Partners Universal International Research Journal (PUIRJ)

Volume: 03 Issue: 03 | July – September 2024 | ISSN: 2583-5602 | www.puirj.com

Measure of Variation in Data of Ratio Type: Standard Multiplicative Deviation

Dhritikesh Chakrabarty

Independent Researcher (Ex Associate Professor of Statistics, Handique Girls' College) Guwahati, Assam, India.

Abstract – A measure of dispersion of data of ratio type has here been developed on the basis of the ratios of the observations to their geometric mean. This measure has here been termed as standard multiplicative deviation (SMD) of data. This measure can also be termed as standard geometric deviation (SGD) and/or standard ratio (SR) of data. The development of this measure has been discussed, in this article, with numerical example (application).

Keywords: Data, Ratio Type, Dispersion, Measure, Standard Multiplicative Deviation.

1. INTRODUCTION

An important characteristic of any set of data is dispersion, also known as variation, [14 , 16 , 17 , 18 , 23] in the data. Measures of dispersion, also known as measures of variation, [1 , 2 , 3 , 12 , 15 , 21 , 22] are statistical tools used to understand the spread or variability of data. In simple terms, they provide insights into how scattered or clustered data points are in a dataset. By quantifying this spread, measures of dispersion allow us to assess the reliability and representativeness of the data.

A measure of statistical dispersion is a nonnegative real number that is zero if all the data are the same and increases as the data become more diverse. Some existing measures of dispersion of data are Standard deviation , Interquartile range, Range , Mean absolute difference , Median absolute deviation, Average absolute deviation, Distance standard deviation etc.

Besides variation, central tendency [14 , 17 , 18 , 23] is another basic characteristics of data. Variation and central tendency are closely related as well as measure of variation is closely related to measure of central tendency [17]. Measure of central tendency [1 , 3 , 11 , 12 , 15 , 19 , 24] is mostly based on measures of average [4 , 7 , 8 , 9 , 20] specifically Pythagorean classical means [6 , 8 , 10 , 13] and others means [5 , 6 , 7 , 8 , 9].

The existing measures of variation are defined on the basis of

the additive differences between the pairs of observations,

the additive deviations of the observations from some arbitrary point

and (3) the additive deviations of the observations from some location parameter i.e. measure of central tendency of data.

Accordingly, these are not suitable for measuring variation in data of ratio type. The thrust in this study is to find out a measure of variation in this type of data. A measure of the same has here been developed on the basis of the ratios of the observations to their geometric mean. This measure has here been termed as **standard multiplicative deviation (SMD)** of data. This measure can also be termed as **standard geometric**

deviation (SGD) and/or **standard ratio (SR)** of data. The development of this measure has been discussed, in this article, with numerical example (application).

2. STANDARD MULTIPLICATIVE DEVIATION AS A MEASURE OF DISPERSION

Let

$$x_1, x_2, \ldots\ldots, x_n$$

be n observations on a variable of ratio type.

Automatically, the observations are positive real and hence their geometric mean exists and is defined by

$$\text{GM}(x_1, x_2, \ldots\ldots, x_n) = (x_1 . x_2 \ldots\ldots\ x_n)^{1/n} = (\textstyle\prod_{i=1}^{n} x_i)^{1/n} = G \text{, say}$$

It is seen that

$$\prod_{i=1}^{n} \frac{x_i}{G} = \prod_{i=1}^{n} \frac{G}{x_i} = 1$$

which implies that the geometric mean of the ratios of

$$\frac{x_1}{G}, \frac{x_2}{G}, \ldots\ldots, \frac{x_n}{G}$$

or equivalently of

$$\frac{G}{x_1}, \frac{G}{x_2}, \ldots\ldots, \frac{G}{x_n}$$

cannot yield a measure of variation in

$$x_1, x_2, \ldots\ldots, x_n$$

and hence cannot be a measure of variation.

Let us define a function $F_{\geq 1}(G : x_i)$ by

$$F_{\geq 1}(G : x_i) = \frac{x_i}{G}, \text{ if } x_i > G,$$

$$F_{\geq 1}(G : x_i) = \frac{G}{x_i}, \text{ if } x_i < G$$

$$\& \quad F_{\geq 1}(G : x_i) = 1, \text{ if } x_i = G$$

so that $F_{\geq 1}(G : x_i) > 1$

which means, $F_{\geq 1}(G : x_i)$ is a non-proper fraction of the ratio of x_i & G. .

$F_{\geq 1}(G : x_i)$ is the multiplier (which is >1) by which x_i is different (greater or smaller) from G.

Therefore, the geometric mean of

$$F_{\geq 1}(G : x_1)\, ,\, F_{\geq 1}(G : x_2)\, , \ldots\ldots\ldots , F_{\geq 1}(G : x_n)$$

$$\text{i.e.} \quad \{\prod_{i=1}^{n} F_{\geq 1}(G : x_n)\}^{\frac{1}{n}}$$

describes the average multiplier by which the observations

$$x_1\, , x_2\, , \ldots\ldots\ldots , x_n$$

are different (greater or smaller) from their geometric mean G

and hence can be regarded as a measure of variation of $x_1\, , x_2\, , \ldots\ldots\ldots , x_n$.

Since the multiplier (i.e. multiplicative deviation) is taken here from geometric mean, this measure of variation can be termed as **standard multiplicative deviation (SMD).**

Also, due to the use of geometric mean in taking deviation, this measure can be termed as **standard geometric deviation (SGD)**.

Moreover, since the ratio is taken here as the deviation of observation from geometric mean, this measure of variation can also be termed as **standard ratio (SR).**

Definition:
The **standard multiplicative deviation** of a set of observations

$$x_1\, , x_2\, , \ldots\ldots\ldots , x_n$$

, denoted here by SR, can be defined by

$$\text{SR}\,(x_1\, , x_2\, , \ldots\ldots\ldots , x_n) = \{\prod_{i=1}^{n} \text{F>1}\,(G : x_n)\}^{\frac{1}{n}}$$

where

$$F_{\geq 1}(G : x_i) = \frac{x_i}{G}\, , \text{ if } x_i > G\, ,$$

$$F_{\geq 1}(G : x_i) = \frac{G}{x_i}\, , \text{ if } x_i < G$$

$$\& \quad F_{\geq 1}(G : x_i) = 1\, , \text{ if } x_i = G$$

or equivalently

$$\text{SR}\,(x_1\, , x_2\, , \ldots\ldots\ldots , x_n) = \{ \frac{\prod_{i=1}^{n} \frac{G}{x_i}\, , \ (\text{for } x_i > G\,)}{\prod_{i=1}^{n} \frac{G}{x_i}\, , \ (\text{for } x_i < G\,)} \prod_{i=1}^{n} \frac{G}{x_i}\, , \ (\text{for } x_i = G\,) \}^{\frac{1}{n}}$$

Note:

In the standard deviation (i.e. standard arithmetic deviation) we take the squares of the arithmetic deviations of the observations from their arithmetic mean and then take the square root of their arithmetic mean.

Note that if we take the squares of

$$F_{\geq 1}(G : x_1),\ F_{\geq 1}(G : x_2),\ \ldots\ldots,\ F_{\geq 1}(G : x_n)$$

and then take the square root of their geometric mean i.e.

$$\sqrt{[\prod_{i=1}^{n} \{F_{\geq 1}(G : x_n)\}2]^{\frac{1}{n}}}$$

then we see that it come down to

$$\{\prod_{i=1}^{n} F_{\geq 1}(G : x_n)\}^{\frac{1}{n}}$$

since $\sqrt{[\prod_{i=1}^{n} \{F_{\geq 1}(G : x_n)\}2]^{\frac{1}{n}}} = \{\prod_{i=1}^{n} F_{\geq 1}(G : x_n)\}^{\frac{1}{n}}$,

Therefore, taking the average of the squares of

$$F_{\geq 1}(G : x_1),\ F_{\geq 1}(G : x_2),\ \ldots\ldots,\ F_{\geq 1}(G : x_n)$$

in measuring multiplicative variation is meaningless.

2.1. Some Properties

(1) For $x_1, x_2, \ldots\ldots, x_n$ which are not all identical,

SR $(x_1, x_2, \ldots\ldots, x_n) > 1$ always

since $F_{\geq 1}(G : x_i) > 1$ for some or all of $x_1, x_2, \ldots\ldots, x_n$ in this case.

SR $(x_1, x_2, \ldots\ldots, x_n) = 1$

if and only if all $x_1, x_2, \ldots\ldots, x_n$ are identical

since $F_{\geq 1}(G : x_i) = 1$ for all of $x_1, x_2, \ldots\ldots, x_n$ in this case.

(2) For any non-zero constant a,

SR $(ax_1, ax_2, \ldots\ldots, ax_n)$ = SR $(x_1, x_2, \ldots\ldots, x_n)$

In particular,

SR $(-x_1, -x_2, \ldots\ldots, -x_n)$ = SR $(x_1, x_2, \ldots\ldots, x_n)$

This follows from the fact that

GM $(ax_1, ax_2, \ldots\ldots, ax_n) = (\prod_{i=1}^{n} a\, x_i)^{1/n} = a.(\prod_{i=1}^{n} x_i)^{1/n}$

$= a.$ GM $(x_1, x_2, \ldots\ldots, x_n) = a.G$

which implies, $F_{\geq 1}(G : ax_i) = F{>}1\,(G : x_i)$

whence, SR $(ax_1, ax_2, \ldots\ldots, ax_n)$ = SR $(x_1, x_2, \ldots\ldots, x_n)$

Putting $a = -1$, one can obtain the result

$$SR(-x_1, -x_2, \ldots\ldots, -x_n) = SR(x_1, x_2, \ldots\ldots, x_n)$$

Remark: This property of standard ratio can be interpreted as follows:

"Standard multiplicative deviation is invariant of change of scale."

(3) For real non-zero m,

$$SR(x_1^m, x_2^m \ldots\ldots, x_n^m) = \{SR(x_1, x_2, \ldots\ldots, x_n)\}^m$$

This follows from the fact that

$$F_{\geq 1}(G : x_i^m) = \{F_{\geq 1}(G : x_i)\}^m,$$

which implies,

$$\{\prod_{i=1}^{n} F_{\geq 1}(G : x_i^m)\}^{\frac{1}{n}} = [\{\prod_{i=1}^{n} F_{\geq 1}(G : x_i)\}^{\frac{1}{n}}]^m$$

and hence the result.

Corollary:

In particular putting m = – 1 in the this result, one can obtain the following result

$$SR\left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots\ldots, \frac{1}{x_n}\right) = \frac{1}{SR\left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots\ldots, \frac{1}{x_n}\right)}$$

3. NUMERICAL EXAMPLE

Let us calculate the standard ratio of the don sex ratio in India. **Table – 3.1** shows the number of females (denoted by (F) per 1000 males in the year 2021 in India. [25].

Table -3.1 : Number of Females **(F)** per 1000 Males in India in 2021 State-wise

State / Union Territory	Value of **F** (x_i)	State / Union Territory	Value of **(F)**
Andaman & Nicobar	876	Lakshadweep	946
Andhra Pradesh	993	Madhya Pradesh	931
Arunachal Pradesh	938	Maharashtra	929
Assam	958	Manipur	985
Bihar	918	Meghalaya	989
Chandigarh	818	Mizoram	976
Chhattisgarh	991	Nagaland	931
Dadra & Nagar Haveli	774	Odisha	979

Daman & Diu	618	Puducherry	1037
Delhi	868	Punjab	895
Goa	973	Rajasthan	928
Gujarat	919	Sikkim	890
Haryana	879	Tamil Nadu	996
Himachal Pradesh	972	Tripura	960
Jammu & Kashmir	889	Uttar Pradesh	912
Jharkhand	948	Uttarakhand	963
Karnataka	973	West Bengal	950
Kerala	1084		

From calculation,

G = GM (x_i) = 927.2544571586795658482621360772 8

Values of $F_{\geq 1}$(G : x_i) obtained from the data in Table – 3.1 have been shown in the following table (Table – 3.2):

Table -3.2 :

State / Union Territory	Value of $F_{\geq 1}$(G : x_i)
Andaman & Nicobar	1.0585096542907300980002992420974
Andhra Pradesh	1.0709034530206302421787999090237
Arunachal Pradesh	1.0115885588452680434679902463889
Assam	1.0331576112726724793628301237107
Bihar	1.0100811080159908124708737865766
Chandigarh	1.1335629060619554594722031003339
Chhattisgarh	1.0687465477788979858931592129 15
Dadra & Nagar Haveli	1.1980031746236170101398735608234
Daman & Diu	1.5004117429752096534761523237496
Delhi	1.0682655036390317578897029217 48
Goa	1.0493344005932258062839600317 02
Gujarat	1.0089819990845261869948445441537
Haryana	1.0548969933545842614883528283018

Himachal Pradesh	1.0482559479718555844892180378359
Jammu & Kashmir	1.043030885442834157309631199187
Jharkhand	1.0223730850589702614154101850498
Karnataka	1.049334400593225806283960031702
Kerala	1.1690426415653204255003213508375
Lakshadweep	1.0202161798162298178259261973176
Madhya Pradesh	1.0040393904956764909047962893263
Maharashtra	1.0018824852529360473153123015942
Manipur	1.062275832049668467820863958095
Meghalaya	1.0665896425351493549998319335593
Mizoram	1.0525697584573364716681860133002
Nagaland	1.0040393904956764909047962893263
Odisha	1.0558051163214471370524119948985
Puducherry	1.1183553683609200011474476391315
Punjab	1.0360384996186363864226392581869
Rajasthan	1.0008040326315658255205703077281
Sikkim	1.0418589406277298492677102652554
Tamil Nadu	1.0741388108847409075630258906219
Tripura	1.0353145165154129229523141114428
Uttar Pradesh	1.0167263784634644362371295351725
Uttarakhand	1.0385498743795235883365400930411
West Bengal	1.024529990301710705004894172782

From this table,

GM {$F_{\geq 1}$ (G : x_i)} = 1.0613636788091281820385036162828

Therefore, the standard ratio of the number of females per 1000 males in India in the year 2021 is **1.0613636788091281820385036162828.**

Now let us see what happens in the case of the ratio of number of female to number of male. The following table (Table – 3.3) has been prepared for the values of the ratio F/M of number of females to the number of males:

Table -3.3 : Ratio F/M of Number of Females to the Number of Males in India in 2021 State-wise

State / Union Territory	Value of F/M ($x_i/$)	State / Union Territory	Value of F/M ($x_i/$)
Andaman & Nicobar	0.876	Lakshadweep	0.946
Andhra Pradesh	0.993	Madhya Pradesh	0.931
Arunachal Pradesh	0.938	Maharashtra	0.929
Assam	0.958	Manipur	0.985
Bihar	0.918	Meghalaya	0.989
Chandigarh	0.818	Mizoram	0.976
Chhattisgarh	0.991	Nagaland	0.931
Dadra & Nagar Haveli	0.774	Odisha	0.979
Daman & Diu	0.618	Puducherry	1.037
Delhi	0.868	Punjab	0.895
Goa	0.973	Rajasthan	0.928
Gujarat	0.919	Sikkim	0.890
Haryana	0.879	Tamil Nadu	0.946
Himachal Pradesh	0.972	Tripura	0.931
Jammu & Kashmir	0.889	Uttar Pradesh	0.929
Jharkhand	0.948	Uttarakhand	0.985
Karnataka	0.973	West Bengal	0.989
Kerala	1.084		0.976

From calculation,

$G^{/}$ = GM ($x_{i^{/}}$) = 0.92725445715867956584826213607728

Calculating the values of F>1 ($G^{/}: x_i{}^{/}$), it has been found that the values of $F_{\geq 1}$ ($G^{/}: x_i{}^{/}$) are exactly same with the corresponding values of $F_{\geq 1}$ (G : x_i) of the Stat / Union Territory.

Accordingly,

GM {$F_{\geq 1}$ ($G^{/}$: $x_{i}{}^{/}$)} = 1.0613636788091281820385036162828

Therefore, the standard ratio of the ratio of number of female to the number male in India in the year 2021 is **1.0613636788091281820385036162828.**

Here, it is to be mentioned that number of females per 1000 males in India overall in the year 2021 is **943** while the ratio of number of female to the number male is **1.0604453870625662778366914103924** as the report published by the Government of India [25].

4. CONCLUSION

The multiplicative deviation between two ratios is convenient for interpreting their dispersion while it is not so convenient for interpreting the same by their additive deviation. Accordingly, standard geometric deviation can be a convenient measure of dispersion of data of ratio type.

From the numerical findings, it has been that all most all the observed values of the number of females per thousand males lie within the interval

$$\left(\frac{GM}{1.11\ \text{SMD}} ,\ GM \times 1.11\ \text{SMD} \right)$$

This fact leads to think of searching for a method/formula, based on geometric mean and standard geometric deviation, of determining the interval within which the observed values fall.

The invariance property namely

"Standard multiplicative deviation is invariant of change of scale."

is an advantageous property of standard multiplicative deviation. This property can help in applying this measure in calculating dispersion of data having large valued numerical observations.

In this article, only three properties of standard multiplicative deviation have been identified while there may be more properties satisfied by this. These are to be identified for the interest of establishing it as a more applicable tool.

REFERENCES

[1] Ali Zulfiqar; Bhaskar, S Bala & Sudheesh, K (2019): "Descriptive Statistics: Measures of Central Tendency, Dispersion, Correlation and Regression", Airway, 2(3), 120 – 125. DOI: 10.4103/ARWY.ARWY_37_19 .
[2] Anderson T. W. & Finn J. D. (1996): "Measures of Variability. In: The New Statistical Analysis of Data", Springer, New York, NY. https://doi.org/10.1007/978-1-4612-4000-6_4 .
[3] Argyrous G. (1997): "Measures of Central Tendency and Measures of Dispersion", In: Statistics for Social Research, Palgrave, London. https://doi.org/10.1007/978-1-349-14777-9_4 .
[4] Bakker Arthur, "The early history of average values and implications for education", Journal of Statistics Education, 2003, 11(1), 17 – 26.
[5] Coggeshall F. (1886): "The Arithmetic, Geometric, and Harmonic Means", The Quarterly Journal of Economics, 1(1), 83–86. https://doi.org/10.2307/1883111 . https://www.jstor.org/stable/1883111 .
[6] Dhritikesh Chakrabarty (2016): "Pythagorean Mean: Concept behind the Averages and Lot of Measures of Characteristics of Data", NaSAEAST- 2016, Abstract ID: CMAST_NaSAEAST (Inv)-1601), 2016. DOI: 10.13140/RG.2.2.27022.57920 .
[7] Dhritikesh Chakrabarty (2017): "Objectives and Philosophy behind the Construction of Different Types of Measures of Average", NaSAEAST- 2017, Abstract ID: CMAST_NaSAEAST (Inv)- 1701. DOI: 10.13140/RG.2.2.23858.17606 .
[8] Dhritikesh Chakrabarty (2019): "Pythagorean Geometric Mean: Measure of Relative Change in a Group of Variables", NaSAEAST- 2016, Abstract ID: CMAST_NaSAEAST-1901 (I). DOI: 10.13140/RG.2.2.29310.77124 .

[9] Dhritikesh Chakrabarty (2020): "Definition / Formulation of Average from First Principle", Journal of Environmental Science, Computer Science and Engineering & Technology, Section C, (E-ISSN: 2278 – 179 X), 9(2), 151 – 163. www.jecet.org . DOI: 10.24214/jecet.C.9.2.15163.
[10] Dhritikesh Chakrabarty (2021): "Four Formulations of Average Derived from Pythagorean Means", International Journal of Mathematics Trends and Technology (IJMTT) (ISSN: 2231 – 5373), 67(6), 97 – 118. http://www.ijmttjournal.org . doi:10.14445/22315373/IJMTT-V67I6P512.
[11] Dhritikesh Chakrabarty (2021): "Measuremental Data: Seven Measures of Central Tendency", International Journal of Electronics and Applied Research (ISSN : 2395 – 0064), 8(1), 15 – 24. http://eses.net.in/online_journal.html .
[12] Fazli K. & Behboodian J. (2002): "A Construction Method for Measures of Central Tendency and Dispersion", International Journal of Mathematical Education in Science and Technology, 33(2), 299 – 302. https://doi.org/10.1080/002073902753586409 .
[13] Geoffrey Hunter (2020): "The Three Classical Pythagorean Means", mbedded.ninja, https://blog.mbedded.ninja › Mathematics › Statistics .
[14] Herbert F. Weisberg (1992): "Central Tendency and Variability, Series: Quantitative Applications in the Social Sciences", Issue 83, ", Chapter- 4, 46 – 75, Sage Publication, London.
[15] Jain Sharad K. & Vijay P. Singh (2019): "Key Statistical Measures of Data", Chap. 18.2 in Engineering Hydrology: An Introduction to Processes, Analysis, and Modeling, McGraw-Hill Education, New York. https://www.accessengineeringlibrary.com/content/book/9781259641978/toc-chapter/chapter18/section/section6
[16] John H. Mc Donald (2024): "Statistics of Dispersion" Section-3.2, Statistics LibreTexts , https://stats.libretexts.org .
[17] Kelly Ivan W. & James E. Beamer (1986): "Central Tendency and Dispersion: The Essential Union", The Mathematics Teacher, 79(1), 59 – 65. JSTOR, http://www.jstor.org/stable/27964757 . Accessed 9 June 2024.
[18] Malakar I. M. (2023): "Conceptualizing Central Tendency and Dispersion in Applied Statistics", Cognition, 5(1), 50 – 62. https://doi.org/10.3126/cognition.v5i1.55408 .
[19] Manikandan S. (2011): "Measures of Central Tendency: Median and mode", Journal of Pharmacology and Pharmacotherapeutics, 2(3), 214 – 215, 2011. DOI: 10.4103/0976-500X.83300 .
[20] Miguel de Carvalho, "Mean, what do you Mean?", The American Statistician, 2016, 70, 764 – 776.
[21] Moore P. G. (2010): "Principles of Statistical Techniques - Measures of Dispersion" Chapter-7, Cambridge University Press.
[22] Murray R. Spiegel & Larry J. Stephens(2018): The Standard Deviation and Other Measures of Dispersion", In the book "Schaum's Outline of Statistics" Chapter-4, ISBN: 9781260011463, McGraw Hill. https://www.accessscience.com › chapter › chapter4.
[23] Weisberg H. F. (1992): "Central Tendency and Variability", Sage University Paper Series on Quantitative Applications in the Social Sciences, ISBN 0-8039-4007-6 pp.2.
[24] Williams R. B. G.(1984):"Measures of Central Tendency", Introduction to Statistics for Geographers and Earth Scientist, Soft cover ISBN978-0-333-35275-5, eBook ISBN978-1-349-06815-9 , Palgrave, London, 51 – 60.
[25] (2022): "Women & Men in India (A Statistical Compilation of Gender Related Indicators in India)", 2021, 23rd Issue, Social Statistics Division, National Statistical Office, Ministry of Statistics and Programme Implementation, Government of India. www.mospi.gov.in , https://mospi.gov.in/web/mospi/reports-publications.