Statistics 1

Notebook

للدكتور : محمد شبول

Statistics deals with : 1. making decisions 2. solve problems 3. Design products (and processes)

Statistics learns information from data, and is useful to describe and under-stand variability

* dot diagram is useful for displaying small body of data that is up to 20 observations.
It allows us to see 2 features of data : 1. location or the middle 2. the scatter or variability

T.parameters → quantities of the true values. → (true parameters := T. parameters)
when the data is variable we can calculate the random variable(x) by $X = \mu + \varepsilon$
(X: R.V, μ: constant, ε: random dist.)

Reasoning

physical laws → product designs

population ← sample → statical inference

*

* statical inference → generalize (give the conclusion to the population)

ways of collecting data

retrospective (past data)

observational: data, presently collected, passive observer

designed experiment: collected in response to process input change

we can use Hypothesis test in researches

The control chart can detect the change over time and shifts and to observe the process
if we had shifts we must adjust the experiment other wise don't adjust. (Deming exp.)

Models

mechanistics models
(From knowledge of physical mechanism)
like Ohm's law : $I = E/R + \underbrace{\varepsilon}_{\text{variability}}$

Empirical model
From Engineering and scientific knowledge of the phenomenon (based on data)

probability models → quantify the risks involved in statistical inference (risk = ...

- probability provides the framework for study
-

Sample space → the set of all possible outcome for a random experement (S)

Sample Space (S)

discreat → finite, countable infinit

interval (continuas) → finite, infinite

To acheive the understanding we quantify and model the variation that affects the physical system's behaviour, the model analyze and predict the physical system's behaviour as a system input affects system outputs. (prediction → experiment, with the physical's system).

physical system

Measurement

Analysis

modle

controlled variables
↓
input → system → output
↑
noise variables. (variables we can't control)

- variation causes distrubtion in the system.

event (E), subset of the sample space

Points of the sample space are called elementary events or realization

Ex 8 1. throw a dice → $S = \{1, 2, 3, 4, 5, 6\}$ x: # shown discrete (finite)

realization

2. throw coin → $S: \{h, t\}$ x: Face shown discrete (finite)

realization

3. 3 machine parts, even classified as either above or bellow specification

$S: \{aaa, aab, aba, baa, bba, abb, bab, bbb\}$ discreate (finite)

realization

4. # of views recorded at a high volume website in a day

$S = \{0, 1, 2 \dots n\}$ discrete (countably infinite)

n is very large

5. selecting a ball from urn wich has three balls green, blue, red

$S: \{G, B, R\}$

6. Rain precipitation

$S: R^+ : \{x \mid x > 0\}$ continuas (infinite)

- if it never goes above 50'' $S: \{x \mid 0 \leq x \leq 50''\}$ continuas (finite)
and didn't rain

7. Two connectors measures ... $S = R^+ \times R^+$ continuas

- if we are intrusted to show if the connecters is conferming or not conforming:
$S: \{yy, yn, ny, nn\}$ descreate (finite)
- if intrested in # of conforming: $S: \{0, 1, 2\}$ discrete (finite)
- if measuring thickness until first fail: $S: \{n, yn, yyn, yyyn, \dots\}$
discrete (countably infinite)

Scanned with CamScanner

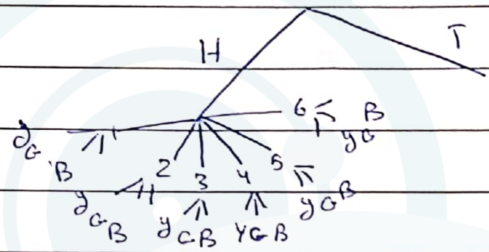
Graphical representation of S using tree diagrams.

we use it when experiment is done in several steps

n_1 ways of 1^{st} step $\rightarrow$ n_1 branches of the tree

n_2 ways " 2^{nd} " $\rightarrow$ n_2 " " " "

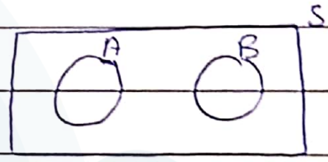
Ex tossing a coin, rolling die and choosing a ball from urn (blue, yellow, green).



$S = \{ H1y, H1G, H1B, H2y, H2G, H2B, \dots \}$

Basics of set theory 8

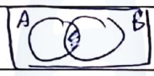
Two events A, B we use venn diagrams $\rightarrow$



• $A \cup B \Rightarrow$ OR

union

• $A \cap B \Rightarrow$ And



intersection

• $\emptyset \Rightarrow$ empty set. or $\{ \}$

• $A^c \Rightarrow A' \Rightarrow$ complement "not"

$(A \cup B)' = A' \cap B'$ $\rightarrow$ $(A \cap B)' = A' \cup B'$

If $A \cap B = \emptyset = \{ \}$ $\Rightarrow$ A and B mutually exclusive or disjoint

that means if A happened B won't happen for sure.

Digital schale provides weights to the nearest gram.

A: event that a weight exceeds 11 g

B: " " " " " is less than or equal to 15 g

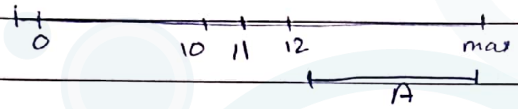
C: " " " " " is greater than or equal to 8 g and less than 12 g

, X is weight

S = { non negative integers from 0 to the largest integer that can be displayed by the scale

discrete (finite)

A.



B.

C.

Find 1. $A \cup B = S$

2. $A \cap B = \{ x \mid 12 \le x \le 15 \} = \{12, 13, 14, 15\}$

3. $A' = \{ x \mid x \le 11 \} = \{0, 1, 2, \dots, 11\}$

4. $A \cup B \cup C = S$

$(A \cup C)' = S - (A \cup C) = S - \{ x \mid x \geq 8 \}$

$= \{ x \mid x < 8 \}$

5. $A \cap B \cap C = \emptyset$

6. $B' \cap C = \emptyset$

7. $A \cup (B \cap C) = \{ x \mid x \geq 8 \}$

Counting techniques:

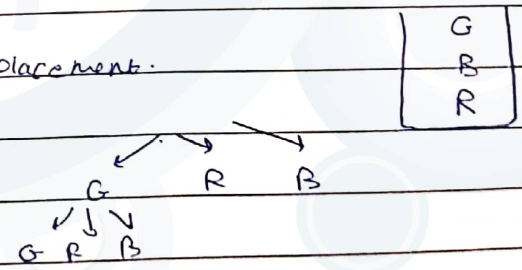
of out comes ↑

① multiplication rule: operation with k steps, each step is done in # of ways

Step 1 with n_1, Step 2 with n_2

total number of ways $= n_1 * n_2 * \dots * n_k$

ex: draw a marble twice from an urn with replacement.

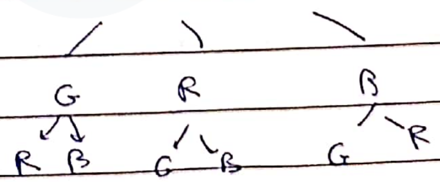


1st draw: 3 choices

2st draw: 3 choices

Tot # of choices: $3 * 3 = 9$

- if without replacement.



1st draw: 3 choices

2st draws: 2 choices

Tot # of choices $= 3 * 2 = 6$

In how many ways can we select r distinct objects from n distinct object

② permutations (order of items arrangment is important) GR ≠ RG

$n * (n-1) * (n-2) * \ldots \ldots \overset{k^{th}}{(n-k+1)}$

without replacing, k objects can be chosen from n with total number of ways

: $n * (n-1) * (n-2) \ldots (n-k+1)$

# of permutations $P^n_k = n(n-1)(n-2) \ldots (n-k+1) \frac{\cancel{(n-k)(n-k-1)\ldots}}{\cancel{(n-k)(n-k-1)\ldots}}$

$P^n_k = \frac{n!}{(n-k)!}$

Ex 5 students we need to choose president, vice, secretary.

5 students (A, B, C, D, E) $\underset{5}{\_}$ $\underset{4}{\_}$ $\underset{3}{\_}$

$= 5 * 4 * 3 = 60$

but the arrangement is important $P^5_3 = \frac{5!}{(5-3)!} = \frac{120}{2} = 60$

$P^n_n = \frac{n!}{(n-n)!} = \frac{n!}{0!} = n!$ ← in Excel = Fact (n)

ordering n ⇒ items multiplication rule

* Set of objects that contains k groups of identical items

$n = n_1 + n_2 + n_3 + \ldots + n_k$

group1 group2 group3 group k

# of ways of ordering n objects is $\frac{n!}{n_1!\, n_2!\, n_3!\, n_k!}$

Ex: we have 5 cars of mercides, 6 of BMW, 3 of Hydai which are identical.

$n = 5 + 6 + 3 = 14$

$P = \frac{14!}{5!\, 6!\, 3!}$

③ combinations (the order is not important)

# of ways of selecting r objects from n without regard to order is

$C^n_r = \frac{n!}{r!\,(n-r)!}$ Excel → Combin (n, r)

Ex: 4 cars Honda, toyota, Ford, BM . # of ways to sellect 2 of them.

HT, HF, HB, TF, TB, FB → 6

$C_r^n = \frac{4!}{2!\,(4-2)!} = \frac{4\times3\times2\times1}{2\times1\times2\times1} = 6$

Ex: group of 7 engineers and 5 staticians we need to form committee of 5, How many ways ~~to~~ committee of 3 engineers and 2 staticians.

$\underset{7}{E}\ \underset{6}{E}\ \underset{5}{E}\ \underset{5}{S}\ \underset{4}{S}$

$C_3^7 * C_2^5 = \frac{7!}{3!\,4!} * \frac{5!}{3!\,2!} = \frac{7\times6\times5}{3\times2} \times \frac{5\times4}{2} = 35\times10 = 350$

Solving problem 41

100 semiconductors , 10 are not comfortable . , 5 of 100

a. $C_5^{100} = \frac{100!}{5!\,(95)!} = 75,287,520$ — $\underset{100}{\square}\ \underset{99}{\square}\ \underset{98}{\square}\ \underset{97}{\square}\ \underset{96}{\square}$

b. $10 * C_4^{90}$ — $\underset{10}{\square}$ | 90

c. $C_5^{100} - C_5^{90} = 31,338,262$ — 90 89 88 87 86

Probability

It is alikehood or chance ~~of~~ that an outcome will occur . P(A) → event.

Ex: P(A) = 0.6

→ the event student achive B in class, likely hood that any random student will achive B is 60%.

→ 60% of students will acheive B

Experement → outcomes → event

random variable ← real

probability & function assignes to each event $A \subset S$, a real number $p(A) \in [0,1]$ as 1 means certainty and 0 impossibillity.

probability have rules like:

1. $p(A) \geq 0$
2. $p(S) = 1$
3. if $A_1, A_2, A_3, \dots$ are disjointed then $P(\cup_n A_n) = \sum_n P(A_n) = p(A_1) + p(A_2) + \dots$

Types of probability

subjective probability (degree of belief)

no accurate or stabel out come to do the assesment and is about what you belive in (in your point of view)

Relative frequancy

how many times the event will occur in a very large # of sampels

Equally likely

N outcomes $p(\text{each outcome}) = \frac{1}{N}$

Ex: urn contains 30 red marbels, 20 green and 50 blue marbels. what is the probability of choosing every marble.

$P(\text{choosen evry marble}) = \frac{1}{100} = .01$

- $p(\text{choosen red marble}) = \frac{30}{100} = .3$ or $P(RM_1) + P(RM_2) + \dots + P(RM_{30})$

$= \frac{1}{100} + \frac{1}{100} + \dots + \frac{1}{100} = \frac{30}{100}$

For a discrete sample space $P(E) = \sum_{1}^{n} P(\text{all outcomes of } E)$

properties of probability if we have event A and B

1. $P(\emptyset) = 0$

2. if $A \subseteq B \rightarrow P(A) \leq P(B)$

3. $0 \leq P(A) \leq 1$

4. $P(A') = 1 - P(A)$

5. if $(A \cap B) = \phi \rightarrow P(A \cup B) = P(A) + P(B)$

6. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Ex: batch contains 6 parts $\{a, b, c, d, e, f\}$ and f is a defective part, 2 parts are selected randomly without replacement what is the probability that f is in the sample

E = f is in the sample , $P(E) = \frac{\text{\# of samples contains f}}{\text{tot \# of any possible sample}}$

Total # of possible sampels choosing 2 from 6 $C^6_2 = \frac{6 \times 5}{2} = 15$

# of samples contains f $C^1_1 \times C^5_1 = 1 \times \frac{5!}{4!} = 5$

$P(E) = \frac{5}{15} = \frac{1}{3}$

Ex. (2-61) Nicle battary , Charge state and probabilites 8

Charge	prob.
0	.17
+2	.35
+3	.33
+4	.15

1. prob at least one of the positive charges.

$P(E_1) = P(+2) + P(+3) + P(+4) = .83$

2. prob that a cell has not a nickel charge greater than (+3)

$P(E_2) = P(0) + P(+2) + P(+3) = .85$

// Addition Rule

for any 2 events A , B

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$

we can think of it as

$A \cup (B \cap A^c)$

Ex: probability of raining today is 20%.

" " " tomorrow is 30%.

" " " today and tomorrow = 6%.

what is the probability of raining today or tomorrow?

$P(A \cup B) = .2 + .3 - 0.06 = .44$

Ex: (2-78) strand of copper wires if we have 100 strand

1. P(high strength and cond.) =
$= P(Hc \cap Hs) = \frac{74}{100}$

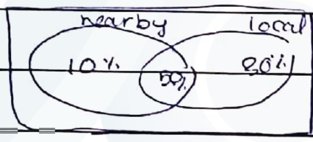
	strength high	strength low
high conductivity	74	8
low conductivity	15	3

2. $P(Ls \cup Lc) = P(Ls \cup Lc)$
$= P(Ls) + P(Lc) - P(Ls \cap Lc) = \frac{3+8+15+3-3}{100} = \frac{26}{100}$

3. E_1 = strand with Lc, E_2 strand have Ls
$P(E_1 \cup E_2) \overset{?}{=} P(E_1) + P(E_2)$
$\frac{26}{100} \neq \frac{8+3}{100} + \frac{15+3}{100} = \frac{29}{100}$

Ex: city 60% of households subscribe to newspaper published in a nerby city and 80% subscribe to local newspaper, 50% subscribe to both.

let A: { household subscribes to near city }
B: { " " " local }



Find
$P(A) = .6$
$P(B) = .8$
$P(A \cap B) = .5$

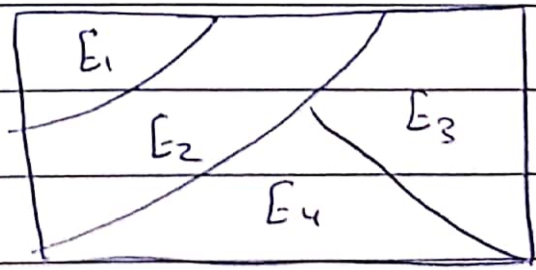
2. $P(A \cup B) = P(A) + P(B) - P(A \cap B) = .6 + .8 - .5 = .9$

3. the house hold subscribes to exactly one of the two papers
$P = .1 + .3 = .4$ or $.9 - .5 = .4$

* $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(A \cap B) - P(A \cap C) - P(B \cap C) + P(A \cap B \cap C)$

if we have $E_1, E_2, \ldots E_k$ ← collection of mutually exlusive events.

$P(E_1 \cup E_2 \cup E_3 \cup \ldots \cup E_k) = P(E_1) + P(E_2) + \ldots + P(E_k) = \sum_{i=1}^{k} P(E_i)$



$E_1, E_2, E_3 \rightarrow ME$

E_1, E_2, E_3, E_4 mutually exclusive and collectively exhasted, so $(E_1 + E_2 + E_3 + E_4 = S)$ because there is no intersection between them

* So $E_i \cap E_j = \emptyset$ for all i and j are mutually exclusive

* and $E_1 \cup E_2 \cup E_3 \cup E_4 = S$ → mat collectively exhausted.

Application of probability concepts (Reliability)

the reliability is the prob. that a system or product will perform its intended function properly.

Reliability systems

series system

a → [1] — [2] — [3] — [4] — b

- where 1, 2, 3, 4 must be success because if one of them fails all the system fails
- each component works / fails independently than others
- system will work if there is a path

A_i : event that component i works

P_i : prob " " " " "

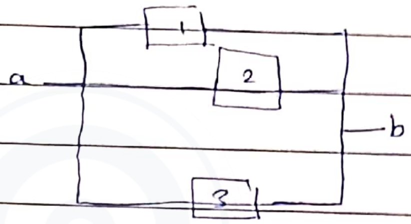
∴ $p(\text{system to work}) = p(A_1 \,\&\, A_2 \,\&\, A_3 \,\&\cdots\&\, A_n)$

$= p(A_1 \cap A_2 \cap A_3 \cdots\cdots \cap A_n)$

$= p(A_1)\, p(A_2)\, p(A_3) \cdots p(A_n)$

$= P_1 * P_2 * P_3 * P_n$

parallel system



$p(\text{system works}) = p(A_1 \text{ or } A_2 \text{ or } A_3 \text{ or} \ldots A_n)$

$= p(A_1 \cup A_2 \cup A_3 \cdots \cup A_n)$

$= 1 - p(\text{system doesn't work})$

$p(A_1^c \cap A_2^c \cap A_3^c \cap \cdots A_n)$

$= (1-P_1)(1-P_2)\cdots(1-P_n)$

as $p(A_1) = P_1$

so $p(\text{system work}) = 1 - [(1-P_1)(1-P_2)\cdots(1-P_n)]$

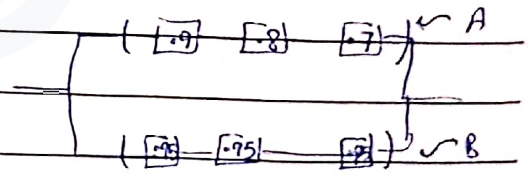
as the system is independent and A_i, P_i are defined as the series.

Summery :

in parallel system $p(\text{system work}) = \prod_{i=1}^{n} P_i$

in parallel system $p(\text{system work}) = 1 - \prod_{i=1}^{n} (1 - P_i)$

Ex 2-136 circuit operate iff there is a functional path from left to right and the devices works independently

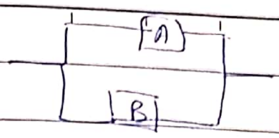


A : denotes that upper devices function

B : " " lower " "

$p(A) = .9 * .8 * .7 = .504$

$p(B) = .95 * .95 * .95 = .8574$



$p(A \cup B) = (1 - [(1 - .504)] * (1 - .8574)]) = .9293$

or $p(A \cap B) = .504 * .8574 = .4321$

$p(\text{circuit works}) = p(A \cup B) = p(A) + p(B) - p(A \cap B) = .9293$

Bayes' theorem

it is the event in terms of conditional probabilities

=> if the condition was present when events occures.

$p(A \cap B) = p(A/B) . p(B) = p(B/A) / p(A)$

$p(A/B) = \dfrac{p(B/A)\, p(A)}{p(B)}$ using the law of total prob. then:

$$p(E_i / B) = \frac{p(B|E_i)\, p(E_i)}{\sum_{i=1}^{n} p(B|E_j)\, p(E_j)}$$

Ex 2-142 $p(A/B) = .07$ $p(A) = .5$ $p(B) = .2$ $p(B/A)$?

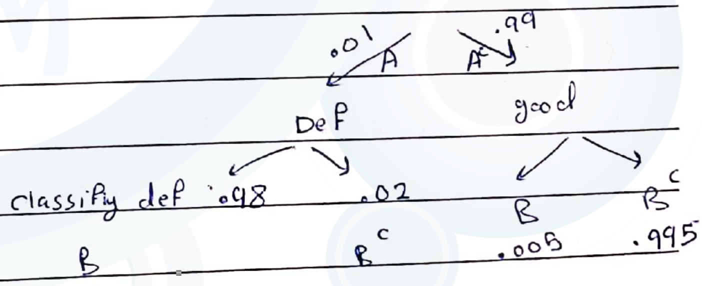
$p(B/A) = \dfrac{p(A/B)\, p(B)}{p(A)} = .28$

if $p(A)$ was not provided in the question then $p(A) = p(A/B)\, p(B) + p(A/B^c)\, p(B^c)$

Ex 2-148

A : item is defective

B : item is classified as defective.



ⓐ $p(B) = p(B/A)\, p(A) + p(B/A^c)\, p(A^c)$

$= .98 * .01 + .005 * .99 = .01475$

ⓑ $p(A^c / B^c) = \dfrac{p(A^c \cap B^c)}{p(B^c)} = \dfrac{p(A^c \cap B^c)}{1 - p(B)} = \dfrac{1 - p(A \cup B)}{1 - p(B)} = \dfrac{1 - (p(A) + p(B) - p(A \cap B))}{1 - p(B)}$

and $p(A \cap B) = p(B/A)\, p(A) = .98 * .01 = .098$

then $P(A^c|B^c) = \frac{1-[.01+.01475-.0098]}{1-.01475} = .9998$

Ex we can handle A, A^c, B, B^c as in the following table:

	A	A^c	
B	$P(A\cap B)$	$P(A^c\cap B)$	$P(B)$
B^c	$P(A\cap B^c)$	$P(A^c\cap B^c)$	$P(B^c)$
	$P(A)$	$P(A)^c$	

which means:

$P(A) = P(A\cap B^c) + P(A\cap B)$

$P(B) = P(A\cap B) + P(A^c\cap B)$

$P(A^c) = P(A^c\cap B) + P(A^c\cap B^c)$

$P(B^c) = P(A\cap B^c) + P(A^c + B^c)$

Ch #3 Discrete R.V & prob distribution.

Random variable : function defined on the sample space

or it is a function that assignes a real number to each outcome.

$X \quad Y \Rightarrow$ R.V

$x \quad y \Rightarrow$ value of R.V

Ex: length is 5 cm

the length is X and 5cm is x

Ex: Area is 20 m^2

the Area is Y and 20 m^2 is y

R.V

Discrete — Continuas.

(spesific amounts)

finite — countable infinite — interval — finite — infinite

X : 0,1,2,3 — X : 0,1,2,3...... max — [0,1] — [0, ∞) — (−∞, ∞)

Ex: throw 2 four face dice, find the sum of the result.

X : sum of values of the 2 dices.

$S = \{(1,1),(1,2) \dots (4,4)\}$

1st die \ 2nd die	1	2	3	4
1	2	3	4	5
2	3	4	5	6
3	4	5	6	7
4	5	6	7	8

$X = \{2, 3, 4, 5, 6, 7, 8\}$

Ex 3-4 batch of 500 machined parts contains 10 non conforming

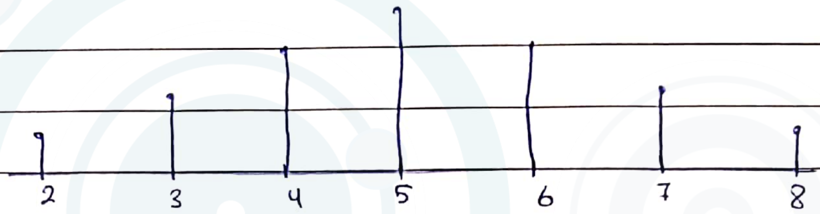
R.V = # of parts in a sample of 5 parts that don't conforms to costomer requirment.

$X = \{0, 1, 2, 3, 4, 5\}$

If we go back to the 2dices examped, we can use the random variable as an event

$p(X=2) = \frac{1}{16}$ $p(Y=4) = \frac{3}{16}$ $p(X=6) = \frac{3}{16}$

$p(Y=3) = \frac{2}{16}$ $p(X=5) = \frac{4}{16}$ $p(X=7) = \frac{2}{16}$



this shape is called distribution of the data of P(R.V) on the R.V