Var $V(\bar{X}) = \frac{\sigma}{n} = \frac{\text{Variance}}{\text{sample size}}$

Standard error $= \sqrt{V(\bar{x})}$

$\sigma^2 = V(\bar{x}) = \frac{\Sigma(\bar{x}-\bar{\bar{x}})^2}{\Sigma f}$

Q $0, 3, 6, 9, 12$

$N = 5$ $n = 2$.

(0, 0)	0	(3, 0)	1.5	(6, 0)	3
(0, 3)	1.5	(3, 3)	3	(6, 3)	4.5
(0, 6)	3	(3, 6)	4.5	(6, 6)	6
(0, 9)	4.5	(3, 9)	6	(6, 9)	7.5
(0, 12)	6	(3, 12)	7.5	(6, 12)	9

(9, 0)	4.5	(12, 0)	6
(9, 3)	6	(12, 3)	7.5
(9, 9)	9	(12, 9)	10.5
(9, 6)	7.5	(12, 6)	9
(9, 12)	10.5	(12, 12)	12

$\bar{X}$	f	$(f)(\bar{X})$	$(\bar{X}-\bar{\bar{X}})^2$	$f(\bar{X}-\bar{\bar{X}})^2$
0	1	0	36	36
1.5	2	3	20.25	40.5
3	3	9	9	27
4.5	4	18	2.25	9
6	5	30	0	0
7.5	4	30	2.25	9
9	3	27	9	27
10.5	2	21	20.25	40.5
12	1	12	36	36
total	25		219	225

$\bar{X} = \frac{f(\bar{X})}{25} = 6$; $\sigma(\bar{x})^2 = \frac{225}{25} = 9$

Standard error $= \sqrt{9} = 3$

Date ______ 20

w/o Replacement

$$V(\bar{x}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$$

$\bar{\bar{x}} = \frac{f(\bar{x})}{n} = \mu_{\bar{x}}$

With Replacement

$\bar{x}$	f	$f\bar{x}$

$$\sigma^2_{\bar{x}} = \frac{\sum f(\bar{x}-\bar{\bar{x}})^2}{\sum f} \longrightarrow \text{Sampling}$$

$$\sigma^2 = \frac{\sum(x-\mu)^2}{N} \longrightarrow \text{Population}$$

Standard error $= \sqrt{V(\bar{x})}$

$$V(\bar{x}) = \frac{\sigma^2}{n} = \frac{225}{25} \quad \frac{18}{2} = 9$$

=

Date ______ 20

Census → ~~Every~~ Study of every unit

Sampling → Process of selecting a sample

→ Survey

Probability Sampling

→ ~~Representative~~ Cluster — group

→ Symmetric → Interval

→ Stratified - sampling

→ subgroup

Non-Probability Sampling

→ Convenient sampling

→ Quota sampling

T distribution.

CI of μ when σ is unknown and dist is normal

$$\bar{x} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}.$$

$n =$ no. of sample

$s =$ sample ~~standard~~ SD

$\sigma =$ population S.D

$df = v =$ degree of freedom $= n-1$

$df =$ no. of independent values

Exampel:

$n = 10$

$df = 10 - 1 = 9$

$\bar{x} = 71$

$s = 6.78$

Example 2

$174.3 \pm \underline{46.5} \mid ~~74.5~~$

Date ________ 20___

$\bar{X}$ = mean

σ = population variance

α = confidence value

n = sample size

z = sample value

Example

$n = 50$

$\bar{x} = 54$

$\sigma = 6.0$ days

$\alpha = 95\%$

$z = \frac{0.05}{2}$ $z_{95\%} = 1.96$ $\quad = 54 \pm \frac{(1.96)(6)}{\sqrt{50}}$

$\frac{0.10}{2}$ $z_{90\%} = 1.65$ $\quad = 54 \pm \frac{(1.65)(6)}{\sqrt{50}}$

$\frac{0.01}{2}$ $z_{99\%} = 2.58$ $\quad = 54 \pm \frac{(2.58)(6)}{\sqrt{50}}$

- Estimate ± (tabulated value)(Standard error)

If σ is unknown then $\bar{X} \pm z_{\frac{\alpha}{2}} \left(\frac{S}{\sqrt{n}}\right)$

Estimation of Hypothesis testing.

Date ______ 20 __

- Process of estimating the ratio of parameters

Estimation

point estimate

$\bar{X} \rightarrow$ estimate

μ - parameter

interval estimate

$a < \mu < b$

if interval estimation is in confidential percentage then it is interval

. 90% or 99% or 95%

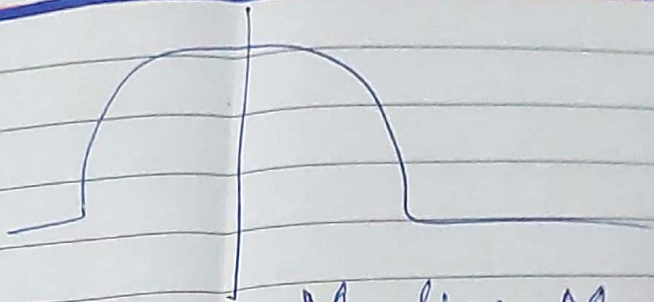
. We assume that σ is known

. population is normal

$$\bar{X} \pm z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

Normal distributions

Date ____ 20__



Mean = Median = Mode
Skewness = 0
Kurtosis = 3

- PDF $= f(x) = \cancel{\dfrac{e^{(x-\mu)^2 / \sigma\sqrt{2\pi}}}{2\pi}} = \dfrac{e^{-(x-\mu)^2/2\sigma^2}}{\sigma 2\pi}$

  $-\infty < x < \infty$

-

- two parameters $= \mu$ & σ

Standard Normal distribution

$$f(y) = \frac{e^{-z^2/2}}{\sqrt{2\pi}}$$

$$z = \frac{x-\mu}{\sigma} \sim (\mu = 0, \sigma = 1)$$

=

t test

Date ______ 20

$u_0 = 36.7$

$\mu_1 = 36.7$

$\alpha = 0.05$

$n = 15$ $\bar{X} = 40.6$ $s = 6$

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

$$t = \frac{(40.6 - 36.7)}{\frac{6}{\sqrt{15}}} = 2.64653$$

$df = n - 1$
$= 15 - 1$
$= 14$

$t_{\frac{\alpha}{2}} = t_{0.05, 14} = 1.761$

$$X \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$$

$$40.6 \pm (1.761)\left(\frac{6}{\sqrt{15}}\right.$$

Two tale T test

Date ______ 20__

Independent Sample

① $\mu_1 \neq \mu_2$ ~~$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n}}}$~~

let

Ho $\mu_1 = \mu_2$ or $\mu_1 - \mu_2 = 0$

$H_1 = \mu_1 \neq \mu_2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

df = ~~(n-1)~~ smaller ($n_1 - 1$ or $n_2 - 1$)

② ~~$\mu_1 = \mu_2$~~

$$(\mu_1 - \mu_2) \pm t_{\frac{\alpha}{2}} \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

ex 1:

$\bar{X}_1 = 191$

$\bar{X}_2 = 199$

$S_1 = 31$ $S_2 = 12$

$n_1 = 8$ $n_2 = 10$

df = 8 - 1
= 7

$$t = \frac{(191 - 199) - (31 - 12)}{\sqrt{\frac{31^2}{8} + \frac{12^2}{10}}} = -1.93$$

Since $t < 2$ SD Ho is accepted

Date ______ 20__

$\mu_1 = \mu_2$

$$t = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n-1)s^2 + (n-1)s^2}{n_1 + n_2 - 2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$= \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1 + n_2 - 2}}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$\boxed{\text{F-test}}$

$H_0 = \sigma_0 = \sigma_2$

$H_1 = \sigma_1 \neq \sigma_2$

$\alpha = 0.05$

$$F = \frac{\text{larger}^2}{\text{smaller}^2} = \frac{38^2}{12^2} = \cancel{0.0927}\ 10.027$$

if $F \geq 4$ then Ho rejected

if $f < 4$ then Ho accepted

df = smaller of $(n_1 - 1$ or $n_2 - 1)$

Dependent Sample

Date ______ 20__

$H_0 = \mu_3 = \mu_7 - \mu_0 = 0$

$H_1 = \mu_7 > \mu_3$

$\alpha = 0.05$

X_1 (3)	X_2 (7)	$d = (X_1 - X_2)$
11.42	16.69	-5.27
8.41	9.44	-1.02
3.95	6.53	-2.55
7.37	5.58	1.79
2.28	2.92	-0.64
1.10	1.88	-0.78
1	1.78	-0.78
0.99	1.5	-0.60
1.35	1.22	0.13

$$t = \frac{d - \mu_0}{\frac{S_d}{\sqrt{n}}}$$

$df = n-1$
$= 8-1$
$= 7$

$$= \frac{(-1.08) - 0}{\frac{(1.525)}{\sqrt{8}}} = 1.775$$

Analysis of Variance (ANOVA) Date ____ 20__

• If there is analysis of two or or samples' mean.

$H_0 = \mu_1 = \mu_2 \cdots \mu_k$

H_1 = one mean is different from others.

$\alpha = 0.05$

• Aova is tool for testing two or more means simeoultaneously.

total Variance = b/w Simple variation + within simple variation

SST = SSPR + SSE.

Sum of square total = Sum of + S.S of error

• Distribution is normal
• populor varaince is same
• Sample must be random and independent

Date______20___

response variable = dependent variable

Example

C	C#	C++	Python

D.V = time

Independent = software/language

In one way we have one dependent and one independent variable.

In two way we have one dependent variable and two ~~multiple~~ I variable

Example

	M		E		D	
	10	100	6	36	5	25
	12	144	8	64	9	81
	9	81	3	9	12	144
	15	225	0	0	8	64
	13	169	2	4	4	16
Total y	59		19		38	⇒
total y^2	719		113		330	
	3681		361		1444	= 5286.

$$\text{total } y^2 = 719 + 113 + 330 \Rightarrow 1162$$

$$\text{total } y = 59 + 19 + 38 = 116.$$

$$\text{Correction factor} = \frac{(116)^2}{15} = 897.066.$$

$$SST = \Sigma\Sigma y^2 - CF$$

$$SST = 1162 - 892.066$$

$$= 264.9330.$$

$$SSTR = \frac{\Sigma y_i^2}{n} - CF$$

$$= \frac{5286}{5} - 892.066$$

Date ______ 20__

SSR - 160.1330

~~SST~~ = ~~160~~ 264.9330 - 160.1330

Variance	SS	df	M.Square	~~Average~~ F-ratio
B/w Sample	160.1330	3-1=2	160/2 = 80.067	$\frac{80.087}{8.733}$ = 9.168
within	264-160 104.8	14-2=12	~~52.4~~ 8.733	
total	264.9330	15-1=14		

As F-ratio is greater than 4 so reject H_0

$F[\alpha]_{df_1 \; df_2} = F(0.05)_{2 \; 12} = 3.89$