LA CORRECTA UTILIZACION DE LOS PROMEDIOS

JULIO CESAR ANGEL GUTIERREZ

INTRODUCCION

La palabra "promedio" es tan antigua como el principio de contar, pero los frecuentes abusos en su utilización nos han conducido a una verdadera crisis en el manejo de la estadística, hasta el punto de pensar que un "mal promedio" puede convertirse en un "buen promedio" o viceversa, mediante argucias matemáticas o estadísticas.

El concepto de promedio está asociado al lenguaje cotidiano de las personas, y a la toma de decisiones en todos los campos de la actividad humana. Como tal, debe dársele una utilización apropiada, para que sus resultados se conviertan en una verdad incuestionable y unificada para quienes los calculan e interpretan.

Cuando decimos "aproximadamente" o "más o menos" nos estamos refiriendo de manera general a un promedio subjetivo que tiene como fundamento la información histórica o la experiencia de cada individuo, pero que nada tiene que ver con los métodos formales de la estadística. En el lenguaje de la empresa, no puede hacerse lo mismo, sino que es necesario utilizar promedios objetivos con base en una correcta recolección, tabulación y análisis de la información.

Las ideas que se describen a continuación tienen como objetivo el mejoramiento en la calidad de los promedios, para los lectores que hacen uso frecuente de ellos.

LA CORRECTA UTILIZACION DE LOS PROMEDIOS

1. ASPECTOS GENERALES

No existe otra palabra más ambigua en estadística que la palabra "promedio". Cuando nos referimos a él, estamos tratando de dar un valor típico o representativo que identifique a todos los datos provenientes de una muestra o de una población; sinembargo, proliferan las equivocaciones al respecto.

Con frecuencia escuchamos a algunos administradores, directores, ejecutivos o políticos, cuando dan a conocer los resultados de su gestión en términos de promedios de calidad, de ventas, de productividad, de empleo, etc., sin que ello refleje un conocimiento estadístico de las variables o atributos que están manejando. En ocasiones hablan de "parámetros" cuando sólo disponen de información muestral y no propiamente muy aleatoria; en otras ocasiones hablan de "estadísticos " mientras están utilizando la totalidad de la información poblacional. Pero no es sólo esto. También es frecuente observar que la media

JULIO CESAR ANGEL GUTIERREZ. Jefe del Departamento de Ciencias Básicas Universidad EAFIT.

aritmética es el promedio para todo, sin tener en cuenta que su uso está restringido a determinadas situaciones, según la distribución de los datos por promediar, el tipo de variable, el tipo de escala y la utilización que se quiera dar al promedio.

Pocas cosas son tan peligrosas como la estadística en manos de un inexperto, porque su mala utilización conduce a la sobrevaloración, subvaloración o encubrimiento de los resultados verdaderos en el manejo de la información.

Infortunadamente el común de la gente es ingenua y desconoce el verdadero sentido de los promedios; pocas veces se cuestionan el hecho de que los promedios sean diferentes cuando son calculados por diferentes personas teniendo como base la misma información, o cuando se toman muestras deliberadas para obtener los promedios deseados. Tampoco les preocupa el método o la fórmula con la cual ha sido calculado el promedio.

Más grave, aún, es el hecho de que tanto quienes calculan los promedios, como quienes los reciben e interpretan, piensan que con él tienen una información muy completa acerca de una población, sin tener en cuenta otros aspectos importantes como la forma de la distribución y su variabilidad.

Los errores anteriormente descritos, y muchos más, pueden complementarse con los errores estadísticos que pueden clasificarse como dependientes del muestreo e independientes del mismo. Por error de muestreo se entiende la diferencia entre la media de la muestra y la del censo, cuando se obtienen ambos resultados usando los mismos procedimientos. Por error independiente del muestreo se entiende a los errores sistemáticos y a las equivocaciones.

Algunos factores que causan el error sistemático o tendencia del promedio a estar por debajo o por encima del valor del parámetro son: población mal definida, falta de definición del cuestionario, métodos imprecisos de entrevistas o inspección, errores en el uso de números aleatorios, uso de estimadores con error sistemático, etc.

Otros todavía más ingenuos creen que el computador tiene la solución para el cálculo de los promedios requeridos y, por ello, consideran que sus resultados son infalibles e incuestionables, sin tener en cuenta que los programas estadísticos para computador no toman decisiones con relación al tipo promedio que se debe utilizar y, en la mayor parte de los casos, sólo calculan la media aritmética, independientemente de si la variable tiene una escala nominal, ordinal, de intervalo o de razón.

Estamos acostumbrados a escuchar por radio y televisión, lo mismo que a leer en revistas y periódicos, cosas como éstas:

a. Después de un estudio exhaustivo (supuestamente muestreo) se encontraron parámetros muy significativos para la actividad económica de la región. En este caso se está abusando de la palabra "parámetro" cuando se debió utilizar la palabra "estadístico"

b. Con los resultados del censo de 1993 se obtuvieron "estadísticos" muy importantes para orientar los planes de desarrollo de los entes territoriales.

   En este caso debe hablarse de "parámetros", aunque ellos no sean confiables, ya que se trata de indicadores poblacionales.

c. Después de otro estudio exhaustivo se encontró que en promedio (promedio aritmético) cada habitante de un municipio posee dos hectáreas de tierra.

   Bien es sabido que la distribución de la posesión de la tierra en Colombia es fuertemente sesgada, en el sentido de que un porcentaje muy reducido de la población posee la mayor parte de la tierra, lo cual indica claramente que la media aritmética es inadecuada.

Pocas cosas son tan peligrosas como la estadística en manos de un inexperto, porque su mala utilización conduce a la sobrevaloración, subvaloración o encubrimiento de los resultados verdaderos en el manejo de la información.

d. Los que tienen la fe ciega en el computador se encuentran a menudo con situaciones como ésta:

   Efectúan una encuesta para conocer la distribución del sexo en las personas de una comunidad, utilizando la notación:

HOMBRE: 1
MUJER: 0

Después de procesar la información encuentran que el promedio del sexo es 0.40, lo cual constituye un error, por cuanto las variables con escala nominal no pueden promediarse con la media aritmética, ya que sus valores sólo constituyen una marca o etiqueta para los elementos de la muestra.

e. Cuando se escucha por radio la información de que en una ciudad, durante el último mes, se han robado en promedio 25 vehículos por día, todo el mundo podría suponer que cada día se roban 25 vehículos, lo cual no es cierto, si se tienen en cuenta que la distribución de los robos de vehículos tiene mucha variabilidad según los días de la semana.

   Esto es lo que se llama la "fe ciega" en la media aritmética.

f. Cuando se requiere dar a conocer un buen promedio como resultado de la gestión de un administrador, como ocurre en política, entonces se toma la muestra en el estrato, grupo, intervalo u hora apropiada. Tal es el caso cuando se desea promediar el ingreso por familia y sólo se toman familias del estrato medio alto.

Casos como los anteriores y muchos más, son los que a menudo se observan en el cálculo de "promedios".

La correcta utilización de los promedios depende básicamente de los siguientes criterios:

a. La forma de la distribución de la variable (simétrica, sesgada, multimodal, etc).

b. El tipo de escala de medición utilizada para la representación de los datos(nominal, ordinal, de intervalo o de razón)

c. El objetivo de promedio. Si se trata de encontrar un valor representativo o un promedio para hacer inferencias.

d. El tipo de variable por promediar (índices, tasas, promedios, velocidades, etc.).

2. ESCALAS DE MEDICION

Un criterio importante para la correcta utilización de los promedios es la determinación de la escala que debe utilizarse. Stevens reconoce cuatro tipos de escalas, a saber:

2.1 Escala Nominal

La escala nominal se utiliza como medida de identidad. Los números pueden servir como indicativos o etiquetas para identificar objetos o clases. Los números que los deportistas llevan en la espalda constituyen una escala nominal, lo mismo que los números utilizados para identificar el sexo de las personas. En estos casos es de interés la determinación del número o porcentaje de casos que están ubicados en cada clase o categoría.

2.2 Escala Ordinal

En esta escala los números reflejan el orden o la jerarquía de los individuos, objetos o elementos. Las medidas ordinales se disponen desde la más alta a las más baja, o viceversa. Ejemplos de mediciones en escala ordinal son los siguientes:

- La dureza de los materiales, en la cual el número 1 caracteriza un material muy suave y fácil de desmenuzar, como el talco. El número 10 en el otro extremo de la escala, corresponde al diamante, que puede rayar a todos los demás y no puede ser rayado por ninguno.

- Los estratos sociales, por ejemplo de uno a seis, en donde el estrato uno identifica a los de menores recursos, y el estrato seis a los de mayores recursos.

Es importante tener en cuenta que en las escalas ordinales dichas medidas no expresan cuánto más o cuánto menos es un individuo o elemento con respecto a otro, pero sí establece una relación de orden entre ellos.

En una escala ordinal no se obtienen grandes resultados estadísticos para los promedios, excepto para la determinación de la mediana y los percentiles.

2.3 Escala de Intervalos

Una escala de intervalos proporciona números que reflejan diferencias entre los elementos o individuos, donde las unidades de medición son iguales. Las escalas de intervalos indican que un elemento o individuo es tantas unidades mayor o menor que otro.

En estas escalas no existe el cero absoluto, porque los puntos cero son relativos o escogidos de manera arbitraria.

Son escalas de intervalos las siguientes:

- Las escalas de temperatura Celsius y Farenheit.
- Las escalas de tiempo según el calendario, como en el caso de las series de tiempo.

La media aritmética es utilizada frecuentemente para este tipo de escala, lo mismo que la desviación estándar.

2.4 Escala de Razones

Una escala de razones es una escala de intervalos en la que además existe un cero absoluto, esto es, la ausencia total de la característica.

Las escalas de razones más comunes corresponden a las medidas de peso, longitud, capacidad, etc.

Cuando se emplea una escala de razones, los números indican razones o cocientes entre ciertas magnitudes de los objetos, y los datos obtenidos pueden ser sometidos a tratamientos estadísticos más rigurosos.

Cuando los datos están en función de kilogramos, se puede decir que un peso es el doble o la mitad de otro; cuando la temperatura máxima de un sitio es 80°F y la de otro sitio es 40°F, no puede decirse que el primer sitio es dos veces más caluroso que el segundo.

A continuación, se analizan los principales promedios utilizados y sus características básicas:

3. PROMEDIO ARITMETICO

La media aritmética, por su facilidad de cálculo y propiedades matemáticas deseables, es el promedio de uso más común, y se conoce sencillamente como la "media".

Si la variable en estudio se define como X, la media aritmética de una muestra de esa población, se denota como $\overline{X}$ (equis barra) y es igual a la suma de los valores individuales de X, dividido por el número de observaciones:

$$\overline{X} = \frac{X_1 + X_2 \; X_3 \; + \ldots + X_n}{n} = \frac{\sum_{i=1}^{n} X_i}{n}$$

Cuando los datos han sido agrupados en intervalos de clase, se supone que cada marca de clase identifica a todos los datos presentes en cada uno de los intervalos, lo cual simplifica considerablemente los cálculos.

Para datos agrupados la media se calcula como:

$$\overline{X} = \frac{\sum n_i X_i}{n} = \sum f_i X_i \text{ , en donde:}$$

X_i: Marca de clase del intervalo i (i= 1, 2,..k)
n_i : Frecuencia absoluta del intervalo i
f_i : Frecuencia relativa del intervalo i

Debe tenerse en cuenta que las medias aritméticas calculadas con datos agrupados y con datos no agrupados, no conducen, generalmente, a los mismos resultados, aunque sí, muy próximos. La razón de esta diferencia radica en el hecho de que, para datos no agrupados, se están sumando individualmente todos los datos, en cambio que para datos agrupados se están sumando las marcas de clase multiplicadas por su respectiva frecuencia absoluta.

3.1 Interpretación de la Media Aritmética por Analogía con la Física (Mecánica)

Una interpretación curiosa de la media aritmética es el hecho de que ella coincide con el punto de equilibrio, el centro de masa o el centro de gravedad de los datos cuando las frecuencias absolutas (n_i) representan las fuerzas o masas aplicadas sobre una barra rígida y homogénea de longitud L.

Para ilustrar este concepto, consideremos una muestra aleatoria de 50 artículos, a los cuales se les midió el peso en libras, con los siguientes resultados:

X_i : 8, 12, 13,9, 19, 15, 10, 11

n_i : 2, 10, 9, 3, 1, 2, 7, 16

En donde:

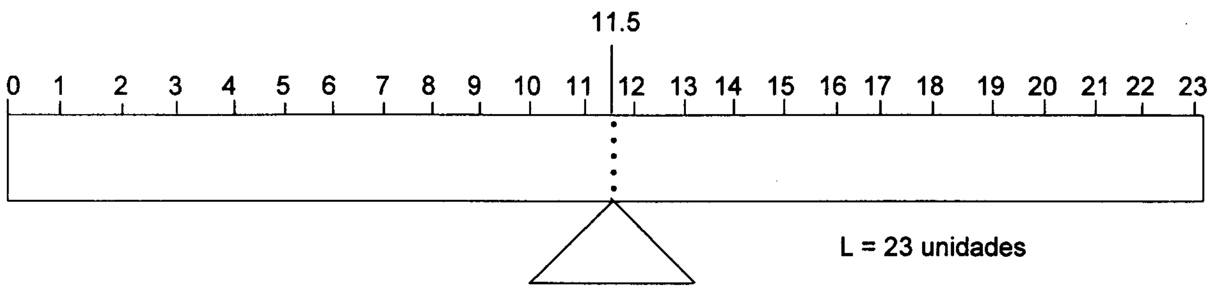
X_i : peso de un artículo (libras)

n_i : Número de artículos con peso X_i

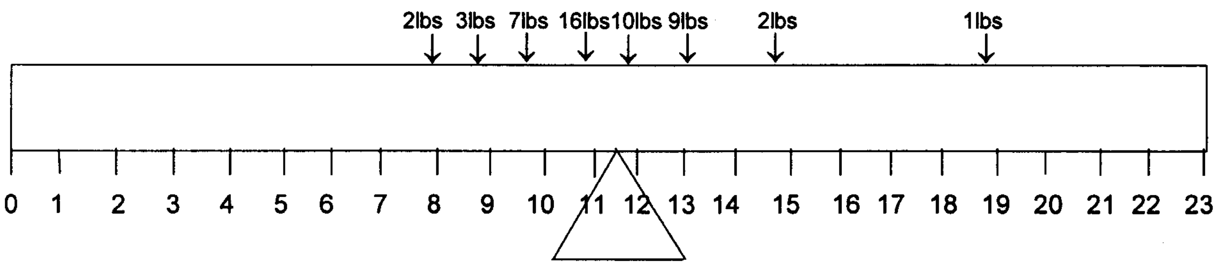
Aplicando la fórmula de la media aritmética tendremos que:

$$\overline{X} = \frac{\sum_{I=1}^{8} n_i X_i}{n} = \frac{575}{50} = 11.5 \text{ libras}$$

Si consideramos una barra libre, rígida y homogénea de longitud L= 2 $\overline{X}$, tendremos que su punto de equilibrio estará ubicado en $\overline{X}$ = 11.5, como se representa en la figura.



Si luego aplicamos masas, iguales a las frecuencias absolutas de cada uno de los datos, ubicadas en cada uno de ellos, obtendremos que la barra continuará en equilibrio.



Desde el punto de vista físico podemos calcular el centro de masa (Cm), como sigue:

$$Cm = \frac{\sum m_i X_i}{\sum m_i} = \frac{(2)(8)+(3)(9)+(7)(10)+(16)(11)+(10)(12)+(9)(13)+(2)(15)+(1)(19)}{2+3+7+16+10+9+2+1} = \frac{575}{50} = 11.5$$

Como era de esperarse, el centro de masa coincide con la media aritmética según las siguientes equivalencias:

$$Cm = \frac{\sum m_i X_i}{\sum m_i} \text{ en donde } m_i = n_i \text{, entonces:}$$

$$Cm = \frac{\sum n_i X_i}{\sum n_i} = \frac{\sum n_i X_i}{n} = \overline{X}$$

Para el caso de datos agrupados por intervalos, los valores de Xi identifican las marcas de clase o puntos medios de cada uno de ellos.

Cuando quiera utilizarse una barra homogénea de longitud mínima (L mín) cuyo punto de equilibrio coincida con la media aritmética y que incluya a todos los datos Xi (abcisas), ésta debe tener una longitud dada por:

$L\ mín = 2\ Max\,[\overline{X} - X\ mín,\ X\ max - \overline{X}]$. La barra deberá estar colocada sobre el eje X de tal manera que uno de sus extremos coincida con el X max ó Xmín, según que la distribución sea sesgada a derecha o a izquierda.

En resumen, se requiere de una barra homogénea que tenga igual longitud antes y después de la media y que además incluya las abcisas correspondientes a todos los datos. Como caso particular de la situación anterior se puede mencionar a la distribución normal, en donde la longitud mínima de la barra coincide con el rango de los datos, esto es: L mín = R=Xmax -Xmín.

Si en el gráfico anterior consideramos las frecuencias absolutas como fuerzas y calculamos los torques con relación al punto $\overline{X}$ = 11.5, habremos derivado una propiedad importante de la media aritmética.

Torque = Fuerza (F) x distancia (d)

$\sum Fd = 2\,(8-11.5) + 3\,(9-11.5) + 7\,(10-11.5) + 16\,(11-11.5) + 10\,(12-11.5) + 9\,(13-11.5) + 2\,(15-11.5) + 1\,(19-11.5) = 0$

Estableciendo la equivalencia con los valores de la variable (X_i) y sus frecuencias absolutas (n_i), hemos obtenido que:

$\sum n_i\,(X_i - \overline{X}) = 0$, lo cual significa que la suma de las desviaciones de los datos con relación a la media aritmética es cero, o también que la suma de los datos es igual a n veces la media aritmética, esto es $\sum n_i X_i = n\overline{X}$.

3.2 Cuando Utilizar la Media Aritmética como Promedio

La media aritmética es el promedio más utilizado ya que se puede calcular con exactitud y se basa en todas las observaciones. Sinembargo, debe utilizarse, así:

a. Cuando la distribución sea simétrica o aproximadamente simétrica.

b. Cuando se quiera hacer un análisis inferencial o se requieran otros estadísticos complementarios como la desviación estándar o el coeficiente de correlación.

c. Cuando las escalas de los datos sean de intervalo o de razón y no sea recomendable otro promedio.

d. Cuando la distribución de los datos sea uniforme.

4. LA MEDIANA (Med)

La mediana es un valor promedio por debajo del cual se encuentran el 50% de los datos. A diferencia de la media aritmética, su valor no es único.

Si se tienen n datos en orden creciente $X_1, X_2, X_3, \ldots, X_n$, entonces la mediana será

$$Med = X_{\frac{n+1}{2}} \quad \text{si n es impar}$$

$$Med = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2} \quad \text{si n es par}$$

No obstante, es bueno observar que para el primer caso con n impar, puede violarse la definición de la mediana y, en el segundo caso con n par, puede reflejarse la no unicidad de ella, característica indeseable para un promedio.

Si tomamos los valores 2,3 y 7, obtendremos que la mediana es Med = 3; por debajo del valor 3 sólo está el 2, que equivale al 33.33% de los datos. Si a la serie anterior le agregamos el número 8 obtendremos que la mediana es Med = $\frac{3+7}{2} = 5$;sinembargo, cualquier número real entre 3 y 7 cumpliría la definición de mediana.

Si los datos están agrupados en intervalos de clase de igual tamaño, la mediana se calcularía así:

$$Med = LIR\,i + c\,\frac{\frac{n}{2} - m_{i-1}}{n_i}$$

En donde:

LIRi Límite inferior real del intervalo que contiene la mediana

c: Tamaño de los intervalos de clase.
n: Número de datos (Tamaño de la muestra)
mi -1: Frecuencia acumulada del intervalo anterior a la clase que contiene la mediana.
ni: Frecuencia absoluta del intervalo que contiene la mediana.

Esta fórmula supone que existe una distribución homogénea de los datos dentro de cada intervalo, lo cual resulta falso en la mayor parte de los casos.

4.1 Cuando Utilizar la Mediana como Promedio

La mediana tiene el inconveniente de que no es única; sinembargo, es útil en algunos casos:

a. Cuando la distribución de los datos es asimétrica.

b. Cuando hay valores extremos que distorsionarían el significado del promedio.

c. Cuando se tienen distribuciones con valores sin determinar, como por ejemplo cuando la primera clase es: menos de X" y la última clase es "más de Y".

d. Cuando los valores extremos no están definidos.

La mediana tiene como propiedades deseables las siguientes:

- La suma de las distancias (valor absoluto) de los datos a la mediana es mínimo.

- La mediana no es sensible a valores extremos. Pueden ocurrir errores por exceso o por defecto sin que el valor de la mediana cambie.

5. LA MODA (Mo)

La moda suele utilizarse como promedio cuando alguno de los datos se destaque claramente sobre los demás, esto es, que su frecuencia absoluta sea considerablemente mayor que todas las demás. Tal es el caso de los estudios de tiempos, en donde el tiempo estándar de una actividad se repite cuando no existen elementos extraños. Al igual que la mediana, es un valor que no se ve afectado por valores extremos de la distribución pero que es poco susceptible para efectuar con él operaciones algebraicas.

Cuando se tienen datos agrupados, la moda se calcula como:

$$Mo = LIR_i + c\,\frac{n_i - n_{i-1}}{2n_i - n_{i-1} - n_{i+1}}$$

Donde n_i, n_{i-1} y n_{i+1} son las frecuencias absolutas de la clase que contiene la moda, la clase anterior a ella y la clase posterior a ella, respectivamente.

5.1 Cuando Utilizar la Moda como Promedio

La moda es el menos confiable de los tres promedios anteriores, y tiene como la mediana el inconveniente de que puede no ser única. Sinembargo, puede utilizarse en algunos casos:

a Cuando haya un dato o un intervalo que tenga una frecuencia considerablemente superior a los demás. Para datos agrupados puede tomarse como moda a la marca de clase del intervalo de mayor frecuencia.

b. Se utiliza propiamente cuando la escala es de tipo nominal, pero también puede utilizarse con escalas ordinal, de intervalo o de razón, siempre y cuando una frecuencia sea muy superior a las demás.

c. Cuando la distribución sea muy asimétrica y una frecuencia sea muy superior a las demás.

d. Cuando la distribución tenga forma de U, o sea cóncava hacia arriba.

e. En distribuciones cuyos valores extremos no están definidos, en cuyo caso deberá escogerse entre la moda y la mediana.

f. Cuando la amplitud de la distribución no es constante.

g. Cuando se quiera encontrar un promedio rápido.

h. En datos multimodales, la moda puede ser útil para dividir la distribución en estratos.

6. LA MEDIA GEOMETRICA (G)

La media geométrica (G) de un conjunto de números positivos es la raíz enésima del producto de esos números, esto es:

$$G = \sqrt[n]{X_1 \; X_2 \; X_3 \; \ldots \; X_n}$$

Para obtenerla, se utiliza el cálculo logarítmico

$$\log G = \frac{1}{n} \log (X_1 \; X_2 \ldots X_n)$$

$$G = \text{antilog } 1/n \;\; (\log X1 + \log X2 + \ldots + \log X n)$$

Para datos agrupados se tiene que:

$$G = \sqrt{X_1^{n\,1}\, X_2^{n\;2} \ldots X_n^{n\,n}}$$

$$\log G = 1/n \;\log \quad (X_1^{n\,1}\, X_2^{n2} \ldots X_n^{n\,n})$$

$$\log G = 1/n \; n\,1 \log X\,1 + n\,2 \log X\,2 + \ldots + n\,n \log X\,n$$

$$\log G = \frac{1}{n} \sum_{i=1}^{n} n_1 \log X_1$$

Donde X_i son las marcas de clase y n_i las frecuencias absolutas.

6.1 Cuando Utilizar la Media Geométrica como Promedio

a. Cuando se quiera dar mayor importancia a los valores pequeños.

b. Cuando los datos tengan un crecimiento geométrico o porcentual.

c. Cuando se quiera promediar razones financieras, contables o números índices.

d. Cuando por cualquier razón haya que utilizar los datos en función de sus logaritmos.

7. LA MEDIA ARMONICA (H)

La media armónica de un conjunto de datos es el inverso de la media aritmética de los inversos de los datos, según la siguiente fórmula:

$$H = \frac{1}{\sum\left(\frac{1/X_i}{n}\right)} = \frac{n}{\sum\left(1/X_i\right)}$$

para datos agrupados la fórmula quedaría:

$$H = \frac{n}{\sum \left(\frac{n_i}{X_i} \right)}$$

En donde:

X_i : Marcas de clase
n_i : Frecuencias absolutas

7.1 Cuando Utilizar la Media Armónica como Promedio

a. Se utiliza preferencialmente para calcular promedios de velocidad.

b. Es de gran utilidad cuando la variable está en forma de tasas de cambio, como por ejemplo galones por minuto, siempre y cuando la variable del numerador (galones) sea constante para todos los valores de la variable.

8. LA MEDIA CUADRATICA (Q)

La media cuadrática es la raíz cuadrada de la media aritmética de los cuadrados de los valores de los datos, así:

$$Q = \sqrt{\frac{\sum X_i^2}{n}}$$

Para datos agrupados, quedaría:

$$Q = \sqrt{\frac{\sum n_i X_i^2}{n}}$$

8.1 Cuando Utilizar la Media Cuadrática como Promedio

El promedio cuadrático está muy restringido a aplicaciones físicas y estadísticas. No obstante, es un promedio que se deja influenciar por valores extremos, especialmente por los grandes.

9. LA MEDIA PONDERADA (X_w)

La media aritmética ponderada está definida como la suma de los productos entre cada valor de la variable y su peso o ponderación, dividida por la suma de las ponderaciones, esto es:

$$\bar{X}_w = \frac{\sum W_i X_i}{\sum W_i}$$

Donde w_i es la ponderación de cada valor X_i

De otra manera, la media ponderada coincide con la media aritmética cuando se cambia la ponderación W_i por la frecuencia absoluta n_i, así:

$$\frac{\sum W_i X_i}{\sum W_i} = \frac{\sum n_i X_i}{\sum n_i} = \frac{\sum n_i X_i}{n} = \bar{X}$$

9.1 Cuando Utilizar la Media Ponderada como Promedio

a. Cuando se quiera calcular un promedio de medias aritméticas. En este caso el factor de ponderación debe ser el tamaño de cada muestra.

b. Para calcular números índices, en cuyo caso los precios, cantidades, costos, valores, etc., se ponderan con magnitudes de un período base ó un periodo dado.

c. Para promediar proporciones según el tamaño de los lotes o estratos.

10. OTROS PROMEDIOS, AUNQUE POCO UTILIZADOS, SON LOS SIGUIENTES:

10.1 La Media Cúbica (M_3)

Definida como la raíz cúbica de la media aritmética de los cubos de los valores de la variable, así:

$$M_3 = \sqrt[3]{\frac{\sum X_i^3}{n}}$$

10.2 El Centro Recorrido (CR)

definido como la media aritmética de los valores extremos de la variable, así

$$CR = \frac{X_{mín} + X_{má}}{2}$$

La relación entre los promedios es lo que muchos utilizan para tergiversar el verdadero sentido de la información.

10.3 La Media Recortada (MR)

definida como la media aritmética de los datos resultantes después de eliminar el n% de los datos superiores y el n% de los datos inferiores.

11. Relación entre los Promedios

La media aritmética, la media armónica y la media geométrica cumplen la siguiente relación:

Media armónica $\leq$ Media geométrica $\leq$ Media aritmética.

A su vez, la mediana siempre estará entre la media y la moda, así:

$$M_o \leq Med \leq \overline{X} \text{ ó } \overline{X} \leq Med \leq M_o$$

Según que la distribución sea sesgada a la derecha o a la izquierda.

De las relaciones anteriores se deduce que para una distribución sesgada a la derecha, la media aritmética constituye una sobrevaloración del promedio, porque:

$$\overline{X} \geq G \geq H \text{ y } \overline{X} \geq Med \geq M_o$$

De igual manera, cuando la distribución es sesgada a la izquierda, se tiene que:

$$H \leq G \leq \overline{X} \leq Med \leq M_o$$

En este caso, la moda constituye una sobrevaloración y la media armónica una subvaloración del promedio.

La relación entre los promedios es lo que muchos utilizan para tergiversar el verdadero sentido de la información.

12. ALGUNAS RECOMENDACIONES SOBRE EL USO DE LOS PROMEDIOS

12.1 Cuando se desea calcular un promedio de datos, que a su vez son promedios calculados, utilice la media ponderada.

12.2 Cuando se desea calcular un promedio de velocidades, utilice la media armónica. También se deberá utilizar cuando se quiera promediar tasas de cambio de la forma X unidades 1/ unidades 2, con la condición de que la magnitud de la unidad 1 sea constante. Si permanece constante la unidad 2, deberá utilizarse la media aritmética.

12.3 Cuando se dispone de datos organizados en una tabla de frecuencias, y la distribución sea aproximadamente simétrica, utilice la media aritmética; cuando la distribución sea sesgada, prefiera la mediana.

12.4 Si en una distribución de frecuencias se observa que el valor modal tienen una frecuencia considerablemente superior a la de todos los demás valores, utilice la moda como promedio.

> Cuando se desea calcular un "buen promedio", es necesario tener una buena información previa:
>
> - El objetivo del promedio.
> - La distribución de frecuencia de los datos.
> - El tipo de escala de medición utilizada.
> - El tipo de variable por promediar.

12.5 Cuando se desea promediar índices, cuyos valores tienen una variación geométrica, utilice el promedio geométrico. Lo mismo para el caso de promediar salarios o variaciones en el tamaño de una población.

12.6 Cuando se desea promediar porcentajes, como por ejemplo los porcentajes de defectuosos obtenidos como resultado de inspeccionar varios lotes, utilice el promedio ponderado, tomando como ponderación el tamaño de los lotes.

12.7 Cuando los datos presenten valores extremos a un lado de la distribución, evite la utilización de la media aritmética. En estos casos puede ser muy útil la media recortada, eliminando el n% de los datos extremos por exceso o por defecto.

12.8 Cuando no haya valores extremos en la distribución, y se quiera un resultado rápido, una buena opción para el promedio puede ser el centro recorrido.

12.9 Cuando se desea hacer inferencias sobre la población bajo estudio, la mejor alternativa es la media aritmética.Con otro promedio se encontrarán obstáculos insalvables.

12.10 Cuando una distribución tenga forma de U no vaya a utilizar la media aritmética, porque obtendría la contradicción de que el promedio es lo menos frecuente. Para estos casos deberá utilizarse la moda.

12.11 Cuando los extremos de la distribución no están definidos es conveniente utilizar la moda o la mediana.

12.12 Cuando se desee utilizar la media cuadrática o la media cúbica, es necesario profundizar en el estudio de la estadística y la física.

12.13 Todo conjunto de datos tiene un promedio correcto. La utilización de otros, constituye una subvaloración o sobrevaloración del mismo que puede conducir a decisiones incorrectas.

Cuando se desea calcular un "buen promedio", es necesario tener una buena información previa:

- El objetivo del promedio.
- La distribución de frecuencia de los datos.
- El tipo de escala de medición utilizada.
- El tipo de variable por promediar.

BIBLIOGRAFIA

Downie, N.M. Y Heath, R.W. Métodos estadísticos aplicados, 5a. ed. México: Editorial Harla, 1986.

Sánchez A., Javier I. Manual práctico de estadística, 4a. ed. Medellín: Universidad Nacional de Colombia, 1990.

Neter, John, y Wasserman, William. Fundamentos de estadística, 3a. ed. México: Compañía editorial Continental, S.A. 1975.

Levin, Richard I. Estadística para administradores, 2a. ed. México: Editorial Prentice-Hall Hispanoamericana, 1988.

Martínez, Biancardino, Ciro. Estadística, 6a. ed.Bogotá: Editorial ECOE, 1992.