

VBEST NOTES

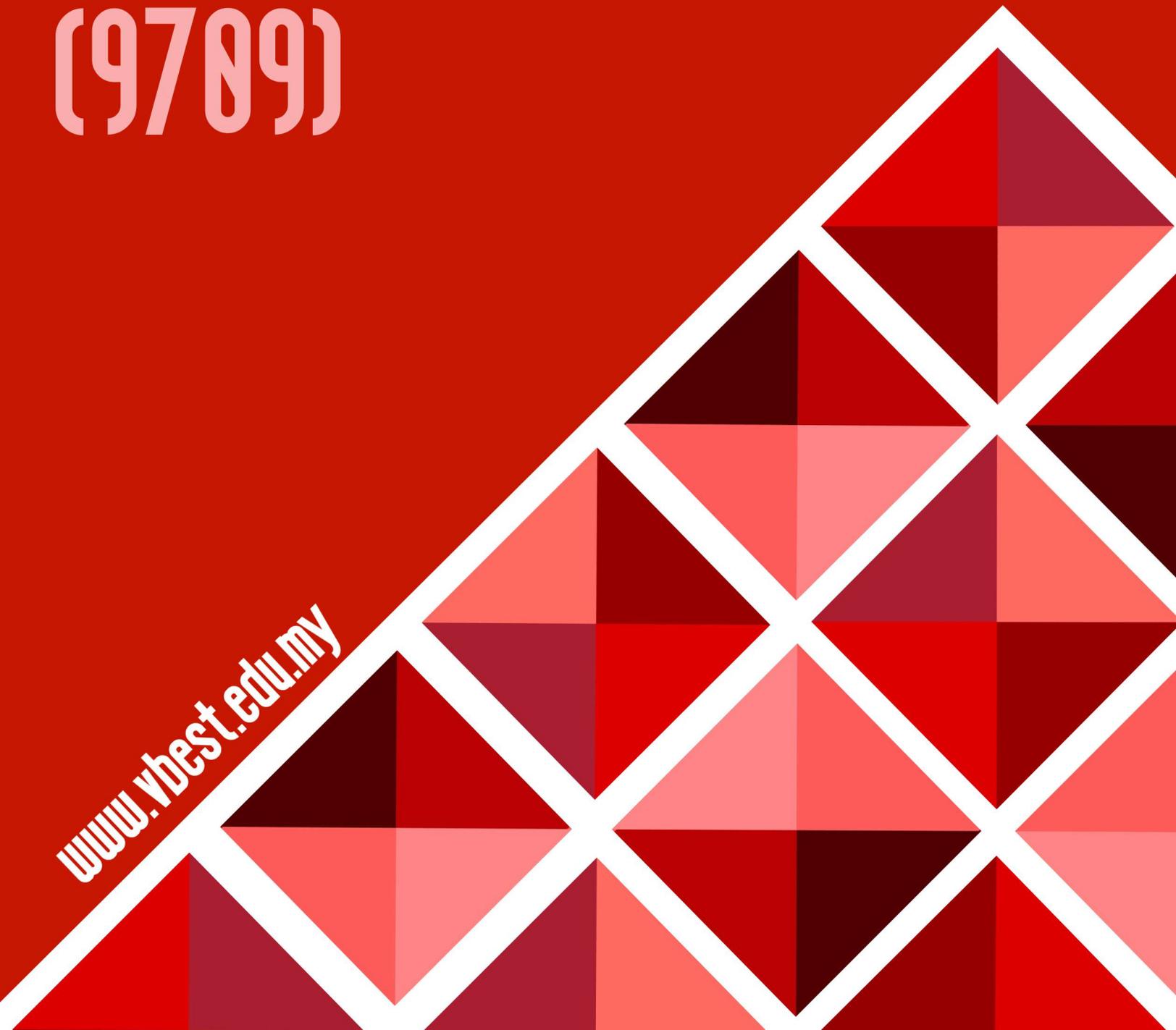


A LEVEL CIE

AS STATISTICS 1

(9709)

www.vbest.edu.my



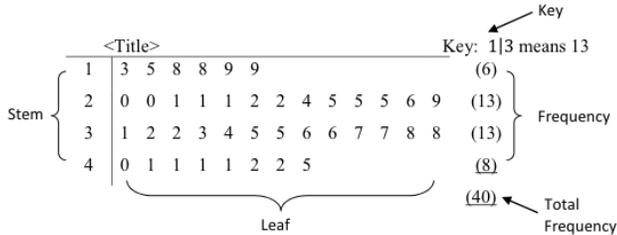
Statistics

- Representation of data
- Permutations and combinations
- Probability
- Discrete random variable
- Normal distribution

Chapter 1 : Representation of data

a) Organising quantitative data

- Stem and leaf diagram



Single set stem and leaf diagram



Back to back stem and leaf diagram

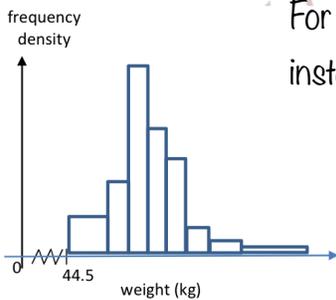
- Frequency distribution for discrete data

No. of cars	Frequency
10	10
20	17
30	21
40	25
50	28
60	30

- Frequency distribution for continuous data

Weight (kg)	Frequency, f	Class Boundaries	Midpoint, x	Cummulative frequency	Class width / size
21-30	5	20.5 - 30.5	25.5	5	10
31-40	4	30.5 - 40.5	35.5	9	10
41-50	2	40.5 - 50.5	45.5	11	10
51-60	3	50.5 - 60.5	55.5	14	10
61-70	6	60.5 - 70.5	65.5	20	10
Σf = 20					

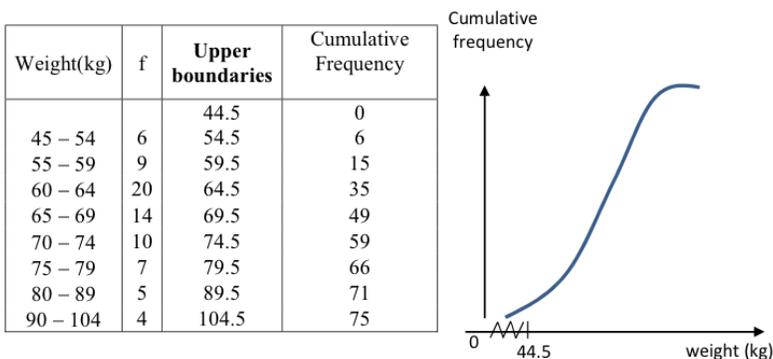
- Histogram



For uneven class size, frequency density is used instead of frequency

$$\text{Frequency density} = \frac{\text{frequency}}{\text{class width}}$$

- Cumulative frequency graph



Cumulative frequency is plotted against upper-class boundaries and first value of the cumulative frequency is zero

b) Measurements of locations and spread

Measurements of central tendency :

- Mode or modal class
- Mean
- Median

Measures of dispersion/spread :

- Range
- Interquartile range
- Standard deviation

- Mean and variance

Ungrouped data

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

n is the total number of datas

Grouped data

$$\text{mean} = \frac{\sum fx}{\sum f}$$

f is the total frequency

$$\text{mean} = \bar{x}$$

$$\text{Standard deviation} = \sqrt{\frac{\sum (x - \bar{x})^2}{n}}$$

OR

$$\sqrt{\frac{\sum x^2}{n} - (\bar{x})^2}$$

$$\text{Standard deviation} = \sqrt{\frac{\sum f(x - \bar{x})^2}{\sum f}}$$

OR

$$\sqrt{\frac{\sum fx^2}{\sum f} - (\bar{x})^2}$$

c) Coding

Eg: A summary of 20 observations of x gave the following information:

$$\sum (x-a) = 100, \quad \sum (x-a)^2 = 1500, \quad \bar{x} = 25$$

Find the value of a and the standard deviation.

$$100/20 = 25 - a$$

$$a = 20$$

$$\sqrt{\frac{1500}{20} - 5^2}$$

$$= \sqrt{50}$$

If x is deducted from every data then x is deducted from the final mean of the datas

*SD is not affected by coding

d) Median and interquartile range

Median is the middle value of an ordered set of data for observation

Q_1 : Lower quartile

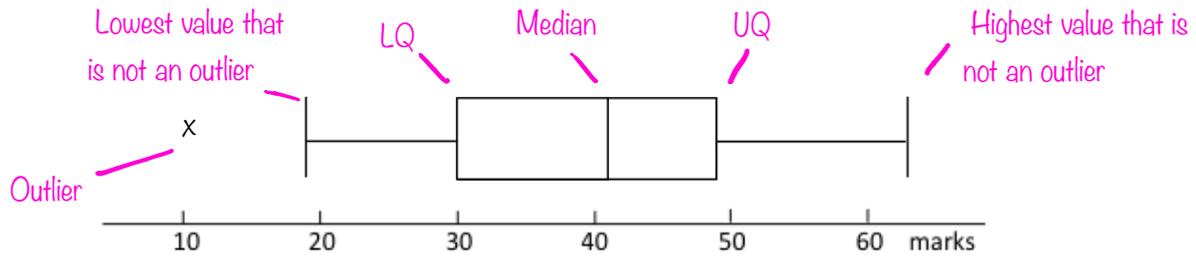
Q_2 : Median

Q_3 : Upper quartile

Inter quartile range (IQR) = $UQ - LQ$

Outlier is an extreme value and is 1.5 times the interquartile range above the upper quartile or below the lower quartile. ($< LQ - 1.5 IQR$ or $> UQ + 1.5 IQR$)

e) Box and whiskers plot



Chapter 2 : Permutation and combination

a) Permutation

Arrangement of n distinct items in online or at a time, order is important

Eg: STAR : $4! = 24$

Eg: AUGUST : $\frac{6!}{2!} = 360$

$2!$ ← Two identical letters

Eg: Arrangement for AUGUST with both of U together

UU _ _ _ _ = $5! = 120$ (Put both U together and count as one)

Eg: Arrangement for AUGUST with both of U not together

$\frac{6!}{2!} - 5! = 240$

Eg: Arrangement for NUMBER with both of U and E together

UE _ _ _ _ = $5! \times 2! = 240$ → U and E can swap places

Eg: Arrangement for CANADA if exactly 2 As are together

Normal - all together - all separate $\frac{6! - 4! - 403}{3!} = 92$

Eg: How many odd numbers can be formed using only three digits from 4,5,6,7 with no digit being used more than once

$\frac{5}{7} \quad 3P2 + 3P2 = 12$

Eg: Find how many numbers between 5000 and 6000 can be formed within the digits 1,2,3,4,5 and 6 if repeated digits are allowed

$5 \text{ _ _ _ _ } \quad 6P4 = 360$

b) Combination

Order of selection is not important

Eg : A committee of 5 people is to be chosen from 4 men and 6 women. William is one of the 4 men and Mary is one of the 6 women. Find the number of different committees that can be chosen if William and Mary refuse to be on the same committee together

$$\begin{array}{r}
 W \text{ ______ } \quad 8C4 \\
 M \text{ ______ } \quad 8C4 \\
 \text{______ } \quad 8C5
 \end{array}
 \qquad
 8C4 + 8C4 + 8C5 = 196$$

Eg : 3 letters from a nine letters of the word EVERGREEN are selected. find the number of selections which contains no Es.

$$\begin{array}{r}
 R \text{ ___ } \quad 3C2 \\
 R R \text{ __ } \quad 3C1 \\
 \text{____ } \quad 3C3
 \end{array}
 \qquad
 3C2 + 3C1 + 3C3 = 7$$

c) Probability

Eg : 4 students are to be selected from 3 female students and 5 male students . Find the probability that the chosen student consist of three male and one female

$$\begin{aligned}
 P(3M1F) &= \frac{5C3 \times 3C1}{8C4} \\
 &= 3/7
 \end{aligned}$$

Chapter 3 : Probability

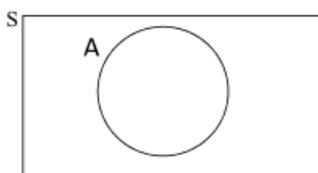
$$P(A) = \frac{n(A)}{n(S)} \leftarrow \begin{array}{l} \text{Number of outcomes of } A \\ \text{Total number of possible outcomes} \end{array}$$

a) Special events

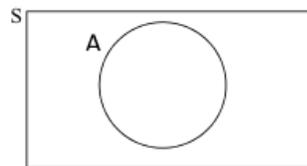
i) Complement of the event A

is denoted by A'

it means that A does not occur



Event A



Event A'

ii) Union of 2 events

is denoted by $A \cup B$

it means that the new set contains all elements that are in at least one of the two sets

iii) Intersection of 2 events

is denoted by $A \cap B$

it means that the new set contains all of the elements that are in both sets

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

b) Mutually exclusive

Two events are said to be mutually exclusive if both events cannot occur at the same time
If A and B are mutually exclusive then $P(A \cap B) = 0$ or $P(A \cup B) = P(A) + P(B)$

c) Conditional probability

$A|B$ means that event A occurs given that B has occurred

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

d) Independent events

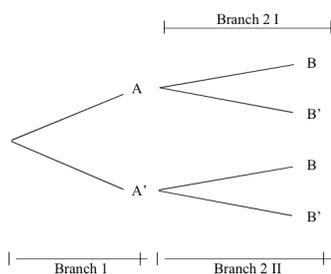
Occurrence or non-occurrence of either event does not effect the other event

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P(B)}{P(B)}$$

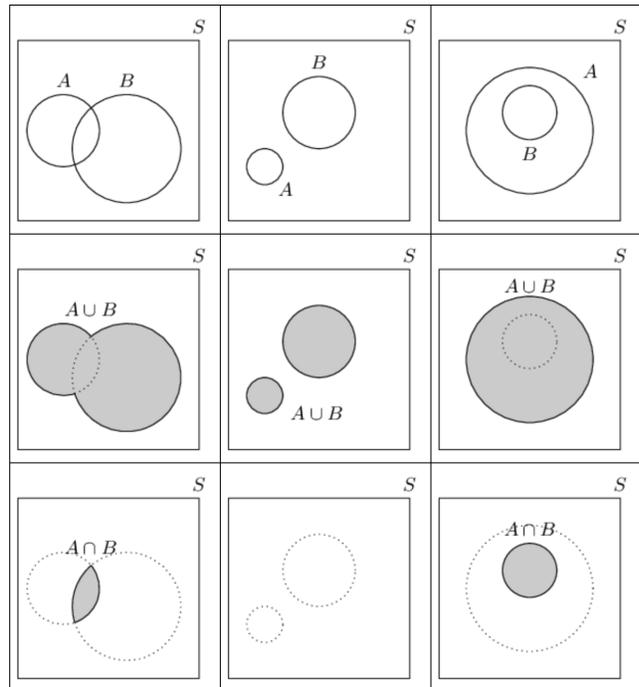
or

$$P(A \cap B) = P(A) \times P(B)$$

e) Tree diagram



Sum of each branch is 1



f) Probability with permutation and combination

$$\text{Probability} = \frac{\text{Restriction}}{\text{No restrictions}}$$

Chapter 4 : Discrete random variables

a) Probability distribution table

x	1	2	3	4
$P(X = x)$	a_1	a_2	a_3	a_4

$$\sum P(X = x) = 1$$

$$a_1 + a_2 + a_3 + a_4 = 1$$

b) Mean and variance

Mean is denoted by $E(X)$ and Variance is denoted by $\text{Var}(X)$

Eg:

x	-2	-1	0	1	2
$P(X = x)$	α	0.2	0.1	0.2	β

(a) Given that $E(X) = -0.2$, find the value of α and the value of β .

(b) Evaluate $\text{Var}(X)$.

$$-0.2 = -2\alpha - 0.2 + 0.2 + 2\beta$$

$$\alpha + 0.2 + 0.1 + 0.2 + \beta = 1$$

$$\alpha = 0.5 - \beta$$

$$-2(0.5 - \beta) - 0.2 + 0.2 + 2\beta = -0.2$$

$$-1 + 4\beta = -0.2$$

$$\beta = 0.2$$

$$\alpha = 0.3$$

c) Binomial distribution

$X \sim B(n, p)$

n is number of trials and p is probability of success at trial

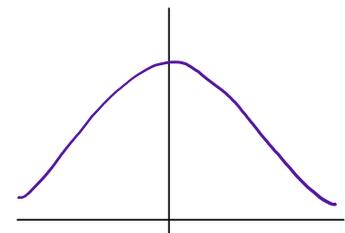
$$P(X = x) = {}^n C_x P^x q^{n-x} ; q = 1 - p$$

$\text{Mean} = np \text{ and Variance} = npq$

Chapter 5 : Normal distribution

a) Standard normal distribution

A normal distribution is symmetrical about its mean



Eg: $X \sim N(39, 3.4^2)$

$$P(X > 45) = P\left(Z > \frac{45 - 39}{3.4}\right)$$

$$P(X < 45) = P\left(Z < \frac{45 - 39}{3.4}\right)$$

b) Approximation

$$P(X \geq 45) = P\left(Z > \frac{44.5 - 39}{3.4}\right)$$

$$P(X \leq 45) = P\left(Z < \frac{45.5 - 39}{3.4}\right)$$

Conditions for approximating binomial to normal : n is large, $np > 5$ and $nq > 5$



1st Edition : Zhi Yee

Producer : Mr. Sai Mun

Copyright by Vbest. All rights reserved. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, without the prior written permission of the publisher. Legal actions could be taken if there is a breach in copyrights of the notes.

Although we make strong efforts to ensure that all information is accurate at time of publication, VBest cannot guarantee that all information on this website or in these notes are always correct, complete, or up to date.