Measuring Association in Contingency Tables

Ryan Miller

Outline

1. Contingency tables and probability
2. Sensitivity and specificity
3. Risk differences, risk ratios (relative risk), and odds ratios

Introduction

- Lately we've been using the Chi-squared test of association to identify *statistically significant* relationships between two categorical variables
- This presentation will focus on how to report the *clinical significance* of the observed effect following a significant Chi-squared test

Contingency tables and probability

- Statisticians define the **probability** of an outcome as it's long-run relative frequency
  - For example, the probability of a coin flip resulting in "heads" is 0.5 because over a very large number of coin flips you'll see "heads" half of the time

Contingency tables and probability

- Statisticians define the **probability** of an outcome as it's long-run relative frequency
  - For example, the probability of a coin flip resulting in "heads" is 0.5 because over a very large number of coin flips you'll see "heads" half of the time
- Under this definition, it's reasonable to *estimate* an outcome's probability using the corresponding *sample proportion* (ie: $\hat{p}$)

Diagnostic tests

Probability is often applied to contingency tables in applications involving *diagnostic testing*:

	Positive	Negative
Present	True Positive	False Positive
Absent	False Negative	True Negative

- Here, subjects are grouped according to the presence or absence of a disease or exposure (rows)
  - Their outcome on a diagnostic test is recorded as either positive (suggesting they have the disease) or negative (suggesting they do not have the disease)

Diagnostic tests

Diagnostic tests:

	Positive	Negative
Present	True Positive	False Positive
Absent	False Negative	True Negative

- If the diagnostic test is effective, you'd expect the proportion of "true positives" in the "present" group to be high
  - The probability corresponding to this proportion is the test's **sensitivity**

Diagnostic tests

Diagnostic tests:

	Positive	Negative
Present	True Positive	False Positive
Absent	False Negative	True Negative

- If the diagnostic test is effective, you'd expect the proportion of "true positives" in the "present" group to be high
  - The probability corresponding to this proportion is the test's **sensitivity**
- You'd also expect the proportion of "true negatives" in the "absent" group to be high
  - The probability corresponding to this proportion is the test's **specificity**

Practice

Radionuclide ventriculography (RNV) is a non-invasive approach to diagnosing coronary artery disease (CAD). Summarized below are data collected from a population at high risk of coronary artery disease:

	positive	negative
CAD Present	302	179
CAD Absent	80	372

1) Use StatKey to perform a Chi-squared test to determine if there's a *statistically significant* relationship between an individual's RNV result and the presence CAD.
2) Estimate the sensitivity and specificity of RNV, then comment upon it's *clinical significance* as a diagnostic test.

Practice (solution)

1) Using StatKey, $X^2 = 195.907$ and the p-value is nearly zero, so there is overwhelming statistical evidence of an association.
2) The sensitivity is $302/481 = 0.628$ and the specificity is $372/452 = 0.823$; so, while the relationship is highly statistically significant, the clinical significance might only be considered moderate

Epidemiology

Epidemiology is a branch of the biomedical sciences that focuses on relationships between health-related exposures and outcomes:

	Disease	Absent
Exposed		
Unexposed		

Two important study designs used in epidemiology are:

1) *Cohort studies* - a single group is followed forward in time and both variables (exposure status and disease status) are directly observed
2) *Case-control studies* - separate groups of disease-positive (cases) and disease-negative (controls) are asked about their past exposures

Practice (cohort study)

The CDC tracked a cohort of 6,168 women born in the 1960s in search of risk factors for breast cancer. One factor they considered was the age at which each woman gave birth to their first child:

	Developed Breast Cancer	Didn't Develop Cancer
Before Age 25	65	4475
After Age 25	31	1157

Practice (cohort study)

The CDC tracked a cohort of 6,168 women born in the 1960s in search of risk factors for breast cancer. One factor they considered was the age at which each woman gave birth to their first child:

	Developed Breast Cancer	Didn't Develop Cancer
Before Age 25	65	4475
After Age 25	31	1157

The Chi-squared statistic is $X^2 = 7.8$ and the p-value is 0.005. So, there's a *statistically significant* association, but is it *clinically significant*?

1) Use the contingency table to estimate the probability that a woman who gives birth *before age 25* develops breast cancer?
2) Use the contingency table to estimate the probability that a woman who gives birth *after age 25* develops breast cancer?
3) Subtract these two probabilities to find the **risk difference**.

Practice (solution)

1) $65/4540 = 0.014$
2) $31/1188 = 0.026$
3) The risk difference is 0.026-0.014 $= 0.012$, or a 1.2% higher risk of breast cancer for women who gave birth after age 25.

Risk ratios (relative risk)

- For uncommon diseases, the **risk difference** may understate the true importance of a risk factor

Risk ratios (relative risk)

- For uncommon diseases, the **risk difference** may understate the true importance of a risk factor
- For example, over a 10-year span, smokers have a 0.00483 probability (0.483%) of developing lung cancer, while non-smokers have a 0.00045 probability (0.045%) of developing lung cancer
  - That's a *risk difference* of 0.004 (or a 0.4% higher risk for smokers)

Risk ratios (relative risk)

- For uncommon diseases, the **risk difference** may understate the true importance of a risk factor
- For example, over a 10-year span, smokers have a 0.00483 probability (0.483%) of developing lung cancer, while non-smokers have a 0.00045 probability (0.045%) of developing lung cancer
  - That's a *risk difference* of 0.004 (or a 0.4% higher risk for smokers)
- In these scenarios, it is generally preferred to look at the **risk ratio** or **relative risk**
  - In this example, the relative risk is $0.00483/0.00045 = 10.73$
  - So, smokers are *10.73 times as likely* to develop lung cancer as non-smokers

Practice

Summarized below are data from the previously described CDC cohort study:

	Developed Breast Cancer	Didn't Develop Cancer
Before Age 25	65	4475
After Age 25	31	1157

Here, the probability a woman who gave birth *before age 25* develops breast cancer is $65/4540 = 0.014$, while that same probability is estimated at $31/1188 = 0.026$ for women who first give birth *after age 25*.

1) Find and interpret the relative risk of breast cancer across these two groups.
2) Would you prefer to report the relative risk or the risk difference in this scenario?

Practice (solution)

1) The relative risk (of breast cancer) is $0.026/0.014 = 1.86$, indicating that women who first gave birth after age 25 were 1.86 times more likely to develop breast cancer than women who first gave birth before age 25 (ie: 86% higher risk).
2) This is debatable, but if you could only choose one it's the relative risk that tends to be more commonly reported.

Case-control studies

A well-known case-control study published in 1969 examined the relationship between oral contraceptive (OC) use and the risk of blood clots. Data from the study is summarized in the table below:

	Cases (blood clots)	Controls (no clots)
Didn't use OC	42	145
Used OC	42	23

- Notice that 42 of 65 individuals (64.6%) in this study that had used OC also had developed blood clots
  - Based upon your prior knowledge, does this seem like an accurate estimate of the risk of developing blood clots for an OC user?

Case-control studies

- 64.6% is nowhere close to the actual probability of an OC user developing blood clots (the real probability is less than 1%)
- In a case-control study, we cannot use conditional proportions to estimate the risk of an outcome given an exposure
  - So, the *risk difference* and *relative risk* cannot be estimated in a case-control study!

Odds

- The **odds ratio** (OR) can be estimated in a case-control study
  - Instead of a ratio of two probabilities (relative risk), the odds ratio is a ratio of two odds

Odds

- The **odds ratio** (OR) can be estimated in a case-control study
  - Instead of a ratio of two probabilities (relative risk), the odds ratio is a ratio of two odds
- The **odds** of an event are the number of times that event occurs relative to the number of times it doesn't occur
  - Suppose the probability of an event is 50%, the odds here are 1 (.5/.5), which people tend to express as "1 to 1 odds"
  - Suppose the probability of an event is 75%, the odds here are 3 (.75/.25), or "3 to 1 odds"

Odds Ratios

Odds ratios have two major advantages:

1. Can be used in case-control studies
2. Symmetry - the OR for survival of cases relative to controls = the OR for death of controls relative to cases

Practice

Below are the results of the OC case-control study:

	Cases (blood clots)	Controls (no clots)
Didn't use OC	42	145
Used OC	42	23

1) Find the *odds* of blood clots for OC users
2) Find the *odds* of blood clots for those not using OC
3) Find the *odds ratio* describing the risk of blood clots for OC users relative to non-users

Practice (solution)

1) The odds of blood clots were 1.83 (42/23) for OC users
2) The odds of blood clots were 0.29 (45/145) for those not using OC
3) The odds ratio for blood clots given OC use is $1.83/0.29 = 6.31$; so the odds of blood clots were 6.31 times higher for OC users relative to those not using OC

Comments - confidence intervals

- As with any measure of effect size, it's a generally a good idea to report it using a *confidence interval estimate*
- Calculating these intervals for a relative risk or odds ratio is beyond the scope of this class
  - However, you should feel comfortable interpreting these intervals (for example, in the article review project)

Summary

This presentation covered several ways to describe the clinical significance of data displayed in a contingency table:

1) Sensitivity and specificity are used to describe the efficacy of a diagnostic test
2) Risk difference and relative risk are used to describe effect sizes in cohort studies
3) Odds ratios are used to describe effect sizes in any type of study (with a particular advantage in case-control studies)