Math 362: Mathmatical Statistics II

Le Chen
le.chen@emory.edu

Emory University
Atlanta, GA

Last updated on March 3, 2020

2020 Spring

Chapter 9. Two-Sample Inferences

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

§ 9.5 Confidence Intervals for the Two-Sample Problem

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

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:

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. ... $\square$

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

§ 9.5 Confidence Intervals for the Two-Sample Problem

§ 9.5 Confidence Intervals for the Two-Sample Problem

Similar to the hypothesis test ...

1. Let $X_1, \cdots, X_n$ be a random sample of size n from $N(\mu_X, \sigma_X^2)$.
2. Let $Y_1, \cdots, Y_m$ be a random sample of size m from $N(\mu_Y, \sigma_Y^2)$.

Prob. 1 Find the $100(1-\alpha)\%$ C.I. for $\mu_X - \mu_Y$

When both σ_X^2 and σ_Y^2 are known

When $\sigma_X^2 = \sigma_Y^2 = \sigma^2$, but is unknown

When $\sigma_X^2 \neq \sigma_Y^2$, both are unknown

Prob. 2 Find the $100(1-\alpha)\%$ C.I. for σ_X^2/σ_Y^2, or σ_X/σ_Y

Prob. 1-1 Find the $100(1-\alpha)\%$ C.I. for $\mu_X - \mu_Y$ with σ_X^2 and σ_Y^2 known.

Sol.

$$\frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \sim N(0,1)$$

$$\mathbb{P}\left(-z_{\alpha/2} \le \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}} \le z_{\alpha/2}\right) = 1 - \alpha$$

$$||$$

$$\mathbb{P}\left((\overline{X} - \overline{Y}) - z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \le \mu_X - \mu_Y \le (\overline{X} - \overline{Y}) + z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right)$$

$$\left((\overline{x} - \overline{y}) - z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}} \quad , \quad (\overline{x} - \overline{y}) + z_{\alpha/2}\sqrt{\frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}}\right)$$

$\square$

Prob. 1-2 Find the $100(1-\alpha)\%$ C.I. for $\mu_X - \mu_Y$ when $\sigma_X^2 = \sigma_Y^2 = \sigma^2$ unknown

Sol.

$$\frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \sim \text{Student t-distribution } (n+m-2)$$

$$\mathbb{P}\left(-t_{\alpha/2,n+m-2} \le \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{S_p\sqrt{\frac{1}{n} + \frac{1}{m}}} \le t_{\alpha/2,n+m-2}\right) = 1 - \alpha$$

$$||$$

$$\mathbb{P}\left((\overline{X} - \overline{Y}) - t_{\alpha/2,n+m-2} S_p\sqrt{\frac{1}{n} + \frac{1}{m}} \le \mu_X - \mu_Y \le (\overline{X} - \overline{Y}) + t_{\alpha/2,n+m-2} S_p\sqrt{\frac{1}{n} + \frac{1}{m}}\right)$$

$$\left((\overline{x} - \overline{y}) - t_{\alpha/2,n+m-2} s_p\sqrt{\frac{1}{n} + \frac{1}{m}} \quad , \quad (\overline{x} - \overline{y}) - t_{\alpha/2,n+m-2} s_p\sqrt{\frac{1}{n} + \frac{1}{m}}\right)$$

□

Prob. 1-3 Find the $100(1-\alpha)\%$ C.I. for $\mu_X - \mu_Y$ when $\sigma_X^2 \neq \sigma_Y^2$ unknown.

Sol.

$$\frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}} \sim \text{Student t-distribution } (\nu)$$

$$\mathbb{P}\left(-t_{\alpha/2,\nu} \leq \frac{\overline{X} - \overline{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}} \leq t_{\alpha/2,\nu}\right) \approx 1 - \alpha$$

$$||$$

$$\mathbb{P}\left((\overline{X} - \overline{Y}) - t_{\alpha/2,\nu}\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}} \leq \mu_X - \mu_Y \leq (\overline{X} - \overline{Y}) + t_{\alpha/2,\nu}\sqrt{\frac{S_X^2}{n} + \frac{S_Y^2}{m}}\right)$$

$$\left((\overline{x} - \overline{y}) - t_{\alpha/2,\nu}\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}} \quad , \quad (\overline{x} - \overline{y}) + t_{\alpha/2,\nu}\sqrt{\frac{s_X^2}{n} + \frac{s_Y^2}{m}}\right)$$

□

Prob. 2 Find the $100(1-\alpha)\%$ C.I. for σ_X^2/σ_Y^2

Sol 1.

$$\frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \sim \text{F-disribution } (n-1, m-1)$$

$$\mathbb{P}\left(F_{\alpha/2,n-1,m-1} \le \frac{S_X^2/\sigma_X^2}{S_Y^2/\sigma_Y^2} \le F_{1-\alpha/2,n-1,m-1}\right) = 1-\alpha$$

$$||$$

$$\mathbb{P}\left(\frac{S_X^2}{S_Y^2}\frac{1}{F_{1-\alpha/2,n-1,m-1}} \le \frac{\sigma_X^2}{\sigma_Y^2} \le \frac{S_X^2}{S_Y^2}\frac{1}{F_{\alpha/2,n-1,m-1}}\right)$$

$$\left(\frac{s_X^2}{s_Y^2}\frac{1}{F_{1-\alpha/2,n-1,m-1}} \quad , \quad \frac{s_X^2}{s_Y^2}\frac{1}{F_{\alpha/2,n-1,m-1}}\right)$$

□

Sol 2. Or equivalently,

$$\frac{S_Y^2/\sigma_Y^2}{S_X^2/\sigma_X^2} \sim \text{F-disribution } (m-1, n-1)$$

$$\mathbb{P}\left(F_{\alpha/2,m-1,n-1} \leq \frac{S_Y^2/\sigma_Y^2}{S_X^2/\sigma_X^2} \leq F_{1-\alpha/2,m-1,n-1}\right) = 1-\alpha$$

$$||$$

$$\mathbb{P}\left(\frac{S_X^2}{S_Y^2}F_{\alpha/2,m-1,n-1} \leq \frac{\sigma_X^2}{\sigma_Y^2} \leq \frac{S_X^2}{S_Y^2}F_{1-\alpha/2,m-1,n-1}\right)$$

$$\left(\frac{s_X^2}{s_Y^2}F_{\alpha/2,m-1,n-1} \quad , \quad \frac{s_X^2}{s_Y^2}F_{1-\alpha/2,m-1,n-1}\right)$$

□

Recall: $$F_{\alpha,m,n} = \frac{1}{F_{1-\alpha,n,m}}$$

Examples from the book...