

**RESTRICCIÓN AL USO DEL COEFICIENTE DE
VARIACIÓN COMO MEDIDA DE VARIABILIDAD DE
LAS SERIES MENSUALES DE LLUVIA**

Jorge Rodríguez G.*

RESUMEN

Tanto en los textos de estadística básica como en los de estadística aplicada (véase, por ejemplo a J. Benjamin y C.Cornell en Probabilidad y Estadística en Ingeniería Civil; McGraw - Hill; 1981; p.9), se afirma que el coeficiente de variación muestral (cv) permite, al comparar dos conjuntos muestrales, cuál de ellas es más variable o cual es más difícil de pronosticar en cuanto a los valores futuros. Aunque en general se pueda estar de acuerdo en que el cv expresa la dispersión relativa de los datos muestrales, en una asesoría --a tesis de pregrado de la Escuela de Geografía-- con relación a series de tiempo de lluvias mensuales, resolvimos que es preciso modificar parcialmente la significación estadística clásica del coeficiente de variación, lo cual constituye el propósito básico de esta nota.

* Profesor Asociado. Escuela de Geografía. UCV.

Antes de exponer la propuesta correspondiente, es conveniente conocer la fórmula usual del **cv**, la cual se transcribe a continuación como ecuación (1)--op.cit., pág.9:

$$cv = \frac{s}{\bar{x}} \quad \text{o} \quad cv(\%) = 100 * cv \quad (1)$$

$s = \text{desviación típica}$

$\bar{x} = m = \text{media aritmética}$

La desviación típica o standard **s** es un estadístico usualmente utilizado para cuantificar el grado de variabilidad de la muestra, interpretación que concuerda con lo prescrito por el Glossary of Meteorology (American Meteorological Society; 1995; p.532), en el cual el término variabilidad se define como: "Spread -1. (Also called variability). The general departure of individual values from central tendency...".

Si **m** (media aritmética) -una medida de la tendencia central- es constante, **cv** aumentará o disminuirá cuando aumente o disminuya **s** (desviación típica), la cual es una medida absoluta de la dispersión muestral de los datos con respecto al estadístico **m**; tal asociación entre **cv** y **s** ha determinado con frecuencia que bajo cualquier condición numérica de **m**, se infiera que el **cv** mide la dispersión de los datos con relación a **m**, o abreviadamente, que **cv** cuantifica la dispersión relativa muestral. En Climatología, el

cv se ha usado tradicionalmente para comparar series mensuales climáticas de lluvia entre sí o comparárlas con otras series de tiempo de otros elementos meteorológicos. No obstante, cuando se compara el **cv** de meses secos (meses con baja cantidad de lluvia) con el **cv** de meses lluviosos, se observa, en general, que **cv (mes seco) > cv (mes lluvioso)** concluyéndose que el mes seco tiene más variabilidad que el mes lluvioso lo cual contradice el resultado del análisis gráfico o visual de los datos, tal como lo evidencia el siguiente registro

	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
E	10	4	0	3	4	1	0	0	0	0	0	0	0	0	11	1	0	0	1	2
N	25	73	44	58	44	6	84	60	20	22	11	52	0	0	35	20	66	26	30	36

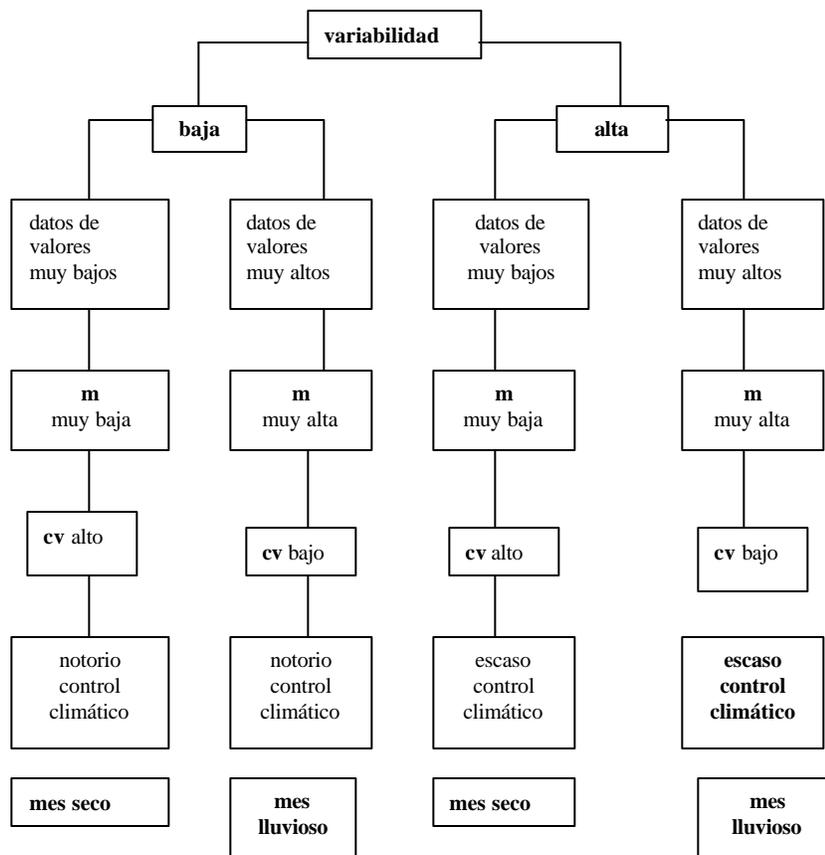
suministrado por los tesisistas Barboza y Blanco. Donde E = lluvia de enero, en mm; N = lluvia de noviembre

Los estadísticos de las series de lluvias mensuales son:

m (E) = 1,9 mm	m (N) = 35,6 mm
s (E) = 3,2 mm	s (N) = 23,6 mm
cv (E) = 171 %	cv (N) = 66 %

Según la forma tradicional de interpretar al **cv** se concluiría en que el mes de enero es más variable que el mes de noviembre. Sin embargo, esta deducción la consideramos no pertinente (véanse gráficos anexos); nótese que enero tiene una desviación típica baja y una media aritmética aún más baja todavía; tales indicadores nos señalan que los totales están siendo ocasionados por factores o controles climáticos que determinan, prácticamente, ausencia de lluvia; así mismo, el pequeño valor de la desviación típica significa que tales montos, relativamente bajos, se agrupan en torno a m, reiterando la notoria influencia del control climático año tras año. De tal manera que dicha serie mensual

no es más que una sucesión de valores casi repetidos y ello indica poca variabilidad. Esa interpretación contradice la forma tradicional de interpretar el *cv*; el esquema que presentamos a continuación muestra las distintas alternativas básicas que podemos hallar en las series climáticas mensuales de lluvia y la correspondiente interpretación.



Como ya se mencionó el término variabilidad, que encabeza al esquema, se considera equivalente al estadístico de dispersión absoluta *s*; en consecuencia las denominaciones variabilidad baja o variabilidad

alta se refieren a valores considerados, relativamente, bajo o alto de la desviación típica: a mayor desviación típica, mayor variabilidad y a menor desviación típica, menor variabilidad. En este diagrama se destaca que una variabilidad baja no conduce, necesariamente, a un cv bajo, pues si la media aritmética m es menor que s se obtendría un cv mayor que 1 (relativamente alto), similar al ejemplo del mes de enero. Esta lluvia de enero tiene baja variabilidad pero un cv alto, significando ello que no siempre el valor del cv se corresponde con un grado de variabilidad único. Todo lo anterior puede explicarse a través de la expresión matemática que relaciona a los tres estadísticos mencionados, como se demuestra seguidamente:

$$s = cv * \bar{x} \quad (2)$$

*si cv es muy grande
y \bar{x} tiende a 0
entonces, s es pequeño*

La ecuación (2) plantea que una muestra puede tener baja variabilidad (pequeño valor en s) pero un alto valor en dispersión relativa (cv) cuando la media aritmética tiende a 0, indicando no coincidencia en las interpretaciones de s y de cv .

En consecuencia, tal como se han bosquejado las opciones, la serie pluvial de noviembre tiene una media aritmética aproximadamente 18 veces mayor a la de enero y una desviación típica aproximadamente 7 veces mayor a la del mes de enero; en ese sentido puede clasificarse, comparativamente, como de alta variabilidad pero con media aritmética mayor que s , lo que implica un cv menor al cv del mes de enero. Aquí de nuevo se presenta una inconsistencia entre la condición muy variable de una serie de tiempo pero con un cv relativamente bajo.

De los cuatro casos, los que clasifican como casos "normales" son los de la siguiente combinación: (1) alta variabilidad pero con menor

valor en m , al cual corresponde un cv alto y (2) baja variabilidad pero con mayor valor en m , el cual se asocia con un cv bajo. En los casos restantes hay incompatibilidad entre el grado de variabilidad y la magnitud del cv .

En resumen, la magnitud del cv no necesariamente se corresponde de manera única con la clase de variabilidad de los datos muestrales y será indispensable hacer la consiguiente interpretación a posteriori de cómo se relacionan la desviación standard y la media aritmética

vers 3/sa30oct99/6:05 pm/do31oct99:6pm

