

MODELO ESTADÍSTICO PARA LA REALIZACIÓN DE ANALOGÍAS ORIENTADAS A PROCESOS DE RECObRO MEJORADO

Robinson Jiménez Díaz ¹, Samuel Fernando Muñoz Navarro ², Claudia Patricia Soto Tavera ³

RESUMEN

Se desarrolló un modelo estadístico que permite analizar campos a nivel mundial en los que se haya implementado métodos de Recobro Mejorado y compararlos con campos colombianos en los que puedan ser aplicados estos métodos. Se implementó una aplicación software denominada EORS, que sirve para realizar analogías usando el modelo desarrollado.

Este modelo puede aplicarse a cualquier método de recobro mejorado o a cualquier área de la ingeniería, en la que se involucren varios parámetros y se tenga que sopesar la importancia de cada uno de ellos.

Palabras claves: Recobro Mejorado; analogías; modelo estadístico; screening.

ABSTRACT

A statistical model that allows analyzing fields at world level in those that EOR techniques has been made and to compare them with Colombian fields in those that one of these methods can be applied was developed. A software application called EORS which allows making analogies using the developed model was implemented.

This model can be applied to any method of enhanced oil recovery or any other engineering area, in which several parameters are involved and it is relevant to consider the importance of each one of them.

Keywords: Enhanced Oil Recovery; analogies; statistical model; screening.

¹Ingeniero de Petróleos. Grupo de Recobro Mejorado. Universidad Industrial de Santander, UIS. Bucaramanga. Colombia.

²M. Sc. en Ingeniería de Hidrocarburos. Ingeniero de Petróleos. Universidad Industrial de Santander, UIS. Bucaramanga. Colombia.

³M. Sc. en Ingeniería de Petróleos. Ingeniera de Petróleos. Instituto Colombiano del Petróleo, ICP. Piedecuesta. Colombia.

INTRODUCCIÓN

Debido a la necesidad manifiesta de aumentar las reservas hidrocarburíferas en Colombia, es conveniente contar con herramientas que permitan el estudio de la aplicabilidad de los métodos de recobro mejorado en campos del ámbito nacional.

Una de estas herramientas son las analogías numéricas, las cuales permiten hacer comparaciones entre experiencias de campo desarrolladas a nivel mundial y nuevos prospectos para la aplicación de un método de recobro mejorado. Este tipo de analogías disminuye la incertidumbre al encontrar yacimientos con propiedades similares y así obtener información adicional que proporcione bases y criterios para la realización de estudios más detallados como por ejemplo pruebas de laboratorio y pilotos en campo.

DESARROLLO DEL MODELO ESTADÍSTICO

El objetivo del modelo estadístico es comparar los principales parámetros de “screening” del campo **n** (nuevo, posible prospecto) con los casos históricos reportados en la literatura en los cuales se llevó a cabo la aplicación de un método de recobro en particular, y de este modo encontrar el más análogo al campo **n**¹.

Se planteó una expresión matemática para poder asignar un *score* o puntuación a cada campo relativa al campo al cual se le va a realizar la analogía (campo **n**).

La ecuación planteada para asignar la puntuación, y que constituye el modelo, es la siguiente:

$$S_{i/n} = \left(\sum_{j=1}^{j=m} \left(I_j^{i/n} * W_j \right) \right) * 100 \quad (1)$$

Donde **i** representa el campo de la base de datos al que se le va asignar un *score* relativo al campo **n**. Si **i** está entre 1 y **k**, quiere decir que **k** es el número de campos actuales en la base de datos¹. Entonces $S_{i/n}$ es el *score* del campo **i** relativo al campo **n**; **j** es el parámetro actual con el que se está realizando la comparación para hacer la analogía, entonces si **j** está entre 1 y **m**, **m** es el número total de parámetros con los que se dispone para efectuar la analogía. $I_j^{i/n}$ es el valor del índice del parámetro **j** para el campo **i** relativo al campo **n**. Este índice tomará valores entre 0 y 1 (ó 0 y 100%). Finalmente, W_j representa el peso de cada parámetro **j** tenido en cuenta para la analogía. El valor del peso se calcula numéricamente a partir de conceptos estadísticos básicos, tales como medidas de tendencia central (media, mediana) y de desviación (coeficiente de variación de Pearson, desviación estándar).

En primera instancia se abordará el problema del índice comparativo $I_j^{i/n}$. Para determinar su valor, se compara cada nuevo valor del parámetro **j** con los valores de ese mismo parámetro de cada uno de los **k** campos existentes en la base de datos. Para hacer tal comparación se desarrolló la siguiente expresión:²

$$\alpha = \left| \frac{j_n - j_i}{j_n + j_i} \right| \quad (2)$$

Donde j_n corresponde al nuevo valor del parámetro y se va a comparar con j_i . La menor diferencia ($j_n - j_i$), será atribuida al campo más análogo al nuevo, respecto al parámetro **j**. En términos generales, a mayor diferencia, menor grado de analogía. Para asignar un valor numérico al índice, se usa la siguiente relación:

¹ Para el desarrollo de este trabajo, los parámetros de screening son equivalentes a los parámetros de analogía: permeabilidad, porosidad, profundidad del yacimiento, espesor neto, temperatura del yacimiento, presión del yacimiento al inicio del proceso de recobro, saturación de petróleo al inicio de proceso de recobro, gravedad y viscosidad del petróleo a condiciones de yacimiento. Cabe anotar que el modelo puede trabajar con otros parámetros siempre y cuando la información que contengan sea de tipo netamente numérico.

² Ibid

$$I_j^{1/n} = 1 - \alpha \quad (3)$$

Cuando j_n sea igual a j_p , el índice tomará su valor máximo: uno (1), ya que $(j_n - j_p)$ se hace cero (0). En ese sentido, a medida que el valor de $(j_n - j_p)$ se aproxime al valor de $(j_n + j_p)$, el índice tenderá hacia su valor mínimo: cero (0). Este hecho se presentará cuando los valores de j_n y j_i sean considerablemente diferentes. De esta manera queda expresado el índice $I_j^{i/n}$.

Posteriormente, se debe determinar el peso global, W_j , de los m parámetros que se disponen para realizar la analogía. Para calcular este peso, se debe evaluar el grado de dispersión que presenten los datos disponibles de cada parámetro; es decir, que con base a tal grado de dispersión se asigna el peso global de cada parámetro partiendo del hecho que a mayor dispersión, menor peso. Para esto se emplea el criterio estadístico conocido como Coeficiente de Variación de Pearson, **CVP**, que se define como la relación entre la desviación estándar, σ , y la media, M , de un conjunto de datos de un parámetro j .

$$CVP_j = \frac{\sigma_j}{M_j} \quad (4)$$

En el anexo se hacen algunas consideraciones importantes para tener en cuenta en el cálculo del **CVP**, la desviación estándar y la media.

Se puede afirmar que a mayor Coeficiente de Variación de Pearson para determinado parámetro, menor peso global tendrá el mismo, es decir, la relación es de proporcionalidad inversa. La meta es plantear una expresión que establezca tal relación. Se propone la siguiente metodología para hallar W_j :

- Calcular el **CVP** para cada parámetro.
- Hacer la sumatoria de los **CVP** para los m parámetros.

- Normalizar el **CVP** de cada parámetro dividiéndolo entre la sumatoria obtenida en el paso anterior. Con esto, los valores del **CVP** normalizado quedan en un rango entre 0 y 1.

$$CVP_{nj} = \frac{CVP_j}{\sum_{j=1}^{j=m} CVP_j} \quad (5)$$

CVP_{nj} : CVP normalizado del parámetro j .
CVP_j: CVP sin normalizar del parámetro j .

- Se hace la siguiente relación, con el propósito de obtener la proporcionalidad inversa:

$$W_j^* = e^{(-CVP_{nj})} \quad (6)$$

- De este modo, los valores de W_j^* también oscilarán entre 0 y 1; pero es preciso realizar una nueva normalización de estos ya que la sumatoria de los W_j^* puede resultar mayor o menor que la unidad.

$$W_{nj}^* = \frac{W_j^*}{\sum_{j=1}^{j=m} W_j^*} \quad (7)$$

- Finalmente, se hace la equivalencia:

$$(100)W_{nj}^* = W_j [\%] \quad (8)$$

W_j es el peso global del parámetro j , que se dispone para la realización de la analogía.

La justificación de la primera normalización es simplemente para darle una mejor distribución a los pesos globales, de lo contrario, se obtendrían valores carentes de sentido ya que se generarían pesos con órdenes de magnitud ínfimos (incluso de 10E-9). Claro que, en todo caso, los resultados estarán fuertemente influenciados por el comportamiento del **CVP** para cada uno de los parámetros de analogía involucrados.

EJEMPLO SINTÉTICO DE APLICACIÓN DEL MODELO ESTADÍSTICO

Se tienen cinco (5) campos, los cuales fueron sometidos a un proceso de recobro mejorado (e.g. inyección WAG); se dispone de cinco (5) parámetros de analogía, como se aprecia en la tabla 1.

Tabla 1. Datos para el ejemplo sintético de aplicación del modelo.

PARÁMETROS DE ANALOGÍA	CAMPO				
	CS_01	CS_02	CS_03	CS_04	CS_05
Permeabilidad (md)	45	13	21	80	19
Porosidad (%)	25	29	19	30	21
Espesor neto (pies)	15	20	9	7	25
Gravedad crudo (°API)	40	38	45	35	41
Viscosidad crudo (cp)	2	1,5	0,98	3,5	0,35

Desarrollo metodológico del ejemplo de aplicación.

1. Se calculan los pesos de los cinco parámetros de analogía, teniendo en cuenta la metodología descrita. La tabla 2 resume los resultados de los criterios estadísticos calculados para cada parámetro y se muestra el peso obtenido por cada uno de estos.

Tabla 2. Criterios estadísticos para los parámetros de analogía y su respectivo peso relativo.

CAMPO	PARÁMETROS DE ANALOGÍA ³				
	(1)	(2)	(3)	(4)	(5)
CS_01	45	25	15	40	2
CS_02	13	29	20	38	1,5
CS_03	21	19	9	45	0,98
CS_04	80	30	7	35	3,5
CS_05	19	21	25	41	0,35

CRITERIOS ESTADÍSTICOS ⁴					
Media	23,60 ^A	24,42 ^G	15,20 ^X	39,53 ^A	0,94 ^A
Mediana	21	25	15	40	1,5
Desv.	36,36	6,12	6,70	5,72	1,74
CVP	1,54	0,25	0,44	0,14	1,85
CVP Norm.	0,36	0,06	0,10	0,04	0,44
Peso Wj [*]	0,69	0,94	0,90	0,97	0,65
Peso Wj [*] Norm.	0,17	0,22	0,22	0,23	0,16
Peso Wj	17 %	22 %	22 %	23 %	16 %

A = Calculada con media armónica.
 G = Calculada con media geométrica.
 X = Calculada con media aritmética.

2. Se calculan los índices comparativos de los cinco campos respecto al campo nuevo, **n**, cuya información se reporta en la tabla 3.

Tabla 3. Información del campo nuevo al cual se le va a realizar la analogía.

CAMPO	PARÁMETROS DE ANALOGÍA ³				
	(1)	(2)	(3)	(4)	(5)
n	40	25	18	42	0,7

Se hará el ejercicio calculando los índices comparativos del campo **CS_01** respecto al campo **n**, empleando las ecuaciones 2 y 3. Se debe tener en cuenta la siguiente notación: **I_k** (índice de permeabilidad), **I_ϕ** (índice de porosidad), **I_h** (índice de espesor neto), **I_γ** (índice de gravedad del crudo), **I_μ** (índice de viscosidad del crudo).

³ Para las tablas 2 y 3 se tiene:

- (1): Permeabilidad, md
- (2): Porosidad, %
- (3): Espesor neto, pies
- (4): Gravedad del crudo, °API
- (5): Viscosidad del crudo, cp

⁴ Para mayor comprensión sobre los cálculos de los criterios estadísticos, se recomienda remitirse al anexo del artículo.

$$I_k = 1 - \frac{|40 - 45|}{40 + 45} = 0,9412$$

$$I_\phi = 1 - \frac{|25 - 25|}{25 + 25} = 1,0000$$

$$I_h = 1 - \frac{|18 - 15|}{18 + 15} = 0,9091$$

$$I_\gamma = 1 - \frac{|42 - 40|}{42 + 40} = 0,9756$$

$$I_\mu = 1 - \frac{|0,7 - 2|}{0,7 + 2} = 0,5185$$

Se concluye que el campo **CS_01** es más análogo al campo **n** en el parámetro de porosidad, puesto que este fue el índice comparativo que arrojó el mayor valor. Los resultados para los índices de todos los campos se resumen en la tabla 4.

Tabla 4. Índices comparativos para los cinco campos del ejemplo.

CAMPO	ÍNDICES COMPARATIVOS RESPECTO AL CAMPO <i>n</i>				
	I_k	I_ϕ	I_h	I_γ	I_μ
CS_01	0,941	1,000	0,909	0,976	0,519
CS_02	0,491	0,926	0,947	0,950	0,636
CS_03	0,689	0,864	0,667	0,965	0,833
CS_04	0,667	0,909	0,560	0,909	0,333
CS_05	0,644	0,913	0,837	0,988	0,667

Los valores en negrilla corresponden a los máximos índices comparativos obtenidos para cada parámetro.

3. Posteriormente, se calcula el *score* de cada campo, para saber cual es el más análogo al campo **n**. Para este propósito, se usa la ecuación 1. Nuevamente se tomará el campo **CS_01** para ejemplificar el cálculo del *score*.

$$S_{CS_01/n} = \frac{[(0,941*17)+(1*22)+(0,909*22)]}{[(0,976*23)+(0,519*16)]} = 87,74$$

4. Finalmente, una vez obtenidos los valores de *score* para cada campo, se observa cual es el más alto y ése será el más análogo al campo nuevo. En la tabla 5 se resumen los *scores* calculados para cada campo.

Tabla 5. Resultados finales del ejemplo de aplicación del modelo.

CAMPO	SCORE (Relativo al campo <i>n</i>)	RANKING
CS_01	87,74	1
CS_02	81,58	3
CS_03	80,91	4
CS_04	69,90	5
CS_05	82,84	2

Para este ejemplo sintético se concluye que el campo más análogo al campo **n** es el **CS_01**.

SOFTWARE EORS

Este software fue desarrollado para la realización de analogías orientadas a procesos de recobro mejorado.

El software permite al usuario realizar las siguientes tareas:

- Consultar datos de algún yacimiento en particular que haya sido sometido a alguno de los métodos de recobro referenciados (Inyección WAG, Inyección Cíclica de Vapor e Inyección Continua de Vapor).
- Realizar analogías o comparaciones entre potenciales prospectos y los campos que se encuentren listados en la base de datos.
- Actualizar los datos existentes de cualquier campo.
- Incorporar información de un campo o yacimiento nuevo.

El software se divide en dos grandes secciones: **Información y Analogías**.

Información.

En esta sección se muestra la información de todos los campos almacenados en la base de datos distribuida en varios paneles. Esta información es modificable con el tiempo. Adicionalmente, se puede insertar datos de nuevos campos a la base de datos.

Analogías.

En esta sección se definen los datos de entrada para realizar una analogía. En una pantalla de salida de resultados, se agrupa en cuatro secciones la descripción final del proceso de analogía resultado:

➤ *Campo resultado:* contiene el campo cuyo puntaje le permite ser el más análogo al campo introducido. Muestra el puntaje final del campo, así como el número de campos utilizados para el proceso y una breve descripción de cómo se calcula tal puntaje. Es preciso recordar que el puntaje máximo para un campo es 100.

➤ *Gráfico resultado:* presenta la descripción gráfica de los resultados del proceso en forma de barras ordenadas de mayor a menor según el puntaje de cada campo.

➤ *Malla de resultados:* describe el puntaje obtenido por cada campo referente a cada parámetro y su respectivo peso.

➤ *Cuadro de definiciones:* permite digitar comentarios y observaciones si se requiere imprimir un reporte.

CONCLUSIONES

1. Con el propósito de llevar a cabo analogías numéricas, se desarrolló un modelo basado en principios básicos de estadística para la asignación del 'peso' o grado de importancia a los parámetros básicos (parámetros de

screening). Este peso depende de la densidad de información disponible del método de recobro que se desea evaluar. Además, el modelo hace una comparación de los nuevos valores de los parámetros del campo prospecto, respecto a cada uno de los valores de tales parámetros de los campos reportados en la literatura, lo cual hace que los resultados tengan un mayor nivel de confianza.

2. La aplicación de este modelo estadístico puede extenderse a cualquier método de recobro mejorado, además de otras ramas de la ingeniería o la investigación en general, donde se tenga que determinar el peso o importancia de ciertos parámetros que influyen en el proceso que se esté evaluando.

NOMENCLATURA

A = Media Armónica.

CVP = Coeficiente Variación de Pearson.

D.m. = Desviación media.

e = Número de Euler ($\approx 2,71828$).

G = Media Geométrica.

$I_j^{i/n}$ = Índice comparativo del parámetro **j** del campo **i** respecto al campo nuevo, **n**.

j = Parámetro actual de analogía.

M = Media (cualquiera de sus tres formas: **A**, **G**, μ).

Me = Mediana.

Mo = Moda.

μ = Media Aritmética.

n = Campo nuevo el cual será objeto de la analogía respectiva.

σ = Desviación estándar o desviación.

V = Varianza.

W_j^* = Peso sin normalizar del parámetro j .

W_{nj}^* = Peso normalizado del parámetro j , (equivale a W_j).

REFERENCIAS

1. Beltrán, Jesús. Análisis e Interpretación de Yacimientos Sometidos a Inyección Continua de Vapor mediante Analogías. UIS. (2005).
2. González, Deisy. Análisis e Interpretación de Yacimientos Sometidos a Inyección Cíclica de Vapor mediante Analogías. UIS. (2005).
3. Jiménez, Robinson. Análisis e Interpretación de Yacimientos Sometidos a Procesos de Inyección Alternada de Agua y Gas (WAG) mediante Analogías. UIS. (2005).
4. www.aulafacil.com/CursoEstadística.htm

RECONOCIMIENTOS

Los autores desean expresar su reconocimiento a Oscar Vanegas Angarita, Ingeniero de Petróleos, Coordinador de la Universidad Industrial de Santander seccional Barrancabermeja, por sus aportes en el desarrollo del modelo estadístico.

También se desea manifestar agradecimiento al Grupo de Recobro Mejorado del Convenio de Cooperación Tecnológica suscrito entre la UIS y ECOPETROL – ICP, por el patrocinio brindado para la ejecución de este trabajo.

ANEXO

Conceptos estadísticos básicos

Medidas de centralización. Son parámetros estadísticos que marcan, bajo distintos criterios, los valores en torno a los cuales se disponen los datos de una distribución. También se llaman medidas de posición central. Estas medidas arrojan información del comportamiento medio de la serie de datos.

Las medidas de centralización más importantes son la media, la mediana y la moda.

Media. Se define como aquel valor que, en una serie estadística de n miembros o datos $j_1, j_2, j_3, \dots, j_n$, representa el valor más probable alrededor del cual se distribuyen el resto de datos de la muestra o de la población, y que por lo tanto se suele tomar como el valor representativo de lo que se pretende medir.

Según las aplicaciones, se calculan diferentes medias. Para el caso que aquí respecta se tienen en cuenta los siguientes criterios:

➤ Si las unidades de los datos recopilados son simples (longitud, masa, tiempo, temperatura, corriente eléctrica) se recomienda emplear la **media aritmética**, la cual es el resultado de sumar todos los elementos del conjunto y dividirlo por el número total de ellos:

$$\mu = \bar{x} = \frac{(j_1 + j_2 + j_3 + \dots + j_n)}{n} \quad (A1)$$

➤ En este grupo estarán los siguientes parámetros: espesor neto [pies], profundidad [pies] y temperatura del yacimiento [°F].

➤ Si las unidades son compuestas, se recomienda el uso de la **media armónica**, la cual se define como el inverso de la media aritmética de los inversos del total de datos que intervienen:

$$A = \frac{1}{\frac{1}{j_1} + \frac{1}{j_2} + \frac{1}{j_3} + \dots + \frac{1}{j_n}} \quad (\text{A2})$$

La media armónica se emplea entonces para la permeabilidad [md], presión del yacimiento [psia], viscosidad del crudo [cp] y gravedad del crudo [° API].

➤ Si las unidades de los datos recopilados vienen expresadas en porcentajes (o fracciones), se emplea la **media geométrica**, que se define como el resultado de multiplicar todos los elementos y extraer la raíz n -ésima del producto.

$$G = (j_1 * j_2 * j_3 * \dots * j_n)^{\frac{1}{n}} \quad (\text{A3})$$

Aquí se ubican la porosidad [%] y la saturación de crudo [%].

Lo más positivo de la media es que en su cálculo se utilizan todos los valores de la serie, por lo que no se pierde ninguna información. Sin embargo, presenta el problema que su valor (en cualquiera de los tres casos) se puede ver influenciado por valores extremos, que se aparten en exceso del resto de la serie. Estos valores anómalos podrían condicionar en gran medida el valor de la media, perdiendo ésta representatividad.

Para efectos de este trabajo, se denomina media, M , a cualquiera de las tres (aritmética, armónica y geométrica) teniendo en cuenta el uso indicado dependiendo de las unidades que se estén manejando.

Mediana. Es el valor de la serie de datos que se sitúa justamente en el centro de la muestra (un 50% de valores son inferiores y otro 50% son superiores a ella). No presenta el problema de estar influida por los valores extremos como la media, pero en cambio no utiliza en su cálculo toda la información de la serie de datos.

Si el número de términos de la distribución es impar, la mediana es el valor del individuo que ocupa el lugar central cuando los datos están ordenados de menor a mayor. Por ejemplo, en la distribución 3, 5, 5, 8, 9, 10, 12, la mediana es $Me = 8$, pues hay tres datos menores que 8 y tres mayores que 8. Si el número de términos de la distribución es par, la mediana es el valor medio de los datos centrales. Así, en la distribución 3, 5, 5, 8, 9, 10, 12, 13, los valores 8 y 9 son los centrales. La mediana será $Me = (8 + 9) / 2 = 8,5$.

Moda. La moda, Mo , de una distribución estadística es el valor que más se repite. Una distribución puede tener más de una moda o no tener ninguna. Por ejemplo, en la distribución 3, 5, 5, 8, 9, 10, 12, 13, la moda sería $Mo = 5$. Ahora bien, si se tiene la siguiente distribución: 2, 3, 3, 3, 4, 5, 8, 9, 9, 9, 11, 13, 14 habrían dos modas, $Mo_1 = 3$ y $Mo_2 = 9$. Sin embargo, en la distribución 2, 4, 5, 6, 7, 9, 12, 13, 15, 16; no habría ninguna moda.

La moda es una medida de centralización poco empleada. En este trabajo, solamente se tuvieron en cuenta la media (en cualquiera de sus tres presentaciones) y la mediana.

Medidas de dispersión. Estos parámetros estadísticos son los encargados de medir que tan diseminados o dispersos se encuentran los datos de una distribución. Los más utilizados se refieren al grado de lejanía de tales datos respecto a la media y son la desviación media, la varianza, la desviación típica (o desviación estándar) y el coeficiente de variación de Pearson.

Desviación media. La desviación media, D.m., es un promedio de los valores absolutos de las desviaciones, $|j_i - M|$, de cada elemento, j_i , de la distribución respecto a su media, M :

$$D.m. = \frac{|j_1 - M| + |j_2 - M| + \dots + |j_n - M|}{n} \quad (\text{A4})$$

Varianza. La varianza, V , es el promedio de los cuadrados de las desviaciones, $(j_i - M)^2$, de cada elemento, j_i , respecto a la media, M :

$$V = \frac{\sum_{i=1}^{i=n} (j_i - M)^2}{n} \quad (\text{A5})$$

Desviación típica (estándar). La desviación típica o desviación estándar, σ , es la raíz cuadrada de la varianza:

$$\sigma = \sqrt{V} = \sqrt{\frac{\sum_{i=1}^{i=n} j_i^2}{n} - M^2} \quad (\text{A6})$$

La razón de ser de la desviación estándar es conseguir que la medida de dispersión se exprese en las mismas unidades que los datos que están siendo evaluados. Por ejemplo, en una distribución de permeabilidades en la que los datos están dados en milidarcys (md), la media viene dada en milidarcys, pero la varianza en milidarcys cuadrados (md²). Para evitar este inconveniente se calcula su raíz cuadrada, obteniéndose así la desviación típica en milidarcys.

Coefficiente de Variación de Pearson. El coeficiente de variación de Pearson, CVP, o simplemente el coeficiente de variación, es el cociente entre la desviación estándar y la media de la distribución:

$$CVP = \frac{\sigma}{M} \quad (\text{A7})$$

Este parámetro sirve para relativizar el valor de la desviación estándar y así poder comparar la dispersión de dos poblaciones estadísticas con gamas de valores muy discretas y diferentes entre sí.

Para el cálculo de la media, la desviación estándar y el coeficiente de variación, hay que tener presentes los criterios relacionados con

la naturaleza de las unidades de los datos en evaluación, planteados anteriormente. Como la desviación estándar se suele asociar con la media aritmética, y esta media no es la única que se tiene en cuenta, la notación de la desviación estándar se puede cambiar simplemente por *desviación*, y esta se calculará de acuerdo a la media correspondiente a cada caso.

El interés del coeficiente de variación es que al ser un porcentaje (o fracción), permite comparar el nivel de dispersión de dos muestras. Esto no ocurre con la desviación estándar, ya que viene expresada en las mismas unidades que los datos de la serie.

Una consideración adicional que debe tenerse en cuenta es que, si al evaluar el coeficiente de variación de Pearson con la media correspondiente para cada caso, se obtiene un valor de este mayor a 0,25 (25%) se recomienda entonces la mediana como medida de centralización para hacer su determinación.