In Neil Salkind (Ed.), ***Encyclopedia of Research Design.***
Thousand Oaks, CA: Sage. 2010

Coefficient of Variation

Hervé Abdi

1 Overview

The coefficient of variation measures the variability of a series of numbers independently of the unit of measurement used for these numbers. In order to do so, the coefficient od variation eliminates the unit of measurement of the standard deviation of a series of numbers by dividing it by the mean of these numbers. The coefficient of variation can be used to compare distributions obtained with different units, such as, for example, the variability of the weights of newborns (measured in grams) with the size of adults (measured in centimeters). The coefficient of variation is meaningful only for measurements with a real zero (*i.e.,* "ratio scales") because the mean is meaningful (*i.e.,* unique) only for these scales. So, for example, it will be meaningless to compute the coefficient of variation of the temperature measured in degrees Fahrenheit, because changing the measurement to degrees Celsius will not change the temperature but will change the value of the coefficient of variation (because the value of zero for Celsius is thirty-two for Fahrenheit and therefore the mean

Hervé Abdi
The University of Texas at Dallas
Address correspondence to:
Hervé Abdi
Program in Cognition and Neurosciences, MS: Gr.4.1,
The University of Texas at Dallas,
Richardson, TX 75083–0688, USA
E-mail: herve@utdallas.edu http://www.utd.edu/∼herve

of the temperature will change from one scale to the other, whereas the standard deviation will not change). In addition, the values of the measurement used to compute the coefficient of variation are assumed to be always positive or null. The coefficient of variation is primarily a descriptive statistic, but it is amenable to statistical inferences such as null hypothesis testing or confidence intervals. Standard procedures are often very dependent of the normality assumption and current work is exploring alternative procedures which are less dependent on this normality assumption.

2 Coefficient of variation

The coefficient of variation denoted C_V (or occasionally V) eliminates the unit of measurement from the standard deviation of a series of number by dividing it by the mean of this series of numbers. Formally, if, for a series of N numbers, the standard deviation and the mean are denoted respectively by S and M, the coefficient of variation is computed as:

$$C_V = \frac{S}{M} \ . \tag{1}$$

Often the coefficient of variation is expressed as a percentage which corresponds to the following formula for the coefficient of variation

$$C_V = \frac{S \times 100}{M} \ . \tag{2}$$

This last formula can be potentially misleading because, as we see later, the value of the coefficient of variation can exceed unity and therefore would create percentages larger than one hundred. Here we will use formula 1 and express C_V as a ratio rather than a percentage.

2.1 Range of the coefficient of variation

In a finite sample of N non-negative numbers with a real zero, the coefficient of variation can take value between 0 and $\sqrt{N-1}$ (the maximum value of C_V is reached when all values but one are equal to zero).

2.2 Estimation of a population coefficient of variation

The coefficient of variation computed on a sample is a *biased* estimate of the population coefficient of variation denoted γ_V. An unbiased estimate of the population coefficient of variation is denoted $\widehat{C}_V$, it is computed as

$$\widehat{C}_V = \left(1 + \frac{1}{4N}\right) C_V \tag{3}$$

(where N is the sample size).

3 Testing the Coefficient of variation

When the coefficient of variation is computed on a sample drawn from a normal population, its standard error denoted σ_{C_V} is known and is equal to

$$\sigma_{C_V} = \frac{\gamma_V}{\sqrt{2N}} \ . \tag{4}$$

When γ_V is not known (which is, in general, the case), σ_{C_V} can be estimated by replacing γ_V by its estimation from the sample. Both C_V or $\widehat{C}_V$ can be used for this purpose ($\widehat{C}_V$being preferable because it is a better estimate). So σ_{C_V} can be estimated as:

$$S_{C_V} = \frac{C_V}{\sqrt{2N}} \quad \text{or} \quad \widehat{S}_{C_V} = \frac{\widehat{C}_V}{\sqrt{2N}} \ . \tag{5}$$

Therefore, under the assumption of normality, the statistic

$$t_{C_V} = \frac{C_V - \gamma_V}{S_{C_V}} \tag{6}$$

follows a Student distribution with $\nu = N - 1$ degrees of freedom. It should be stressed that this test is very sensitive to the normality assumption. Work is still being done to minimize the effect of this assumption.

By rewriting Equation 6, confidence intervals can be computed as

$$C_V \pm t_{\alpha,\nu} S_{C_V} \tag{7}$$

Table 1 Example for the coefficient of variation. Daily commission (in dollars) of ten salespersons.

Saleperson	1	2	3	4	5	6	7	8	9	10
Commission	152	155	164	164	182	221	233	236	245	248

(with $t_{\alpha,\nu}$ being the critical value of Student's t for the chosen α level and for $\nu = N - 1$ degrees of freedom). Again, because $\widehat{C}_V$ is a better estimation of γ_V than C_V, it makes some sense to use $\widehat{C}_V$ rather than C_V.

4 Example

Table 1 lists the daily commission in dollars of ten car salespersons. The mean commission is equal to \$200 with a standard deviation of \$40. This gives a value of the coefficient of variation of:

$$C_V = \frac{S}{M} = \frac{40}{200} = 0.200 \ , \tag{8}$$

which corresponds to a population estimate of

$$\widehat{C}_V = \left(1 + \frac{1}{4N}\right) C_V = \left(1 + \frac{1}{4 \times 10}\right) \times 0.200 = 0.205 \ . \tag{9}$$

The standard error of the coefficient of variation is estimated as:

$$S_{C_V} = \frac{C_V}{\sqrt{2N}} = \frac{0.200}{\sqrt{2 \times 10}} = 0.0447 \tag{10}$$

(the value of $\widehat{S}_{C_V}$ is equal to 0.0458).

A t-criterion testing the hypothesis that the population value of the coefficient of variation is equal to zero is equal to:

$$t_{C_V} = \frac{C_V - \gamma_V}{S_{C_V}} = \frac{0.2000}{0.0447} = 4.47 \ . \tag{11}$$

This value of $t_{C_V} = 4.47$ is larger than the critical value of $t_{\alpha,\nu} = 2.26$ (which is the critical value of a Student's t distribution for $\alpha = .05$ and $\nu = 9$ degrees of freedom). Therefore we can reject the null hypothesis and conclude that γ_V is larger than zero. A 95% corresponding confidence interval gives the values of

$$C_V \pm t_{\alpha,\nu} S_{C_V} = 0.2000 \pm 2.26 \times 0.0447 = 0.200 \pm 0.1011 \qquad (12)$$

and therefore we conclude that there is a probability of .95 that the value of γ_V lays in the interval $[0.0989 \quad 0.3011]$.

Related entries

Mean, standard deviation, variance, variability.

Further readings

Abdi, H., Edelman, B., Valentin, D., & Dowling, W.J. (2009). *Experimental design and analysis for psychology.* Oxford: Oxford University Press.

Curto, J.D., & Pinto, J.C. (2009). He coefficient of variation asmptotic distribution in the case of non-*iid* random variables. *Journal of Applied Statistics*, **36**, 21–32.

Nairy, K.S. & Rao, K.N. (2003). Tests of coefficients of variation of normal populations. *Commnunications in Statistics Simulation and Computation*, **32**, 641-661.

Martin, J.D., & Gray, L.N. (1971). Measurement of relative variation: Sociological examples. *American Sociological Review*, **36**, 496–502.

Sokal, R.R, & Rohlf, F.J. (1995). *Biometry (3rd Edition).* New York: Freeman.