

Estimation of Instability as a Measure of Risk in Actuarial Science

Sudhakar T. Chinchane¹, S.L. Sananse²,

Research Scholar, Department of Statistics, Dr. Babasaheb Ambedkar Marathwada University,
Aurangabad (M.S.), India

Professor, Department of Statistics, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad (M.S.), India

ABSTRACT: In this paper, we present various methods of coefficient of variation (C.V.) for various types of data in case of estimation of instability. Generally, researchers use usual coefficient of variation (C.V.) around mean for trend data in case of over time and over locations. In this study, we show significant result between estimation of instability by usual coefficient of variation method and Pal & Bisaria's coefficient of variation (C.V.) method for trend data over time & over locations. By usual coefficient of variation (C.V.) method, it shows more coefficient of variation (C.V.) than Pal & Bisaria's coefficient of variation (C.V.) for such type of trend data. Hence, we suggest that Pal & Bisaria's method shows less instability i.e. present less variability than usual coefficient of variation (C.V.) method. Obtained result shows that, when instability is less than risk chances also less. Positive correlation shows between instability and risk.

In this paper, we show that such types of misleading results occurs when researchers use usual coefficient of variation (C.V.) method instead of Pal & Bisaria's coefficient of variation (C.V.) method for trend data. By using Sach's t-test procedure for testing significance between two coefficients of variations for each year pooled over time period and locations shows stability for trend data.

KEYWORDS: Coefficient of Variation (CV), Stability, Standard Errors, Instability, Pooling

I. INTRODUCTION

The knowledge of coefficient of variation (CV) is essential on many occasions. It is also essential for estimation of sample size which is required for conducting sample surveys in agriculture and other fields. The value of CV is not generally known before conduct of the survey.

Instability generally measures through by coefficient of variation. In actuarial science, coefficient of variation plays vital role to test the variability in the crop yield data and it shows that when variability is more then risk also more that means more instability presence in that data. Generally, many researchers find coefficient of variation (C.V.) about mean. We have studied and observe that this is not sufficient because the data in the form of linear or non linear trend. So, in this study we present appropriate coefficient of variation method for based on nature of data. For this purpose, a pilot survey can be conducted. However, it would be time-consuming. Other way is to make use of the past available data, where it is assumed that the coefficient of variation (C.V.) remains stable over location and time period. In this connection it is noteworthy that Murthy (1977) has remarked "the population coefficient of variation (C.V.) is usually fairly stable over time period and locations of similar nature.

In this paper, an attempt has been made to examine this assumption of stability of coefficient of variation. For this, we determine stability of coefficient of variation over time period as well as locations. The secondary crop yield data of seven districts of Marathwada Region of Maharashtra for the period from 2000-01 to 2009-10 are used. The stabilities of coefficient of variation were tested by using the procedure described by Sachs (1984). It is noted that coefficient of variation of crop yield data remained stable. This is in accordance with the statement made by Murthy

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

(1977). It is also noted from the present analysis that the coefficient of variation (C.V.) pooled over locations was within the Deming (1966) range.

In describing the amount of variation in a population a measure often used is the coefficient of variation (C.V.). It is also a good index of reliability of an experiment. It expresses experimental error as a percentage of mean. The higher its value, the lower is the reliability of the experiment.

The stability of coefficient of variation (C.V.) over time period and over related locations makes it possible to determine the sample size for estimating a parameter with a specified margin of error. It also helps in forecasting the percentage error in the estimate of a parameter for a given sample size. However, the value of CV is not available before conduct of the survey. For obtaining its value, a pilot survey has to be conducted. In some cases past data are available, which also help to estimate the value of CV.

Not much research has been carried out regarding stability of CV. In this connection it is noteworthy that Murthy (1977) has noted.

“The population coefficient of variation is usually fairly stable over time and location or characteristics of similar nature.”

Deming (1966) has also observed that the coefficient of variation (C.V.) of number of bushels per acre of rice, wheat and likewise the yields for many other crops, measured between fields within a village is usually somewhere between 1/4(25%) to 1/2(50%). The coefficient of variation (C.V.) measured between village is somewhat smaller 1/5(20%) to 2/5 (40%).

In this paper an attempt has been made to examine the assumption of stability of coefficient of variation (C.V.). Further, the statement of Deming (1966) about range of coefficient of variation (C.V.) is also examined.

II. A REVIEW OF METHODS

Estimation of Coefficient of Variation

For the yield data, the following quantities were calculated for each year pooled over locations and for each location pooled over time period and are presented in Tables 2.1 and 2.2:

i) Mean (\bar{x}):

Mean of a set of observations is their sum divided by the number of observations i.e. the arithmetic mean \bar{x} of a set of N numbers X_1, X_2, \dots, X_N is defined as

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^N x_i$$

ii) Standard Deviation (S.D.):

Standard deviation, usually denoted by the Greek letter small sigma (σ), is the positive square root of the arithmetic mean of the squares of the deviations of the given values from their arithmetic mean. For the frequency distribution $x_i | f_i, i = 1, 2, \dots, n$,

$$\sigma = \sqrt{\frac{1}{N} \sum_i f_i (x_i - \bar{x})^2}$$

Where, \bar{x} is the arithmetic mean of the distribution and $\sum f_i = N$.

iii) Coefficient of Variation (C.V.):

In order to study the variability in the time series data, coefficient of variation (C.V.) is used as an index of instability.

$$C.V.(\%) = \frac{\text{Standard Deviation}}{\text{Arithmetic Mean}} = \frac{\delta}{\bar{x}} \times 100$$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

According to Prof. Karl Pearson who suggested this measure, coefficient of variation (C.V.) is the percentage variation in the mean, standard deviation being considered as the total variation in the mean.

For comparing the variability of two series, we calculate the coefficient of variations for each series. The series having grater coefficient of variation (C.V.) is said to be more variable than the other and the series having lesser coefficient of variation (C.V.) is said to be more consistent (or homogeneous) than the other.

Martin and Gray (1971) proposed relative coefficient of variation (C.V.) and defined it as

$$C.V. = \frac{\sigma}{\bar{x}} \times 100 \quad \text{for all } x_i > 0, i = 1, 2, \dots, n$$

which can take on values between 0 to 100 percent only.

iv) Coefficient of Variation Proposed by Pal and Bisaria:

In case of linear or non-linear trend exists in the data, the coefficient of variation around mean does not give the instability. Here, we have used the coefficient of variation developed by Pal and Bisaria, the coefficient of variation has been estimated. The magnitude of risk in yield and gross returns per hectare for major crop was measured by coefficient of variation (CV).

$$\text{Coefficient of Variation} = \frac{\left(\sum (X_t - \hat{X}_t)^2 / n \right)^{1/2}}{\bar{X}} \times 100$$

where,

X_t is actual value in t^{th} year,

\hat{X}_t is trend value in t^{th} year,

\bar{X} is mean of X,

n is number of years in the study period.

Year to year fluctuations of a character from its trend represent variability or risk. Coefficient of variation is commonly used tool to quantify the risk. Different types of trend curves were attempted and the one giving highest R^2 was finally chosen for obtaining data adjusted for trend and for computation of risk.

Comparison of Coefficient of Variation

Sachs (1984) has described the procedure for testing significance between two coefficients of variation for sample sizes for each year pooled over locations and for each location pooled over time period.

The standard error (S.E.) of the CV is given by,

$$S.E. = \frac{CV}{\sqrt{2n}} \sqrt{1 + \frac{(CV)^2}{(10)^4}} \approx \frac{CV}{\sqrt{2n}} \dots\dots\dots(1)$$

The difference between two coefficients of variation with sample sizes can be tested by

$$t = \frac{|CV_1 - CV_2|}{\sqrt{\frac{(CV_1)^2}{2n_1} + \frac{(CV_2)^2}{2n_2}}} \dots\dots\dots(2)$$

where,

CV_1 : Coefficient of Variation for ' n_1 ' observations.

CV_2 : Coefficient of Variation for ' n_2 ' observations.

The value of t is judged in terms of test statistic.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

III. ANALYSIS AND APPLICATIONS AS CASE STUDY

The secondary yield data of crop cutting experiments conducted in seven districts of Marathwada Region of Maharashtra for the period from 2000-01 to 2009-10 were used for the present study and data collected from Directorate of Agriculture, Govt. of Maharashtra.

The yield data of seven districts (Locations) were pooled for each year to examine stability of CV over time period. Similarly, yield data for the period of 2000-01 to 2009-10 were pooled for each location to examine stability of CV over different locations.

The aim is to examine the stability of coefficients of variation over time period as well as locations.

Table 2.1: Coefficient of Variation for the Data Pooled Over Time Period 2001-2010

Sr. No.	Time Period (Year)	Number of Observations	Mean (q/ha)	Standard Deviation	Coefficient of Variation (%)
1	2000-2001	7	270.7143	94.8081	35.0215
2	2001-2002	7	430.8571	101.5372	23.5663
3	2002-2003	7	443.0000	140.1832	31.6441
4	2003-2004	7	522.0000	112.6617	21.5827
5	2004-2005	7	535.0000	98.6070	18.4312
6	2005-2006	7	562.4286	152.9639	27.1970
7	2006-2007	7	769.2857	302.5677	39.3310
8	2007-2008	7	1132.0000	379.0866	33.4882
9	2008-2009	7	723.2857	195.4138	27.0175
10	2009-2010	7	775.8571	210.0075	27.0678

It is noted from Table 2.1 that coefficient of variation for the data pooled over time period 2001-2010. For seven districts data sets, statistical measures such as mean, standard deviation and coefficient of variation are calculated. Their corresponding results are presented in Table 2.1.

Table 2.2: Coefficients of Variation for the Data Pooled Over Seven Locations

Sr. No.	Name of Location	Number of Observations	Mean (q/ha)	Standard Deviation	Coefficient of Variation (%)
1	Aurangabad	10	720	360.3184	50.04422
2	Beed	10	523.1	152.3253	29.11972
3	Jalna	10	756.8	292.0947	38.59602
4	Parbhani	10	640.8	312.3466	48.74323
5	Hingoli	10	746.5	405.5913	54.33239
6	Latur	10	532.1	184.8594	34.74147
7	Nanded	10	395.8	140.5132	35.50106

It is noted from Table 2.2 that coefficient of variation for the data pooled over time period 2001-2010 for seven districts is presented. For ten years data sets, statistical measures such as mean, standard deviation and coefficient of variation are calculated. Their corresponding results are presented in Table 2.2.

Here, it is interesting to note the observation of Murthy (1977) about stability of coefficient of variation over time period. Murthy (1977) in his book has quoted the example of CV of number of persons in villages in 1951 and in 1961 for supporting his statement. The CV obtained by him for 1951 and 1961 were the same (60%). However, the CV of related characteristics was 62% for geographical area and 57% for cultivated area. Eventhough there is a difference of 5%, Murthy (1977) considered them stable. Similarly the coefficients of variation (CVs) obtained by him over time period may not be always same.

We tested significance among twenty one different combinations of coefficients of variation presented in Table 2.2.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

Instability of Coefficient of Variation along with Standard Errors

When linear or non-linear trend exists in the data, the CV around mean does not give the instability. Here, we have used the coefficient of variation developed by Pal and Bisaria, the CV has been estimated. The magnitude of risk in yield and gross returns per hectare for major crop was measured by coefficient of variation (CV).

Table 2.3: Coefficients of Variation and Standard Error by Different Methods over Locations

Sr. No.	Name of the Districts	Coefficient of Variation and Standard Error by Usual Method		Coefficient of Variation and Standard Error by Pal & Bisaria's Method	
		Coefficient of Variation	Standard Error	Coefficient of Variation	Standard Error
1	Aurangabad	50.04	11.19	11.31	2.53
2	Beed	29.12	6.51	13.65	3.05
3	Jalna	38.60	8.63	20.55	4.60
4	Parbhani	48.74	10.90	21.64	4.84
5	Hingoli	54.33	12.15	28.80	6.44
6	Latur	34.74	7.77	13.72	3.07
7	Nanded	35.50	7.94	22.77	5.09

It is noted from Table 2.3 that coefficient of variation and standard error for the data of time period 2001-2010 over locations presented. This gives coefficient of variation and standard errors by using as usual and method proposed by Pal and Bisaria are calculated. Their corresponding results are presented in Table 2.3. Here, we test instability of coefficient of variation, the table shows coefficient of variation by using Pal and Bisaria is more consistent i.e. less variability than the usual method. Similarly, for standard error also less by method proposed by Pal and Bisaria.

Stability of Coefficient of Variation over Time Period

The difference between twenty one different combinations of coefficients of variation pooled over time period is tested as per the procedure given by Sachs (1984) and are presented in Table 2.4.

Table 2.4: Testing Difference between Coefficients of Variations (CVs) over Time Period

Sr. No.	Time Period	Calculated values of test statistic t									
		2001	2002	2003	2004	2005	2006	2007	2008	2009	2010
1	2001	-	0.8581	0.2263	1.0330	1.3256	0.5580	0.2588	0.1000	0.5722	0.5682
2	2002		-	0.6474	0.1963	0.5428	0.3190	1.0873	0.7662	0.3044	0.3085
3	2003			-	0.8306	1.1410	0.3370	0.4815	0.1266	0.3516	0.3475
4	2004				-	0.3511	0.5113	1.2510	0.9450	0.4970	0.5010
5	2005					-	0.8437	1.5216	1.2456	0.8302	0.8340
6	2006						-	0.8024	0.4612	0.0148	0.0107
7	2007							-	0.3577	0.8160	0.8122
8	2008								-	0.4756	0.4715
9	2009									-	0.0042
10	2010										-

Table value of test statistic t at 5% L.S. for 9 d.f. = 2.262

Table value of test statistic t at 1% L.S. for 9 d.f. = 3.25

It is noted from the results that the coefficient of variation are stable over time period 2001 to 2010. Thus the test of significance indicated Non-significant at 5% & 1% level of significance that is out of 45 different combinations, CV is stable in 45 combinations. In general, it can be noted that the CV remained stable over time period 2001 to 2010 and instability in zero combinations, which agrees with the statement made by Murthy (1977).

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

We shall now test the stability of CV pooled over locations.

Stability of Coefficient of Variation over Different Locations

We also have attempted to find out stability of CV over different locations for the same characteristic.

Table 2.5: Testing Difference between Coefficients of Variations (CVs) over Time Period

Sr. No.	Locations	Calculated values of test statistic t						
		Aurangabad	Beed	Jalna	Parbhani	Hingoli	Latur	Nanded
1	Aurangabad	-	1.1428	0.5728	0.0589	0.1836	0.7943	0.7495
2	Beed		-	0.6198	1.0929	1.2934	0.3922	0.4395
3	Jalna			-	0.5161	0.7467	0.2347	0.1866
4	Parbhani				-	0.2421	0.7397	0.6944
5	Hingoli					-	0.9545	0.9175
6	Latur						-	0.0484
7	Nanded							-

Table value of test statistic t at 5% L.S. for 6d.f. = 2.45

Table value of test statistic t at 1% L.S. for 6d.f. = 3.70

But unless the differences among them are insignificant or non-significant stability cannot be assured.

We therefore tested significance among 21 different combinations of coefficients of variation as per the procedure described in 2.3.1. The results are presented in Table 2.5. It is noted from the results that, out of all possible combinations of coefficients of variation, which are 21, stability was observed in 21 combinations and instability in zero combinations.

It may, therefore, be noted in general that the coefficients of variations remain stable over different locations.

It can also be noted from present analysis that the coefficients of variation pooled over time period (Table 2.4) and pooled over locations (Table 2.5) are within the Deming (1966) range.

The results obtained will help to estimate sample size required for computing sample surveys in the district and similar other areas.

IV. CONCLUSION

It can be concluded from the present analysis that the coefficient of variation for Cotton crop yield data of Marathwada region of Maharashtra State is observed to be

Table 2.1 shows that, coefficient of variation for the data pooled over time period 2001 to 2010. Among this coefficient of variation (CV) of over time period, in year 2005 coefficient of variation (CV) (18.43%) is very less as compare to other time period. This shows that year 2005 Cotton crop yield is more consistent than other years Cotton crop yield. In case of test of variability, year 2007 have more variability (39.33%) in Cotton crop yield than other years over time period 2001 to 2010.

On the basis of Table 2.2, coefficient of variation for the data pooled over seven locations. Among this coefficient of variation (CV) of over locations, in Beed district coefficient of variation (CV) (29.12%) is verly less as compare to other locations. This shows that year Beed district crop yield is more consistent than other district Cotton crop yield. In case of test of variability, Hingoli district have more variability (54.33%) in Cotton crop yield than other districts.

On the basis of Table 2.3, coefficients of variation and standard error by usual and Pal & Bisaria's method. Beed district Cotton crop yield is more consistent because it's coefficient of variation (CV) (29.12%) is less than other district and also standard error is less (S.E.= 6.51). But by using Pal & Bisaria's method Aurangabad district Cotton

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Website: www.ijirset.com

Vol. 6, Issue 5, May 2017

crop yield is more consistent because its coefficient of variation (CV) (11.31%) is less than other district and also standard error is less (S.E.= 2.53). Some consequences are occurs in computation of coefficient of variation such as in this case study Aurangabad district coefficient of variation (CV) by Pal & Bisaria's method (CV=11.31%) is less than Beed district coefficient variation by usual method (CV= 29.12%). Similarly, all districts calculated coefficient of variations (CV's) shows consequences results by usual and Pal & Bisaria's methods which shows in Table 2.3.

On the basis of Table 2.4, it is noted from the results that the coefficient of variation are stable over time period 2001 to 2010. Thus, the test of significance indicated non-significant at 5% & 1% level of significance that is out of 45 different combinations of time period, coefficient of variation (CV) is stable in 45 combinations. In general, it can be noted that the coefficient of variation (CV) remained stable over time period 2001 to 2010 and instability in zero combinations.

But unless the differences among them are insignificant or non-significant stability can not be assured. We therefore tested significance among 21 different combinations of coefficients of variation as per the procedure described in above. The results are presented in Table 2.5. It is noted from the results that, out of all possible combinations of coefficients of variation, which are 21, stability was observed in 21 combinations and instability in zero combinations. It may, therefore, be noted in general that the coefficients of variations remain stable over different locations. It can also be noted from present analysis that the coefficients of variation pooled over time period (Table 2.4) and pooled over locations (Table 2.5) are within the Deming (1966) range. The results obtained will help to estimate sample size required for computing sample surveys in the district and similar other areas.

Generally, researchers are not use significance test for comparison between coefficient of variations (CV's). By using Sachs proposed methods for comparing two coefficients of variations (CV's) with the help of t-test statistic. Which agrees with the statement made by Murthy (1977) "Population coefficient of variation (CV) is usually fairly stable over time and characteristic of similar nature".

REFERENCES

- [1] Deming, W.E. (1966): Some Theory of Sampling, Dover Publications, INC., New York. pp. 106-108.
- [2] Murthy, M.N. (1977): Sampling Theory and Methods, Statistical Publishing Society, Calcutta. pp. 96-97.
- [3] Sachs, Lothar (1984): Applied Statistics, a Handbook of Techniques; 2nd Edition, Springer-Verlag, New York.
- [4] Sananse S.L. & S.L. Maidapwad (2009): On Estimation of Growth Rates using Linear Models; International Journal of Agricultural Sciences, Vol. 5, No.2 : pp. 463-469.
- [5] Sananse S.L. & Borude S.G. (1992): Measurement of Risk Yield of Rice Crop in the Kokan Region of Maharashtra; Journal of Maharashtra Universities, 17(3): 455-457.
- [6] Sananse S.L., Thakare G.G. & Thakare R.P. (1995): Measurement of Risks Yield of Sugarcane Crop in the Maharashtra; Maharashtra Journal of Agricultural Economics, Vol. 7, No. 1.
- [7] S.T. Chinchane, S.L. Sananse, C.D. Sonar, S.V. Saste, "A Study on Estimation of Probability of Crop Failure and Crop Loss Ratio of Cotton Crop in Marathwada Region of Maharashtra", International Journal of Innovative Research in Science, Engineering and Technology, pp 12732 – 12740 Vol.5, Issue 1, January 2016.
- [8] Suresh Pal & Geeta Bisaria (1990): Risk Consideration in Product Prices: An Expected Deviation Approach; Indian Journal of Agricultural Economics, Vol. 45, No.4, Oct.-Dec. 1990.
- [9] Chinchane S.T., Sananse S.L., Sonar C.D., "An Analysis Yield-Price Risk Associated with Cereal Crops", International Journal of Statistika and Matematika, Volume 6, Issue 2, pp. 73-78, 2013.