Recap

Comparing two population means:

① Paired samples: Apply one-sample procedures to the differences.

② Independent samples:

Confidence interval for $\mu_1 - \mu_2$

↓

Population normal

- Yes
  - $\sigma_1^2 = \sigma_2^2$ → $(\bar{X} - \bar{Y}) \pm t_{n_1+n_2-2, \alpha/2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
  - No assumption regarding σ_1^2 & σ_2^2 → $(\bar{X} - \bar{Y}) \pm t_{\nu, \alpha/2} \sqrt{\frac{s_X^2}{n_1} + \frac{s_Y^2}{n_2}}$, where ν is obtained by Satterthwaite approx
- NO
  - Are n_1 and n_2 large?
    - Yes → $(\bar{X} - \bar{Y}) \pm z_{\alpha/2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$ or $(\bar{X} - \bar{Y}) \pm z_{\alpha/2} \sqrt{\frac{s_X^2}{n_1} + \frac{s_Y^2}{n_2}}$
    - NO → Take advanced course

Confidence interval for $p_1 - p_2$

Are n_1 and n_2 large?

YES

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

NO

Take advanced course

Confidence interval for a normal variance

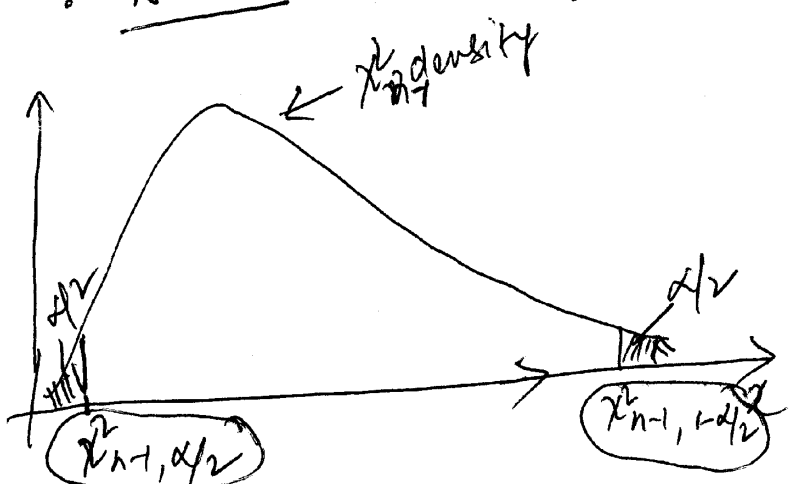
Set up: $X_1, \ldots, X_n \overset{iid}{\sim} N[\mu, \sigma^2]$ — both unknown

$100(1-\alpha)\%$ CI for σ^2:

- Result: $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$

  χ^2_{n-1}: chi-square distribution with $(n-1)$ d.f.

  — A right-skewed distribution
  — Approaches normality as d.f. $\uparrow \infty$.
  — $\chi^2_k \overset{d}{=} \sum_{i=1}^{k} Z_i^2$, where $Z_1, \ldots, Z_k \overset{iid}{\sim} N(0,1)$



- Pivot: $\dfrac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$.

Want: $P[L \le \sigma^2 \le U] = 1-\alpha$.

Have: $1-\alpha = P\left[\chi^2_{n-1,\alpha/2} \le \dfrac{(n-1)S^2}{\sigma^2} \le \chi^2_{n-1,1-\alpha/2}\right]$

χ^2_{n-1}

$= P\left[\dfrac{\chi^2_{n-1,\alpha/2}}{(n-1)S^2} \le \dfrac{1}{\sigma^2} \le \dfrac{\chi^2_{n-1,1-\alpha/2}}{(n-1)S^2}\right]$

$= P\left[\dfrac{(n-1)S^2}{\chi^2_{n-1,1-\alpha/2}} \le \sigma^2 \le \dfrac{(n-1)S^2}{\chi^2_{n-1,\alpha/2}}\right]$

- How to get a CI for σ^2?
- How about a CI for SD σ?

$\Rightarrow$ $100(1-\alpha)\%$ CI for σ^2: $\left[\dfrac{(n-1)s^2}{\chi^2_{n-1,1-\alpha/2}}, \dfrac{(n-1)s^2}{\chi^2_{n-1,\alpha/2}}\right]$.

Confidence interval for ratio of two normal variances

set up:

$X_1, \ldots, X_{n_1} \overset{\text{ris}}{\sim} N[\mu_X, \sigma_X^2]$

$Y_1, \ldots, Y_{n_2} \overset{\text{ris}}{\sim} N[\mu_Y, \sigma_Y^2]$

> indep.

$100(1-\alpha)\%$ CI for σ_X^2/σ_Y^2:

Result:

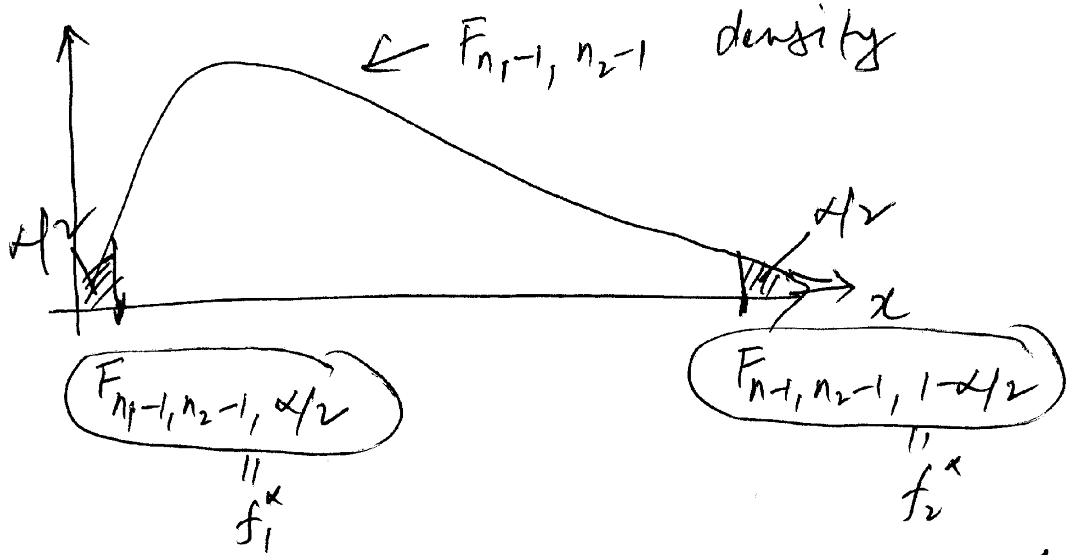
$$\frac{(n_1-1)S_X^2/\sigma_X^2}{(n_2-1)S_Y^2/\sigma_Y^2} \sim F_{n_1-1,\, n_2-1}$$

- F-distribution with (n_1-1) and (n_1-1) degrees of freedom

- If $F \sim F_{\nu_1, \nu_2}$ then $\frac{1}{F} \sim F_{\nu_2, \nu_1}$

- Pivot: $$\frac{(n_1-1)S_X^2/\sigma_X^2}{(n_2-1)S_Y^2/\sigma_Y^2} = \frac{(n_1-1)S_X^2}{(n_2-1)S_Y^2} \cdot \frac{\sigma_Y^2}{\sigma_X^2}.$$

- How to get a CI for σ_X^2/σ_Y^2?



$$\Rightarrow \quad 1-\alpha = P\left[f_1^\alpha \le F_{n_1-1,\, n_2-1} \le f_2^\alpha \right]$$

$$= P\left[f_1^\alpha \le \frac{(n_1-1)\, s_X^2}{(n_2-1)\, s_Y^2} \cdot \frac{\sigma_Y^2}{\sigma_X^2} \le f_2^\alpha \right]$$

$$= P\left[f_1^\alpha \frac{(n_2-1)\, s_Y^2}{(n_1-1)\, s_X^2} \le \frac{\sigma_Y^2}{\sigma_X^2} \le f_2^\alpha \frac{(n_2-1)\, s_Y^2}{(n_1-1)\, s_X^2} \right]$$

$$= P\left[\frac{(n_1-1)\, s_X^2}{f_2^\alpha (n_2-1)\, s_Y^2} \le \frac{\sigma_X^2}{\sigma_Y^2} \le \frac{(n_1-1)\, s_X^2}{f_1^\alpha (n_2-1)\, s_Y^2} \right].$$

CI for σ_X^2/σ_Y^2.

Note:

Suppose ~~we~~ we have a CI for θ:

$$1-\alpha = P\left[L \le \theta \le U \right] \quad \text{— } (*)$$

$$= P\left[\frac{1}{U} \le \frac{1}{\theta} \le \frac{1}{L} \right]$$

$\uparrow$ Ex.

Suppose g is a monotonically $\uparrow$ fn. of θ then:

$$(*) \Rightarrow 1-\alpha = P\left[g(L) \le g(\theta) \le g(U) \right],$$

Suppose g is monotonically $\downarrow$ fn. of θ. Then:

$$(*) \Rightarrow 1-\alpha = P\left[g(U) \le g(\theta) \le g(L) \right].$$

Homework 6

(solution)

#1]. 95 % CI : $1-\alpha = 0.95$

$n = 30$, $s^2 = 25$.

Using the formula derived earlier: the CI is.

$$\left[\frac{(29)(25)}{\chi^2_{29,\,0.975}},\ \frac{(29)(25)}{\chi^2_{29,\,0.025}} \right]$$

$\chi^2_{29,\,0.975} \rightarrow$ qchisq(0.975, 29) $\qquad \chi^2_{29,\,0.025} \rightarrow$ qchisq(0.025, 29)

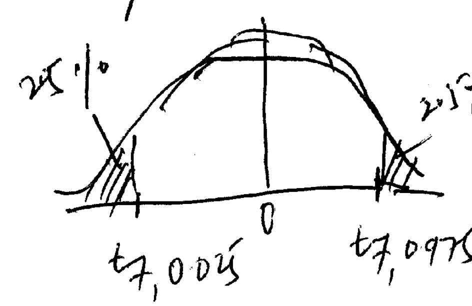
$$= \left[\frac{(29)(25)}{45.72},\ \frac{(29)(25)}{16.05} \right]$$

$$= [\ \_\_\ ,\ \_\_\_\_\ ].$$

#2.] Paired data: work with diff. (After - before)

95% CI for: $\mu_D = \mu_{after} - \mu_{before}$

Can use a t-interval under the assumption that differences are normal.



$$\bar{D} \pm t_{7,0.975} \frac{S_D}{\sqrt{8}}$$

↑

$qt(0.975, 7) = 2.364$.

$$= -6.75 \pm 2.364 \, \frac{10.067}{\sqrt{8}}$$

$$= [-15.17, \ 1.67]$$

#4] 95% CI : $p_1 - p_2$

p_1: single parent. p_2: two-parent

— Under the random sample assumption:

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2} \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

$z_{\alpha/2}$: 1.96

$$\hat{p}_1 = \frac{61}{414}$$

$$\hat{p}_2 = \frac{74}{501}$$

Plug in and get the answer.

#5] 95% CI for $\mu_{small} - \mu_{big}$.

$$(\bar{x}_{small} - \bar{x}_{big}) \pm t_{v, 1-\alpha/2} \sqrt{\frac{S_x^2}{n_1} + \frac{S_y^2}{n_2}}$$

$t_{v, 1-\alpha/2}$: R

→ Need normality assumption.

#6 (a) ~~Compare~~ Look at side-by-side box plot.
and see if equal variance
assumption is reasonable

~~(b)~~

(b) $(\bar{X}-\bar{Y}) \pm t_{?,\alpha/2} \sqrt{(\quad)}$

appropriately.