Statistics ①

↓

two processes

- describing data
  - Descriptive Statistics
    - → Involves
      - collecting, presenting and characterizing data.
- Drawing conclusion
  - Inferential statistics.
    - → Estimation, Hypothesis test

① ⇒ Calculating <u>Mean</u>, <u>Median</u>, <u>Mode</u> and <u>Range</u>

12, 7, 14, 5, 7, 11, 9

→ Mean $= \frac{\text{Sum}}{n} = \frac{12+7+14+5+7+11+9}{7} = \frac{65}{7}$

$= 9.286$

→ Median = ~~5~~, ~~7~~, ~~7~~, 9, ~~11~~, ~~12~~, ~~14~~ → ordered

Median = 9

→ Mode = 5, 7, 7, 9, 11, 12, 14

Mode = 7

→ Range 5, 7, 7, 9, 11, 12, 14

$= 14 - 5$

$= 9$

(2) ⇒ Ali recieved the following score on the first test:
what score will he need to get her next exam
to have an average score of 90?

Handling Missing value with Mean.

82, 89, 85, 96, ? $\bar{X} = 90$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{5}$$

$$5 \times \frac{90}{1} = \frac{82 + 89 + 85 + 96 + x_5}{\cancel{5}} \times \cancel{5}$$

$$450 = 352 + x_5$$

$$\boxed{98 = x_5}$$

(3) ⇒ Standard deviation.

32, 111, 138, 28, 59, 77, 97

$(32 + 111 + 138 + 28 + 59 + 77 + 97)/7 = 77.4$

$32 - 77.4 = -45.4$	$(-45.4)^2 = 2061.16$
$111 - 77.4 = 33.6$	$(33.6)^2 = 1128.96$
$138 - 77.4 = 60.6$	$(60.6)^2 = 3672.36$
$28 - 77.4 \quad -49.4$	$(-49.4)^2 = 2440.36$
$59 - 77.4 \quad -18.4$	$(-18.4)^2 = 338.56$
$77 - 77.4 \quad -0.4$	$(-0.4)^2 = 0.16$
$97 - 77.4 \quad 19.6$	$(19.6)^2 = 384.16$

$(2061.16 + 1128.96 + 3672.36 + 2440.36 + 338.56 + 0.16 + 384.16)/7$

$$\boxed{= 1432.2}$$

④ ⇒ Mean Absolute deviation.

7, 11, 14, 19, 22, 29

X	$\bar{X}$	$X-\bar{X}$	$\lvert X-\bar{X}\rvert$
7	17	-10	10
11	17	-6	6
14	17	-3	3
19	17	2	2
22	17	5	5
29	17	12	12
102			38

$$MD = \frac{\sum |x-\bar{x}|}{n}$$

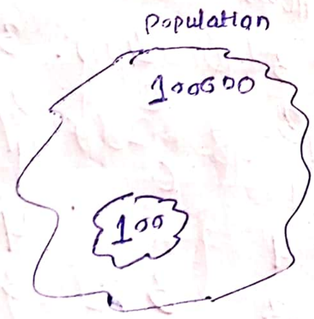
$$= 38/6$$

$$= \frac{36}{6} + \frac{2}{6}$$

$$= 6 + 1/3$$

$$\boxed{= 6.\bar{3}}$$

⑤ ⇒ Sample vs Population.



Sample Mean

$n = 100$

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

↓

'Statistics'

Population Mean

$N = 100000$

$\mu = \frac{Sum}{n} \rightarrow$ "Parameter"

$\because \bar{X} \approx \mu \quad \bar{X} \rightarrow \mu$

$$\bar{X} = \sum_{i=1}^{n} x_i = \frac{x_1 + x_2 \cdots x_n}{n}$$

$$\mu = \frac{\sum_{i=1}^{N} x_i}{N} = \frac{x_1 + x_2 \cdots x_N}{N}$$

(6) ⇒ Calculate the weighted Mean

16, 20, 12, 16, 16, 10, 16, 20, 24, 20

$$\bar{x} = \frac{\sum x}{n} = \frac{16+20+12+16+16+10+16+20+24+20}{10}$$

$$= \frac{170}{10} = \boxed{17}$$

$$\bar{x}_w = \frac{\sum wx}{\sum w} = \frac{1(10)+2(12)+4(16)+3(20)+1(24)}{10}$$

$$\frac{170}{10} = \boxed{17}$$

(7) ⇒ Calculate Weighted Average

Class	Hours	Grade	Points
chemistry	3	B	3
Physics	3	C	2
Lab	1	C	4
calculus	4	A	4
English	3	B	3

$$\bar{x}_w = \frac{w_1x_1 + w_2x_2 + \cdots w_5x_5}{w_1 + w_2 + \cdots w_5}$$

$$\bar{x}_w = \frac{3(3)+3(2)+(1)(4)+(4)(4)+(3)(3)}{3+3+1+4+3}$$

$$\bar{x}_w = \frac{9+6+4+16+9}{19} = \frac{44}{14} = \boxed{3.14}$$

(8) ⇒ How to Calculate Variance of Sample

6, 9, 14, 10, 5, 8, 11

$$\bar{x} = \frac{\sum x_i}{n} = 63/7 = 9$$

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

$(x_i - \bar{x})$		$(x_i - \bar{x})^2$
5 − 9	−4	16
6 − 9	−3	9
8 − 9	−1	1
9 − 9	0	0
10 − 9	1	1
11 − 9	2	4
14 − 9	5	25

$$s^2 = \frac{56}{n-1} = \frac{56}{6}$$

$$\boxed{= 9.3}$$

(9) ⇒ Scale of Measurements.

- Nominal Scale.
  (1) Quantitive / catagorical.
  (2) Name, colors, label, gender etc.
  (3) order does not matter.

- Ordinal Scale
  (1) Ranking / placement
  (2) the order metters
  (3) Difference can not measure

- Interval Scale Data.
  (1) the order Matters.
  (2) Difference can be measured.
  (3) No true "0" starting point?

- Ratio Scale
  (1) the order Matter.
  (2) Differen are measured
  (3) Contains "0" starting point.

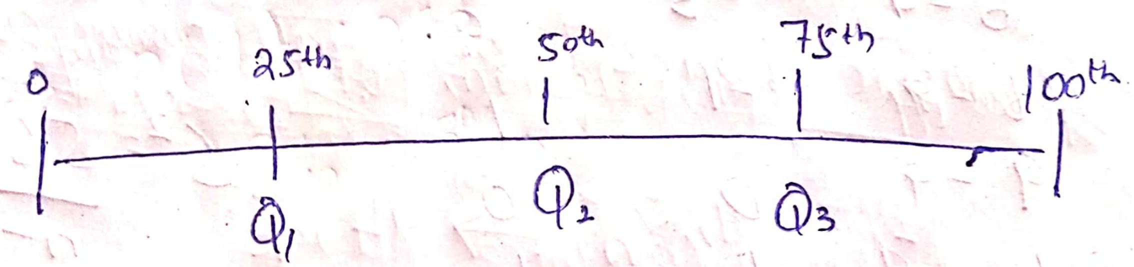
(10) ⇒

Quartiles: Divide data into 4 equal types

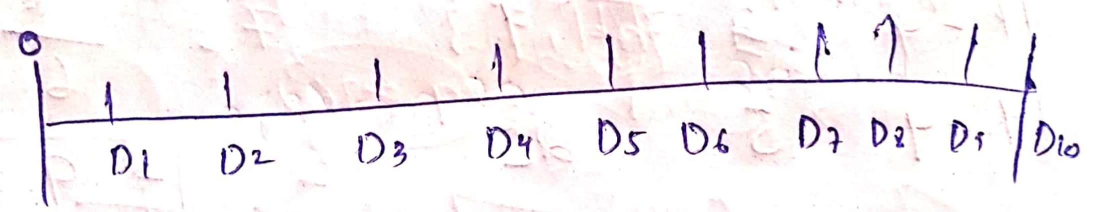
Deciles: Divide data into 10 equal type

Percentiles: Divide data into 100 equal parts

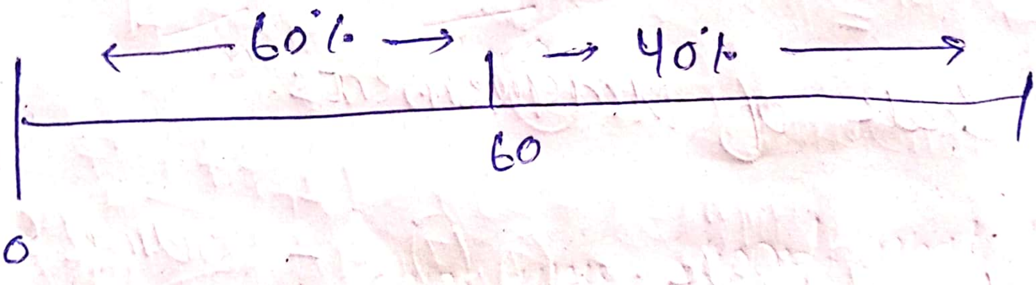
Quantile



Deciles



Percentile



Probability

$P(A) \rightarrow = \dfrac{\text{Outcomes}}{\text{total outcomes}}$.

Probabity of event a occuring.

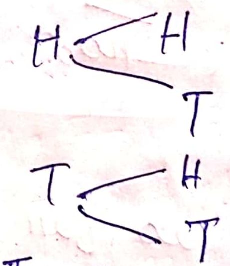
How to calculate probability?

sample Space: set of all possible outcomes.

eg 1 coin → H, T

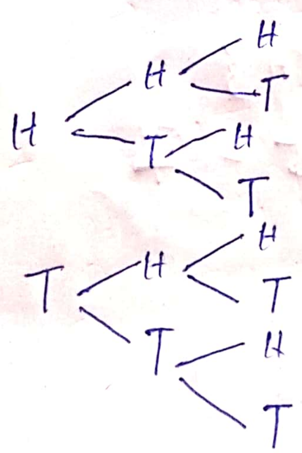
$S = \{H, T\}$

2 coins → 1 2



$S = \{HH, HT, TH, TT\}$

① ⇒ practicle example for three Coins.



$S = \{HHH, HHT, HTH, HTT,$
$THH, THT, TTH, TTT\}$

$= 2^3$

$= 2 \cdot 2 \cdot 2$

$= 8$

(2) ⇒ $0 \leq p(A) \leq 1$

0 ↓ Never happen

1 ↓ always happen.

⇒ $p(A) = 0.3 \rightarrow 30\%$ → chance of occurrence

⇒ $p(A) = 0.3$

$\{ 3/10 \quad 30/10 \}$

⇒ example situations

$p(B) = 0.20 \quad 20\%$

$\left\{ \frac{200}{100} \quad \frac{2000}{1000} \right\}$

(3) ⇒ if two fair coins were flopped. what is the probability at least one head?

$S = \{HH, HT, TH, TT\}$

$A \rightarrow \{HH, HT, TH\}$

$P = 3/4 = 0.75 = 7$

4 ⇒ Marginal probabity

$S = \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ $A\{1, 2, 3, 4\}$

$B = \{6, 7, 8\}$ $C = \{1, 7, 8, 9\}$

$$P(A) = \frac{\text{Successful outcomes}}{\text{total possible outcomes.}}$$

$P(A) = \boxed{4/9}$

$P(B) = 3/9 = \boxed{1/3.}$

$P(C) = 4/9 = \boxed{0.444}$

5 ⇒ Union probality

$P(A \cup B)$

$A \cup B = \{1, 2, 3, 4, 6, 7, 8\}$

$= 7/9 = \boxed{0.7} = 77.7\%$

$P(A \cup B) = P(A) + P(B) - P(A \text{ and } B)$

$ME = P(A \text{ and } B) = 0$

$ME = P(A \text{ or } B) = P(A) + P(B)$

$P(B \text{ or } C) = P(B) + P(C) - P(B \text{ and } C)$

$= 3/9 + 4/9 - 2/9 = \boxed{5/9}$

⑥ ⇒ How to calculate expected value.

1. Lisa plays a game in which there are only two outcomes. the cost play the game 100$ if she wins, she recieved 500$. the probability of wing is 20% . what is the expected value for winning a single game on average.

$E(x) = x_1 p_1 + x_2 p_2$

$E(x) = 500(0.20) + (-100)(0.80)$

$E(x) = 100 - 80$

$\boxed{E(x) = \$ 20.}$

outcomes:	wins	loss
values:	\$ 500	-\$100
Probability	20%	80%

7) Normal distribution

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

example

if $X \sim N(4, 9)$, find $P(X > 6)$

Solution

$X \sim N(\bar{x}, \sigma^2)$

let $Z = \frac{X - \mu}{\sigma} = \frac{X - 4}{3}$

Hence.

$P(X > 6) = 1 - P(X < 6)$

$= 1 - \Phi\left(\frac{6-4}{3}\right)$

$= \Phi(0.67)$

$= 1 - 0.74857$

$\boxed{= 0.25143}$

example what will be the probabity density function of normal distribution of data.

$x = 3$, $\mu = 4$ and $\sigma = 2$?

Solution

$x = 3$

Mean $= 4$

Std $= 2$

$$F(3, 4, 2) = \frac{1}{2\sqrt{2\pi}} e^{\frac{-(3-2)^2}{2 \times 2^2}}$$

$F(3, 4, 4) = \boxed{1.106}$

(8) ⇒ Binomial distribution: $P_x = \binom{n}{x} p^x q^{n-x}$.

example:

if a coin is tossed 5 times find the probability of:

(a) Exactly two head.

(b) At least 4 head.

Solution

$n = 5$

P(head) $p = 1/2$ P(tail) $q = 1/2$.

(a):

$x = 2$.

$P(x=2) = {}^5C_2\, p^2 q^{5-2} = 5!/2!\,3! \times (1/2)^2 \times (1/2)^3$

$\boxed{P(x=2) = 5/16}$

(b):

$x \geq 4$, $P(x \geq 4) = P(x=4) + P(x=5)$

Hence

$P(x=4) = {}^5C_4\, p^4 q^{5-4} = 5!/4! \times (1/2)^5 = 5/32$

$P(x=5) = {}^5C_5\, p^5 q^{5-5} = (1/2)^5 = 1/32$.

therefore

$P(x \geq 4) = 5/32 + 1/32 = 6/32 = \boxed{3/16}$

(9) ⇒ Passion Distribution

$$F(x) = (e^{-\lambda} \lambda^{x})/x!$$

example:

A random variable x has a passion distribution with parameter λ such that $P(x=1) = (0.2)$ $P(x=2)$. Find $P(x=0)$.

Solution

$$P(X=x) = (e^{-\lambda} \lambda^{x})/x!$$

Given that

$$P(x=1) = (0.2)\ P(x=2)$$

$$(e^{-\lambda} \lambda^{1})/1! = (0.2)\ (e^{-\lambda} \lambda^{2})/2!$$

$$\Rightarrow \lambda = \lambda^{2}/10$$

$$\lambda = 10$$

$$P(x=0) = (e^{-\lambda} \lambda^{0})/0!$$

$$P(x=0) = e^{-10} = 0.000454$$

$$\text{thus } P(x=0) = \boxed{0.0000454}$$

(10) ⇒ Uniform Distribution.

$$P_x = 1/n \quad / \quad F(x) = 1/(b-a) \text{ for } a \le x \le b$$

example:

Using the Uniform Probality density function of a random variable $X - X \sim (0, 23)$, Find $P(2 < X < 18)$

Solution

$P(2<X<20) = (20-2) = \frac{1}{23-0} = 18/23$

$\boxed{P = 18/23}$

(b) I have taken a full table in excel.

Solution

Sample mean $= 4.83$.

distribution lies between 0 and 14.

thus

$X = 0$ and $Y = 14$.

theoritical Mean $= \mu = (x+y)/2 = (0+14)/2 = 7$

theoritical Std $= \sigma = \sqrt{[\{x-y\}2/12]} =$

$\sqrt{[[0-14]2/12}$

$= ?(196/2)$

$= ?\ 98$

$\boxed{= 9.899}$