Inference for population proportion p.

想推测全校学生中，男生的占比.

猜一个 p 的大致范围	对 p 的某 claim 进行判定

p 大致在 $(0.45, 0.48)$

$p > 0.5$ 是不对的

Confidence Interval

Hypothesis Testing.

做推断就是算"概率"

Hypothesis Testing for P.

*要做 Hypothesis testing. 先满足条件.

Settings: eg. 别人对全校学生男性占比
提出个一个想法 or claim. $P>0.5$
你想检验这个想法是否有
证据支撑.

所以，你去抽了一个 SRS 的样本.

$n = 50$

构建了 Hypothesis.

原假设 null hypo. $H_0: \ p = 0.5$ ← 等于.

alternative hypothesis. $H_A: \ p > 0.5$ ← $<$, $>$ or $\neq$.

null hypothesis ← 假定已成立，
看看能不能推翻

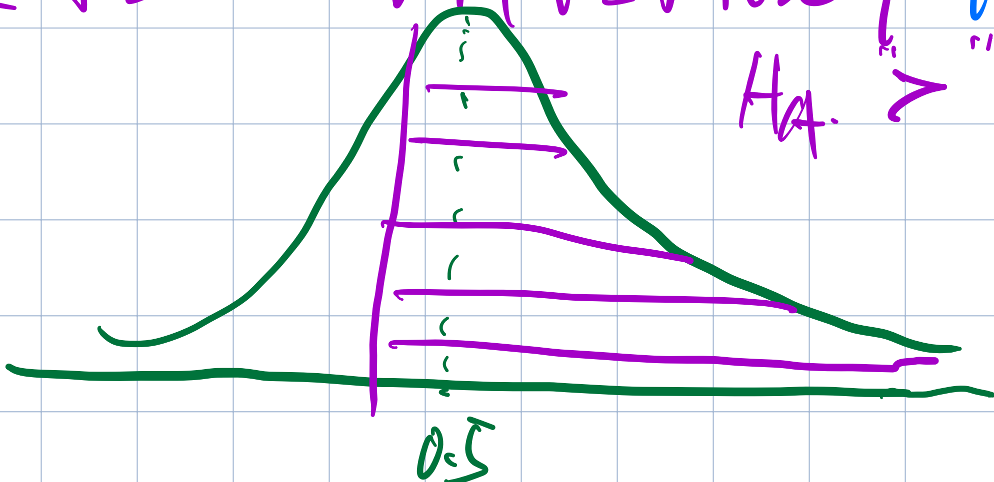
$p = 0.5$ 已成立.

$$\Rightarrow \hat{p} \sim N\left(0.5, \sqrt{\frac{0.5(1-0.5)}{50}}\right)$$

senario ①

我们取到的样本是 $\hat{p} = 0.48$

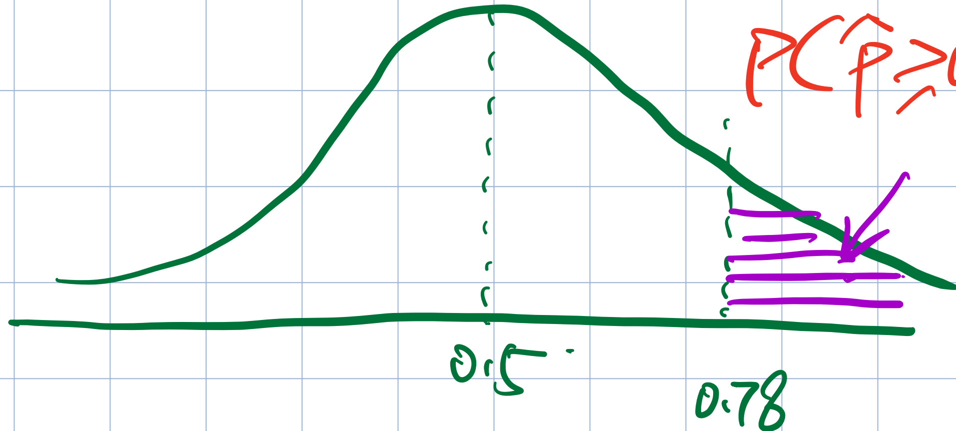
算取到此样本以及比取到此样本更符合 H_a 的样本的概率 P-value



$P > 0.5 \Rightarrow$ 取到此样本是个大概率

无法推翻 $H_0 \Leftarrow$ 事件. 没什么好奇怪的

Senario ② 抽到样本 $\hat{p} = 0.78$



$P(\hat{p} \geq 0.78) = 0.023$

推翻了 H_0
支持 H_A.

抽到一个小概率样本 $\Rightarrow$ 见鬼了

日常情况见不到鬼. 有猫腻.

原假设不符合情况!! 所以导致见鬼!

① 定 Hypothesis. H_0 $p = ?$

H_a $p <$ or $>$ or $\neq$?

② 验 3 conditions. 1) SRS.

2) 10%

3) Large count.

$\hat{p} \sim N$:

③ 算 p-value.

④ 做判断 p-value > 0.05 H_0 成立.

p-value < 0.05 H_a 成立.

State ① 解决什么问题 做 H.T.

② 用什么变量 P is ......

Plan ① 想用什么方法.

② 用此方法是否符合条件

Do ① 记算关键指标.

Conclude ① 做出结论.

Hypothesis test. 有claim 验证对不对.

Confidence Interval 求 pop. parameter

一个大致范围.

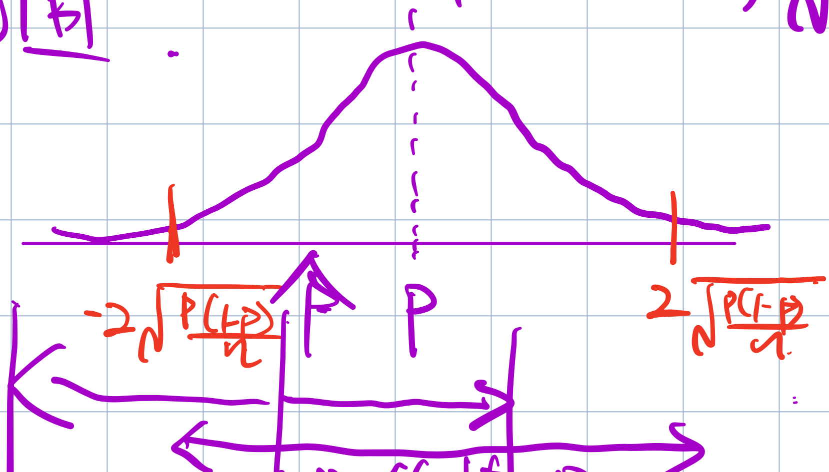
全体参数

全体均值 μ_x

全体比例 P.

$\hat{p} = \frac{9}{20} = 0.45 \quad n = 20$ 问全校的 p 在什么范围

$\hat{p} \sim N\left(p, \sqrt{\frac{p(1-p)}{n}}\right)$



95%的情况

$\hat{p}$ 在这个范围

95%的情况下.

随这个范围

$$\longleftarrow \hat{p} \longrightarrow$$

$$-2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \qquad +2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

95%的情况 $\left(\hat{p}-2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p}+2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$

包含 P. $(0.227, 0.672)$

$$\underline{1.344} \Rightarrow \frac{0.344}{1.344}$$

$p = 70\% = 0.7$

控制半径.

$$\doteq 2\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < 0.03$$

$$\frac{2\sqrt{\hat{p}(1-\hat{p})}}{0.03} < \sqrt{n}$$

$$\sqrt{n} > \frac{2\sqrt{\hat{p}(1-\hat{p})}}{0.03}$$

$$n > \frac{2}{0.03^2}\boxed{\hat{p}(1-\hat{p})}$$

$\hat{p} = 0.5$ 时，$\hat{p}(1-\hat{p})$ 达最高值

$$n > \frac{2}{0.03^2} \times 0.5 \times 0.5$$

$$n > 555.55$$

confidence interval → 区间. C.I.

区间中心、sample statistics.

区间半径. Margin of Error MOE

Margin of Error = Critical Value C.V.

× Standard Error S.E.

置信度

Confidence Level 95%

Confidence level 不是概率.

C.L = 95% ⇒ CV = 1.96

Standard Error $\sqrt{\dfrac{\hat{p}(1-\hat{p})}{n}}$.

Sample 值.

$n=60$ $\hat{p}=0.52$. $p=0.8$? $p>0.8$?

1) 定假设.

2) 验条件.

3) 算 p-value

4) 做结论.

D定假设.

Ta说: $p>0.8$ or $p=0.8$.

你觉得: $p<0.8$

变成 ⇓ H_A.

H_A: $p<0.8$

⇓ 将 $>$ or $<$ or $\neq$ 改成 "="

用作 H_0

H_0: $p=0.8$

H_0 is called Null hypothesis. 原假设.

假定已经成立. 就把 $p=0.8$ 作为 pop. proportion.

⇒ 可构建 sampling distribution

$$\hat{p} \sim N\left(0.8, \sqrt{\frac{0.8 \times 0.2}{60}}\right).$$

⇒ 查一查取到这样 $(\hat{p}=0.52)$ 一个样本.

或比此样本更符合 H_A 方向样本 $(\hat{p}<0.52)$

的概率 → p-value.

如果 p-value 很大（$p\text{-value} > 0.05$）

⇒ 很轻松取到这样样本，没什么大不了。

⇒ 瞎情况，无法推翻 H_0

如果 p-value 很小（$p\text{-value} < 0.05$）

⇒ 取到这样样本非常不可能

⇒ 你取到了这样的样本，见鬼了。

⇒ 不该见鬼有东西出问题了。

⇒ 原假设出问题了。

⇒ 推翻原假设 H_0（$P = 0.8$）

计算

$\hat{p} \sim N\left(0.8, \sqrt{\frac{0.8\times0.2}{60}}\right)$

$P\text{value} = P(\hat{p} \le 0.52) = \text{normalcdf}\left(\begin{array}{l} \text{lower}: 0 \\ \text{upper}: 0.52 \\ \mu: 0.8 \\ \sigma: \sqrt{\frac{0.8\times0.2}{60}} \end{array}\right)$

≈ 0

结论: $p\text{-value} \approx 0 < 0.05$ ∴ 推翻 H_0.

∴ H_A 成立.

4-step process

State ← 目的，目标，settings定义

Plan ← 用什么方法，验证条件

Do ← 关键指标计算

Conclude ← 根据关键指标作出结论。

4-step for HT.

1). State: We are testing

做的工作

$H_0: p = 0.8$

$H_A: p < 0.8$

细节

at $\alpha = 0.05$

先定 H_A，将 H_A 设置为你想证明的
再定 H_0。H_A 是用 ">" or "<" or "≠"
H_0 用 "=".

α or α-level 是你设定的小概率
的标准。

变量定义

where p is the true proportion of female student among all students in our school.

Plan：1)声明所用方法． 2)验证方法适用条件

We are using 1-sample z-test for p.

其它情况 2-sample t test for $\mu_1-\mu_2$

1-sample 因为只有p不是 p_1-p_2.

若是 p_1-p_2 则为 2-sample. 如例 $\mu_1-\mu_2 \Rightarrow$ 2-sample

z-test 中的 z 对应 for p. 若是 for μ 则是 t

原因是 sampling distribution 形状不同.

z-test 中的 test 对应我们做 hypothesis test.

1-sample
- z系列 → p
- t系列 → μ

(μ, p)

2-sample
- z系列 → $p_1 - p_2$
- t系列 → $\mu_1 - \mu_2$

multiple sample (≥3)
- χ^2-test
- ANOVA

→ test

interval or test

验条件：3条件（大部分）

1) SRS the sample of 60 students was randomly selected. ✓

(It is reasonable to assume)

2) 10% condition the sample size (60) is less than 10% of the population size (600) ✓

test for p

3) $np_0 = 60 \times 0.8 = 48 \geq 10$

$n(1-p_0) = 60 \times 0.2 = 12 \geq 10$ ✓

Do：记算关键指标（for test 计算 p-value）

用计算器，写明参数

By using calculator

Hu里的P值

1-propZ-test

函数名

P_0 : 0.8

x : $\hat{p} \cdot n = 0.47 \times 60$ ← 4舍5入

n : samplesize 60

prop : "<"

Ha中不等号方向

We get p-value ≈ 0

$z = -6.4549$.

Conclude：

Because our p-value (≈ 0) is less than our α-level (0.05). We can reject Ho. We do have convincing evidence that the true proportion of female student among all student is less than 0.8

Type I Error : R O T

Reject Ho

Ho True.

Probability of making type I error: α-level

Type II error : F O F.

fail to reject H0. H0 false.

Probability of Type II error:

→ β-level.

= 1 − Power of test.

Power of test ⇐

1) Sample size.

n↑ power↑ β↓

2) α-level

α↑ → power↑ → β↓

3). effective gap.

d.

Confidence Interval

$\alpha = P_0 - P$

$\alpha\uparrow$ Power$\uparrow \to \beta\downarrow$

找 pop. parameter 的大致范围

有 $\hat{p}$，找 $P \in C$ ， How?

→ 围绕 $\hat{p}$ 构建 alternative sampling distribution.

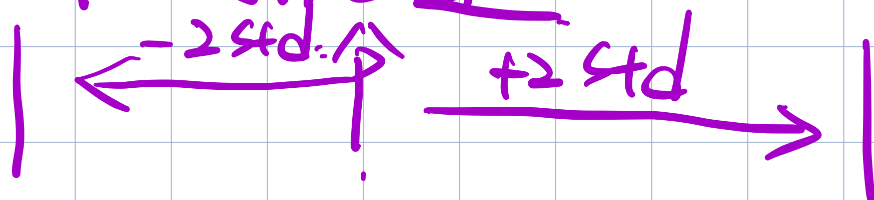
why 叫 alternative? 不是 real sampling dist.

real sampling distri ← 以 pop. param 为中心。

alternative sampling distri ← 以 sample statistics ($\hat{p}$ or $\bar{x}$) 为中心。

1) 以 $\hat{p}$ 为中心、以 $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ 为 std. 构建 alt-sampling distri: $\tilde{p} \sim N\left(\hat{p}, \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$

$\tilde{p}$ 是 p 的可能值



lower bound of confidence level

0.47

upper bound of confidence level:

$0.47 - \sqrt{\frac{0.47 \times 0.53}{60}}$

$0.47 + 2 \times \sqrt{\frac{0.47 \times 0.53}{60}}$

0.341

0.598

$p \in [0.341, 0.598)$

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

95% 的 $\hat{p}$.

$SRR \Rightarrow E(\hat{p}) = p \Rightarrow \hat{p}$ 和 p 应该很像.

$\hat{p}$ is the best guess for p.

Confidence interval 置信区间.

Sample Statistics $\pm$ Margin of Error

$\hat{p}$

Std of clt-sampling distri

Critical Value $\times$ Standard Error

用 confidence Level 算

(c-level)

用公式

c-level 作为 [illegible] $\xrightarrow{\text{invNorm}}$ z 值.