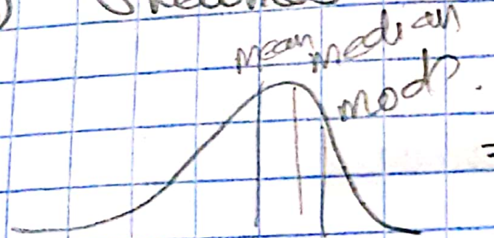
1) Relative Frequency

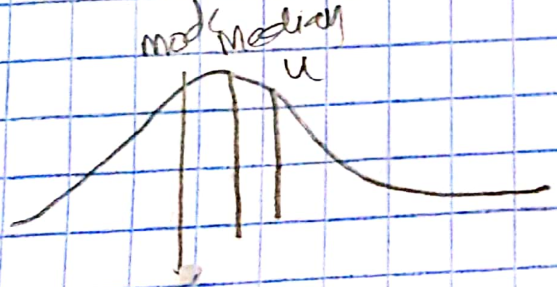
$$\frac{\text{Frequency in that class interval}}{n \text{ (total } f)} \times 100$$

Bar Charts
Line Graphs

2) Skewness



= Left Skew Mean < Median < Mode
-ve Skew



= Right Skew
+ve Skew

Mean > Median > Mode

3) Stem + Leave Plot

01 23 46 58 67 103 114

Stem	Leave
0	1
2	3
4	6
5	8
6	7
10	3
11	4

4) Sample Mean $= \frac{\Sigma x}{n}$

$$\bar{x} = \frac{\Sigma x}{n}$$

5) Mid Range $= \frac{Q_3 + Q_1}{2}$

6) IQR $= Q_3 - Q_1$

7) $\frac{\Sigma (x-\bar{x})^2}{n-1}$ = Variance

$$\sigma^2 = \frac{\Sigma (x-\bar{x})^2}{n-1}$$

$\underline{\underline{\sigma}} = \underline{\underline{sd}}$ $\qquad \sigma^2$ = Variance

$$\sqrt{\sigma^2} = sd$$

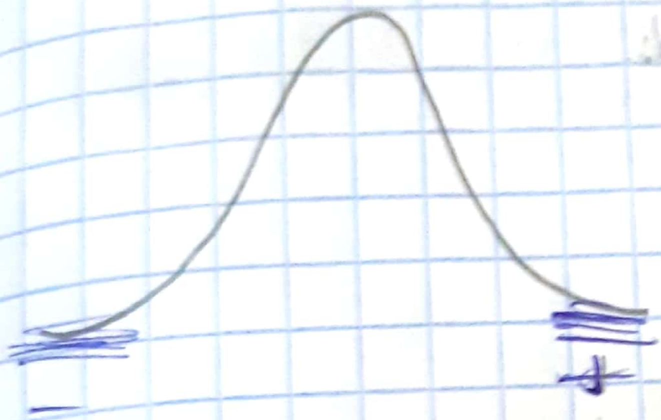
8) Coefficient of Variation

$$\frac{Sd}{mean}$$

9) Z score = $\frac{\text{Value} - \text{Mean}}{\text{std deviation}}$

$\frac{x_1 - \mu}{\sigma}$ = Score

> = 1 − Zscore Value from



< = Same Value

−ve = Shade −ve side
+ve = Shade +ve side

10

• frequency

✓ Coded data x_c

$$\frac{\text{Midpoint of group} - \text{Midpoint of median group}}{\text{Class Width}}$$

✓ Coded Mean $= \overline{x_c}$

$$= \frac{\sum f x_c}{\sum f}$$

✓ Coded Variance $= \dfrac{\sum f(x_c)^2 - n(\overline{x_c})^2}{N-1}$

✓ Reverse Coding

Actual Mean

$(\bar{x}_c \times CW)$ + Mid point of Median group

Actual Variance

$S_c^2 \times (CW)^2$

Actual Sd

Coded Sd $\times$ CW

Class Limit	Frequency	Midpoint (x)	Cf	x_c	$\Sigma f x_c$	$f(x_c)^2$
15 - 20	55	17.5	55	-2	-110	220
20 - 25	69	22.5	124	-1	-69	69
25 - 30	84	27.5	208	0	0	0
30 - 35	47	32.5	255	1	47	47
35 - 40	26	37.5	281	2	52	104
	Σf 281				$\Sigma f x_c = -80$	$\Sigma f(x_c)^2 = 440$

$x^c = \frac{17.5 - 27.5}{5}$

$\frac{22.5 - 27.5}{5}$

$\frac{27.5 - 27.5}{5}$

$\frac{32.5 - 27.5}{5}$

$\frac{37.5 - 27.5}{5}$

Coded Mean

$= \frac{\Sigma f x_c}{\Sigma f}$

$= \frac{-80}{281}$

$= -0.2847$

Actual Mean

(Coded Mean $\times$ CW) + Mid

$(-0.2847 \times 5) + 27.5$

$= 26.0765$

Coded Variance

$\frac{\Sigma f(x_c)^2 - n(\bar{x}_c)^2}{N-1}$

$440 - 281(-0.2847)^2$

$= 417.224$

$\frac{417.224}{281-1} = 1.49$

Coded sd : $\sqrt{1.49} = 1.22$

Act variance $= 1.49 \times 5^2 = 37.25$

Act sd $= 1.22 \times 5 = 6.1033$

Class Limit

Frequency

Calculating Grad Median

Find n^{th} value $= \frac{n}{2} = \frac{25}{2} = 12.5$

Median = Lower Median Class Boundary

$+$

$CW \times \left(\frac{(n^{th} \text{ value of median} - cf \text{ of prev class})}{\text{frequency of median class}} \right)$

$24.5 + \left(\frac{12.5 - 11}{8} \times 5 \right)$

$= 25.4375$

Mutually Exclusive

2 ~~3~~ times

p atleast 1 head

HH

HT

(HH) + (HT)

BINOMIAL

$n =$

π

$p =$ win

100 ~~2~~ people popn
30 of them smoke = win $= \frac{30}{100}$

6 random

p = all smoke = ~~so~~ y = 6

~~p = 6~~

$p = (x = 6)$

$\binom{n}{y} \pi^{y} (1 - \pi)^{n-y}$

* $\binom{n}{y} \pi^{y} (1-\pi)^{n-y}$

$n = 10$ $\quad p = \pi = 22\%$
$y = 5$

$\binom{10}{5} 0.22^{5} (1-0.22)^{10-5}$

$252 \times (5.15 \times 10^{-4}) \times 0.289 = 0.0375$ or 3.7

* 2> $\quad 2 \leq 0$

$Y=0 \quad Y=1 \quad Y=2$

* atleast 1

$1 - \cancel{P} \, P(Y=0)$

* Mean = $n\pi$
Variance = $n\pi(1-\pi)$

* If mean $= n\pi \geq 5$: = Normal $\Downarrow$ Z score

and $n(1-\pi) \geq 5$ = Normal $\Downarrow$ ~~N$~~ Zscore.

Poisson

~~$P(Y = y)$~~ =

$\lambda = 1.8$ (rate) $y = 4$ = 4 births

$$\frac{\lambda^y e^{-\lambda}}{y!} = \frac{1.8^4 e^{-1.8}}{4!}$$

$= 0.0723$

Type 1 = There is an association when there is.

Ho = No associat ⇒ Ac
HA = Associat ✓

Type 2 .

CI

Mean and Proportion

Sample > 20

~~Sample~~ .

1. Calculate the SEM

$$SEM = \frac{sd}{\sqrt{n}} = \sqrt{\frac{sd^2}{n}} = \text{Mean}$$

$$SEM = \sqrt{\frac{p \times (1-p)}{n}} = \text{Proportion}$$

= Mean $\pm$ Zscore $\times$ SEM
(95%
1.96)

CI Proportion $\pm$ Zscore $\times$ SEM

Using t Sample < 20

95%

$\bar{x} = 34.1$ $n = 25$ $\alpha = 1 - 0.95$
$s = 4.3$ $= 0.05$

$v = df = 25 - 1 = 24$

Mean $\pm$ tvalue $\times$ SEM

Mean $\pm t_{(n-1)\frac{\alpha}{2}} \times$ SEM

Signifiance Test

~~Zscore = sample mean / std error of mean~~

$$\cancel{= \frac{\bar{x} - \mu_0}{SEM}}$$

$$\cancel{= 34.1}$$

Z score $= \frac{\bar{x} - \mu_0}{SEM}$

$n > 20$

$$= \frac{34.1 (\text{Africa}) - 37.1 (\text{US})}{\frac{sd(\text{Africa})}{\sqrt{n : \text{Africa}}} = \frac{4.3}{\sqrt{25}}}$$

$$= -3.49$$

95%: $\alpha = 0.05$ $= 0.00024 \times 2$

$= 0.00048$

Reject H_0

Use t table = Same formula

when $n < 20$

=> You get probability and go look where it falls

0.001	0.0005
0.002	0.001

= -3.49

< 0.05

Comparing two means

$n < 40$ = used

$n > 40$

A) CI for 2 different population

$$\bar{x}_1 - x_2 \pm Z_{\alpha/2} \times \sqrt{\frac{sd_1^2}{n_1} + \frac{sd_2^2}{n_2}}$$

=> You will get a range

* $-$ and +ve = zero is involved

$\Downarrow$

Hence no real diff in mean

* +ve and +ve = zero not involved

Hence diff in mean of popn

Sign Test comparing means $n > 40$

$$Z = \frac{\bar{x}_1 - \bar{x}_2}{SEM}$$

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

⇒
from Z table if value is less than < 0.05
there is a real difference in mean in two popr

$n < 40$

~~Sign test~~ mean = Use t cz
Mean compared when $n < 40$

1) $n < 40$
2) Common variance

Compare using CI   Common variance

$$CI = \bar{x}_1 - \bar{x}_2 \pm t^* \sqrt{Sp^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

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

Again use + and -ve range
No diff as zero is include

+ve + +ve
Difference is there as no zero.

Sign

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{Sp^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Use same

< 0.05

Unequal Variance

F ~~test~~

F=

$$F(v_1, v_2) = \frac{S_1^2}{S_2^2} = \text{Calculated}$$

$v_1 = n_1 - 1$
$v_2 = n_2 - 1$

If Calculated is more than Critical do welch t ~~test~~

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Paired t test

= ~~Mean~~ difference for each obs

$\bar{d}$ Mean difference ⇒ Also include +ve + -ve.

Sd of mean = Calc

$SEd = \frac{sd}{\sqrt{n}}$

$T = \frac{\bar{d}}{SE(\bar{d})}$ = value at df (n-1)

Compare = ~~Should give~~ < 0.05

~~if no diff~~

Reject = There is a diff

$$\cancel{Z = \frac{p_1 - p_2}{\sqrt{\bar{p}}}}$$

CI $= p \pm 1.96 * \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$

Sign

$$Z = \frac{p_1 - p_2}{\sqrt{\bar{p}q\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

Beef Hot dog = Mean = 162.2

~~Variance~~

$\sqrt{x}$ = sd = 23.125

Poultry = $\bar{x}$ = 117.583

sd = 24.239

F test

$F(v_1, v_2)$ $\therefore$ No Welch

$n_1 = 15$

$n_2 = 12$

CI

$$x_1 - x_2 \pm t^* \sqrt{S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$$

$$S_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}$$

= ~~162.2~~

$S_p^2 = 557.98$

$162.2 - 117.583 \pm 2.060 \sqrt{557.98\left(\frac{1}{15} + \right.}$

44.617 ± 18.846

$63.463 \qquad 25.8$

25.8

S.g

$$t = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{S_{p^2}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

$$= \frac{44.617}{557.98\left(\frac{1}{15} + \frac{1}{12}\right)}$$