Math 362: Mathematical Statistics II

Le Chen
le.chen@emory.edu

Emory University
Atlanta, GA

Last updated on April 13, 2021

2021 Spring

Chapter 9. Two-Sample Inferences

§ 9.1 Introduction

§ 9.2 Testing $H_0 : \mu_X = \mu_Y$

§ 9.3 Testing $H_0 : \sigma_X^2 = \sigma_Y^2$

§ 9.4 Binomial Data: Testing $H_0 : p_X = p_Y$

§ 9.5 Confidence Intervals for the Two-Sample Problem

Plan

§ 9.1 Introduction

§ 9.2 Testing $H_0 : \mu_X = \mu_Y$

§ 9.3 Testing $H_0 : \sigma_X^2 = \sigma_Y^2$

§ 9.4 Binomial Data: Testing $H_0 : p_X = p_Y$

§ 9.5 Confidence Intervals for the Two-Sample Problem

By the central limit theorem, when n and m are large

$$\frac{\frac{X}{n} - \frac{Y}{m} - \mathbb{E}\left(\frac{X}{n} - \frac{Y}{m}\right)}{\sqrt{\mathrm{Var}\left(\frac{X}{n} - \frac{Y}{m}\right)}} \overset{approx.}{\sim} N(0,1)$$

Under $H_0 : p_X = p_Y$,

$$\mathbb{E}\left(\frac{X}{n} - \frac{Y}{m}\right) = 0$$

$$\mathrm{Var}\left(\frac{X}{n} - \frac{Y}{m}\right) = \frac{p(1-p)}{n} + \frac{p(1-p)}{m}$$

The MLE for p under H_0 is

$$p_e = \frac{x+y}{n+m}$$

By the central limit theorem, when n and m are large

$$\frac{\frac{X}{n} - \frac{Y}{m} - \mathbb{E}\left(\frac{X}{n} - \frac{Y}{m}\right)}{\sqrt{\operatorname{Var}\left(\frac{X}{n} - \frac{Y}{m}\right)}} \overset{approx.}{\sim} N(0,1)$$

Under $H_0 : p_X = p_Y$,

$$\mathbb{E}\left(\frac{X}{n} - \frac{Y}{m}\right) = 0$$

$$\operatorname{Var}\left(\frac{X}{n} - \frac{Y}{m}\right) = \frac{p(1-p)}{n} + \frac{p(1-p)}{m}$$

The MLE for p under H_0 is

$$p_e = \frac{x+y}{n+m}$$

By the central limit theorem, when n and m are large

$$\frac{\frac{X}{n} - \frac{Y}{m} - \mathbb{E}\left(\frac{X}{n} - \frac{Y}{m}\right)}{\sqrt{\mathrm{Var}\left(\frac{X}{n} - \frac{Y}{m}\right)}} \overset{approx.}{\sim} N(0,1)$$

Under $H_0 : p_X = p_Y$,

$$\mathbb{E}\left(\frac{X}{n} - \frac{Y}{m}\right) = 0$$

$$\mathrm{Var}\left(\frac{X}{n} - \frac{Y}{m}\right) = \frac{p(1-p)}{n} + \frac{p(1-p)}{m}$$

The MLE for p under H_0 is

$$p_e = \frac{x+y}{n+m}$$

By the central limit theorem, when n and m are large

$$\frac{\frac{X}{n} - \frac{Y}{m} - \mathbb{E}\left(\frac{X}{n} - \frac{Y}{m}\right)}{\sqrt{\mathrm{Var}\left(\frac{X}{n} - \frac{Y}{m}\right)}} \overset{approx.}{\sim} N(0,1)$$

Under $H_0 : p_X = p_Y$,

$$\mathbb{E}\left(\frac{X}{n} - \frac{Y}{m}\right) = 0$$

$$\mathrm{Var}\left(\frac{X}{n} - \frac{Y}{m}\right) = \frac{p(1-p)}{n} + \frac{p(1-p)}{m}$$

The MLE for p under H_0 is

$$p_e = \frac{x+y}{n+m}$$

Testing $H_0 : p_X = p_Y$

v.s.

(at the α level of significance)

$$z = \frac{\frac{x}{n} - \frac{y}{m}}{\sqrt{p_e(1-p_e)\left(\frac{1}{n}+\frac{1}{m}\right)}}, \qquad p_e = \frac{x+y}{n+m}$$

$H_1 : p_X < p_Y$:	$H_1 : p_X \neq p_Y$:	$H_1 : p_X > p_Y$:
Reject H_0 if	Reject H_0 if	Reject H_0 if
$z \leq -z_\alpha$	$\lvert z\rvert \geq z_{\alpha/2}$	$z \geq z_\alpha$

E.g. Nightmares among men and women:

Is 34.4% significantly different from 31.1% ($\alpha = 0.05$)?

Sol. ... □

E.g. Nightmares among men and women:

Table 9.4.1 Frequency of Nightmares

	Men	Women	Total
Nightmares often	55	60	115
Nightmares seldom	105	132	237
Totals	160	192	
% often:	34.4	31.3	

Is 34.4% significantly different from 31.1% ($\alpha = 0.05$)?

Sol. ... $\square$