

Statistics II :-

* مزید ملاحظہ کیجئے
 کہ فارمولہ ان کے لیے
 غالباً 2 بجوٹ ہے

Problem 4 :-

Lower
 & upper

8-4 :-

Gain in a circuit is $N(\mu, 20^2)$

@ 95% C.I for μ when $n=10, \bar{X}=1000$

$$\alpha = 0.05$$

$$1000 - 2 \frac{20}{\sqrt{10}} \leq \mu \leq 1000 + 2 \frac{20}{\sqrt{10}}$$

$$987.6 \leq \mu \leq 1012.4$$

(b) 95% C.I for μ when $n=25, \bar{X}=1000$

$$1000 - 1.96 \frac{20}{\sqrt{25}} \leq \mu \leq 1000 + 1.96 \frac{20}{\sqrt{25}}$$

$$992.2 \leq \mu \leq 1007.8$$

(c) 99% $n=10$

$$\alpha = 0.01$$

2) دقیق
 precision

$$1000 - Z \leq \mu$$

(d) 99% $n=25$

8-78-

The same Gain estimation

① $h = ?$ to achieve 95% c.I. of 402

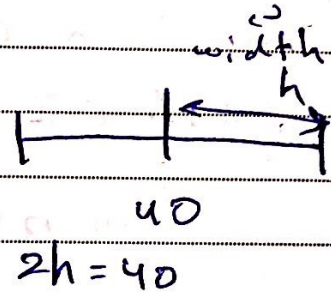
$$Z_{\frac{\alpha}{2}} = 1.96$$

$$h = Z_{\frac{\alpha}{2}} \frac{6}{\sqrt{n}}$$

$$20 = 1.96 \frac{20}{\sqrt{n}}$$

$$\Rightarrow h = 3.84$$

$$h \approx 4$$

② $h = ?$ 99%

$$c.I. = 40$$

$$Z_{.005} \quad n = \left(\frac{2.58 \cdot 20}{20} \right)^2 = 6.66$$

$$h \approx 7$$

Confidence
precision
c.I. is the upper tail 100% percentage pt.

C.I. on μ, σ^2 unknown

$t_{\alpha, k}$ is the upper tail 100% percentage pt.

$$Z_{\alpha, k} = -t_{\alpha, k}$$

8-148-

X: Life in hours of a light bulb.

$$X \sim N(\mu, 25^2), \quad n=20, \quad \bar{X} = 1014 \text{ hours}$$

(a) 95% C.I. two-sided on the mean.

$$1014 - \leq \mu \leq 1014 + Z_{0.025} \frac{25}{\sqrt{20}}$$

$$1003.5 \leq \mu \leq 1025$$

(b) Lower sided

$$L \leq \mu$$

$$1014 - Z_{\alpha} \frac{25}{\sqrt{20}} \leq \mu$$

$$1.65 \quad 1005 \leq \mu$$

$$P\left\{ -t_{\frac{\alpha}{2}, n-1} \leq \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \leq t_{\frac{\alpha}{2}, n-1} \right\} = 1 - \alpha$$

$$\left(\mu \leq \bar{X} + t_{\frac{\alpha}{2}, n-1} \frac{s}{\sqrt{n}} \right) = 1 - \alpha$$

8-248- one-sided

(a) 95%, $df=14$

$$t_{0.05, 14} = 1.761$$

5.12
1.761
1.11

(b) 99%, $df=14$

$$t_{0.01, 14} = 2.539$$

99.9%, $df=24$

$$t_{0.001, 24} = 3.467$$

8-250-

$$\begin{aligned} \text{(a) mean} &= \frac{\text{sum}}{n} = \frac{251.845}{10} = 25.1848 \\ &= (1.605)^2 = 2.576 \end{aligned}$$

$$SE_{\text{mean}} = \frac{\text{Std. dev.}}{\sqrt{n}}$$

(b) 95% C.I on μ .

$$\bar{x} - t_{\frac{\alpha}{2}, n} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\frac{\alpha}{2}, n} \frac{s}{\sqrt{n}}$$

$$2.262 \pm \frac{1.605}{\sqrt{10}}$$

(26) practice

8-26

$$a) SE_{\text{mean}} = \frac{\text{Std. dev.}}{\sqrt{n}} = \frac{6.11}{\sqrt{n}} = 1.58$$

$$\Rightarrow N=5$$

$$\text{mean} = \frac{\text{sum}}{n} = 50.0433$$

b) 95% C.I on μ

$$46.709 < \mu \leq 53.977$$

8-31

$$\bar{y} = 118.3, S = 9.9, n =$$

$$t_{.01, 13} = 2.65$$

$$\mu \leq 125.312$$

~~8-36~~

8-36 a-

@ Normal

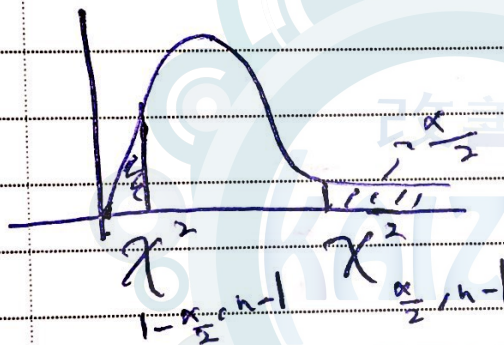
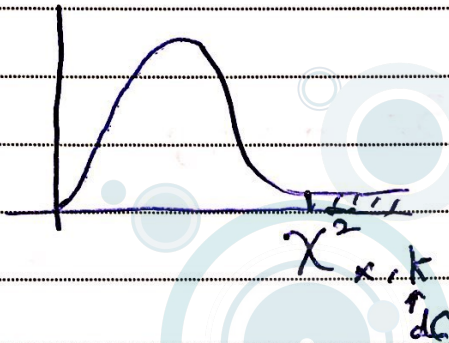
$$\bar{x} = 16.98$$

$$(b) P\left(\bar{x} - t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}} < \mu \leq \bar{x} + t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right)$$

$$16.455 < \mu \leq 17.505$$

C. I σ^2 of a normal dist.

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$



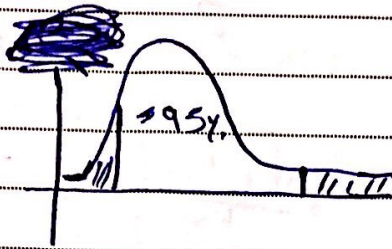
$$P\left(\frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}, n-1}} \leq \sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}, n-1}}\right) = 1-\alpha$$

$$\sigma^2 \leq \frac{(n-1)S^2}{\chi^2_{1-\alpha, n-1}}$$

$$\frac{(n-1)S^2}{\chi^2_{\alpha, n-1}} \leq \sigma^2$$

8-430-

(a)



13.848

b) 21.67

$$c) \chi^2_{\frac{\alpha}{2}, df} = \chi^2_{.05, 19} = 30.14$$

$$\chi^2_{1-\frac{\alpha}{2}, df} = \chi^2_{.95, 19} = 10.117$$

8-510-95% for σ , $n=15$, $s=.00831$

$$\sigma^2 \leq \frac{(n-1)s^2}{\chi^2_{1-\alpha, df}}$$

$$\chi^2_{.95, 14}$$

$$\chi^2_{.95, 14} = 6.571$$

$$\sigma^2 \leq 0.000145$$

$$s = \frac{\sum (x - \bar{x})^2}{n-1}$$

n-1

It's not valid
because it's
not normally
distributed.

* Large sample confidence interval for pop. proportion

$$\hat{p} = \frac{X}{n}$$

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

$$P\left(\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \leq \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 1 - \alpha$$

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{E}\right)^2 \hat{p}(1-\hat{p})$$

if we don't have $\hat{p}$ $\Rightarrow p = 0.5$

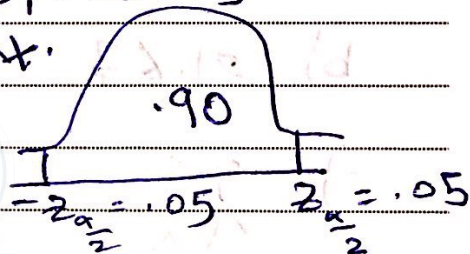
Practice 8-57

success $\Rightarrow$ underweight.

$$X = 12$$

$$n = 30$$

$$\hat{p} = \frac{12}{30} = .4$$



$$P\left(\hat{p} - z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq \mu \leq \hat{p} + z_{\frac{\alpha}{2}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right) = 1 - \alpha$$

a) 90%

$$P\left(.4 - 1.65 \sqrt{\frac{.4(.6)}{30}} \leq \mu \leq .4 + 1.65 \sqrt{\frac{.4(.6)}{30}}\right) = .90$$

$$z_{.95} =$$

$$P(.25$$

b) $E < .02$

$$n = \left(\frac{1.96}{.02}\right)^2 (.4) * .6 = 2304.96 \approx 2305$$

c) At least 95% confident.

$$n = \left(\frac{1.96}{.02} \right)^2 * 0.5 * 0.5 = 2401$$

انتظاری علی آخر اشی قیاسی (self-study)

sample

10 =

estimation C.I
تجزیه رقم 11

$$L \leq X_{n+1} \leq u$$

$$\bar{X} - t_{\frac{\alpha}{2}, n-1} S \sqrt{1 + \frac{1}{n}} \leq X_{n+1} \leq \bar{X} + t_{\frac{\alpha}{2}, n-1} S \sqrt{1 + \frac{1}{n}}$$

prob. 8-61

که تان راجه کیست بنظر آید n

$$\bar{X} = 60139.7, s = 3645.94, n = 16$$

$$\alpha = .05$$

$$t_{\frac{\alpha}{2}, n-1} = t_{.025, 15} = 2.131$$

$$60139.7 - 2.131 (3645.94) \sqrt{1 + \frac{1}{16}} \leq X_{n+1} \leq 60139.7 + 2.131 (3645.94) \sqrt{1 + \frac{1}{16}}$$

CH. 98- Hypotheses testing :-

population parameter can be estimated point estimate
C.I

Two claims "statements" about pop. parameter :-

Hypotheses

Let's say we are interested in the length of ~~product~~ ^α

$\mu = 5$

H_0 : null hypothesis

H_1 : Alternative

$\left. \begin{array}{l} H_0 : \mu = 5 \\ H_1 : \mu \neq 5 \end{array} \right\} \Rightarrow$ Two sided Hypotheses testing.

$\mu < 5$
OR $\mu > 5$

one-sided H_1

$H_0 : \mu = 5$

$H_1 : \mu > 5$

$H_0 : \mu = 5$

$H_1 : \mu < 5$

Sample from pop. $\rightarrow$ test statistic to get information

$\rightarrow$ if Information is cons. with H_1 we reject H_0 .

* The acceptance or rejection should be accompanied with uncertainty ~~prob.~~ b.c of the sampling, prob. of wrong conclusion with the decision.

$\bar{X} \Rightarrow$ estimate μ

(i) close to μ . ✓ μ .

(ii) if $\bar{X}$ is significantly different than μ .

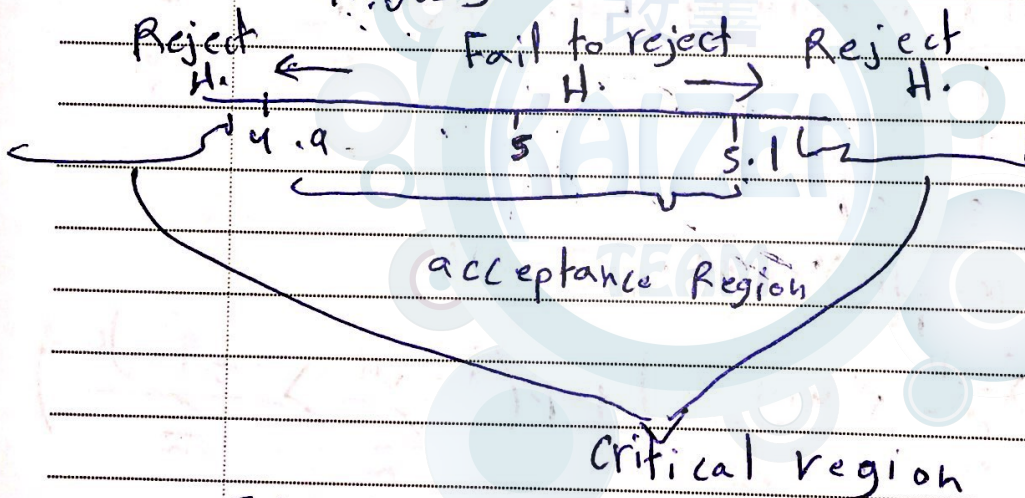
$\Rightarrow$ supports H_1

We need acceptance region

ex.:

$$4.9 \leq \bar{X} \leq 5.1$$

$$H_0: \mu = 5$$



5% α level

Two wrong conclusions :-

① To reject H_0 while it's true = Type I error
if $\bar{X}$ falls in the critical region.

② Fail to reject H_0 while it is false = Type II error
 $\bar{X}$ falls in the acceptance region

Decision H_0 is true

✓

Type I

 H_0 is false

Type II

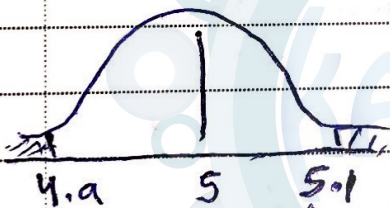
✓

Decisionfail to reject H_0 reject H_0

$P(\text{Type I error}) = \alpha = P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$

$\alpha \Rightarrow$ significance level.

Ex: Suppose the true mean is $\mu = 5$ inches, dist. of $\bar{x}$ is normally dist. $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$



~~$$\alpha = P(\bar{x} < 4.9) + P(\bar{x} > 5.1)$$~~

$$P\left(\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} < \frac{4.9 - 5}{\frac{\sigma}{\sqrt{n}}}\right) + P\left(\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{5.1 - \mu}{\frac{\sigma}{\sqrt{n}}}\right)$$

$$P(Z < z_1) + P(Z > z_2)$$

$\rightarrow$ if you have 1000 ^{samples} _{in 1000 trials}

you will make 57 wrong conclusions

57 samples will lead to wrong rejection of H_0

rejection.

acceptance region

No.

$$\downarrow \alpha \Rightarrow 1 \rightarrow n$$

acceptance region

β = Type II error = fail to reject H_0 when H_0 is false.

Ex: Randomly sample of 30 values from a normal distribution

$\sigma = 15$ & μ is unknown

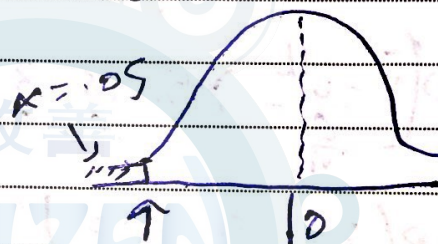
$$H_0: \mu = 10$$

نمونه گیری از جامعه
معمولی با 30
نمونه

at $\alpha = 0.05$

$$H_1: \mu < 10$$

Lower

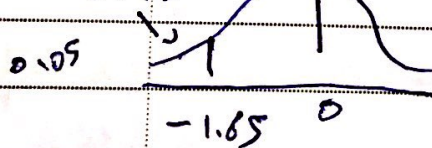
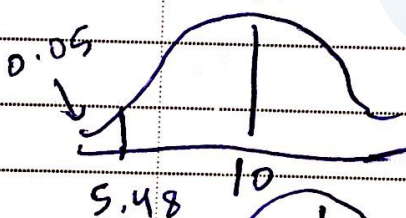


sample $\rightarrow \bar{X} \rightarrow Z \rightarrow$ compare $Z_{\alpha} = -1.65$

$Z < Z_{\alpha} \Rightarrow$ Reject H_0

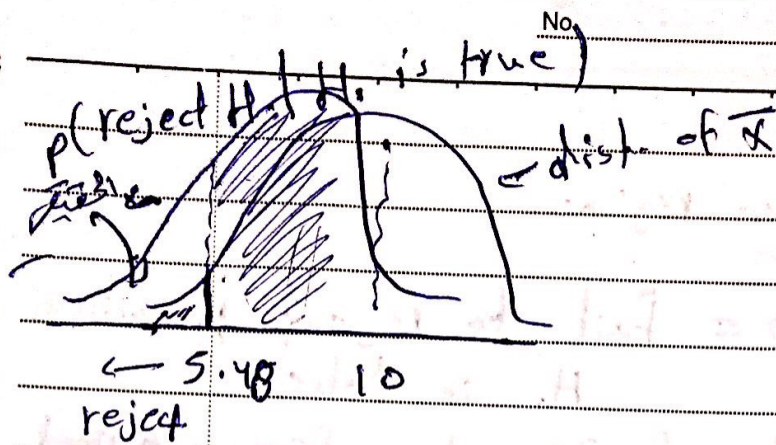
$$Z_{\alpha} = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} \Rightarrow \bar{X} = \mu_0 + Z_{\alpha} \frac{\sigma}{\sqrt{n}} = 10 + (-1.65) \frac{15}{\sqrt{30}}$$

$$= 5.48$$



$$\bar{X} < 5.48 \text{ reject}$$

$$\bar{X} \rightarrow Z < -1.65 \text{ reject}$$



What if $\mu=3$

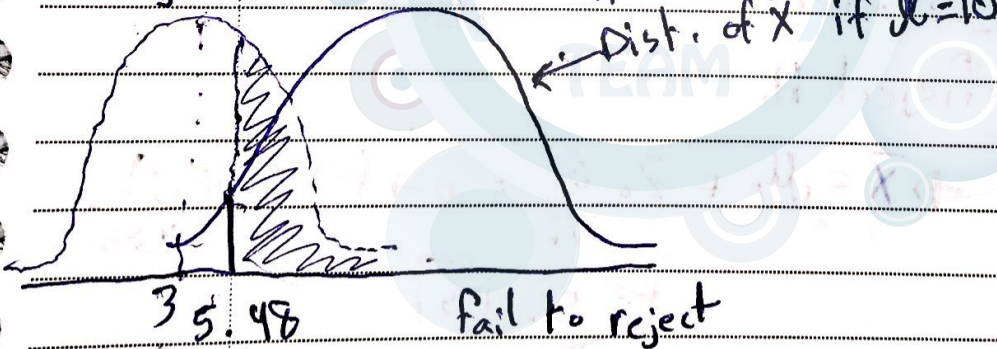
$p(\text{Type II error}) = p(\text{fail to reject } H_0 \text{ is false})$

$$= p(\bar{x} > 5.48 | \mu=3)$$

$$= p\left(\frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{5.48 - 3}{\frac{\sigma}{\sqrt{n}}}\right)$$

Dist. of $\bar{x}$
if $\mu=3$

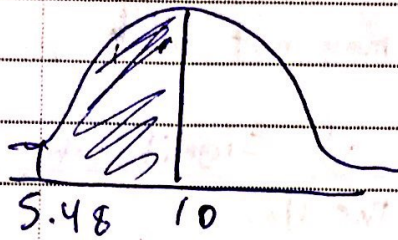
$$p(z > \frac{5.48 - 3}{\frac{\sigma}{\sqrt{n}}})$$



$$p(\bar{x} > 5.48 | \mu=3)$$

بمقدار μ منفرجه عن μ_0
يكون الخطر احتمال انه يوافق بالخطأ.

$$P(\text{fail to reject} | \mu = \mu_1)$$



$$P(\bar{X} > 5.48 | \mu = 3)$$

$$P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{5.48 - 3}{\frac{6}{\sqrt{n}}}\right)$$

 $1 - \beta$

power of the test

$$\sigma = 6$$

$$n = 30$$

$$\beta = P(Z > \frac{5.48 - 3}{\frac{6}{\sqrt{30}}}) = P(Z > 0.9056)$$

$$1 - P(Z \leq 0.9056)$$

$$= 0.171$$

$$\text{power of the test} = 1 - \beta = 0.829$$

(To reject when you have to reject.)

Fail to reject

Reject.

α & Type I error $\Rightarrow$ Risk of Consequences.

(Prob. of wrongly rejecting H_0)

We can control $\alpha \Rightarrow \therefore$ Rejection is a strong conclusion.

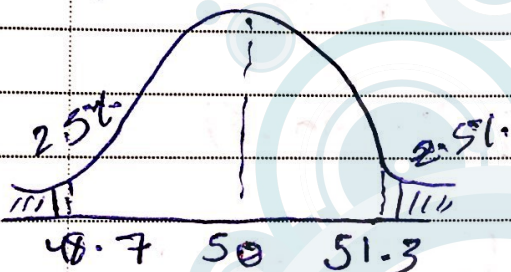
β depends on sample size $\Rightarrow$ Accepting is a weak conclusion.

P Value in H.T:-

H_0 may be rejected or may not, with specific α
 P-value is ^{the} smallest level of significance that
 would lead to rejection of the H_0 .

EX:- $H_0 \rightarrow \mu = 50$

$H_1 \rightarrow \mu \neq 50$ $n=16, \sigma=2.5, \bar{X}=51.3$



$$P = \text{prob. (above 51.3 + prob. below 48.7)}$$

$$= 1 - P(48.7 < \bar{X} < 51.3)$$

$$= 1 - P\left(\frac{48.7 - 50}{\frac{2.5}{\sqrt{16}}} < Z < \frac{51.3 - 50}{\frac{2.5}{\sqrt{16}}}\right) = 0.038$$

$$= 3.8\%$$

H_1 is correct.

$$\leq 0.05$$

test $\mu \neq 50$
 Mean = 51.3

P-value $< \alpha$
 Reject H_0

if H_0 is true $\Rightarrow$ prob. of obtaining a random
 sample whose mean is at least as far from

50 as 51.3 or 48.7 is 0.038

$\therefore$ it is rare to have a sample with mean = 51.3

if $\alpha = .05 \Rightarrow H_0$ will be rejected.

$P\text{-value} < \alpha \rightarrow \bar{x}$ will be in the critical region.

② p-value is the smallest

but indicates the risk of wrongly rejecting H_0 .

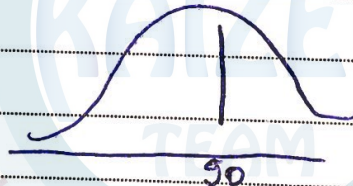
↳ when p-value is small go ahead and reject.

Two approaches for Hypothesis testing.

★ 1- α - value: fixed significant level.

2- p-value: calculated significant level.

H_0 is either true or false



α p-value $\rightarrow$ to indicate the risk of wrongly calculating.

[L u] C.I

—————

Risk of wrongly rejecting.

No.

EX.: p. 9.18-

Q H.: $\mu = 25$

H.: $\mu \neq 25$ ✓

Q H.: $G > 10$

H.: $G = 10$

No

Q H.: $\bar{X} = 50$

H.: $\bar{X} \neq 50$

No

$$\textcircled{1} H_0: p = .1$$

$$p = .5$$

$$N_0$$

$$\textcircled{2} H_0: s = 30$$

$$H_1: s > 30$$

$$N_0$$

$$9.68 -$$

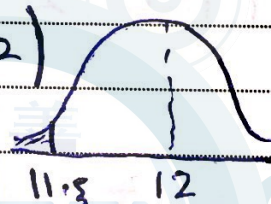
$$6 = .5 \quad \mu = 12$$

$$H_0: \mu = 12$$

$$n = 16$$

$$H_1: \mu < 12$$

$$a) \alpha = P(\bar{X} \leq 11.5 | \mu = 12)$$

$$= P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \leq \frac{11.5 - 12}{\frac{.5}{4}}\right)$$


$$= P(Z \leq -4) = 0$$

Risk
 . $\frac{120}{100} = 1.2$

$$\textcircled{b} \beta = P(\text{accept } H_0 \text{ when } \mu = 11.25)$$

$$\beta = P(\bar{X} > 11.5 | \mu = 11.25)$$

$$= P(Z > 2) = 1 - P(Z < 2)$$

$$= 1 - .97725$$

$$= .02275$$

$$\textcircled{c} \beta = P(\text{accept } H_1 | \mu = 11.5)$$

$$= P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} > \frac{11.5 - 11.5}{\frac{3}{\sqrt{16}}}\right)$$

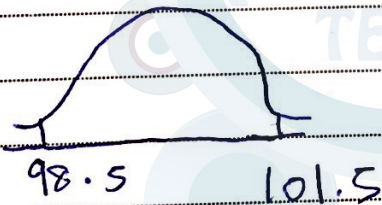
$$= P(Z > 0) = 0.5$$

+ 100000

9.11 % -

height $\sim N(100, 2^2)$ $H_0: \mu = 100$
 $H_1: \mu \neq 100$
 $n = 5$

$$\textcircled{a} P(\text{Type I error}) = \alpha = P(\bar{X} \leq 98.5) + P(\bar{X} \geq 101.5)$$



$$= P\left(Z \leq \frac{98.5 - 100}{\frac{2}{\sqrt{5}}}\right) + P\left(Z \geq \frac{101.5 - 100}{\frac{2}{\sqrt{5}}}\right)$$

$$= P(Z \leq -1.68) + P(Z \geq 1.68) = 0.09296$$

$$\textcircled{b} P(\text{Type II error} | \mu = 103)$$

$$\beta = P(\text{type II error}) \text{ if } \mu = 103$$

$$\beta = P(98.5 \leq \bar{X} \leq 101.5 | \mu = 103)$$

$$\beta = P(-5.03 \leq Z \leq -1.68)$$

$$= P(Z \leq -1.68) - P(Z \leq -5.03)$$

$$= 0.04648$$

$$\textcircled{c} P(\text{Type II error}) \text{ if } \mu = 105$$

$$\beta = 0.00005$$

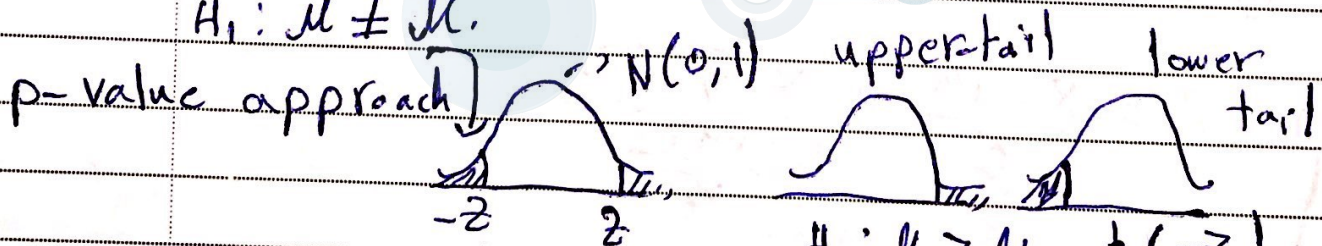
100 = mean of $\bar{X}$ is μ

* Tests on mean of a normal dist. ; σ^2 known.

$\bar{X}$: unbiased estimator of μ with $\frac{\sigma^2}{n}$

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0$$



$$H_1: \mu > \mu_0 \quad \phi(1-z)$$

$$P\text{-value} = 2[1 - \phi(z)] \quad 1 - \phi(z)$$

$$H_1: \mu < \mu_0$$

Z-test

H.W كورنيل - كورنيل
Ch. 8

mini I tab. mini I tab. على الارتفاع والوقت

Fixed significance level approach

$\alpha \Rightarrow$ critical region,

$$1 - \alpha = P\left(-z_{\frac{\alpha}{2}} \leq Z \leq z_{\frac{\alpha}{2}}\right)$$

sample $\rightarrow \bar{X} - z_{\frac{\alpha}{2}} \Rightarrow z. > z_{\frac{\alpha}{2}}$ } Reject.
test statistic. $z. < -z_{\frac{\alpha}{2}}$

$$\bar{X} \in \begin{cases} a = \mu_0 - z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \\ b = \mu_0 + z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \end{cases}$$

q. 368-

$$a) SE = \frac{\sigma}{\sqrt{n}} = \frac{1.475}{\sqrt{25}} = 0.295$$

$H_0: \mu = 35$

$H_1: \mu \neq 35$

(b) Two-sided.

$$z. = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}} = \frac{35.7 - 35}{\frac{1.475}{\sqrt{25}}} = \dots$$

$$p\text{-value} = 2[1 - \phi(z.)] = 2[1 - \phi(1)] = 0.0486$$

$$\alpha = 0.05$$

© 95% c.I two sided.

$$\bar{x} - z_{\alpha/2} \leq \mu \leq \bar{x} + z_{\alpha/2} \frac{s}{\sqrt{n}}$$

$$35.71 - (1.96) \frac{1.8}{\sqrt{25}} \leq \mu \leq \dots$$

$$35.0044 \leq \mu \leq 36.4156$$

(d) if $H_1: \mu > 35$

$$P\text{-value} = 1 - \Phi(z) = 1 - \Phi(1.9722) = 0.0243$$

9-39

$$H_0: \mu = 99 \quad H_1: \mu > 99 \quad \sigma = 2.5$$

a) $n = 12, \mu = 100.039, s = 2.365$

$$SE = \frac{s}{\sqrt{n}} = \frac{2.365}{\sqrt{12}} = 0.6827$$

b) one sided

c) if $H_1: \mu = 98, \mu > 98 \Rightarrow \dots$ p-value

$$z = \frac{100.039 - 98}{\frac{2.5}{\sqrt{12}}} = 2.8253 \Rightarrow \text{Reject or not}$$

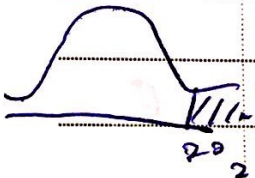
$$\alpha = 0.05$$

$$2.8253$$

95% lower c.I

Reject

$$p\text{-value} = 1 - \Phi(2.8253) = 0.002$$



(d)

$$\bar{x} - z_{\alpha/2} \frac{s}{\sqrt{n}} = 100.039$$

$$100.039$$

Q If $H_1: \mu \neq 99$
p-value.

α & p-value لا الفرق بين

Type II error & sample size

$H_0: \mu = \mu_0$

α : selected if H_0 is false

β : calculated

$\mu = \mu_0 + \delta \Rightarrow \delta = \mu - \mu_0$

$$\beta = P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2} | H_0 \text{ is true})$$

$$\beta = \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right)$$

من الطريقة القوية بس بتكون

بالطريقة
بتكون

$$\beta \approx \Phi\left(z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}\right) \Rightarrow z_{\beta} \approx z_{\alpha/2} - \frac{\delta\sqrt{n}}{\sigma}$$

$$n = \frac{(z_{\alpha} + z_{\beta})^2 \sigma^2}{\delta^2}$$

$$n = \frac{(z_{\alpha} + z_{\beta})^2 \sigma^2}{\delta^2} \quad \text{for the upper } \beta^{\text{th}} \text{ percentile. } z_{\beta}$$

$\beta \Rightarrow$ احتمال
يكون

$q = 42\%$

$n = 10, \bar{X} = 154.2^\circ \text{F}$ normally dist. with $\sigma = 1.5^\circ \text{F}$

Q 1- mean of milking pt. μ

2- $H_0: \mu = 155$

3- $H_1: \mu \neq 155$

4- $z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$

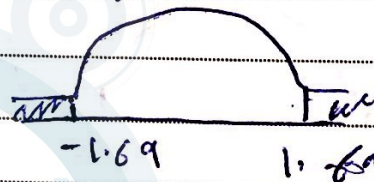
5. Reject H_0 if $Z < -Z_{\frac{\alpha}{2}}$ or $Z > Z_{\frac{\alpha}{2}}$ $\alpha = .01$
 $Z_{.005} = -2.58$

$$6. Z = \frac{154.2 - 155}{\frac{1.5}{\sqrt{10}}} = -1.69$$

7. $-1.69 > -2.58$ ∴ fail to reject.

there's not enough evidence to reject
 sufficient

b) p-value = ~~0.0455~~



$$2 * p(Z < -1.69) = 2 * 0.045514 = 0.091028$$

$$\begin{aligned} c) \beta &= \Phi\left(\frac{Z_{.005} - \frac{\delta\sqrt{n}}{6}}{1.5}\right) - \Phi\left(\frac{-Z_{.005} - \frac{\delta\sqrt{n}}{6}}{1.5}\right) \\ &= \Phi\left(\frac{2.58 - (155 - 150)\sqrt{10}}{1.5}\right) - \Phi\left(\frac{-2.58 - (155 - 150)\sqrt{10}}{1.5}\right) \\ &= \Phi(-7.96) - \Phi(-13.12) = 0 - 0 = 0 \end{aligned}$$

d) $h \Rightarrow \beta = 0.1$

$$\begin{aligned} h &= \left(\frac{Z_{\frac{\alpha}{2}} + Z_{\beta}}{\frac{\delta}{\sigma}}\right)^2 \sigma^2 = \left(\frac{Z_{.005} + Z_{.1}}{\frac{\delta}{\sigma}}\right)^2 \sigma^2 \\ &= 1.35 \approx 2 \end{aligned}$$

$\beta \uparrow$ $h \downarrow$

P. 7/7

OC curves

$d \Rightarrow$ كذا
وجود الكتاب

في
section 11

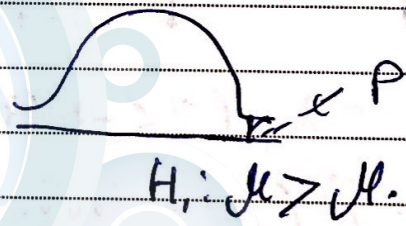
* Tests on the mean σ^2 unknown.

Test statistic $\Rightarrow$

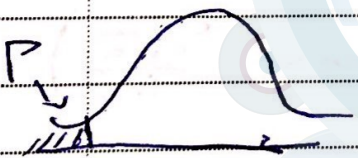
$$T = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$$



$H_1: \mu \neq \mu_0$



$H_1: \mu > \mu_0$



$H_1: \mu < \mu_0$

Ex. prob. 9-55

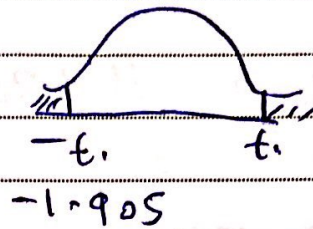
@ $n=10$

$dh = a$

$$SE = \frac{s}{\sqrt{n}} = 0.296 = \frac{s}{\sqrt{10}} \Rightarrow s = 0.9$$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{12.564 - 12}{\frac{0.9}{\sqrt{10}}} = 1.905$$

p-value =



$$t_1 = 1.905$$

$$\alpha = 0.05$$

$$1.833$$

$$\alpha = 0.025$$

$$2.262$$

$$0.025 < p\text{-value} < 0.05$$

$$0.03 < p\text{-value} < 0.05$$

if $\alpha = 0.05$

fail to reject $p\text{-value} > 0.05$

① 95% C.I on μ .

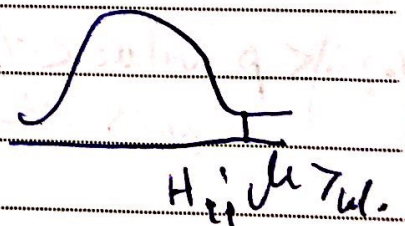
$$\bar{x} \pm t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

$$11.8945 < \mu < 13.2335$$

② If $H_0: \mu > 12$ $t_1 = 1.905$

$$= 1.833$$

Reject H_0 $t_1 > t_{\alpha/2, n-1}$ ✓



(f) $H_0: \mu = 11.5$, based on c.i. $\Rightarrow$ Rejected.

9-578 - $\mu = 300$, $\mu \neq 300$

a) ~~the~~ mean

2) $H_0: \mu = 300$

3) $H_1: \mu \neq 300$

4) $t = \frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}}$

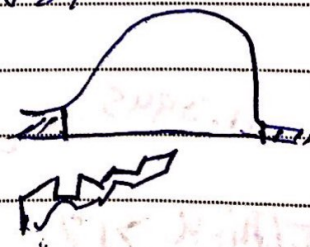
5) Reject H_0 if $|t| > t_{\frac{\alpha}{2}, n-1}$ where $\alpha = 0.05$ &
 $t_{\frac{\alpha}{2}, n-1} = 2.056$ for $n = 27$

6) $\bar{X} = 325.496$, $s = 198.786$, $n = 27$

$$t = \frac{325.496 - 300}{\frac{198.786}{\sqrt{27}}} = 0.6665$$

$$0.6665 < 2.056$$

$\Rightarrow$ fail to reject H_0



$0.23 < p\text{-value} < 0.4$
 0.6665

$$.5 \rightarrow .256$$

$$x \rightarrow .6665$$

$$.8 \rightarrow .684$$

19

$$\frac{.8 - .5}{x - .5} = \frac{0.684 - 0.256}{0.6665 - 0.256}$$

95% two-sided c.i

$$= \frac{0.428}{0.4105}$$

$$\bar{x} - t_{\alpha/2, n-1} \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

$$= .3$$

$$x - .5$$

$$246.8409 < \mu < 404.1511$$

$$= .3 + .4105$$

$$.428$$

section

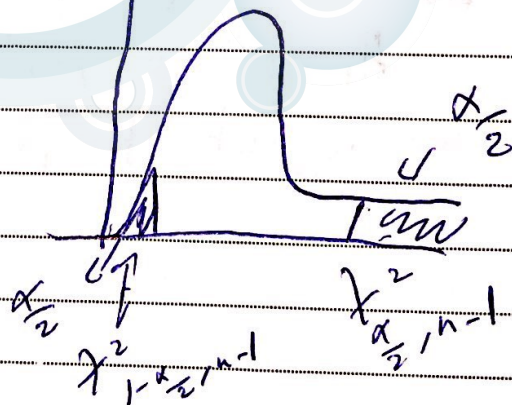
part 1

Tests on σ^2

$$H_0: \sigma^2 = 6.2$$

$$H_1: \sigma^2 \neq 6.2$$

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$



$$\chi^2 < \chi^2_{1-\alpha/2, n-1}$$

Reject

$$\chi^2 > \chi^2_{\alpha/2, n-1}$$

Reject

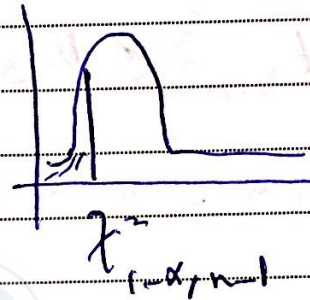
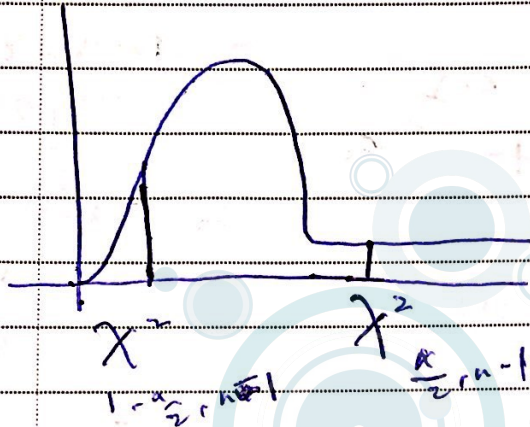
$$\sigma^2 \neq 6.2$$

24

$$H_0: \sigma^2 = 6.2$$

$$\sigma^2 \neq 6.2$$

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2}$$



(انتقل الى الرساتين (1) و (2))

$$\sigma^2 < 6.2$$

$$\alpha = 77\%$$

$$n = 17, S = 0.09$$

⊙ A) parameter of int. is 6

$$H_0: \sigma^2 = (0.75)^2$$

$$H_1: \sigma^2 > 0.75^2$$

we reject H_0 if $\chi^2 > \chi^2_{\alpha, n-1}$

$$\chi^2_{0.05, 16} = 26.3$$

$$\chi^2 = \frac{16 * 0.09^2}{0.75^2}$$

$$= 0.23$$

$\Rightarrow$ fail to reject

$$0.23 < 26.3$$

B) How to ~~prove~~ answer the same question in a different way?

95% one sided C.I

$$\frac{(n-1)S^2}{\chi^2_{\alpha, n-1}} < \sigma^2 \quad \frac{16 \times 0.09^2}{26.3} < \sigma^2$$

$$H_0: \sigma^2 \leq 0.75^2$$

$$H_1: \sigma^2 > 0.75^2$$

$$0.07 < \sigma^2$$

$$0.07 < 0.6$$

Fail to reject.



لا يمكن أن نقول
بأنه

Tests on a pop. proportion & -
sample size
of "n"

X: random variable of interest.

$$\hat{p} = \frac{X}{n}$$

Parameters of a binomial distribution n, p
sampling dist. of $\hat{p}$.

$$np > 5 \quad \text{normal}$$

$$n(1-p) > 5$$

and n it's large enough
dist. $\rightarrow$ Normal

انتظمت هجوت

$$H_0: p = p_0$$

$$H_1: p \neq p_0$$

$$X \sim N(np_0, np_0(1-p_0))$$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

$$p\text{-value} = 2[1 - \Phi(Z)] \quad p = p_0$$

$$p = 1 - \Phi(Z) \quad p > p_0$$

$$p = \Phi(Z) \quad p < p_0$$

Type II error (β) $p \neq p_0$

$$\beta = \Phi\left[\frac{p_0 - p + Z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right] - \Phi\left[\frac{p_0 - p - Z_{\alpha/2} \sqrt{\frac{p_0(1-p_0)}{n}}}{\sqrt{\frac{p(1-p)}{n}}}\right]$$

$$\text{if } p < p_0$$

$$\beta = 1 - \Phi\left[Z_{\alpha/2} \frac{\text{الجزء الثاني}}{\text{بها}}\right]$$

$$p > p_0$$

$$\beta = \Phi\left[Z_{\alpha/2} \frac{\text{الجزء الثالث}}{\text{بها}}\right]$$

Sin test $\Rightarrow$ non parametric test.
 جہت سے، sec. $\hat{f}$

Testing for goodness of fit:-

To know the underlying prob. dist.
 based on chi square distribution.

n observations $\Rightarrow$ Histogram with k bins.

$$\chi^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i}$$

(مربع)

Hypothesis: H_0 : dist. is Poisson
Normal

H_1 : " " is not

if $\chi^2 > \chi^2_{k-p-1}$

ok! χ^2_{k-p-1} is the critical value
 if $\chi^2 > \chi^2_{k-p-1}$ then reject H_0
 if $\chi^2 \leq \chi^2_{k-p-1}$ then do not reject H_0

p : # of parameter estimated from sample
 k : # of bins

reject to

Type 2 error

① Using OC curve

$$d = \frac{(\mu_1 - \mu_2) - \Delta_0}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

② Using formula

$$\beta = \Phi \left[\frac{z_{\alpha/2} - \frac{(\Delta - \Delta_0)}{\sqrt{\sigma_1^2 + \sigma_2^2}}}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right] - \Phi \left[\frac{-z_{\alpha/2} - \frac{(\Delta - \Delta_0)}{\sqrt{\sigma_1^2 + \sigma_2^2}}}{\sqrt{\sigma_1^2 + \sigma_2^2}} \right]$$

$$h = \frac{(z_{\alpha/2} + z_{\beta})^2 (\sigma_1^2 + \sigma_2^2)}{(\Delta - \Delta_0)}$$

Q-128-

no. of defect in printed circuit board
is hypo to follow poisson dist.!

n = 60 circuit board

of def. observed from

h = rate of mean defect.	0	32
↓	1	15
estimation from sample.	2	9
	3	4

$$\bar{h} = \frac{32 \times 0 + 15 \times 1 + 9 \times 2 + 3 \times 3}{60}$$

$$\bar{h} = 75 \rightarrow \text{avg}$$

To calculate expected # of defects we need p_i for each category.

$$f(x) = \frac{e^{-\lambda} \lambda^x}{x!} \quad X: \# \text{ of defects}$$

$$p_1 = p(X=0) = \frac{e^{-.75} (.75)^0}{0!} = .472$$

$$p_2 = p(X=1) = .354$$

$$p_3 = p(X=2) = .133$$

$$p_4 = p(X \geq 3) = 1 - p(X < 3) = 1 - p(p_1 + p_2 + p_3)$$

$$= 0.041$$

Expected E

60% $p_1 = 28.32$	0	32
60% $p_2 = 21.24$	1	15
60% $p_3 = 7.98$	2	4
60% $p_4 = 2.45$	3	4
60% $p_{\text{total}} = 10.00$		

$$\chi^2 \rightarrow E \geq 5$$

یا 5 سے زیادہ

یا 5 سے زیادہ

یا 5 سے زیادہ

Contingency table test

Two variables (classifications/categories/criteria)

^ we want to test if they are independent.

H_0 : The two categories are independent.

H_1 : " " " " " dependent.

n elements of sample from pop.:

Two methods of classification.

1st method has r levels.

2nd " " c " ".

O_{ij} : observed frequency in row i & column j .

method 2

	1	2	...	c
method 1	1 O_{11}	O_{12}	...	O_{1c}
	2 O_{21}	O_{22}	...	O_{2c}
	$\vdots$	$\vdots$		$\vdots$
	r O_{r1}	...	...	O_{rc}

P_{ij} : prob. that a randomly selected element falls in the ij th cell.

$$P_{ij} = U_i V_j$$

U_i : prob. an element falls in row i .

V_j : " " " " " column j .

$$U_i = \frac{1}{n} \sum_{j=1}^c O_{ij}$$

$$V_j = \frac{1}{n} \sum_{i=1}^r O_{ij}$$

Expected freq. in each cell is $E_{ij} = np_{ij}$

$$\chi^2_o = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{1}{n} \sum_{j=1}^c O_{.j}^2 \sum_{i=1}^r O_{ij}$$

$\chi^2_o > \chi^2_{\alpha, df}$ reject independence.
 $\downarrow$
 dependent $(c-1)(r-1)$

non parametric tests

If you know dist. or data interval or ratio. } parametric procedures
 $\sum \frac{1}{n} \chi^2$

* If dist. is unknown or ordinal or Nominal.

scale 1-5 HP, Acel } Non parametric
 def. باقی
 معین لا یقین

یا: ان معین لا یقین

1. Sign test $\rightarrow$ median
 2. welcoxon - Signal rank test
 3. Kruskal wallis $\Rightarrow$ Anova
- Parametric

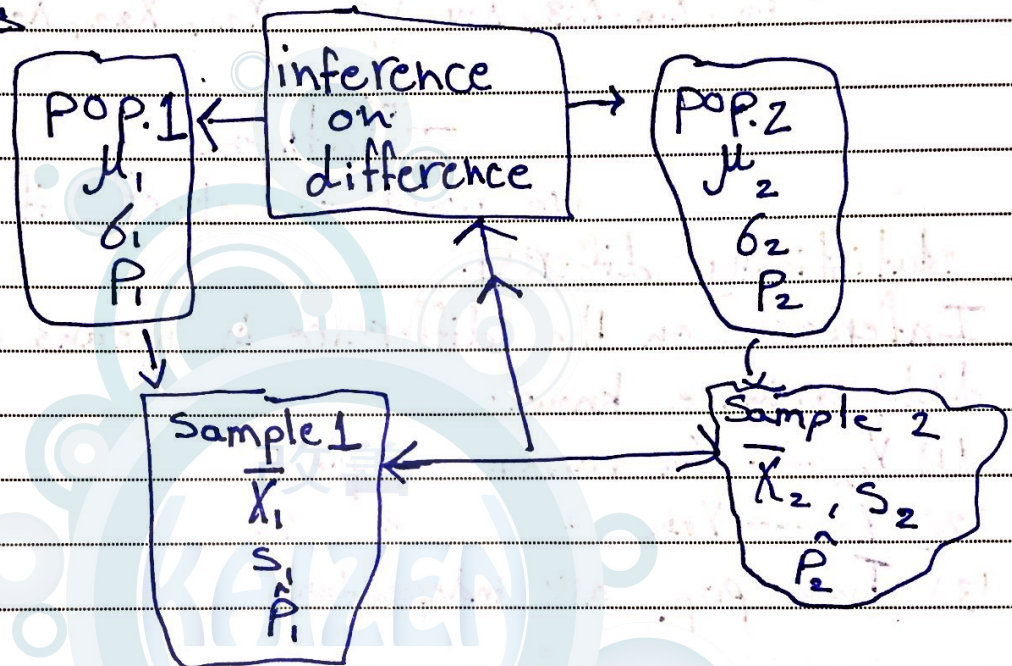
Statistical inferences for 2 samples.

Goal: To compare parameter (μ, σ^2)

of two different populations.

Quiz
ch 10

~~inference~~



producing juice

↳ To test hypo that people like it.

sample if we don't develop another recipe

بطلع $\bar{x}_1$ $H_0: \mu_1 - \mu_2 = 0$

$H_1: \mu_1 - \mu_2 > 0$

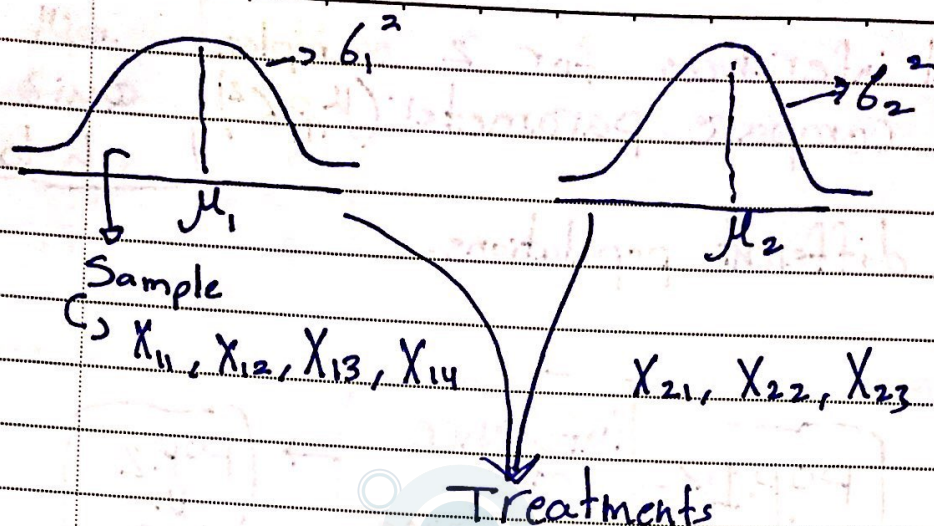
بعد، صحتنا تست

حسب القيمة الأكبر

بالسؤال.

↳ New sample

$\bar{x}_2$



we want to check if they are statistically different.

Inference on the difference in means at 2 normal dist., variance known.

- (1) $X_{11}, \dots, X_{1n}$ random sample 1
- (2) $X_{21}, \dots, X_{2n}$ " " 2
- (3) Two pop. are independent

expected: $E(\bar{X}_1 - \bar{X}_2) = E(\bar{X}_1) - E(\bar{X}_2) = \mu_1 - \mu_2$

Variance: $V(\bar{X}_1) + V(\bar{X}_2) = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

Two sided

$$P\text{-Value} = 2[1 - \Phi(Z)]$$

$$H_0: \mu_1 - \mu_2 = D_0$$

$$H_1: \mu_1 - \mu_2 \neq D_0$$

$$P\text{-Value} \rightarrow Z \text{ critical value}$$

الخاتمة

C.I Difference between two means.

σ^2 known.

$$\sigma_1 = \sigma_2$$

$$\sigma \neq \sigma_2$$

$$P(-z_{\frac{\alpha}{2}} < Z < z_{\frac{\alpha}{2}}) = 1 - \alpha$$

$$P\left(-z_{\frac{\alpha}{2}} < \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} < z_{\frac{\alpha}{2}}\right) = 1 - \alpha$$

مباين
هون

$$\bar{X}_1 - \bar{X}_2 - z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + z_{\frac{\alpha}{2}} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$n_1 \geq 40$$

$$n_2 \geq 40$$

$$s_1^2 \rightarrow \sigma_1^2$$

$$s_2^2 \rightarrow \sigma_2^2$$

To obtain error E for estimating difference

$$(\mu_1 - \mu_2)$$

$$n_1 = n_2 = n$$

$$n = \left(\frac{z_{\frac{\alpha}{2}}}{E} \right)^2 (\sigma_1^2 + \sigma_2^2)$$

One sided C.I

$$\mu_1 - \mu_2 < \bar{X}_1 - \bar{X}_2 + z_{\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\bar{X}_1 - \bar{X}_2 - z_{\alpha} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2$$

H.T

~~on~~ on μ_1, μ_2 variances unknown.

$$\text{Case, } \sigma_1^2 = \sigma_2^2 = \sigma^2$$

$$H_0: \mu_1 - \mu_2 = D_0$$

$$H_1: \mu_1 - \mu_2 \neq D_0$$

$$E(\bar{X}_1 - \bar{X}_2) = \mu_1 - \mu_2$$

$$V(\bar{X}_1 - \bar{X}_2) = \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2} = \sigma^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

pooled estimator of σ^2 is S_p^2

(variance of variances of all pop.)

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$T_0 = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{S_p \left(\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \right)} \quad \text{with } n_1 + n_2 - 2 \text{ df.}$$

$$H_1: \mu_1 - \mu_2 \neq D_0 \Rightarrow \text{p-value} = \text{prob. above } |t_0|$$

reject if $t_0 > t_{\frac{\alpha}{2}, n_1 + n_2 - 2}$ or below $-t_0$

$$\text{or } t_0 < -t_{\frac{\alpha}{2}, n_1 + n_2 - 2}$$

$$H_1: \mu_1 - \mu_2 > D_0 \quad \text{p-value} = \text{prob. above } t_0$$

reject if

$$t_0 > t_{\alpha, n_1 + n_2 - 2}$$

$H_1: \mu_1 - \mu_2 < \Delta_0$ p-value = prob. below t_0

reject if
 $t_0 < -t_{\frac{\alpha}{2}, n_1+n_2-2}$

T-test

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

$$T_0^* = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad r = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\left(\frac{s_1^2}{n_1}\right)^2 + \left(\frac{s_2^2}{n_2}\right)^2}$$

Meaning of degrees of freedom. Round down,
 on $(\mu_1 - \mu_2)$

C.I. $\hat{\sigma}^2$ unknown.
 Variances

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

$$\bar{X}_1 - \bar{X}_2 - t_{\frac{\alpha}{2}, n_1+n_2-2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{X}_1 - \bar{X}_2 + t_{\frac{\alpha}{2}, n_1+n_2-2} \cdot s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

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

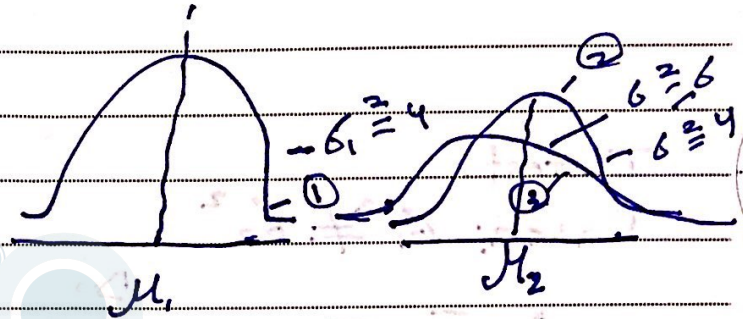
$$\bar{X}_1 - \bar{X}_2 - t_{\frac{\alpha}{2}, r} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \bar{X}_1 - \bar{X}_2 + t_{\frac{\alpha}{2}, r} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

$r \text{ df} \Rightarrow$ before.

Inferences on the variances of two normal dist.

$$H_0: \sigma_1^2 = \sigma_2^2$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$



$$F = \frac{\frac{W}{u}}{\frac{Y}{v}}$$

W & Y $\Rightarrow$ Chi-squared quantities.

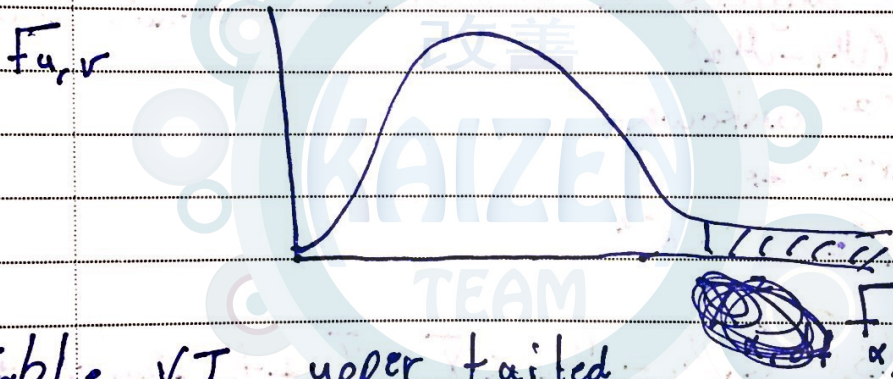


Table V.I upper tailed prob.

$$P(F > f_{\alpha, u, v}) = \int f(x) dx$$

$$f_{1-\alpha, u, v} = \frac{1}{f_{\alpha, u, v}}$$

$$f_{.95, 5, 10} = \frac{1}{f_{.05, 5, 10}}$$

Hypothesis tests on the ratio of two variances.

$$F = \frac{\frac{S_1^2}{\sigma_1^2}}{\frac{S_2^2}{\sigma_2^2}} \quad \text{with } \begin{matrix} n_1 - 1 \\ \text{df} \\ n_2 - 1 \end{matrix}$$

$$H_0: \sigma_1^2 = \sigma_2^2 \quad F_0 = \frac{S_1^2}{S_2^2}$$

$$H_1: \sigma_1^2 \neq \sigma_2^2$$

$$f_0 > f_{\frac{\alpha}{2}, n_1-1, n_2-1} \quad \text{or} \quad f_0 < f_{1-\frac{\alpha}{2}, n_1-1, n_2-1}$$

$$\text{if } H_1: \sigma_1^2 > \sigma_2^2 \quad f_0 > f_{\alpha, n_1-1, n_2-1}$$

$$H_1: \sigma_1^2 < \sigma_2^2 \quad f_0 < f_{1-\alpha, n_1-1, n_2-1}$$

C.I on the ratio of two variances.

$$\frac{S_1^2}{S_2^2} f_{1-\frac{\alpha}{2}, n_2-1, n_1-1} < \frac{\sigma_1^2}{\sigma_2^2} < \frac{S_1^2}{S_2^2} f_{\frac{\alpha}{2}, n_2-1, n_1-1}$$

inferences on two pop. proportions:

Two binomial populations
parameters: P_1, P_2

Two independent samples $n_1 \leftarrow X_1$
 $P_1 = \frac{X_1}{n_1} \quad P_2 = \frac{X_2}{n_2} \quad n_2 \leftarrow X_2$

$$Z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$\hat{p} = \frac{X_1 + X_2}{n_1 + n_2}$$

$$H_0: p_1 = p_2$$

$$H_1: p_1 \neq p_2$$

$$H_1: p_1 > p_2$$

$$H_1: p_1 < p_2$$

$$2[1 - \Phi(|z_0|)]$$

p-value

$$1 - \Phi(z_0)$$

$$\Phi(z_0)$$

Case 5: ratios
difference

C.I.:

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

Case 5: Why do we use Ratios when testing Hypothesis on two populations variances.

non-parametric procedures:

Tala Alrameli

man-wetni u-test.

No. _____

2 pop. not normal

1-man - whitly u-Test.

Ch. 11: simple linear regression.

* Analysis to study of the relationship between two or more variables.

↑
[chsz]

Two model types

1- Deterministic : $F = ma$ distance = $d_0 + Vt$

Once I determined the mass & acceleration,
I can calculate the $\downarrow$ exact value of F .

(predict force) (predict distance.)

2- Non deterministic: Fuel usage of a vehicle "y" is related to vehicle weight X.

$\Rightarrow$ Because y (value of the response of interest) can't be predicted perfectly from knowledge of the corresponding X .

One independent variable X $\longrightarrow$ response variable y .
or
predictor or regressor

Deterministic
(Theory.)

Non deterministic/
experimental
empirical.

example in the text book.

purity / Hydration level (X).

mean / expected value
(X) is

$$E(y|X) = \mu_{y|X} = \beta_0 + \beta_1 X$$

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

SLR

Intercept

regression coeff.

mean of $\varepsilon = 0$

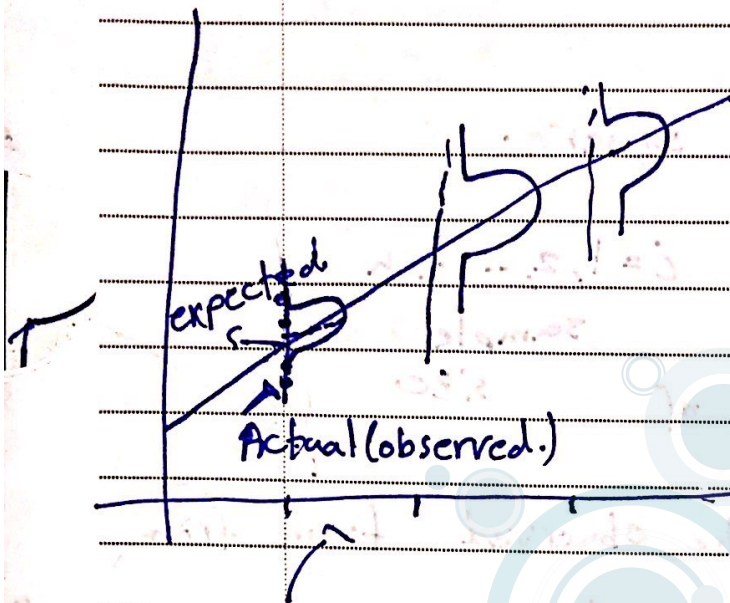
variance of $\varepsilon = \sigma^2$

$$E(y|X) = E(\beta_0 + \beta_1 X + \varepsilon) = \beta_0 + \beta_1 X$$

$$V(y|X) = V(\beta_0 + \beta_1 X + \varepsilon) = 0 + \beta_1^2 V(\varepsilon) + \sigma^2$$

$$= \sigma^2$$

$\mu_{y|X} = \beta_0 + \beta_1 X$ is the line of mean values.



Distribution of y -values at each x with the same variance for each x .

In other words, there is a distribution of y -values at each value of x $\Rightarrow$ Normal distribution with mean $\beta_0 + \beta_1 x$ & Variance of σ^2 .

~~Example~~ expected
error) σ^2

CH 10 نویز و خطا

$$M_{y|x} = 11.25$$

$\hookrightarrow y$ has a mean value of 93.75 & Variance of 2.

*To estimate β_0 & β_1 .

Gauss: method of least square.

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad i = 1, 2, \dots, n$$

sample size

Sum of the squares of i.

the deviations of the observed from from the regression line. $L = \sum_{i=1}^n \epsilon_i^2 = \sum (Y_i - \beta_0 - \beta_1 X_i)^2$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n Y_i X_i - \frac{\sum_{i=1}^n Y_i \sum_{i=1}^n X_i}{n}}{\sum_{i=1}^n X_i^2 - \frac{(\sum_{i=1}^n X_i)^2}{n}} S_{xy} / S_{xx}$$

$$\bar{Y} = \frac{1}{n} \sum Y_i \quad \bar{X} = \frac{1}{n} \sum X_i$$

Fitted or estimated regression line is

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

$$e_i = Y_i - \hat{Y}_i \text{ residual.}$$



$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

Example

oxygen purity in terms of carbon level.

$$n=20 \quad \sum_{i=1}^{20} X_i = 23.92 \quad \sum y_i = 1,843.21$$

$$\sum y_i^2 = 170,044.5321 \quad \sum X_i^2 = 29.289$$

$$\hat{\beta}_1 = \frac{S_{xy}}{S_{xx}} = \frac{\sum (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2} \quad \sum X_i y_i = 2,214.6566$$

$$S_{xx} = \sum X^2 - \frac{(\sum X)^2}{n} = 29.2892 - \frac{23.92^2}{20}$$

کلی نقل ہو گیا!

Estimating σ^2 ← "Variance of the error ϵ_i "

$\epsilon_i = y_i - \hat{y}_i$ → Sum of squares of the residuals
→ Called error sum of squares.

$$SS_E = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$\sum (SS_E) = (n-2) \sigma^2$$

$$\Rightarrow \sigma^2 = \frac{SS_E}{n-2}$$

$$SS_E = SS_T - \hat{\beta}_1 S_{xy}$$

$$SS_T = SS_R + SS_E$$

↑
The total Reg.

sum of

sq.

انضامی ہون گت شغلہ ناسفہ .

$$SS_T = \sum (y_i - \bar{y})^2 = \sum y_i^2 - n\bar{y}^2$$

Hypothesis testing SLR.

Total variability

$$SS_T = \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{..})^2$$

$$= n \sum_{i=1}^a (\bar{y}_{i.} - \bar{y}_{..})^2 + \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2$$

between treatments

within treatments

$$SS_T = SS_{Treat} + SS_E$$

$$\sum (SS_{Treat.}) = (a-1)\sigma^2 + n \sum_{i=1}^a \bar{y}_{i.}^2$$

$$\sum (SS_E) = a(n-1)\sigma^2$$

تالا الزماني

SS_T

N-1 (a-1)

SS_{Treat}

a-1

SS_{Err}

a(n-1)

between treatments

within treatments

Mean square

$$MS_{Treat} = \frac{SS_{Treat.}}{(a-1)}$$

$$MS_E = \frac{SS_E}{[a(n-1)]} = \frac{a(n-1)\sigma^2}{a(n-1)} = \sigma^2$$

if H_0 is true: $MS_{Treat.} = \frac{(a-1)\sigma^2 + n\sum T_i^2}{(a-1)} = \sigma^2$

if H_0 is false: $MS_{Treat.} > \sigma^2$

$\frac{MS_{Treat.}}{MS_E} = \frac{\sigma^2}{\sigma^2} = 1$

$\frac{MS_{Treat.}}{MS_E} > 1$

MS_E

$F_0 = \frac{SS_{Treat.}}{(a-1)} \div \frac{SS_E}{[a(n-1)]}$

Reject H_0 if $F_0 > F_{\alpha, a-1, a(n-1)}$

Shortcuts

$SS_T = \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{y_{..}^2}{N}$

$SS_{Treat.} = \sum_{i=1}^a y_{i.}^2 - \frac{y_{..}^2}{N}$

$SS_E = SS_T - SS_{Treat.}$

Some of sum of df
variance squares mean

Treatment, SS Treat: $a-1$ MS

Error SS E: $a(n-1)$ MS E

Total SS T: $an-1$

Ex. 1: Hardwood concentration:

Hypothesis: H_0 Conc. = μ

$H_1: T_1 = T_2 = T_3 = T_4 = 0$

Sol:

$$SS_T = \sum_{i=1}^n y_{ij}^2 - \frac{y_{..}^2}{n} = 7^2 + 8^2 + \dots + 20^2$$

$$= \left(\frac{383^2}{24} \right) = 512.96$$

$$SS_{\text{Treat.}} = \sum_{i=1}^4 y_i^2 - \frac{y_{..}^2}{N} = 60^2 + 94^2$$



$$102^2 + 127^2 - \frac{383^2}{24} = 382.79$$

$$SS_E = 512.96 - 383.79 = 130.17$$

Table

• ~~190~~

ANNOVA Table

Source of variance	Sum of Squares	df	mean square	F _{0.05}
H w concentration	382.79	3	127.6	19.6
error	130.17	20	6.51	
Total	512.96	23		

$$f_{0.05, 3, 20} = 4.94$$

∴ **Reject.**