Basic Statistics & Epidemiology

Dr. Joe Nolan

Northern Kentucky University

Department of Mathematics & Statistics

Overview

- Basic Design Principles

- Bias, Confounding, and Sample Size

- P-values, Confidence Intervals, and Clinical Relevance

- Statistical Practice Specific to Epidemiology

Basic Design Principles

- Basic Design Principles
    - Randomized Controlled Trials
    - Outcome Variables: Level of Measurement Drives Analysis
    - Study Population vs. Sample
- Bias, Confounding, and Sample Size
- P-values, Confidence Intervals, and Clinical Relevance
- Statistical Practice Specific to Epidemiology

RCT: Randomized, Controlled Trials

- Randomization: Participants are assigned to treatments by random selection. Additional constraints may be used:

  - Stratification: Control for a known confounding variable

  - Balance: Ensure each treatment group gets (approximately) same number of subjects

- Controlled: One of the treatments represents a control group, or a "standard" to which we will compare

  - Typically either *placebo* or *standard of care*.

Other Typical Design Features

- Open-label: Everyone knows the treatment being given.

- Blind: Participant does not know the treatment being given.

- Double-blind: Neither participant nor study staff know the treatment being given.

- Repeated-measures: Baseline measurement is taken and then participants are measured over time (likely following one or more treatments). Participants in this case are their own controls.

- Retrospective vs. Prospective

Outcome Variables

- Best choices are variables that are quantitative (interval/ratio)
  - Units, and distance, are meaningful!
  - Statistical methods are robust and confidence intervals are available.
  - Methods include: T-tests, ANOVA, and Regression
- Another good choice is a variable that is binary (represents success/failure)
  - Estimate "chance" of success for given set of predictors
  - Statistical methods robust, confidence intervals available.
  - Methods include: Proportions, Logistic Regression Models

Outcome Variables (Caution!)

- Not everything can be measured "nicely".
    - Ordinal Measurements
    - Likert Scales
    - Nominal Measurements that are not binary
- Statistical methods for these are not as robust, and it is often impossible to obtain usable confidence intervals without converting to something binary.
    - Interval-ratio methods often applied to Likert Scales, but it should be in a very limited fashion.
    - Repeated measures can be very helpful, to combat "interpretation variability".

Global vs. Individual

- Global goals: Assess whether a treatment is generally effective across many people (even though it may be ineffective for some)

  - Statistical models → Comparison of means, medians, or proportions

- Individual goals: Estimate or predict something for an individual patient

  - Chance something will occur (e.g. readmission)

  - Some particular (future) measurement of interest (e.g. lifetime)

  - Statistical models → Risk intervals, Prediction intervals

Population vs. Sample

- Population: Generally, the group about whom we want to draw inference

  - Defined by Inclusion Criteria

  - Goal of statistics is to generalize from the sample to this group.

- Sample: the group for whom we actually collect data

  - Statistical ideal is "random" but almost never feasible.

  - Goal is to collect sample in such a way that bias is unlikely.

  - Participant must meet both inclusion and exclusion criteria to be in sample.

- Literature Caution: If all that's being discussed is descriptive stats, that isn't sufficient for population inference!

Bias, Confounding, & Sample Size

- Basic Design Principles

- Bias, Confounding, and Sample Size

- P-values, Confidence Intervals, and Clinical Relevance

- Statistical Practice Specific to Epidemiology

Bias

- Property of the SAMPLE; often a result of poor design choices

  - May stem from inappropriate investigator choices in selecting the sample

  - May result from missing data / non-response

  - May simply come from misfortune in sampling

- Practical Definition: Sample misrepresents population <u>in some way relating</u> to study variables.

- Implications of Bias: Generalizability is compromised.

Confounding

- Property of GROUPINGS, though very similar to bias

- Practical Definition: Treatment groups are imbalanced with respect to an extraneous variable which <u>in some way relates</u> to outcome variable(s) being compared.

- Implications: Comparison for treatments won't be "fair". If differences between groups are found, won't know if treatment-related or a result of confounder.

- Mitigation: Only confounders that are known and measurable can be controlled. Strategies to reduce confounding include randomization, matching, stratification.

Sample Size: Common Misconception

- The misconception: larger sample size solves everything.

- The facts: sample size has nothing to do with bias or confounding. In fact, a biased sample that is large is in some sense more dangerous – due to the misconception that sample size somehow makes it better.

Overview

- Basic Design Principles

- Bias, Confounding, and Sample Size

- P-values, Confidence Intervals, and Clinical Relevance

- Statistical Practice Specific to Epidemiology

P-value

- Measures the likelihood that a study result would be achieved based on random chance (i.e. assuming that there are NO differences in treatments)

- Small p-value suggests the result is not random chance (e.g. supports the inference that there are some differences in treatment averages)

- Small is relative, but often taken to be < 0.05

  - If NO difference in treatments, then 1 in 20 studies would produce a p-value less than 0.05 and falsely identify some difference (Type 1 error or False Positive)

  - If doing MANY statistical tests, use something more stringent than 0.05!

15

Clinical Relevance

- Small p-values are generally necessary, but NOT usually sufficient for clinical relevance.

- When differences are found, the natural question is which treatments are best, and by how much? Effect size matters!

  - Confidence Intervals are needed to assess "how much?"

  - P-values for hypothesis tests that are structured to identify only effects of a specific size are possible, but not at all typical.

Confidence Intervals

- Usually "95%". The percentage attached to a CI is effectively the chance that the CI for a given sample will "capture" the truth.

  - Not feasible to get to 100%.

- Think of every value "within" the interval as a plausible value of the parameter being estimated.

- CI's offer far more information than standard hypothesis tests. **Use them to think about how small (or large) the effect might be and assess clinical relevance.**

Sample Size == Precision

- Greater sample size yields greater precision (or statistical power)

- Allows for the detection of smaller differences between groups

  - Good, if those differences will be clinically meaningful

- More importantly, reduces the width of confidence intervals

  - Direct role in assessment of clinical relevance

Stats in Epidemiology

- Basic Design Principles

- Bias, Confounding, and Sample Size

- P-values, Confidence Intervals, and Clinical Relevance

- Statistical Practice Specific to Epidemiology

  - Risk, Absolute Risk, Relative Risk

  - NNT and NNH

  - Odds, Odds Ratio

  - Understanding Diagnostic Tests

Risk

- **Absolute Risk**: Chance of an event occurring within a specific time period.
  - Express as fraction (8/100), decimal (0.08), or percentage (8%).
- **Absolute Risk Difference**: The *difference* in size of risk for two groups.
- **Relative Risk**: *Ratio* of risks for two groups (with the purpose of comparison, typically treatment divided by control)

Example

- Suppose that 1 in 1000 smokers develop disease X. Suppose that 3 in 500000 nonsmokers develop disease X.

  - Absolute risk for smokers is 0.1%, absolute risk for nonsmokers is 0.0006%.

  - Risk for smokers relative to that of nonsmokers is 0.1 / 0.0006 = 167. Smokers are 167 times as likely to develop disease X.

Simplify for Patients!

- Risk (and absolute risk difference) is easier to understand.

- Relative Risk is harder to interpret and prone to exaggeration

  - "167 times as likely" needs the context of the absolute risk.

  - In this example, 167 times "very low risk" is still "very low risk". Not necessarily a reason to treat.

- Can look at this also in terms of Number Needed to Treat (NNT) or Number Needed to Harm (NNH).

  - Absolute Risk Reduction (ARR) = |Treatment Group Risk – Control Group Risk|

  - Number Needed to Treat (NNT) = 1 / ARR   (if Control > Treatment)

  - Number Needed to Harm (NNH) = 1 / ARR   (if Treatment > Control)

Smokers vs. Nonsmokers Example

- ARR = |0.1 – 0.0006| = 0.0994

- NNH = 1 / 0.0994 = 10.06

- For every 1006 additional smokers, approximately 100 more with disease X.

- Or, flip to NNT: for every 1006 we convince to stop smoking, approximately 100 fewer will develop disease X. (admittedly this is simplistic)

Odds & Odds Ratio

- Odds will be even more confusing for patients than risk!
  - Risk is $\frac{a}{a+b}$ while odds is $\frac{a}{b}$
  - Example: Odds of 2/3 is Risk of 2/5 = 40%. Think of Odds as a ratio of "plus" to "minus"; it doesn't correspond directly to the idea of a fraction.
  - Odds ratio is the ratio of odds in the experimental group to odds in the control group. Greater than ONE means greater likelihood in the first group (or top group in forming the ratio).
  - Smoking Example:
    - Smoker odds: 1/999   Nonsmoker odds: 3/499997
    - Odds Ratio: (1/999) / (3/499997) = 166.8
    - Interpretation: Odds of Disease X for Smokers are 166.8 times the odds for nonsmokers (a bit misleading)

Where do we see these things?

- Contingency Tables

  - Estimate Risk or Odds for specific groups.

- Logistic Regression

  - Odds Ratio is the basic inferential output

  - Better practice (?): Can use model to estimate risk for specified inputs (individualized care)

Diagnostic Testing

- Diagnostic Testing 2x2 Tables
- Likelihood Ratios and Posttest Probabilities

Diagnostic Testing Terminology

- Prevalence: Proportion or percentage of patients who have disease

- Sensitivity: Proportion *who have disease* who test positive

- Specificity: Proportion *who do not have disease* who test negative

  - Ideal: Sensitivity = Specificity = 1 = 100%   (highly unlikely)

- Positive Predictive Value: Proportion *who test positive* that do indeed have the disease

- Negative Predictive Value: Proportion *who test negative* that indeed do not have the disease

Example

	Diseased	Not Diseased	
Positive Test	75	66	141
Negative Test	3	1297	1300
	78	1363	

Prevalence = 78 / 1441 = 5.4%
Sensitivity = 75 / 78 = 96%
Specificity = 1297 / 1363 = 95%
PPV = 75 / 141 = 53%
NPV = 1297 / 1300 = 99.7%

Likelihood Ratios

- Positive Likelihood Ratio (LR+)

  - Ratio of % correct positives to % incorrect positives

  - LR+ = Sensitivity / (1 – Specificity)

  - Ideally quite large
    $$LR+ = 0.96 / (1 - 0.95) = 19.2$$

- Negative Likelihood Ratio (LR-)

  - Ratio of % incorrect negatives to % correct negatives

  - LR- = (1 – Sensitivity) / Specificity

  - Ideally quite small
    $$LR+ = (1-0.96) / 0.95 = 0.042$$

Pretest / Posttest Statistics

- Pretest probability: chance of having disease before a test is performed (prevalence if no other info, or could be based on a model incorporating other known risk factors)

- Posttest probability: chance of having disease after test is performed

  - Calculate based on odds: posttest odds / (1+posttest odds)

- Pretest odds: pretest probability / (1 – pretest probability)

- Posttest odds: Pretest Odds * LR

  - LR = LR+ if positive test or LR- if negative test

Example

- Suppose a patient presents with pretest probability of 1/3 (perhaps model based).

- Pretest odds = (1/3) / (1 – 1/3) = 0.5

- Patient tests positive (LR+ was 19.2)

- Posttest odds are 0.5*19.2 = 9.6

- Posttest probability = 9.6 / (1+9.6) = 90.6%

  - The post-test probability is a reasonable number for patient understanding (the rest is not).

Handout

- https://imath.nku.edu/
  - Username: mercy_health   PW: nkustats
  - Will contain these slides
  - Also contains PDF copy of my Biostats text
- Stats Direct
- Glossary of Terms

Questions