} \approx 4.24$ units, $PC = \sqrt{26} \approx 5.10$ units

- c i Q lies on the edge between cells A and C. Q is equally closest to sites A and C.
ii $QA = \sqrt{10} \approx 3.16$ units, $QC = \sqrt{10} \approx 3.16$ units, $QB = 3\sqrt{2} \approx 4.24$ units

- 2 a All points in the green cell are closest to site A than any other site.
b All points on the blue edge are equally closest to sites D and C.
c The red vertex is equally closest to sites A, B, and C.
- 3 a i site B ii site D iii site C iv site A

- b i distance to site A = $\sqrt{5}$ units
distance to site D = $\sqrt{5}$ units
ii distance to site A = $\sqrt{13}$ units
distance to site D = $\sqrt{13}$ units
- c $(-3, 0)$ is equally closest to sites A and D, hence it lies on an edge. $(-3, 2)$ is closest to site B and does not lie on an edge.
- d 32 units^2
- 4 a true b true c not necessarily true
d not necessarily true
- 5 If the circle passed through another site then P would lie on an edge. If another site was contained in the circle, P would be closer to that site than X . P would not lie in cell X .
- 6 a i post office E ii post office B iii post office A
iv post office C
b $(-1, \frac{1}{2})$ and $(-1, -2)$

- 7 a i school C ii school A iii school D iv school B



- iii distance to site A = $\sqrt{13}$ km
distance to site D = $\sqrt{17}$ km

- c i $(0, 2)$ ii 5 km

- 8 A vertex of a Voronoi diagram is equally closest to at least 3 sites (whose cells meet at that vertex). The circle's edge passes through site X , which is one of the sites V is closest to, and since every point on the edge is equidistant from the centre, the other closest sites must also lie on the edge of the circle.

EXERCISE 17B

