Expected Value (2A)

Young Won Lim
3/5/18

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

This document was produced by using LibreOffice and Octave.

Expected Value

Suppose random variable X can take
value x_1 with probability p_1,
value x_2 with probability p_2, and so on, up to
value x_k with probability p_k.

Then the expectation of this random variable X is defined as

$$\mathrm{E}[X] = x_1 p_1 + x_2 p_2 + \cdots + x_k p_k.$$

Since all probabilities pi add up to one (p_1 + p_2 + ... + p_k = 1),
the expected value can be viewed as the weighted average,
with p_i's being the weights:

$$\mathrm{E}[X] = \frac{x_1 p_1 + x_2 p_2 + \cdots + x_k p_k}{1} = \frac{x_1 p_1 + x_2 p_2 + \cdots + x_k p_k}{p_1 + p_2 + \cdots + p_k}.$$

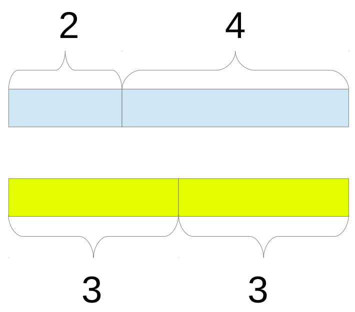
Arithmetic Mean

2 elements $\{a, b\}$ $$A = \frac{(a+b)}{2}$$

3 elements $\{a, b, c\}$ $$A = \frac{(a+b+c)}{3}$$

n elements $\{a_1, a_2, \ldots, a_n\}$ $$A = \frac{(a_1+a_2+\cdots+a_n)}{n} = \frac{1}{n}\sum_{i=1}^{n} a_i$$

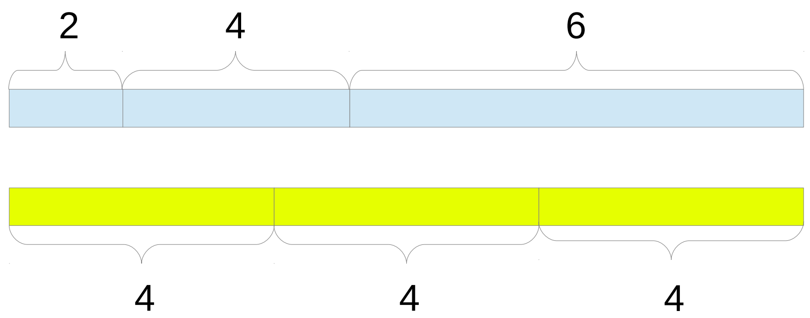
Arithmetic Mean - Example



same length

$$(2+4)=(3+3)=2\cdot 3=2\cdot A$$

Arithmetic Mean: A = 3



same length

$$(2+4+6)=(4+4+4)=3\cdot 4=3\cdot A$$

Arithmetic Mean: A = 4

Geometric Mean

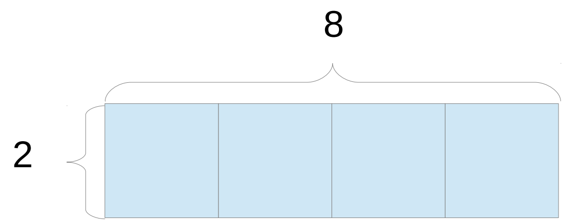
2 elements $\{a, b\}$ $\qquad G = \sqrt{a \cdot b}$ $\qquad (a > 0,\ b > 0)$

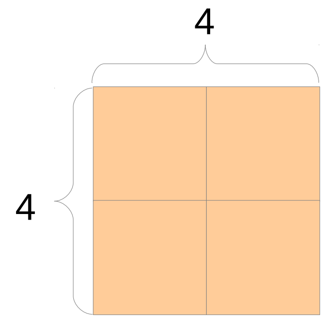
3 elements $\{a, b, c\}$ $\qquad G = \sqrt[3]{a \cdot b \cdot c}$ $\qquad (a > 0,\ b > 0,\ c > 0)$

n elements $\{a_1, a_2, \ldots, a_n\}$ $\qquad G = \sqrt[n]{a_1 \cdot a_2 \cdots a_n} = \left(\prod_{i=1}^{n} a_i\right)^{\frac{1}{n}}$

$(a_i > 0)$

Geometric Mean - Example

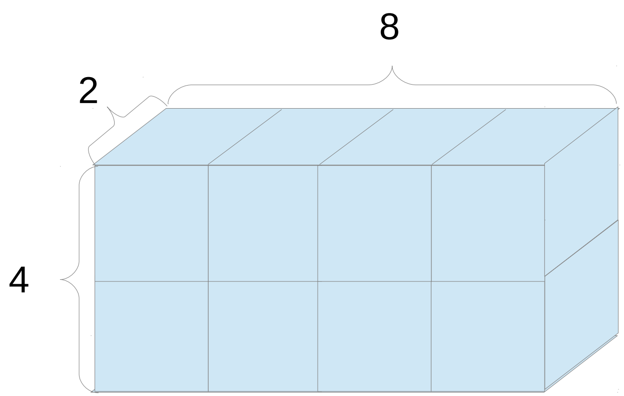


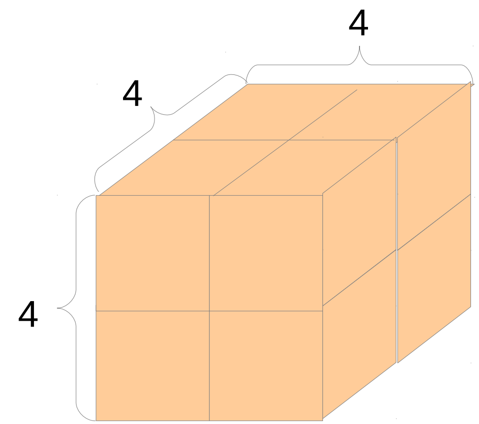


$$(2 \cdot 8) = (4 \cdot 4) = 4^2 = G^2$$

same area

Geometric Mean: G = 4





$$(2 \cdot 4 \cdot 8) = (4 \cdot 4 \cdot 4) = 4^3 = G^3$$

same volume

Geometric Mean: G = 4

Harmonic Mean

2 elements (a, b)

$$H = \frac{2}{\left(\frac{1}{a} + \frac{1}{b}\right)} = \frac{2ab}{a+b}$$

3 elements (a, b, c)

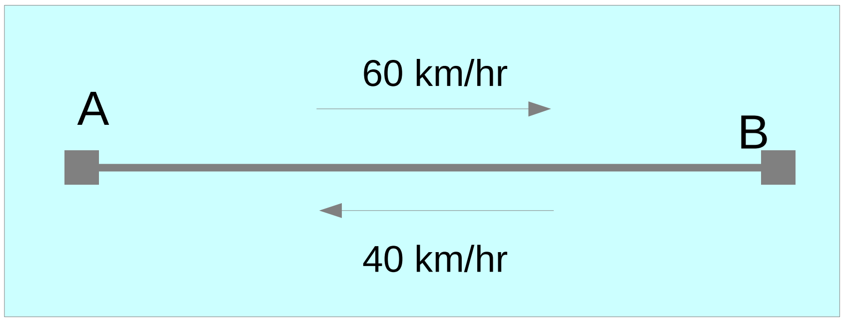
$$H = \frac{3}{\left(\frac{1}{a} + \frac{1}{b} + \frac{1}{c}\right)} = \frac{3abc}{ab+bc+ca}$$

n elements $(a_1, a_2, \ldots, a_n)$

$$H = \frac{n}{\left(\frac{1}{a_1} + \frac{1}{a_2} + \cdots + \frac{1}{a_n}\right)} = \frac{n}{\sum_{i=1}^{n} \frac{1}{a_i}}$$

Harmonic Mean

same distance = D



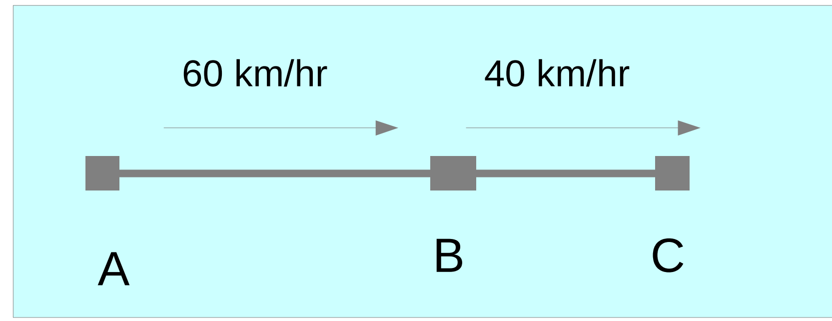
Average Speed

$$= \frac{(\text{total distance})}{(\text{total time})}$$

$$= \frac{2 \cdot D}{\frac{D}{60} + \frac{D}{40}} = \frac{2 \cdot 60 \cdot 40}{60 + 40} = 48\,\text{km/hr}$$

Harmonic Mean: H = 48

same duration = T



Average Speed

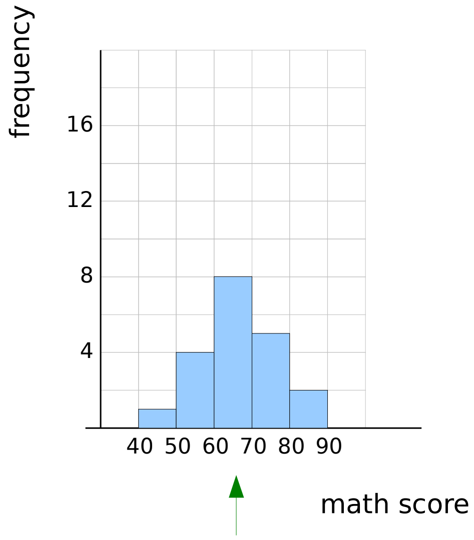
$$= \frac{(\text{total distance})}{(\text{total time})}$$

$$= \frac{60 \cdot T + 40 \cdot T}{2T} = 50\,\text{km/hr}$$

Arithmetic Mean: A = 50

Absolute Frequency

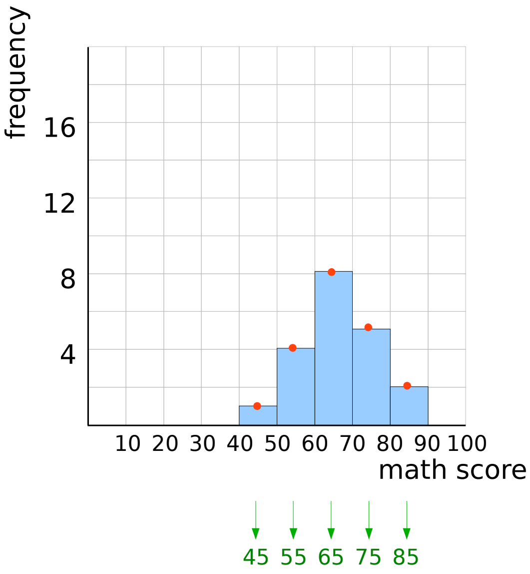
math score	# of students
40 ~ 50	1
50 ~ 60	4
60 ~ 70	8
70 ~ 80	5
80 ~ 90	2
Total	20

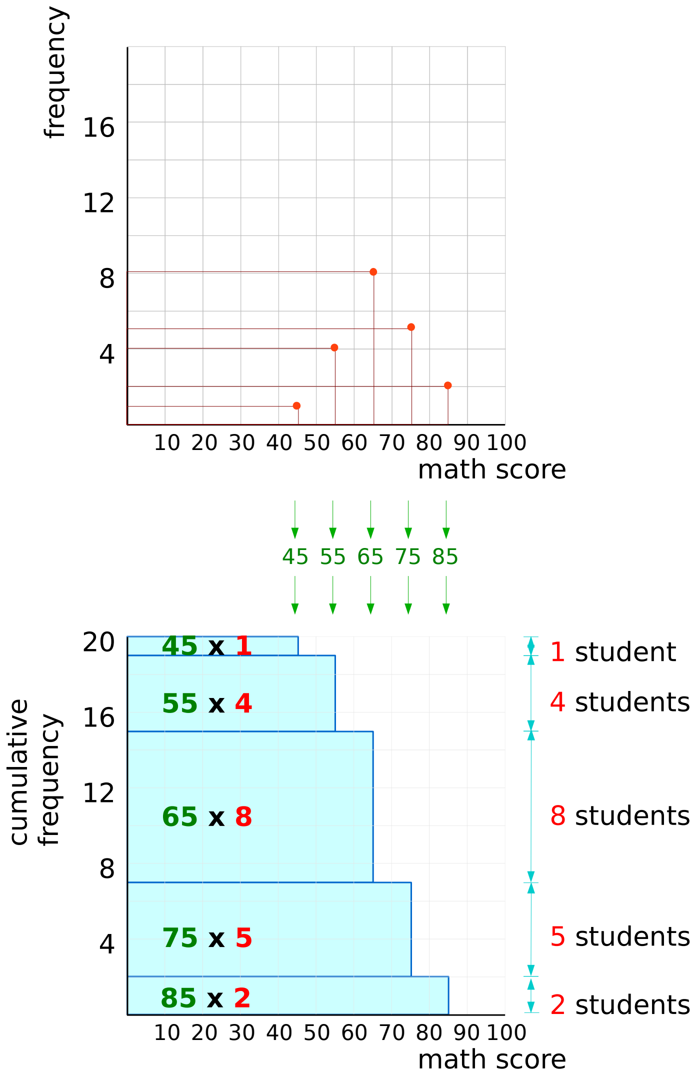


Average Score?

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

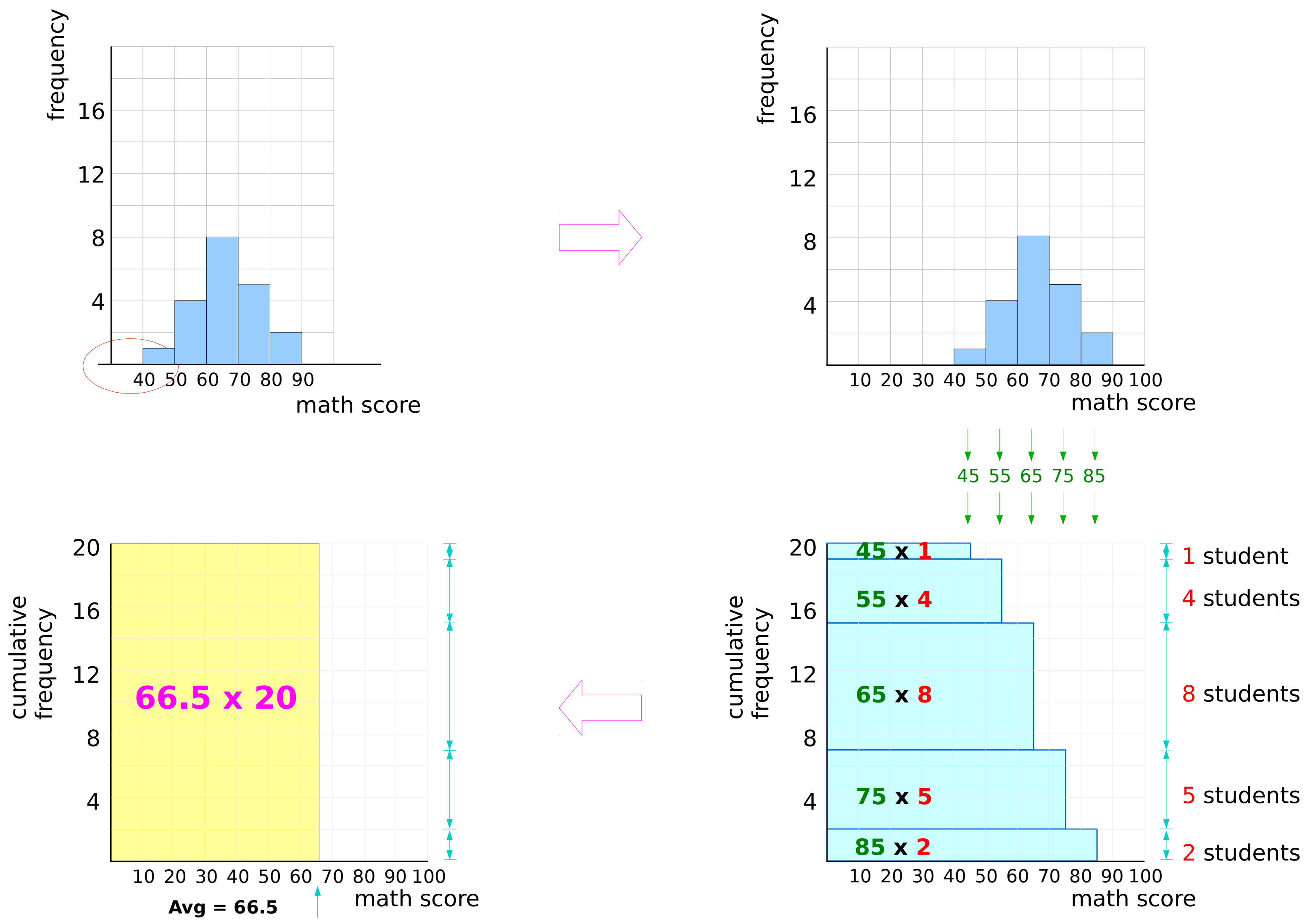
Absolute Frequency & Average - Rectangles





Absolute Frequency & Average

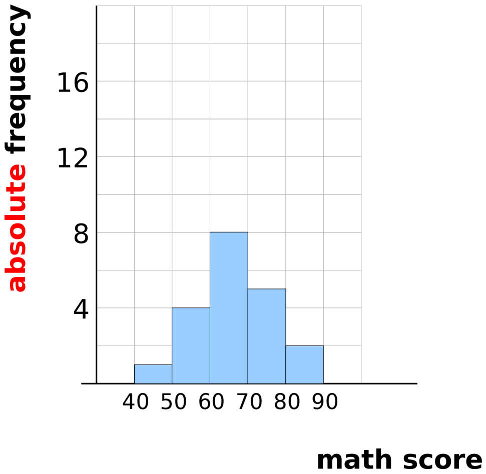
Absolute Frequency



Comparison of Histograms – Absolute Frequency

Group A

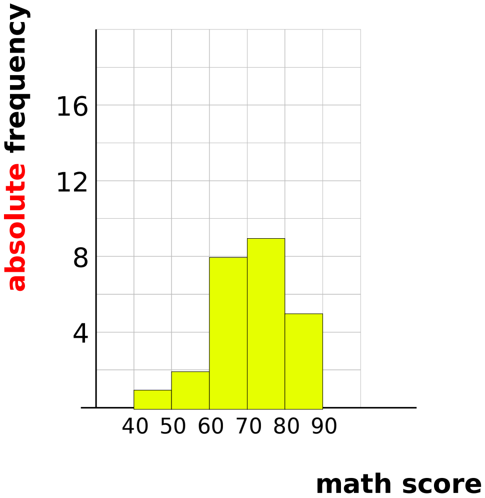
math score	# of students
40 ~ 50	1
50 ~ 60	4
60 ~ 70	8
70 ~ 80	5
80 ~ 90	2
Total	20



Group A

Group B

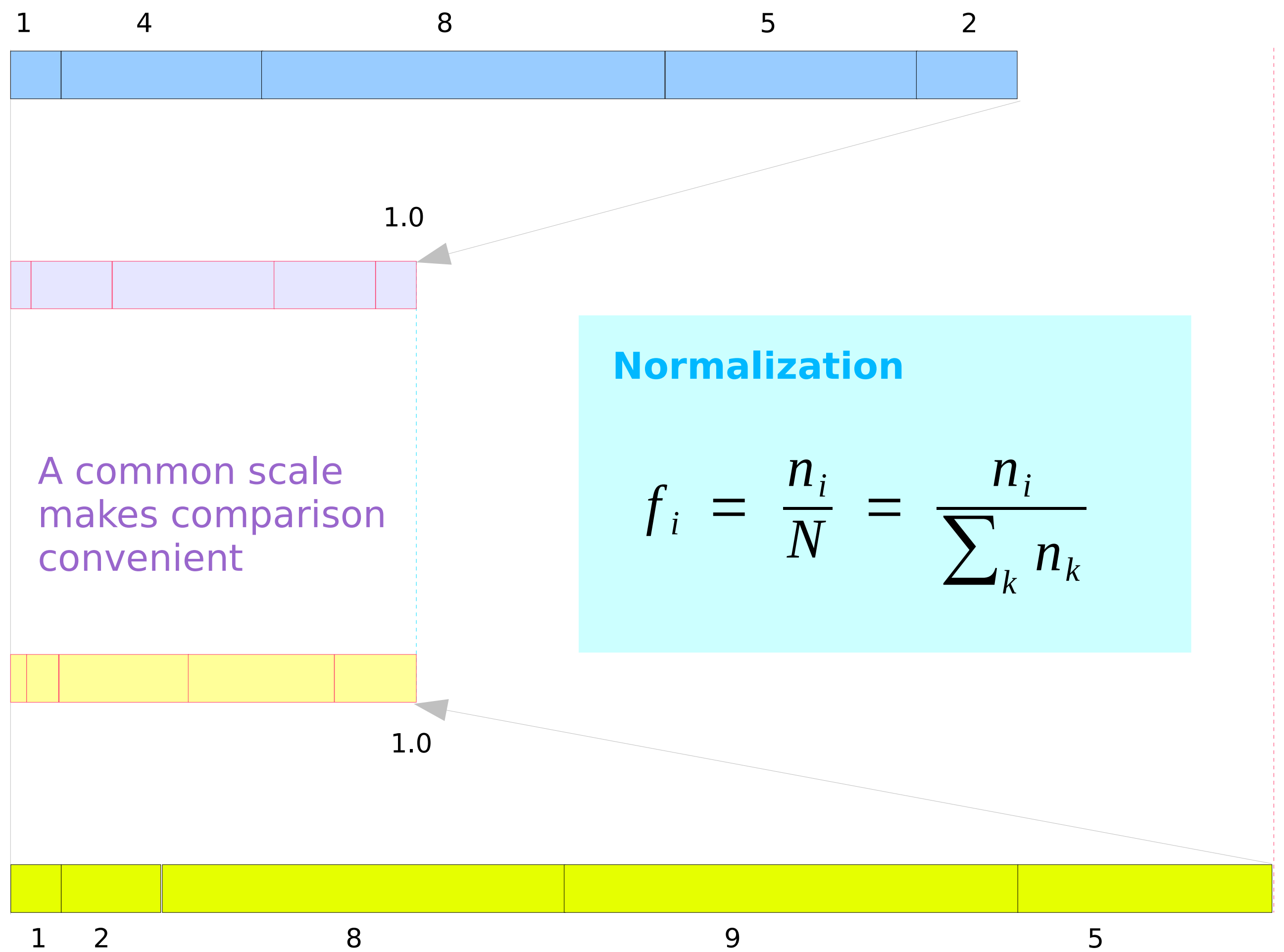
math score	# of students
40 ~ 50	1
50 ~ 60	2
60 ~ 70	8
70 ~ 80	9
80 ~ 90	5
Total	25



Group B

Normalization

Group A (20 students)



Group B (25 students)

Normalization

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

Total 20 student counts

f_i relative count

n_i (absolute) count

N total $\left(= \sum_j n_j \right)$ counts

Total 25 student counts

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

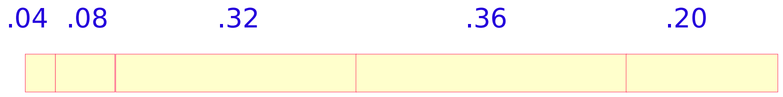
$$N = \sum_i n_i = 20$$

$$\sum_i f_i = 1.00$$

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

$$N = \sum_i n_i = 25$$



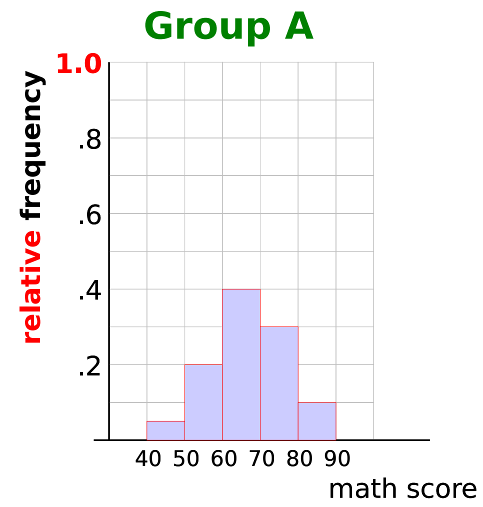
$$\sum_i f_i = 1.00$$

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

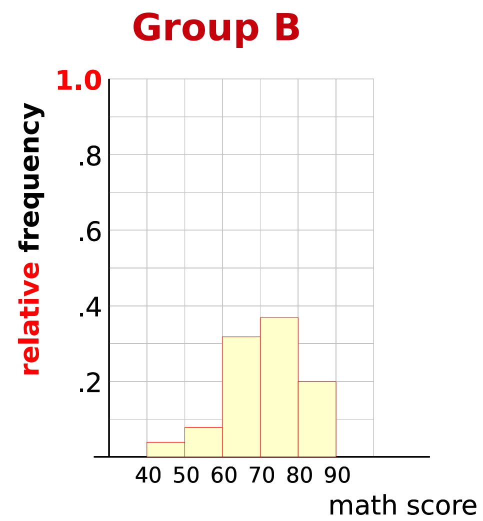
$$N = \sum_i n_i = 20$$



Group B n_i f_i

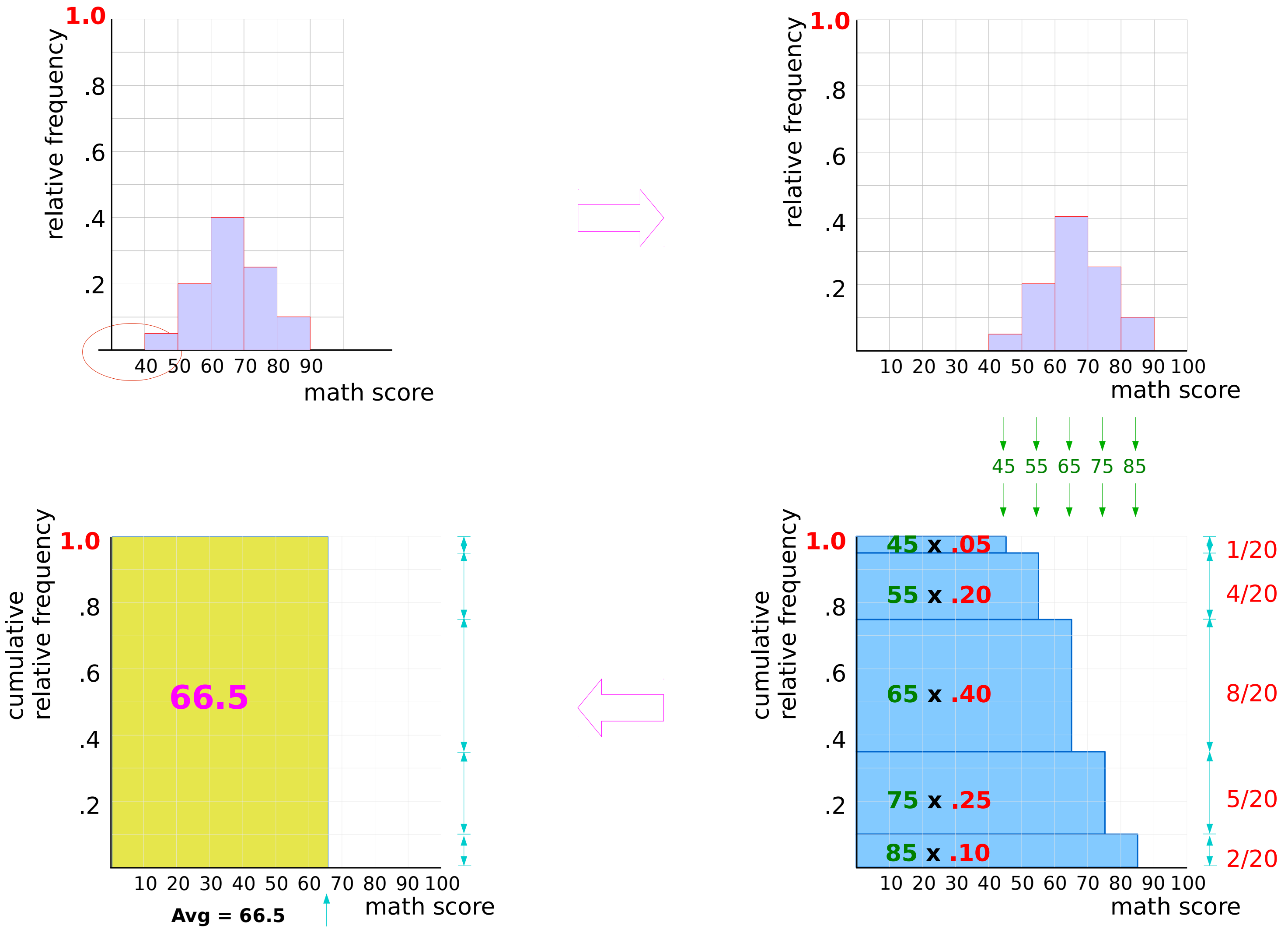
math score	absolute frequency	relative frequency
40 ~ 50	1	1/25 = 0.04
50 ~ 60	2	2/25 = 0.08
60 ~ 70	8	8/25 = 0.32
70 ~ 80	9	9/25 = 0.36
80 ~ 90	5	5/25 = 0.20
Total	25	25/25 = 1.00

$$N = \sum_i n_i = 25$$



Relative Frequency & Average

Relative Frequency **Group A**

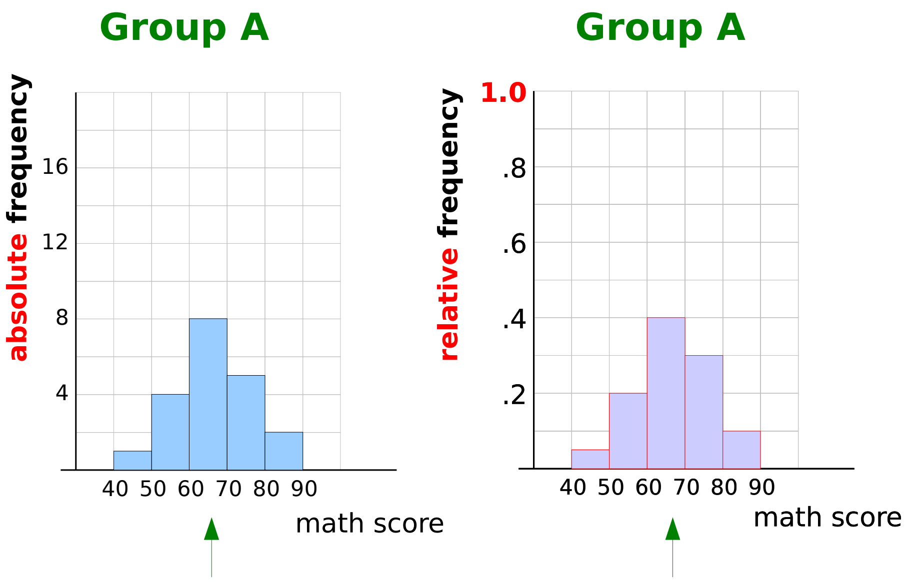


Young Won Lim
3/5/18

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?

$$\frac{45\times1 + 55\times4 + 65\times8 + 75\times5 + 85\times2}{1+4+8+5+2}$$

$$= \frac{45\times1}{20} + \frac{55\times4}{20} + \frac{65\times8}{20} + \frac{75\times5}{20} + \frac{85\times2}{20}$$

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

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

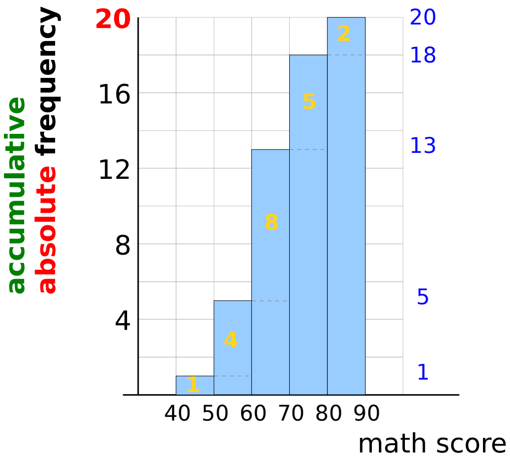
$N = \sum_i n_i = 20$

Group A f_i

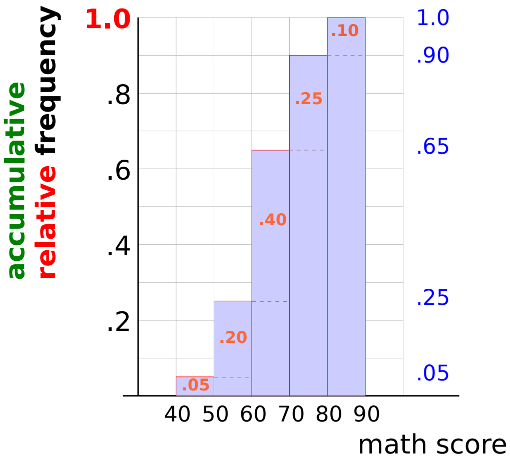
math score	relative frequency	**accumulative** relative frequency
40 ~ 50	0.05	0.05
50 ~ 60	0.20	0.25
60 ~ 70	0.40	0.65
70 ~ 80	0.25	0.90
80 ~ 90	0.10	1.00
Total	1.00	1.00

$1.0 = \sum_i f_i = 25$

Group A



Group A

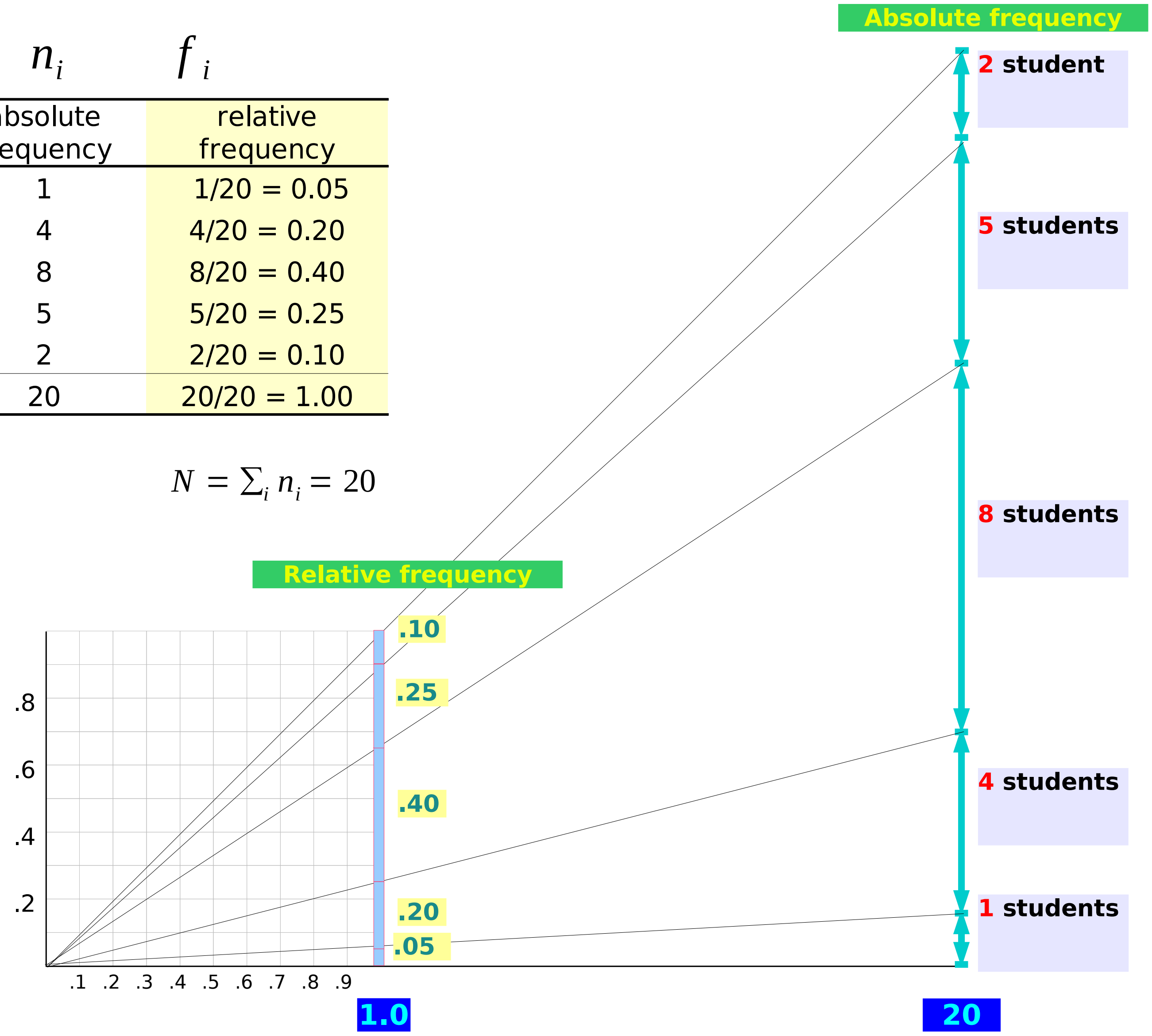


Relative Frequency & Ratio

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

$$N = \sum_i n_i = 20$$



Normalization Equations (1)

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

relative frequency = count / total count

fraction = part / total

0.xxx = △△ / □□□□□

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

total count = count / relative freq

total = part / fraction

□□□□□ = △△ / 0.xxx

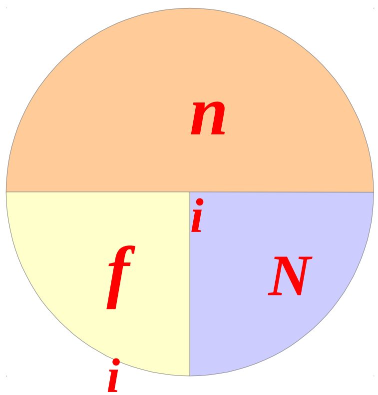
$$n_i = N \cdot f_i$$

count = total count • relative frequency

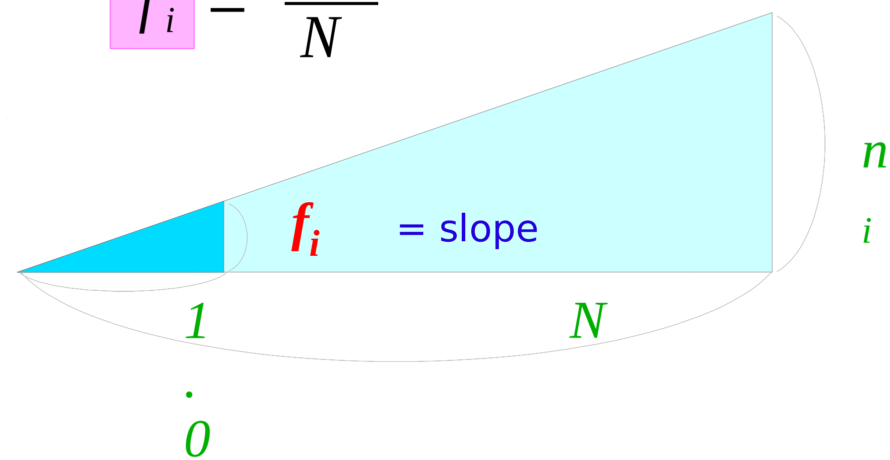
part = total • part

△△ = □□□□□ • 0.xxx

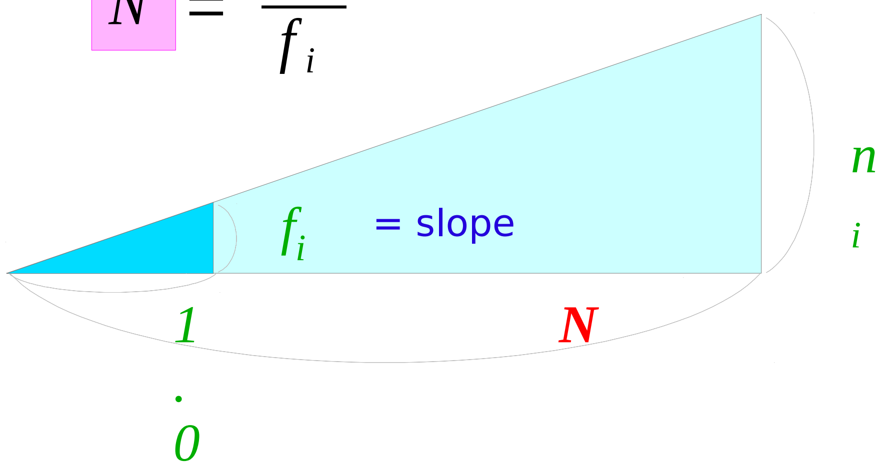
Normalization Equations (2)



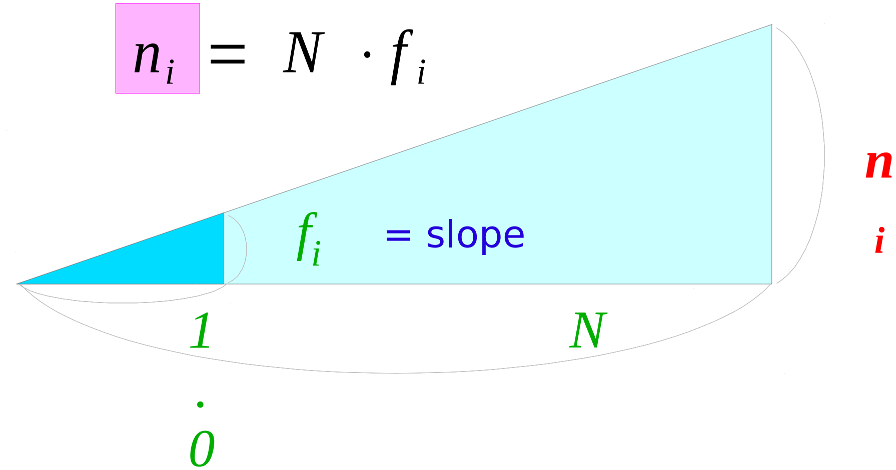
$$f_i = \frac{n_i}{N}$$



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

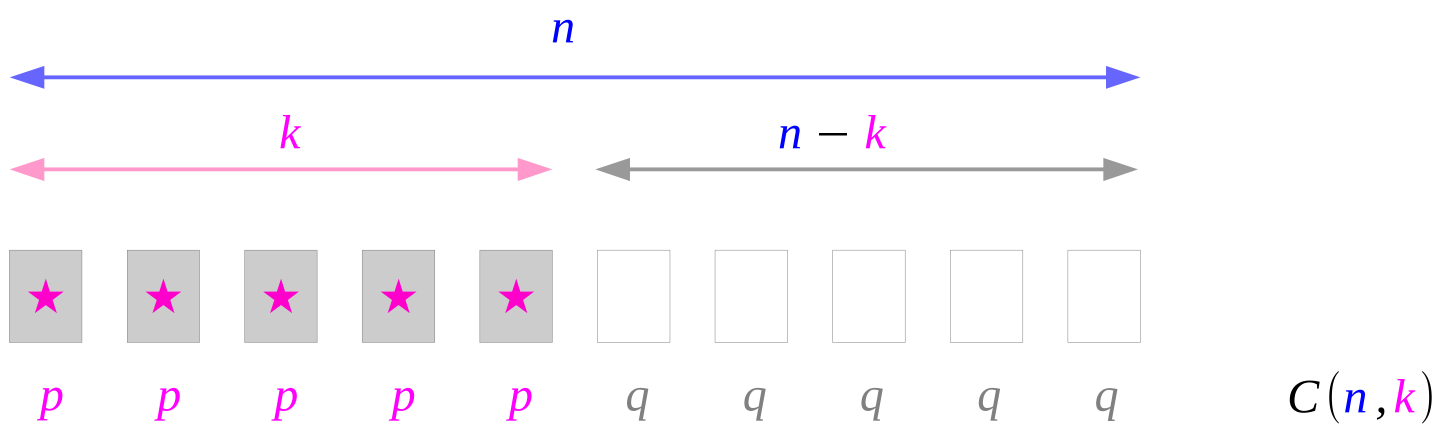


$$n_i = N \cdot f_i$$



Binomial Probability Distribution

$$f(k;n,p) = \Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$



Expected Value of the Binomial Distribution

$$f(k;n,p) = \Pr(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\begin{aligned}
\mu &= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\
&= np \sum_{k=0}^{n} k \frac{(n-1)!}{(n-k)!k!} p^{k-1} (1-p)^{(n-1)-(k-1)} \\
&= np \sum_{k=1}^{n} \frac{(n-1)!}{((n-1)-(k-1))!(k-1)!} p^{k-1} (1-p)^{(n-1)-(k-1)} \\
&= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)} \\
&= np \sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^{\ell} (1-p)^{(n-1)-\ell} \\
&= np \sum_{\ell=0}^{m} \binom{m}{\ell} p^{\ell} (1-p)^{m-\ell} \\
&= np(p+(1-p))^m \\
&= np
\end{aligned}$$

$$\sum_k k \cdot P(X=k)$$

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = n \cdot \frac{(n-1)!}{(n-k)!k!}$$

$$(n-k) = (n-1)-(k-1)$$

$$\binom{n-1}{k-1} = \frac{(n-1)!}{(k-1)!((n-1)-(k-1))!}$$

with $\ell := k-1$

with $m := n-1$

$$\sum_{l=0}^{m} \binom{m}{l} p^l (1-p)^{m-l} = (p+(1-p))^m$$

References

[1] http://en.wikipedia.org/
[2] https://en.wikiversity.org/wiki/Discrete_Mathematics_in_plain_view
[3] https://upload.wikimedia.org/wikiversity/en/6/60/Sequence.1.A.pdf
[4] https://upload.wikimedia.org/wikiversity/en/b/bf/Histogram.1.A.pdf