

Com estudiar correctament l'edat en articles científics

Cómo estudiar correctamente la edad en artículos científicos

Pedro A. Galilea

Departament de Fisiologia i Valoració Funcional, Centre d'Alt Rendiment, Sant Cugat del Vallés.

RESUM

En aquest article s'analitza com i per què s'acostuma a fer un mal ús de l'edat en la literatura científica. S'expliquen quines són les característiques d'aquesta variable i s'aporten dues solucions per a un ús adequat de l'edat: o l'ús de taules de decimals d'any, o sumar 0.5 a l'edat legal del subjecte en estudi.

Paraules clau

Estadística, Edat, Índexs Univariants.

RESUMEN

En el presente artículo se analiza cómo y por qué se suele hacer mal uso de la edad en la literatura científica. Se explican cuáles son las características de esta variable y se aportan dos soluciones para su uso adecuado de la edad: o el uso de tablas de decimales de año, o el sumar 0.5 a la edad legal del sujeto en estudio.

Palabras clave

Estadística, Edad, Índices Univariantes.

L'edat és una de les variables més usades i, a la vegada, més maltractades a l'hora de ser estudiada per la literatura científica.

Cap article no s'oblida de referir l'edat dels subjectes que han estat estudiats, per tal de poder inferir unes conclusions que puguin ser extrapolades a subjectes d'edat semblant a la dels estudiats per altres investigadors. De fet, l'edat és una variable amb molta influència en la resposta a un determinat agent per part dels subjectes estudiats, però no ens correspon a nosaltres demostrar-ho en aquest article.

Quan en un treball d'investigació es refereix l'edat d'un grup, el més habitual és indicar els *anys complets* de cada subjecte, o edat legal, i a partir d'aquí calcular la mitjana i la desviació estàndard, o qualsevol altre índex univariant d'aquesta variable.

Aquesta metodologia té l'inconvenient de no informar-nos de manera correcta de les caracterís-

La edad es una de las variables más usadas y a la vez más maltratadas en su estudio por la literatura científica.

Ningún artículo se olvida de referir la edad de los sujetos que han sido estudiados, a fin de poder inferir unas conclusiones que puedan ser extrapoladas a sujetos de edad parecida a la de los estudiados por otros investigadores.

De hecho, la edad es una variable con mucha influencia en la respuesta a un determinado agente por parte de los sujetos estudiados, no correspondiéndonos a nosotros el demostrarlo en el presente artículo.

Cuando en un trabajo de investigación se refiere la edad de un grupo, lo más habitual es indicar los años *cumplidos* de cada sujeto, o edad legal, y a partir de aquí calcular la media y la desviación standard, o cualquier otro índice univariante, de esa variable.

tiques de la distribució de freqüències de la variable edat en el grup en estudi.

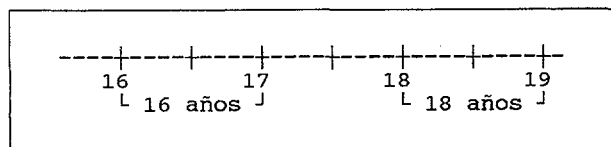
Com hem d'expressar l'edat per obtenir i donar una bona informació? Primer de tot, hem d'entendre quin tipus de variable és. Tal com habitualment és usada, anys complerts, la variable edat és una variable quantitativa discreta.

Quan ens disposem a estudiar la distribució de l'edat en un grup, de fet estem dividint la variable edat en classes, els intervals de la qual tenen com a límits tabulats de classe el dia en què es compleix l'edat e i el dia anterior a aquell en què es compleix l'edat $e+1$.

Quan llegim que un subjecte té una edat determinada, en realitat entenem que la seva edat és superior al límit tabulat inferior de l'interval de classe, sense saber si està prop d'aquest límit, prop del límit superior o bé en el centre de l'interval de classe.

L'estadística clàssica ensenya que, per calcular alguns índexs univariants (mitjana, desviació estàndard...) d'una variable a partir de la seva distribució de freqüències, és essencial la determinació del punt mitjà de la classe (Colton, 1979; Schwartz, 1985). Per al càlcul d'aquests índexs, a cada individu que es trobi dins de cada interval de classe se li adjudicarà el valor mitjà de l'interval de classe corresponent (Colton, 1979).

I això *no* és el que es fa habitualment, sinó que a cada individu se li adjudica el límit inferior de la classe (Ah!, quin mal costum aquest de treure's anys!) (Taula 1).



Taula 1. Intervals de classe, amb els seus límits de tabulació i els seus punts mitjans.

Tabla 1. Intervalos de clase, con sus límites de tabulación y sus puntos medios.

Més adequat seria adjudicar als individus inclosos en un interval determinat el valor corresponent al punt central d'aquest interval, i que en el cas present seria igual a sumar 0.5 a la seva edat legal. Però aleshores ens queda l'inconvenient que estiguem usant com a mesures decimals unes dades que en realitat no ho són: 17.3 anys no significa 17 anys i tres mesos, sinó 17 anys i 109 dies, no fàcilment comprensible.

Una altra possible opció seria treballar amb els límits tabulats de classe a la meitat de l'any, de manera que l'edat e correspongués al centre de l'interval entre $e-0.5$ i $e+0.5$, però aleshores un

Esta metodología tiene el inconveniente de no informarnos de manera correcta de las características de la distribución de frecuencias de la variable edad en el grupo en estudio.

¿Cómo debemos expresar la edad para obtener y dar una buena información?

Ante todo debemos entender qué tipo de variable es.

Tal como habitualmente es usada, años cumplidos, la variable edad es una variable cuantitativa discreta.

Quando nos disponemos a estudiar la distribución de la edad en un grupo, de hecho estamos dividiendo la variable edad en clases, cuyos intervalos tienen como límites tabulados de clase el día en que se cumple la edad e y el día anterior a aquel en que se cumple la edad $e+1$.

Quando leemos que un sujeto tiene una edad determinada, en realidad estamos entendiendo que su edad es superior al límite tabulado inferior del intervalo de clase, sin saber si se halla cerca de este límite, cerca del límite superior o bien en el centro del intervalo de clase.

La estadística clásica enseña que, para calcular algunos índices univariantes (media, desviación standard,...) de una variable a partir de su distribución de frecuencias, es esencial la determinación del punto medio de la clase (Colton, 1979; Schwartz, 1985). Para el cálculo de esos índices, a cada individuo que se halle dentro de cada intervalo de clase se le adjudicará el valor medio del intervalo de clase correspondiente (Colton, 1979).

Y eso *no* es lo que se hace habitualmente, sino que a cada individuo se le adjudica el límite inferior de la clase (Ah!, qué mala costumbre esa de quitarlos años.) (Tabla 1).

Más adecuado sería adjudicar a los individuos incluidos en un intervalo determinado el valor correspondiente al punto central de ese intervalo, y que en el caso presente sería igual a sumar 0.5 a su edad legal.

Pero entonces nos queda el inconveniente de que estamos usando como medidas decimales unos datos que en realidad no lo son: 17.3 años no significa 17 años y 3 meses, sino 17 años y 109 días, no fácilmente comprensibles.

Otra posible opción sería trabajar con los límites tabulados de clase a la mitad del año, de forma que la edad e correspondiera al centro del intervalo entre $e-0.5$ y $e+0.5$, pero entonces un sujeto tendría dos edades, una a nivel legal y otra a nivel científico, lo cual no haría más que aportar confusión al tema.

En realidad, la edad es una variable cuantitativa continua, pues está variando cada día. Su discontinuidad es sólo aparente, pues la falta de variación es tan sólo debida a la precisión con que nos expresamos.

Considero que lo más adecuado, en un trabajo científico, sería expresar la edad ayudándose de tablas de decimales de año (Ross y Marfell-Jones,

subjecte tindria dues edats, una a nivell legal i una altra a nivell científic, la qual cosa no faria més que aportar confusió al tema.

En realitat, l'edat és una variable quantitativa contínua, ja que cada dia varia. La seva discontinuïtat és solament aparent puix que la manca de variació és tan solament causada per la precisió amb què ens expressem.

Considero que, en un treball científic, el més adequat fóra expressar l'edat amb l'ajuda de taules de decimals d'any (Ross i Marfell-Jones, 1982), que permeten realitzar tota mena d'operacions matemàtiques amb ells.

Ho comprovarem amb un exemple senzill.

Suposem que tinguéssim un grup de cinc individus, les edats legals i decimals dels quals es corresponguessin amb les indicades a la Taula 2. Hi afegim una tercera columna adjudicant a cada individu el punt mitjà de l'interval de classe a què pertany ($e+0.5$).

Si calculem la mitjana i la desviació estàndard per a cada columna, obtenim diferents resultats amb significat distint, fins i tot estadísticament (tot i que no és correcte realitzar un estudi estadístic amb aquestes dades, ja que les dades de la primera columna no poden ser utilitzades per a aquest menester).

Més relleu tindrien aquestes diferències si, en base a aquests càlculs haguéssim de classificar aquests individus en categories d'acord amb la seva edat, per exemple júniors si són menors de 18 anys i sèniors si són més grans d'aquesta edat, o bé si haguéssim de comparar-los amb els d'un grup de referència en el qual l'edat fos expressada de manera correcta. Si observem les dades de la primera columna, és absolutament evident que és il·lògic que la mitjana del grup sigui inferior a 18 anys, atès que 4 sobre 5 individus estan per damunt d'aquesta edat.

Conclusió

Considero que fins ara no s'ha fet sempre un ús adequat de la variable edat en la bibliografia científica, i que aquesta variable caldria que hom procurés analitzar-la com una variable contínua, sobretot en el cas que s'hagués d'indicar índexs univariants per aportar informació sobre les característiques del grup en estudi.

Per això caldria procurar que fos expressada per mitjà de taules de decimals d'any o, en el seu defecte, considerant com a millor aproximació a l'edat real del subjecte l'edat legal actual, més de 0.5 anys.

1982), que permeten realitzar toda clase de operaciones matemáticas con ellos.

Vamos a comprobarlo con un sencillo ejemplo.

Supongamos que tuviésemos un grupo de cinco individuos, cuyas edades legales y decimales se correspondiesen con las indicadas en la Tabla 2. Añadimos una tercera columna adjudicando a cada individuo el punto medio del intervalo de clase al que pertenece ($e+0.5$).

	1) Edad legal	2) Edad decimal	3) Edad + 0.5
	18	18.874	18.5
	18	18.468	18.5
	18	18.745	18.5
	18	18.663	18.5
	17	17.482	17.5
Media	17.80	18.446	18.30
D.std.	0.45	0.559	0.45

Taula 2. Dades imaginàries d'un grup de subjectes.

Tabla 2. Datos imaginarios de un grupo de sujetos.

Si calculamos la media y la desviación standard para cada columna, obtenemos diferentes resultados con distinto significado, incluso a nivel estadístico (aunque no es correcto realizar un estudio estadístico con estos datos, pues los datos de la primera columna no pueden ser utilizados para este menester).

Mayor relevancia tendrían estas diferencias, si en base a estos cálculos debiésemos clasificar estos individuos en categorías acordes a su edad, por ejemplo juniors si menores de 18 años y seniors si mayores de esa edad, o bien debiésemos compararlos con los de un grupo de referencia en el que la edad fuese expresada de forma correcta. Si observamos los datos de la primera columna, es de todo punto evidente que es ilógico que la media del grupo sea inferior a 18 años, dado que 4 sobre 5 individuos está por encima de esa edad.

Conclusión

Considero que hasta la fecha no se ha hecho siempre un uso adecuado de la variable edad en la bibliografía científica, y que esta variable debería procurar ser analizada como una variable contínua, sobre todo en el caso de que debieran indicarse índices univariantes para aportar información sobre las características del grupo en estudio.

Por ello debería procurar ser expresada por medio de tablas de decimales de año, o, en su defecto, considerando como mejor aproximación a la edad real del sujeto la edad legal actual más 0.5 años.

Bibliografía

COLTON, TH.: *Estadística en medicina*. Ediciones Científicas y Técnicas. Barcelona, 1979.
ROSS, W.D.; MARFELL-JONES, M.J.: Kinanthropometry. En J.D. MacDougall, Wenger y H.J. Green (Eds.)

Physiological testing of the elite athlete. Canada Association of Sport Sciences, pag. 107, 1982.
SCHWARTZ, D.: *Métodos estadísticos para médicos y biólogos*. Editorial Herder, 1985.