Measuring association in contingency tables

Patrick Breheny

April 2

Which association?   Risk difference
Confidence intervals   Relative risk
Summary   Odds ratio

Hypothesis tests and confidence intervals

- Fisher's exact test and the χ^2-test can be used to calculate p-values and assess evidence against the null

- This null hypothesis can be loosely described as saying that there is "no association" between treatment/exposure and the outcome, or more formally as the hypothesis that the two events are independent

- However, we also need to be able to measure and place confidence intervals on effect sizes – otherwise, we have no way of assessing the practical and clinical significance of the association

- So we want a confidence interval ... but what do we want a confidence interval of?

Which association?   Risk difference
Confidence intervals   Relative risk
Summary   Odds ratio

Difference in proportions

- One way of measuring the strength of an association for categorical data is to look at the difference in proportions
- For example, in Lister's experiment, 46% of the patients who received the conventional surgery died, but only 15% of the patients who received the sterile surgery died
- The difference in these percentages is 31 percentage points
- For the purposes of this class, I'll refer to this measure of association as the *risk difference*, but it is also known as the "excess risk" or "attributable risk"

Which association?
Confidence intervals
Summary

Risk difference
Relative risk
Odds ratio

Risk difference: Advantages and shortcomings

- The major advantage of the risk difference is that it directly translates to impact: the Lister experiment indicates that by implementing sterile surgical techniques, doctors would save 31 lives for every 100 surgeries

- However, differences in proportions have their shortcomings, particularly when used to describe rare events

- For example, it has been estimated that over a 10 year time span, for smokers the probability of developing lung cancer is 0.483%, while for nonsmokers, the probability is 0.045%

- The risk difference here is just 0.4 percentage points

- Quantifying the association in this manner makes it appear as though smoking has an extremely minor impact on lung cancer

Which association?   Risk difference
Confidence intervals   Relative risk
Summary   Odds ratio

The relative risk

- Instead, especially for rare events (like most diseases), we often describe the strength of an association using ratios

- If we quantify the association between smoking and lung cancer using ratios, we find that the probability of developing lung cancer is over 10 times higher for smokers: $0.483/0.045 = 10.7$

- This measure more accurately (and dramatically) conveys the strong association between smoking and lung cancer

- As another example, Lister's experiment: the risk of dying from surgery is three times lower ($46/15 = 3.1$) if sterile technique is used

- This ratio is called the *relative risk*, and it is often more informative than the risk difference

Which association?   Risk difference
Confidence intervals   **Relative risk**
Summary   Odds ratio

Absolute versus relative risks

- This is a good place to discuss absolute versus relative risks
- Smoking has a huge relative effect on lung cancer, but a fairly small absolute effect: even among smokers, lung cancer is quite rare
- As a contrast, let us consider as well smoking's effect on heart disease, where the U.S. health department has estimated that over a 10 year time span, for smokers the probability of developing coronary artery disease is 2.947%, while for nonsmokers, the probability is 1.695%
- This is a much larger risk difference (1.25 vs. 0.44), but a much smaller relative risk (1.7 vs 10.7)

Which association?   Risk difference
Confidence intervals   Relative risk
Summary   Odds ratio

Absolute versus relative risks (cont'd)

- Both measures convey important, and different, information
- In terms of understanding the biological causes of lung cancer, smoking is critical; on the other hand, lots of things cause heart disease
- In terms of the calculating the effect of preventative measures on a population, however, the risk difference is more meaningful
- For example, suppose a smoking cessation program caused 1,000 people to quit smoking
- This would likely save 4.4 lives by preventing lung cancer, but 12.5 lives by preventing coronary artery disease

Which association?   Risk difference
Confidence intervals   Relative risk
Summary   Odds ratio

Shortcomings of the relative risk

- The relative risk is a widely used measure of the strength of an association, but it too has shortcomings
- One is that it's asymmetric
- For example, the relative risk of dying is $46/15 = 3.1$ times greater with the nonsterile surgery, but the relative risk of living is only $85/54 = 1.57$ times greater with the sterile surgery

Which association?   Risk difference
Confidence intervals   Relative risk
Summary   Odds ratio

Relative risks and retrospective studies

- Another problem is that it doesn't work with retrospective studies
- For example, consider the results of a classic case-control study of the relationship between oral contraceptives and thromboembolic disease (blood clots) published in 1969:

	Cases	Controls
Used OC	42	23
Didn't use OC	42	145

- Is the probability of developing thromboembolic disease given that a person used oral contraceptives $42/(42 + 23) = 65\%$?

Which association?
Confidence intervals
Summary

Risk difference
Relative risk
Odds ratio

Relative risks and retrospective studies (cont'd)

- Absolutely not; this isn't even remotely accurate
- The whole design of the study is based around the idea of intentionally getting a large number of subjects with the disease; the risk of thromboembolism is not even close to 65% (in reality, it's well under 1% even for OC users)
- For retrospective and cross-sectional studies, then, we cannot calculate a relative risk
- This would require an estimate of the probability of developing a disease given that an individual was exposed to a risk factor, which we can only get from a prospective study
- Instead, retrospective studies yield the probability of exposure to a risk factor given that you have developed the disease

Which association?   Risk difference
Confidence intervals   Relative risk
Summary   **Odds ratio**

Odds

- A slightly different measure of association, the *odds ratio*, gets around both of these flaws

- Instead of taking the ratio of the probabilities, the odds ratio is a ratio of the odds of developing the disease given risk factor exposure to the odds given a lack of exposure

- The *odds* of an event is the ratio of the number of times the event occurs to the number of times the event fails to occur
  - For example, if the probability of an event is 50%, then the odds are 1; in speech, people usually say that "the odds are 1 to 1"
  - If the probability of an event is 75%, then the odds are 3; "the odds are 3 to 1"

Which association?
Confidence intervals
Summary

Risk difference
Relative risk
Odds ratio

The symmetry of the odds ratio

- As advertised, the odds ratio possesses the symmetry that the relative risk does not

- For example, in Lister's experiment the odds of dying were $6/34 = .176$ for the sterile group and $16/19 = .842$ for the control group

- The relative odds of dying with the control surgery is therefore $.842/.176 = 4.77$

- On the other hand, the odds of surviving were $34/6 = 5.67$ for the sterile group and $19/16 = 1.19$ for the control group

- The relative odds of surviving with the sterile surgery is therefore $5.67/1.19 = 4.77$

Which association?
Confidence intervals
Summary

Risk difference
Relative risk
Odds ratio

An easier formula for the odds ratio

- Summarizing this reasoning into a formula, if our table looks like

a	c
b	d

  then

$$\widehat{\mathrm{OR}} = \frac{ad}{bc}$$

- Because of this formula, the odds ratio was originally called the "cross-product ratio"

Which association?   Risk difference
Confidence intervals   Relative risk
Summary   Odds ratio

There are two odds ratios

- Keep in mind that there are two odds ratios, depending on how we ordered the rows and columns of the table, and that they will be reciprocals of one another
- When calculating and interpreting odds ratios, be sure you know which group has the higher odds of developing the disease
- In Lister's experiment, the odds ratio for surviving with the sterile surgery was 4.77, but the odds ratio for surviving with the control surgery was $1/4.77 = 0.210$
- NOTE: When writing about an odds ratio less than 1, it is customary to write, for example, that "the sterile procedure reduced the odds of death by 79%"

Which association? | Risk difference
Confidence intervals | Relative risk
Summary | Odds ratio

Odds ratios and retrospective studies

- The symmetry of the odds ratio works wonders when it comes to retrospective studies
- In our case-control study, the odds ratio for OC use given thromboembolism is

$$\widehat{\text{OR}} = \frac{42 \cdot 145}{23 \cdot 42}$$
$$= 6.3$$

- However, this is also the odds ratio for thromboembolism given OC use
- This is a minor miracle: we have managed to obtain a prospective measure of association from a retrospective study!
- Hence the popularity of the odds ratio: it can be used for any study design (prospective, retrospective, cross-sectional) that results in a 2x2 contingency table

Which association?
Confidence intervals
Summary

Risk difference
Relative risk
Odds ratio

Interpretation of odds ratios

- For the sake of this course, we are going to focus on odds ratios, although it is worth noting that many researchers consider relative risks easier to interpret, and prefer reporting them when possible

- Indeed, odds ratios are always larger in magnitude (i.e., further away from 1) than relative risks, something to keep in mind when interpreting clinical significance

- For example, we saw that for the Lister study, the relative risk was 3.1, while the odds ratio was 4.8

- An even more extreme example is the Nexium trial (recall that the healing rates in the two groups were 93% vs. 89%), where the relative risk is 1.04 but the odds ratio is 1.6

Which association?
Confidence intervals
Summary

Risk difference
Relative risk
Odds ratio

# The χ^2-test and measures of association

- Note that when the difference between two proportions equals 0, the relative risk equals 1 and the odds ratio equals 1

- Furthermore, when relative risk of disease given exposure equals 1, the relative risk of exposure given disease equals 1

- Indeed, all of these statements are equivalent to saying that exposure and disease are independent

- Thus, any one of these may be thought of as the null hypothesis of the χ^2-test

- This is why the null hypothesis is often loosely described as there being no "association" between exposure and disease

The odds ratio and the central limit theorem

- So far in this class, we've calculated confidence intervals for averages and percentages

- For both statistics, the central limit theorem guarantees that if the sample size is big enough, their sampling distribution will look relatively normal

- The central limit theorem, however, does not directly establish that the distribution of odds ratio will be normally distributed, and indeed, it tends to be skewed to the right
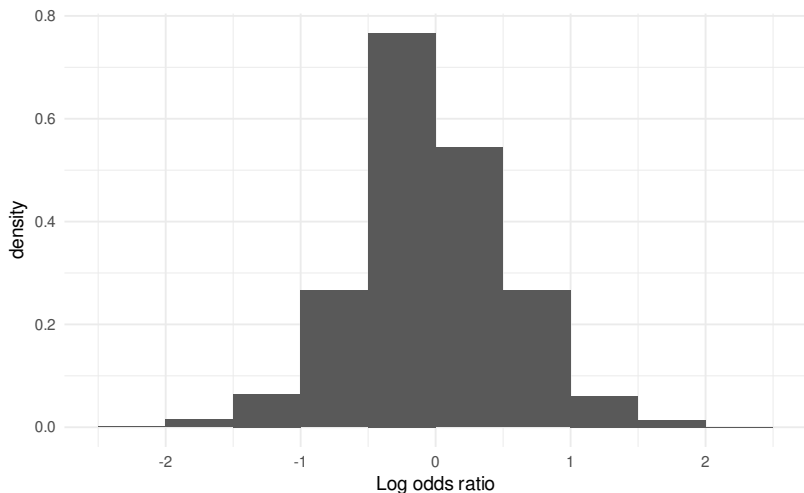
Simulation: The sampling distribution of the odds ratio

$n = 25$ per group
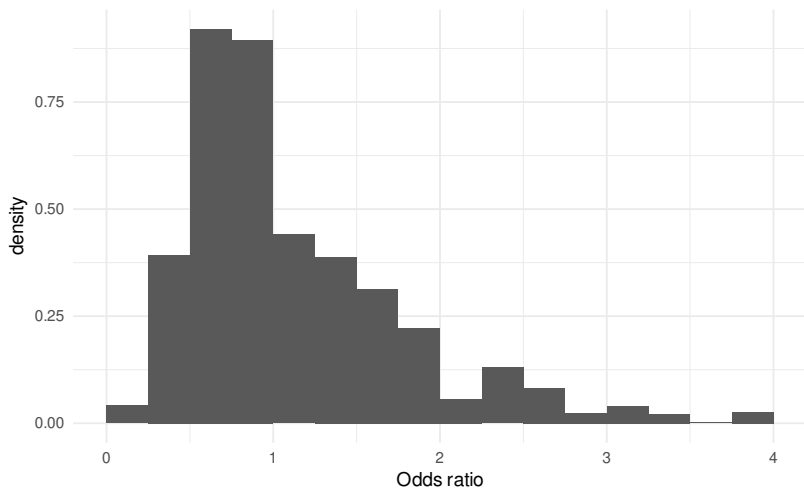
Sampling distribution: After a log transform

However, if we take the logarithm of the odds ratio:

Confidence intervals for the log odds ratio

- The log of the odds ratio is quite normal-looking and amenable to finding confidence intervals for using the central limit theorem/normal distribution
- Thus, the procedure that has been developed for constructing approximate confidence intervals for the odds ratio actually constructs confidence intervals for the log of the odds ratio
- Getting a confidence interval for the odds ratio itself then requires an extra step of converting the confidence interval back to the odds ratio scale

The standard error of the log odds ratio

It can be shown that a good estimate of the standard error of the log of the odds ratio is

$$\mathrm{SE}_{lOR} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

where a, b, c, and d are the four entries in the contingency table

Confidence intervals for the odds ratio: procedure

The first three steps for constructing confidence intervals for the odds ratio should look familiar; the fourth will be new:

(1) Estimate the standard error of the log OR,
$$\mathrm{SE}_{lOR} = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

(2) Determine the values that contain the middle $x\%$ of the normal curve, $\pm z_{x\%}$

(3) Calculate the confidence interval for the log OR:

$$(L, U) = \left(\log(\widehat{\mathrm{OR}}) - z_{x\%} \cdot \mathrm{SE}_{lOR}, \, \log(\widehat{\mathrm{OR}}) + z_{x\%} \cdot \mathrm{SE}_{lOR} \right)$$

(4) Convert the confidence interval from step 3 back into the odds ratio scale to obtain the confidence interval for the odds ratio:

$$(e^L, \, e^U)$$

NOTE: "log" above refers to the natural log (sometimes abbreviated "ln"), not the base-10 logarithm

Confidence intervals for the odds ratio: example

To see how this works, let's calculate a 95% confidence interval for the OR of surviving with sterile surgery in Lister's experiment

(1) Estimate the standard error of the log odds ratio:

$$\mathrm{SE}_{lOR} = \sqrt{\frac{1}{34} + \frac{1}{6} + \frac{1}{19} + \frac{1}{16}}$$
$$= 0.56$$

(2) As usual, 1.96 contains the middle 95% of the normal distribution

(3) Recall that the sample odds ratio was 4.77, so the log of the sample odds ratio is 1.56; the 95% confidence interval for the log odds ratio is therefore

$$(1.56 - 1.96 \times 0.56, 1.56 + 1.96 \times 0.56) = (0.47, 2.66)$$

Confidence intervals for the odds ratio: example (cont'd)

(4) The 95% confidence interval for the odds ratio is therefore

$$\left(e^{0.47}, e^{2.66}\right) = (1.60, 14.2)$$

- Note that the confidence interval doesn't include 1; this agrees with our test of significance
- Note also that this confidence interval is asymmetric (its right half is much longer than its left half) – this would be impossible to achieve without the log transform
- Now we have an idea of the possible clinical significance of sterile technique: it may be lowering the odds of surgical death by a factor of about 1.6, or by a factor of 14, with a factor of around 5 being the most likely

Reverse example

Note also the symmetry that arises if we had decided to calculate a
confidence interval for the other odds ratio (the relative odds of
dying on the sterile surgery):

- $\widehat{\text{OR}} = 0.21$
- $\log(\widehat{\text{OR}}) = -1.56$
- 95% CI for $\log(\widehat{\text{OR}})$: (-2.66, -0.47)
- 95% CI for $\widehat{\text{OR}}$: (0.07, 0.62)

Supplemental: CIs for RRs

- The same procedure is used to construct confidence intervals for relative risks; the only difference is that

$$\mathrm{SE}_{lRR} = \sqrt{\frac{b}{a(a+b)} + \frac{c}{c(c+d)}}$$

- So, in the Lister example, the 95% CI for the relative risk is $[1.3, 6.9]$ (risk of death in control vs. sterile surgery)

- This slide is just an extra, for-your-information slide; for the purposes of quizzes/tests, I will only ask about CIs for odds ratios

Summary

- There are three natural ways to measure the association present in a 2×2 table:
    - Risk difference
    - Relative risk
    - Odds ratio
- One big advantage of the odds ratio is that it works equally well for both prospective and retrospective studies, unlike the other two
- Be able to construct confidence intervals for odds ratios