2 Medias y porcentajes. La salsa de todas las noticias

Una de las dificultades de describir datos es informar sobre su variabilidad. Cuando se habla de cuáles son nuestros salarios, cuanto gastamos en lotería o cuanto tardamos en llegar al trabajo, es evidente que no todos ganamos ni gastamos ni tardamos lo mismo, aunque esa variabilidad a menudo se ignora usando solo la media, como si se tratara de un valor típico generalizable, cuando en muchos casos claramente no lo es.

Y tan frecuentes como las medias son los porcentajes. Están en las páginas de economía, en las de política, en las encuestas o, por supuesto, en la publicidad. Estamos tan acostumbrados a verlos que nos parecen lo más natural y facilón. No hace falta que nadie nos explique qué es un porcentaje o cómo interpretarlo, pero los porcentajes tienen un lado oscuro y a veces no son lo que parecen.

Hablando de salarios. Medias que no son típicas

Citando como fuente al INE, *ElPeriódico.com* (14-11-2014), publica un artículo sobre salarios que titula: “Los indefinidos ganan 2.048 euros de media, frente a 1.282 de los temporales” y el subtítulo dice: “El salario medio en el sector público, 2.530 euros, es superior al del sector privado, 1.700 euros”. El texto del artículo incluye también porcentajes por encima y por debajo de determinados valores según los sectores, pero lo que se destaca es la media como medida.


SEGÚN EL INE

Los indefinidos ganan 2.048 euros de media, frente a 1.282 de los temporales

El salario medio en el sector público, 2.530 euros, es superior al del sector privado,1.700 euros

Descripción de los salarios usando medias. (El Periódico versión web, 14 nov. 2014)

El artículo tiene cinco comentarios de lectores, todos críticos e incrédulos con estas cifras por considerarlas exageradas. Uno de ellos, parece que no muy creyente en la estadística, dice que: “Más inspectores de trabajo y menos encuestadores, que no sirven para nada”.

Una semana más tarde, citando al Ministerio de Hacienda, *El Mundo* (21-11-2014) destacaba en su portada el titular: “El 34% de los trabajadores españoles gana menos de 645 €”, cifra que corresponde al salario mínimo. En el texto también se comenta que el 46,4% de los trabajadores gana menos de 1.000 euros al mes. Aquí no se destacan las medias y estas cifras sí nos parecen razonables.

El 34% de los trabajadores españoles gana menos de 645 €

Sólo el 0,7% de los asalariados ingresa más de 6.500 euros al mes, según Hacienda

Descripción de los salarios usando percentiles (portada de El Mundo, 21 nov. 2014)

El problema es que "medio" no significa "típico" cuando la distribución de los valores es muy asimétrica y eso es lo que ocurre con los salarios, donde la mayoría cobra muy poco pero hay unos pocos que cobran mucho –algunos muchísimo– y son esos los que hacen subir la media dejando a la mayoría por debajo. Por ejemplo, en una empresa de 10 trabajadores si 9 cobran 500 euros al mes y uno cobra 5.500, se puede decir que cobran un promedio de 1.000 euros cada uno. Para sorpresa de muchos de ellos.

La media y la mediana

Cuando la variable que se analiza presenta una variabilidad simétrica respecto a su valor medio, como en el caso de la altura de las personas, en que cuanto más nos alejamos de la media menos individuos encontramos, tanto hacia el lado de los altos como de los bajos, sí es verdad que la media es el valor más frecuente y puede considerarse representativa. Aun así, naturalmente, no informa sobre la variabilidad, lo cual en algunas situaciones puede ser tan importante o más que la propia media. Así, si unos extraterrestres nos informan de que van a venir a visitarnos y para saber cómo son les preguntamos cuál es su altura, si nos responden que es de 1,5 metros y nos los imaginamos más pequeñitos que nosotros podemos estar cometiendo un grave error, puesto que podría haber algunos de varios centímetros y otros de dos o tres metros (que seguramente serían los que nos vendrían a visitar).

Cuando los datos son asimétricos como los salarios (aquí hay muchos "bajitos" y pocos muy altos) es mejor utilizar la mediana para describirlos. La mediana es el valor que queda en el centro cuando los datos están ordenados. Si el número de datos es par se toma el promedio de los dos centrales. Si la mediana de los salarios fuera de 800 euros significaría que una mitad cobra menos y otra mitad cobra más que esa cantidad. Así que si a usted le dicen la mediana de los salarios ya sabe si se encuentra en la mitad de los afortunados o en la mitad de los que todavía no lo son. Evidentemente esto no pasa con la media ya que la inmensa mayoría puede estar por debajo.

Para completar la mediana se pueden usar los cuartiles. Siguiendo con los valores ordenados de menor a mayor, el primer cuartil es el que deja por debajo el 25% y por encima el 75%. Para el tercer cuartil será justo al revés, deja por debajo el 75% y por encima el 25%. La mediana es el segundo cuartil pero es un valor tan singular que se utiliza un nombre específico para ella ("mediana"). De forma análoga los deciles se refieren a grupos de 10 en 10 así, el decil del 40% deja por debajo el 40% de los datos y la generalización de este sistema son los percentiles: el percentil del 22% es el valor que deja por debajo el 22% de las observaciones. Así de fácil.

Cuando se habla de salarios, o de distribución de la riqueza, es mejor usar este método de descripción aunque no hace falta citar explícitamente la mediana o los percentiles, se dice que el 50% está por debajo de tal valor y listo.

El medio pollo y la renta per cápita

Describir datos solo dando la media se puede prestar a malos entendidos, como los que se utilizan para hacer chistes sobre la estadística, como aquel que dice que cuando un señor se come un pollo y otro no come nada la estadística lo explica diciendo que en promedio se han comido medio pollo cada uno. O que si vas a la cocina de tu casa y metes la cabeza en el horno y los pies en el frigorífico un estadístico te dirá que tienes el cuerpo a la temperatura media ideal. Sin ser tan exagerados, si nos dicen que la temperatura media diaria en un lugar es de 22 grados podemos pensar que es un lugar agradable para pasar las vacaciones y, efectivamente, puede ser una bonita ciudad costera con una temperatura de 24 grados durante el día y 20 por la noche, pero también podría ser un lugar desértico donde la temperatura pasara de 50 grados durante el día a -6 por la noche.

Por cierto, una pequeña digresión sobre temperaturas, cuando se habla de media diaria muchas veces se está haciendo referencia al promedio del valor máximo y el valor mínimo –y eso no es la media–. Otras veces se dice que para calcular la temperatura media se toma varias veces al día y se hace el promedio. El problema de este sistema es que no conduce a un único valor (depende de a qué horas se tome). Seguramente existen sistemas más rigurosos pero no son los que aparecen normalmente.

Si en un país la esperanza de vida es muy baja (en algunos países está todavía por debajo de los 50 años) muchos piensan que no hay ancianos en esos países porque la gente muere en torno a los 50 años, pero esto no es así, puede haberlos y, de hecho, los hay. La esperanza de vida puede ser baja porque mueren muchos niños pero siguen habiendo ancianos. Igual que hay ricos en los países con renta muy baja. Si usted pudiera elegir el país donde nacer no debería fijarse solo en la renta per cápita. Hay países con una renta alta porque hay una minoría muy rica pero donde la mayor parte son pobres. Le convendría más elegir otro país con una renta media más modesta pero donde incluso los más pobres vivan dignamente. La muy usada renta per cápita no es más que el medio pollo que se come cada uno.

En el ámbito de la educación preocupan los resultados de las pruebas PISA que valoran diferentes aspectos del rendimiento de los estudiantes. Se realizan cada 3 años y los resultados siempre aparecen en la prensa, casi siempre destacando que todavía tenemos bastante que mejorar. Lo curioso es que solo se presentan los valores medios de cada país y nunca la dispersión. El país que obtuvo los resultados más altos en la última prueba realizada en 2012 fue Singapur con 562 puntos. La media de la OCDE fue de 496 y España obtuvo 483. En España también se presentan los resultados por comunidad autónoma, pero nunca se informa sobre la dispersión. Podría ser que en España tengamos los mejores colegios del mundo –aquellos en los que los estudiantes sacan las calificaciones más altas- pero también tengamos muchos muy malos. O quizá todos los colegios han obtenido unos valores muy similares. La dispersión es una información muy relevante, pero nunca se da.

En el ámbito industrial, si una empresa lanza a la atmósfera un producto tóxico pero 15 g al día apenas se notan, quizá la administración, siendo prudente, permita lanzar solo una media de 10 g al día. Si esa fuera la única instrucción la empresa podría no lanzar nada durante todo el año y el 31 de diciembre dejar ir 3650 g. Todos muertos pero no se podría decir que no se han cumplido las normas.

Hablando solo con medias se pueden decir cosas sorprendentes como que cada español pasa una media de 12 horas al año en la cárcel (no, esto no lo he leído en ningún periódico). Lo que sí he visto es que a veces se informa sobre lo que se supone que son valores medios pero como si esos fueran los valores que corresponde a cada individuo, como una noticia publicada en *ElNortedeCastilla.es* (14-11-2014) se lee que "Cada soriano gasta 235 euros en décimos del Gordo de Navidad". Parece que se les ha olvidado decir que eso debe ser en promedio, y lo de que los décimos son del Gordo parece muy optimista. No vi que después del sorteo se publicara que "Todos los sorianos son millonarios".

Cada soriano gasta 235 euros en décimos del Gordo de Navidad

Castilla y León vuelve a ser la comunidad que más dinero destina por habitante este sorteo, con 90,84 euros de media

ElNortedeCastilla.es, 13-11-2014

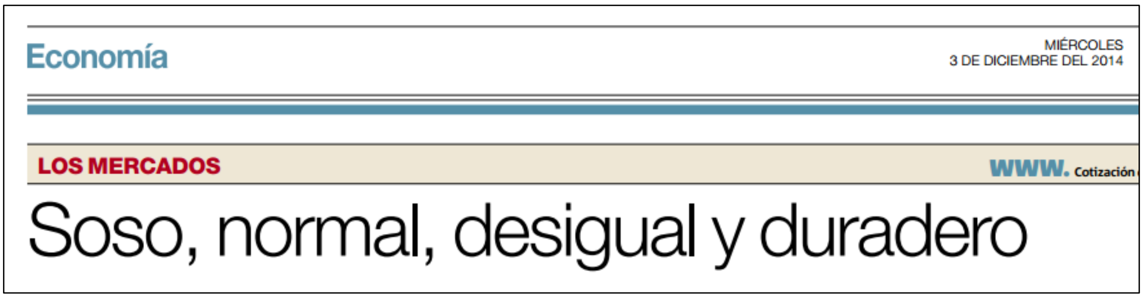
También hay que andarse con cuidado cuando se hacen cálculos con medias. Si 10 librerías dan su lista de libros más vendidos, el que aparece más veces en primer lugar no es necesariamente el más vendido. Si a una librería llega un cliente cada 5 minutos y se tarda también 5 minutos en atender a cada uno, todo funcionará como un reloj con un solo vendedor (suponiendo que no se tome ni un momento para descansar), pero a ninguna tienda llega un cliente cada 5 minutos exactamente, ni se tarda exactamente 5 minutos en atenderlo. Si lo que ocurre es que llegan cada 5 minutos en promedio y se tarda en atenderlos otros 5 minutos también en promedio, si solo tiene un vendedor el tiempo de espera crecerá rápidamente, si es que sigue entrando gente.

¿Qué significa "normal"?

Relacionada con la media aparece la palabra "normal" con un significado que no siempre está claro.

En el terreno de la estadística hemos malgastado esta palabra utilizándola para nombrar la "ley normal" o campana de Gauss, un patrón de variabilidad muy común pero que no tiene la exclusiva de la normalidad. Existen muchos otros patrones como los que explican los tiempos de vida (tanto para las personas como para las bombillas), el número de coches que llegan cada minuto a una gasolinera, o los números premiados en la lotería de Navidad, que no siguen una ley normal y no por ello se pueden calificar de anormales o raros, son como son y siempre son así. En el mundo de las matemáticas se prefiere hablar de distribución Gaussina para evitar el término "normal" con sus connotaciones.

Pero en el mundo del periodismo parece que no se afina tanto. A veces "normal" significa previsible, "como siempre", como cuando decimos que "es normal que haga frio" en el mes de enero. Ese significado parece tener en el titular "Soso, normal, desigual y duradero" (*El Periódico*, 3-12-2014) refiriéndose al comportamiento de la bolsa que parece que esos días era aburrido y sin muchos altibajos.


Economía
MIÉRCOLES
3 DE DICIEMBRE DEL 2014
LOS MERCADOS
www. Cotización
Soso, normal, desigual y duradero

"Normal" en el contexto de la bolsa usado como sinónimo de previsible. El Periódico, 3-12-2014

También puede ser sinónimo de deseable. Si le hacen un análisis de sangre y el resultado es normal todo está perfecto. Pero si usted está ilusionado con que su hijo haga carrera en el futbol profesional, si los expertos le dicen que juega "normal" la noticia ya no es tan buena. Otras veces puede significar "correcto" o "tolerable", como en el titular: "Barcelona sitúa en un 'mínimo normal' del 3% los pisos que están vacíos" (*El Periódico*, 17-1-2014)


ACTUACIONES DE POLÍTICA RESIDENCIAL
BCN sitúa en un «mínimo normal» del 3% los pisos que están vacíos

"Normal" como sinónimo de aceptable. El Periódico, 17-1-2014

Un médico comentaba que determinada dolencia es frecuente pero no es normal ¿hay que preocuparse? Por otro lado, lo que significa "anormal" tampoco está claro. Puede ser defectuoso o atípico, pero también puede ser excelente. Como ejemplo de lenguaje que se presta a varias interpretaciones se cuenta la anécdota de un profesor al que un alumno muy poco dotado para el estudio, pero voluntarioso, le pidió que le hiciera una carta de recomendación; el profesor no se quiso negar pero tampoco quería decir mentiras, así que escribió frases del estilo de “es un alumno singular, que destaca tanto por sus conocimientos como por su capacidad…”, “es uno de esos alumnos de los que te sigues acordando muchos años después…”.

En fin, la palabra “normal” se utiliza mucho pero no está siempre claro lo que se quiere decir. He consultado los libros de estilo de *El País* y de *The New York Times*, verdaderos diccionarios sobre el uso y el significado que debe darse a las palabras, pero ninguno de los dos incluye “normal”. No sé si será una manía de alguien preocupado por evitar malos entendidos en la información cuantitativa, pero creo que en algunos casos hay que andarse con cuidado con el significado de esta palabra.

Porcentajes: Poner los números en su contexto

Los valores absolutos pueden ser engañosos. Un número que parece pequeño puede ser en realidad muy grande situado en su contexto, y otro que de entrada parece muy grande puede ser pequeño. Después de una negociación laboral el titular puede ser “Los salarios subirán 30 céntimos la hora” (¡vaya miseria!) o “El dinero destinado al pago de salarios aumentará en 3 millones de euros” (¡qué barbaridad!). Las dos cifras pueden ser ciertas, aunque la impresión que dan es muy distinta. Estaría más claro si se dijera que los salarios han aumentado un 5%.

Los beneficios de algunas empresas pueden parecer escandalosos (y también pueden serlo), pero dar su valor absoluto no informa sobre si son razonables o no. Una gran empresa puede producir varios miles de millones de beneficio, pero al tener un enorme capital invertido quizá su rentabilidad sea pequeña y los accionistas busquen mejores oportunidades de forma que las acciones caerán y la empresa se irá a pique. Quizá otra empresa con unos beneficios de apenas unos millones puede ser mucho más rentable y una excelente inversión.

Algo parecido ocurre cuando se habla de la incidencia de enfermedades. *ElPais.com* (5-12-2014) publica un artículo sobre la vacuna para prevenir el cáncer de cuello de útero. Dice la entradilla que “Cada año más de 500.000 mujeres desarrollan la enfermedad en el mundo, y alrededor de 266.000 mujeres mueren por esta causa. Una cada dos minutos.” De acuerdo, pero también es cierto que según datos de la Organización Mundial de la Salud, esta enfermedad causa el 1% de las muertes entre las mujeres, lo cual no suena tan dramático (no digo que no sea importante).

> **Cáncer de cuello uterino: un desafío de género**
>
> - Cada año más de 500.000 mujeres desarrollan la enfermedad en el mundo, y alrededor de 266.000 mujeres mueren por esta causa. Una cada dos minutos. El 90% de estas muertes ocurren en países en vías de desarrollo

ElPais.com, 5-12-2014

En una línea similar, aunque en este caso se percibe un aroma publicitario, *LaVanguardia.com* (13-1-2015) publica un artículo sobre una campaña denominada “Pro-mujeres sin fractura” para concienciar a la población femenina de la importancia de prevenir las fracturas óseas causadas por la osteoporosis. El artículo empieza de la forma habitual cuando se quiere impresionar: “Cada tres segundos se produce una fractura ósea causada por la osteoporosis en el mundo, lo que representa casi nueve millones de fracturas anuales.” Sería más informativo, aunque seguramente menos espectacular, decir qué porcentaje de personas sufre una fractura debido a ese problema. Y todavía sería más interesante saber cómo cambia ese porcentaje siguiendo los consejos que se dan, pero de eso no habla.

La vida no se rompe a los 70

La campaña 'Pro-mujeres sin fractura' conciencia a la población femenina de la importancia de prevenir las fracturas óseas causadas por la osteoporosis, una enfermedad que afecta a tres millones de españoles

Sanidad | 13/01/2015 - 09:00h

LaVanguardia.com, 13-1-2015

Cálculo de porcentajes

El porcentaje siempre se refiere al valor inicial. Si antes había 4 millones de parados y ahora hay 5, el paro ha aumentado un 25%. Cuando pase de 5 a 4 disminuirá un 20%. Si la popularidad de un ministro era de 4,83 (signifique eso lo que signifique) y pasa a 3,05, ha disminuido un 36,9% (100x(4,83-3,05)/4,83) no un 58,4% tal como publicó *El País* (29-12-2013) y que es el porcentaje que debería incrementar para pasar del valor actual al que tenía antes de bajar (he de decir que *El País* informó del error, y que me enteré leyendo su fe de errores).

Los porcentajes de caída no están bien calculados. El País, 29-12-2013

Las operaciones con porcentajes no son triviales y, desde luego, no se pueden manejar como si se tratara de simples números. Hacer un 50% de descuento y después otro 20% no es lo mismo que hacer un 70%. Con números redondos: supongamos que el precio inicial es 100, con el 50% de descuento se queda en la mitad, es decir, 50, y si ahora le quitamos el 20% queda 50 – 50·0,2 = 40. Sin embargo, si se hubiera hecho el 70% de descuento habría quedado en 30.

Cuando se trabaja con porcentajes a veces aparecen resultados curiosos. Si una empresa de la competencia está vendiendo al mismo precio que usted, puede hacer un 33% de descuento y decir que la competencia cobra un 50% más (puede comprobarlo fácilmente suponiendo que el precio inicial es, por ejemplo, de 300).

La publicidad de alguna tienda de electrodomésticos dice que sus productos están de oferta y que no cobran el IVA, destacando que hacen un 21% de descuento, pero no es lo mismo no cobrar el IVA (del 21%) que hacer un 21% de descuento. Si un producto cuesta 100 euros con el IVA incluido, ¿cuánto costará sin IVA? Paso a paso: Sea X el precio sin IVA, tenemos que X + X*0,21 = 100, luego X*(1+0,21)=100; X=100/1,21 = 82,6 (¡no es 100-21=79!).

No es lo mismo devolver el importe del IVA (21%) que hacer el 21% de descuento (anuncio publicado en La Vanguardia el 12-09-2012)

A veces en los diarios se ven curiosas operaciones con porcentajes. El País (9-11-2014) publicaba un artículo que titula: “El hospital de Mataró aumenta en un 145% las derivaciones a la privada” con la entradilla: “El centro sanitario del Maresme ha duplicado en dos años los desvíos de pacientes”. El texto incluye los datos: “Mientras que en 2012 Salud desvió 155 pacientes […] la previsión para 2014 es enviar otros 380”. Efectivamente de 155 a 380 el incremento es del 145%, pero ¿de dónde sale el doble?


El hospital de Mataró aumenta en un 145% las derivaciones a la privada

El centro sanitario del Maresme ha duplicado en dos años los desvíos de pacientes

El 145% no es el doble. El País, 9-11-2014

Sara Fontdecaba y Maria Montón, en el trabajo antes citado, comentan un artículo publicado en *El Periódico* (5-1-2006) alertando sobre la desprotección infantil ante videojuegos violentos. La entradilla dice: “El 65% de los menores de 10 a 17 años admiten que acceden a programas para mayores de edad”, y en el texto aclaran que se trata del “50% de los niños y el 15% de las niñas”. Eso no es el 65 sino el 32,5%.

Alerta por la desprotección infantil ante videojuegos violentos

Amnistía Internacional dice que el sector no está bien regulado y pide al Gobierno que intervenga

El 65% de los menores de 10 a 17 años admiten que acceden a programas para mayores de edad

El 65% que se cita corresponde al 50% de los niños y el 15% de las niñas. El Periódico, 5-1-2006.

Porcentaje, ¿sobre qué?

A veces no está claro a que se refieren los números. Una conocida marca de pasta dentífrica dice que “ayuda a reducir el sangrado por inflamación hasta en un 88%”. Aparte de la ambigüedad del texto (¿”ayuda a reducir” es lo mismo que “reduce”?) y de que “hasta en un 88%” solo dice que no será mayor pero puede ser cualquier número menor, no está claro sobre qué se ha calculado ese 88%. ¿Quiere decir que si antes se sangraban 100 ml (por poner una cantidad) ahora se sangrarán 22?, ¿o que de cada 100 personas que sangraban ahora sangraran 22?

Algunos diarios publican las audiencias de los diferentes canales de TV. Dan porcentajes de audiencia pero no está claro lo que representan. ¿Qué significa que la audiencia de una cadena un determinado día ha sido del 15%? De forma similar, cuando se busca en la hemeroteca de la página web de algunos periódicos aparece un porcentaje de coincidencia para cada uno de los resultados pero no está claro lo que significa ese número. Ni siquiera está claro que sea un porcentaje.

ESPAÑA			MILES DE TELESPECTADORES
1	Tele 5	La que se avecina	3.842
2	TVE-1	Masterchef junior	3.731
3	Tele 5	La que se avecina: making of	2.867
4	Tele 5	Informativos (21.00 h.)	2.714
5	TVE-1	El tiempo 2	2.711
6	A-3	Jane Eyre (2011) (cine)	2.513
7	Tele 5	Pasapalabra	2.506
8	TVE-1	Telediario 2	2.158
9	A-3	El secreto de Puente Viejo	2.110
10	A-3	Noticias 1	2.079

El Periódico, 2 de Enero de 2014

Porcentajes y puntos porcentuales

Si las encuestas decían que un partido sacaría el 5% de los votos y ha obtenido el 10%, no ha obtenido el 5% más de lo previsto ("10%-5%") sino que ha sacado el doble, es decir, el 100% más. Cuando se comparan porcentajes la diferencia son puntos porcentuales, que no es lo mismo que porcentajes. *ElMundo.com* (23-4-2014) publica: "Wert rebaja un 10% la exigencia para no tener que devolver la beca" y en el texto se lee: "Hasta ahora, la ley decía que todos los estudiantes tenían que reintegrar el dinero de las ayudas si no aprobaban el 50% de los créditos matriculados. Y la intención de Wert es que, en estas carreras más difíciles, este porcentaje se reduzca hasta el 40%." La reducción no es del 10 sino del 20%.


EDUCACIÓN Nuevo plan de becas del Ministro de Educación

Wert rebaja un 10% la exigencia para no tener que devolver la beca

ElMundo.com, 23-4-2014

En ciertos contextos ya se entiende que se suman porcentajes. Por ejemplo, si nos dicen que la hipoteca será al Euribor más el 1% no quieren decir que si con el interés del Euribor teníamos que pagar 100 nos cobrarán 101. Quiere decir que si el Euribor es del 1% nos cobrarán el 2%, exactamente el doble. Supongo que en la letra pequeña debe estar claro, de lo contrario se podría marear a los abogados del banco.

A veces también se puede jugar con porcentajes basados en niveles o en cambios de nivel. Si ahora se dieran 9 millones de euros en becas y el gobierno se comprometiera a aumentarlas en un millón adicional pero finalmente solo aumentan 200.000 se habrían quedado solo en el 20% de lo prometido, pero también podrían decir que se comprometieron a 10 y han dado 9,2 que es quedarse en el 92%, que ya no suena tan mal (es un ejemplo, son números ficticios).

Nivel de referencia y comparaciones

Naturalmente, el porcentaje depende del valor de referencia. *ElPeriodico.com* (17-10-2014) publica que la deuda pública española hasta agosto de 2014 equivalía al 96,1% del PIB, quedando por debajo del límite fijado que era del 97,6%. El porcentaje ha bajado no porque se haya reducido la deuda sino porque ha aumentado el PIB debido a sus nuevos criterios de cálculo, que también incluye actividades ilegales como drogas y prostitución.

ADMINISTRACIONES PÚBLICAS

La deuda pública en agosto alcanza el 96,1%

- **Los datos tienen en cuenta el nuevo cálculo del PIB, que incluye drogas y prostitución, lo que ha reducido el porcentaje de endeudamiento**

ElPeriodico.com (17-10-2014)

ElMundo.com publica que solo un 6,6% de las ofertas de empleo se dirige a mayores de 45 años. Seguro que es poco, pero se podría valorar mejor si se supiera también cual es el porcentaje de demandantes que está en esa franja de edad.

ECONOMIA Informe de Adecco e Infoempleo

Sólo un 6,6% de las ofertas de empleo se dirige a mayores de 45 años

ElMundo.com, 13-10-2014

Elegir la referencia que más conviene puede ser todo un arte. Dados los beneficios de una empresa se puede decir que son el 3% de las ventas, el 15% de capital invertido, o que han disminuido un 50% respecto al año anterior.

Tal como explica S. K. Campbell en su interesante libro “Equívocos y falacias en la interpretación de estadísticas”, un fabricante de coches podría decir que el 90% de las unidades vendidas en un país en los últimos 25 años están todavía en la carretera, y ser verdad, pero también puede ser verdad que sus coches duren muy poco. Puede ser que cuando se introdujo en este país hace 25 años apenas vendía pero la red comercial se actualizó y reforzó hace 5 años, y en este periodo ha vendido el 90% de los coches.

Haciendo malabares se puede pretender justificar que es más seguro conducir con niebla que con el tiempo despejado, puesto que hay muchos más accidentes con tiempo despejado. O que si el 40% de los accidentes los protagonizan personas que han bebido, entonces el 60% -¡un porcentaje mayor!- lo tienen personas que no han bebido, luego es más seguro haber bebido. Claro que si un día circulan 1000 conductores de los cuales 4 han bebido, puede ser que de los 996 que no han bebido, seis tengan un accidente (un 0,6%) mientras que los 4 que han bebido acaben accidentados (el 100%).

Quizá en un hospital muere el 10% de los que llegan accidentados al servicio de urgencias mientras que en otro ese porcentaje es solo del 5%. Parece que si nos accidentamos rogaríamos que nos llevaran al segundo hospital, pero podría ocurrir que el primero sea un hospital grande, con muchos más medios materiales y humanos, y que ahí lleven los casos más graves mientras que el segundo es un hospital comarcal donde solo llevan los leves. Quizá un 10% en el primero es un porcentaje razonable mientras que el 5% del segundo no está justificado en absoluto.

Resumen

- La media no es buena representante de los datos si están distribuidos de forma asimétrica, como los salarios, donde hay muchos bajos y pocos altos.
- En estos casos de distribución asimétrica es mejor describir los datos haciendo referencia a porcentajes por debajo de cierto valor. Ya hay términos acuñados para esto, como la mediana, que es el valor que deja por debajo el 50%, o los percentiles.
- La media no informa sobre la variabilidad de los datos y en muchos casos, como en la distribución de la riqueza, la variabilidad también es importante.
- La palabra "normal" tiene diversas interpretaciones. No ser normal no necesariamente es malo.
- Los valores absolutos pueden ser engañosos. Es mejor contextualizarlos indicando también el porcentaje que representan.
- El porcentaje siempre debe referirse al valor inicial.
- A veces se habla de porcentajes sin que esté claro sobre qué se han calculado ni que representan.
- Las operaciones con porcentajes no son triviales. No se puede operar con porcentajes como si fueran números simplemente.
- No es lo mismo porcentajes que puntos porcentuales. Si antes los beneficios eran del 6% y han subido al 9% no han aumentado un 3% sino un 50%. Eso sí, han aumentado 3 puntos porcentuales.

Guía para políticos. Con qué compararse para poder decir que los resultados no son tan malos [posible recuadro]:

Que los políticos aseguren que van a ganar las elecciones no nos sorprende, aunque algunos se nos antojan muy optimistas. Lo que resulta más sorprendente es que se muestren entusiasmados con los resultados obtenidos cuando acaban de perder. La clave está en elegir una buena referencia con que compararse. Algunas posibilidades:

- Compararse con las previsiones: "Somos los que hemos obtenido un mejor resultado respecto a lo que decían las encuestas" o "Las encuestas decían que nos íbamos a hundir y aquí estamos, solo hemos perdido 4 diputados" (pueden ser la mitad o más de los que tenían).

- Compararse con las elecciones anteriores: "Somos los que más han subido respecto a las últimas elecciones". Pueden ser las últimas celebradas o las últimas que se hicieron del mismo tipo (europeas, generales,…), lo que más convenga. Un partido pequeño que solo tuviera 2 diputados y pase a 3 defraudando todas las expectativas de crecimiento puede decir que incrementó el número de diputados un 50%, algo que seguramente ningún partido grande ha sido capaz de conseguir.

- Compararse con un partido análogo que todavía ha perdido más. Puede ser del mismo o de otro país. El mensaje sería "no lo hemos hecho tan mal".

A veces se sacan relativamente pocos votos pero se es decisivo para decidir el color del gobierno, y en ese caso ya está claro lo que hay que destacar aunque se haya perdido un puñado de votos.

Si no hay donde cogerse, seguramente lo mejor es reconocer la derrota.

Otras medias [posible recuadro]

Supongamos que una ciudad tenía una población de 1000 habitantes pero que en los dos últimos años ha crecido mucho. El primer año un 20% y el segundo un 80%, lo cual significa que ahora tiene:

$$(1000 \times 1{,}2) \times 1{,}8 = 2160 \text{ habitantes}$$

Podría decirse que en promedio han crecido un 50% (media aritmética de 20 y 80) pero si sustituimos los incrementos que hemos considerado por el 50% tenemos:

$$(1000 \times 1{,}5) \times 1{,}5 = 2250 \text{ habitantes}$$

Este resultado es muy raro, porque de la media –como valor representativo- esperamos que se pueda sustituir por cada uno de los valores y dé el mismo resultado final, cosa que en este caso no ocurre. La media aritmética no es una buena media para resumir una tasa de crecimiento, ya sea de la bolsa, de riqueza en general o del número de parados.

Observe que los valores que indican el crecimiento se van multiplicando y, por tanto, para poderlos sustituir por un valor único habrá que hacer la raíz enésima del producto de los n

valores. Si solo tenemos dos será tan fácil como hacer la raíz cuadrada del producto. En nuestro ejemplo tendremos que:

$$\left(1000 \times \sqrt{1{,}2 \times 1{,}8}\right) \times \sqrt{1{,}2 \times 1{,}8} = 1000 \times 1{,}2 \times 1{,}8 = 2160$$

Esto es lo que llamamos media geométrica y en nuestro ejemplo es igual a $\sqrt{1{,}2 \times 1{,}8} = 1{,}47$.

Otra situación, no tan presente en la prensa pero también curiosa es el promedio de velocidades. Si dos ciudades están a 100 km si usted va a 80 km/h y vuelve a 120, ¿a qué velocidad media ha hecho todo el recorrido? La velocidad es igual al espacio recorrido partido por el tiempo tardado en recorrerlo. En nuestro caso el espacio son 200 km (la ida y la vuelta) y los tiempos son 100/80 en la ida y 100/120 a la vuelta lo cual da una velocidad global de:

$$\frac{200}{\frac{100}{80} + \frac{100}{120}} = 96 \text{ km/h}$$

A este promedio se le llama media armónica.

Contenido