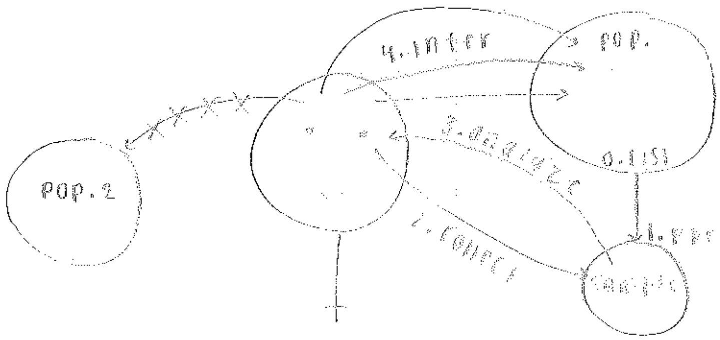
hypothesis testing

one sample

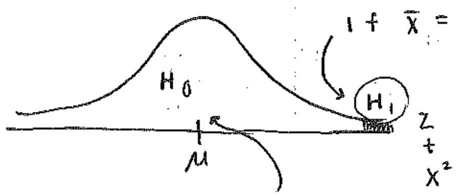


- **hypothesis testing**: a procedure, based on sample info by which one "accepts" or "rejects" the hypothesis
  ↳ "fail to reject"
- **Null hypothesis**: H_0 the hypothesis set up to test against
  ↳ the standard
  - "there is no change / difference"
- **Alternate hypothesis**: H_1 the hypothesis to be accepted if the the null is rejected
  ↳ the alternative
  - "there is a difference"
  - 3 looks of the alternate hypothesis
    - ↳ one-tailed test to the right if $\mu > 100$, critical region on right
    - ↳ one-tailed test to the left if $\mu < 100$, critical region on left
    - ↳ two-tailed test if $\mu \neq 100$, has critical region on both sides

$H_0 : \mu = 100$

$H_1 : \mu \gtrless 100$

★ μ will not always be 100, this was just an example ★



if $\bar{x} = 135.79$, $z = 4.56$
reject H_0

if $\bar{x} = 100.57$, $z = .123$ null hypothesis is not rejected

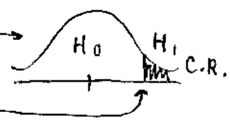
- hypothesis testing errors
  - **type I error**: rejecting the null hypothesis when it's true
    - ↳ worst error; "α" = level of significance = .1, .05, .01
  - **type II error**: accepting the null when it is false
    - ↳ "β"

increase sample size = increase power
similar to probability

	"fail to reject"	"reject"
H_0 : true	☺	type I α = .1, .05, .01
H_0 : false	type II β	power of a test $(1-\beta)$

→ uncomfortable but correct

4 ingredients for a statistical test:

1. H_0 $p = .97$
2. H_1 $p > .97$ → pic


3. critical value
   $z_c = 1.645, 1.96...$
4. $\hat{p} \rightarrow z =$

also p-value, but not necessary

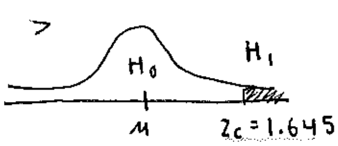
print out project checklist
  don't need samples / graphs but do bivariate data ☆
  ↳ scatterplot...

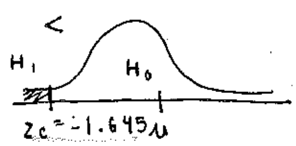
project due end of April

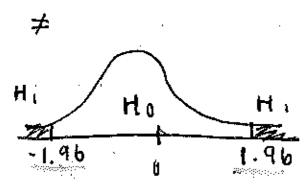
9.2 (11-14, 16) print template

SOH - CAH - TOA
"five is one and one is two"

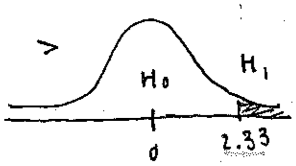
$\alpha = .05$

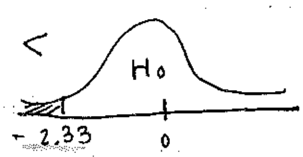


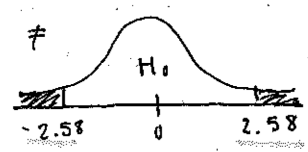




$\alpha = .01$ ← usually for risky topics







test μ when σ is known
  SRS, get $\bar{x}$

ex. sun spots $\bar{x} = 47.0$ $\sigma = 35$ (prev. studies) $\mu = 41$ (avg. from thousands of years) $\alpha = .05$ (default) $n = 40$
  "higher than"
  H_1: >

→

inference template

name of test

1-sample mean z-test

$H_0: \mu = 41$ sun spot cycles have an avg of 41 per cycle

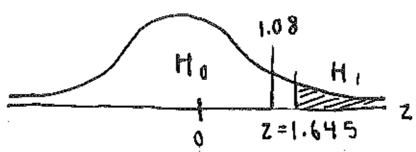
$H_1: \mu > 41$

conditions:

- SRS
- independent
- $n = 40 > 30$ ∴ CLT invoked →
- σ known

"it's reasonable to assume blah blah blah"

pic.:



$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$ $\bar{x} = 47$ $z = \frac{47-41}{35/\sqrt{40}} = 1.08$

conclusion:

- r H_0 ~~H_1~~ $\alpha = .05$
- ∄x insufficient statistical evidence suggesting there are higher than 41 sun spots ∴ $\mu \approx 41$

ex. sprinkler system for fires

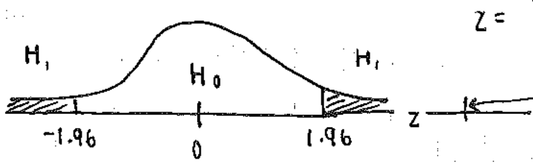
$\mu = 130°F$ $n = 81$ $\bar{x} = 131.08°F$ $\sigma = 1.5°F$ a contradiction of H_0?

$H_0 = 130$ $H_1: \mu > 130$ $\alpha = .05$ ≠

1-sample mean test w/z curve

$H_0: \mu = 130°F$ Sprinklers are activated at an avg of 130°F

$H_1: \mu \neq 130°F$



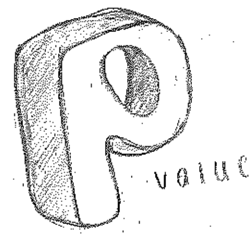
$z = \frac{131.08 - 130}{1.5/\sqrt{81}} = 6.48$

conclusion:

- r ~~H_0~~ H_1 $\alpha = .05$
- ∃x statistical evidence suggesting that the sprinklers do activate at a temp. different than 130°F also > 130°F

9.2 #11 $\mu = 16.4$ ft $n = 36$ $\bar{x} = 17.3$ $\sigma = 3.5$ ft $\alpha = .01$

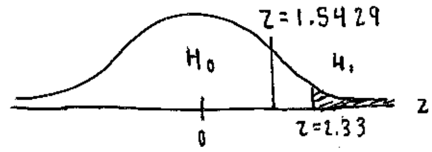
is it increasing? is $H_1: \mu > 16.4$ ft



√'s
- RRS
- indep.
- $n = 36 > 30$ ∴ CLT invoked
- σ known, z

$H_0: \mu = 16.4$ ft

$H_1: \mu > 16.4$ ft



$$z = \frac{17.3 - 16.4}{3.5/\sqrt{36}} = 1.5429$$

H_0 $\cancel{H_1}$ $\alpha = .01$

∃x insufficient evidence suggesting the storm is increasing above severe rating of 16.4 ft ∴ $\mu \approx 16.4$ ft

p-value shows prob. of getting results if the null is true

$\mu = 68.7$g $\mu = 70$g $p = .18$ $\alpha = .05$

$p > \alpha$ so accept null

$H_0: \mu = 70$g 18% chance of getting this result (mean of 68.7 or less) from a sample of this size and variation

education, psych, sociology: $\alpha = .20, .10$ bc less serious

if p is small, there is sufficient evidence against null

read stat 9.2 peck 9.2 add p-values

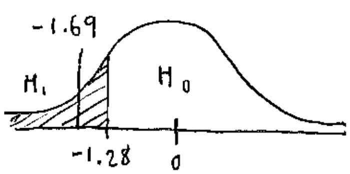
ex. pepperdine cafe

$\mu = 41$ min $n = 40$ $\bar{x} = 38$ min $\sigma = 11.2$ min $\alpha = .10$

$H_0: \mu = 41$ min avg time spent in pepp. cafe was 41 min at lunch

$H_1: \mu < 41$ min

use z-table $z = -1.28$ shows $\alpha = .10$
or t-table with one-tail



Z = SNC

$$z = \frac{38 - 41}{11.2/\sqrt{40}} = -1.69 \quad p = .0455$$

$\cancel{H_0}$ H_1 $\alpha = .10$

$p < \alpha$, strong evidence against the null

∃x sufficient statistical evidence that pepp. students spend < 41 min...

testing μ when σ unknown

- get $\bar{x}$ and s from a sample

1. H_0, H_1, α
2. $t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ with degrees of freedom $n-1$
3. use t table to find p-value

9.2(17-20, 22)

checks:

independent

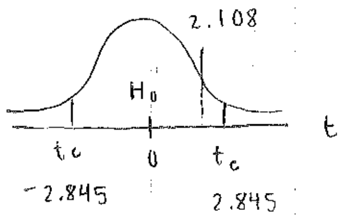
RRS

$n < .1N \rightarrow$ "it is reasonable to assume ..."

ex. leukemia drugs

$\bar{x} = 17.1$ $s = 10.0$ $\mu = 12.5$ $\alpha = .01$ $n = 21$

$t = 2.108$



$\bar{x} \rightarrow t = \frac{\bar{x} - \mu}{s/\sqrt{n}}$ t-score

$t_c = 2.845$

9.3(9-12, 18-20)

H_0 $\cancel{H_1}$ $\alpha = .01$

$\exists x$ insufficient statistical evidence suggesting that the mean remission time of 6-mp is different from 12.5 weeks $\therefore \mu \approx 12.5$

go to t table looking for $t = 2.108$ $\therefore$ p-value between .05 and .02

$.02 < pvlu < .05 \rightarrow$ $\alpha = .01 < .02 < pvlu < .05 \rightarrow$ fail to reject null

"weak evidence against null"

25. c = confidence level, α = significance level for a two-tailed test

null hypothesis $H_0: \mu = k$

reject H_0 whenever k falls outside of $c = 1 - \alpha$

fail to reject if in the interval $c = 1 - \alpha$

$H_0: \mu = 20$, $H_1: \mu \neq 20$ $n = 36$ $\bar{x} = 22$ $\sigma = 4$ $\alpha = .01$

a) $c = 1 - \alpha = 1 - .01 = .99$

$ME = z_c\left(\frac{\sigma}{\sqrt{n}}\right) = 2.58\left(\frac{4}{\sqrt{36}}\right) = 1.72$

$22 - 1.72 < \mu < 22 + 1.72 \rightarrow 20.28 < \mu < 23.72$

$k = 20$, not in interval: reject

b) $z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = 3.0000$ $pvlu \approx .0026 < \alpha = .01$, reject null

1 sample proportion z-test

"binomial"

$np > 5$ and $nq > 5$ → $n(\pi) > 5$ $\quad n(1-\pi) \geq 10$

$\hat{p} = \frac{r}{n}$ → $Z = \dfrac{\hat{p} - \pi}{\sqrt{\pi(1-\pi)/n}}$

ex. eye surgery

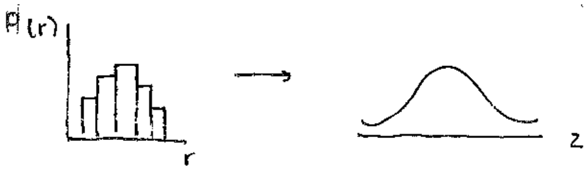
$\pi = .30$ $\quad n = 225$ $\quad r = 88$ $\quad \hat{p} = 88/225 = .3911$ $\quad \alpha = .01$

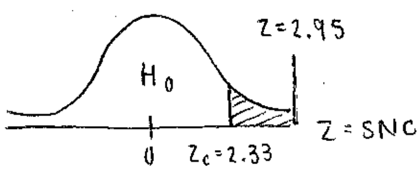
$H_0 : \pi = .30$

$H_1 : \pi > .30$

RRS, independent, $n\pi > 5$ $\quad n(\pi - 1) > 5$

↳ $67.5 > 5$ $\quad$ ↳ $157.5 > 5$ $\quad \therefore$ ANAIJ





$Z = \dfrac{\hat{p} - \pi}{\sqrt{\dfrac{\pi(1-\pi)}{n}}} = 2.95$

pvlu $= .0016 < .01$ → strong evidence against the null

~~H0~~ $\quad$ H_1 $\quad \alpha = .01$

∃x strong statistical evidence suggesting a higher rate of restoration

If $n\pi > 5$ but not > 10, say "possibly concerned that sample is not large enough"

10.1 (6, 7, 9, 12, 13, 15)

9 review; mileage

μ = 11.1 thousand   11,100

σ = 600

n = 36

$\bar{x}$ = 10.8   10,800

α = .05

"different from" $\neq$

stat → calc → z-test $\bar{x} \to z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} = \frac{10.8 - 11.1}{.6/\sqrt{36}} = -3$

z_c = 1.96 and -1.96   large z-value, small p-value

pvlu < α "SEVN"

.0027 < .05

$\uparrow n \ \downarrow \sigma \to \uparrow \frac{z \text{ or}}{T} \ \downarrow p$

π = .35

n = 81

r = 39

$\hat{p}$ = .4815

" more than"

α = .05

pvlu < α "SEVN"

.0066 < .05

μ = 48

n = 10

$\bar{x}$ = 44.2

s = 8.61

α = .05

"less than"

t = -1.3952

p = .0982

p vlu > α "WEVN"

.098 > .05

hypothesis testing

two samples

hypothesis testing, dependent groups

paired data samples that help draw conclusions about the difference of 2 groups
before / after, matching even without before / after
- using matched data reduces extraneous factors; reduces variance, increases accuracy

$\bar{d}$ (basically $\bar{x}$): mean difference between the paired data
s_d: sample standard deviation
↳ $\sqrt{\frac{\Sigma(d-\bar{d})^2}{n-1}}$

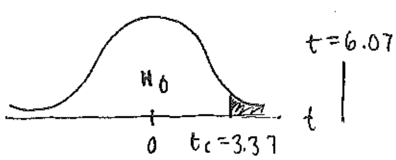
$\boxed{\text{2 dep. mean t-test}}$

$H_0: \mu_d = 0$ "no difference"
$\bar{d} \to t = \frac{\bar{d}-\mu_d}{s_d/\sqrt{n}}$

$H_0: \mu_d = 0$ avg diff. of creativity training − control ≈ 0
$H_1: \mu_d > 0$

$\bar{d} \to t = \frac{\bar{d}-\mu_d}{s_d/\sqrt{n}} = \frac{9.83-0}{3.97/\sqrt{6}} = 6.06 \approx 6.07$

$\alpha = .01$
$n = 6$
$df = 5$
$t_c = 3.365$



$\cancel{H_0}$ H_1 $\alpha = .01$

pvlu ≤ 0.005 < α .01 "SEVN"

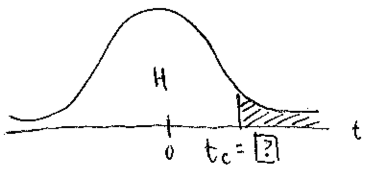
checks:

RRS
independent / dependent
↳ all independent from one another ↳ within each thing being tested
n < .1N "it is reasonable to assume..."

baby IQ example:

$H_0: \mu_d = 0$ no diff between age 3 and age 8

$H_0: \mu_d > 0$



p value < α

.001 < .05

"SEVN"

$t = 3.17$

confidence interval: $\bar{d} - E < \mu_d < \bar{d} + E$

9.91

$2.39 < \mu_d < 17.43$

$E = t_c(s_d / \sqrt{n})$

$E = t_c \frac{22.11}{\sqrt{50}}$

98% C

$t_c \approx 2.412$

$E = 7.52$

population average difference ... 98 times hypothesis testing σ unknown

2 independent sample means t-test

checks:

2 independent, random, n_1 and $n_2 \geq 30 \therefore$ CLT (or mound-shaped) $n_1 < .1N_1$ and $n_2 < .1N_2$

σ_1 and σ_2 unknown

degrees of freedom:

- $n_1 - 1$ and $n_2 - 1 = n_1 + n_2 - 2$
- smaller $n-1$
- satterthwaite's formula

Null / Alternate Hypothesis

H_0: $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$ identify μ_1 and μ_2

H_1: $\mu_1 < \mu_2$

$\mu_1 > \mu_2$

$\mu_1 \neq \mu_2$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

use t-table for this

1. find d.f. and then t_c

confidence interval

$(\bar{x}_1 - \bar{x}_2) - E < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + E$

$$E = t_c \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

if only negative values, $\mu_1 < \mu_2$

if only positive, $\mu_1 > \mu_2$

if both, no conclusion

↳ $\mu_1 \approx \mu_2$

ex. 5

advil: $\bar{x}_1 = 20.1$ min $s_1 = 8.7$ $n_1 = 12$

tylenol: $\bar{x}_2 = 11.2$ min $s_2 = 7.5$ $n_2 = 18 \rightarrow$ d.f. $= 7$

H_0: $\overset{A}{\mu_1} = \overset{T}{\mu_2}$ the time it takes to enter the bloodstream is the same

H_1: $\mu_1 > \mu_2$

$t = 2.437$ $t_c = 1.895$

$.010 < \text{pvlu} < .025 < \alpha = .05$ "SEVN"

$$E = 1.895 \sqrt{\frac{8.7^2}{12} + \frac{7.5^2}{12}} = 6.92$$

$1.98 < \overset{A}{\mu_1} - \overset{T}{\mu_2} < 15.82$ "if we took..."

$+ \rightarrow + \Rightarrow \mu_1 > \mu_2$

population difference of means $\mu_1 - \mu_2$

2 independent sample proportion z test

n_1, r_1, p_1 $\quad$ n_2, r_2, p_2

for large samples, distribution is approximation of normal

$$\hat{p}_1 - \hat{p}_2 = \frac{r_1}{n_2} - \frac{r_2}{n_2}$$

H_0: $p_1 = p_2$ "no difference"

H_1: $p_1 < p_2$

$p_1 > p_2$

$p_1 \neq p_2$

$$\bar{p} = \frac{r_1 + r_2}{n_1 + n_2} \qquad \bar{q} = 1 - \bar{p}$$

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\frac{\bar{p}\bar{q}}{n_1} + \frac{\bar{p}\bar{q}}{n_2}}}$$

checks:

RRS

Indep.

$n_1 < .1N_1 \quad n_2 < .1N_2$

$$\left.\begin{array}{l} n_1\bar{q} > 5 \\ n_1\bar{p} > 5 \\ n_2\bar{q} > 5 \\ n_2\bar{p} > 5 \end{array}\right] \text{ANAIJ}$$

ex.7 voting

$n_1 = 625$ $\quad r_1 = 295$ $\quad p_1 = .472$ $\quad \alpha = .05$

$n_2 = 625$ $\quad r_2 = 350$ $\quad p_2 = .56$

H_0: $p_1 = p_2$

H_1: $p_1 < p_2$

$z = -3.11$

$p = .0009 < \overset{\alpha}{.05}$ ∴ "SEUN"

$$\hat{\sigma} = \sqrt{\frac{\hat{p}_1\hat{q}_1}{n_1} + \frac{\hat{p}_2\hat{q}_2}{n_2}}$$

Find z_c for confidence interval

$\text{Error} = z_c \cdot \hat{\sigma}$

$(\hat{p}_1 - \hat{p}_2) - E < p_1 - p_2 < (\hat{p}_1 - \hat{p}_2) + E$ $\quad$ pop. difference of proportions $p_1 - p_2$

10.3 (4, 9, 11, 12, 17, 22)