

Histogram (1A)

- Absolute Frequency
- Relative Frequency
- Accumulative Frequency
- Average
- Normalization

Copyright (c) 2008 Young W. Lim.

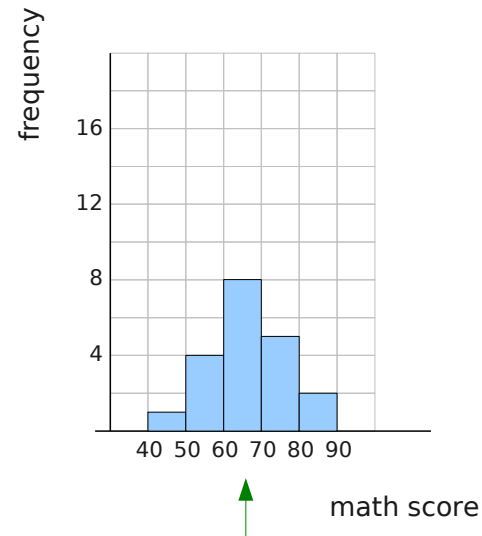
Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

Please send corrections (or suggestions) to youngwlim@hotmail.com.

This document was produced by using OpenOffice and Octave.

Absolute Frequency

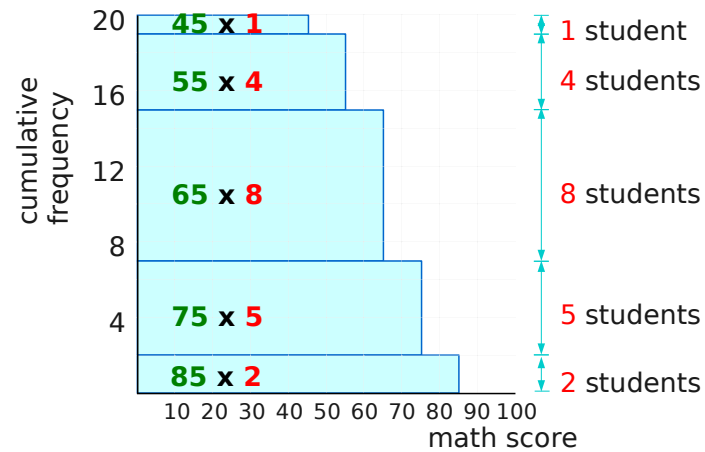
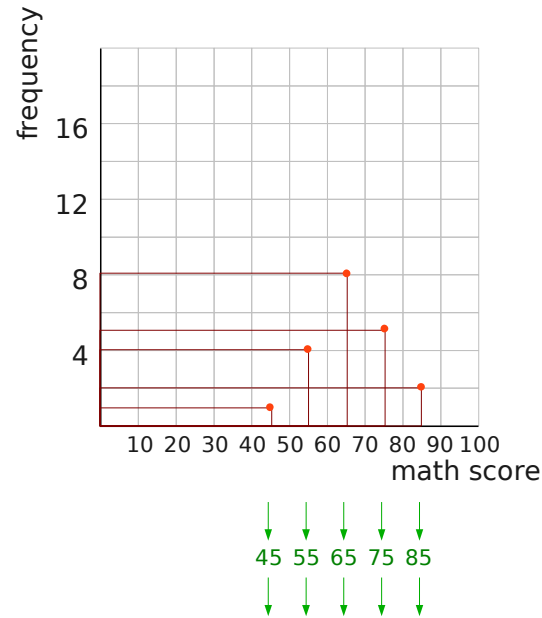
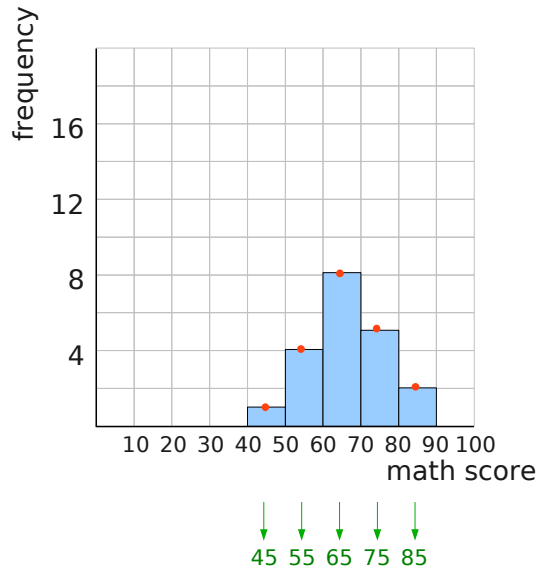
math score	# of students
40 ~ 50	1
50 ~ 60	4
60 ~ 70	8
70 ~ 80	5
80 ~ 90	2
Total	20



Average Score?

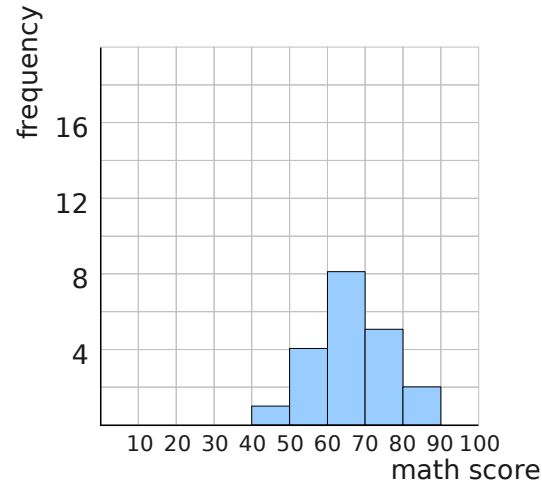
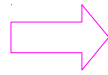
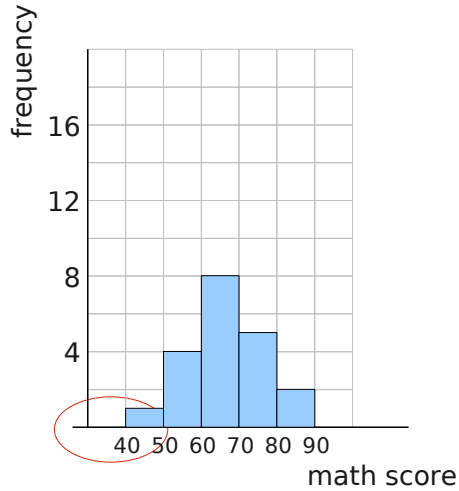
$$\frac{45 \times 1 + 55 \times 4 + 65 \times 8 + 75 \times 5 + 85 \times 2}{1 + 4 + 8 + 5 + 2} = 66.5$$

Absolute Frequency & Average - Rectangles

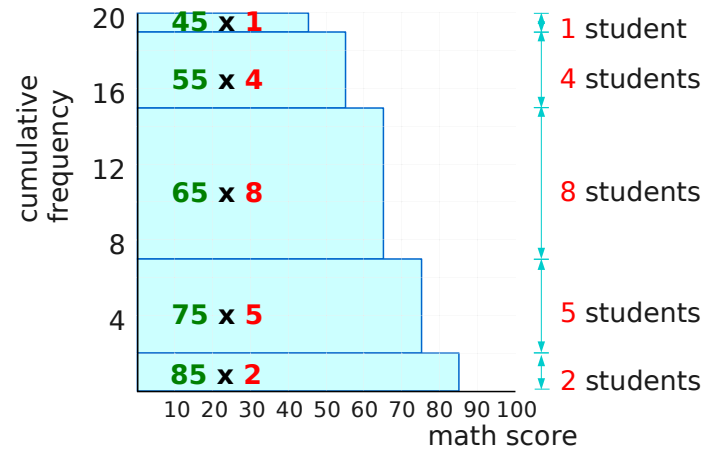
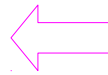
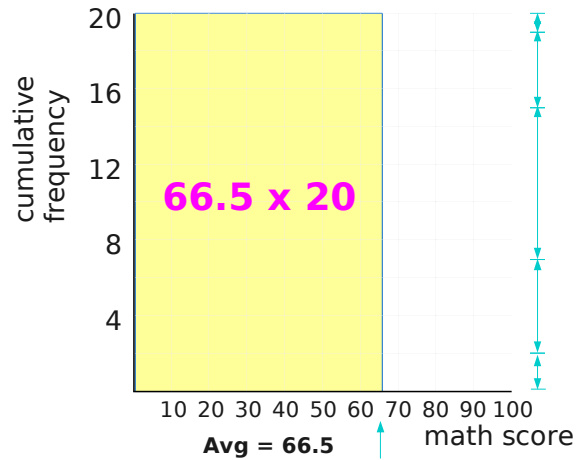


Absolute Frequency & Average

Absolute Frequency



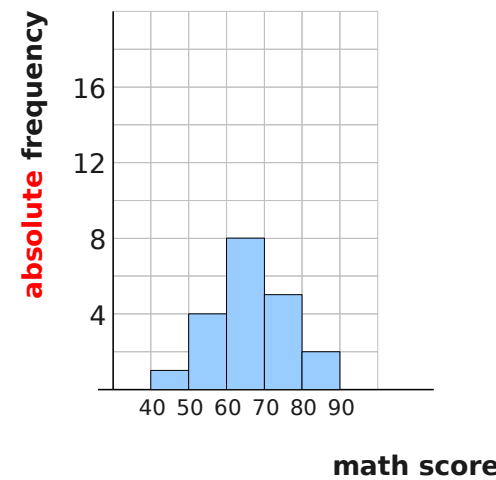
45 55 65 75 85



Comparison of Histograms - Absolute Frequency

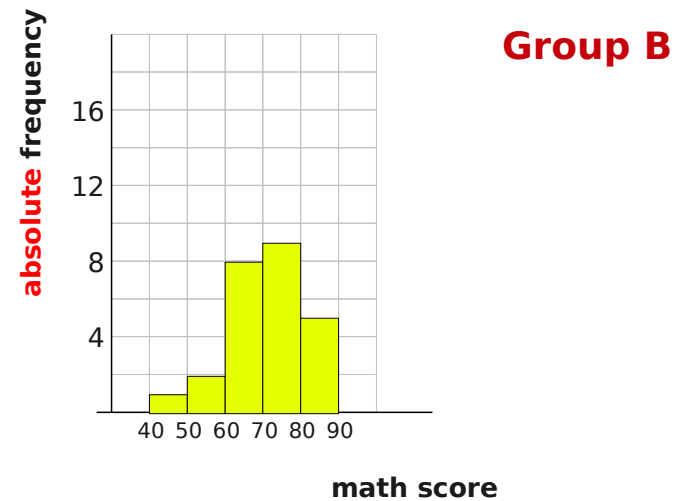
Group A

math score	# of students
40 ~ 50	1
50 ~ 60	4
60 ~ 70	8
70 ~ 80	5
80 ~ 90	2
Total	20



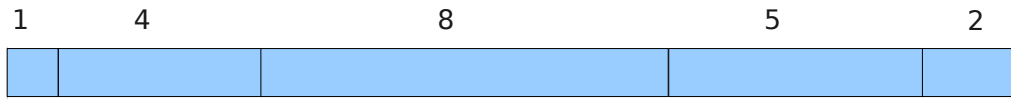
Group B

math score	# of students
40 ~ 50	1
50 ~ 60	2
60 ~ 70	8
70 ~ 80	9
80 ~ 90	5
Total	25

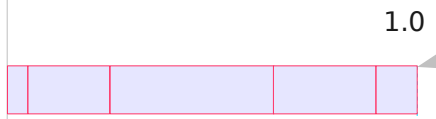


Normalization

Group A (20 students)



Total 20 student counts

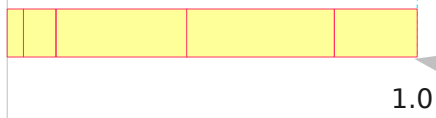


A common scale makes comparison convenient

Normalization

$$f_i = \frac{n_i}{N} = \frac{n_i}{\sum_k n_k}$$

f_i relative count
 n_i (absolute) count
 N total ($= \sum_j n_j$) counts



Total 25 student counts

Group B (25 students)



Relative Frequency

Normalization

$$f_i = \frac{n_i}{N} = \frac{n_i}{\sum_i n_i}$$

Group A	n_i	f_i
math score	absolute frequency	relative frequency
40 ~ 50	1	1/20 = 0.05
50 ~ 60	4	4/20 = 0.20
60 ~ 70	8	8/20 = 0.40
70 ~ 80	5	5/20 = 0.25
80 ~ 90	2	2/20 = 0.10
Total	20	20/20 = 1.00

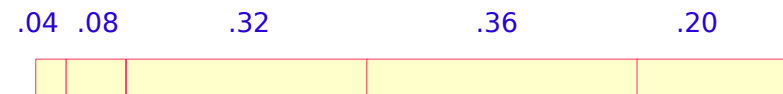
$$N = \sum_i n_i = 20$$



$$\sum_i f_i = 1.00$$

Group B	n_i	f_i
math score	absolute frequency	relative frequency
40 ~ 50	1	1/25 = 0.04
50 ~ 60	2	2/25 = 0.08
60 ~ 70	8	8/25 = 0.32
70 ~ 80	9	9/25 = 0.36
80 ~ 90	5	5/25 = 0.20
Total	25	25/25 = 1.00

$$N = \sum_i n_i = 25$$



$$\sum_i f_i = 1.00$$

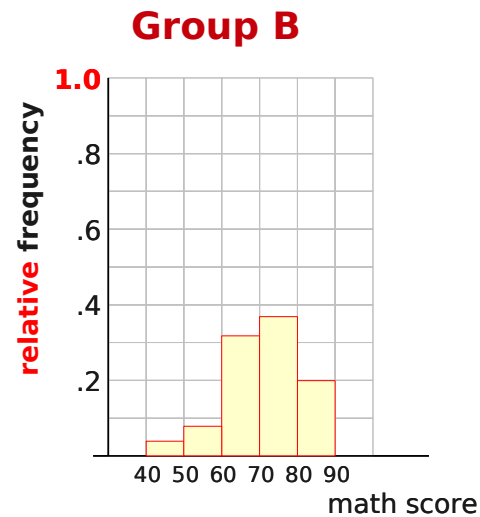
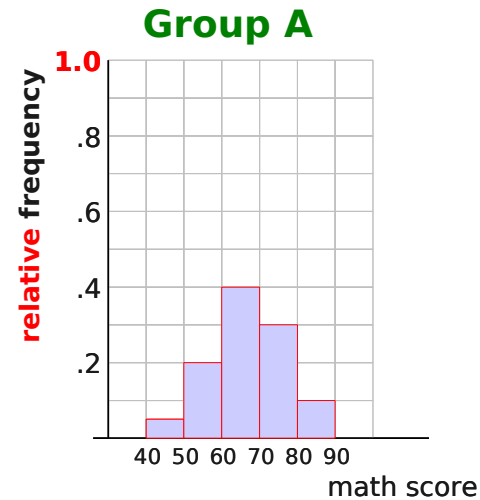
Comparison of Histograms - Relative Frequency

Group A	n_i	f_i
math score	absolute frequency	relative frequency
40 ~ 50	1	$1/20 = 0.05$
50 ~ 60	4	$4/20 = 0.20$
60 ~ 70	8	$8/20 = 0.40$
70 ~ 80	5	$5/20 = 0.25$
80 ~ 90	2	$2/20 = 0.10$
Total	20	$20/20 = 1.00$

$$N = \sum_i n_i = 20$$

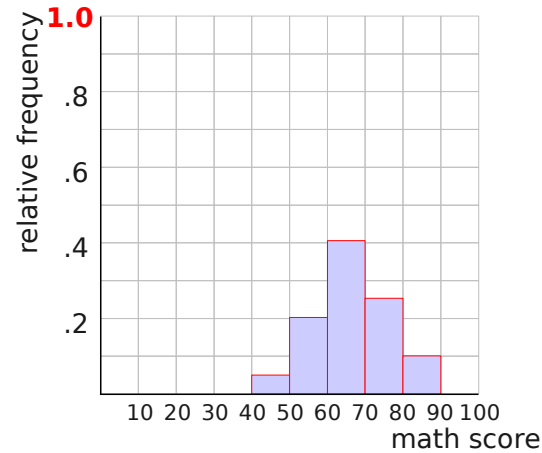
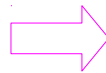
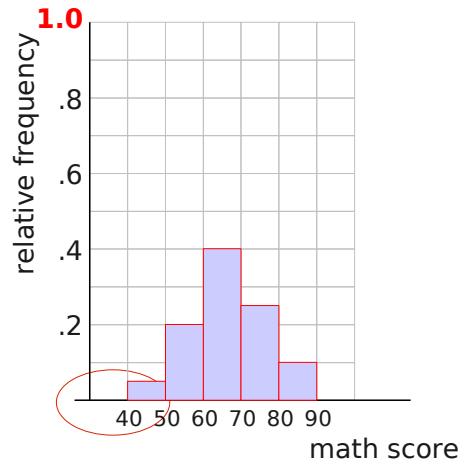
Group B	n_i	f_i
math score	absolute frequency	relative frequency
40 ~ 50	1	$1/25 = 0.04$
50 ~ 60	2	$2/25 = 0.08$
60 ~ 70	8	$8/25 = 0.32$
70 ~ 80	9	$9/25 = 0.36$
80 ~ 90	5	$5/25 = 0.20$
Total	25	$25/25 = 1.00$

$$N = \sum_i n_i = 25$$

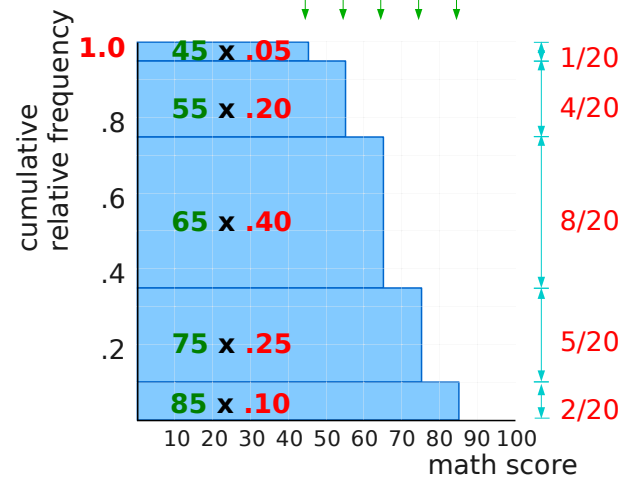
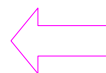
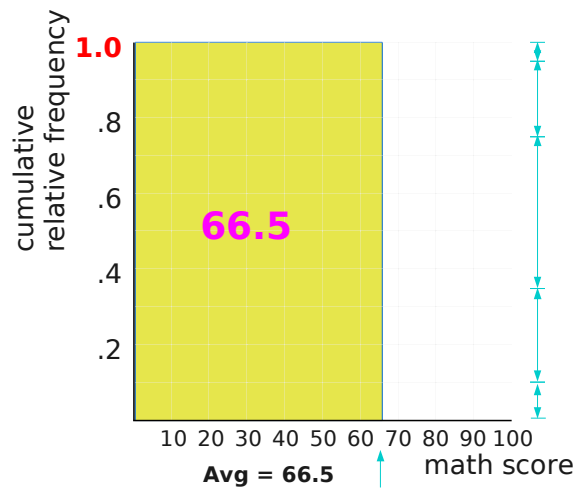


Relative Frequency & Average

Relative Frequency **Group A**



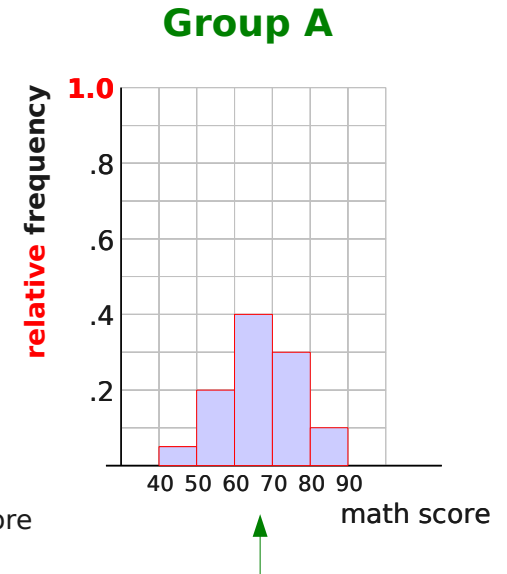
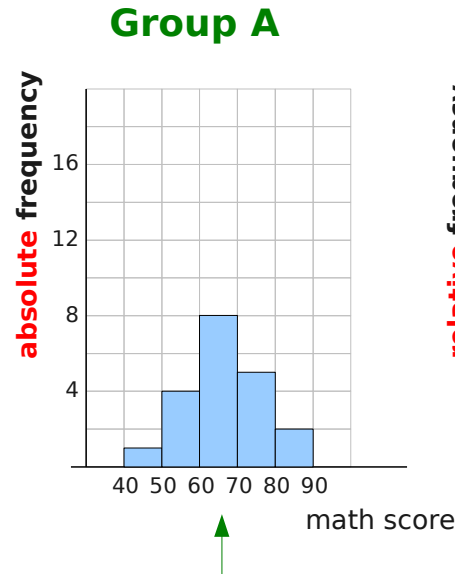
45 55 65 75 85



Average

Group A	n_i	f_i
math score	absolute frequency	relative frequency
40 ~ 50	1	$1/20 = 0.05$
50 ~ 60	4	$4/20 = 0.20$
60 ~ 70	8	$8/20 = 0.40$
70 ~ 80	5	$5/20 = 0.25$
80 ~ 90	2	$2/20 = 0.10$
Total	20	$20/20 = 1.00$

$$N = \sum_i n_i = 20$$



Average Score?

$$\frac{45 \times 1 + 55 \times 4 + 65 \times 8 + 75 \times 5 + 85 \times 2}{1 + 4 + 8 + 5 + 2}$$

$$= \frac{45 \times 1}{20} + \frac{55 \times 4}{20} + \frac{65 \times 8}{20} + \frac{75 \times 5}{20} + \frac{85 \times 2}{20}$$

$$= 45 \times \frac{1}{20} + 55 \times \frac{4}{20} + 65 \times \frac{8}{20} + 75 \times \frac{5}{20} + 85 \times \frac{2}{20} = 66.5$$

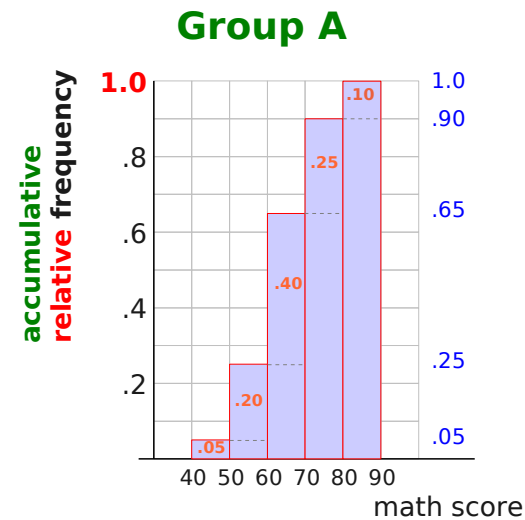
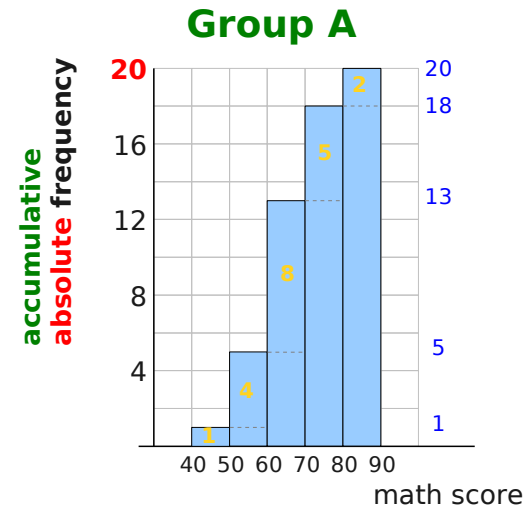
Accumulative Frequency

Group A		n_i
math score	absolute frequency	accumulative absolute frequency
40 ~ 50	1	1
50 ~ 60	4	5
60 ~ 70	8	13
70 ~ 80	5	18
80 ~ 90	2	20
Total	20	20

$$N = \sum_i n_i = 20$$

Group A		f_i
math score	relative frequency	accumulative relative frequency
40 ~ 50	0.05	0.05
50 ~ 60	0.20	0.25
60 ~ 70	0.40	0.65
70 ~ 80	0.25	0.90
80 ~ 90	0.10	1.00
Total	1.00	1.00

$$1.0 = \sum_i f_i = 25$$

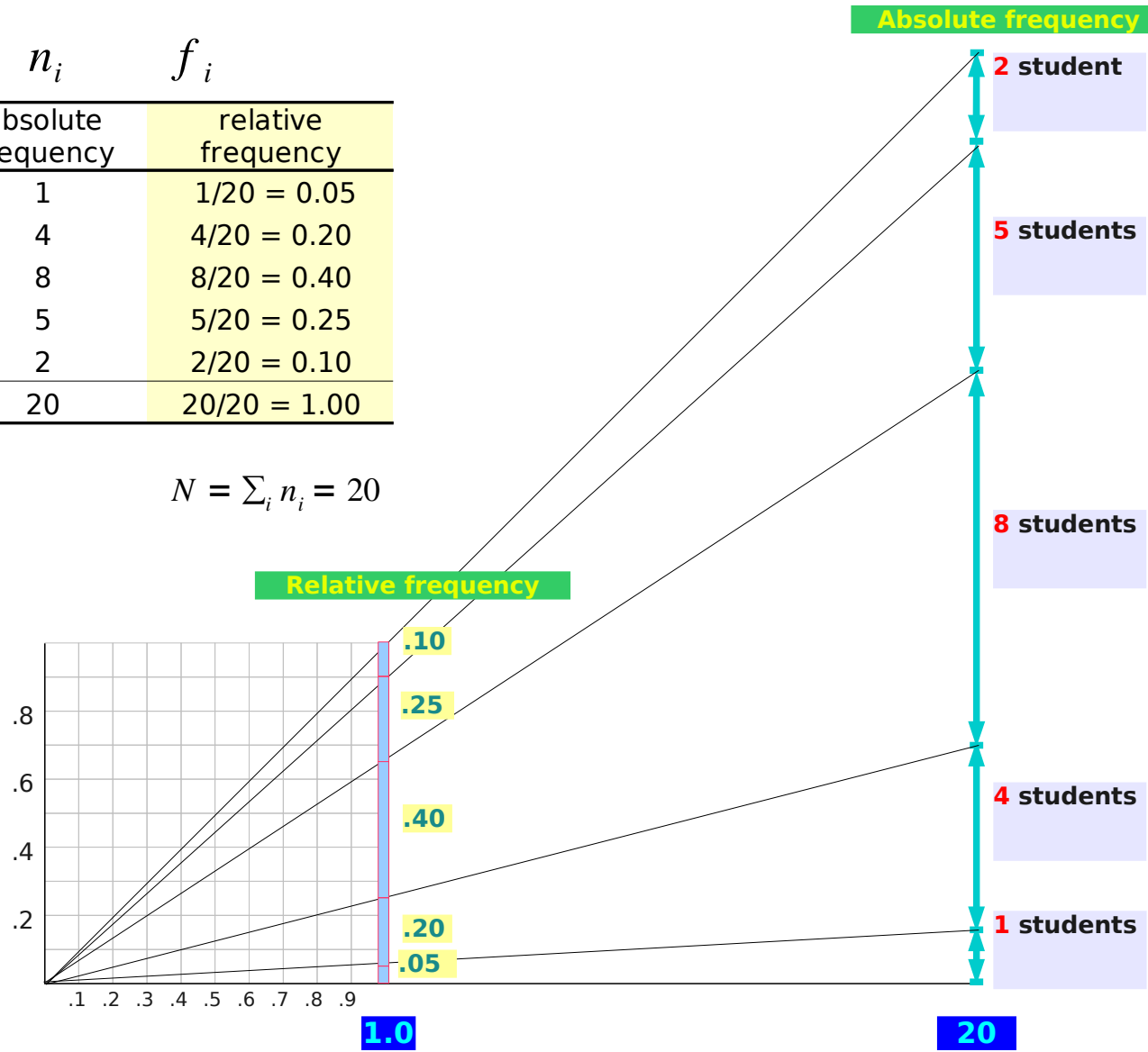


Relative Frequency & Ratio

Group A

math score	n_i absolute frequency	f_i relative frequency
40 ~ 50	1	$1/20 = 0.05$
50 ~ 60	4	$4/20 = 0.20$
60 ~ 70	8	$8/20 = 0.40$
70 ~ 80	5	$5/20 = 0.25$
80 ~ 90	2	$2/20 = 0.10$
Total	20	$20/20 = 1.00$

$$N = \sum_i n_i = 20$$



Normalization Equations (1)

$$f_i = \frac{n_i}{N}$$

relative frequency = $\frac{\text{count}}{\text{total count}}$

fraction = $\frac{\text{part}}{\text{total}}$

0.xxx = $\frac{\triangle\triangle}{\square\square\square\square}$

$$N = \frac{n_i}{f_i}$$

total count = $\frac{\text{count}}{\text{relative freq}}$

total = $\frac{\text{part}}{\text{fraction}}$

$\square\square\square\square$ = $\frac{\triangle\triangle}{0.xxx}$

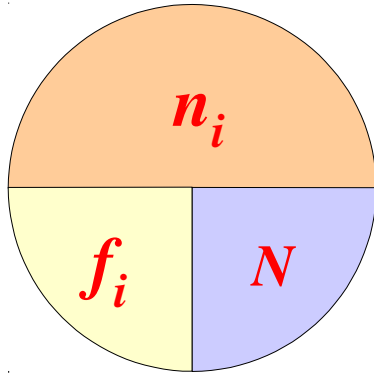
$$n_i = N \cdot f_i$$

count = total count \cdot relative frequency

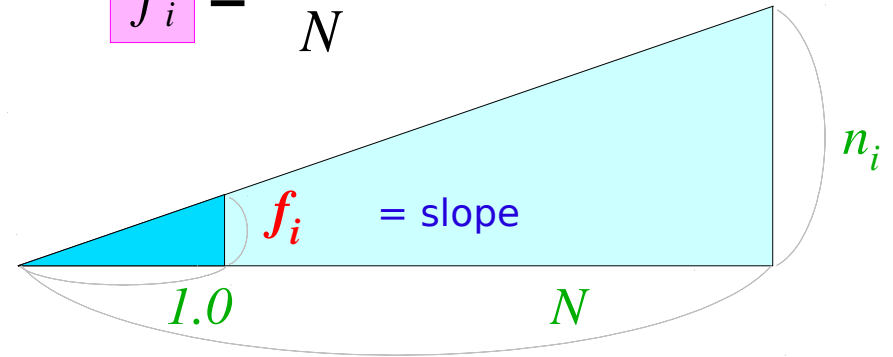
part = total \cdot fraction

$\triangle\triangle$ = $\square\square\square\square \cdot 0.xxx$

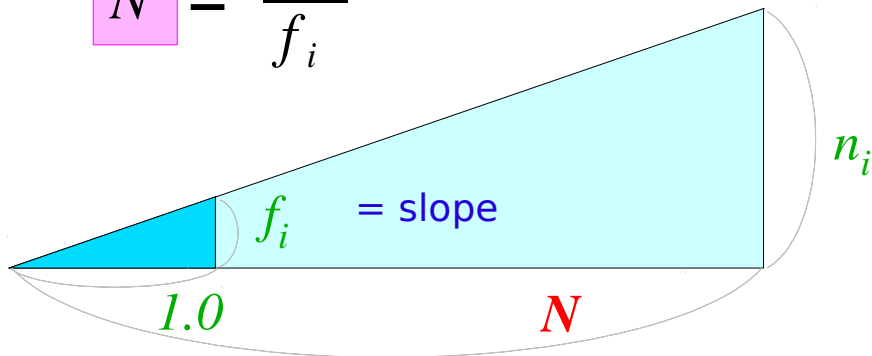
Normalization Equations (2)



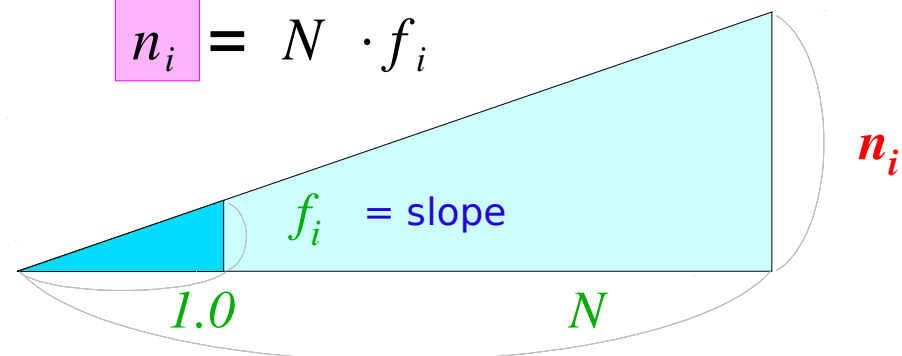
$$f_i = \frac{n_i}{N}$$



$$N = \frac{n_i}{f_i}$$



$$n_i = N \cdot f_i$$



References

- [1] <http://en.wikipedia.org/>
- [2] <http://planetmath.org/>
- [3] “최상위수학” 디딤돌
- [4] “센 수학”
- [5] “개념과 유형” 비상