# An Introduction to tests for Categorical Outcomes

In all the exercises so far this afternoon focussed singularly on scenarios where our dependent (AKA outcome) variable is a **continuous variable**. For example, the **number of items** recalled correctly during a memory assessment, the **standardised score** on an IQ test.

We have, of course, come across **categorical variables** too. However, up until now we’ve only considered how to use these as independent (AKA predictor) variables in our analyses. For example: the effects of **Gender** (being *Male* or *Female*); **Group** (tested in the *Morning* versus tested in the *Afternoon*)*.*

In this final tutorial we’re going to look at the analysis options open to you when your study is interested in analysing the influence of other variables upon a **categorical dependent variable**. For example, often in medically-orientated research there is a particular interest in understanding what factors in the present predict whether someone will or won’t go on to develop a certain health problem in the future (e.g. Depression, Schizophrenia, Obesity)? This type of dependent variable is what we’d call a *dichotomous (or binomial) categorical outcome*. This isn’t a continuous variable because there are only two possible response values (i.e. two groups a participant can fall into): a person either will develop the specific health problem or they won’t. For these sorts of analyses we need to look to different types of statistical tests, which are suited to evaluating how variables influence this specific type of **categorical**outcome.

As you might imagine, we as psychologists are often as interested in these types of outcomes as we are continuous ones. For example: imagine we want to improve our understanding of how our behaviour has an impact on categorical states we may fall into regarding our educational attainment (e.g. passing or failing), relationship status (e.g. single, in a relationship, married etc...), mental well-being (high-, medium- or low-risk of suicide) etc… .

The relevant chapters in the Andy Field Textbook for this exercise is:

**Chapter 19: Categorical Outcomes: Chi-square & Log-linear analysis**

In this week’s exercises we’ll first focus on a basic type of categorical analysis that will establish some key concepts in this field of analysis. Then the second exercise will introduce a variant of the standard regression model we’ve worked with previously, which with a few minor modifications works as well for analysing effects upon categorical dependent variables as it does continuous ones. Specifically, we’ll focus on:

* + - 1. How to test if **a single categorical independent variable** has an influence on a categorical dependent variable
      2. How to test if **more than one independent variable (either continuous or categorical)** has an influence on a categorical dependent variable

**Exercise 1: 1 Categorical outcome and 1 Categorical predictor**

Let’s begin with the simplest form of categorical analysis: in which we have only two variables that are both categorical, each only has two levels and we only wish to predict whether one variable has an influence on the other. In other words, we identify one of these two categorical variables as being the independent variable and the other as the dependent variable in our design

**Note:** the theory I’m about to discuss is covered in a more detail in the **Andy Field textbook, Chapter 19, Section 19.3**

As an example, let’s imagine I’m interested in determining if elderly individuals who are identified as having ‘high’ blood pressure are at greater risk of being diagnosed with Vascular Dementia at some point in the 5 years following that blood pressure test, compared to those where categorised as having ‘normal’ blood pressure. The best way to describe such data is in the form of a **contingency table**, which displays the frequencies for each of these two variables across rows and columns (see example below).

|  |  |  |  |
| --- | --- | --- | --- |
| *Table 1 (observed)* | **DV: Vascular Dementia** | |  |
| **IV: Blood Pressure** | *Yes* | *No* | **Total** |
| *High* | 28 | 48 | 76 |
| *Normal* | 10 | 114 | 124 |
| **Total** | 38 | 162 | 200 |

**Null Hypothesis testing using contingency tables**

Considering our hypothesis in terms of the contingency table, we can think of it as taking the form of a prediction that: *the category that a participant falls into within the independent variables alters the likelihood of which category they will fall into within the dependent variable*. In other words, being a ‘High’ for blood pressure increases your odds of also being a ‘Yes’ for Vascular Dementia compared to if you were a ‘Normal’ for blood pressure. We test this hypothesis by testing how significantly different from **chance** the frequency distribution is across the cells highlighted in the yellow in the table.

**Extra tips/pointers:**

These values are sometimes referred to as the ‘counts’ in each cell of the table (e.g. 28 participants were ‘counted’ as having high blood-pressure and vascular dementia)

To flip our hypothesis around, to consider it as a null hypothesis that we will actually test our data against, we ask if the distribution of observations we see in the table are likely due to ‘random chance’ allocation. For example, even if there was no relationship between blood pressure and vascular dementia and we looked at the 200 elderly individuals in this example we might expect some of these people who happened to have high blood pressure would go on to develop dementia. Similarly, we’d expect some with normal blood pressure to go on to develop dementia too. If there was no systematic effect of blood pressure on your risk of dementia we’d subsequently anticipate that the proportion of people who *overall* go on to develop Dementia would be very similar to the proportions *within each Blood pressure group* that go on to develop dementia (i.e. the risk of developing dementia is approximately the same irrespective if you’re ‘High’ or ‘Normal’ for your blood pressure).

The question we’re therefore interested in asking statistiscally is whether the actual **observed frequency distribution** (illustrated in the table above) is significantly different enough from an **expected frequency distribution** based on the null-hypothesis, as described in the previous paragraph.

Can you illustrate in the table below on the next page the values you would expect in each cell were distributions to be consistent with 20% of the sample going on to develop vascular dementia and the null hypothesis also being true:

|  |  |  |  |
| --- | --- | --- | --- |
| *Table 2 (expected)* | **DV: Vascular Dementia** | |  |
| **IV: Blood Pressure** | *Yes* | *No* | **Total** |
| *High* | 15 | 61 | 76 |
| *Normal* | 25 | 99 | 124 |
| **Total** | 40 | 160 | 200 |

Now comparing the values you entered into the yellow boxes in this ‘expected’ frequency distribution with the ‘observed’ distribution in the contingency table on the previous page. What key differences do you notice? Also, if you were to speculate (without doing any further analysis) would you be in favour of rejecting the null hypothesis?

If you compare these expected values to the actual (observed) data there’s quite clear deviation from the actual counts. A much higher proportion of those individuals with high blood pressure appear to also have developed vascular dementia than we might otherwise have anticipated if there was no relationship (i.e. 36% [28/76] rather than 20% [15/76]). Conversely, far fewer people with normal blood pressure develop vascular dementia than expected (i.e. 8% [10/124] rather than 20% [25/124]).

This discrepancy would have me doubtful that the data is consistent with the null hypothesis.

**Chi-squared testing and Expected Frequencies**

To formally test the statistical significance of the null hypothesis in this example we need to use a certain family of tests called **chi-squared tests (**χ2**).** The easiest of these to understand (in computational terms) is the Pearson Chi-square, which we’ll apply to the example we’ve been working through.

In the previous example I implied that we could work out the expected frequencies in the null hypothesis if we knew 20% of our sample was likely to go on and develop dementia. This was a convenient ‘sleight of hand’ though, to allow us to then compare expected against observed frequencies. Of course, in practice we may not be privy to this sort of information[[1]](#footnote-1) prior to collecting our sample. Instead, all we will have is our frequency distributions in the ‘margins’ of our contingency table (i.e. proportion of total sample with high blood pressure and proportion of total sample with vascular dementia) but we can work back from these to estimate the expected frequency in each cell using the following equation:

Above Table 3 (on the next page) I’ve used this equation to calculate the expected frequency for having high blood pressure and vascular dementia, based on the Margins in Table 1. Can you complete the rest of the blanks in Table 3 using the same method?

|  |  |  |  |
| --- | --- | --- | --- |
| *Table 3* | **DV: Vascular Dementia** | |  |
| **IV: Blood Pressure** | *Yes* | *No* | **Total** |
| *High* | 14.4 | 61.6 | 76 |
| *Normal* | 23.6 | 100.4 | 124 |
| **Total** | 38 | 162 | 200 |

**Extra tips/pointers:**

Notice how similar these are to the expected frequencies you worked out in Table 2. This illustrates the fact that when can also simply use observed data about the distribution of response within our Independent and Dependent variables to work out what the expected frequency distribution under the null. Once we have this we can then contrast this with the observed distribution, as well now do…

To quantify the degree of difference between the expected and observed frequencies we calculate the Chi-squared (χ2)statistic[[2]](#footnote-2). This is difference between an Expected versus Observed frequency squared, for each ‘count’ cell, divided by that cell’s expected value and then summed together for all count cells. This is essentially a measure of how much deviance from the null hypothesis there is in your observed frequencies.

To demonstrate this in action I’ve calculated below the value χ2 statistic based on the observed cell counts (i.e. Table 1) versus the expected cell counts (i.e. Table 3):

This pearson chi-squared test has 1 degree of freedom[[3]](#footnote-3) and the value we have for χ2 here is therefore statistically significant because this value (i.e. 26.52) is much larger than the critical value for χ2 at one degree of freedom for rejecting the null at the 5% level (if χ2 ≥ 3.84 then p < .05). This means we would reject the null hypothesis in this scenario and find support for our alternative hypothesis: that blood pressure level did influence the risk of going on to develop vascular dementia.

Can you now calculate the value you would get for χ2 if, instead of using the cell counts from Table 1 you used the values you calculated and entered in Table 2 (i.e. this would represent a scenario where the observed frequencies were very similar to the expected frequencies. Can you also interpret the statistical significance of this χ2 (i.e. based on the cut-off value mentioned for rejecting null hypothesis at the p<.05 level in the previous paragraph)?

The χ2 test here still only has one degree of freedom so the critical value we need our χ2 statistic to be above for statistical significance (i.e. p < .05) is still 3.84.

The value we have for χ2 now isn’t larger than that value though (i.e. 0.269). So in this scenario we would not reject the null-hypothesis.

This is hardly surprising given the fake data you created for Table 2 is intended to represent a rough estimate of what the cell-counts would be if they were to be in-line with the expected cell counts under the null hypothesis!

**Assumptions and limitations for Chi-Squared Testing**

We’ll discuss the several different types of chi-squared tests SPSS can do for us in a moment but before doing this it is best to highlight some assumptions that need to be taken into account when running any of the following types of analysis:

1. Observations must be independent in your sample. This rules out using this method of analysis if you have a repeated measures experimental design.

**Extra tips/pointers:**

An easier way of thinking about this is that each participant **can only be ‘counted’ once in the contingency table**. That is, they either get counted as a ‘High+Yes’, ‘High+No’, ‘Normal+Yes’ or a ‘Normal+No’ in the yellow cells in the example tables on the previous pages. They can’t both be a ‘High+Yes’ and a ‘High+No’ but just at different points in the time course of your study.

1. We have to have the expected frequencies for each possible response that are greater than 5 or (in large contingency tables, no more than 20% of the expected frequencies can be below 5 but all should be above 1). Think of this as having enough data that could conceivably be consistent with the null hypothesis for it to be testable
2. Another point Field Makes (**Chapter 19, Section 19.5.3**) is that in interpreting statistical significance we should look at row and column ***percentages*** rather than raw frequency scores as the latter can be misleading because their magnitudes are influenced by the actual number of observations we have in our sample. This is not an assumption as such but something to bear in mind to avoid risk of over-interpretation!

Lastly, bear in mind that the results of a chi-squared test can be interpreted much like the results of an F-value for an interaction in an ANOVA. That is, it is a test of whether the interaction between the two variables is statistically significant but it tells us nothing of the direction of this relationship (e.g. does High Blood Pressure increase or reduce the risk of Vascular Dementia and/or does Low Blood Pressure do the same or opposite?). To explore this further we need to look at the **standardised residuals** from our output (something I’ll come back to in the following example).

**Using SPSS to analyse Categorical Data**

The example we’re going to work uses the data in **Exercise1\_dataset.sav** file, which represents the following study:

**Using Crosstab Contingency Tables in SPSS**

Imagine you are a social psychologist interested in superstitious beliefs in sports people. Specifically, you are interested in what influences football managers to believe (or not) in the concept of home-field advantage1 (i.e. believing you are more likely to win when playing at home). You ask a sample of football managers whether or not they believe in home-advantage (Believe vs. Don’t Believe) and see whether their likelihood of believing in this effect depends on a number of different factors, including whether:

1. Their mentor (i.e. the first manager they worked under) believed in this advantage
2. Whether in the past season they have won more times than they have lost at home
3. Whether they consider the team to have a clear ‘local rival’ they play ‘derby’ matches against on a regular basis

1inspired by Van Den Ven, N. (2011), Supporters are not necessary for Home Advantage: Evidence from same-stadium derbies and games without an audience. *Journal of Applied Psychology,* 41(12). 2785-2792

To test the first hypothesis in the box on page 5 (i.e. whether their mentor believed in home-advantage) we’re going to use the **Crosstabs…** contingency tables function. To run this analysis:

1. Go to **Analyze > Descriptive Statistics > Crosstabs…**
2. In the Crosstabs window, select your Independent Variable (*Mentor\_Blf*) and move it to the **Row(s):** box.
3. Next select your Dependent Variable (*Man\_Blf*) and move it to the **Column(s):** box.
4. Open the **Statistics…** sub-window and in here select the following tests to allow us to test our hypotheses regarding the frequency distributions within this contingency table:
   * Chi-square
   * Contingency Coefficient
   * Phi and Cramer’s V

Then click **Continue** to close this sub-window.

1. We’ll also ask for **Fisher’s Exact Test**, which is a variation on the regular Chi-square tests. Whilst not all that necessary in this example, it is a test recommended when you have a 2x2 contingency table like in our example but you have a much smaller dataset. To select this click on **Exact…** and in the sub-window that pops up change to the **Exact** option, then click **Continue** to close this sub-window.
2. Finally, using the **Cells…** sub-window we need to request some additional output to help check assumptions and aid interpretation. Open this up and make sure the following are selected:
   * **Counts**: both observed and expected
   * **Percentages:** Row, Column and Total

To explore further any significant effect of mentor belief’s (should we get one) also ensure the following are selected:

* + **Residuals:** Standardized

Then click **Continue** to close this sub-window.

1. To run the analysis click **OK**.

To interpret the output take the following steps:

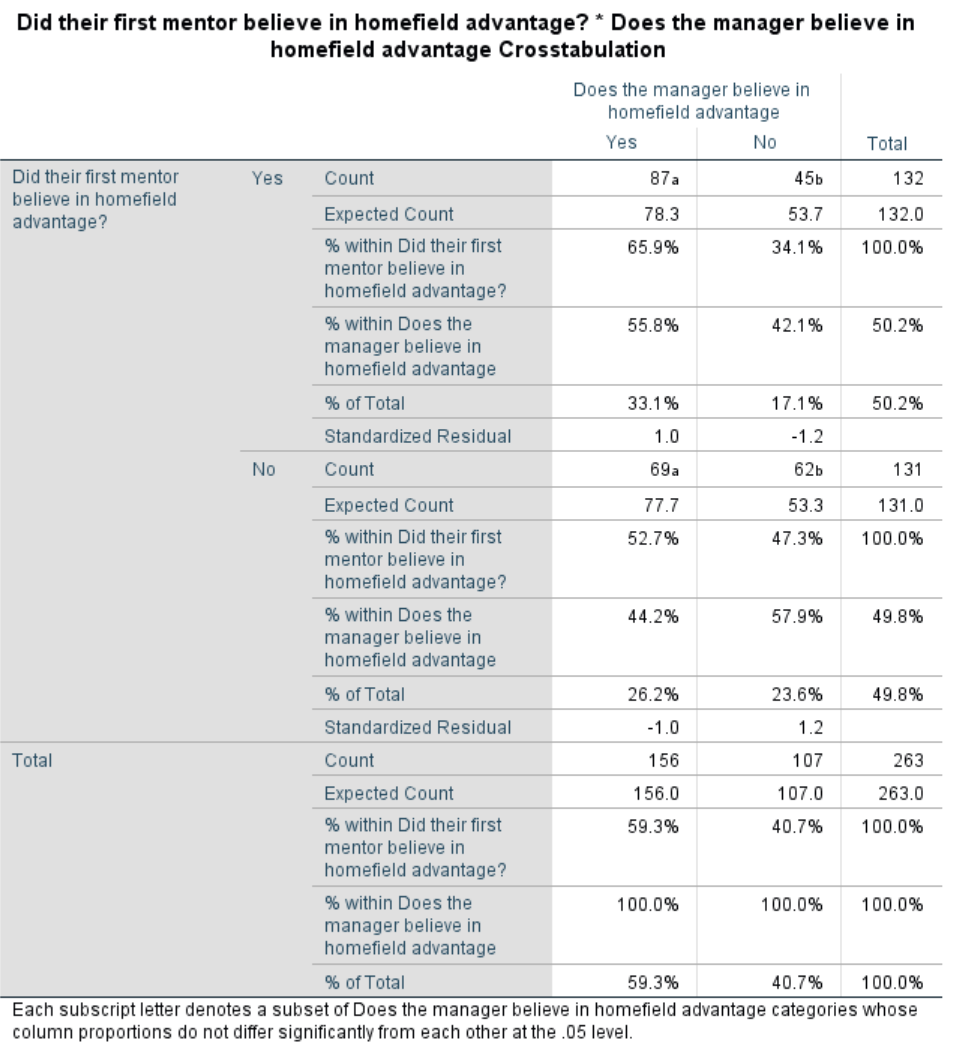
* **Check Assumptions**

In the **Crosstabulation** table (which shows *Mentors’ Beliefs [Mentor\_Blf]* by *Managers’ Beliefs [Man\_Blf]* in home advantage) first look at the Expect Count values for each cell (see figure on the next page illustrating where you can find them) and evaluate whether our assumptions for running this type of analysis have been met? If you think they have why this is (*Hint:* refer back to point 2 on page 5)?

All the expected counts here are well above 5 so we’re okay to assume we’ve met this assumption

What assumption would we violate if we changed the research question to ask whether football players’ beliefs in home advantage were influenced by their manager’s belief? In such a scenario we’d likely as teams of player from each club to participate and what problems would this present in terms of assumptions?

We’d violate the assumption of independence if we looked at more than one player from the same football club because our independent variable (i.e. the manager’s belief) would get counted more than once in the contingency table. In a similar vein, we’d have to make sure that out of all of the managers in this sample none of them considered the same individual as having been their mentor (i.e. each manager would need to have had a *different* mentor for the assumption of independence to have been met for this current analysis.



Look Along this row…

… and this row.

… and this row…

* **Investigate Chi-Squared Results: Testing our Hypothesis**

In the **Chi-Square Tests** table the row we’re primarily interested in is the *Pearson Chi-Square* and in particular its *Asymptotic Significance*. As with all other kinds of significance testing we’ve come across, the fact that p < .05 here (i.e. p = .029) tells us that we have strong grounds to reject the null hypothesis that a manager’s belief in home advantage is not influence by their mentor’s beliefs. In other words, we have grounds to go on and explore further the possible patterns of difference between manager and mentor belief to ascertain whether (i) managers are more likely to believe in home advantage if their mentors did or (ii) managers are less likely to believe in home advantage if their mentors did or

You’ll notice we also get a number of other versions of the chi-square test reported in this table:

* The Yates’ **Continuity Correction**, which is an adjusted version of the Pearson Chi-Square that corrects for the fact that the Pearson test can often be overly lenient in 2x2 contingency table and thus is a greater risk of making a type-1 error. However, Field suggests (see **Chapter 19, Section 19.3.5**) that Yates’s is perhaps too conservative a correction (i.e. it goes too far in the opposite direction).
* The **Likelihood Ratio**, which is a type of Chi-square we came across last week when discussing how to compare model fit in multi-level linear models.
* **Fisher’s Exact test**, which is provided as both a 2-sided and 1-sided test. Although, as I’ve stated earlier, this variation is probably unnecessary given how large our sample size is.

Reassuringly, all these variations of the χ2 arrive at the same conclusion, that we should reject the null hypothesis (i.e. p < .05)

* **Exploring further the Nature of Association**

To explore the nature of the significant association further we could look to describe it using the observed percentages in the **Crosstabulation**. Specifically we want to summarise our results in terms of either the percentages within the rows ***or*** the percentages within the columns (i.e. in instances where the percentages sum to 100% in the margins of the table). Below I’ve provided an example of how to provide this interpretation along the rows (i.e. **% of Managers believing if mentors did**):

Of the managers whose first mentor believed in home advantage a higher proportion of them also believed in this advantage (65.9%) than didn’t (34.1%). Meanwhile, of the managers whose first mentor didn’t in home advantage there was a much more marginal bias in favour of also not believing in this advantage (i.e. 47.3% did believe, whilst 52.7% did).

Can you provide an equivalent statement looking down the **columns (i.e. % of Mentors believing if Mentee’d Manager did)**

Of the managers who believed in home advantage, approximately 10% more of them had a mentor that also believed in home advantage (55.8%) than didn’t (44.2%). Meanwhile, in managers that didn’t believe in home advantage approximately 16% more of them had a mentor that also didn’t believe (57.9%) compared to those that did (42.1%).

We can be more objective in our interpretations by also looking at the **Standardized Residuals** in this table also. These tell us the strength of relationship between each cell within the contingency table and our overall significance results for our Chi-Square test (i.e. which cell(s), out of the four cells within this 2x2 table, are contributing more to the conclusion that there is systematic variability in the frequency distribution due to the independent variable?).

In terms of its absolute size a standardised residual can be treated like a z-score, meaning that a value >±1.96 indicates a cell is having a significant (at the p<.05 level) influence, whilst a residual value of >±2.58 is significant at the p < .01 level. Are there any clear relationships emerging here?

Unfortunately not really, there is systematic variation but the standardised residuals suggest it is not clearly being driven by any one cell or group of cells, making it difficult to draw firm conclusions. In other words, I’d be cautious about inferring too much here given the residuals aren’t giving us a clear picture.

* **Exploring further the strength of Associations**

So far we have some ambiguous evidence for a meaningful relationship between a manager’s belief in home-advantage and their mentors. The overall significance of the χ2 test suggests a relationship exists but the Standardised Residuals are inconclusive as to the nature of that relationship, suggesting that whilst there may be a relationship its **effect size** is likely small, and thus difficult to reliably detect (i.e. whatever the influence mentors have had it’s a relatively small contributing factor in influencing whether a manager also believes in home advantage). To properly assess effects sizes here we need to go beyond statistical significance and look a few additional things in follow-up testing:

Firstly, we can get a quick sense of the effect size for the overall chi-square result by looking at the **Symmetric Measures** table (this is a bit like interpreting R2 or . This table presents adjusted versions of the original chi-square test, which take into account sample size and degrees of freedom and try to restrict the range of the test statistic (i.e. the *Value* in this table) to fall between 0 to 1 (to make it similar to interpreting a correlation coefficient). Looking at these test statistics and reflecting on what you know about correlations what conclusions would you draw about the strength of the relationship between manager and mentor belief here?

If we evaluate these values in the same way we would pearson’s r correlation coefficients then we can see that the actual effect size is small (i.e. only a little over .1). These are still statistically significant relationships between a mentor’s and a manager’s belief in home advantage but this represents only a small proportion in the overall variability in managers what influences whether managers believe in home advantage. This makes sense, as it is logical that lots of other factors, besides their mentor’s views, will have helped to form these manager’s views over time about whether it is valid to believe in the phenomenon of a home advantage.

**Odds Ratio’s**

A more common method of reporting effect sizes for categorical outcomes is to use a statistic we call the **odds ratio**. To calculate odds ratios (**ORs**) of believing in home advantage if your mentor also did, compared to if they didn’t we need to do the following calculation, using the *Count* valuesin the **Crosstabulations** Table:

Odds Manage Yes AND Mentor Yes  = Manger ‘Yes’ & Mentor ‘Yes’ / Manger ‘No’ & Mentor ‘Yes’

= 87/45 = 1.93·

Odds Manager Yes BUT Mentor No = Manger ‘Yes’ & Mentor ‘No’ / Manger ‘No’ & Mentor ‘No’

= 69/62 = 1.11290323

Odds Ratio = Odds Manage Yes AND Mentor Yes  / Odds Manager Yes BUT Mentor No

= 1.93·/1.11290323 = **1.74**

This value of 1.74 for this Odds Ratio represents the odds of one outcome versus another (i.e. a manager believing in home advantage versus not believing) under each level of the independent variable (i.e. whether their mentor *did* or *did not* also believe in home advantage). We then look at the ratio of these two odds to see if our outcome is more likely (i.e. has higher odds) under one level of the IV compared to the other.

Lastly, we now need to understand what an odds ratio of 1.74 means! To do this follow these rules:

* Odds ratios are expressed as a single numerical value[[4]](#footnote-4)
* **An odds ratio (OR) > than 1** (with 95% CI around it that have a lower bound >1) suggest there is a **significantly elevated chance of the predicted outcome**, whereby the odds of falling in the category of interest are greater than into its comparator (i.e. you’re more likely believe in home advantage) if you exhibit the ‘risk factor’ (i.e. you’re mentor believed in home advantage). The size of the odds ratio is proportional to the increase in risk. For example, an OR of 2 would mean participants were twice as likely to believe in home advantage if their mentor also did, whilst an OR of 100 means they’re a hundred times more likely.
* **An odds ratio (OR) = 1** (and a 95% CI that has its lower bound <1 whilst its upper bound is >1) suggests there is no systematic relationship because you are as likely to fall within the category of interest as you are within the comparator group (i.e. you’re as likely to believe in home advantage as not) irrespective of if you fall within the ‘risk factor level’ for your predictor (i.e. your mentor believing in home advantage).

**Extra tips/pointers:**

It might help to think of this as the odds being at “fifty:fifty” or ‘chance’ level. After all what value do you get when you divide 50 by 50 (i.e. 1!)?

* **An odds ratio (OR) < than 1** (and a 95% CI around it that has an upper bound <1) suggests there is a significantly reduced chance of the predicted outcome, whereby the odds of falling in the category of interest are less than into its comparator (i.e. you’re less likely to believe in home-advantage) if you exhibit the ‘risk factor’ (i.e. your mentor believe in home advantage). The size of the odds ratio is proportional to the reduction in risk. For example, an OR of 0.5 would mean participants were half as likely to believe in home advantage if their mentor believed in home advantage, an OR of 0.01 means you are a hundred times less likely.

**Extra tips/pointers:**

Due to the inverse nature of this relationship you often hear ‘risk factors’ in such analysis instead be referred to as ‘protective factors’. After all, they reduce the likelihood of the (often negative) outcome you’re interested when running this type of analysis (e.g. regular exercise is a protective risk factor for heart disease because taking regular exercise reduces your odds of having a heart attack compared to those individuals who don’t take regular exercise). .

Given these interpretive points, how would you interpret the odds of a manager believing in home advantage if you’ve worked under a mentor that also believed in home advantage?

Also, more generally how would you evaluate the evidence for claiming that mentors’ beliefs in a home advantage influence the likelihood of managers’ also believing in a home-advantage?

The odds ratio here is greater than 1 suggesting managers who train under a mentor who believe in home advantage are 1.7 times more likely to also report believing in this advantage.

However, in making claims based on this finding though I’d be cautious in over-interpreting this result, as we know the effect size is small (from the standardised residual) and therefore whilst I might propose that there seems to be some weak evidence here for mentors influencing a manager’s opinion it is far from the only factor that plays into their beliefs about whether a home advantage effect exists.

To conclude Exercise 1, can you repeat the same analysis but this time test whether winning more often than not at home in the last year influences beliefs about a home advantage. Summarise what you find in a format similar to the summary below, which describes the effect of whether the team has a close local rival on belief in home advantage.

There was a significant association between having a close local rival and the manager believing in home advantage χ2(1) = 8.68, p = .003. The odds ratio indicated the odds of a manager believing in home advantage if you had a local rival, compared to if they didn’t have such a rival, were 0.46. In other words managers with a local rival were less than half as like as those that didn’t to believe in home advantage.

Odds Manage Yes AND Record Yes  = Manger ‘Yes’ & Record ‘Yes’ / Manger ‘No’ & Record ‘Yes’

= 125/23 = 5.435

Odds Manager Yes BUT Record No = Manger ‘Yes’ & Record ‘No’ / Manger ‘No’ & Record ‘No’

= 31/84 = 0.369

Odds Ratio = Odds Manage Yes AND Record Yes  / Odds Manager Yes BUT Record No

= 5.435/0.369

= **14.72**

There was a significant association between whether a team had a winning record at home and their manager believing in home advantage χ2(1) = 88.67, p < .001. The odds ratio this indicated the odds of a manager believing in home advantage if they had a winning home record, compared to if they didn’t, were 14.72. In other words managers with a winning record were at least 14 times more likely to believe in home advantage that those with losing home record.

Are these three results in line with your expectation and what would you infer from them about the factors that affect a manager’s belief in the concept of home advantage?

* The finding that manager’s beliefs are influenced by their mentors is perhaps unsurprising. It stands to reason that Managers my model their beliefs and behaviours on others working in the same profession, particularly those whom they aspire to become like.
* Similarly, the finding that a winning home record increases the odds of believing in home advantage is logical, as the team’s results are consistent with this belief that the team is likely to do well when playing at home.
* The finding that belief in home advantage is less common amongst managers who play regularly against local rivals is more interesting. This perhaps suggest that managers who regularly have to play important home and away fixtures against local rivals see home-advantage as a potentially detrimental belief to ascribe to in these circumstances. After all, in local derbies home advantage is (anecdotally) viewed as less important.

Having said all that, in the paper that inspired these example (Van Den Ven, 2011) they did actually find evidence of home advantage in derby fixtures (i.e. teams were more likely to win when playing at home, even when playing a rival that shared the same stadium as themselves). So whether belief actually influences performance is another questions entirely!

1. i.e. the overall expected prevalence of our the outcome (i.e. vascular dementia), irrespective of its association with the independent variable, is not a known value [↑](#footnote-ref-1)
2. which we can then use to compare against an expected sampling distribution to derive a p-value [↑](#footnote-ref-2)
3. Calculated as df = (Number of rows - 1)(Number of columns - 1) [↑](#footnote-ref-3)
4. As we’ll see in later exercises, we often reported a 95% confidence interval around this value too because, as ever, we are estimating the odds in the wider population based on the odds we’ve seen in our sample. [↑](#footnote-ref-4)