Original Research Article

Relationship between Karl Pearson's Coefficient of Skewness (S_k) and Ginni's Coefficient of Concentration (G) Under the purview of Digital Payments in India

Abstract:

Aims: Statistics is one of the most important fields. It carries out observation, collection, analysis and interpretation of collected data with necessary techniques and tools for analysis. There are many parameters which help in analyzing the data. Amongst them, Karl Pearson's Coefficient of Skewness and Ginni' Coefficient of Concentration are very important. Karl Pearson's Coefficient of Skewness plays an important role in measuring the symmetry. Symmetry and its measure help in knowing the manner of distribution carried out by the population. Karl Pearson's Coefficient of Skewness is represented by S_k. Apart from Karl Pearson's coefficient of Skewness, Ginni's Coefficient of Concentration is also one more term which is the ratio of Ginni's Coefficient of Mean deviation and two times of Mean of the distribution. The range of Ginni's Coefficient of Concentration is from 0 to 1. It is represented as G. This paper tries to find the relationship between Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration.

Methodology: Theoretical framework is being constructed with the help of known formulas of Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration. A new formula is being generated between Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration by mediating mean as the common factor between the two parameters.

Result: There is a positive relationship between Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration.

Conclusion: With the help of the generated formula, Karl Pearson's Coefficient of Skewness can be calculated once Ginni's Coefficient of Concentration is known where all other parameters are considered to be known.

Keywords: Karl Pearson's Coefficient of Skewness, Ginni's Coefficient of Concentration

Introduction:

Giini Coefficient of Concentration and Karl Pearson's Coefficient of Slewness are most important parameters in analyzing the given set of data**.** Ginni's Coefficient of Concentration is a parameter which tells about the inequality with respect to Lorenz curve. The range of Ginni's Coefficient of Concentration is from 0 to 1. Inequality measure helps in getting an estimate of how the population members are being dispersed. If Ginni's Coefficient of Concentration is 0, it means that the all are equal in the given data set. It means all data points exist in equal combination with each other. If the Ginni's coefficient of Concentration is equal to 1 then it means that the inequality is maximum and all the data points are unequal quantitatively. For any data set, symmetry is important. Measure of symmetry and its importance lies with the fact that symmetrical distribution generally build a composite and regular data set which not only helps in understanding the population but also brings out distribution pattern of the data set. Considering both the parameters into account it can be said that Karl Pearson's Coefficient of Concentration and Ginni's Coefficient of Concentration have to be analyzed in comparison to each other. While calculating the formula of Karl Pearson's Coefficient of Concentration, average mean is used. Similarly to calculate Ginni's Coefficient of Concentration average mean is used. So "Average mean" is the factor common between the two parameters. By taking this into account, it can be understood that the Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration can be linked with the help of common factor Average Mean. This research is an attempt to understand the relationship between Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration with the common factor "Average Mean".

Literature Review:

1. Contribution of K Pearson with respect to calculation of Skewness: Skewness is the parameter which tells about the distribution of the population or it tells about the symmetry of the population. There are different ways of measuring the skewness. In the earlier days, many methods have been employed in order to analyze skewness. K. Pearson's contribution had thrown light on skewness with a measurable formula [Pearson, K. (1894].

2. Situational Formulas for Skewness: Skewness can be measured accordingly with the help of different formulas under different situations. Different Formulas build the foundational understanding of Skewness. [Bowley, A. L. (1901)]

3. Need of Statisitcal tools for measuring SKewness: With the advance in statistical tools and techniques, different methods have been adopted in order to measure symmetry of the distribution. Need of different methods is because even though the parameters might have same quantitative value but the symmetry may still differ [Hossain and Adnan (2007)].

4. List of formulas: There are many formulas which given a different picture of skewness when it comes to its measurement. Measuring the skewness with the help of quartile has been in use. Renowned statistician Sir Arthur Lyon Bowley has framed formula for measuring skewness. In his formula he has employed quartile Q_1, Q_2 and Q_3 namely first quartile, second quartile and third quartile [Arthur Lyon Bowley (1901)].

$$Sk = \frac{Q_3+Q_1-2Q_2}{Q_3-Q_1}$$

When it comes to skewness, one of the most striking works done was by pearson. There have been many formulas framed by pearson in order to measure the skewness. Basically, usage of the concept of normal distribution has been a breakthrough in measuring the skewness. Basically there are two formulas used to measure skewness which are one is

$$Sk = \frac{Mean - Mode}{\sigma}$$

Apart from this, another formula used for measuring Skewness is

$$Sk = \frac{3(Mean - Median)}{\sigma}$$

There are no minimum and maximum values for Coefficient of skewness measure. But it helps in taking maximum value for the skewness while integrating with respect to σ.There are many other formulas which can be employed in order to calculate the value of skewness.

5. Formula for Rank Skewness by Yule , Udny: One such formula is Rank Skewness. Rank Skewness takes ranks of the data set into consideration before calculating the skewness of the data set. Based on rank of different data sets, skewness is calculated [Yule, Udny, G. (1912)].

6. Contribution of formula by Groeneveld , Meeden and Gin: On the path of evolution, there have been many formulas which have given substantial proofs to solve the problems related to Coefficient of skewness. Many formulas use mean as a parameter in order to measure coefficient of skewness. One of the formulas is

$$\frac{(\mu\text{-}v)}{E(IX\text{-}VI)}$$

[Groeneveld, Meeden and Gin (1984)]

7. Development of Statistical Tables for Calculating Skewness: Not only developing the formula but also different supporting statistical tables have been developed in order to calculate the coefficient of skewness [Pearson, E.S. and Hartley, H.O. (1966)]

8. Future Aspects of Karl Pearson’s Coefficient of Skewness: Works of Pearson has significantly motivated different people to take up statistics as a most important field. In the same way it has helped to introduce personalities like Mahalanobis, who has contributed tremendously [Tapan K. Nayak (2009)]

9. Measure of Inequality by Ginni’ S coefficient of Concentration: When it comes to Ginni’s Coefficient of Concentration, it is generally used measure the inequality. Under the comparison with Lorenz curve is done then it is measured between 0 and 1. Construction of Ginni’s Indice has been a tale of curiosity for statistician. This will bring the idea of usage of mean and Ginni’s Coefficient of Mean In order to bring the application of Ginni’s Coefficient of Concentration; it was used in estimating the inequalities measures. Interpolation of the inequalities measures

help in assessing the it's difference with respect to Lorenz curve.The Ginni's Coefficient of Concentration's main purpose was to measure inequality with respect to distribution of income and wealth. The most important part of Ginni's Coefficient of Concentration is the ease of measurement. The measurement of Ginni's Coefficient of Concentration is easy to understand. [Atkinson, T. (1970)]

10. Utility of Ginni's Coefficient of Concentration: There are various uses of Ginni's Coefficient of Concentration. Apart from measuring the inequality, Ginni's Coefficient of Concentration is used in simple ordered alternative wherein homogeneity is tested.[Bansal, P., Arora, S. and Mahajan, K. (2011)].

11. Usage of Ginni's Coefficient of Concentration: Goodness Of Fit is very important from analysis point of view. It is also used to measure the goodness of fit. This is generally done in with the help of Ginni's spacing index [Jammalamadaka, S. R. and Goria, M. (2004)

12. Application of Ginni's Coefficient of Concentration in wealth Distribution: Ginni's Coefficient of Concentration has emerged out of family and growth distribution with respect to wealth. It has mainly brought the concept of equal distribution of wealth in comparison with morenz curve. [Fei, J., Ranis, G. and Kuo, S. (1978)]

13. Usage in Income Index: There are many methods which are employed in order to measure the deviation in income distribution. One such method is using Ginni's Coefficient of Concentration[Lorenz, M. O. (1905) Methods of measuring the concentration of wealth. Publications of the American Statistical Association, 9, 209–219]

14. Combination of Moments and Ginni's Coefficient of Concentration: In Statistics, Moments and their utility has helped in order to develop a theory on moments with respect to Concentration of distribution. This is very well summarizwd with the help of Ginni's Coefficient of Concentration. [Raghavachari, M. (1974)]

15. Variance Estimators: There are many estimators with respect to each method and Parameter. One the same line variance estimators are prevalent for Ginni's Coefficient of Concentration. Especiall those variance are known to be Jackknife Variance estimators.[Yitzhaki, S. (1991)]

16. Graphical Representation of Ginni's Coefficient of Concentration: Graphically also Ginni's Coefficient of Concentration can be represented. This is done with the help of Histogram. With the help of histogram, interpolation is done and the Lorenz curve is estimated.[Till`e, Y. and Langel, M. (2012)]

17. Future aspects of Ginni's Coefficient of Concentration: Initially Ginni's Coefficient of Concentration was being used as a descriptive measure of statistics. In the coming days, it will be used as a index for determining inference. Under the context of Inferential Statistics, it can will be used for inferential analysis [Giovanni Maria Giorgi, Chiara Gigliarano (2016)]

18. Utility of Ginni's Coefficient Of Concentration in Coming times: Anova and its application has brought many possibilities for Ginni's Coefficient of Concentration. One such utility is breaking up of Ginni's index in the form of ANOVA [Giovanni Maria Giorgi (2019)]

Theoretical Framework:

Ginni's Coefficient of Concentration (G) = $\Delta_1/2\bar{X}$

$\bar{X} = \Delta_1/2G$---------------- (1)

Karl Pearson's Coefficient of Skewness (S_k) = 3(mean-median)/σ

S_k=3(Mean-Median)/σ

$S_k\sigma$ = 3(Mean-Median)

$S_k\sigma$ = 3Mean-3Median

3 Mean= $S_k\sigma$ + 3 Median

Mean= [($S_k\sigma$ + 3 Median)/3]--(2)

Substituting (2) in (1)

[($S_k\sigma$ + 3 Median)/3]= $\Delta_1/2G$

$S_k\sigma$+ 3Median= $3\Delta_1/2G$

$$S_k = \frac{3\Delta_1}{2\sigma G} - \frac{3\ \text{Median}}{\sigma} \quad \text{------------------------------} \quad (3)$$

Proofs:

Problem No 1: The following are the digital payments at different years in India:

YEAR	TOTAL DIGITAL PAYMENTS (Billion)
2015-16	1015390
2016-17	1276250
2017-18	1369867.34
2018-19	1637134.25
2019-20	1620894.13
2020-21	1414851.73

Source:
1.https://rbidocs.rbi.org.in/rdocs/AnnualReport/PDFs/9IXPAYMENT5467258C18F0496FA39CA6EA7B2BF9B9.PDF

2. https://www.rbi.org.in/Scripts/PSIUserView.aspx

a. Calculating Ginni's Coefficient of Mean Deviation:

Differences	Differences	Differences	Differences	Differences
-206042.4	-16240.1	267266.9	93617.34	260860
-222282.52	251026.8	360884.3	354477.3	
44984.39	344644.1	621744.3		
138601.73	605504.1			
399461.73				
154722.93	1184935	1249895	448094.7	260860

$$\Delta_1 = \frac{\text{Total Sum of Differences}}{\frac{n(n-1)}{2}}$$

Δ_1= 3298508/15

Δ_1= 219900.5

b. Calculating Mean of the Distribution

$\overline{X}$= (8334387.45)/6

$\overline{X}$= 1389064.575

c. Calculating Ginni's Coefficient of Concentration:

(G) = $\Delta_1/2\ \overline{X}$

G= 219900.5 /2*1389064.575

G=0.079154

Karl Pearson's Coefficient of Skewness (S_k)

a. Calculating Median= 1392360

b. Calculating Mean= 1389064.575

c. Standard Deviation: 231808.9

According to theoretical formula:

Karl Pearson's Coefficient of Skewness (S_k) =3(Mean-Median)/ Standard Deviation

Karl Pearson's Coefficient of Skewness= 3(1389064.575-1392360)/ 231808.9

Karl Pearson's Coefficient of Skewness = -0.04264

According to the generated formula:

Karl Pearson's Coefficient of Skewness (S_k) = $\frac{3\Delta_1}{2\sigma G} - \frac{3\ Median}{\sigma}$

Karl Pearson's Coefficient of Skewness (S_k) =

[3*219900.5/2*231808.9*0.079154]-3*1392360/231808.9

Karl Pearson's Coefficient of Skewness (S_k) =

[659701.5/36697.2033412]-18.0194979571

Karl Pearson's Coefficient of Skewness (S_k) = 17.9768876081-18.0194979571

Karl Pearson's Coefficient of Skewness (S_k) = -0.0426103489

Problem 2: The following are the volume of digital financial transactions

Year	Transactions (in Lakhs)
2015-16	59471
2016-17	97808
2017-18	147144
2018-19	232602
2019-20	341240
2020-21	437118

Source:

1.https://rbidocs.rbi.org.in/rdocs/AnnualReport/PDFs/9IXPAYMENT5467258C18F0496FA39CA6EA7B2BF9B9.PDF

2.https://www.rbi.org.in/Scripts/PSIUserView.aspx

a. Calculating Ginni’s Coefficient of Mean Deviation:

Differences	Differences	Differences	Differences	Differences
95878	108638	85458	49336.00	38337.00
204516	194096	134794.00	87673.00	
289974	243432.00	173131.00		
339310.00	281769.00			
377647.00				
1307325	827935	393383	137009	38337

Δ_1= Total Sum of Differences

n(n-1)

2

Δ_1= 2703989/15

Δ_1= 180265.9

b. Calculating Mean of the Distribution

$\bar{X}$= (8334387.45)/6

$\bar{X}$= 219230.50

c. Calculating Ginni’s Coefficient of Concentration:

(G) = $\Delta_1/2\ \bar{X}$

G=180265.9 /2*219230.50

G= 0.41113325928

Karl Pearson’s Coefficient of Skewness (S_k)

a. Calculating Median= 78,640

b. Calculating Mean= 219230.50

c. Standard Deviation: 27108.35

According to theoretical formula:

Karl Pearson’s Coefficient of Skewness (S_k) =3(Mean-Median)/ Standard Deviation

Karl Pearson’s Coefficient of Skewness= 3(219230.50-78,640)/ 27108.35

Karl Pearson's Coefficient of Skewness = 15.5587300591

According to the generated formula:

Karl Pearson's Coefficient of Skewness (S_k) = $\frac{3\Delta_1}{2\sigma G} - \frac{3\ Median}{\sigma}$

Karl Pearson's Coefficient of Skewness (S_k) =

[3*180265.9 /2*27108.35*0.41113325928]-3*78640/27108.35

Karl Pearson's Coefficient of Skewness (S_k) =

[540797.7/22290.2885784]-8.70285354881

Karl Pearson's Coefficient of Skewness (S_k) = 24.2615836083-8.70285354881

Karl Pearson's Coefficient of Skewness (S_k) = 15.5587300595

Problem No 3: The following are the Transactions of Payments with respect to Prepaid Payment Instruments (Billion) in India.

Year	Prepaid Payment Instruments (Billion)
2016-17	838
2017-18	1416.34
2018-19	2133.23
2019-20	2155.58
2020-21	1976.95

Source:

1.https://rbidocs.rbi.org.in/rdocs/AnnualReport/PDFs/9IXPAYMENT5467258C18F0496FA39CA6EA7B2BF9B9.PDF

2.https://www.rbi.org.in/Scripts/PSIUserView.aspx

Ginni's Coefficient of Mean Deviation:

2155.58-2133.23=22.35	2133.23-1976.95=156.28	1976.95-1416.34=560.61	1416.34-838= 578.34
2155.58-1976.95=178.63	2133.23-1416.34=716.89	1976.95-838=1138.95	
2155.58-1416.34=739.24	2133.23-838=1295.23		
2155.58-838=1317.58			
Sum=2257.8	Sum=2168.4	Sum=1699.56	Sum=578.34

Ginni's Coefficient of Mean Deviation = $\frac{\text{Total Sum of Differences}}{\frac{n(n-1)}{2}}$

Ginni's Coefficient of Mean Deviation = 6704.1/10= 670.41

Mean= 1704.02

Ginni's Coefficient of Concentration=670.41/2*1704.02

Ginni's Coefficient of Concentration= 0.19671424044

According to theoretical formula

Karl Pearson's Coefficient of Skewness (S_k) = 3(Mean-Median)/σ

Karl Pearson's Coefficient of Skewness (S_k) = 3(1704.02-1976.95)/ 569.0713631

Karl Pearson's Coefficient of Skewness (S_k) = -1.4388177882

According to Generated Formula

Karl Pearson's Coefficient of Skewness (S_k) = $\frac{3\Delta_1}{2\sigma G} - \frac{3 \text{ Median}}{\sigma}$

Δ_1 = 670.41

G = 0.19671424044

Median=1976.95

σ = 569.0713631

Karl Pearson's Coefficient of Skewness (S_k) =

[3*670.41/2*569.0713631*0.19671424044]-3*1976.95/569.0713631

Karl Pearson's Coefficient of Skewness (S_k) = [2011.23/223.888881896]-10.4219793589

Karl Pearson's Coefficient of Skewness (S_k) = 8.98316157090-10.4219793589

Karl Pearson's Coefficient of Skewness (S_k) = -1.438817787

Special Case: 1. Taking into account Total number of Digital Transactions for consecutive two years.

Year	Transactions (in Lakhs)
2019-20	**341240**
2020-21	**437118**

Ginni's Coefficient of Mean Deviation:

437118-341240 = 95878
Sum=95878

Ginni's Coefficient of Mean Deviation = Total Sum of Differences

n(n-1)

2

Ginni's Coefficient of Mean Deviation = 95878/1= 95878

Mean= 389179

Ginni's Coefficient of Concentration=95878/2*389179

Ginni's Coefficient of Concentration= 0.12317982213

According to theoretical formula

Karl Pearson's Coefficient of Skewness (S_k) = 3(Mean-Median)/σ

Karl Pearson's Coefficient of Skewness (S_k) = 3(389179-389179)/ 67795.98

Karl Pearson's Coefficient of Skewness (S_k) = 0

According to Generated Formula

Karl Pearson's Coefficient of Skewness (S_k) = $\frac{3\Delta_1}{2\sigma G} - \frac{3\ \text{Median}}{\sigma}$

Δ_1 = 95878

G = 0.12317982213

Median=389179

σ = 67795.98

Karl Pearson's Coefficient of Skewness (S_k) =

[3*95878/2*67795.98*0.12317982213]-3*389179/67795.98

Karl Pearson's Coefficient of Skewness (S_k) = [287634/16702.1935150]-17.2213308222

Karl Pearson's Coefficient of Skewness (S_k) = 17.2213308222-17.2213308222

Karl Pearson's Coefficient of Skewness (S_k) = 0

2. For Consecutive three years

Year	**Transactions (in Lakhs)**
2018-19	232602
2019-20	341240
2020-21	437118

Ginni's Coefficient of Mean Deviation:

437118-341240 = 95878	341240-232602=108638	
437118-232602=204516		
Sum=95878	Sum=108638	Total Sum= 204516

Ginni's Coefficient of Mean Deviation = $\frac{\text{Total Sum of Differences}}{\frac{n(n-1)}{2}}$

Ginni's Coefficient of Mean Deviation = 204516/3= 68172

Mean= 336986.7

Ginni's Coefficient of Concentration=68172/2*336986.7

Ginni's Coefficient of Concentration= 0.10114939254

According to theoretical formula

Karl Pearson's Coefficient of Skewness (S_k) = 3(Mean-Median)/σ

Karl Pearson's Coefficient of Skewness (S_k) = 3(336986.7-341240)/ 102324.3

Karl Pearson's Coefficient of Skewness (S_k) = -0.1247005843

According to Generated Formula

Karl Pearson's Coefficient of Skewness (S_k) = $\frac{3\Delta_1}{2\sigma G} - \frac{3\text{ Median}}{\sigma}$

Δ_1 = 68172

G = 0.10114939254

Median=341240

σ = 102324.3

Karl Pearson's Coefficient of Skewness (S_k) =

[3*68172/2*102324.3*0.10114939254]-3*341240/102324.3

Karl Pearson's Coefficient of Skewness (S_k) = [204516/20700.0815741]-1023720/102324.3

Karl Pearson's Coefficient of Skewness (S_k) = 9.87996106526-10.0046616492

Karl Pearson's Coefficient of Skewness (S_k) = -0.1247005839

3. For consecutive four years

Year	Transactions (in Lakhs)
2017-18	147144
2018-19	232602
2019-20	341240
2020-21	437118

Ginni's Coefficient of Mean Deviation:

437118-341240 = 95878	341240-232602=108638	232602-147144=85458	
437118-232602=204516	341240-147144=194096		
437118-147144=289974			
Sum=590368	Sum= 302734	Sum= 85458	T.S=978560

$$\text{Ginni's Coefficient of Mean Deviation} = \frac{\text{Total Sum of Differences}}{\frac{n(n-1)}{2}}$$

Ginni's Coefficient of Mean Deviation = 978560/6= 163093.3333

Mean= 289526

Ginni's Coefficient of Concentration=163093.33/2*289526

Ginni's Coefficient of Concentration= 0.28165518813

According to theoretical formula

Karl Pearson's Coefficient of Skewness (S_k) = 3(Mean-Median)/σ

Karl Pearson's Coefficient of Skewness (S_k) = 3(289526-286921)/ 126452.5

Karl Pearson's Coefficient of Skewness (S_k) = 0.06180186235

According to Generated Formula

$$\text{Karl Pearson's Coefficient of Skewness } (S_k) = \frac{3\Delta_1}{2\sigma G} - \frac{3\text{ Median}}{\sigma}$$

Δ_1 = 163093.3333

G = 0.28165518813

Median=286921

σ = 126452.5

Karl Pearson's Coefficient of Skewness (S_k) =

[3*163093.3333/2*126452.5*0.28165518813]-3*286921/126452.5

Karl Pearson's Coefficient of Skewness (S_k) = [489279.9999/71232.005354]-860763/126452.5

Karl Pearson's Coefficient of Skewness (S_k) = 6.86882248321-6.80700658349

Karl Pearson's Coefficient of Skewness (S_k) = 0.06181589972

4. For 5 Consecutive years

Year	Transactions (in Lakhs)
2016-17	97808
2017-18	147144
2018-19	232602
2019-20	341240
2020-21	437118

Ginni's Coefficient of Mean Deviation:

95878	108638	85458	49336	
204516	194096	134794		
289974	243432			
339310				
Sum=929678	Sum= 546166	Sum= 220252	Sum=49336	T.S=1745432

Ginni's Coefficient of Mean Deviation = $\dfrac{\text{Total Sum of Differences}}{\frac{n(n-1)}{2}}$

Ginni's Coefficient of Mean Deviation = 1745432/10= 174543.2

Mean= 251182.4

Ginni's Coefficient of Concentration=174543.2/2*251182.4

Ginni's Coefficient of Concentration= 0.34744273484

According to theoretical formula

Karl Pearson's Coefficient of Skewness (S_k) = 3(Mean-Median)/σ

Karl Pearson's Coefficient of Skewness (S_k) = 3(251182.4-232602)/ 139082.1

Karl Pearson's Coefficient of Skewness (S_k) = 0.40077910816

According to Generated Formula

Karl Pearson's Coefficient of Skewness (S_k) = $\frac{3\Delta_1}{2\sigma G} - \frac{3\ \text{Median}}{\sigma}$

Δ_1 = 174543.2

G = 0.34744273484

Median=232602

σ = 139082.1

Karl Pearson's Coefficient of Skewness (S_k) =

[3*174543.2/2*139082.1*0.34744273484]-3*232602/139082.1

Karl Pearson's Coefficient of Skewness (S_k) = [523629.6/96646.1303825]-697806/139082.1

Karl Pearson's Coefficient of Skewness (S_k) = 5.41800895625-5.01722363985

Karl Pearson's Coefficient of Skewness (S_k) = 0.40078531640

Data Collection:

Sl.No	Theoretical Value	Generated Formula Value
1	-0.04264	-0.0426103489
2	15.5587300591	15.5587300595
3	-1.4388177882	-1.4388177880

Sl.No	G	Generated Formula Value
1	0.079154	-0.0426103489
2	0.41113325928	15.5587300595
3	0.19671424044	-1.4388177880

Special Case:

Sl.No	Theoretical Value	Generated Formula Value
1	0	0
2	-0.1247005843	-0.1247005839
3	0.06180186235	0.06181589972
4	0.40077910816	0.40078531640
5	15.5587300591	15.5587300595

Sl.No	G	Generated Formula Value
1	0.12317982213	0
2	0.10114939254	-0.1247005839
3	0.28165518813	0.06181589972
4	0.34744273484	0.40078531640
5	0.41113325928	15.5587300595

Data Analysis:

T Test:Paired T Test is done in order to understand the difference between two groups.

Null Hypothesis: There is no significant difference between Theoretical Value and Generated Formula value.

Alternate Hypothesis: There is a significant difference between Theoretical Value and Generated Formula Value.

Sl.No	Theoretical Value	Generated Formula Value
1	-0.04264	-0.0426103489
2	15.5587300591	15.5587300595
3	-1.4388177882	-1.4388177880

Calculated T value = 0.211314 Table T Value $(_{2,0.05})$ = 2.132

NON-PARAMETRIC TEST

Mann Whitney U Test:

H_0: The two populations are equal. : H_1: The Two populations are not equal.

1. Collection of Data:

Theoretical Value	Generated Formula Value
-0.04264	-0.0426103489
15.5587300591	15.5587300595
-1.4388177882	-1.4388177880

2. Arranging the Data in increasing order.

Sl.No	**Order**
1	-1.4388
2	-1.4388
3	-0.0426
4	-0.0426
5	15.5587
6	15.5587

3. Assigning the Ranks to the values arranged in increasing order.

Sl.No	Order	Rank
1	-1.4388	1.5
2	-1.4388	1.5
3	-0.0426	3.5
4	-0.0426	3.5
5	15.5587	5.5
6	15.5587	5.5

4. Adding up different ranks of each group

Sl.No	**Theoretical Value**	**Ranks**
1	-0.0426	3.5
2	15.5587	5.5
3	-1.4388	1.5
Total		10.5

Sl.No	**Generated Formula Value**	**Ranks**
1	-0.0426	3.5
2	15.5587	5.5
3	-1.4388	1.5
Total		10.5

5. Calculating U Statistic

$$U_1 = n_1 * n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 * n_2 + \frac{n_1(n_1+1)}{2} - R_2$$

Substituting the values

$$U_1 = 3*3 + \frac{(3*4)}{2} - 10.5 = 4.5$$

$$U_2 = 3*3 + \frac{(3*4)}{2} - 10.5 = 4.5$$

6. Comparing the calculated U value with table U Value

$U_{tab\ 0.05,3,3} = 0$

$$U_1 = 3*3 + \frac{(3*4)}{2} - 10.5 = 4.5$$

$$U_2 = 3*3 + \frac{(3*4)}{2} - 10.5 = 4.5$$

4.5>0

4.5>0

This implies that

$U_1 > U$ and

$U_2 > U$

$U_{cal} > U_{tab.}$

Correlation:

Y (Karl pearson's Coefficient of Skewness	X (Ginni's Coefficient of Concentration)
-0.0426	0.07915
15.5587	0.41113
-1.4388	0.19671

Coefficient Of Correlation = 0.908647

T Test:Paired T Test is done in order to understand the difference between two groups.

Null Hypothesis: There is no significant difference between Theoretical Value and Generated Formula value.

Alternate Hypothesis: There is a significant difference between Theoretical Value and Generated Formula Value.

Sl.No	Theoretical Value	Generated Formula Value
1	0	0
2	-0.1247005843	-0.1247005839
3	0.06180186235	0.06181589972
4	0.40077910816	0.40078531640
5	15.5587300591	15.5587300595

Calculated T value = 0.217829477 Table T Value $_{(4,0.05)}$ = 2.776

NON-PARAMETRIC TEST

Mann Whitney U Test: H_0: The two populations are equal. : H_1: The Two populations are not equal. Collection of Data:

Sl.No	Theoretical Value	Generated Formula Value
1	0	0
2	-0.1247005843	-0.1247005839
3	0.06180186235	0.06181589972
4	0.40077910816	0.40078531640
5	15.5587300591	15.5587300595

1. Arranging the Data in increasing order.

Sl.No	**Order**
1	-0.1247
2	-0.1247
3	0
4	0
5	0.0618
6	0.06182
7	0.40078
8	0.40079
9	15.5587
10	15.5587

3. Assigning the Ranks to the values arranged in increasing order.

Sl.No	Order	Rank
1	-0.1247	1.5
2	-0.1247	1.5
3	0	2.5
4	0	2.5
5	0.0618	5
6	0.06182	6
7	0.40078	7
8	0.40079	8
9	15.5587	9.5
10	15.5587	9.5

2. Adding up different ranks of each group

Sl.No	**Theoretical Value**	**Ranks**
1	0	2.5
2	-0.124700584	1.5
3	0.061801862	5
4	0.400779108	7
5	15.55873006	9.5
Total		25.5

Sl.No	**Generated Formula Value**	**Ranks**
1	0	2.5

2	-0.124700584	1.5
3	0.0618159	6
4	0.400785316	8
5	15.55873006	9.5
Total		27.5

3. Calculating U Statistic

$$U_1 = n_1 * n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$U_2 = n_1 * n_2 + \frac{n_1(n_1+1)}{2} - R_2$$

Substituting the values

$$U_1 = 5*5 + \frac{(5*6)}{2} - 25.5 = 14.5$$

$$U_2 = 5*5 + \frac{(5*6)}{2} - 27.5 = 12.5$$

4. Comparing the calculated U value with table U Value

$U_{tab\ 0.05,5,5} = 4$

$$U_1 = 5*5 + \frac{(5*6)}{2} - 25.5 = 14.5$$

$$U_2 = 5*5 + \frac{(5*6)}{2} - 27.5 = 12.5$$

$14.5 > 4$

$12.5 > 4$

This implies that

$U_1 > U$ and

$U_2 > U$

$U_{cal} > U_{tab.}$

Correlation:

Sl.No	G	Generated Formula Value
1	0.12317982213	0
2	0.10114939254	-0.1247005839
3	0.28165518813	0.06181589972
4	0.34744273484	0.40078531640
5	0.41113325928	15.5587300595

Correlation Coefficient = 0.66617403

Application of Differentiation Principles

$$S_k = \frac{3\Delta_1}{2\sigma G} - \frac{3\ \text{Median}}{\sigma}$$

where Δ_1, σ and median are considered to be constants.

Differentiating S_k with respect to 'G'.

$$\frac{dS_k}{dG} = \frac{d}{dG}\left(\frac{3\Delta_1}{2\sigma G} - \frac{3\ \text{Median}}{\sigma}\right)$$

$$\frac{dS_k}{dG} = \frac{3\Delta_1}{2\sigma}\frac{d}{dG}\left(\frac{1}{(G)}\right) - \frac{d}{dG}\left(\frac{3\ \text{Median}}{\sigma}\right)$$

$$\frac{dS_k}{dG} = \frac{3\Delta_1}{2\sigma}\frac{(-2)}{(G)^2} - 0$$

$$\frac{dS_k}{dG} = \frac{-3\Delta_1}{\sigma\,(G)^2}$$

Application of Integration Principles

$$S_k = \frac{3\Delta_1}{2\sigma G} - \frac{3\ \text{Median}}{\sigma}$$

Integrating Sk with respect to G

$$\int \mathbf{SkdG} = \int \frac{3\Delta_1}{2\sigma G}\, dG - \int \frac{3\ \text{Median}}{\sigma}\, dG$$

$$\int \mathbf{SkdG} = \frac{3\Delta_1}{2\sigma} \log G - \frac{3\ \text{Median}}{\sigma} (G)$$

Discussion and Conclusion

1. Under Digital Payments, Number of digital transactions and PPI, T Test Shows that study has failed to reject null hypothesis which means that there is no difference between Theoretical values of Karl Pearson's Coefficient of skewness and values of generated formula.

2. With respect to Digital Payments, Number of digital transactions and PPI, Mann Whitney U Test shows that the study has failed to reject null hypothesis as there is no difference between mean values of Karl Pearson's Coefficient of Skewness and Values of generated formula.

3. For Digital Payments, Number of digital transactions and PPI, Coefficient of Correlation between Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration is 0.908647

4. The relationship between Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration is that Karl Pearson's Coefficient of Skewness is inversely proportional to Ginni' coefficient of Concentration.

5. The analysis of Digital Payments, Number of digital transactions and PPI presents that the Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration are positively correlated to each other

6. Under the special case of digital payments, T Test Shows that study has failed to reject null hypothesis which means that there is no difference between Theoretical values of Karl Pearson's Coefficient of skewness and values of generated formula.

7. For the special case of digital payments, Mann Whitney U Test shows that the study has failed to reject null hypothesis as there is no difference between mean values of Karl Pearson's Coefficient of Skewness and Values of generated formula.

8. With respect to Special Case of Digital Payments, Coefficient of Correlation between Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration is 0.66617403.

9. Special Case analysis brings the relationship between Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration where Karl Pearson's Coefficient of Skewness is inversely proportional to Ginni' coefficient of Concentration

10. . Special Case analysis of Digital Payments presents that the Karl Pearson's Coefficient of Skewness and Ginni's Coefficient of Concentration are positively correlated to each other.

References:

1.https://rbidocs.rbi.org.in/rdocs/AnnualReport/PDFs/9IXPAYMENT5467258C18F0496FA39CA6EA7B2BF9B9.PDF

2. https://m.rbi.org.in/Scripts/AnnualReportPublications.aspx?Id=1322

3. Pearson, K. (1894): Contributions to the mathematical theory of evolution. I. In: Karl Pearson's Early Statistical Papers, Cambridge University Press, Cambridge, pp. 1–40.

4. Bowley, A. L. (1901): Elements of Statistics, P.S. King and Son, Laondon.

5. Hossain and Adnan (2007): "A New Approach to Determine the Asymmetry of a Distribution," Journal of Applied St atistical Science, Vol.15, pp. 127-134.

6. Yule, Udny, G. (1912): An introduction to the theory of statistics. C. Griffin, limited.

7. Hossain, M.F. and Adnan, M.A.S.A (2007): "A New Approach to Determine the Asymmetry of a Distribution," Journal of Applied St atistical Science, Vol.15, pp. 127-134.

8. Pearson, E.S. and Hartley, H.O. (1966): "Biometrika Tables for Statisticians", Vols. I and II. Cambridge University Press, Cambridge.

9. Tapan K. Nayak (2009), Impact Of Karl Pearson's Work on Statistical Developments in India, International Statistical Review, 77,1,72-80 doi:10.1111/j.1751-5823.2009.00067.x

9. Atkinson, T. (1970) On the measurement of inequality. Journal of Economic Theory, 2, 244–263.

10. Cowell, F. and Metha, F. (1982) The estimation and interpolation of inequality measures. Review of Economic Studies, 49, 273–290.

11.Jammalamadaka, S. R. and Goria, M. (2004) A test of goodness-of-fit based on Gini's index of spacings. Statistics and Probability Letters, 68, 177–187.]

12.Fei, J., Ranis, G. and Kuo, S. (1978) Growth and the family of distribution of income by factor components. Quarterly Journal of Economics, 92, 17–53.

13. Lorenz, M. O. (1905) Methods of measuring the concentration of wealth. Publications of the American Statistical Association, 9, 209–219]

14. Raghavachari, M. (1974) A remark on the paper "on the moments of concentration ratio". Sankhya series B, 36, 466–467.

15. Yitzhaki, S. (1991) Calculating jackknife variance estimators for parameters of the Gini method. Journal of Business and Economic Statistics, 9, 235–239.

16. Till`e, Y. and Langel, M. (2012) Histogram-based interpolation of the Lorenz curve and Gini index for grouped data. The American Statistician, 66, 225–231.

17. Giorgi, G. M. and Pallini, A. (1987a) About a general method for the lower and upper distribution-free
bounds on Gini's concentration ratio from grouped data. Statistica, 67, 171-184.

18. Giovanni Maria Giorgi ,Chiara Gigliarano (2016), The Gini concentration index: a review of the inference literature, Journal of Economic Survey, DOI: 101111/joes.12185.

19. Giovanni Maria Giorgi (2019), THE GINI CONCENTRATION RATIO: BACK TO THE FUTURE Rivista Italiana di Economia Demografia e Statistica Volume LXXIII n. 2 Aprile-Giugno 2019 Invited paper to the 56th SIEDS Scientific Meeting – Ascoli Piceno