August 1, 2019

Two-Sample Test for Difference of Proportions

Setup:

$H_0: \pi_1 = \pi_2$

$H_1: \left.\begin{array}{l}\pi_1 > \pi_2 \\ \pi_1 < \pi_2\end{array}\right\}$ one-tailed

$\pi_1 \neq \pi_2$ (two-tailed)

DeMoivre-Laplace Theorem extends to a two-sample context:

If $n_1\hat{\pi}_1 \geq 10$, $n_1(1-\hat{\pi}_1) \geq 10$, $n_2\hat{\pi}_2 \geq 10$, $n_2(1-\hat{\pi}_2) \geq 10$,

then given $H_0: \pi_1 = \pi_2$,

$$Z = \frac{(\hat{\pi}_1 - \hat{\pi}_2) - (\pi_1 - \pi_2)}{\sqrt{\dfrac{\hat{\pi}(1-\hat{\pi})}{n_1} + \dfrac{\hat{\pi}(1-\hat{\pi})}{n_2}}} \sim N(0,1)$$

approximately

pooled sample proportion

$$\boxed{\text{Define } \hat{\pi} = \frac{\text{total \# of successes over groups 1 and 2}}{\text{total \# of trials}}}$$

$$\boxed{\cancel{\text{You might be tempted to use } \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}}}$$

however, you do use this for CI

not appropriate for hypothesis test

<u>EX</u>: MSDS students all run a marathon (literally) at the end of bootcamp. They're not really designed to run marathons, but we want to see if there is any difference between men and women with respect to completion rates.

<u>MALE</u>

$n_M = 52$

# of successes = 10

<u>FEMALE</u>

$n_F = 50$

# of successes = 15

$H_0: \pi_M = \pi_F$

$H_1: \pi_M \neq \pi_F$

Check: In both groups, sample successes and sample failures are ≥ 10.

NOTE: $\hat{\pi} = \dfrac{10+15}{52+50} \approx 0.245$

$$Z = \frac{\left(\frac{15}{50} - \frac{10}{52}\right) - \overset{\to 0}{\cancel{(\pi_F - \pi_M)}}}{\sqrt{\frac{0.245(1-0.245)}{52} + \frac{0.245(1-0.245)}{50}}} \approx \frac{0.10769}{\sqrt{0.00355 + 0.00369}}$$

$$\approx \frac{0.10768}{0.085} \approx 1.266 \sim N(0,1)$$

if $H_0: \pi_M = \pi_F$ is true

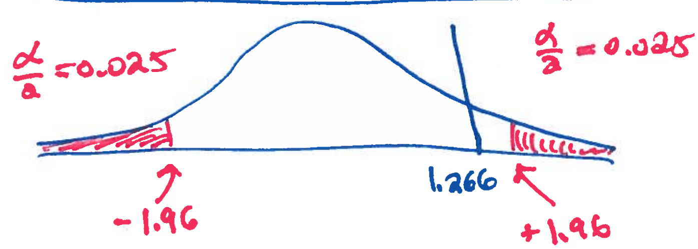
<u>P-VALUE APPROACH</u>

1-pnorm(1.266,0,1)

≈ 0.102

Because $2(0.102) = 0.204 > 0.05 = \alpha$, we conclude that there is not enough evidence at this time to reject $H_0: \pi_M = \pi_F$, i.e., there's not enough evidence right now to suggest different marathon completion rates.

<u>CRITICAL VALUE APPROACH</u>



Because $-1.96 < 1.266 = z < 1.96$, we do not have enough evidence at this time to conclude.

NOTE: The $100(1-\alpha)\%$ CI for the difference of two proportions is therefore given by

$$(\hat{\pi}_1 - \hat{\pi}_2) \pm z^*_{\frac{\alpha}{2}} \sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}$$

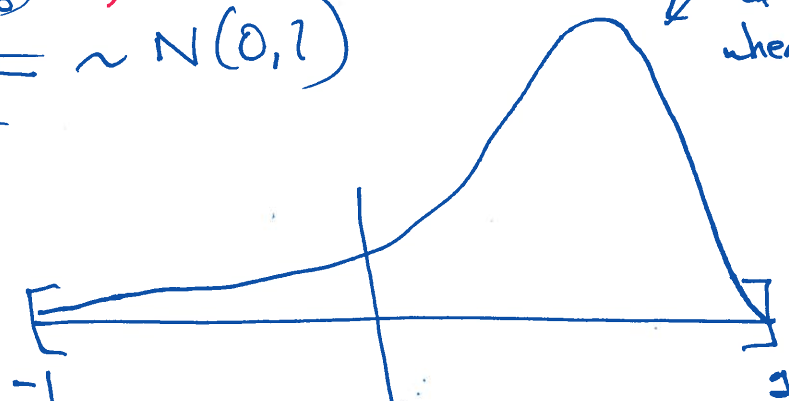
HYPOTHESIS TEST FOR DIFFERENCE IN RHOS:

From the bonus notes, you learned to if $H_0: \rho = \rho_0$ is true, then

$$Z = \frac{F(\hat{\rho}) - F(\rho_0)}{\sqrt{\frac{1}{n-3}}} \sim N(0,1)$$

approximately

EX: sampling dist. of $\hat{\rho}$ when $\rho_0 = 0.85$



Fisher determined that $F(\hat{\rho})$ was closer to normal that just $\hat{\rho}$, but if $|\hat{\rho}| < 0.5$, $F(\hat{\rho}) \approx \hat{\rho}$ and so the transformation is less necessary.

$H_0: \rho_{X,Y} = \rho_{A,B}$

$H_1: \rho_{X,Y} \neq \rho_{A,B}$

Then the appropriate test statistic is

$$Z = \frac{\left(F(\hat{\rho}_{X,Y}) - F(\hat{\rho}_{A,B})\right) - \cancel{\left(F(\rho_{X,Y}) - F(\rho_{A,B})\right)}^{\,0}}{\sqrt{\frac{1}{n_{X,Y}-3} + \frac{1}{n_{A,B}-3}}} \sim N(0,1)$$

when $H_0: \rho_{X,Y} = \rho_{A,B}$ is true.

The $100(1-\alpha)\%$ CI for the difference of the two rhos

$$a = \left(F(\hat{\rho}_{X,Y}) - F(\hat{\rho}_{A,B})\right) - z^{*}_{\frac{\alpha}{2}} \sqrt{\frac{1}{n_{X,Y}-3} + \frac{1}{n_{A,B}-3}}$$

$$b = \left(F(\hat{\rho}_{X,Y}) - F(\hat{\rho}_{A,B})\right) + z^{*}_{\frac{\alpha}{2}} \sqrt{\frac{1}{n_{X,Y}-3} + \frac{1}{n_{A,B}-3}}$$

$\Rightarrow$ Actually, we need: $\left[F^{-1}(a), F^{-1}(b)\right]$ – is the $100(1-\alpha)\%$ CI for the true difference between $\rho_{X,Y}$ and $\rho_{A,B}$.

Two-Sample Tests for Means

Case #1: $\sigma_1^2 = \sigma_2^2$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\underbrace{\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}}_{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}} \sim t(n_1 + n_2 - 2)$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

pooled variance

Case #2: $\sigma_1^2 \neq \sigma_2^2$

$H_0: \mu_1 = \mu_2$

$H_1: \mu_1 \neq \mu_2$

$$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \sim t(\text{Satt})$$

Satterthwaite

$$\text{Satt} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{1}{n_1 - 1}\left(\frac{s_1^2}{n_1}\right)^2 + \frac{1}{n_2 - 1}\left(\frac{s_2^2}{n_2}\right)^2}$$

Qn: In practice, how do I decide whether or not σ_1^2 should be treated as if it were equal to σ_2^2?

Approach #1: Nothing is equal, ever. Always use case #2.

Approach #2: Use a formal statistical test to check the null hypothesis $H_0: \sigma_1^2 = \sigma_2^2$.

Approach #3: Use the "quick and dirty" heuristic... basically, check if $\frac{1}{2} \leq \frac{s_1^2}{s_2^2} \leq 2$ ... if so, use case #1. Otherwise, use case #2.

RECALL: If $X_1, \ldots, X_n \sim N(0,1)$ and independent, then $\chi^2 \sim X_1^2 + \ldots + X_n^2$ is a chi-squared r.v. with n degrees of freedom.

RESULT: Suppose $X \sim \chi^2(n)$ and $Y \sim \chi^2(m)$ and X and Y are independent.

n numerator degrees of freedom

$$F = \frac{X/n}{Y/m} \sim F(n, m)$$

m: denominator degrees of freedom

"F ratio distribution"

DERIVATION: Suppose that $X_1, \ldots, X_n$ is i.i.d. and normal and $Y_1, \ldots, Y_m$ is i.i.d. and normal.

We know that

$$\frac{(n-1)S_X^2}{\sigma_X^2} \sim \chi^2(n-1) \quad \text{-and-} \quad \frac{(m-1)S_Y^2}{\sigma_Y^2} \sim \chi^2(m-1)$$

Then

$$\frac{\dfrac{\cancel{(n-1)}S_X^2}{\sigma_X^2 \cancel{(n-1)}}}{\dfrac{\cancel{(m-1)}S_Y^2}{\sigma_Y^2 \cancel{(m-1)}}} = \underbrace{\left(\frac{S_X^2}{S_Y^2}\right) \cdot \left(\frac{\sigma_Y^2}{\sigma_X^2}\right)}_{= \dfrac{S_X^2 / S_Y^2}{\sigma_X^2 / \sigma_Y^2}} \sim F(n-1, m-1)$$

NOTE: If $H_0: \sigma_X^2 = \sigma_Y^2$ is assumed to be true, then the test statistic is $\frac{S_X^2}{S_Y^2}$.

Recall: If $X_1, \ldots, X_n \sim N(0,1)$ and independent, then $\chi^2 = X_1^2 + \ldots + X_n^2$ is defined as a chi-squared r.v. with n degrees of freedom.

Result: Suppose $X \sim \chi^2(n)$ and $Y \sim \chi^2(m)$ and X and Y are independent.

n numerator degrees of freedom

$$F = \frac{X/n}{Y/m} \sim F(n, m)$$

m denominator degrees of freedom

DERIVATION: Suppose $X_1, \ldots, X_n$ is i.i.d. and normal and $Y_1, \ldots, Y_m$ is i.i.d. and normal; we know that

$$\frac{(n-1)S_X^2}{\sigma_X^2} \sim \chi^2(n-1) \quad \text{-and-} \quad \frac{(m-1)S_Y^2}{\sigma_Y^2} \sim \chi^2(m-1)$$

Then

$$\frac{\dfrac{\cancel{(n-1)} S_X^2}{\sigma_X^2 \cancel{(n-1)}}}{\dfrac{\cancel{(m-1)} S_Y^2}{\sigma_Y^2 \cancel{(m-1)}}} = \frac{S_X^2}{S_Y^2} \cdot \frac{\sigma_Y^2}{\sigma_X^2} \sim F(n-1, m-1)$$

$$= \frac{\left(\dfrac{S_X^2}{S_Y^2}\right)}{\left(\dfrac{\sigma_X^2}{\sigma_Y^2}\right)}$$

Note: If $H_0: \sigma_X^2 = \sigma_Y^2$ is assumed to be true, then the test statistic is $\dfrac{S_X^2}{S_Y^2}$.

NOTE:

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

$$\Rightarrow \mathbb{P}\left(\frac{S_X^2}{S_Y^2} \frac{1}{F^*_{\frac{\alpha}{2}, n-1, m-1}} \geq \frac{\sigma_X^2}{\sigma_Y^2} \geq \frac{S_X^2}{S_Y^2} \frac{1}{F^*_{1-\frac{\alpha}{2}, m-1, m-1}} \right) = 1-\alpha$$

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

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

A quick review of errors:

DEF: A type I error occurs when we reject the null hypothesis when it is actually true.

DEF: A type II error occurs when we do not reject the null hypothesis when it is actually false.

TRUTH

Your decision	H_0 is false	H_0 is true
Reject H_0	☑ $1-\beta$	type I error α
Don't reject H_0	type II error β	☑ $1-\alpha$

WAYS TO INCREASE POWER

1) Increase α.
2) Increase $|\mu_0 - \mu_1|$, i.e., the effect size.
3) Increase n, the sample size.

DEF: The power of a statistical test is the probability that you reject the null hypothesis when the null hypothesis is actually false.

In general, statisticians have decided to fix (in advance) the type I error rate, and then they design likelihood ratio tests (i.e., test statistics) to minimize β, i.e., maximize $1-\beta$, i.e., maximize power.