
Chapter 8.

The χ^2 Tests

The distribution of a categorical variable in a sample often needs to be compared with the distribution of a categorical variable in another sample. For example, over a period of 2 years a psychiatrist has classified by socioeconomic class the women aged 20-64 admitted to her unit suffering from self poisoning sample A. At the same time she has likewise classified the women of similar age admitted to a gastroenterological unit in the same hospital sample B. She has employed the Registrar General's five socioeconomic classes, and generally classified the women by reference to their father's or husband's occupation. The results are set out in [Table 8.1](#) .

Table 8.1 Distribution by socioeconomic class of patients admitted to self poisoning (sample A) and gastroenterological (sample B) units				
Socioeconomic class	Samples		Total	Proportion in group A
	A	B		
	a	b	n = a + b	p = a/n
I	17	5	22	0.77
II	25	21	46	0.54
III	39	34	73	0.53
IV	42	49	91	0.46
V	32	25	57	0.56
Total	155	134	289	

The psychiatrist wants to investigate whether the distribution of the patients by social class differed in these two units. She therefore erects the null hypothesis that there is no difference between the two distributions. This is what is tested by the chi squared (χ^2) test. By default, all χ^2 tests are two sided.

It is important to emphasize here that χ^2 tests may be carried out for this purpose only on the *actual numbers* of occurrences, not on percentages, proportions, means of observations, or other derived statistics. Note, we distinguish here the Greek (χ^2) for the test and the distribution and the Roman (χ^2) for the calculated statistic, which is what is obtained from the test.

The χ^2 test is carried out in the following steps:

For each observed number (O) in the Table find an "expected" number (E); this procedure is discussed below.

Subtract each expected number from each observed number	$O - E$
Square the difference	$(O - E)^2$
Divide the squares so obtained for each cell of the Table by the expected number for that cell	$(O - E)^2 / E$
χ^2 is the sum of $(O - E)^2 / E$	

To calculate the expected number for each cell of the Table consider the null hypothesis, which in this case is that the numbers in each cell are proportionately the same in sample A as they are in sample B. We therefore construct a parallel Table in which the proportions are exactly the same for both samples. This has been done in columns (2) and (3) of [Table 8.2](#). The proportions are obtained from the totals column in [Table 8.1](#) and are applied to the totals row. For instance, in [Table 8.2](#), column (2), $11.80 = (22/289) \times 155$; $24.67 = (46/289) \times 155$; in column (3) $10.20 = (22/289) \times 134$; $21.33 = (46/289) \times 134$ and so on.

Thus by simple proportions from the totals we find an expected number to match each observed number. The sum of the expected numbers for each sample must equal the sum of the observed numbers for each sample, which is a useful check. We now subtract each expected number from its corresponding observed number.

Class (I)	Expected numbers		O - E		$(O-E)^2/E$	
	A (2)	B (3)	A (4)	B (5)	A (6)	B (7)
I	11.80	10.20	5.20	-5.20	2.292	2.651
II	24.67	21.33	0.33	-0.33	0.004	0.005

III	39.15	33.85	-0.15	0.15	0.001	0.001
IV	48.81	42.19	-6.81	6.81	0.950	1.009
V	30.57	26.43	1.43	-1.43	0.067	0.077
Total	30.57	134.00	0	0	3314	3.833

$$\chi^2 = 3.314 + 3.833 = 7.147. \text{ d.f.} = 4. 0.10 < P < 0.50.$$

The results are given in columns (4) and (5) of [Table 8.2](#) . Here two points may be noted.

1. The sum of these differences always equals zero in each column.
2. Each difference for sample A is matched by the same figure, but with opposite sign, for sample B.

Again these are useful checks.

The figures in columns (4) and (5) are then each squared and divided by the corresponding expected numbers in columns (2) and (3). The results are given in columns (6) and (7). Finally these results, $(O - E)^2 / E$ are added. The sum of them is χ^2 .

A helpful technical procedure in calculating the expected numbers may be noted here. Most electronic calculators allow successive multiplication by a constant multiplier by a short cut of some kind. To calculate the expected numbers a constant multiplier for each sample is obtained by dividing the total of the sample by the grand total for both samples. In [Table 8.1](#) for sample A this is $155/289 = 0.5363$. This fraction is then successively multiplied by 22, 46, 73, 91, and 57. For sample B the fraction is $134/289 = 0.4636$. This too is successively multiplied by 22, 46, 73, 91, and 57.

The results are shown in [Table 8.2](#) , columns (2) and (3).

Having obtained a value for $\chi^2 = \sum[(O - E)^2 / E]$ we look up in a Table of χ^2 distribution the probability attached to it ([Appendix Table C](#)). Just as with the t Table, we must enter this Table at a certain number of degrees of freedom. To ascertain these requires some care.

When a comparison is made between one sample and another, as in [Table 8.1](#) , a simple rule is that the degrees of freedom equal (number of columns minus one) x (number of rows minus one) (not counting the row and column containing the totals). For the data in [Table 8.1](#) this gives $(2 - 1) \times (5 - 1) = 4$. Another way of looking at this is to ask for the minimum number of Figures that must be supplied in [Table 8.1](#) , *in addition to all the totals*, to allow us to complete the whole Table. Four numbers disposed anyhow in samples A and B provided they are in separate rows will suffice.

Entering [Table C](#) at four degrees of freedom and reading along the row we find that the value of χ^2 (7.147) lies between 3.357 and 7.779. The corresponding probability is: $0.10 < P < 0.50$. This is well above the conventionally significant level of 0.05, or 5%, so the null hypothesis is not disproved. It is therefore quite conceivable that in the distribution of the patients between socioeconomic classes the population from which sample A was drawn were the same as the population from which sample B was drawn.

Quick method

The above method of calculating χ^2 illustrates the nature of the statistic clearly and is often used in practice. A quicker method, similar to the quick method for calculating the standard deviation, is particularly suitable for use with electronic calculators⁽¹⁾.

The data are set out as in [Table 8.1](#). Take the left hand column of figures (Sample A) and call each observation a. Their total, which is 155, is then Σa .

Let p = the proportion formed when each observation a is divided by the corresponding figure in the total column. Thus here p in turn equals $17/22$, $25/46$... $32/57$.

Let \bar{p} = the proportion formed when the total of the observations in the left hand column, Σa , is divided by the total of all the observations.

Here $\bar{p} = 155/289$. Let $\bar{q} = 1 - \bar{p}$, which is the same as $134/289$.

Then

$$\chi^2 = \frac{\Sigma pa - \bar{p}\Sigma a}{\bar{p}\bar{q}}$$

Calculator procedure

Working with the figures in [Table 8.1](#), we use this formula on an electronic calculator (Casio fx-350) in the following way:

$$17 \text{ Inv } x^2 \div 22 \text{ Min}$$

$$25 \text{ Inv } x^2 \div 46 \text{ M+}$$

$$39 \text{ Inv } x^2 \div 73 \text{ M+}$$

$$42 \text{ Inv } \chi^2 \div 91 \text{ M+}$$

$$32 \text{ Inv } \chi^2 \div 57 \text{ M+}$$

$$155 \text{ Inv } \chi^2 \div 289 \text{ M+}$$

Withdraw result from memory on to display screen

MR (1.7769764)

We now have to divide this by $\bar{p} \times \bar{q}$ Here $\bar{p} = 155/289$ and $\bar{q} = 134/289$.

$$1.776975 \times (289 \div 155) \times (289 \div 134)$$

This gives us $\chi^2 = 7.146$.

The calculation naturally gives the same result if the figures for sample B are used instead of those for sample A. Owing to rounding off of the numbers the two methods for calculating χ^2 may lead to trivially different results.

Fourfold tables

A special form of the χ^2 test is particularly common in practice and quick to calculate. It is applicable when the results of an investigation can be set out in a "fourfold table" or "2 x 2 contingency table".

For example, the practitioner whose data we displayed in believed that the wives of the printers and farmers should be encouraged to breast feed their babies. She has records for her practice going back over 10 years, in which she has noted whether the mother breast fed the baby for at least 3 months or not, and these records show whether the husband was a printer or a sheep farmer (or some other occupation less well represented in her practice). The figures from her records are set out in [Table 8.3](#)

The disparity seems considerable, for, although 28% of the printers' wives breast fed their babies for three months or more, as many as 45% of the farmers' wives did so. What is its significance?

Table 8.3 Numbers of wives of printers and farmers who breast fed their babies for less than 3 months or for 3 months or more

	Breast fed for		
	Up to 3 months	3 months or more	Total
Printers' wives	36	14	50
Farmers' wives	30	25	55
Total	66	39	105

The null hypothesis is set up that there is no difference between printers' wives and farmers' wives in the period for which they breast fed their babies. The χ^2 test on a fourfold table may be carried out by a formula that provides a short cut to the conclusion. If a, b, c, and d are the numbers in the cells of the fourfold table as shown in Table 8.4 (in this case Variable 1 is breast feeding (<3 months 0, ≥ 3 months 1) and Variable 2 is husband's occupation (Printer (0) or Farmer (1)), χ^2 is calculated from the following formula:

$$\chi^2 = \frac{(ad - bc)^2 (a + b + c + d)}{(a + b)(c + d)(b + d)(a + c)}$$

With a fourfold table there is one degree of freedom in accordance with the rule given earlier.

Table 8.4 Notation for two group χ^2 test

		Variable 1		
		0	1	Total
Variable 2	0	a	b	a + b
	1	c	d	c + d
Total		a + c	b + d	a + b + c + d

As many electronic calculators have a capacity limited to eight digits, it is advisable not to do all the multiplication or all the division in one series of operations, lest the number become too big for the display.

Calculator procedure

Multiply a by d and store in memory

Multiply b by c and subtract from memory

Extract difference from memory to display

Square the difference

Divide by a + b

Divide by c + d

Multiply by a + b + c + d

Divide by b + d

Divide by a + c

$$ad - bc$$

$$(ad - bc)^2$$

$$\frac{(ad - bc)^2}{(a + b)}$$

$$\frac{(ad - bc)^2}{(a + b)(c + d)}$$

$$\frac{(ad - bc)^2(a + b + c + d)}{(a + b)(c + d)}$$

$$\frac{(ad - bc)^2(a + b + c + d)}{(a + b)(c + d)(b + d)}$$

$$\frac{(ad - bc)^2(a + b + c + d)}{(a + b)(c + d)(b + d)(a + c)}$$

From [Table 8.3](#) we have

$$\frac{[(36 \times 25) - (30 \times 14)]^2 \times 105}{33 \times 39 \times 55 \times 50} = 3.418$$

Entering the χ^2 [Table C](#) with one degree of freedom we read along the row and find that 3.418 lies between 2.706 and 3.841. Therefore $0.05 < p < 0.01$. So, despite an apparently considerable difference between the proportions of printers' wives and the farmers' wives breast feeding their babies for 3 months or more, the probability of this result or one more extreme occurring by chance is more than 5%.

We now calculate a confidence interval of the differences between the two proportions, as described in [Chapter 6](#). In this case we use the standard error based on the observed data, not the null hypothesis. We could calculate the confidence interval on either the rows or the columns and it is important that we compare proportions of the outcome variable, that is, breast feeding.

$$P_1 = 14/50 = 0.28, P_2 = 25/55 = 0.45, P_1 - P_2 = 0.17.$$

$$SE(P_1 - P_2) = \left(\sqrt{\frac{0.28 \times 0.72}{50} + \frac{0.45 \times 0.55}{55}} \right) = 0.0924$$

The 95% confidence interval is

$$0.17 - 1.96 \times 0.0924 \text{ to } 0.17 + 1.96 \times 0.0924 = -0.011 \text{ to } 0.351$$

Thus the 95% confidence interval is wide, and includes zero, as one might expect because the χ^2 test was not significant at the 5% level.

Increasing the precision of the P Value in 2 x 2 tables

It can be shown mathematically that if X is a Normally distributed variable, mean zero and variance 1, then X^2 has a χ^2 distribution with one degree of freedom. The converse also holds true and we can use this fact to improve the precision of our P values. In the above example we have $X = 1.85$, and from we find P to be about 0.065. However, we do need the χ^2 tables for more than one degree of freedom.

Small numbers

When the numbers in a 2 x 2 contingency table are small, the χ^2 approximation becomes poor.

The following recommendations may be regarded as a sound guide⁽²⁾. In fourfold tables a χ^2 test is inappropriate if the total of the Table is less than 20, or if the total lies between 20 and 40 and the smallest expected (not observed) value is less than 5; in contingency tables with more than one degree of freedom it is inappropriate if more than about one fifth of the cells have expected values less than 5 or any cell an expected value of less than 1. An alternative to the χ^2 test for fourfold tables is known as Fisher's Exact test and is described in [Chapter 9](#)

When the values in a fourfold table are fairly small a "correction for continuity" known as the "Yates' correction" may be applied⁽³⁾. Although there is no precise rule defining the

circumstances in which to use Yates' correction, a common practice is to incorporate it into χ^2 calculations on tables with a total of under 100 or with any cell containing a value less than 10.

The χ^2 test on a fourfold table is then modified as follows:

$$\frac{[(|ad - bc|) - 0.5(a + b + c + d)]^2 (a + b + c + d)}{(a + b)(c + d)(b + d)(a + c)}$$

The vertical bars on either side of $ad - bc$ mean that the smaller of those two products is taken from the larger. Half the total of the four values is then subtracted from that the difference to provide Yates' correction. The effect of the correction is to reduce the value of χ^2 .

Applying it to the figures in [Table 8.3](#) gives the following result:

$$\frac{[(36 \times 25) - (30 \times 14) - (105 \div 2)^2 \times 105]}{(66 \times 39 \times 55 \times 50)} = 2.711$$

In this case $\chi^2 = 2.711$ falls within the same range of P values as the $\chi^2 = 3.418$ we got without Yates' correction ($0.05 < P < 0.1$), but the P value is closer to 0.1 than it was in the previous calculation. In fourfold tables containing lower frequencies than [Table 8.3](#) the reduction in P value by Yates' correction may change a result from significant to non-significant; in any case care should be exercised when making decisions from small samples.

Comparing proportions

Earlier in this chapter we compared two samples by the χ^2 test to answer the question "Are the distributions of the members of these two samples between five classes significantly different?" Another way of putting this is to ask "Are the relative proportions of the two samples the same in each class?"

For example, an industrial medical officer of a large factory wants to immunize the employees against influenza. Five vaccines of various types based on the current viruses are available, but nobody knows which is preferable. From the work force 1350 employees agree to be immunized with one of the vaccines in the first week of December, 50 the medical officer divides the total into five approximately equal groups. Disparities occur between their total numbers owing to the layout of the factory complex. In the first week of the following March he examines the records he has been keeping to see how many employees got influenza and how many did not. These records are classified by the type of vaccine used ([Table 8.5](#)).

Table 8.5 People who did or did not get influenza after inoculation with one of five vaccines				
Type of vaccine	Numbers of employees			
	Got influenza	Avoided influenza	Total	Proportion got influenza
I	43	237	280	0.18
II	52	198	250	0.21
III	25	245	270	0.09
IV	48	212	260	0.18
V	57	233	290	0.20
Total	2255	1125	1350	

In [Table 8.6](#) the figures are analyzed by the χ^2 test. For this we have to determine the expected values. The null hypothesis is that there is no difference between vaccines in their efficacy against influenza. We therefore assume that the proportion of employees contracting influenza is the same for each vaccine as it is for all combined. This proportion is derived from the total who got influenza, and is 225/1350. To find the expected number in each vaccine group who would contract the disease we multiply the actual numbers in the Total column of [Table 8.5](#) by this proportion. Thus $280 \times (225/1350) = 46.7$; $250 \times (225/1350) = 41.7$; and so on. Likewise the proportion who did not get influenza is 1125/1350.

The expected numbers of those who would avoid the disease are calculated in the same way from the totals in [Table 8.5](#), so that $280 \times (1125/1350) = 233.3$; $250 \times (1125/1350) = 208.3$; and so on.

The procedure is thus the same as shown in [Table 8.1](#) and [Table 8.2](#) .

The calculations made in [Table 8.6](#) show that χ^2 with four degrees of freedom is 16.564, and $0.001 < P < 0.01$. This is a highly significant result. But what does it mean?

Table 8.6 Calculation of χ^2 test on figures in Table 8.5						
Type of vaccine	Expected numbers		O - E		$(O - E)^2/E$	
	Got influenza	Avoided influenza	Got influenza	Avoided influenza	Got influenza	Avoided influenza
I	46.7	233.3	-3.7	3.7	0.293	0.059
II	41.7	208.3	10.3	-10.3	2.544	0.509
III	45.0	225.0	-20.0	20.0	8.889	1.778
IV	43.3	216.7	4.7	-4.7	0.510	0.102
V	48.3	241.7	8.7	-8.7	1.567	0.313
Total	225.0	1125.0	0	0	13.803	2.761

$\chi^2 = 16.564$, d.f. = 4, $0.001 < P < 0.01$.

Splitting of χ^2

Inspection of [Table 8.6](#) shows that the largest contribution to the total χ^2 comes from the figures for vaccine III. They are 8.889 and 1.778, which together equal 10.667. If this figure is subtracted from the total χ^2 , $16.564 - 10.667 = 5.897$. This gives an approximate figure for χ^2 for the remainder of the Table with three degrees of freedom (by removing the vaccine III we

have reduced the Table to four rows and two columns). We then find that $0.1 < P < 0.5$, a non-significant result. However, this is only a rough approximation. To check it exactly we apply the χ^2 test to the figures in Table 8.4 minus the row for vaccine III. In other words, the test is now performed on the figures for vaccines I, II, IV, and V. On these figures $\chi^2 = 2.983$; d.f. = 3; $0.1 < P < 0.5$. Thus the probability falls within the same broad limits as obtained by the approximate short cut given above. We can conclude that the figures for vaccine III are responsible for the highly significant result of the total χ^2 of 16.564.

But this is not quite the end of the story. Before concluding from these figures that vaccine III is superior to the others we ought to carry out a check on other possible explanations for the disparity. The process of randomization in the choice of the persons to receive each of the vaccines should have balanced out any differences between the groups, but some may have remained by chance. The sort of questions worth examining now are: Were the people receiving vaccine III as likely to be exposed to infection as those receiving the other vaccines? Could they have had a higher level of immunity from previous infection? Were they of comparable socioeconomic status? Of similar age on average? Were the sexes comparably distributed? Although some of these characteristics could have been more or less balanced by stratified randomization, it is as well to check that they have in fact been equalized before attributing the numeral discrepancy in the result to the potency of the vaccine.

χ^2 Test for trend

Table 8.1 is a 5 x 2 table, because there are five socioeconomic classes and two samples. Socioeconomic groupings may be thought of as an example of an ordered categorical variable, as there are some outcomes (for example, mortality) in which it is sensible to state that (say)

social class II is between social class I and social class III. The χ^2 test described at that stage did not make use of this information; if we had interchanged any of the rows the value of χ^2 would have been exactly the same. Looking at the proportions p in Table 8.1 we can see that there is no real ordering by social class in the proportions of self poisoning; social class V is between social classes I and II. However in many cases, when the outcome variable is an ordered categorical variable, a more powerful test can be devised which uses this information.

Table 8.7 Change in eating poultry in randomized trial ⁽⁴⁾					
	Intervention	Control	Total	Proportion in intervention	Score
	a	b	n	$p=a/n$	x
Increase	100	78	178	0.56	1
No change	175	173	348	0.50	0
Decrease	42	59	101	0.42	-1
Total	317	310	627	0.51	

Consider a randomized controlled trial of health promotion in general practice to change people's eating habits⁽⁵⁾. Table 8.7 gives the results from a review at 2 years, to look at the change in the proportion eating poultry.

If we give each category a score x the χ^2 test for trend is calculated in the following way:

$$E_{xp} = \frac{\sum ax - \sum a \sum nx}{N}$$

and

$$E_{xx} = \frac{\sum nx^2 - (\sum nx)^2}{N}$$

then

$$\chi^2 = E_{xp}^2 / (E_{xx} \bar{p} \bar{q})$$

where:

N is the total sample size

$\bar{p} = \sum / an$ and $\bar{q} = \sum / bn$

$\sum a = 317$

$N = 627$

$\sum ax = 100 \times 1 + 175 \times 0 - 42 \times 1 = 658$

$\sum nx = 178 \times 1 + 348 \times 0 - 101 = 77$

$\sum xp = 58 - 317 \times 77 / 627 = 19.07$

$\sum nx^2 = 178 \times 1^2 + 348 \times 0^2 + 101 \times (-1)^2 = 279$

$(\sum nx)^2 / N = 77^2 / 627 = 9.46$

Thus

$$\Sigma xx = 279 - 9.46 = 269.54$$

$$\bar{p} = 317/627 = 0.5056 \quad \bar{q} = 310/627 = 0.4944$$

$$\chi^2 = 19.07^2 / 279.54 \times 0.5056 \times 0.4944 = 5.20$$

This has one degree of freedom because the linear scoring means that when one expected value is given all the others are fixed, and we find $p = 0.02$. The usual χ^2 test gives a value of $\chi^2 = 5.51$; d.f. = 2; $0.05 < P < 0.10$. Thus the more sensitive χ^2 test for trend yields a significant result because the test used more information about the experimental design. The values for the scores are to some extent arbitrary. However, it is usual to choose them equally spaced on either side of zero. Thus if there are four groups the scores would be -3, -1, +1, +3, and for five groups -2, -1, 0, +1, +2. The χ^2 statistic is quite robust to other values for the scores provided that they are steadily increasing or steadily decreasing.

Note that this is another way of splitting the overall χ^2 statistic. The overall χ^2 will always be greater than the χ^2 for trend, but because the latter uses only one degree of freedom, it is often associated with a smaller probability. Although one is often counseled not to decide on a statistical test after having looked at the data, it is obviously sensible to look at the proportions to see if they are plausibly monotonic (go steadily up or down) with the ordered variable, especially if the overall χ^2 test is nonsignificant.

Table 8.8 Calculation of χ^2 for comparison between actual distribution and theoretical distribution					
Mice	Observed cases	Theoretical proportions	Expected cases	O - E	(O - E)²/E
Entirely white	380	0.510	400	-20	1.0000
Small brown patch	330	0.408	320	10	0.3125
Large brown patch	74	0.082	64	10	1.5625
Total	784	1.000	784	0	2.8750

Comparison of an observed and a theoretical distribution

In the cases so far discussed the observed values in one sample have been compared with the observed values in another. But sometimes we want to compare the observed values in one sample with a theoretical distribution.

For example, a geneticist has a breeding population of mice in his laboratory. Some are entirely white, some have a small patch of brown hairs on the skin, and others have a large patch. According to the genetic theory for the inheritance of these colored patches of hair the population of mice should include 51.0% entirely white, 40.8% with a small brown patch, and 8.2% with a large brown patch. In fact, among the 784 mice in the laboratory 380 are entirely white, 330 have a small brown patch, and 74 have a large brown patch. Do the proportions differ from those expected?

The data are set out in [Table 8.8](#) . The expected numbers are calculated by applying the theoretical proportions to the total, namely 0.510×784 , 0.408×784 , and 0.082×784 . The degrees of freedom are calculated from the fact that the only constraint is that the total for the expected cases must equal the total for the observed cases, and so the degrees of freedom are the number of rows minus one. Thereafter the procedure is the same as in previous calculations of χ^2 . In this case it comes to 2.875. The χ^2 Table is entered at two degrees of freedom. We find that $0.2 < P < 0.3$. Consequently the null hypothesis of no difference between the observed distribution and the theoretically expected one is not disproved. The data conform to the theory.

McNemar's test

McNemar's test for paired nominal data was described in , using a Normal approximation. In view of the relationship between the Normal distribution and the χ^2 distribution with one degree of freedom, we can recast the McNemar test as a variant of a χ^2 test. The results are often expressed as in [Table 8.9](#) .

Table 8.9 Notation for the McNemar test				
		First subject of pair		
		Variable 1		
	Variable 2	0	1	Total
2nd subject of pair	0	e	f	e + f
	1	g	h	g + h
Total		e + g	f + h	n

Table 8.10 Data from for McNemar's test			
		First subject of pair	
		Responded	Did not respond
2nd subject of pair	Responded	16	10
	Did not respond	23	5

McNemar's test is then

$$X^2 = \frac{(f - g)^2}{(f + g)} \text{ with 1 d.f.}$$

or with a continuity correction

$$X_c^2 = \frac{(|f - g| - 1)^2}{(f + g)} \text{ (with 1 d.f.)}$$

The data from are recast as shown in [Table 8.10](#) . Thus

$$X^2 = \frac{(10 - 23)^2}{(10 + 23)} = 5.12$$

or

$$X_c^2 = \frac{(|10 - 23| - 1)^2}{(10 + 23)} = 4.36$$

From [Table C \(Appendix\)](#) we find that for both X^2 values $0.02 < P < 0.05$. The result is identical to that given using the Normal approximation described in [Chapter 6](#), which is the square root of this result.

Extensions of the χ^2 test

If the outcome variable in a study is nominal, the χ^2 test can be extended to look at the effect of more than one input variable, for example to allow for confounding variables. This is most easily done using *multiple logistic regression*, a generalization of *multiple regression*, which is described in [Chapter 11](#). If the data are matched, then a further technique (*conditional logistic regression*) should be employed. This is described in advanced textbooks and will not be discussed further here.

Common questions

I have matched data, but the matching criteria were very weak. Should I use McNemar's test?

The general principle is that if the data are matched in any way, the analysis should take account of it. If the matching is weak then the matched analysis and the unmatched analysis should agree. In some cases when there are a large number of pairs with the same outcome, it would appear that the McNemar's test is discarding a lot of information, and so is losing power. However, imagine we are trying to decide which of two high jumpers is the better. They each jump over a bar at a fixed height, and then the height is increased. It is only when one fails to jump a given height and the other succeeds that a winner can be announced. It does not matter how many jumps both have cleared.

References

1. Snedecor GW, Cochran WG. In: *Statistical Methods*, 7th ed. Iowa: Iowa State University Press, 191,0:47.
2. Cochran WG. Some methods for strengthening the common χ^2 tests. *Biometrics* 1956; 10 :417-51.
3. Yates F. Contingency tables involving small numbers and the χ^2 test. *J Roy Stat Soc Suppl* 1934; 1:217-3.
4. Capples ME, McKnight A. Randomized controlled trial of health promotions in general practice for patients at high cardiovascular risk. *BMJ* 1994;309:993-6.

Exercises

Exercise 8.1 In a trial of new drug against a standard drug for the treatment of depression the new drug caused some improvement in 56% of 73 patients and the standard drug some improvement in 41% of 70 patients. The results were assessed in five categories as follows:

New treatment: much improved 18, improved 23, unchanged 15, worse 9, much worse 8;

Standard treatment: much improved 12, improved 17, unchanged 19, worse 13, much worse 9.

What is the value of χ^2 which takes no account of the ordered value of data, what is the value of the χ^2 test for trend, and the P value? How many degrees of freedom are there? What is the value of P in each case?

Exercise 8.2 An outbreak of *pediculosis capitis* is being investigated in a girls' school containing 291 pupils. Of 130 children who live in a nearby housing estate 18 were infested and of 161 who live elsewhere 37 were infested. What is the χ^2 value of the difference, and what is its significance? Find the difference in infestation rates and a 95% confidence interval for the difference.

Exercise 8.3 The 55 affected girls were divided at random into two groups of 29 and 26. The first group received a standard local application and the second group a new local application. The efficacy of each was measured by clearance of the infestation after one application. By this measure the standard application failed in ten cases and the new application in five. What is the χ^2 value of the difference (with Yates' correction), and what is its significance? What is the difference in clearance rates and an approximate 95% confidence interval?

Exercise 8.4 A general practitioner reviewed all patient notes in four practices for 1 year. Newly diagnosed cases of asthma were noted, and whether or not the case was referred to hospital. The following referrals were found (total cases in parentheses): practice A, 14 (103); practice B, 11 (92); practice C, 39 (166); practice D, 31 (221). What are the χ^2 and P values for the distribution of the referrals in these practices? Do they suggest that any one practice has significantly more referrals than others?