Inference for Binomial Proportions

Chapter 8, Lab 1

OpenIntro Biostatistics

Topics

- Inference for a single proportion
- Inference for the difference of two proportions

This lab introduces methods of inference for binomial proportions, as a parallel to the one-sample and two-sample methods for means introduced in Chapters 4 and 5. In the setting where a binary outcome is recorded for a single group of participants, inference about the binomial probability of success provides information about a population proportion p. Just as inference can be done for the difference of two population means, inference can also be done in the setting of comparing two population proportions p_1 and p_2.

The material in this lab corresponds to Sections 8.1 and 8.2 in *OpenIntro Biostatistics*.

Introduction

Inference for a single proportion

Suppose that X is binomial with parameters n (total number of trials) and p, the population parameter of success, where x represents the number of successes. Inference about p is based on the sample proportion $\hat{p}$, where $\hat{p} = x/n$; $\hat{p}$ is the point estimate of p.

Inference for p can be made using the normal approximation to the binomial, or directly using the binomial distribution.

- *Inference with the normal approximation*
  - The sampling distribution of $\hat{p}$ is approximately normal when 1) the sample observations are independent, 2) $np \geq 10$, $n(1-p) \geq 10$.[1] Under these conditions, the sampling distribution of $\hat{p}$ is approximately normally distributed with mean p and standard deviation $\sqrt{\frac{p(1-p)}{n}}$. For confidence intervals, substitute $\hat{p}$ for p; for hypothesis testing, substitute p_0 for p.
  - The approximate two-sided 95% confidence interval for p is given by

$$\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

  - The test statistic z for the null hypothesis $H_0 : p = p_0$ based on a sample size of n is

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{(p_0)(1-p_0)}{n}}}$$

[1] The second condition is commonly referred to as the **success-failure condition**, since it can be effectively restated as the number of successes is greater than 10 and the number of failures is greater than 10.

- *Inference with exact methods*
  - Confidence intervals and p-values based on the binomial distribution are best calculated via R.
  - The logic behind calculating a p-value from the binomial distribution: Let X be a binomial random variable with parameters n and p_0, where $\hat{p} = x/n$ and x is the observed number of events. For a test of $H_0 : p = p_0$ versus $H_A : p \neq p_0$, the p-value equals $2 \times P(X \geq x)$.

Inference for the difference of two proportions

The normal model can be applied to $\hat{p}_1 - \hat{p}_2$ if the sampling distribution for each sample proportion is nearly normal, and if the samples are independent random samples from the relevant populations and independent of each other.

Each sample proportion approximately follows a normal model when $n_1 p_1$, $n_1(1-p_1)$, $n_2 p_2$, and $n_2(1-p_2)$ are all ≥ 10. To check success-failure in the context of a confidence interval, use $\hat{p}_1 and \hat{p}_2$.

The standard error of the difference in sample proportions is

$$\sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}.$$

For hypothesis testing, an estimate of p is used to compute the standard error of $\hat{p}_1 - \hat{p}_2$: $\hat{p}$, the weighted average of the sample proportions $\hat{p}_1$ and $\hat{p}_2$,

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2} = \frac{x_1 + x_2}{n_1 + n_2}.$$

To check success-failure in the context of hypothesis testing, check that $\hat{p}n_1$ and $\hat{p}n_2$ are both ≥ 10.

Inference for a single proportion

1. Advanced melanoma is an aggressive form of skin cancer that until recently was almost uniformly fatal. In rare instances, a patient's melanoma stopped progressing or disappeared altogether when the patient's immune system successfully mounted a response to the cancer. Those observations led to research into therapies that might trigger an immune response.

   A 2013 report in the New England Journal of Medicine by Wolchok, et al. reported the results of a study in which patients were treated concurrently with two new therapies, nivolumab and ipilimumab. Of 52 patients, 21 experienced an immune response.

   a) What is the estimate of the population probability of an immune response after concurrent treatment with nivolumab and ipilimumab, $\hat{p}$?

   b) Evaluate whether the assumptions for using the normal approximation are met.

   c) Calculate an approximate 95% confidence interval for p and interpret the interval in context of the data.

   d) Confirm the answer to part c) using `prop.test( )`.

2. Suppose that out of a cohort of 120 patients with stage I lung cancer at the Dana-Farber Cancer Institute (DFCI), 80 of the patients survive at least 5 years. National Cancer Institute statistics indicate that the 5-year-survival probability for stage I lung cancer patients nationally is 0.60. Do the data collected from the 120 patients support the claim that the DFCI population with stage I lung cancer has a different 5-year-survival probability than the national population? Let $\alpha = 0.05$.

   a) State the null and alternative hypotheses.

   b) Evaluate whether the assumptions for using the normal approximation are met. If so, conduct the hypothesis test based on the normal approximation.

   c) Compare the results in part b) to those from conducting the hypothesis test based on exact binomial methods.

3. In 2009, the FDA Oncology Drug Advisory Committee (ODAC) recommended that the drug Avastin be approved for use in glioblastoma, a form of brain cancer. Tumor shrinkage after taking a drug is called a response; out of 85 patients, 24 exhibited a response. Historically, response probabilities for brain cancer drugs were approximately 0.05, or about 5%. Assess whether there is evidence that the response probability for Avastin is different from previous drugs. Report both a p-value and 95% confidence interval; summarize your findings.

Inference for the difference of two proportions

4. The use of screening mammograms for breast cancer has been controversial for decades because the overall benefit on breast cancer mortality is uncertain. A 30-year study to investigate the effectiveness of mammograms versus a standard non-mammogram breast cancer exam was conducted in Canada with 89,835 female participants.[2] Each woman was randomized to receive either annual mammograms or standard physical exams for breast cancer over a 5-year screening period.

   By the end of the 25 year follow-up period, 1,005 women died from breast cancer. The results are summarized in the following table.[3]

	Death from breast cancer?	
	Yes	No
Mammogram Group	500	44,425
Control Group	505	44,405

   a) Calculate $\hat{p}_1$ and $\hat{p}_2$, the two sample proportions of interest.

   b) Analyze the results; do the data suggest that annual mammography results in a reduction in breast cancer mortality relative to standard exams? Be sure to check the assumptions for using the normal approximation.

   c) Calculate and interpret a 95% confidence interval for the difference in proportions of deaths from breast cancer. Be sure to check the assumptions for using the normal approximation.

[2] Miller AB. 2014. *Twenty five year follow-up for breast cancer incidence and mortality of the Canadian National Breast Screening Study: randomised screening trial. BMJ* 348 (2014): g366.

[3] During the 25 years following the screening period, each woman was screened for breast cancer according to the standard of care at her health care center.

5. Remdesivir is an antiviral drug previously tested in animal models infected with coronaviruses like SARS and MERS. As of May 2020, remdesivir had temporary approval from the FDA for use in severely ill COVID-19 patients and was the subject of numerous ongoing studies.

   A randomized controlled trial conducted in China enrolled 236 patients with severe COVID-19; 158 were assigned to receive remdesivir and 78 to receive a placebo. In the remdesivir group, 103 patients showed clinical improvement; in the placebo group, 45 patients showed clinical improvement.[4]

   a) Calculate $\hat{p}_1$ and $\hat{p}_2$, the two sample proportions of interest.

   b) Conduct a formal comparison of the clinical improvement rates and summarize your findings. Be sure to check the assumptions for using the normal approximation.

   c) Report and interpret an appropriate interval estimate. Be sure to check the assumptions for using the normal approximation.

[4] Wang, Y, et al. Remdesivir in adults with severe COVID-19: a randomised, double-blind, placebo-controlled, multi-centre trial. *Lancet* 395(10236). 16 May 2020.