

**CENTRO DE INVESTIGACIONES Y REFERENCIA DE
ATEROSCLEROSIS DE LA HABANA**

**UNIVERSIDAD DE CIENCIAS MÉDICAS DE
LA HABANA**

**ENSAYO CRÍTICO ACERCA DE LA MEDICIÓN DE
LAS DESIGUALDADES SOCIALES EN SALUD**

Jorge Bacallao Gallestey Dr.C

**TESIS PARA OPTAR POR EL TÍTULO DE
DOCTOR EN CIENCIAS**

La Habana 2013

Prólogo

En 1999 trabajé durante poco más de tres meses como consultor temporero en el Programa de Análisis de la Situación de Salud, en la sede de OPS-OMS, en Washington. Ya en ese momento la medición y el análisis de las desigualdades se consideraba un componente esencial del análisis de la situación de salud a partir de gran cantidad de indicios en el sentido de que la mejoría global de los indicadores de salud podría ocurrir concurrentemente con incrementos en las brechas entre grupos socioeconómicos y de que las desigualdades sociales y económicas, *per se*, ejercen una influencia negativa sobre la salud, con independencia de los niveles económicos, y del desarrollo social promedio.

Para esa fecha, el tema de las desigualdades ya se había instalado en la agenda de los salubristas, y desde entonces ha adquirido cada vez mayor visibilidad. Pocos ponen en duda en la actualidad que entre sociedades con niveles de desarrollo económico comparables, las que tienen un desarrollo más homogéneo exhiben mejores indicadores de salud.

Como fruto de aquella consultoría y de la motivación personal que ella originó, se realizaron dos publicaciones en un número especial de la Revista Panamericana de Salud Pública^{1 2} dedicadas al tema de la medición. También se realizaron varios talleres de generación de evidencias sobre desigualdades (Córdoba, Argentina, 2005; La Paz, Bolivia, 2007; Lima, Perú, 2008; La Habana, Cuba, 2008; Managua, Nicaragua, 2011) y dos nuevas publicaciones^{3 4} en la Revista Cubana de Salud Pública.

Para los estadísticos, cultivadores de un recurso analítico cuya principal materia de operación racional es la variabilidad, resulta un hecho natural que la expresión de la salud resulte incompleta si los indicadores clásicos, como las tasas y las frecuencias relativas, no se acompañan de las medidas correspondientes de dispersión. No obstante, si la convicción sobre la necesidad de medir las desigualdades es unánime, lo es mucho menos el cómo medir.

A lo largo de estos años, he ido consolidando la opinión de que la medición de las desigualdades entraña dificultades metodológicas, que no pocas veces remiten a juicios de valor. En algunos casos, estas dificultades se presentan como paradojas, que aunque no afectan la racionalidad de los recursos métricos, obligan a usarlos con cautela, para

¹ Bacallao J, Castillo-Salgado C, Schneider MC, Mujica OJ, Loyola E, Vidaurre M. Índices para medir las desigualdades de salud de carácter social basados en la noción de entropía. Rev Panam Salud Pública 2002; 12: 429-435.

² Schneider MC, Castillo-Salgado C, Bacallao J, Loyola E, Mujica OJ, Vidaurre M, Roca A. Métodos de medición de las desigualdades de salud. Rev Panam Salud Pública 2002; 12: 398-414.

³ Bacallao J. Dos temas polémicos en el contexto de la medición de las desigualdades sociales en salud. Rev Cub Salud Pública 2007, 33 (2). Disponible en http://bvs.sld.cu/revista/spu/vol33_3_07/spu307su.html.

⁴ Bacallao J. Indicadores basados en la noción de entropía para la medición de las desigualdades sociales en salud. Rev Cub Salud Pública 2007, 33 (4). Disponible en: http://bvs.sld.cu/revistas/spu/vol33_4_07/spu407su.html.

desentrañar la sustancia matemática de algunas contradicciones aparentes, saber interpretar los resultados, y sobre todo, escoger el modo adecuado de divulgarlos.

Documentar estos contenidos polémicos es un propósito importante de esta tesis. Otros no menos relevantes son, presentar una visión general del repertorio de procedimientos para medir desigualdades sanitarias y desigualdades sociales en relación con la salud, y exponer algunos recursos originales que mejoran o refinan los ya existentes o que concretan posturas teóricas personales.

La estructura del texto, poco ortodoxa tal vez, obedece a esos propósitos. En el primer capítulo se realiza una introducción a la temática y se exponen discursivamente los términos clave que harán posible una comunicación fluida con el lector. Tres cuestiones centran el interés de este capítulo: la medición de la desigualdad total o la desigualdad asociada a una clasificación socioeconómica, las disputas en torno al efecto de la desigualdad (un atributo agregado), sobre la salud (como atributo individual) y la posibilidad de la medición multidimensional de la desigualdad. El segundo capítulo se dedica a exponer los aspectos conceptuales y metodológicos básicos en la elección o evaluación de una medida de desigualdad en relación con la salud. En muchos casos, estos aspectos son de naturaleza axiológica. En el tercer capítulo se presenta un repertorio clásico de medidas de desigualdad y se destacan las cuestiones que merecen especial consideración crítica, en punto a alcance y limitaciones.

En el cuarto capítulo se exponen las novedades metodológicas de la tesis, que constituyen propiamente lo que, bajo una estructura más usual, serían los resultados y su discusión. Varias de estos aportes son adiciones o refinamientos a procedimientos ya existentes, que los complementan o permiten superar algunas de sus limitaciones. Este capítulo contiene también un novedad teórico-conceptual: la medición de la desigualdad con respecto a dimensiones socioeconómicas latentes mediante el empleo del análisis de conglomerados basados en modelos y del análisis de conglomerados basados en regresión, cuyos supuestos subyacentes se avienen mejor a la filosofía de la medición de las desigualdades sociales que el análisis de conglomerados clásico. El capítulo se complementa con un par de ejemplos ilustrativos. El quinto capítulo se dedica a mostrar algunos resultados de la medición de las desigualdades en Cuba, destacando hechos relevantes que han tenido una escasa divulgación hasta el presente y a ilustrar con ellos la aplicación de dos de los recursos metodológicos conceptualmente novedosos del capítulo precedente. En el sexto y último capítulo se realiza una presentación general de un módulo para la medición de desigualdades incluido en el software para el análisis de datos epidemiológicos tabulados Epidat en su versión 4.0 y se aplican varios de sus recursos métricos a estadísticas sanitarias nacionales. Finalmente se formulan conclusiones y recomendaciones generales. En un anexo se incluyen demostraciones importantes que no aparecen en el texto principal.

El lector advertirá, a lo largo de todo el texto, la presencia de un aliento crítico que aspira a promover una mejor apreciación del alcance y las limitaciones de todos los procedimientos.

El autor

SOBRE LA TESIS Y SUS OBJETIVOS

Este trabajo es un **ensayo sobre la medición de las desigualdades sociales en salud**,⁵ que incluye aportes metodológicos y conceptuales. Dentro de esta vasta temática, he seleccionado el tema de la medición porque es el más afín con mi formación profesional, y porque tiene incentivos desde el punto de vista metodológico y trascendencia práctica.

También dentro del campo de la medición he hecho mi propia selección de subtemas. He evitado la presentación y discusión de algunas métricas complejas con interesantes propiedades matemáticas, pero de escasa y dudosa aplicación práctica en el ámbito de la salud. Esta selección personal entraña el riesgo de alguna omisión importante, que he debido necesariamente asumir. Aunque no se hace explícito en el texto, el trabajo contiene una particular taxonomía de las medidas de desigualdad que adopté inicialmente como parte de un colectivo de autores y que años después he propuesto como coautor de un módulo para la medición de desigualdades que es parte del software libre Epidat 4.0 que se expone en un capítulo de este texto. (1-2)

El análisis de las desigualdades sociales en salud, los determinantes sociales de las desigualdades -que no necesariamente coinciden con los determinantes sociales de la salud- y las acciones en marcha en muchos países para enfrentar las desigualdades sociales, son temas de gran interés que desbordan los propósitos de este texto.

Tampoco abundan los resultados, en el sentido familiar de un producto elaborado a partir de la captura y procesamiento de datos en escenarios reales. Dentro de un enfoque casi netamente metodológico, los resultados son justamente los aportes metodológicos y conceptuales. No obstante, aquellos no están del todo ausentes: en el capítulo V se exponen aplicaciones de un enfoque multidimensional novedoso, con mediciones de las desigualdades sociales en relación con los factores de riesgo para enfermedades asociadas a la aterosclerosis, en varios consultorios de La Habana. Por otra parte, todo el repertorio seleccionado de recursos métricos se emplea en dos contextos reales de medición, uno de ellos también en Cuba, que se aprovechan para mostrar que la elección de una medida no es sólo de carácter técnico, sino que involucra también consideraciones políticas o de valor.

Otros resultados, a partir de algunos datos reales y otros ficticios, se incluyen también como aplicación del módulo de Epidat 4.0 dedicado a las desigualdades, en el sexto y último capítulo del documento.

⁵ La expresión “desigualdades sociales en salud” de dudosa corrección semántica pero ampliamente utilizada en la literatura en español, alude a desigualdades cuya medición se realiza sobre variables sanitarias (tasas de morbilidad o mortalidad, servicios de salud o recursos destinados a la salud) estratificadas según indicadores sociales. En el texto también se tratan, aunque más colateralmente, las “desigualdades sociales en relación con la salud”. La medición en este caso se realiza sobre variables sociales, con la intención de relacionar las desigualdades que éstas exhiban, con las variables sanitarias.

A estas breves consideraciones iniciales sólo resta añadir que los objetivos que se insertan a continuación guardan correspondencia con la estructura del documento por capítulos.

Objetivos (todos en relación con las desigualdades sociales en salud y su medición)

1. Exponer y discutir tres aspectos polémicos, uno de los cuales se relaciona con la necesidad de medir y dos de ellos con el modo de hacerlo.
2. Examinar los criterios y condiciones a los que se subordina la elección de una métrica.
3. Evaluar críticamente los índices que se emplean con mayor frecuencia, con énfasis especial en sus condiciones de uso, su alcance y sus limitaciones.
4. Exponer y fundamentar varias novedades metodológicas de orden instrumental y conceptual.
5. Exponer y discutir los resultados de la aplicación de procedimientos basados en el análisis de clases latentes con datos provenientes del ámbito nacional.
6. Presentar el módulo específico de Epidat 4.0 para la medición de desigualdades e ilustrar su uso con estadísticas sanitarias del ámbito nacional.

Referencias:

1. Schneider MC, Castillo-Salgado C, Bacallao J, Loyola E, Mujica OJ, Vidaurre M, Roca A. Métodos de medición de las desigualdades de salud. Rev Panam Salud Pública 2002; 12: 398-414.
2. Santiago MI, Hervada, X, Naviera G, Silva LC, Fariñas H, Vázquez E, Bacallao J, Mujica OJ. El programa EPIDAT: usos y perspectivas. Rev Panam Salud Pública 2010; 27: 80-82.

Capítulo I

Introducción

Resumen

En este capítulo se introduce la definición más aceptada de dos términos clave: desigualdad e inequidad⁶ y se describen los contextos y aplicaciones más comunes de la medición de la desigualdad. A continuación se exponen varias propiedades que la literatura sobre el tema considera atributos esenciales o deseables de las medidas de desigualdad. Finalmente la exposición se centra en tres aspectos que aun hoy día son objeto de debate: (a) el efecto de las desigualdades sobre la salud a escalas poblacional e individual, (b) la medición de la desigualdad entre individuos o entre clases, (c) la medición multidimensional de la desigualdad.

1.1 Desigualdad e inequidad

La desigualdad -que en la literatura de Norteamérica se conoce como disparidad (disparity)- es una diferencia objetiva que puede constatarse mediante el simple y único expediente de comparar dos o más objetos. En cualquier país, hay diferencias poblacionales en la distribución de la salud. Los factores genéticos y biológicos inducen diferencias en la salud individual, tanto como en los rasgos físicos. Las personas de mayor edad tienden a ser menos saludables que las más jóvenes, y las mujeres, en casi todas las sociedades, exhiben mayores expectativas de vida que los hombres.

Hay tres factores, no obstante, que transforman una simple desigualdad en una inequidad. Las inequidades son desigualdades *sistemáticas, producidas socialmente* (y por tanto modificables) *e injustas*. (1)

Las inequidades son *sistemáticas* porque no se distribuyen al azar, sino que exhiben patrones consistentes en la población. Uno de los más ostensibles es el que se presenta, en la abrumadora mayoría de los países en los que la salud no se reconoce como un derecho, en relación con la jerarquía socioeconómica: hay una relación monótona y no un simple efecto de umbral (pobres vs. ricos) entre la posición socioeconómica relativa y las condiciones de salud.

Son *inducidas por los procesos sociales*, el ejemplo siguiente basta para argumentarlo. Si María es una niña de seis años, que vive en la zona rural de Guatemala, tiene cuatro hermanos y hermanas y es hija de una campesina abandonada por su esposo y analfabeta, tiene menos oportunidades de llegar a ser una profesional o una profesora universitaria,

⁶ En español existe el término iniquidad, que en su segunda y menos frecuente acepción se refiere a las injusticias y a la falta de equidad que son objeto de estudio de la moderna Epidemiología Social. Aunque el término “inequidad” no ha sido aún admitido por la Real Academia de la Lengua Española, será el que emplearemos a lo largo de este texto para no contrariar lo que se ha convertido ya en un uso común.

que Pedro, un niño de su misma edad, que vive en Ciudad Guatemala, en compañía de su mamá, su papá (ambos con estudios secundarios y buenos ingresos) y un solo hermano.

Por último, son *injustas*, porque en general existen y se mantienen como producto de órdenes sociales injustos o que atentan contra nociones ampliamente compartidas de justicia, como por ejemplo, la noción de que todos los niños, con independencia de su grupo social de origen, deberían tener la misma probabilidad de sobrevivir o de alcanzar una educación universitaria. A causa de este último rasgo, las inequidades presuponen juicios éticos.

1.2 Contextos y aplicaciones comunes de medición de la desigualdad

La medición de la desigualdad entraña siempre una intención comparativa. Cualquier medida de desigualdad (de las que se consideran en este texto o de otras más o menos complejas) supone la existencia de un referente o de términos de comparación. Tal como ocurre con un coeficiente de correlación, un valor promedio o una varianza, las medidas de desigualdad dicen muy poco por sí mismas. Quienes se empeñen en encontrar un significado en una medición de la desigualdad *per se*, deberían tomar en cuenta este hecho.

Las circunstancias usuales de comparación son las siguientes:

- Comparaciones diacrónicas en un único escenario: estas comparaciones implican la descripción de la evolución de las desigualdades en un mismo país o en un mismo contexto. Como se verá oportunamente, esta aplicación es particularmente problemática en relación con variables dicotómicas como la muerte o la enfermedad, porque la variabilidad total cambia sólo si cambian los indicadores globales de salud.
- Comparaciones concurrentes en distintos escenarios: se refiere a comparaciones que se hacen de rutina entre países o entre regiones de un mismo país. La interpretación de estas comparaciones es muy problemática por varios motivos: (a) la existencia de factores “confusores” de las diferencias,⁷ (b) la no comparabilidad debida a diferencias en los valores promedio de los indicadores de salud, que puede ser un caso especial de “confusión” si los indicadores son tasas (porque la media y la varianza están correlacionadas entre sí), y (c) sesgos de medición en los indicadores (como cuando se comparan países en vías de desarrollo, donde las fuentes de información son no fiables, con países desarrollados, que cuentan con fuentes de información fiables).
- Comparaciones concurrentes con respecto a diferentes criterios de estratificación socioeconómica o a diferentes indicadores de salud: alude a comparaciones de las desigualdades que se practican con el fin de identificar los factores socioeconómicos a los cuales se asocian las mayores brechas, o los

⁷ El término “confusor” se utiliza en su sentido técnico, bien conocido en la Estadística: factor que se asocia simultáneamente con la variable independiente y con la dependiente y que no se encuentra en el camino causal entre ambas.

indicadores de salud en los cuales se manifiestan dichas brechas. Esta aplicación exige la estandarización previa de las variables. Su mayor escollo es, sin embargo, la dificultad de disponer de información con una calidad homogénea en relación con los factores de estratificación y con las variables de salud.

1.3 Atributos deseables para los índices de desigualdad en salud

Se exponen y fundamentan a continuación los principales atributos que hacen de un índice un buen recurso métrico. No es posible aspirar a que un índice posea todos los atributos. Ni siquiera es posible pretender identificar algún índice incondicionalmente preferible a todos los demás, debido por una parte, a que la evaluación de las bondades de un índice cualquiera se hace en función de varios atributos no siempre conciliables entre sí⁸, y por otra, como ya se ha señalado, a que el uso de un recurso métrico cualquiera se subordina a juicios de valor, se hagan o no estos explícitos en el momento de la medición.

No obstante, para cualquier índice:

- **Es indispensable que sea sensible a las desigualdades en salud:** este es un atributo elemental e inexcusable que remite a la noción básica de validez: un índice diseñado para medir desigualdades, tiene que ser capaz de ponerlas de manifiesto.
- **Es deseable que refleje la dimensión social de las desigualdades en salud:** de otro modo no sería útil para medir desigualdades sociales en salud, sino sólo desigualdades sanitarias. El vínculo con la dimensión social no queda siempre explícito en la estructura analítica del índice, sino que se consigue indirectamente cuando se usan variables sociales para caracterizar los objetos de medición. Por ejemplo: hay un vínculo explícito cuando se miden desigualdades entre grupos dados por deciles de ingreso, pero hay un vínculo indirecto cuando los municipios de un país se ordenan o clasifican según el índice de desarrollo humano o el porcentaje de población en estado de pobreza.
- **Es deseable que utilice toda la información disponible y no sólo la correspondiente a determinados grupos:** algunos índices utilizan sólo algunos grupos de la jerarquía o clasificación social, por ejemplo, los grupos extremos, o algún grupo especial que se toma como referente. Estos índices ocultan espacios de desigualdad que pueden tener especial relevancia. En general es preferible utilizar índices que abarquen a todos los grupos poblacionales definidos por el indicador social de interés.
- **Es deseable que sea sensible a cambios en la distribución de la población:** algunos índices se calculan sólo a partir del indicador de salud, sin tomar en cuenta el tamaño de los grupos. Si la composición de la población cambia, la relación entre la cantidad de personas expuestas y las no expuestas puede variar en un sentido positivo o negativo, y el índice debería ser sensible a la magnitud y a la dirección del cambio. Por ejemplo, el riesgo relativo de desnutrición crónica entre hijos de

⁸ Por ejemplo, los índices que satisfacen un mayor número de propiedades métricas suelen ser más difíciles de calcular, y sobre todo, de interpretar.

familias pobres con respecto a los hijos de familias no pobres, podría mantenerse invariante (o incluso aumentar), pese a que tanto la prevalencia de desnutrición crónica como los porcentajes de pobreza hayan disminuido. Hay índices que toman en cuenta la distribución de la población y otros que la ignoran. En general los primeros son preferibles, pero -de nuevo- en dicha preferencia intervienen juicios de valor.

- **Es deseable que sea sencillo de calcular y fácil de interpretar:** Puesto que normalmente las medidas de desigualdad se calculan para el consumo de quienes deben tomar decisiones, el uso de medidas complejas y de difícil interpretación, no es aconsejable independientemente de sus propiedades analíticas. Algunas medidas que se usan con frecuencia en el campo de la economía son menos populares en el ámbito de la salud porque son difíciles de interpretar y de representar mediante los recursos gráficos y descriptivos más comunes.
- **Es deseable que sea sensible a los cambios en el tiempo.** La evolución temporal es la aplicación por excelencia de la medición de las desigualdades. Las métricas no ponderadas y las que no toman en cuenta todo el dominio de las clases no son sensibles a cambios en el tiempo que no ocurren en los grupos extremos, o que se producen a expensas de la distribución poblacional de los grupos sociales y no de la magnitud de la variable de salud en los grupos.
- **Es deseable que el índice pueda descomponerse:** Esta propiedad puede entenderse en dos sentidos. Uno de ellos se refiere a la descomposición aditiva de la desigualdad total en un componente "dentro de grupos" y otro "entre grupos". El otro alude a la descomposición de la desigualdad como suma de las contribuciones de las diferentes dimensiones que la integran. En el primer caso, la descomposición de la desigualdad permitiría estimar qué fracción de la desigualdad es atribuible al criterio de estratificación elegido y qué porcentaje puede considerarse desigualdad marginal o residual, con lo cual se puede evaluar la pertinencia del criterio de clasificación. (2,3) En el segundo caso, la descomposición sentaría las bases para un ordenamiento de las variables en relación con su contribución a la desigualdad. (4-7)

Otros atributos interesantes pero de pertinencia más discutible son:

- **La invarianza:** es la propiedad de independencia con respecto a la escala de medición del indicador de salud. Para quienes prefieren las medidas relativas de desigualdad este es un atributo importante porque implica que si la variable de salud se multiplica por una constante (por ejemplo, se triplican uniformemente las tasas) el índice no cambia. Para quienes prefieren las medidas absolutas, la invarianza no es un atributo deseable.
- **La sensibilidad con respecto al estado de salud promedio de la población:** este es un atributo mucho más discutible. En un pasaje previo de este texto se argumentó que, debido a que la enfermedad y la muerte son atributos binarios para los fines de las estadísticas sanitarias, un cambio en la distribución de estos eventos sin una reducción en la carga total de enfermedad o muerte, implica admitir que algún grupo social tiene una carga de enfermedad o muerte inferior a la que merece, lo cual contraría cualquier noción elemental de justicia.

Incluir la salud promedio de la población en la estructura algorítmica del índice, lo haría sensible a este aspecto, pero al precio de integrar dos nociones que normalmente se evalúan por separado.

- **El principio de las transferencias o principio de Pigou-Dalton:** En el ámbito de la economía el principio de Pigou-Dalton postula que una transferencia de ingreso de una persona más rica a una más pobre, o de un grupo más rico a uno más pobre, debe resultar en una disminución de la desigualdad, siempre que la transferencia no alcance a cambiar la posición relativa de ninguna persona o grupo en la jerarquía. (8) Este principio, que en el contexto económico responde a la idea intuitiva de desigualdad, es más difícil de aceptar en el campo de la salud. La idea de redistribución o de transferencia de la salud es contraria al sentido común, debido a la dificultad de imaginar mecanismos sociales plausibles mediante los cuales podría modificarse la distribución en términos de dispersión o desigualdad, sin cambiar la medida global de salud. Hay un sentimiento general de rechazo a la idea de que algunas personas atesoren mucha riqueza mientras otras viven en pobreza extrema. De nuevo, sin embargo, es mucho más difícil aceptar la idea de injusticia intrínseca en el hecho de que alguna persona o algún grupo posean una cuota de salud mayor de la que justamente merecen, y que, por consiguiente la medida general de salud o enfermedad deba redistribuirse. Lo inaceptable no son las desigualdades en salud, sino el patrón sistemático de covariación de dichas desigualdades con el ordenamiento socioeconómico en clases sociales. El principio de las transferencias sí sería aplicable a los recursos destinados a la salud o a los servicios de salud.
- **La precisión:** Algunos autores (9-11) sugieren incorporar sistemáticamente a la medición de la desigualdad, intervalos de confianza o medidas del error estándar de los índices. En la mayoría de las aplicaciones en que los indicadores de salud son el resultado de encuestas probabilísticas, esto es posible y deseable; no así, sin embargo, cuando la medición de la desigualdad se hace con datos poblacionales en que las fuentes de información está sujeta a errores de naturaleza no aleatoria..

1.4 Las desigualdades sociales y la salud

Hace ya algunos años, el tema de las desigualdades quedó definitivamente incorporado en la agenda de los salubristas, y en las metas de las políticas públicas en gran cantidad de países. (12-15) Todo indica que las desigualdades sociales tienen un impacto negativo sobre la salud, que se añade al que ejercen las privaciones económicas y las contracciones en el desarrollo social promedio.

El tema es importante porque si no existiera tal impacto el problema de las desigualdades sociales quedaría confinado a los planos ético y político, pero no trascendería al campo de la salud pública.

Si la hubiese, la influencia podría darse a un nivel poblacional (poblaciones desiguales como agente causal de bajos niveles de salud), o a un nivel mixto (poblaciones desiguales como ambiente o contexto adverso para la salud del individuo). En el primer caso hay mayor consenso: en sociedades con un desarrollo económico similar, las menos desiguales tienen un mejor desempeño en términos de salud. (16-19)

Una interesante discrepancia tiene su origen en la existencia de un posible artefacto matemático. Algunos autores han señalado que al menos parte de la relación entre desigualdades sociales y salud podría ser sólo aparente, y debida a la asociación no lineal entre el ingreso y la salud, de modo que la transferencia de recursos económicos de los pobres a los ricos, incrementa la desigualdad total, al tiempo que produce un deterioro en la salud de los primeros mucho mayor que el incremento correspondiente en la de los segundos, con una resultante neta de disminución de los niveles poblacionales de salud. Esa asociación entre desigualdad y salud -sostienen estos autores- sería espuria y atribuible al tipo de relación funcional entre ingreso y salud. (20-22)

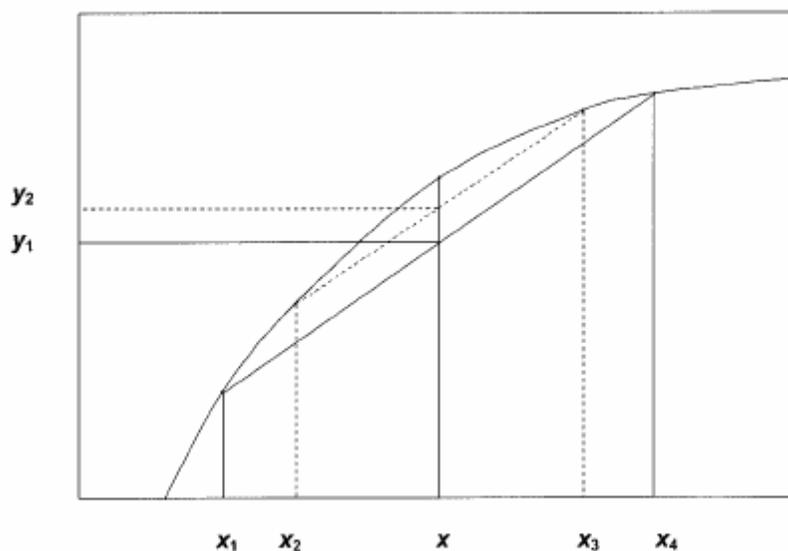


Fig. I.1 Relación (no lineal) entre \underline{x} (ingreso) y \underline{y} (salud).

En la figura 1, $x_2 - x_1 = x_4 - x_3$. Si el caudal económico de los ricos se incrementa desde x_3 hasta x_4 a expensas de un deterioro idéntico (de x_2 a x_1) en el de los pobres, el impacto negativo sobre la salud de estos últimos es mucho mayor que el efecto positivo sobre la salud de los primeros, y ello, no debido a que se ha incrementado la brecha entre pobres y ricos, como en efecto ha ocurrido, sino debido a la relación no lineal entre ingreso y salud.⁹

⁹ Si la relación fuese lineal, cualquier transferencia, con un resultado positivo para unos y negativo para otros, tendría una resultante nula en el indicador de salud. A lo largo de todo el recorrido para X, un cambio de x_1 a x_2 produciría un cambio igual a $y_2 - y_1$. (El signo dependería de la dirección del cambio).

No obstante, parece haber sobradas razones teóricas y una apreciable cantidad de sustento empírico que demuestran que el argumento de la relación no lineal es insuficiente para explicar la sistemática asociación entre desigualdades sociales y salud. (23-26)

En el segundo caso, las divergencias entre los autores son mayores, lo cual es explicable debido a la dificultad para controlar la gran cantidad de factores individuales que podrían confundir el efecto contextual de la desigualdad sobre la salud individual y derivar en la conocida "falacia ecológica" que consiste en extrapolar una asociación del nivel agregado, al nivel individual. (27)

En el orden metodológico, el modo de verificar la influencia de la desigualdad sobre la salud individual es la aplicación de modelos jerárquicos, que incluyen el efecto contextual de la desigualdad, al tiempo que controlan para otros efectos contextuales y para factores individuales. Con ese enfoque, varios trabajos recientes confirman la existencia de ese efecto contextual de las desigualdades sociales sobre la salud individual. (28-29)

Esta influencia podría deberse a que en las sociedades con grandes brechas socioeconómicas, segmentos sustanciales de la población viven en condiciones de pobreza y ésta condiciona los niveles de salud. También podría deberse a la condición económica relativa y a las reacciones que ella desencadena cuando las personas se comparan, conscientemente o no, con otras de su entorno. Esta percepción subjetiva de la condición económica o la posición social repercute negativamente sobre la salud. (30-32)

En dos investigaciones recientes en el ámbito nacional, la percepción subjetiva de la condición socioeconómica fue la variable más relevante para la formación de conglomerados con claras diferencias con respecto a factores de riesgo para las enfermedades asociadas a la aterosclerosis. En esta tesis se suscribe el punto de vista de que las desigualdades sociales tienen un efecto sobre la salud, tanto a nivel individual como poblacional, y que son por tanto de interés para la salud pública, en el plano científico y en el plano práctico. (33-34)

1.5 ¿Desigualdad entre individuos o desigualdad entre grupos?

Otro tema arduamente debatido es el dilema entre la medición de las desigualdades totales entre los miembros de una población, o las desigualdades entre grupos de esa población, definidos de acuerdo a indicadores socioeconómicos.

El elemento esencial en cualquier debate en torno a este dilema remite a juicios de valor. En las diferencias entre los seres humanos intervienen factores aleatorios, biológicos,

sociales y la interacción de todos ellos. No obstante, desde la perspectiva de una acción consecuente para eliminarlas o reducirlas, sólo los factores sociales ofrecen un incentivo para el análisis de las desigualdades. Tanto los factores biológicos como los aleatorios están determinados por leyes genéticas o estocásticas que no es posible cambiar (salvo tal vez mediante los recursos de la ingeniería genética que desbordan los propósitos de esta discusión); no así los factores sociales, que determinan desigualdades injustas y evitables, ya sea por sí mismos o en su interacción con los otros factores. Sin embargo, a partir de esta convicción compartida por los partidarios de una y otra opción, cada uno toma por senderos opuestos. Por un lado, los que proponen la medición de las desigualdades totales argumentan que el componente ético, indispensable para una definición de los términos de justicia y evitabilidad, desvirtúa la neutralidad del enfoque científico del problema; por el otro, quienes defienden el estudio de las desigualdades sociales sostienen que tal definición ética es ineludible y que es la única posibilidad de conferir algún sentido y justificación prácticos a lo que de otro modo sólo sería una mera curiosidad científica. (35)

A partir de esta toma de posición esencial, dos hechos emergen claramente a la vista. En primer lugar, que en la práctica, virtualmente cualquier elección de un criterio de estratificación social o socioeconómica, inevitablemente dejará fuera una proporción sustancial de fuentes de desigualdad. En términos más técnicos, que gran parte de las diferencias interindividuales serán diferencias "dentro de grupos" y por tanto no explicables por el criterio de agrupación elegido. En segundo lugar, y en contraposición con lo anterior, que las desigualdades totales enmascaran las desigualdades sociales (36) que son las que poseen verdadera relevancia práctica, y que no necesariamente existe una relación entre la desigualdad total y las desigualdades entre grupos. Dos estudios recientes, uno en niños (37) y otro en adultos (38) encontraron poca correspondencia entre uno y otro tipo de desigualdad en relación con la mortalidad.

Consideremos una población ficticia de 100 personas, de las cuales 20 están enfermas y tomemos como medida de desigualdad o variabilidad a la varianza.¹⁰ Puesto que estamos tratando con una variable Bernoulli que toma el valor 1 para un sujeto enfermo y 0 para uno sano, la varianza viene dada por la expresión

$$V(p) = p*(1-p) = 0.2*0.8 = 0.16 \quad (\text{I.I})$$

...en donde p es la proporción de sujetos enfermos.

Supongamos que 16 de los 20 enfermos (80%) fuesen desempleados y sólo 4, sujetos con empleo fijo, y que no hubiera ningún desempleado entre los sujetos sanos: es obvio que la

¹⁰ Desigualdad y variabilidad no son sinónimos, pero para los efectos prácticos de este análisis podemos suponer que son intercambiables. En los casos extremos de que no haya ningún enfermo o de que todos lo estén, no habría ninguna desigualdad y la variabilidad sería también nula.

estratificación y la consecuente medición de la diferencia confieren un significado completamente diferente, que la desigualdad total es incapaz de capturar.

Si se opta por la medición global de la desigualdad, se renuncia a lo que podría ser la verdadera razón de medir: poner de manifiesto desigualdades injustas y evitables. Lo que importa de la desigualdad no es ella per se, sino su covariación con factores socioeconómicos. La práctica habitual consiste en medir las desigualdades con respecto a criterios de estratificación normativos: sexo, ingreso, ocupación, raza o color de la piel, estatus migratorio y educación.

Dentro de esta opción, es inevitable que emerjan criterios subjetivos en torno a lo que es justo e injusto, evitable o inevitable. Por ejemplo, ¿es posible desligar la educación y la escolaridad de las condiciones de salud? ¿Se puede aspirar a conseguir que los segmentos poblacionales con menor educación tengan la misma conciencia de riesgo y los mismos niveles de respuesta a las políticas y programas de salud? ¿O se puede aspirar a eliminar radicalmente diferencias de género, que tienen probablemente una base predominantemente biológica?

Algunos autores han propuesto el uso de indicadores como la expectativa de vida saludable, que pretende integrar en una medida sintética todo el potencial de salud a escala individual y social, y en la que confluyen las influencias de todos o una mayoría de los indicadores socioeconómicos. Esta propuesta, que pretende incorporar implícitamente la dimensión socioeconómica en la medición, no responde a la cuestión central que se refiere a la posibilidad de utilizar la medición como base para acciones concretas dirigidas a disminuir o eliminar las desigualdades. (39)

Gakidou y King proponen una alternativa integradora que implica, en primer lugar, la elección de un indicador que resume el estado de salud de la población¹¹, y luego el empleo de un modelo beta-binomial que incluye algunos confusores como la edad y el nivel educacional de la madre, el número de hijos, y el intervalo entre nacimientos. (40) Con este enfoque, demuestran que es posible estimar los dos componentes de la desigualdad total: el componente "entre" y el componente "dentro" de grupos, cualquiera sea el criterio de estratificación que se elija.

Aunque esta última es una propuesta razonable, tropieza con una dificultad de índole práctica que la literatura no parece tomar en cuenta: ¿cómo diseñar un estudio que permita medir de manera confiable la desigualdad total? La propuesta no responde algunas interrogantes clave: ¿cómo se resuelven los problemas de factibilidad para el estudio de toda la población o la captura de información fiable y representativa? Si se recurre al muestreo, ¿cómo se muestrea de suerte que los criterios de formación de conglomerados no impliquen, a su vez, un sesgo de selección o una selección arbitraria de

¹¹ Los autores mencionan la expectativa de vida saludable al nacimiento y la probabilidad de supervivencia a los dos años de vida. Por razones prácticas, usan esta última en su artículo.

grupos poblacionales que lleve, en última instancia, a medir desigualdades sociales en lugar de desigualdades totales?

La decisión depende del propósito y las circunstancias en que se ejecuta el acto de medir. Para el análisis de la situación de salud que se hace a escala de un país o una región, con el fin de diseñar políticas de salud, lo usual es realizar el análisis agregado por estratos definidos según criterios normativos representados por indicadores socioeconómicos como el ingreso, la educación, el género, la ocupación, el color de la piel y otros similares. En este ensayo se sostiene el punto de vista de que, independientemente de la utilización de estos criterios habituales, la elección de la variable socioeconómica que genera la jerarquía o la clasificación, debe ser idealmente el resultado de hipótesis concretas sobre las causas que generan la desigualdad y debe orientarse según la perspectiva de un repertorio de acciones previsibles para reducirla o eliminarla.¹²

Casi todos los recursos métricos que se describen en esta tesis, incluyendo a la mayoría de las propuestas metodológicas novedosas, se enfocan desde esta perspectiva. Sin embargo, en estudios que tengan la finalidad central o adicional de identificar causas, es necesario el análisis de la desigualdad desagregado a nivel del individuo o de la familia.

Como se expondrá oportunamente, en este texto se recomienda un enfoque alternativo, que consiste en encontrar la estratificación que minimice la variación dentro de los grupos, con el objetivo de aproximar todo cuanto sea posible la medición de la desigualdad total y la medición asociada a grupos.

1.6 La medición global o multidimensional de la desigualdad

Otra vía que han explorado algunos autores consiste en medir las que llamaremos *desigualdades globales*, entendiendo como tales, aquellas que integran, sin distinciones en punto a relevancia, los indicadores socioeconómicos y los de salud en una única dimensión de desigualdad. La racionalidad, tanto epistemológica como práctica, que fundamenta este enfoque, es que las dimensiones de la desigualdad (económica, social, cultural, sanitaria...) no son independientes entre sí, y que sería ilusorio intentar modificar el perfil de desigualdades actuando sobre una sola de sus dimensiones, al tiempo que se dejan intactas otras que forman parte de la trama causal que las genera. (41-43)

En términos métricos, la consecuencia de este enfoque sería la medición de la desigualdad con respecto a todas sus dimensiones (o en la práctica a las más ostensibles) sin la distinción artificial de asumir subordinaciones de unas con respecto a otras.

¹² En la práctica, sin embargo, suelen emplearse de rutina criterios normativos como los grupos étnicos o los niveles de ingreso., entre otras cosas, porque las desigualdades en relación con estas dimensiones socioeconómicas serían manifiestamente injustas y evitables.

Esta alternativa es compleja. En la práctica consiste en reunir tanto los posibles criterios de estratificación como los indicadores de salud de interés, en una medición multidimensional de la desigualdad. La complejidad radica en la definición de relaciones de orden, cuando cada unidad de observación viene definida por dos o más variables y éstas no tienen una relación monótona para cualquier pareja de variables, una de las cuales sea la variable de salud.

Por ejemplo, consideremos las cuatro distribuciones ficticias que siguen, en las que se describen tres sujetos (en las filas) caracterizados por tres variables (en las columnas) de las cuales, supongamos que la de la tercera columna representa a una variable de salud.

$$X_1 = \begin{bmatrix} 9 & 10 & 11 \\ 5 & 7 & 2 \\ 3 & 5 & 1 \end{bmatrix} \quad X_2 = \begin{bmatrix} 5 & 10 & 1 \\ 3 & 5 & 11 \\ 9 & 7 & 2 \end{bmatrix}$$

$$Y_1 = \begin{bmatrix} 7.0 & 8.4 & 7.2 \\ 5.4 & 7.2 & 3.6 \\ 4.6 & 6.4 & 3.2 \end{bmatrix} \quad Y_2 = \begin{bmatrix} 5.4 & 8.4 & 3.2 \\ 4.6 & 6.4 & 7.2 \\ 7.0 & 7.2 & 3.6 \end{bmatrix}$$

No es difícil constatar que Y_1 es menos desigual que X_1 , y que Y_2 es menos desigual que X_2 , porque los rangos de variación en todas las variables, incluyendo a la variable de salud son considerablemente mayores en X_1 que en Y_1 y en X_2 que en Y_2 . Todo se complica cuando se trata de comparar X_1 con X_2 y Y_1 con Y_2 .

La literatura en torno a la medición de la desigualdad multidimensional se ha orientado hacia la búsqueda de criterios de dominancia que especifican las condiciones bajo las cuales una distribución puede considerarse al menos "tan desigual" como otra. (44-47) Como a menudo estos criterios sólo permiten ordenamientos parciales, los empates se decidirían mediante la definición de índices de desigualdad, que son funciones de un espacio multidimensional en R (el conjunto de los números reales), y que por consiguiente inducen un ordenamiento total del conjunto de distribuciones. Los índices de desigualdad se basan en principios de ordenamiento, que no sólo incorporan la noción de dispersión, sino también la noción de correlación entre los atributos.

Además de su intrínseca complejidad, estos procedimientos no toman en cuenta el rol relativo de las variables de salud (como variable dependiente) y de las variables socioeconómicas (como independientes) en el contexto de la medición de las desigualdades sociales relativas a la salud.

Referencias

1. Whitehead M, Dahlgren G. Concepts and principles for tackling social inequalities in health. World Health Organization. Studies on social and economic determinants of population health No. 2. Copenhagen 2007.
2. Bacallao J. Indicadores basados en la noción de entropía para la medición de las desigualdades sociales en salud. *Rev Cub Salud Pública* 2007; 33 (4). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-34662007000400007&lng=es&nrm=iso&tlng=es.
3. Clarke PM, Gerdtham UE, Connelly LB. A note on the decomposition of the health concentration index. *Health Economics* 2003; 12: 511-516.
4. Tsuy KY. Multidimensional poverty indices. *Social Choice and Welfare* 2002; 19: 69-93.
5. Gundgaard J, Lauridsen J. Decomposition of sources of income related health inequalities applied on SF-36 summary scores: a Danish health survey. *Health and Quality of Life Outcomes* 2006; 4: 53-60.
6. Yiengprugsawan V, Lim LLY, Carmichael GA, Sidorenko A, Sleigh AC. Measuring and decomposing inequity in self-reported morbidity and self-assessed health in Thailand. *Int J for Equity in Health*. 2007; 6: 23-40.
7. Fernández A, Costa M. Descomposición de los índices de Gini y entropía generalizada. *Estadística Española* 1998; 40: 233-256.
8. Lasso de la Vega C, Urrutia A. Characterizing multidimensional inequality measures which fulfill the Pigou-Dalton bundle principle. *Social Choice and Welfare* 2011; 35: 319-329.
9. Keppel K, Pamuk E, Lynch J et al. Methodological issues in measuring health disparities. National Center for Health Statistics. *Vital Health Stat* 2005; 2 (141).
10. APHO. Technical briefing 3. Commonly used public health statistics and their confidence intervals. Mayo 2008. Disponible en: www.apho.org.uk
11. Paoli B, Haggard L, Gulzar S. Confidence intervals in public health. Office of Public Health Assessment. Utah Department of Health. 2002
12. Woodward A, Kawachi I. Why reduce health inequalities? *J Epidemiol Community Health* 2000; 54: 923-929.
13. Marmot. M. Does money really matter? Or is it a marker for something else? *Health Aff* 2002; 21: 31-46.
14. Subramanian SV, Belli P, Karachi I. The macroeconomic determinants of health. *Annu Rev Public Health* 2002; 23: 287-302.
15. Wagstaff A, van Doorslaer E. Income inequality and health: what does the literature tell us? *Annu Rev Public Health* 2000; 21: 543-567.
16. Backlund E, Rowe G, Lynch J, Wolfson MC, Kaplan GA, Sorlie PD. Income inequality and mortality: a multilevel prospective study of 521,248 individuals in 50 US States. *Int J Epidemiol* 2007; 36: 590-596.
17. Macinko JA, Leyhu S, Starfield B, Wulu Jr JT. Income inequality and health: a critical review of the literature. *Med Care Res and Rev* 2003; 60: 407-452.

18. Subramanian SV, Kawachi I. Chasing the elusive null-the story of income inequality and health. *Int J Epidemiol* 2007; 36: 596-599.
19. Wilkinson RG, Pickett KE. *The spirit level: why more equal societies almost always do better*. New York. Penguin Books. 2009.
20. Gravelle H. How much of the relation between population mortality and unequal distribution of income is a statistical artefact? *BMJ* 1998; 316: 382-385.
21. Fiscella K, Franks P. Poverty or income inequality as predictor of mortality: longitudinal cohort study. *BMJ* 1996; 314: 1724-1727.
22. Jusot F. Income, income inequality and mortality: what do we know from the French data? *Equity III Workshop*. Helsinki. June 3-4, 2004.
23. Senn S. Societies with narrower income distributions are healthier. In: *Mortality and distribution of income (Letter)*. *BMJ* 1998; 316: 1611.2.
24. Wilkinson RG. Low relative income affects mortality. In: *Mortality and distribution of income (Letter)*. *BMJ* 1998; 316: 1611.2.
25. Sutton GC. Widening income inequalities cause poorer health. In: *Mortality and distribution of income (Letter)*. *BMJ* 1998; 316: 1611.2.
26. Gravelle H. Author's reply. In: *Mortality and distribution of income (Letter)*. *BMJ* 1998; 316: 1611.2.
27. Robinson WS. Ecological correlations and the behavior of individuals. *Int J Epidemiol* 2009; 38: 337-341
28. Kondo N, Sembajwe G, Kawachi I, van Dam RM, Subramanian SV, Yamagata Z. Income inequality, mortality and self-rated health: meta-analysis of multilevel studies. *BMJ* 2009; 339:b4471. doi: 10.1136/bmj.b4471.
29. Subramanian SV, Kawachi I. Whose health is affected by income inequality? A multilevel interaction analysis of contemporaneous and lagged effects of state income inequality on individual self-rated health in the United States. *Health Place* 2006; 12: 141-156.
30. Kondo N, Kawachi I, Subramanian SV, Takeda Y, Yamagata Z. Do social comparisons explain the association between income inequality and health? Relative deprivation and perceived health among male and female Japanese individuals. *Soc Sci Med* 2008; 67: 982-987.
31. Kondo N, Kawachi I, Hirai H, Kondo K, Subramanian SV, Hanibuchi T, et al. Relative deprivation and incident functional disability among older Japanese women and men: prospective cohort study. *J Epidemiol Comm Health* 2009; 63: 461-467.
32. Subramanian SV, Kawachi I. Income inequality and health: What have we learned so far? *Epidemiol Rev* 2004; 26: 78-91.
33. Díaz-Perera G. *Determinantes contextuales de las enfermedades asociadas a la aterosclerosis y sus factores de riesgo*. Tesis presentada en opción a la categoría de Doctora en Ciencias de la Salud. La Habana. 2011.
34. Díaz-Perera G, Bacallao J, Alemañy E. Relación entre la dimensión socioeconómica y la dimensión salud en familias cubanas. *Rev Cub Salud Pública* 2012, 38: 403-413.
35. Wagstaff A, van Doorslaer E. Overall versus socioeconomic health inequality: a measurement framework and two empirical illustrations. *Health Economics* 2004; 13: 297-301.

36. Asada Y, Hedemann T. A problem with the individual approach in the WHO health inequality measurement. *Int J Equity in Health* 2002. Disponible en: <http://www.equityhealthj.com/content/pdf/1475-9276-1-2.pdf>.
37. Braveman P, Starfield B, Geiger HJ, Murray CJL. World Health Report 2000: How it removes equity from the agenda for public health monitoring and policy commentary: Comprehensive approaches are needed for full understanding. *BMJ* 2001; 323: 678-681.
38. Houweling TAJ, Kunst AE, Mackenbach JP. World health report 2000: Inequality index and socioeconomic inequalities in mortality. *The Lancet* 2001; 357: 1671-1672.
39. Gakidou EE, Murray CJ, Frenk J. Defining and measuring health inequality: an approach based on the distribution on health expectancy. *Bull of the World Health Organization* 2000; 78: 42-54.
40. Gakidou EE, King G. Measuring total health inequality: adding individual variation to group-level differences. *Int J Equity in Health* 2002; 1, 3. Disponible en: <http://www.equityhealth.com/content/1/1/3>.
41. Araar A. The hybrid multidimensional index of inequality. *Cahier de Recherche. Working Paper 09-45. Centre Interuniversitaire sur le Risque, les Politiques Economiques et l'Emploi. CIRPEE.* 2009.
42. Weymark JA. The normative approach to measurement of multidimensional inequality. En: Farina F, Savaglio E. (Eds). *Inequality and Economic Integration.* Londres. Routledge 2006: 323-328.
43. Lahelma E, Martikainen P, Laaksonen M, Aittomäki A. Pathways between socioeconomic determinants of health. *J Epidemiol Comm Health* 2004; 58: 327-332.
44. Maasoumi E. The measurement and decomposition of multidimensional inequality. *Econometrica* 1986; 54:771-779.
45. Koshevly G, Mosler K. Multivariate Gini indices. *Journal of Multivariate Analysis* 1997; 60: 252-276.
46. Tsui KY. "Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson-Kolm-Sen approach", *Journal of Economic Theory*, 1995; 67: 251 - 265.
47. Tsui KY. "Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation", *Social Choice and Welfare* 1999; 16:145 -157.

Capítulo II

Aspectos clave en la elección de una métrica

Una paradoja es una verdad parada de cabeza para atraer la atención.

Nicholas Falletta (The Paradoxicon, pxvii, Doubleday, New York. 1983)

Resumen

En este capítulo se presentan y discuten los aspectos que deben tenerse en cuenta en la selección y evaluación de una métrica de las desigualdades con respecto a la salud. Estos aspectos representan disyuntivas metodológicas y conceptuales que en todos los casos remiten a juicios de valor o a consideraciones políticas de salud.

2.1 ¿Desigualdades absolutas o desigualdades relativas?

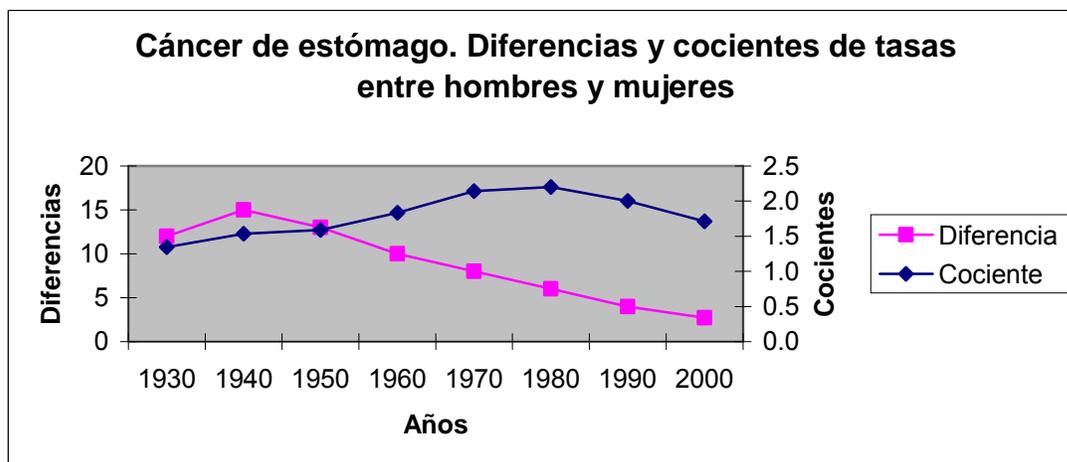
Las diferencias absolutas (diferencias propiamente) y relativas (cocientes) en los indicadores de salud de grupos sociales, constituyen el lenguaje primario del análisis de las desigualdades en Epidemiología y Salud Pública. En Epidemiología los riesgos relativos se usan con mayor frecuencia como medida de efecto. (1). Las desigualdades absolutas y relativas pueden no estar correlacionadas e incluso estarlo negativamente, como en el ejemplo de la tabla II.1, lo cual implica que del empleo de una u otra opción pueden derivarse conclusiones radicalmente diferentes. (2-5)

Los datos de la tabla y el gráfico siguientes, tomados de Wingo et al. (2003)¹³ muestran tendencias claramente diferentes en la evolución de la desigualdad en la mortalidad por cáncer de estómago entre mujeres y hombres, según se miren en términos absolutos o relativos.

Tabla II.1 Mortalidad por cáncer de estómago (tasas x 100,000) en hombres y mujeres. EEUU (1930-2000)

Año	Hombres	Mujeres	Diferencia	Cociente
1930	47	35	12	1.3
1940	43	28	15	1.5
1950	35	22	13	1.6
1960	22	12	10	1.8
1970	15	7	8	2.1
1980	11	5	6	2.2
1990	8	4	4	2.0
2000	6.5	3.8	2.7	1.7

¹³ Wingo R et al. Cancer 2003; 97:3133-3275.



Fuente: Tabla II.1

Es fácil advertir que la desigualdad relativa tiende a crecer cuando la menor de las tasas se hace más pequeña, y que si existe una diferencia entre los grupos que se comparan, dicha desigualdad relativa tiende a aumentar cuando las tasas decrecen.

Más problemático aún es el hecho que muestra el cuadro siguiente (tomado de Bacallao, 2007). (6)

Tabla II.2. Distribución de la calidad de vida en zonas urbana y rural. Cociente de tasas

Puntos de corte	Por debajo del punto de corte		Por encima del punto de corte		Cociente rural-urbano	
	Rural (%)	Urbano (%)	Rural (%)	Urbano (%)	Por debajo	Por encima
pc1	85,1	75,1	14,9	24,9	1,13	0,60
pc2	81,2	71,6	18,8	28,4	1,13	0,66
pc3	77,4	68,2	22,6	31,8	1,14	0,71
pc4	64,9	56,9	35,1	43,1	1,14	0,81
pc5	58,3	51,0	41,7	49,0	1,14	0,85
pc6	44,6	38,6	55,4	61,4	1,15	0,90
pc7	36,6	31,4	63,4	68,6	1,16	0,92
pc8	31,1	26,5	68,9	73,5	1,17	0,94
pc9	24,9	20,9	75,1	79,1	1,19	0,95
pc10	16,2	13,1	83,8	86,9	1,24	0,96
pc11	11,8	9,1	88,2	90,9	1,29	0,97

La columna de la extrema izquierda representa 11 puntos de corte, relativos a una variable latente subyacente que mide calidad de vida. El bloque de las dos primeras columnas contiene los porcentajes de población urbana y rural, respectivamente, que se encuentran por debajo del punto de corte, y las dos columnas siguientes, los respectivos porcentajes por encima del punto de corte.¹⁴ Las dos columnas de la extrema derecha contienen los cocientes rural/urbano para los porcentajes de los dos bloques anteriores. Supongamos que pc9 representa un umbral convencional para la clasificación de la "mala calidad de vida". Para ese punto de corte específico, el cociente de tasas rural/urbano para el evento adverso "tener mala calidad de vida" es 1,19.

Una intervención no discriminatoria que lograra mejorar la calidad de vida de toda la población de modo que sólo los que están por debajo del pc11 no alcanzan a mejorar su calidad de vida hasta sobrepasar el pc9, reduciría las tasas rural y urbana, respectivamente de 24,9 a 11,8 y de 20,9 a 9,1, pero incrementaría la desigualdad relativa de 1,19 a 1,29. Sin embargo, si consideramos el evento "tener buena calidad de vida", el cociente mejoraría de 0,95 a 0,97, puesto que se acercaría a 1.¹⁵

En términos absolutos, la desigualdad se reduce: en efecto, para el evento adverso la brecha se reduce, de 4,0 (24,9 - 20,9) a 2,7 (11,8 - 9,1). Obviamente, lo mismo ocurre para el evento positivo complementario.¹⁶

Algunos autores sostienen que las desigualdades relativas "casi inevitablemente" se incrementan cuando disminuye el nivel total de la variable de salud (eg. tasas de mortalidad para una enfermedad) y que, por el contrario, aumentan cuando se usa el complemento de dicha variable (eg. tasas de supervivencia para la misma enfermedad). (5). Otros afirman que las desigualdades absolutas "casi inevitablemente" disminuyen cuando disminuye el total de la variable de salud. (7)

Hay resultados empíricos que muestran que tanto las medidas absolutas como las relativas se asocian con los niveles poblacionales de salud: los niveles altos de la variable de salud en la población se acompañan de una desigualdad relativa pequeña, y de igual modo la desigualdad absoluta tiende a ser pequeña en niveles altos o bajos de la salud poblacional. (8)

Lo que para algunos es un dilema de pura sustancia matemática, para otros tiene implicaciones axiológicas, porque los juicios morales sobre la existencia de inequidad - sostienen estos últimos- se hacen sentir desde el momento en que se elige una medida de desigualdad. (9) Consecuentemente, la elección no sería neutral en punto a valores: habría juicios normativos que subyacen a la elección de una estrategia de medición en lugar de otra, en las variables que deben incluirse, y en la forma en que dichas variables

¹⁴ Por consiguiente, las columnas 2 y 4, y las columnas 3 y 5 suman 100%

¹⁵ Notar que para el evento "estar por encima del punto de corte" la desigualdad se reduce a medida que el cociente se acerca a 1 por la izquierda.

¹⁶ El lector puede comprobar que en la escala de los "odds" la desigualdad se incrementa de 1,25 a 1,33.

tendrían que ser comparadas y ponderadas. Lo usual -añaden- no es que estos criterios normativos se hagan explícitos antes de medir la desigualdad, sino que estén implícitos en la propia métrica elegida.

El uso de medidas absolutas o relativas supone un pronunciamiento acerca de si la igualdad tiene "un significado normativo independiente", (10) es decir, si es un fin en sí misma. Habría que preguntarse, por tanto, si la disminución de la desigualdad relativa es una meta deseable *per se*, al margen de los indicadores de salud globales de la población y de sus valores absolutos en cada grupo. El uso de las medidas relativas implícitamente se alinea con la posición estrictamente igualitaria de que lo que importa es la desigualdad en sí misma, más allá de cualquier otra consideración.

Bajo otros principios de equidad, la elección normativa de las medidas relativas no sería aceptable. (11) Por ejemplo, el llamado "principio débil" de la equidad en relación con la salud postula que "todas las personas o los grupos de personas deberían tener iguales niveles de salud" salvo si:

- a. Para lograr la igualdad se necesita disminuir la salud de alguien, o si
- b. Hay límites técnicos o prácticos para mejorar la salud global.

El principio del "balance justo" incluye otras dos reservas a la búsqueda a ultranza de la igualdad:

- c. Si una mayor reducción de la desigualdad entraña sacrificios inaceptables de la salud promedio de la población.
- d. Si una mayor reducción de la desigualdad entraña sacrificios inaceptables de otros bienes igualmente importantes como la educación, el empleo o la seguridad social.

En este ensayo no se toma partido por una u otra alternativa (así sucederá igualmente con las otras opciones metodológicas) sino que se defiende el punto de vista de que ambas medidas pueden ser útiles, pero que en todos los casos debe hacerse referencia a los niveles poblacionales de salud. La decisión es menos de naturaleza técnica que política o axiológica. Como aporte metodológico al debate se han calculado expresiones algebraicas simples que describen las condiciones en que pueden mejorar los niveles poblacionales de salud sin que se incrementen las desigualdades relativas o absolutas (ver apartado 4.3). (6)

2.2 ¿Cómo elegir un grupo de referencia?

Como se verá oportunamente, varias de las medidas de desigualdad que se exponen en este texto, suponen la selección de un grupo de referencia. Para ubicar al lector en el más general de los escenarios, consideremos k clases (que pueden ser unidades geodemográficas como municipios, provincias o países, o estratos de una variable socioeconómica, como deciles de ingreso o niveles de educación), y una variable de salud que caracteriza a cada clase. Si se trata de unidades geodemográficas, tendremos,

además de la variable de salud (que suele ser una tasa de mortalidad o morbilidad), una variable socioeconómica que caracteriza a cada unidad.¹⁷

En tales circunstancias, cada grupo de la clasificación puede caracterizarse en relación con un referente. Es aquí donde aparecen diferentes opciones. Las más usuales son:

- El promedio poblacional
- El mejor de los grupos
- El grupo más numeroso
- El promedio del conjunto de los grupos mejores que el grupo en cuestión.
- Un referente externo fijo

El promedio poblacional es intuitivamente es consecuente con el concepto mismo de desigualdad. Si se deja de lado cualquier consideración acerca de la salud promedio en un contexto de medición cualquiera, la ausencia de desigualdad vendría dada por una diferencia cero de cada clase con respecto a dicho promedio poblacional.

Usar como referente *al mejor de los grupos* implica asumir que todos los grupos de la sociedad tienen el potencial de alcanzar los indicadores de tal grupo. Es importante notar que “el mejor de los grupos” en relación con los indicadores de salud, puede no coincidir con el mejor de los grupos con arreglo a la variable socioeconómica. Cuando el criterio de clasificación elegido tiene una correlación por rangos desigual a 1 con el criterio basado en la salud, la diferencia con respecto al mejor grupo socioeconómico puede subestimar la verdadera magnitud de la desigualdad en salud. Por otro lado, el mejor de los grupos puede tener un tamaño relativo pequeño con respecto al resto, con lo cual las estimaciones de la desigualdad son inestables, en especial cuando se desea describir las tendencias. Con todo, en la práctica, ésta es la opción más frecuente.

Para describir las tendencias de un grupo específico de la población y trazar metas sobre bases objetivas, la elección del *conjunto de la población en mejor situación que el grupo en cuestión*, puede ser una opción preferible al promedio poblacional, porque combina una tendencia a reducir las brechas y a mejorar los indicadores de salud de la población. De nuevo es importante tomar en cuenta la pertinencia del indicador socioeconómico elegido en relación con su grado de asociación con la salud.

Elegir el *grupo más numeroso* como referente, confiere estabilidad a las estimaciones de la desigualdad, pero ignora la dimensión socioeconómica. En sociedades pobres, que suelen ser notablemente inmóviles en relación con las condiciones sociales y sanitarias, esta medida de desigualdad no es sensible a los efectos de concentración, que ocurren, por ejemplo, cuando hay una mejoría en los indicadores que sólo afecta a los grupos con mejores condiciones, pero no a la mayoría.

¹⁷ En rigor, no es necesario suponer la existencia de esta variable socioeconómica. Hacerlo revela la intención explícita de vincular la dimensión socioeconómica con las desigualdades en relación con la salud.

El empleo de un *referente fijo* da mayor estabilidad a la medición, en particular en estudios de tendencia y cuando se usan comparaciones dos a dos, pero es contrario a la idea intuitiva de desigualdad como algo privativo de un contexto interno de medición.

La disyuntiva de elegir un grupo de referencia se presenta principalmente en las medidas basadas en las comparaciones dos a dos, pero no es privativa de este tipo de medidas. La elección más frecuente es el mejor de los grupos, probablemente porque es la que más se aproxima a la práctica epidemiológica basada en el cociente de riesgos con respecto a la categoría a la que a priori se asocia el menor riesgo. Por ejemplo, en el caso del ingreso, un percentil superior extremo de la distribución;¹⁸ en el caso de la escolaridad, la clase de los universitarios.

Como en todos los casos, la elección se subordina a juicios de valor, se hagan o no explícitos dichos juicios. Por ejemplo, si se supone que es injusta e inaceptable cualquier desigualdad en los niveles de salud asociada al ingreso, el único referente consecuente sería justamente el grupo con mayores ingresos, y la desigualdad quedaría eliminada sólo cuando la diferencia con respecto a este grupo fuese cero, o el cociente fuese uno.

Si se elige el promedio poblacional, emerge de inmediato el fenómeno característico de las desigualdades sanitarias, que no se presenta en las desigualdades asociadas al ingreso. El fenómeno consiste en que, si se trata de indicadores negativos como la mortalidad, la morbilidad o la carga de enfermedad, no es posible reducir la desigualdad sin que cambie el nivel poblacional de salud, es decir, a través de una simple redistribución de las muertes o las enfermedades. Si se tratara de indicadores positivos como los servicios de salud (eg. camas hospitalarias, personal médico o inmunizaciones) sí sería posible concebir distribuciones menos desiguales que no cambien los niveles promedio en la población y en ese caso, el promedio poblacional sería un referente razonable.

Con las mismas reservas que el promedio poblacional, "el conjunto de la población en mejor situación que el grupo en cuestión" tiene dos ventajas, como grupo de referencia, que compensan la desventaja de ser una medida inusual. La primera es que su nivel de salud no está influido por el grupo con el que se compara, como sí ocurre con la media poblacional. La segunda, que al igual que el mejor grupo representa, para el grupo en cuestión, una brecha a reducir.

2.3 ¿Todos los grupos o sólo los grupos extremos?

En un contexto de medición cualquiera con k ($k > 2$) clases, es siempre deseable que la medida de desigualdad incluya la información contenida en todos los grupos. Las métricas basadas en comparaciones por pares, que incluyen, por ejemplo, sólo a los

¹⁸ Usualmente el primer quintil, el primer cuartil o el primer decil.

grupos extremos, pueden esconder fuentes importantes de desigualdad. La distorsión es aún mayor cuando los grupos son poco representativos, particularmente en términos de sus tamaños poblacionales. Sin embargo, esto no implica desentenderse de ciertos grupos específicos que pueden tener un significado especial en lo que concierne a la desigualdad. Es el caso, por ejemplo, de grupos marginales o de unidades geodemográficas con deterioro manifiesto en los indicadores de salud con respecto al resto de la población, tengan o no representatividad en cuanto al tamaño de sus poblaciones.

2.4 ¿Cómo se definen o construyen los grupos?

Mayor interés aún tiene el debate acerca de los criterios para la formación o definición de los grupos. La tendencia predominante parece ser contextual y subordinarse a fines políticos o juicios de valor. Mientras que en los Estados Unidos, los criterios de clasificación preferidos son la etnia o raza, el sexo, la condición urbana o rural y el ingreso, en Europa se concede preferencia a la ocupación y la educación (12-16). No debe olvidarse que la medición de las desigualdades cumple ante todo un propósito de abogacía, es decir, un afán de poner de manifiesto un problema para hacer visible la necesidad de afrontarlo. En tal sentido, más que las desigualdades en relación con la salud, *per se*, y sus tendencias, lo que interesa es su covariación con la condición socioeconómica. En la elección de los criterios de medición y estratificación de la condición socioeconómica, parecen predominar los sesgos culturales, las tradiciones o las coyunturas políticas, sobre otras consideraciones más objetivas.

Mi propia posición al respecto -que desarrollo con mayor profundidad en el capítulo 5- es que una crucial tarea de investigación debe ser identificar la variable latente que exhibe la mayor correlación con las desigualdades en relación con la salud. En Cuba, por ejemplo, donde el acceso a la salud se rige por principios cardinales de equidad y de igualdad de derechos, los criterios cuasi-normativos que se emplean en otros países, tendrían una débil covariación con la salud, por lo que es imperativo, en primer lugar, desvelar los factores que explican dichas desigualdades y sólo entonces hacer efectiva la medición. Aunque ese empeño desborda los propósitos de este texto, en el capítulo 6 incluyo algunos resultados preliminares en esa dirección. (17)

2.5 ¿Es importante ponderar las clases por su tamaño?

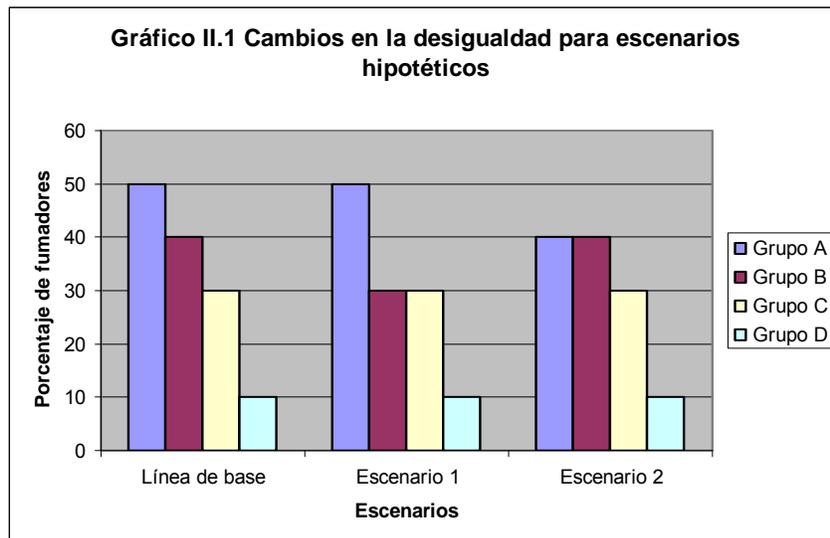
Una pregunta clave es: ¿se debe incluir el tamaño de las clases en la construcción de una métrica de la desigualdad? Algunas de las más conocidas medidas de desigualdad asignan ponderaciones idénticas a las clases, y no toman en cuenta su tamaño. Es el caso, por ejemplo, del cociente o la diferencia de tasas y del índice de Pearcy-Keppel. (18) De nuevo un juicio de valor parece estar implícito en la elección de una medida que preste o no atención al tamaño de los grupos. Para algunos, los indicadores relativos de salud de un grupo o clase poblacional, importan con independencia de su tamaño, o sea, de la cantidad de personas que están expuestas a las condiciones que prevalecen en dicho grupo. Sin embargo, resulta completamente contrario a la intuición, por ejemplo, que el cociente de tasas extremas tenga la misma importancia cuando el grupo con la tasa peor

tiene el doble de la población que el grupo con la tasa mejor, que cuando sucede exactamente lo contrario.

Más aún, si las desigualdades se miran desde la perspectiva de sus variaciones en el tiempo y de sus tendencias, un cambio positivo no se consigue sólo reduciendo las brechas entre las tasas de mortalidad o morbilidad, sino también modificando las proporciones relativas de personas expuestas a las condiciones adversas que generan, no sólo las malas condiciones de salud, sino también las propias desigualdades. Una de las novedades metodológicas que propone esta tesis consiste en incorporar la simple idea de la ponderación en algunas de las más comunes medidas de desigualdad. El tema se aborda en el capítulo 4.

2.6 ¿Se toma en cuenta la ubicación de la clase en la jerarquía socioeconómica?

Tanto o más polémico en relación con la interpretación de las medidas de desigualdad es el hecho que muestran claramente los datos ficticios del gráfico 2.1.



En la línea basal hay una clara desigualdad en relación con el porcentaje de personas sin acceso a agua potable entre cuatro grupos de una clasificación socioeconómica no ordenada. Se describen dos escenarios en los que se ha conseguido reducir la tasa global de personas sin acceso, y disminuir la brecha.¹⁹ En este ejemplo se han utilizado el índice de desigualdad de Pearcy-Keppel (IPK) y la desviación logarítmica media, cuyas expresiones analíticas se exponen en el Capítulo III. El IPK se define como el promedio de las desviaciones de la tasa de cada grupo con respecto a la menor de las tasas, con ponderaciones iguales para cada desviación y expresado dicho promedio como porcentaje con respecto a la menor de las tasas. La desviación logarítmica media pondera a cada

¹⁹ Es fácil constatar que si se emplea el cociente de tasas extremas, el escenario 1 exhibe el mismo nivel de desigualdad que la línea de base.

grupo con su tamaño, y se calcula como el promedio de las diferencias en escala logarítmica, de cada tasa con respecto a la tasa promedio de todos los grupos.

Los cambios relativos en la desigualdad de acuerdo a estos índices aparecen en la tabla 2.3.

Tabla II.3. Cambios relativos en los niveles de desigualdad de acuerdo al índice de desigualdad y a la desviación logarítmica media.

Índice	Línea basal	Escenario 1	Cambio (%)	Escenario 2	Cambio (%)
Índice de desigualdad	225	200	-11.1 %	200	-11.1 %
Desviación logarítmica media	15.5	14.7	-5.2%	13.1	-15.6%

Para simplificar la interpretación y el análisis, hemos supuesto que los grupos son del mismo tamaño, con lo cual se elimina el problema de las ponderaciones. Como puede constatarse, la única diferencia entre los escenarios 1 y 2 consiste en que la disminución de la brecha (y de la tasa global de fumadores) se ha conseguido a expensas del grupo B, en el escenario 1, y a expensas del A en el escenario 2. El índice de desigualdad es insensible a este hecho, mientras que la desviación logarítmica media, que pondera de manera diferente el cambio según el tramo de la distribución del cual provenga, y asigna una mayor contribución a la reducción de la desigualdad que proviene de los grupos con peores niveles de salud, se reduce tres veces más en el escenario 2.

Parecería una vez más, que son juicios valorativos a priori y no consideraciones estrictamente técnicas las que permitirían optar por una u otra variante: atenerse a la reducción de la desigualdad sin más, o tomar en cuenta el origen de los cambios que dan lugar a dicha reducción.

Referencias

1. Walter SD. Choice of effect measure for epidemiological data. *J Clin Epidemiol* 2000; 53: 931-939.
2. Eikemo TA, Skalická V, Avendano M. Variations in relative health inequalities: are they a mathematical artifact? *International J Equity and Health*; 2009; 8: 32-36.
3. Clarke PM, Gerdtham UG, Johannesson M, Bingefors K, Smith L. On the measurement of relative and absolute income-related health inequality. *Soc Sci Med* 2002; 55: 1923-1928.
4. Boström G, Rosen M. Measuring social inequalities in health-politics or science. *Scand J Pub Health* 2003; 31: 211-215.
5. Scanlan JP. Can we actually measure health disparities? *Chance* 2006; 19: 47-51.

6. Bacallao J. Dos temas polémicos en el contexto de la medición de las desigualdades sociales en salud. *Rev Cub Salud Pública* 2007, 33 (2). Disponible en http://bvs.sld.cu/revista/spu_vol33_3_07/spu307su.html.
7. Victora CG, Vaughan JP, Barros FC, Silva AC, Tomasi E. Explaining trends in inequities: evidence from Brazilian child health studies. *Lancet* 2000; 356: 1093-1098.
8. Houweling TAJ, Kunst AE, Huisman M, Mackenbach JP. Using relative and absolute measures for monitoring health inequalities: experiences from cross-national analyses on maternal and child health. *Int J for Equity in Health* 2007; 6:15. doi:10.1186/1475-9276-6-15.
9. Harper S. et al. Implicit value judgments in the measurement of health inequalities. *The Milbank Quarterly* 2010; 88: 4-29.
10. Norheim OS., Asada Y. The ideal of equal health revisited: definitions and measures of inequity in health should be better integrated with theories of distributed justice. *Int J for Equity in Health* 2009; 8: 40-49.
11. Temkin LS. Equality, priority or what? *Economics and Philosophy* 2003; 19: 61-87.
12. Lynch J, Davey Smith G, Harper S, Brainbridge K. J Explaining the social gradient in coronary heart disease. *J Epidemiol Community Health* 2006; 60: 436-441.
13. Mackenbach JP, Stirbu I, Roskam AJ, Schaap MM, Menvielle G, Leinsalu M, Kunst AE. Socioeconomic inequalities in health in 22 European Countries. *New England J Med* 2008; 358: 2468-2481.
14. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part I). *J Epidemiol Community Health* 2006; 60: 7-12.
15. Wagstaff A, Paci P, van Doorslaer E. On the measurement of inequalities in health. *Soc Sci Med* 1991; 33: 545-557.
16. Zhang Q, Wang Y. Socioeconomic inequality in obesity in the United States: Do gender, age and ethnicity matter? *Soc Sci Med* 2004; 58: 1171-1180.
17. Borowy I. Similar but different: Health and economic crisis in 1990's Cuba and Russia. *Soc Sci Med* 2011; 72: 1489-1498.
18. Pearcy JN, Keppel KG. A summary measure of health disparity. *Public Health Rep* 2002; 117: 273-280.

Capítulo III

Examen crítico de las medidas de desigualdad

The job of a "measure" or an "index" is to distill what is particularly relevant for our purpose, and then focus specifically on that... The central issues in devising an index relate to systematic assessment of importance. Measurement has to be integrated with evaluation. This is not an easy task"

Amartya Sen

Resumen:

En este capítulo se analiza críticamente más de una docena de índices de desigualdad social con relación a la salud. Aunque el listado no es exhaustivo, contiene aquellos que con mayor frecuencia se emplean en la literatura. Otros índices podrían descartarse, ya sea porque no satisfacen requerimientos mínimos en términos de propiedades básicas o porque resultan demasiado complejos y por tanto no aplicables en la práctica. La selección corresponde -con pocas variaciones- al inventario de estos índices incluidos en un artículo reciente de Harper y Lynch. (1) En todos los indicadores que aparezcan acompañados de un asterisco, se proponen variantes mejoradas o refinamientos metodológicos en el Capítulo IV.

I. Índices basados en comparaciones dos a dos

3.1 (*) El cociente de tasas.

Como su nombre indica, este índice mide la desigualdad entre dos grupos como el cociente de sus tasas respectivas. Formalmente

$$CT = \frac{t_1}{t_2} \quad (\text{III.1})$$

en donde t_1 y t_2 son las tasas de los grupos 1 y 2, respectivamente. El cociente de tasas es un índice fácil de calcular y de interpretar, e invariante por cambio de escala: si el numerador y el denominador se multiplican por una constante, el cociente no cambia. Sus más obvias limitaciones se enumeran a continuación:

- Si $k > 2$, CT ignora la fuente de desigualdad contenida en los $k - 2$ grupos no involucrados en su cálculo. Por ejemplo, el cociente de las tasas extremas de una jerarquía socioeconómica ignora lo que ocurre en el resto de la clasificación.
- No es sensible a cambios en la estructura poblacional porque no toma en cuenta los tamaños de los grupos: el cociente de tasas es idéntico si el grupo más pobre de la población es 10 veces mayor que el más rico o si ocurre exactamente lo contrario. Esta segunda dificultad es particularmente grave cuando se comparan diversos escenarios que no tienen la misma composición y, en especial, cuando se realizan

comparaciones en el tiempo. Es intuitivamente claro que aun cuando se mantengan constantes las tasas del grupo mejor y del peor, y por consiguiente CT permanezca invariante, la situación empeora considerablemente si aumenta el número o, en especial, la proporción de personas en el grupo peor.

- En relación con indicadores socioeconómicos en escala ordinal, el cociente de tasas extremas subestima la magnitud de la desigualdad si, a causa de las desviaciones de la linealidad, los grupos extremos según el indicador socioeconómico y el indicador de salud no coinciden.

3.2 (*) La diferencia de tasas. $(DT = t_1 - t_2)$ (III.2)

Las limitaciones señaladas con respecto al cociente, también son válidas en relación con la diferencia, que también es fácil de calcular e interpretar, pero no es invariante a cambios de escala.

3.3 (*) El riesgo atribuible poblacional. $(RAP = t_{total} - t_{(N)})$ (III.3)

En la expresión anterior $t_{(N)}$ representa la tasa del grupo con la mejor situación socioeconómica, o la tasa menor de toda la clasificación, y t_{total} , la tasa de toda la población. El RAP se interpreta como la cantidad de eventos (enfermedades o muertes (por cada 1000, 10000 o 100,000 habitantes, según la expresión de la tasa) que podrían evitarse si la tasa de la población se redujese a la del mejor de los grupos. Este es un índice fácil de calcular y formalmente análogo a otro índice homónimo de uso muy frecuente entre epidemiólogos y salubristas. (2)

A los efectos del cálculo de la desigualdad el RAP sólo utiliza la tasa global y un referente de comparación y pasa por alto la magnitud de la dispersión entre los grupos, sean éstos clases de una variable socioeconómica o unidades geodemográficas. De este modo se pasa por alto el atributo ya examinado en el apartado 2.3 del capítulo anterior. Otro tanto ocurre con la versión relativa del RAP que se define como:

3.4 Riesgo atribuible poblacional relativo: $RAP(\%) = \frac{t_{total} - t_{(N)}}{t_{total}} * 100\%$ (III.4)

Varias limitaciones de los índices basados en comparaciones dos a dos se ponen de manifiesto en las distribuciones que muestran las 4 tablas siguientes. Las distribuciones 2 a 4 son pequeñas variaciones de la distribución original, en la que 10 clases ficticias ordenadas según la magnitud de sus necesidades básicas insatisfechas se describen en términos de sus tasas de mortalidad infantil.

Tabla III.1 Tasas de mortalidad infantil en 10 clases ficticias ordenadas según necesidades básicas insatisfechas. (Distribución 1)

CLASES	NV	TMI	NBI
A	230	26,0	14,1
B	250	39,0	14,8
C	1.000	50,0	15,4
D	12.000	57,0	15,5
E	550	58,0	15,6
F	10.000	59,0	15,8
G	2.400	69,0	17,0
H	1.200	73,0	17,4
I	2.500	78,0	17,6
J	4.500	109,0	27,9

Tabla III.2 Tasas de mortalidad infantil en 10 clases ficticias ordenadas según necesidades básicas insatisfechas. (Distribución 2)

CLASES	NV	TMI	NBI
A	2.300	26,0	14,1
B	2.500	39,0	14,8
C	1.000	50,0	15,4
D	12.000	57,0	15,5
E	550	58,0	15,6
F	10.000	59,0	15,8
G	2.400	69,0	17,0
H	1.200	73,0	17,4
I	250	78,0	17,6
J	450	109,0	27,9

Tabla III.3 Tasas de mortalidad infantil en 10 clases ficticias ordenadas según necesidades básicas insatisfechas. (Distribución 3)

CLASES	NV	TMI	NBI
A	230	26,0	14,1
B	250	39,0	14,8
C	1.000	40,0	15,4
D	12.000	41,0	15,5
E	550	58,0	15,6
F	10.000	59,0	15,8
G	2.400	76,0	17,0
H	1.200	77,0	17,4
I	2.500	78,0	17,6
J	4.500	109,0	27,9

Tabla III.4 Tasas de mortalidad infantil en 10 clases ficticias ordenadas según necesidades básicas insatisfechas. (Distribución 4)

CLASES	NV	TMI	NBI
A	230	13,0	14,1
B	250	19,5	14,8
C	1.000	25,0	15,4
D	12.000	28,5	15,5
E	550	29,0	15,6
F	10.000	29,5	15,8
G	2.400	34,5	17,0
H	1.200	36,5	17,4
I	2.500	39,0	17,6
J	4.500	54,5	27,9

La diferencia entre las distribuciones 1 y 2 radica sólo en la composición poblacional de las 4 provincias extremas A, B, I y J. No cambia el valor del indicador socioeconómico, ni del indicador de salud. Se mantienen, por consiguiente, invariantes, tanto el cociente

como la diferencia de las tasas extremas. Con respecto a la distribución 1, en la 2 las tasas son más bajas porque el tamaño de la población de nacidos vivos (NV) expuesta a las condiciones más desfavorables ha disminuido considerablemente, al mismo tiempo que ha aumentado la población de NV en condiciones más favorables.

En la distribución 3, con respecto a la 1, todo el cambio ha ocurrido en las tasas de las clases C, D, G y H. En las dos primeras han disminuido las tasas y en las dos últimas han aumentado. La brecha se ha incrementado, pero ni el cociente ni la diferencia de tasas extremas son sensibles a este cambio, porque no afecta los extremos de la distribución.

En la distribución 4, con respecto a la 1, todas las tasas, incluyendo por supuesto la tasa poblacional, se han reducido a la mitad. El cociente de tasas extremas no cambia, pero la diferencia de tasas también se reduce a la mitad.

La tabla que sigue resume las estadísticas de desigualdad según los índices basados en comparaciones dos a dos.

Tabla III.5 Medidas de desigualdad dos a dos (CT, DT, RAP y RAP%) correspondientes a las distribuciones 1 a 4 de las tablas III.1 a III.4

Índices	Distribución 1	Distribución 2	Distribución 3	Distribución 4
Cociente	<i>4.2</i>	<i>4.2</i>	<i>4.2</i>	<i>4.2</i>
Diferencia	<i>83.0</i>	<i>83.0</i>	<i>83.0</i>	<i>41.5</i>
RAP	<i>40.7</i>	<i>30.2</i>	<i>35.5</i>	<i>20.4</i>
RAP (%)	<i>61.0</i>	<i>53.7</i>	<i>57.7</i>	<i>61.0</i>
Tasa global	<i>66.7</i>	<i>56.2</i>	<i>61.5</i>	<i>33.4</i>

Si fuese posible elegir una evolución temporal a partir de la distribución 1, sería natural aspirar a llegar a la distribución 4, en la que las tasas se han reducido a la mitad, pese a que la desigualdad relativa no ha cambiado.

La distribución 2 y la 3 son indudablemente mejores que la 1, aunque por razones diferentes, en ambos casos pese a que el cociente y la diferencia no han cambiado. En la 2, los valores de los indicadores socioeconómico y de salud se mantienen idénticos, pero al variar la composición poblacional en términos de las proporciones relativas de nacidos vivos sometidos a condiciones extremas (favorables y desfavorables) cambia la tasa global y con ella los riesgos atribuibles, tanto el absoluto como el relativo. La distribución 3 ha experimentado una reducción de más de 6 puntos porcentuales en la tasa global, pero ello ha sido a expensas de los grupos C y D, con menos privaciones que el G y el H, en los que las tasas se han incrementado. Como ello no ocurrió en los grupos extremos, el cociente y la diferencia no han variado, pero sí los riesgos atribuibles absoluto y relativo. En este caso, los índices de desigualdad elegidos, en su conjunto, dan sustento a la preferencia de la distribución 3 con respecto a la 1.

Mucho más problemático es el caso que describe la distribución 5.

Tabla III. 6 Tasas de mortalidad infantil en 10 clases ficticias, ordenadas según necesidades básicas insatisfechas. (Distribución 5)

CLASES	NV	TMI	NBI
A	2.300	20,0	14,1
B	2.500	39,0	14,8
C	1.000	50,0	15,4
D	12.000	57,0	15,5
E	550	58,0	15,6
F	10.000	59,0	15,8
G	2.400	69,0	17,0
H	1.200	73,0	17,4
I	250	78,0	17,6
J	450	109,0	27,9

Con respecto a la distribución inicial, la tasa global ha disminuido en casi 11 puntos porcentuales (55.8), pero salvo el RAP, todos los índices han empeorado:

- el cociente de tasas ($109 / 20 = 5.5$)
- la diferencia de tasas ($109 - 20 = 89$)
- el RAP ($55.8 - 20 = 35.8$)
- el RAP porcentual ($100 * 35.8 / 55.8 = 64.2$)

¿Qué respuestas podrían darse a las siguientes dos preguntas naturales? ¿Qué situación es preferible: la que describe la distribución 1 o la que describe la distribución 5? ¿Han aumentado, o han disminuido las desigualdades en el tránsito de la distribución 1 a la 5?

Las tasas globales podrían aportar un argumento taxativo a favor de la distribución 5; sin embargo, hay que notar que esta distribución es más desigual con respecto a casi todos los indicadores. Incluso la respuesta podría ser sutilmente diferente si les analiza como dos escenarios independientes, o si se considera que se ha llegado a la distribución 5 a partir de la 1. Una mejoría en las condiciones socioeconómicas podría cambiar la distribución relativa de la población bajo condiciones básicas insatisfechas, del modo favorable que supone el tránsito de la distribución 1 a la 5; es natural que esa evolución termine induciendo una reducción de las tasas de mortalidad infantil, que es más ostensible cuando se logra alcanzar el tope de la escala socioeconómica.

En cuanto a la magnitud de la desigualdad. ¿A cuál de los índices se presta mayor atención para formular un juicio en este sentido?

En este momento ya es oportuno adelantar dos observaciones cruciales, aun sin examinar otros índices. En primer lugar, que cuanto mayor sea la diferencia en el valor global del

indicador de salud en la población, más difícil es formular juicios comparativos en relación con la magnitud de la desigualdad. En segundo lugar, que en el ámbito de la medición de las desigualdades sociales en salud, probablemente más que en ningún otro, las diferentes opciones métricas no siempre covarían positivamente, sino que pueden conducir a valoraciones diferentes, en particular en relación con la evolución de las desigualdades.

II. Índices basados en modelos de regresión.

3.5 El índice de efecto absoluto basado en la regresión.

El índice de efecto absoluto es la pendiente β que corresponde al modelo $t_i = \alpha + \beta s_i$ en el cual

t_i : es el valor de la variable de salud en la clase i -ésima

s_i : es el valor de la variable socioeconómica en la clase i -ésima

Para la estimación de β se recomienda el uso de los mínimos cuadrados, ponderados por el tamaño relativo de las clases. La interpretación de β es la usual en el modelo de regresión: el incremento que experimenta la variable dependiente (la variable de salud) por cada unidad de incremento en la variable independiente (la variable socioeconómica).

Schneider y colaboradores (3) relacionaron las tasas de mortalidad infantil (TMI) con el producto interno bruto (PIB) ajustado por poder adquisitivo en países del área andina y encontraron la relación:

$$TMI = 75,7 - 0,007 \cdot PIB$$

De acuerdo a la interpretación clásica de los parámetros de la regresión el modelo sugiere que por cada mil dólares americanos (USD) de incremento en el producto interno bruto, se esperaría una reducción de 7 muertes por cada 1000 niños nacidos vivos.

La utilidad del índice de efecto es meramente descriptiva, salvo si existiese una relación causal entre el indicador socioeconómico y el de salud, algo que normalmente no ocurre en la práctica debido a la complejidad de los mecanismos de causalidad. Como índice de desigualdad la utilidad de β es muy limitada. Aun suponiendo que el modelo sea válido, β puede emplearse para comparar la desigualdad con respecto a varios indicadores socioeconómicos sólo a condición de que éstos vengán expresados en las mismas unidades de medición.

3.6 El índice de desigualdad de la pendiente (IDP)

El IDP es un índice relativo de desigualdad debido a que se basa en la asociación entre el indicador de salud y la posición relativa con respecto a una clasificación socioeconómica

ordenada. Una vez que las clases han sido ordenadas (de la peor a la mejor) según el indicador socioeconómico elegido, se ajusta un modelo de regresión en el que la variable dependiente es el indicador de salud, y la variable independiente es el rango del punto medio de cada clase en una escala entre 0 y 1.

Si t_i es el valor de la variable de salud y r_i la posición del punto medio de la clase i -ésima en la serie ordenada de dichas clases,²⁰ entonces el IDP se define como la pendiente (estimada por el método de los mínimos cuadrados ponderados) del modelo de regresión:

$$t_i = \alpha + \beta r_i \quad (\text{III.5})$$

El IDP puede interpretarse como la diferencia de tasas entre los puntos extremos de la escala que corresponde a los valores de $r_i = 1$ y $r_i = 0$, lo que representa la brecha teórica máxima en salud. En efecto:

Si en la expresión (I) se sustituye a r_i por los valores extremos de su escala de variación, es decir, 0 y 1 y luego se calcula la diferencia, se obtiene:

$$t_{r_i=0} = \alpha; \quad t_{r_i=1} = \alpha + \beta \quad (\text{III.6})$$

$$t_{r_i=1} - t_{r_i=0} = \beta \quad (\text{III.7})$$

El IDP no es invariante por cambios de escala en la variable de salud. Si todas las tasas se multiplican por un valor k , no cambia la desigualdad relativa y sin embargo

$$S = ks = ka + kb * r_i \quad (\text{III.8})$$

Por tal motivo, algunos autores han propuesto versiones estandarizadas de estos índices, que se conocen como índices relativos de desigualdad:

²⁰ A los valores de los r_i , se les conoce como *ridits*.

²¹ Si las t_i son, en efecto, tasas el valor de β será negativo porque las tasas más altas se presentan en los grupos con más bajo desarrollo socioeconómico. Si se tratara de un indicador de salud positivo (como el acceso a los servicios, el valor de β será positivo.

3.7 (*) El índice relativo de desigualdad de Pamuk (IRD_P)

El IRD_P se define como:

$$IRD_P = \frac{IDP}{t_{total}} \quad (\text{III.9}) \quad (4)$$

en donde t_{total} es el promedio del indicador de salud en la población (tasa promedio o tasa global de la población). Elegir entre IDP e IRD_P replantea la disyuntiva fundamental entre medidas absolutas y medidas relativas de la desigualdad. Es interesante examinar la relación (III.9): si se mantiene constante la desigualdad absoluta (la brecha absoluta entre los extremos de la jerarquía socioeconómica), la desigualdad relativa se incrementa sólo si disminuye la tasa total de la población. Por otro lado, si disminuye la desigualdad absoluta, la desigualdad relativa podría aumentar si hay una disminución suficientemente grande de la tasa total. Sin calar en la sustancia matemática que explica estas relaciones, no deja de parecer contrario a la intuición el hecho de que, al disminuir la brecha absoluta, el modo de evitar que aumente la brecha relativa es manteniendo en cotas bajas la disminución de la variable de salud en la población.

3.8 (*) El índice relativo de desigualdad de Kunst y Mackenbach (IRD_{KM})

A partir del IDP Kunst y Mackenbach (5) proponen la siguiente versión relativa:

$$IRD_{KM} = 1 - \frac{\beta}{t_{(r_i=1)}} \quad (\text{III.10})$$

En donde $t_{r_i=1}$ representa la tasa límite más baja alcanzable (en el extremo más alto de la jerarquía socioeconómica). La racionalidad de este puede apreciarse mejor a partir de la expresión

$$IRD_{KM} = \frac{\alpha}{\alpha + \beta} \quad (\text{III.11})$$

a la que se llega tras un álgebra mínima en la expresión (III.10). Es fácil ver ahora que la expresión (III.11) es justamente el cociente entre las tasas teóricas extremas.²²

El IDP y el IRD_{KM} son conceptualmente análogos a la diferencia y al cociente de tasas extremas, respectivamente, sólo que se basan en valores estimados en lugar de valores observados. La estimación se realiza a partir de un modelo de regresión que toma en

²² En el capítulo siguiente se proponen variantes mejoradas de estos dos índices.

cuenta toda la distribución y no sólo sus extremos, superando así un elemento crítico ya señalado en los índices basados en comparaciones dos a dos.

Los tres índices anteriores (IDP , IRD_p , IRD_{KM}), particularmente el último de ellos, son considerados entre las mejores opciones métricas de las desigualdades sociales en salud. Ello se debe a sus propiedades:

- Incorporan explícitamente la dimensión socioeconómica de las desigualdades en salud.
- Son fáciles de calcular e interpretar.
- Usan la información de todos los grupos socioeconómicos, incluyendo al tamaño de dichos grupos.

Estos índices, no obstante, pueden usarse únicamente en clasificaciones ordinales, cuando dicha clasificación se hace a partir de una variable socioeconómica ordinal por naturaleza (ingreso, educación) o cuando corresponde a grupos geodemográficos que han sido caracterizados y posteriormente ordenados según una variable socioeconómica que permite dicho ordenamiento.

III. Índices basados en el concepto de disparidad.

Si bien los indicadores basados en rangos son los más intuitivos y los más fáciles de calcular e interpretar, los que se basan en el concepto de disparidad son los de mayor arraigo estadístico, porque reproducen, tanto formal como conceptualmente, la noción de variabilidad, esencial en el ámbito de la Estadística.

Pese a su respaldo intuitivo estos índices están lejos de ser los que más aparecen en la literatura o los que más se emplean en la práctica. Su aplicación está indicada para la medición de la desigualdad en grupos no ordenados como los grupos étnicos o los que vienen dados por la preferencia sexual. En clasificaciones ordenadas es preferible utilizar los índices basados en el modelo de regresión o los basados en la noción de desproporcionalidad.

Los índices de este grupo toman en cuenta toda la clasificación, por lo que constituyen medidas de desigualdad global y no simples comparaciones dos a dos.

En este texto se incluyen los dos más populares: el índice de disparidad o índice de Pearcy-Keppel (IPK) y la varianza entre grupos (VEG). Los lectores con un conocimiento elemental de estadística descriptiva, encontrarán una gran similitud de estos índices con la desviación media y con la varianza, respectivamente. En el capítulo siguiente se incluyen también un par de variantes de ambos índices que podrían utilizarse indistintamente en la práctica.

3.9 (*) El índice de disparidad o índice de Pearcy-Keppel (IPK)

Se define del siguiente modo:

$$IPK = \frac{\left(\frac{1}{N} \sum_{i=1}^N |t_i - t_{ref}| \right)}{t_{ref}} \quad (III.12)$$

N es el número de clases

t_i es el valor de la tasa en el grupo i-ésimo

t_{ref} es una tasa de referencia que normalmente coincide con el mínimo de las tasas de los grupos o clases.²³

3.10 (*) La varianza entre grupos (VEG)

Se define como:

$$VEG = \sum_{i=1}^N W_i (t_i - t_{total})^2 \quad (III.13)$$

t_{total} es la tasa global de los N grupos o clases

t_i es la tasa en el grupo i-ésimo

W_i es el tamaño relativo de la i-ésima clase o grupo (la población del grupo i-ésimo dividida entre la población total)

3.11 La desviación logarítmica media (*)

Se define mediante la expresión:

$$DLM = - \sum_{i=1}^K W_i [\ln(t_i) - \ln(t_{total})] \quad (III.14)$$

Los índices de disparidad tienen el inconveniente de que no incorporan explícitamente la condición socioeconómica en su definición. En efecto, para ninguno de los tres índices de este tipo se necesita el indicador socioeconómico que caracteriza a los estratos y, por

²³ Otras opciones frecuentes son la tasa promedio o global o un referente externo.

consiguiente, también se puede prescindir del ordenamiento que genera dicho indicador. (6) Esto quiere decir que si se realizara una permutación arbitraria del ordenamiento de las clases, no cambiaría el valor de ninguno de los dos índices de disparidad. No importan, por tanto, ni el valor real, ni el relativo del indicador socioeconómico.

Se puede conseguir un vínculo indirecto con el indicador socioeconómico si éste se utiliza para construir los estratos, como ocurriría con el ingreso o la escolaridad. Sin embargo, en estos casos que generan una clasificación ordinal, los índices basados en regresión son preferibles.

El índice de Percy-Keppel, sólo toma en cuenta el valor del indicador de salud en el grupo, y no el tamaño de éste: esto implica que a la contribución de un estrato geodemográfico pequeño en términos poblacionales, se le asigna la misma importancia que a la de un estrato grande. La VEG sí tiene en cuenta el tamaño del estrato. Su peor inconveniente es que no es invariante a cambios de escala, es decir, que es sensible a cambios en el valor absoluto del indicador de salud.

En ciertas circunstancias, sin embargo, esta propiedad podría ser una ventaja en lugar de un inconveniente. Ello ocurre cuando el énfasis se pone en las desigualdades absolutas, en lugar de las relativas. (7) En efecto, la brecha relativa entre grupos se mantiene invariante si el indicador de salud se multiplica por una constante k en todos los grupos; sin embargo, la brecha absoluta disminuye si $k < 1$, y aumenta si $k > 1$. Desde la perspectiva de sus implicaciones en términos de salud pública podría ser deseable disponer de un indicador que reflejase este hecho, y en tal caso la varianza entre grupos sería una opción favorable.

La desviación logarítmica media también toma en cuenta el tamaño relativo de los grupos o estratos. En el capítulo IV se introducirá una mejora en la desviación logarítmica media para incorporar lo que Wagstaff (8) denomina "aversión" a la desigualdad, que consiste en asignar una mayor contribución a las diferencias que ocurren a expensas de los grupos con peor situación de salud.

IV. Índices basados en el concepto de entropía. (*)

En el campo de la termodinámica la entropía es una función de estado que mide el desorden de un sistema físico o químico, y por tanto su proximidad al equilibrio térmico. En cualquier transformación que se produce en un sistema aislado, la entropía nunca disminuye. Cuando un sistema aislado adquiere una configuración de entropía máxima, ya ha alcanzado el equilibrio.

Más concretamente, la entropía es la magnitud que mide la parte de la energía que no puede utilizarse para producir un trabajo, lo que equivale al grado de desorden que poseen las moléculas del sistema cerrado. De modo análogo, en la teoría de la información, la noción de entropía se emplea para medir la uniformidad en la distribución de la información en el tiempo. (9)

No es difícil advertir que estas nociones permiten construir una analogía verosímil, aplicable al ámbito de las desigualdades en salud. En efecto, si una enfermedad se distribuye uniformemente entre todas las clases de una población, hay una entropía máxima, lo cual significa que no hay desigualdades y que se ha agotado todo margen de acción para la redistribución de las condiciones de salud.

Cuando disminuyen las desigualdades en las condiciones de salud, la entropía del sistema aumenta y, por el contrario, disminuye cuando se incrementan las desigualdades. En ese caso, si se utiliza la analogía termodinámica o informática, se diría que aumenta la redundancia del sistema. Los sistemas que tienen una alta concentración de la enfermedad o en el acceso a los servicios de salud, son sistemas altamente redundantes. La "energía acumulada" en el sistema físico puede entonces homologarse con la magnitud de las acciones de salud necesarias para corregir o disminuir la redundancia del sistema.

No obstante, las primeras aplicaciones fuera de los campos originales de la termodinámica y la teoría de la información se deben a los economistas, cuando advirtieron que un sistema caracterizado por una alta concentración de la riqueza en algunos subgrupos poblacionales, es un sistema redundante, con una amplia capacidad potencial para la redistribución de la riqueza. (10-11)

El siguiente paso es de índole metodológica y consiste en definir indicadores con buenas propiedades que sean adecuados para medir el grado de uniformidad (o de falta de ella) en la distribución, ya sea del recurso económico, de la carga de enfermedad o del acceso a servicios de salud, asociada a una clasificación en grupos sociales. (12)

En el caso de la salud, este paso metodológico puede representarse del siguiente modo:

Cuadro III.1 Distribución de la población y de la carga de enfermedad en K grupos

Clases	Población	Variable de salud
1	Π_1	P_1
2	Π_2	P_2
.	.	.
.	.	.
.	.	.
K	Π_k	P_k
Total	1	1

Hay K clases socioeconómicas o K unidades geodemográficas. En cada una de ellas, la población es N_K , y la cantidad de enfermos S_K ,²⁴ de modo que:

$$\begin{aligned}
 \circ \quad N &= \sum_{i=1}^K N_i \\
 \circ \quad S &= \sum_{i=1}^K S_i \\
 \circ \quad \pi_i &= \frac{N_i}{N}; \quad P_i = \frac{S_i}{S}
 \end{aligned} \tag{III.15}$$

De lo que se trata ahora es de medir la discrepancia entre las series π_i y P_i , que son, en rigor, dos funciones de probabilidad, ya que:

$$\sum_{j=1}^K \pi_j = \sum_{j=1}^K P_j = 1 \tag{III.16}$$

Aunque no son las únicas²⁵, las dos alternativas que se presentan en este texto son el índice de disimilitud o índice de Hoover y el índice de Theil, que se definen del siguiente modo:

3.12 Índice de Hoover
$$I_H = \frac{1}{2} \sum_{i=1}^K |d_i| \tag{III.17}$$

3.13 Índice de Theil
$$I_T = \sum_{i=1}^K P_i \ln \left(\frac{P_i}{\pi_i} \right) \tag{III.18}$$

En donde:
$$d_i = \pi_i - P_i \tag{III.19}$$

Para ambos índices existen versiones estandarizadas, que toman valores en el intervalo [0 1[

Son las siguientes:

²⁴ La interpretación de S y de S_K como cantidad total de enfermos en la población y en el estrato K, respectivamente puede generalizarse a cualquier variable de salud no binaria. La argumentación en términos de una variable binaria conduce a las tasas y se hace más intuitiva y fácil de interpretar.

²⁵ Como se verá en el capítulo siguiente y en el Anexo 2, la DLM puede considerarse también un índice de la familia del Theil y el índice de Hoover. Otro bien conocido es el índice de Kullback-Liebler.

$$Z_{HOOVER} = 1 - \exp(-I_H)$$

(III.20)

$$Z_{THEIL} = 1 - \exp(-I_T)$$

A los índices basados en el concepto de entropía les falta un vínculo explícito con la condición socioeconómica. Su resultado sería el mismo si los estratos se ordenan en forma creciente, o decreciente, o si se les asigna un orden aleatorio. Cuando se aplican al estudio de las desigualdades sociales en salud, no capturan la relación funcional entre la salud y lo socioeconómico (ya sea en sentido absoluto o relativo), a diferencia de los indicadores basados en regresión.

Tal como ocurre con varios de los índices para medir desigualdades (salvo los que se basan en operaciones directas con tasas u otros indicadores de salud), los índices basados en la noción de entropía no vienen expresados en unidades de medición conocidas, ni tienen una interpretación inmediata. Ello se debe a que no hay un referente o un umbral de normalidad que permita formular juicios acerca de la magnitud de la desigualdad que estos índices están diseñados para medir, sino que su evaluación es autorreferencial. Sus condiciones fundamentales de aplicación son las comparaciones en el tiempo en escenarios similares, o las comparaciones entre diferentes variables de salud con respecto a los mismos estratos, en un escenario único.

No obstante, el índice de Hoover sí tiene una interpretación inmediata en su expresión no transformada: la proporción de observaciones que tendrían que redistribuirse para alcanzar la completa igualdad. Este hecho lo hace especialmente aplicable al estudio de las desigualdades en el campo de la salud, cuando éstas se refieren al acceso a servicios o a otro tipo de recursos. Como hacen notar Schneider y colaboradores (3) su aplicación al análisis de las desigualdades en salud, cuando se trata de indicadores negativos, como los clásicos de mortalidad o morbilidad, es cuestionable porque carece de sentido, tanto ético como práctico, hablar de redistribuciones de las enfermedades o las muertes.

El índice de Theil tiene varias propiedades importantes:

1. *Es invariante a cambios de escala:* si se incrementa la población proporcionalmente en todos los grupos o cambia la carga total de enfermedad de modo que no se modifiquen las cargas relativas de los grupos, la desigualdad se mantiene constante y, en correspondencia, también los dos índices, como es deseable, ya que no se ha modificado la diferencia entre los grupos (12)
2. *Es sensible a cambios en cualquiera de las clases o estratos* y no sólo en los grupos extremos, como los índices basados en rangos (12)
3. *La descomponibilidad:* El índice de Theil permite relacionar la desigualdad total en una población con la desigualdad atribuible a grupos sociales, mediante una relación

aditiva semejante a la que caracteriza al análisis de la varianza. Si una población P se divide en k clases, es posible demostrar que:

$$I_T(\text{total}) = I_T(\text{"entre"}) + I_T(\text{"dentro"}) \quad (\text{III.21})$$

En esta expresión, el término de la izquierda se refiere a la desigualdad entre todos los individuos de la población, que puede deducirse de la fórmula del índice teniendo en cuenta que cada individuo sería una clase que representa una fracción poblacional igual a $1/n$, cuando n es el tamaño de la población en cuestión.

Por otro lado, el I_T dentro de grupos puede calcularse como el promedio ponderado de los I_T de cada uno de los k grupos. Normalmente, se calculan el I_T total y el I_T entre grupos aplicando la expresión del índice, y el I_T dentro de grupos se obtiene por diferencia (13-14).²⁶

4. *La equivalencia en entropía:* (10) Un sistema de k clases socioeconómicas con un nivel de desigualdad dado por Z (que representa la versión transformada a escala [0,1] de cualquiera de los índices anteriores) puede homologarse con un sistema de 2 clases en el que una de las clases o estratos, que representa el $100 \cdot p\%$ de la población, padece el $100 \cdot (1-p)\%$ de la carga de enfermedad, en tanto que la otra, que representa el $100 \cdot (1-p)\%$ sufre el $100 \cdot p\%$ de la carga, donde:

$$p = \left(\frac{1}{\pi} \right) \left[\text{sen}^{-1} (1-Z)^{(0.06 \cdot Z + 0.6)} \right] \quad (\text{III.22})$$

Esta función es aproximadamente la inversa de:

$$Z = 1 - \left(\frac{1}{p-1} \right)^{(2p-1)} \quad (\text{III.23})$$

Por ejemplo, si $Z_{\text{THEIL}}=0,75$, entonces $p=0,13$ (Tabla 13), que corresponde a una partición en dos clases en la que el 13% de la sociedad acumula el 87% de la carga de enfermedad o muerte (o el 87% de los recursos en salud), mientras que el 87% restante acumula el 13% de la carga de enfermedad o muerte (o el 13% de los recursos en salud). Sin embargo, si $Z_{\text{THEIL}}=0,00$, entonces $p=0,50$ y ello equivale a la completa igualdad de una sociedad de dos clases de igual tamaño en la que cada una sufre el 50% de la carga de enfermedad o muerte o posee el 50% de los recursos. La tabla siguiente muestra varios ejemplos de esta relación de equivalencia en el sentido de la entropía.

²⁶ La descomponibilidad del Theil es aplicable en cualquier nivel de una agrupación jerárquica. Por ejemplo, si las unidades de análisis son municipios, y éstos están agrupados en provincias, es posible calcular la desigualdad total como la suma de la desigualdad entre provincias + la desigualdad dentro de provincias. Esta última sería una desigualdad entre municipios.

Tabla III.7 Equivalencia en entropía y potencial de redistribución para eliminar la desigualdad.

Índice Z	Transformación p	Equivalencia en entropía		Potencial de redistribución
0	0,50	50	50	0%
		50	50	
0,14	0,36	64	36	14%
		36	64	
0,29	0,30	70	30	20%
		30	70	
0,44	0,24	76	24	26%
		24	76	
0,50	0,22	78	22	28%
		22	78	
0,56	0,20	80	20	30%
		20	80	
0,75	0,13	87	13	37%
		13	87	

A medida que aumenta el valor de Z, la distribución se hace más desigual y su potencial de redistribución se incrementa. Un sistema de clases socioeconómicas con una desigualdad de $Z=0,56$ (Tabla 13) necesita una redistribución de 30% para eliminar la desigualdad, porque el nivel de desigualdad existente corresponde al de una distribución en que 20% de la población soporta 80% de la carga de enfermedad o muerte (o posee 80% de los recursos), mientras que el 80% restante sólo experimenta 20% de la carga de enfermedad o muerte (o posee 20% de los recursos).

V. Índices basados en la noción de desproporcionalidad (*)

Formalmente estos índices se asemejan a los que se basan en la noción de entropía. La diferencia esencial es que la comparación ahora involucra dos distribuciones de probabilidad acumulativa y no simplemente a las funciones de probabilidad, como en el caso de la entropía. Ello supone que la variable con respecto a la cual se realiza la clasificación se mide al menos en una escala ordinal lo cual permite la construcción de una jerarquía socioeconómica.

La desproporcionalidad a que alude el título de esta categoría de índices se refiere a dos distribuciones acumuladas: una de sujetos en una población, la otra, de un indicador de interés. Los sujetos de la primera distribución han sido previamente ordenados. El índice particular de desproporcionalidad depende del criterio de ordenamiento y del indicador de interés.

El índice de Gini es, casi con seguridad, el más usado de los indicadores de desigualdad. Un recuento muy detallado sobre la evolución de la teoría sobre el índice y sus aplicaciones puede encontrarse en el libro de Silber (15).

El índice de Gini, se basa en la curva de Lorenz, que representa la relación entre la población ordenada y acumulada, y un indicador socioeconómico acumulado (casi siempre

ingreso o gastos). El índice resulta de comparar la curva de Lorenz con la bisectriz del ángulo recto del primer cuadrante, que corresponde a la plena igualdad de ambas distribuciones acumuladas, y es justamente dos veces el área comprendida entre las dos curvas.

La figura 1, tomada de Schneider y colaboradores (3), ilustra gráficamente la definición anterior y la racionalidad de su cálculo.

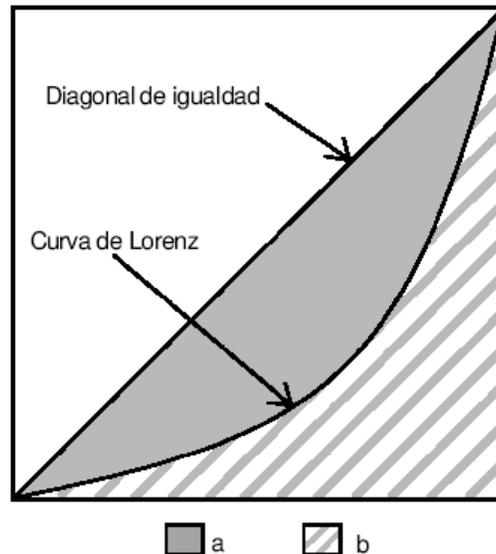


Figura 1. Curva de Lorenz y diagonal de igualdad.

Si los individuos de una población se ordenan en forma creciente con respecto a su ingreso, los más pobres se ubicarán cerca del origen de coordenadas y los más ricos hacia el extremo derecho del eje de las abscisas. Si representamos en este eje la proporción de individuos que resulta de acumular el ordenamiento, y en el eje de las ordenadas la proporción también acumulada de su ingreso, la curva de Lorenz de la figura mostraría claramente que una proporción alta de sujetos sólo posee una baja proporción de los ingresos, siempre menor que la que corresponde a la situación de plena igualdad, o de plena proporcionalidad entre población e ingreso. La diferencia o desproporcionalidad será mayor cuanto mayor sea el área (a) comprendida entre las dos curvas, sombreada en gris. Debido a la conveniencia de un re-escalamiento del índice en el intervalo 0-1, el área se expresa como proporción con respecto al área (b) por debajo de la diagonal. Esto equivale a dividir entre $\frac{1}{2}$ o, lo que es lo mismo, a multiplicar por 2. El valor de 0 corresponde al caso de plena igualdad, cuando la curva y la bisectriz se confunden, y el valor de 1, al caso extremo en que un solo individuo concentra todo el ingreso.²⁷ A partir de la Figura 1 el índice de Gini puede calcularse como:

²⁷ Debe notarse que el área del cuadrado es 1 y que, por tanto, el área del triángulo bajo la diagonal es 0,5.

$$I_G = \frac{a}{a+b} = 2a$$

puesto que $a + b = \frac{1}{2}$

Debido a su definición, el índice de Gini es siempre positivo y la curva se ubica siempre debajo de la diagonal. La lógica de la curva, y del índice y su escala de valores, no cambian si en lugar de ordenar sujetos se ordenan grupos y se representan gráficamente la población acumulada (siempre en el eje de las abscisas) y la proporción acumulada de ingreso (en el eje de las ordenadas).

La práctica habitual en el ámbito económico para medir las desigualdades sociales, ha consistido en ordenar de acuerdo al ingreso. No obstante, el índice de Gini se ha expuesto sólo por razones argumentativas, porque no es una métrica de desigualdades sociales en salud, y ni siquiera de desigualdades en salud.

Sin embargo, la racionalidad del índice hizo concebir su extensión al ámbito de la salud (16, 17). Si en lugar de ordenar la población según ingreso (del más pobre al más rico) se ordena según una variable de salud (del más enfermo al más saludable) se obtiene una nueva curva de Lorenz, pero ahora con la particularidad de que, si en efecto el indicador es de naturaleza negativa (como la mortalidad o la morbilidad) la curva se ubicará por encima de la diagonal, mientras que si se trata de un indicador positivo (cobertura de inmunizaciones o acceso a los servicios de salud), se ubicará por debajo. Por convención, al valor del índice de Gini, que se define del mismo modo, se le asigna un signo positivo si la curva de Lorenz se localiza por debajo de la diagonal y un signo negativo, si la curva se localiza por encima. De este modo el rango de variación se extiende desde -1 hasta +1, aunque los valores próximos a 0 siguen siendo expresión de poca desigualdad. A la curva que se obtiene al ordenar de acuerdo a la variable de salud y al índice resultante, se les suele llamar "pseudo-curva de Lorenz" e "índice de pseudo-Lorenz", respectivamente.

Este índice mide desigualdades en salud, pero no desigualdades sociales, porque ignora el sustrato socioeconómico de la desigualdad. (18 - 20)

La lógica interna del índice de Gini puede conservarse sin renunciar a la medición de las desigualdades sociales en salud, si el ordenamiento de la población se realiza de acuerdo a una variable socioeconómica, y luego se mide la proporcionalidad de la distribución acumulada que genera dicho ordenamiento con la de un indicador de salud. Los sujetos o los grupos se ordenan del más pobre al más rico (o en general, del peor al mejor de acuerdo al indicador socioeconómico elegido) y las proporciones poblacionales acumuladas se representan en el eje de las abscisas. En el de las ordenadas, se representa la proporción acumulada del indicador de salud, de acuerdo a cuya identidad, la curva así obtenida (ahora llamada curva de concentración) se encontrará encima de la diagonal si el indicador es negativo, o debajo de ella si el indicador positivo. Si a partir de esta representación gráfica se mide el área comprendida entre la curva de concentración y la diagonal, y se multiplica por 2, se obtiene el llamado índice de concentración, que toma valores entre -1 y +1. Normalmente la curva de concentración se

encuentra de un solo lado de la diagonal (encima o debajo de acuerdo a la naturaleza del indicador de salud). Ello ocurre cuando existe una relación monótona (creciente o decreciente) entre el indicador de salud y el indicador socioeconómico que genera la jerarquía social ordenada. Si ello no ocurre, la curva puede cruzar la diagonal, como corresponde al caso en que la mortalidad o morbilidad son más frecuentes en lo más bajo y en lo más alto de la escala social, y menor en los grupos intermedios.

Por sus propiedades, el índice de concentración ha desplazado al de Gini como medida de las desigualdades sociales en salud y en él se concentrará la atención en este texto.

Formalmente el índice de concentración (IC) se define del siguiente modo:

$$3.14 \quad I_C = 1 - 2 \int_0^1 L(r) dr \quad (\text{III.24})$$

IC toma el valor 0 cuando $L(r) = r$, es negativo cuando $L(r) > r$, y positivo cuando $L(r) < r$ para todo r en el intervalo (0 1).

A nivel individual, IC se puede calcular fácilmente: si se ordenan los sujetos en forma creciente según una variable socioeconómica, IC viene dado por: (21-22)

$$I_C = \frac{2}{nS} \sum_{i=1}^n s_i r_i - 1 \quad (\text{III.25})$$

En la expresión anterior, S es el valor promedio de la variable de salud en la población, n el número total de sujetos, s_i , el valor de la variable de salud en el sujeto i -ésimo y r_i el rango relativo²⁸ que corresponde al sujeto i -ésimo en el ordenamiento.

En datos agrupados según clases socioeconómicas, el índice de concentración toma la forma siguiente: (23)

$$I_C = \frac{2}{t_{total}} \sum_{i=1}^K p_i t_i r_i - 1 \quad (\text{III.26})$$

en donde

K es el número de clases

t_{total} es la tasa global de la población

t_i es la tasa de la clase i -ésima

r_i es el rídit (o rango promedio) de la clase i -ésima

²⁸ El rango relativo o rídit es la posición en el arreglo ordenado comenzando en 0 dividida entre el número de casos más 1 ($n+1$)

p_i es el tamaño relativo de clase i -ésima

El I_C y el IDP ya expuesto pueden obtenerse uno a partir del otro mediante la relación:

$$I_C = \frac{2IDP \text{var}(r)}{t_{total}} = 2IRD_{PAMUK} \text{var}(r) \quad (\text{III.27}) \quad (\text{Anexo 1})$$

en donde $\text{var}(r)$ es la varianza de los ridits y viene dada por la expresión:

$$\text{var}(r) = \sum_{i=1}^K p_i \left(r_i - \frac{1}{2} \right)^2 \quad (\text{III.28})$$

El I_C posee tres atributos fundamentales. (3, 19)

1. Refleja el componente socioeconómico de las desigualdades en salud.
2. Hace uso de la información contenida en toda la jerarquía social.
3. Es sensible a los cambios, tanto en la distribución del indicador de salud, como en la composición de la población en los grupos que integran la escala socioeconómica.

El I_C es, además invariante por cambios de escala, es decir que si el valor del indicador de salud en los grupos se multiplica por una constante arbitraria k ($k \neq 0$) no se modifica el valor del índice.

Una variante del I_C es el llamado índice de concentración generalizado (21-22) que se obtiene como el producto del I_C por el valor global del indicador de salud en la población (t_{total}).

A pesar de las excelentes propiedades métricas del I_C -e incluso de los elogios de la literatura- (21) la extrapolación del terreno económico al de la salud no está exenta de dificultades que complican, o incluso hacen paradójica la interpretación de los resultados. (24) Una de ellas -ya expuesta en relación con los índices basados en rangos- es la asimetría intrínseca de los índices a la disyuntiva de considerar la variable de salud en sentido negativo (por ejemplo, presencia de la enfermedad) o en sentido positivo (ausencia de ella). La otra es que la noción de concentración puede distorsionarse cuando en lugar del ingreso, se considera mecánicamente la carga de enfermedad.

Referencias:

1. Harper S, Lynch J. Methods for measuring cancer disparities: using data relevant to Healthy People 2010 Cancer Related Objectives. Technical Report. Center for Social Epidemiology and Population Health. University of Michigan. 2010.
2. Schoenbach VJ. Comprendiendo los fundamentos de la Epidemiología. (Traducción al español). Chapel Hill, North Carolina. 2000.

3. Schneider C, Castillo-Salgado C, Bacallao J, Loyola E, Mujica OJ, Vidaurre M, Roca A. Métodos de medición de las desigualdades de salud. *Rev Panam Salud Pública* 2002; 12: 398-415.
4. Keppel K, Pamuk E, Lynch J et al. Methodological issues in measuring health disparities. National Center for Health Statistics. *Vital Health Stat* 2005; 2: 141-147.
5. Mackenbach JP, Kunst AE. Measuring the magnitude of socioeconomic inequalities in health: an overview of available measures illustrated with two examples from Europe. *Soc Sci Med* 1997; 44: 757-771.
6. Pearcy JN, Keppel KG. A summary measure of health disparity. *Public Health Rep* 2002; 117: 273-280.
7. Chakravarti SR. The variance as a subgroup decomposable measure of inequality. *Soc Indic Res* 2001; 53: 79-95.
8. Wagstaff A. Inequality aversion, health inequality and health achievement. The World Bank Development Research Group Public Services and Human Development Network Health, Nutrition and Population Team. Policy Research Working Paper 2765. 2002.
9. García L. Introducción a la termodinámica clásica. México DF. Editorial Trillas. 1986.
10. Kluge G. Entropywise equivalent societies. Disponible en: <http://ourworld.compuserve.com/homepages/SMIPP/entequiv.htm>.
11. Conceição P, Ferreira P. The young person's guide to the Theil index: suggesting intuitive interpretations and exploring analytical applications. University of Texas International Press. Working Paper No. 14. 2000).
12. Bacallao J, Castillo-Salgado C, Schneider C, Mujica OJ, Loyola E, Vidaurre M. Indices para medir las desigualdades de salud de carácter social, basados en la noción de entropía. *Rev Panam Salud Pública* 2002; 12: 429-435.
13. Conceição P, Galbraith JK, Bradford P. The Theil index in sequences of nested and hierarchic group structures: implications for the measurement of inequality through time, with data aggregated at different levels of industrial classification. *Eastern Economic Journal* 2000; 27: 61-74.
14. Milanovic B. Measuring international and global inequality. Princeton and Oxford. Princeton University Press 2005; 20-27.
15. Silber J. Handbook of income inequality measurement. Boston. Kluwer Academic. 1999.
16. Leclerc A, Lert F, Fabien C. Differential mortality: some comparisons between England and Wales, Finland and France based on inequality measures. *Int J Epidemiol* 1990; 19: 1001-1010.
17. Le Grand J. Inequalities in health: the human capital approach. Welfare State Program Discussion Paper No. 1. London School of Economics 1985.
18. Houweling TAJ, Kunst AE, Mackenbach JP. World Health Report 2000: Inequality index and socioeconomic inequalities in mortality. *Lancet* 2001; 357: 1671-1672.
19. Wagstaff A, Paci P, van Doorslaer E. On the measurement of inequalities in health. *Soc Sci Med* 1991; 33: 545 - 557.

20. Mackenbach JP, Kunst AE. Measuring the magnitude of socioeconomic inequalities in health: an overview of available measures illustrated with two examples from Europe. *Soc Sci Med* 1997; 44: 757-771.
21. Kakwani N, Wagstaff A, van Doorslaer E. Socioeconomic inequalities in health. Measurement, computation and statistical inference. *Journal of Econometrics* 1997; 77: 87 - 103.
22. Erreygers G. Correcting the concentration index. *J Health Econ* 2008; doi:10.1016/j.jhealeco.2008.02003.
23. Van Doorslaer E, Wagstaff A, Bleichrodt H et al. Income-related inequalities in health: some international comparisons. *J Health Econom* 1997; 16: 93-112.
24. Koolman X, van Doorslaer E. On the interpretation of the concentration index of inequality. *Health Economics* 2004; 13: 649-656.

Capítulo IV

Novedades metodológicas en la medición de las desigualdades sociales en salud

El triunfo supremo de la razón consiste en arrojar dudas sobre su propia validez.

Miguel de Unamuno

Resumen:

En este capítulo se introducen varias novedades metodológicas en los índices (marcados con asterisco) del capítulo anterior. Estas novedades no resuelven completamente las insuficiencias de los índices, pero mejoran sus propiedades mediante recursos simples, que -sorprendentemente- no aparecen en la literatura revisada ni se emplean en la práctica. Se proponen también nuevos índices junto a su lógica subyacente, a partir de una heurística que se describe en el texto. El primero ("índice de concentración ajustado") aporta criterios simples para una medición de la desigualdad social, depurada de factores de sesgo por confusión; el segundo, un "índice integrado" basado en la noción de entropía con mejores propiedades que los que se emplean de rutina en la literatura. Por último, se expone un enfoque alternativo para identificar y medir desigualdades sociales, especialmente aplicable al caso en que éstas se ocultan tras dimensiones latentes o no explícitas. Al final del capítulo se incluyen ejemplos que ilustran la aplicación de los índices y procedimientos en general.

4.1 *El cociente ponderado de tasas*

$CPT = f * CT$ en donde

$$f = \frac{n_1}{n_2} \quad (IV.1)$$

es el cociente de tasas multiplicado por una ponderación igual al cociente de los tamaños respectivos de las clases. En la expresión

$$CT = \frac{t_1}{t_2}$$

suponemos que $t_1 \geq t_2$. $CPT = CT$ si y sólo si $n_1 = n_2$. A diferencia de CT , en relación con CPT se puede conseguir disminuir la desigualdad, no sólo mediante una disminución en el

cociente de las tasas, sino también mediante una redistribución más favorable de las personas expuestas a las condiciones de las clases.

Bajo la misma lógica puede definirse

4.2 La diferencia ponderada de tasas

$$DPT = f * (t_1 - t_2)$$

en donde (IV.2)

$$f = \frac{n_1}{n_2}$$

DPT puede disminuir también no sólo por una reducción en DT, sino también por una redistribución poblacional, o por supuesto, por una combinación favorable del producto de ambos factores.

4.3 Condiciones de cambio en las tasas y su efecto sobre las desigualdades según CT y DT.

Supongamos un dominio A cualquiera en el que X e Y son unidades o grupos (por ejemplo, percentiles) con valores extremos de las tasas x e y, respectivamente, y tales que $x > y$.

i. Supongamos, además, que las tasas experimentan una disminución de magnitud k fija ($k > 0$ y, por supuesto, $k < y < x$)²⁹. Entonces:

a. La diferencia entre las tasas ($x - k$ e $y - k$) se mantiene constante puesto que:

$$(x - k) - (y - k) = x - y \quad (IV.3)$$

b. El cociente entre las tasas, sin embargo, aumenta³⁰. En efecto:

$$\frac{x - k}{y - k} > \frac{x}{y} \Leftrightarrow (x - k)y > x(y - k) \Leftrightarrow xy - yk > xy - xk \Leftrightarrow x > y \quad (IV.4)$$

Por lo tanto, cualquier intervención que no priorice al grupo con peores condiciones, tiende a incrementar la desigualdad relativa.

²⁹ La condición $k < y < x$ es necesaria para que las tasas finales (después del cambio) $x - k$ e $y - k$ se mantengan positivas.

³⁰ Este resultado es extremadamente importante, porque demuestra que cualquier intervención *no discriminatoria* (*no focalizada*) que disminuya la tasa global en la población, incrementa la desigualdad relativa.

ii. Supongamos que se produce una intervención focalizada que produce en X una disminución absoluta de magnitud k_x y en Y una de magnitud k_y ($k_x > k_y$)³¹. Entonces:

a. La diferencia entre las tasas disminuye, ya que:

$$(x - k_x) - (y - k_y) = (x - y) - (k_x - k_y) < (x - y) \quad (\text{IV.5})$$

puesto que $k_x - k_y > 0$

b. El cociente entre las tasas $\frac{x - k_x}{y - k_y}$ disminuye si y sólo si $\frac{k_x}{k_y} > \frac{x}{y}$. En efecto:

$$\begin{aligned} \frac{x - k_x}{y - k_y} < \frac{x}{y} &\Leftrightarrow (x - k_x)y < x(y - k_y) \Leftrightarrow xy - k_x y < xy - k_y x \\ &\Leftrightarrow k_x y > k_y x \Leftrightarrow \frac{x}{y} < \frac{k_x}{k_y} \end{aligned} \quad (\text{IV.6})$$

Este resultado implica que para que una reducción en la tasa global produzca disminución de la desigualdad relativa, el cociente de las reducciones en el grupo con la tasa más alta y el grupo con la tasa más baja, tiene que ser superior al cociente de sus tasas respectivas. Por consiguiente, la mejoría en la salud global de la población tiende a aumentar la desigualdad a menos que se realicen intervenciones discriminatorias o focalizadas, y que éstas produzcan una disminución suficientemente mayor en el grupo con la tasa más alta.

iii. Supongamos, por último, que se realiza una intervención que produce una disminución relativa de k_x en el grupo con la tasa más alta y de k_y en el grupo con la tasa más baja, y que $k_x \geq k_y$. Las nuevas tasas serán, respectivamente, $x - xk_x$, y $y - yk_y$. Entonces:

a. La diferencia entre las tasas se mantiene constante sólo si $xk_x = yk_y$ ³², y disminuye sólo si $xk_x \geq yk_y$.

b. El cociente entre las tasas, por su parte se mantiene constante si $k_x = k_y$ y disminuye sólo si $k_x > k_y$.³³

³¹ O sea, la intervención se "focaliza" en el grupo con peor situación para producir en él la mayor disminución de la tasa de enfermedad.

³² Esta condición es equivalente a que la reducción absoluta en las tasas sea idéntica en ambos grupos.

³³ La demostración es inmediata como en los casos i y ii.

El interés de las condiciones i-iii radica en que proporcionan una guía conceptual, con un fundamento algebraico simple, para interpretar adecuadamente el significado de las dos medidas de desigualdad (diferencia y cociente) y para no confundirse con la paradoja aparente que puede generar su uso indistinto. En particular, se quiere llamar la atención sobre el hecho de que ciertas intervenciones exitosas (en el sentido de que disminuyen la tasa global y la tasa de los grupos) y focalizadas a favor de los más necesitados, pueden incrementar la desigualdad, ya sea en sentido absoluto o relativo.

4.4 El índice de Pearcy-Keppel ponderado y la varianza entre grupos relativa

La expresión

$$I_{PK}^* = \frac{\sum_{i=1}^N W_i |t_i - t_{ref}|}{t_{ref}} \quad (IV.7)$$

es una versión ponderada del I_{PK} que toma en cuenta el tamaño relativo de los grupos y con ello permite superar una de sus insuficiencias.

También

$$VEG^* = \frac{\sum_{i=1}^N W_i (t_i - t_{total})^2}{t_{total}} \quad (IV.8)$$

proporciona una versión relativa de la varianza entre grupos, que permitiría realizar comparaciones en el tiempo cuando cambia el valor total del indicador de salud. En ambos casos las W_i se definen como en (III.13).

4.5 Los índices relativos de desigualdad corregidos

A partir de la ecuación $t_i = \alpha + \beta r_i$, la figura que sigue ilustra la racionalidad que subyace al IDP y al IRD_{KM} (ver 3.6 y 3.8). El IDP es la diferencia entre los puntos extremos de la escala $t_i | r_i = 0 - t_i | r_i = 1 = -\beta$, mientras que el IRD_{KM} es el cociente entre dichos puntos, es decir,

$$\frac{\alpha}{\alpha + \beta} \quad (IV.9)$$

Sin embargo, como ilustra la figura que sigue, estos son sólo extremos teóricos, inalcanzables en la práctica, que dan lugar a valores sobrestimados de la desigualdad.

Índice de desigualdad de la pendiente

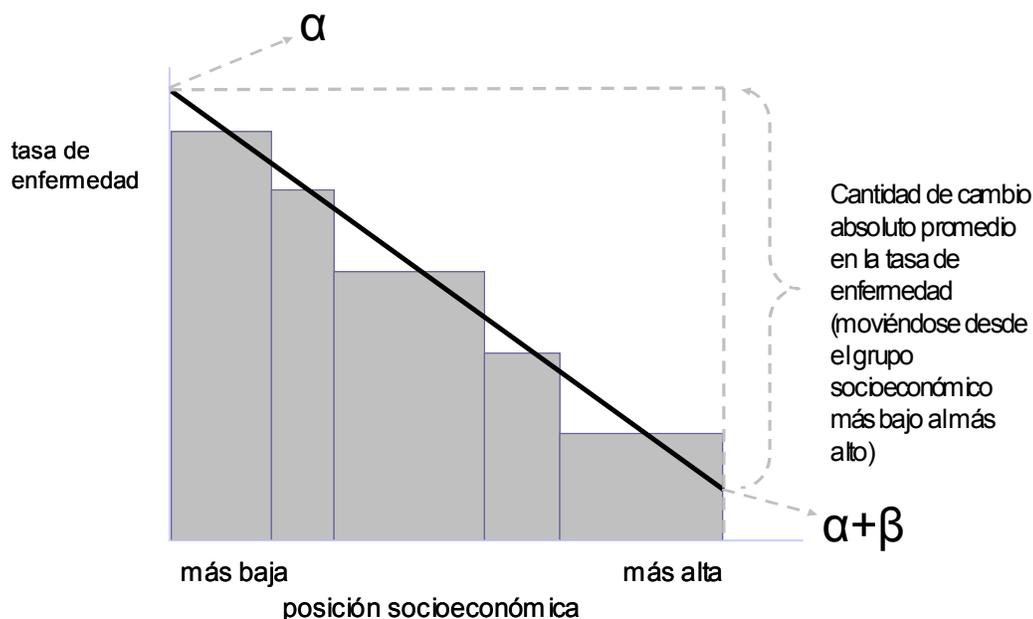


Figura IV.1 Lógica geométrica del índice de desigualdad de la pendiente

Mucho más realista sería definir los índices de desigualdad absoluto y relativo a partir del valor de las tasas estimadas para los rangos de los grupos extremos en la escala socioeconómica. Esto daría lugar al "índice de desigualdad de la pendiente corregido" ($IDP_{\text{corregido}}$) y al "índice relativo de desigualdad corregido" ($IRD_{\text{corregido}}$) que se definen respectivamente como:

$$IDP_{\text{corregido}} = t_{\text{peor}} - t_{\text{mejor}} \quad ; \quad IRD_{\text{corregido}} = \frac{t_{\text{peor}}}{t_{\text{mejor}}} \quad (IV.10)$$

4.6 El índice de concentración ajustado.

El caso que se examina en el presente apartado corresponde al cálculo de la concentración a nivel del individuo o la familia. La unidad de análisis se ordena de peor a mejor de acuerdo a una variable socioeconómica, que tradicionalmente (aunque no necesariamente) es el ingreso, y se construye la curva de concentración a partir de la posición relativa de dicha unidad de análisis y la magnitud acumulada de la variable de salud. Dado que habitualmente se trabaja con variables de salud con una connotación negativa (como los indicadores de mortalidad y morbilidad), éstas se concentran en los más pobres, el índice de concentración se ubica por encima de la diagonal principal y toma -convencionalmente- valores negativos dentro de un recorrido que va, teóricamente, desde -1 a +1. (Ver Capítulo III, apartado V).

El apartado 1.5 del Capítulo 1 terminaba con el siguiente pasaje: *"En este ensayo se sostiene el punto de vista de que, independientemente de la utilización de estos criterios habituales, la elección de la variable socioeconómica que genera la jerarquía o la clasificación, debe ser idealmente el resultado de hipótesis concretas sobre las causas que generan la desigualdad y debe orientarse según la perspectiva de un repertorio de acciones previsibles para reducirla o eliminarla."*

La lógica detrás de esta afirmación es simple: identificar la dimensión socioeconómica correlacionada con la salud, hacer visible la desigualdad social en salud derivada de dicha correlación y realizar las intervenciones pertinentes. Sin embargo, la correlación podría sólo reflejar una asociación espuria debida a efectos de confusión, y las intervenciones orientadas por dicha asociación, podrían ser por tanto, inefectivas o estériles.

Consideremos, por ejemplo, una población en la que el desempleo influye sobre la variable de salud, la que a su vez depende de la edad. Si el desempleo es más frecuente entre los más jóvenes, la edad podría confundir el efecto del estatus laboral sobre la salud.³⁴ De estas consideraciones se deriva la idea de ajustar el índice de concentración para posibles covariantes, depurándolo de efectos confusores que sesgan su medición. Como se verá oportunamente, la elección de las covariantes es clave, en la misma medida que en el contexto de los modelos lineales, cuando se desea ajustar el efecto de un factor cualquiera.

El caso particular que se examinará es aquel en que la unidad de análisis viene dada por n individuos o familias, ordenados de acuerdo a una variable socioeconómica y , que genera concentración de la variable de salud s , pero que se supone "confundida" por una variable z , que se asocia con ella y también con la variable de salud. El ajuste del efecto de y para el confusor z , se realiza mediante el modelo:

$$s_i = \alpha + \alpha_y y_i + \alpha_z z_i + \varepsilon_i \quad (\text{IV.11})$$

en donde

s : variable de salud

y : variable socioeconómica que define el gradiente de salud de interés

z : confusor de la asociación entre y y s

³⁴ Otro ejemplo interesante es el de las poblaciones que emigran a EE UU y Europa en muchos países latinoamericanos; los hijos de los emigrantes suelen presentar tasas más altas de enfermedades infecciosas y más bajo rendimiento escolar. Las desigualdades entre las personas que emigran y las que no lo hacen en relación con la salud y la educación podrían, sin embargo, estar "confundidas" por la condición socioeconómica, que es, al mismo tiempo, una causa poderosa de emigración, y un factor que influye en la salud y en el bajo rendimiento escolar.

$\alpha, \alpha_y, \alpha_z$ parámetros del modelo
 ε error

Para datos no agrupados (ver III.25), el índice de concentración puede calcularse según la expresión

$$I_C = \frac{2}{nS} \sum_{i=1}^n s_i r_i - 1 \quad (\text{IV.12})$$

A partir de (IV.11) y (IV.12) se puede obtener la siguiente descomposición del índice de concentración (ver Anexo 1):

$$I_C = IC_y \frac{\alpha_y \mu_y}{S} + IC_z \frac{\alpha_z \mu_z}{S} + \frac{IC_\varepsilon}{S} \quad (\text{IV.13})$$

en donde

μ_y, μ_z : medias de las variables y, z respectivamente

S : media de la variable de salud en la población

IC_y, IC_z : índices de concentración de las variables y, z , respectivamente.

IC_ε : índice de concentración generalizado³⁵ para ε

Si se supone además que todo el potencial confusor para la asociación entre y y \underline{s} , está contenido en \underline{z} , resulta que ε está incorrelacionado con cualquiera de los factores que afectan la salud y por consiguiente:

$$C_\varepsilon = 0 \quad ^{36} \quad (\text{IV.14})$$

Eso reduciría (IV.13) a

$$I_C = IC_y \frac{\alpha_y \mu_y}{S} + IC_z \frac{\alpha_z \mu_z}{S} \quad (\text{IV.15})$$

que propone una descomposición del índice de concentración como una suma ponderada de las concentraciones de y y el confusor contenidos en (IV.11). El primer sumando de la derecha en (IV.15) representa el índice de concentración ajustado.

³⁵ El índice de concentración generalizado es el índice de concentración multiplicado por la media poblacional, y equivale, como ya se expuso, al índice de desigualdad de la pendiente

³⁶ Bajo este supuesto, ε contiene componentes de desigualdad en salud, pero carecen de interés porque se suponen incorrelacionadas con la dimensión socioeconómica.

Si y es ingreso, entonces IC_y es el índice de Gini, y por tanto la concentración de la salud con respecto al ingreso, se expresa como la suma ponderada de la concentración del ingreso y de la concentración debida a la variable z .

La expresión (IV.15) es útil para explicar los cambios en el tiempo, que es la aplicación de mayor interés en el estudio de las desigualdades.

El cambio en el tiempo de la concentración total vendría dada por:

$$\delta I_C = \left(IC_{y(t)} \frac{\alpha_{y(t)} \mu_{y(t)}}{S_{(t)}} - IC_{y(t-1)} \frac{\alpha_{y(t-1)} \mu_{y(t-1)}}{S_{(t-1)}} \right) + \left(IC_{z(t)} \frac{\alpha_{z(t)} \mu_{y(t)}}{S_{(t)}} - IC_{z(t-1)} \frac{\alpha_{z(t-1)} \mu_{z(t-1)}}{S_{(t-1)}} \right) \quad (IV.16)$$

La expresión (IV.16) muestra que los cambios en la concentración pueden deberse a: (a) cambios en la salud total de la población [$S(t) \neq S(t-1)$]; (b) cambios en la concentración de la salud con respecto a y [$IC_{y(t)} \neq IC_{y(t-1)}$]; (c) cambios en los valores medios de z [$\mu_{z(t)} \neq \mu_{z(t-1)}$] o en los valores medios del indicador socioeconómico y [$\mu_{y(t)} \neq \mu_{y(t-1)}$] y (d) cambios en el efecto de y [$\alpha_{y(t)} \neq \alpha_{y(t-1)}$] o de los confusores, sobre S [$\alpha_{z(t)} \neq \alpha_{z(t-1)}$].

Aunque en la práctica sería extremadamente difícil explicar el cambio en la desigualdad, particularmente cuando hay cambios en las tasas globales de salud de la población, o en los parámetros del modelo (IV.11) el interés de (IV.16) radica en que se podrían suponer invariantes algunos de los componentes mencionados en el párrafo anterior y verificar la influencia de los otros.

De gran importancia, tanto teórica como práctica es la selección de la variable z . Esta selección descansa, en parte, en los mismos principios que se aplican en el contexto de los diagramas causales para elegir los factores que deben controlarse antes de estimar el efecto de las presuntas causas sobre la variable de respuesta.

Si una variable Z se encuentra en la trayectoria causal de y sobre z , no debe ser incluida en (IV.11) y por consiguiente tampoco en (IV.13) ni en (IV.15), porque parte del efecto de y sobre z , está mediado por Z y la estimación ajustada de la concentración estaría sesgada. Por ejemplo, al calcular la concentración de la salud para el ingreso, no sería aconsejable ajustar para el nivel educacional que se encuentra probablemente en la vía causal entre ingreso y salud, debido a que la escolaridad tiene una distribución desigual en relación con el ingreso, y además la baja escolaridad es causa de mala salud.

Si por el contrario, la variable Z puede "confundir" la asociación entre y y z , entonces sí debe controlarse, porque de lo contrario la estimación de la concentración de z sobre y estaría sesgada.

El otro criterio para elegir una variable de control en la medición de la desigualdad por concentración, es que dicha variable sea o no modificable mediante intervenciones. Si una variable mediadora no es modificable, debe controlarse porque de ese modo puede descomponerse la concentración en su componente directo (no mediado) y su componente indirecto (mediado por la variable controlada). Por ejemplo, si los inmigrantes tienen una edad promedio superior a los nativos y la edad influye en el riesgo de enfermar, la concentración de la enfermedad en los inmigrantes podría deberse en mayor medida a su edad que a su condición migratoria. Si no se controla para la edad, la concentración sobrestimaría la desigualdad evitable mediante intervenciones o programas selectivos a favor de los inmigrantes.³⁷

El ejemplo que sigue, basado en una población ficticia de 100 casos, ilustra la descomposición de la concentración que se presenta en este apartado. El primer cuadro muestra las correlaciones de una variable de salud s con una variable y que define una jerarquía socioeconómica, y una variable z , que probablemente confunde la asociación entre ambas. Se ve claramente que z tiene altas correlaciones, tanto con y como con s . La correlación entre y y s , ajustada para z , disminuye considerablemente (de -0,66 a -0,20). Ambos hechos combinados son sugerentes de confusión o de mediación parcial.

Tabla IV.1 Correlaciones (y valores de "p") entre una variable de salud (s), un indicador socioeconómico (y) y una covariante (z)

		z	y	s
z	Correlación	1	,905**	-,659**
	Valor de "p"		,000	,000
	N	100	100	100
y	Correlación	,905**	1	-,661**
	Valor de "p"	,000		,000
	N	100	100	100
s	Correlación	-,659**	-,661**	1
	Valor de "p"	,000	,000	
	N	100	100	100

**Correlación significativa al nivel 0,01

Tabla IV.2 Correlación entre el indicador de salud (s) y el indicador socioeconómico (y) controlando para la covariante (z)

Variable de control		s	y	
z	s	Correlación	1,000	-,200
		Valor de "p"	.	,047
y		Correlación	-,200	1,000
		Valor de "p"	,047	.

³⁷ Este ejemplo y otros del apartado 4.6 tienen valor argumentativo. Son ejemplos que se basan en supuestos hipotéticos que podrían aplicarse en algunos contextos y tal vez no en otros.

La tabla que sigue, contiene los resultados del ajuste del modelo (IV.11), y en particular las estimaciones de los parámetros α_y y α_z .

Tabla IV.3 Estimación de los parámetros del modelo (IV.11)

Fuente	Valor	t	Pr > t	Intervalo de confianza (95%)	
Intercepto	1,881	4,707	< 0,0001	1,088	2,674
α_y	-0,001	-2,047	0,043	-0,002	0,000
α_z	-0,386	-1,910	0,059	-0,788	0,015

A partir de estas estimaciones y de las relaciones (IV.13) y (IV.15), se calculan el índice de concentración y sus componentes, así como el índice de concentración ajustado, que muestra la siguiente tabla.

Tabla IV.4 Descomposición de la concentración y cálculo del índice de concentración ajustado

Componentes de la concentración	Valor del índice	Contribución porcentual
Variable y	-0,054	51,4%
Variable z	-0,041	39,1%
Concentración residual	-0,010	9,5%
Total	-0,105	100,0%

Se presenta ahora una disyuntiva que es imposible resolver sin un conocimiento teórico sobre la naturaleza de las variables en cuestión, del cual, obviamente no se dispone en este ejemplo ficticio: ¿es conveniente o no ajustar la concentración para la variable \underline{z} ?

Todo depende de la naturaleza de \underline{z} , y del rol que desempeña en el modelo estructural que conecta a las tres variables: \underline{z} , \underline{y} y \underline{s} . Si \underline{z} es una variable mediadora modificable del efecto de \underline{y} sobre \underline{s} , su efecto no debe controlarse, y el índice de concentración sería -0,095, que se obtiene sustrayendo la concentración residual del índice de concentración total.

Si por el contrario, \underline{z} es una variable confusora, entonces parte del efecto de concentración socioeconómica de la salud se debe a \underline{z} , por lo cual se debe ajustar para \underline{z} , y la concentración debida a \underline{y} , quedaría reducida al índice de concentración ajustado (-0,054).

También debe ajustarse si \underline{z} ejerce un efecto mediador no modificable, aunque ahora por una razón diferente: si no se ajusta para \underline{z} , el margen de acción para reducir la desigualdad estaría indebidamente hipertrofiado.

4.7 Un índice integrado basado en la noción de entropía.

Un poco de álgebra a partir III.14 permite comprobar que la desviación logarítmica media (DLM) puede expresarse como sigue:

$$DLM = \sum_{i=1}^K \pi_i \ln \left(\frac{\pi_i}{p_i} \right) \quad (\text{IV.17})$$

En (IV.17) π_i designa la proporción del tamaño poblacional total contenido en la clase i -ésima y p_i designa la proporción de los casos (enfermos o fallecidos) incluidos en la clase i -ésima. Tanto π_i como p_i son distribuciones de probabilidad, como se describe en (III.15), de modo que I_T y DLM son dos maneras de calcular la divergencia (o entropía relativa) entre las dos distribuciones.

Si se compara la expresión (IV.17) con la del índice de Theil (III.18)

$$I_T = \sum_{i=1}^K p_i \ln \left(\frac{p_i}{\pi_i} \right)$$

se constata que los dos índices sólo difieren en la ponderación que asignan a la divergencia en escala logarítmica entre $\{p_i\}$ y $\{\pi_i\}$. I_T puede interpretarse como un estadístico de la razón de verosimilitudes para medir la desviación de la distribución empírica π_i con respecto a la distribución teórica p_i . Por tal motivo, I_T está muy influenciado por las altas proporciones relativas de la enfermedad. Del mismo modo, DLM puede entenderse como un estadístico de la razón de verosimilitudes para medir la desviación de la distribución empírica p_i con respecto a la distribución teórica π_i . Es fácil ver que I_T y DLM cumplen sólo uno de los tres axiomas de distancia, a saber: $d(P, \Pi) = 0 \Leftrightarrow P \equiv \Pi$. Tanto I_T como DLM son asimétricos ($d(P, \Pi) \neq d(\Pi, P)$).

El promedio de I_T y DLM cumple dos de los axiomas de distancia, y al igual que sus componentes puede demostrarse que tiene la propiedad de descomponibilidad (ver Anexo 1). Denominaremos "índice de entropía integrado" a este nuevo índice que puede expresarse del siguiente modo:

$$I_{\text{int}} = \frac{I_T + DLM}{2} = \sum_{i=1}^K (p_i - \pi_i) * \ln \left(\frac{p_i}{\pi_i} \right) \quad (\text{IV.18})$$

A partir de (IV.18) (ver Anexo 1) es posible llegar a la expresión

$$I_{\text{int}} = \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) * \ln \left(\frac{s_i}{n_i} \right) \quad (\text{IV.19})$$

en la que s_i y n_i son, respectivamente, el número de casos y el tamaño de la clase i -ésima. Pero $s_i/n_i = t_i$ es la tasa (base 1) de la clase i -ésima, por lo que I_{int} tiene una interesante interpretación: es la suma ponderada de las diferencias de las dos distribuciones empíricas. El factor de ponderación son las tasas log-transformadas. I_{int} asigna una importancia mayor a las diferencias que corresponden a clases con tasas altas, de modo que el exceso de casos con respecto a lo esperado contribuye más a la desigualdad cuanto más altas son las tasas a las que corresponde dicho exceso. Esta es una propiedad importante, y es congruente con el principio de aversión a la desigualdad de Wagstaff. (1)

4.8 La medición de la desigualdad según dimensiones latentes

Casi toda la literatura sobre desigualdades sociales en salud se relaciona con indicadores socioeconómicos prácticamente normativos, como el ingreso, la educación, el empleo, el sexo, los grupos raciales, o la condición migratoria, por sólo citar un grupo no exhaustivo, pero bien representativo de estos indicadores. (2,3) Con frecuencia se usan también indicadores sintéticos como el índice de desarrollo humano o el porcentaje poblacional con necesidades básicas insatisfechas.

Estas elecciones no son arbitrarias. En primer lugar, responden a un consenso bastante general -aunque no necesariamente unánime- de que las desigualdades sanitarias generadas por cualquiera de los indicadores mencionados suelen ser sistemáticas, injustas y evitables, lo que las convierte en inequidades y por lo tanto, en objetos relevantes para intervenciones programáticas. En segundo lugar, porque debido a la práctica derivada del mencionado consenso, se pueden realizar comparaciones entre países o entre regiones dentro de un mismo país. (4)

No obstante, con frecuencia no describen una relación de causalidad útil para orientar intervenciones exitosas, porque su efecto está "confundido" o mediado por otros factores (este aspecto de la cuestión ya se trató en un apartado anterior) y casi siempre enmascaran gran cantidad de fuentes de desigualdad "dentro" de las clases, por lo que subestiman la magnitud de la desigualdad social en salud.

Por otra parte, y por diferentes razones, las variables socioeconómicas clásicas podrían no dar lugar a gradientes de salud, y no ser, por consiguiente, fuentes de desigualdades sociales, lo cual no implica que éstas no existan, sino que están enmascaradas en dimensiones latentes -no directamente observables- que es importante identificar y caracterizar.

La propuesta que se desarrolla en el presente apartado se basa en la premisa de que existen subgrupos poblacionales singulares en términos de salud. Estos grupos no son explícitos, sino que deben buscarse inductivamente a partir de indicadores socioeconómicos relevantes, presumiblemente asociados con la condición de salud, para

luego caracterizarlos y validar su pertinencia como fuentes de desigualdades sanitarias.³⁸

Hay al menos tres opciones analíticas para confirmar formalmente esta conjetura. La primera es el análisis de conglomerados clásico ("cluster analysis") cuyo propósito es encontrar grupos internamente compactos y distinguibles entre sí, a partir de criterios de distancia o disimilitud. Una variante del análisis de conglomerados, es el algoritmo de las k-medias, cuya racionalidad de naturaleza estadística consiste en maximizar la relación entre la dispersión entre conglomerados con respecto a la dispersión dentro de los conglomerados. (5-6)

La segunda es el análisis de conglomerados basado en modelos (model-based cluster analysis) que puede considerarse como un caso particular del análisis de clases latentes (7-8) y que se origina en la suposición de que las observaciones provienen de una mezcla finita de subpoblaciones, que tienen una función de verosimilitud conjunta de la forma

$$p(x) = \prod_{i=1}^N \sum_{k=1}^K \pi_k f(x_i / \theta_k, \varphi_k) \quad (\text{IV.20})$$

en la cual las π_k designan las proporciones de cada subpoblación en la mezcla de todas ellas, θ_k, φ_k son los parámetros de la función de densidad f de cada una, y las x_i son variables socioeconómicas.³⁹ El fundamento lógico de este enfoque es que existe una variable \underline{g} , discreta y no observada, que define una clasificación relevante. \underline{g} puede considerarse como una variable faltante y sus valores pueden estimarse mediante alguno de los algoritmos de imputación conocidos. (9,10)⁴⁰

La tercera opción es el análisis de conglomerados basado en regresión (11), que es en rigor un caso particular de la opción precedente, bajo el supuesto de que existe una variable dependiente y de que cada subpoblación viene definida por un modelo de asociación diferente (con parámetros variables) entre dicha variable y un conjunto de predictores. Para el propósito que nos ocupa, la variable dependiente es una variable de salud relevante, y el conjunto de predictores está definido por un vector de variables socioeconómicas.

En el apartado siguiente se presentan con mayor detalle los dos últimos enfoques. El primero es mucho mejor conocido. (6)

³⁸ Considero que es eso lo que ocurre en Cuba. En el capítulo que sigue se exponen algunos resultados preliminares que apoyan esta presunción.

³⁹ Si f es la normal multivariada, entonces θ_k, φ_k representan el vector de medias y la matriz de covarianzas de cada subpoblación.

⁴⁰ El más común es el algoritmo EM, del que se habla más adelante en este mismo apartado.

4.8.1 El análisis de conglomerados basado en modelos

En los últimos años se ha publicado mucho sobre modelos de clases latentes y otros tipos de modelos basados en mezclas finitas. (12-14) Esto se debe en parte al desarrollo de algoritmos de cómputo que reducen considerablemente la complejidad del problema aunque éste involucre a una gran cantidad de variables, y aún en mayor medida a que estos modelos representan un avance sustancial con respecto a los enfoques tradicionales del análisis de conglomerados, el análisis factorial y los modelos de regresión. Todos estos modelos clásicos involucran relaciones entre variables observadas. Los modelos de clases latentes, sin embargo, incluyen una o más variables discretas no observadas.

Por otra parte, no descansan en las suposiciones, casi siempre no realistas, de los modelos clásicos (relaciones lineales, distribución Gaussiana y homogeneidad), y por consiguiente, son menos vulnerables a los sesgos asociados a las violaciones de los supuestos. Recientemente estos modelos se han extendido y permiten incluir mezclas de variables medidas en distintas escalas (ordinal, nominal, continua y de tipo conteo), e incorporar la relación de las clases latentes con covariantes y otras variables exógenas, con lo que se elimina una segunda fase del análisis para relacionar la composición de los conglomerados con variables demográficas o de otro tipo. (12-14)

El análisis de conglomerados clásico emplea algoritmos de clasificación no supervisados que agrupan casos cercanos de acuerdo a algún criterio ad hoc de distancia. (6) El análisis de conglomerados basado en modelos pertenece a una clase especial de modelos "difusos" (fuzzy models) que implican la estimación de probabilidades de pertenencia para clasificar los casos en el conglomerado apropiado. El más conocido de ellos es el modelo de mezclas finitas en donde cada clase latente representa un conglomerado oculto. (15-17)

Para contextualizar los modelos:

Si $X = (X_1, X_2, \dots, X_N)$ es una muestra de N personas o viviendas, supondremos que X fue extraída de una mezcla de K densidades de probabilidad del tipo

$$f(x) = \sum_{k=1}^K \pi_k f_k(x/\theta_k, \varphi_k) \quad (\text{IV.21})$$

La verosimilitud de las N observaciones viene dada por:

$$L = f(X) = \prod_{i=1}^N f(X_i) = \prod_{i=1}^N \sum_{k=1}^K \pi_k f_k(x_i/\theta_k, \varphi_k) \quad (\text{IV.22})$$

Es posible calcular la probabilidad de que la k -ésima persona o vivienda pertenezca al conglomerado k , si se conocen o pueden estimarse los parámetros de (IV.22), mediante la expresión

$$p_{ik} = \frac{\pi_k f_{ik}}{\sum_{k=1}^K \pi_k f_{ik}} \quad (\text{IV.23})$$

Los modelos de mezclas finitas Gaussianos, tienen funciones de densidad de la forma:

$$\phi(x_i) = \sum_{k=1}^K \pi_k \phi_k(x_i / \mu_k, \Sigma_k) \quad (\text{IV.24})$$

en la cual μ_k es el vector de medias y Σ_k la matriz de covarianzas, y la función de densidad ϕ tiene la forma:

$$\phi_k(x_i / \mu_k, \Sigma_k) = \frac{\exp\left[-\frac{1}{2}(x_i - \mu_k)^T \Sigma_k^{-1} (x_i - \mu_k)\right]}{\sqrt{\det(2\pi\Sigma_k)}} \quad (\text{IV.25})$$

Si el número de conglomerados (K) es fijo, π_k , μ_k y Σ_k se estiman mediante el algoritmo EM (expectation-maximization) (10, 18-20), que es la herramienta por excelencia en los modelos de mezclas finitas. El algoritmo se implementa bajo el supuesto de que hay datos faltantes en la variable que identifica los conglomerados.

Los datos completos son (G_i, x_i) en donde $G_i = (G_{i1}, G_{i2}, \dots, G_{iK})$ es la parte no observada de los datos y

$$G_{ik} = 1 \text{ si } x_i \text{ pertenece al grupo } k \\ 0 \text{ en caso contrario}$$

Supondremos además, que las G_i son independientes y tienen todas la misma distribución multinomial con probabilidades p_1, p_2, \dots, p_K .

El algoritmo EM es un procedimiento iterativo que alterna entre un paso de cálculo de valores esperados y un paso de maximización. En la iteración l -ésima el paso E consiste en la actualización las probabilidades de (IV.23) de modo que

$$p_{ik}^{(l)} = \frac{\pi_k^{(l-1)} f_{ik}(x_i / \theta_k^{(l-1)}, \phi_k^{(l-1)})}{\sum_{j=1}^K \pi_j^{(l-1)} f_{ik}(x_i / \theta_j^{(l-1)}, \phi_j^{(l-1)})} \quad (\text{IV.26})$$

Y el paso M, recalcula los parámetros de la distribución:

$$p_k^{(l)} = \frac{1}{n} \sum_{i=1}^n p_{ik}^{(l)}; \quad \mu_k^{(l)} = \frac{\sum_{i=1}^n p_{ik}^{(l)} x_i}{\sum_{i=1}^n p_{ik}^{(l)}} \quad \text{y (IV.27)}$$

$$\Sigma_k^{(l)} = \frac{\sum_{i=1}^n p_{ik}^{(l)} (x_i - \mu_k^{(l)})(x_i - \mu_k^{(l)})^T}{\sum_{i=1}^n p_{ik}^{(l)}} \quad \text{(IV.27')}$$

Estas estimaciones son las que maximizan la función de verosimilitud total (IV.22).

Varios autores han propuesto una descomposición basada en el cálculo de los autovectores de las matrices de covarianza, que permite una reparametrización (y con ella una simplificación) de los modelos (21,22) a partir de diferentes supuestos en relación con Σ_k .

La descomposición es de la forma:

$$\Sigma_k = \lambda_k \Delta_k A_k \Delta_k' \quad \text{(IV.28)}$$

en la cual:

- λ_k es una constante
- Δ_k es la matriz de los autovectores de Σ_k
- A_k es una matriz diagonal cuyos elementos son proporcionales a los autovalores de Σ_k

Habitualmente $\lambda_k = |\Sigma_k|^{1/v}$ en donde v es el número de variables observadas y A_k es convertida a una escala tal que $|A_k| = 1$. Una propiedad interesante de la descomposición dada por (IV.28) es que sus parámetros tienen una interpretación geométrica inmediata que regula la construcción de los conglomerados: λ_k representa el volumen, Δ_k la orientación y A_k la forma. Las alternativas se resumen en el cuadro siguiente:

Tabla IV. 5 Suposiciones acerca de la estructura de los conglomerados

Modelo	Descripción			Distribución de los conglomerados
	Volumen	Forma	Orientación	
λI	Igual	Igual	No	Esférica
$\lambda_k I$	Variable	Igual	No	Esférica
λA	Igual	Igual	Eje coordinado	Diagonal
$\lambda_k A$	Variable	Igual	Eje coordinado	Diagonal
λA_k	Igual	Variable	Eje coordinado	Diagonal
$\lambda_k A_k$	Variable	Variable	Eje coordinado	Diagonal
$\lambda \Delta \Delta \Delta^T$	Igual	Igual	Igual	Elipsoidal
$\lambda \Delta_k A \Delta_k^T$	Igual	Igual	Variable	Elipsoidal
$\lambda_k \Delta_k A \Delta_k^T$	Variable	Igual	Variable	Elipsoidal
$\lambda_k \Delta_k A_k \Delta_k^T$	Variable	Variable	Variable	Elipsoidal

Una de las desventajas de algoritmos más heurísticos para la formación de conglomerados es la falta de un criterio formal para decidir el número de éstos. El método de conglomerados basado en modelos puede descansar en reglas estadísticas de decisión. Una de ellas es el factor Bayesiano que se define como el cociente entre los "odds" a posteriori y los "odds" a priori, que puede aplicarse en conjuntos de modelos no anidados. En lugar del factor Bayesiano, que suele ser difícil de calcular, se usa el criterio de información Bayesiano (CIB) para decidir entre dos modelos cualesquiera. CIB se define del siguiente modo:

$$CIB = 2 \times (\log \text{verosimilitud}) - \eta \times \log(N) \quad (IV.29)$$

...en donde N es el número de casos (sujetos o viviendas) y η es el número de parámetros del modelo. Comúnmente diferencias de 2 o menos unidades en el CIB se consideran muy pequeñas para justificar una preferencia clara de un modelo sobre otro, y diferencias mayores que 10 unidades se consideran muy relevantes. (23)

Los métodos que en términos generales se han descrito en este apartado corresponden al programa MClust que es una rutina en software libre escrita en lenguaje de programación R. (20) En el próximo capítulo se expone un ejemplo con datos tomados de un estudio realizado en La Habana en el que se utiliza este programa.

4.8.2 El análisis de conglomerados basado en regresión.

La lógica de este enfoque radica en que el contexto -en su sentido más amplio, que lo identifica como las circunstancias en que los individuos nacen, se desarrollan y mueren-

no sólo tiene poderosos efectos sobre la salud individual o familiar, sino que modifica el efecto de otros factores. Diferentes contextos pueden configurar asociaciones particulares entre las variables socioeconómicas y las variables de salud. En este sentido, el enfoque del presente apartado puede considerarse un caso particular del anterior, que concibe al contexto como una dimensión representada por una variable discreta latente, para los cuales cambian los valores de los parámetros de la relación que vincula a la salud con la condición socioeconómica.

Una diferencia esencial con respecto al enfoque anterior basado en mezclas finitas de poblaciones, es que, en este caso, el modelo subyacente identifica de manera explícita a una variable dependiente y a un vector de variables independientes. Otra diferencia esencial es que el método propuesto en 4.8.1 es un método difuso ("fuzzy clusters") mientras que los conglomerados basados en regresión generan una partición. (24)

En el análisis de conglomerados basados en regresión consideramos un conjunto de n personas o viviendas $\Theta = \{1, 2, \dots, n\}$ cuyos datos tienen la forma (y_j, x'_j) , ($j = 1, 2, \dots, n$) donde $x_j \in R^v$ es un vector columna no aleatorio y $y_j \in R$ es una variable aleatoria dependiente para el sujeto (o la familia) j -ésima. Se supone además que existe una partición fija latente $\Omega = (\gamma_1, \gamma_2, \dots, \gamma_K)$ de Θ , que hay recuperar a partir de los datos.

El modelo básico es:

$$y_j = x'_j \beta_i + \varepsilon_j; \quad \varepsilon_j \square N(0, \sigma_i^2) \quad \text{para todo } j \in \gamma_i \quad i = 1, 2, \dots, K \quad (\text{IV.30})$$

Para el modelo (IV.30) la función de verosimilitud viene dada por

$$\log L_n(\gamma_i, (\beta_i, \sigma_i^2)_{i=1,2,\dots,K}) = -\frac{1}{2} \sum_{i=1}^K \sum_{j \in \gamma_i} \left(\log(2\pi) + \log(\sigma_i^2) + \frac{(y_j - \beta'_i x_j)^2}{\sigma_i^2} \right) \quad (\text{IV.31})$$

cuyo máximo para $(\hat{\beta}_i, \hat{\sigma}_i^2)_{i=1,2,\dots,K}$ dado, es:

$$\hat{\gamma}_i = \arg \min_i \left[\log(\sigma_i^2) + \frac{(y_j - \beta'_i x_j)^2}{\sigma_i^2} \right] \quad (\text{IV.32})$$

Por otra parte, para un $\hat{\gamma}_i$ dado (IV.31) es la suma de los logaritmos de las funciones de verosimilitud de funciones de regresión homogéneas dentro de los conglomerados. Por

tanto el máximo de (IV.31) se obtiene para el estimador máximo verosímil $\widehat{\beta}_i$ basado en los valores (y_j, x_j) con $j \in \gamma_i$ y el estimador de la varianza dado por:

$$\sigma_i^2 = \frac{\sum_{j \in \gamma_i} (y_j - \widehat{\beta}_i' x_j)^2}{\widehat{n}_i}; \quad i = 1, 2, \dots, K \quad (\text{IV.33})$$

en donde \widehat{n}_i es el número de casos en el conglomerado $\widehat{\gamma}_i$.

El ejemplo que se expone en el capítulo siguiente se ha resuelto mediante el algoritmo de conglomerados basados en regresión contenido en el software NCSS, versión 2007.

4.9 Presentación y discusión de ejemplos

En lo que resta de este capítulo se presentan dos ejemplos, el primero de ellos a partir de datos reales tomados de fuentes del Ministerio de Salud de Nicaragua/Dirección General de Planificación y Desarrollo/Oficina Nacional de Estadísticas. 2007 (25), y el segundo con datos ficticios. El primer ejemplo ilustra el empleo de todos los índices y en particular las novedades metodológicas expuestas; el segundo se detiene en otros aspectos generales de interés sobre la medición de las desigualdades y su interpretación. Se pondrá especial énfasis en mostrar que este empeño es ante todo un ejercicio de razonamiento crítico y no meramente un asunto técnico que se contrae a la aplicación de los índices. En el proceso emergen apreciaciones de valor, que también desbordan lo exclusivamente técnico.

La medición de la desigualdad latente se ilustra en el capítulo siguiente con ejemplos tomados de una investigación realizada en varios municipios habaneros en el contexto de un proyecto con otros propósitos más abarcadores. (26)

Ejemplo No. 1

A partir de la población de nacidos vivos (pobnvr) y de muertes de niños menores de un año registrados (m<1r) se calculó la tasa de mortalidad infantil (TMI). Las variables socioeconómicas (porcentaje de analfabetos, porcentaje poblacional con educación secundaria o superior, porcentaje de viviendas con agua potable y porcentaje de viviendas sin servicio sanitario en su interior) fueron transformadas convenientemente para que todas tuvieran la misma polaridad.⁴¹ Con estas cuatro variables se practicó un análisis de componentes principales y se calculó un índice sintético (la primera componente principal) que captura más de 80% de la variabilidad en las variables originales. Este índice sintético (que denominamos F1) se utilizó para ordenar los

⁴¹ Las variables positivas (educación secundaria y agua potable dentro de la vivienda) fueron transformadas por complemento con respecto al 100%, de modo que las cuatro variables tuvieran una polaridad negativa.

municipios en forma creciente (de peor a mejor).⁴² Con respecto al arreglo ordenado que genera este índice sintético se calculan los índices de desigualdad para variables en escala ordinal o continua. La tabla IV.6 contiene las tasas departamentales de mortalidad infantil en 2001 y 2007.

Tabla IV.6 Tasas de mortalidad infantil (por mil nacidos vivos) en departamentos nicaragüenses en 2001 y 2007 ordenadas según un índice sintético de la condición socioeconómica (F1)

2001			2007		
Departamento	TMI	F1	Departamento	TMI	F1
Jinotega	13.3	3.3	RAAN	19.3	2.8
RAAN	21.9	2.6	Jinotega	14.7	2.4
Rio San Juan	8.6	2.3	Rio San Juan	15.7	2.1
RAAS	15.8	2.3	RAAS	13.7	2.0
Boaco	17.3	1.2	Boaco	17.0	1.6
Madriz	15.2	0.9	Matagalpa	14.8	1.1
Matagalpa	18.0	0.8	Madriz	11.8	0.8
Nueva Segovia	20.6	0.1	Chontales	11.5	0.7
Chinandega	26.3	-0.6	Nueva Segovia	19.1	0.5
Chontales	9.6	-0.7	Estelí	15.2	-0.8
Rivas	8.6	-0.7	Chinandega	18.7	-1.1
Estelí	12.6	-0.8	Rivas	13.6	-1.2
Masaya	74.4	-1.5	Masaya	73.9	-1.5
León	17.0	-1.6	León	20.1	-1.6
Granada	15.1	-1.9	Granada	19.0	-1.9
Carazo	19.1	-1.9	Carazo	18.3	-2.0
Managua	3.5	-3.9	Managua	2.4	-4.1
Total nacional	16.7		Total nacional	15.6	

Entre 2001 y 2007, la TMI en el país disminuyó en 1.1 puntos porcentuales: desde 16.7 hasta 15.6. En 10 de los 17 departamentos la TMI disminuyó. Vale la pena notar además, que hubo mejorías discretas en casi todos los indicadores socioeconómicos. El lector puede verificar (ver datos en Anexo 3) que los porcentajes de analfabetismo, población adulta sin educación secundaria, viviendas sin agua potable y sin condiciones de saneamiento básico (que son las variables que se integran en F1) se comportaron del siguiente modo entre 2001 y 2007:

- analfabetismo: de 22.9% a 18.8%
- sin educación secundaria: de 80.0% a 76.3%
- sin agua potable en el interior de la vivienda: de 69.9% a 70.3%
- sin condiciones de saneamiento básico: de 14.8% a 12.9%

⁴² También son índices sintéticos muy usados en el contexto de la medición de desigualdades, el índice de desarrollo humano (IDH) y las necesidades básicas insatisfechas (NBI).

La pregunta clave es: ¿han aumentado o han disminuido las desigualdades en ese lapso de 6 años?

Antes de aplicar toda la batería métrica de desigualdades a los datos del cuadro anterior, conviene echarles una ojeada crítica. Varias cosas saltan a la vista. La primera es que Masaya tiene un comportamiento atípico, con una TMI (tanto en 2001 como en 2007) considerablemente más alta que la tasa total del país y que la de cualquier otro departamento. ¿Se incluye a Masaya en el cálculo de las desigualdades asumiendo que las causas de esos valores excepcionalmente altos de la TMI tienen poco que ver con las condiciones socioeconómicas descritas en el análisis, o se mantiene este departamento?⁴³ La opinión de este autor es que Masaya debería excluirse porque su presencia introduce una disociación entre la dimensión socioeconómica y la dimensión salud, que resultaría ajena al propósito central de medir desigualdades sociales en salud. En el recuadro siguiente se incluye el análisis de los residuos de una regresión de TMI sobre F1, que muestra claramente que Masaya es un caso atípico con un seguro efecto de apalancamiento sobre varios de los índices de desigualdad.

Las TMI del país, sin incluir a Masaya, fueron respectivamente 13.5 y 12.6 en 2001 y 2007 y en 9 de los 16 departamentos la TMI disminuyó.

Recuadro IV.1 Valores predichos, residuos y residuos estandarizados de una regresión de TMI sobre F1.

SILAIS	F1	tmi	Pred(tmi)	Residual	Std. residual
Jinotega	3.279	13.266	22.193	-8.927	-0.560
RAAN	2.562	21.940	21.206	0.733	0.046
Rio San Juan	2.345	8.602	20.908	-12.306	-0.772
RAAS	2.258	15.787	20.789	-5.002	-0.314
Boaco	1.245	17.262	19.397	-2.135	-0.134
Madriz	0.936	15.222	18.971	-3.749	-0.235
Matagalpa	0.789	17.974	18.769	-0.795	-0.050
Nueva Segov	0.138	20.601	17.874	2.726	0.171
Chinandega	-0.551	26.339	16.928	9.411	0.590
Chontales	-0.655	9.560	16.785	-7.225	-0.453
Rivas	-0.683	8.553	16.746	-8.194	-0.514
Estelí	-0.795	12.622	16.592	-3.970	-0.249
Masaya	-1.482	74.362	15.648	58.714	3.683
León	-1.622	16.994	15.456	1.538	0.096
Granada	-1.937	15.148	15.022	0.125	0.008
Carazo	-1.942	19.079	15.016	4.063	0.255
Managua	-3.884	3.495	12.347	-8.853	-0.555

⁴³ Cualquiera sea la respuesta, es importante tomar en consideración que la desigualdad, relativa a la jerarquía socioeconómica, subestima considerablemente la desigualdad en salud, si el cociente de las tasas extremas incluye a los valores de Masaya en el numerador

El segundo aspecto a subrayar es que, aunque hay una notable correspondencia entre los ordenamientos socioeconómicos que genera el indicador sintético, la coincidencia no es total. Por ejemplo, en términos socioeconómicos (de acuerdo a la métrica F1) la peor situación en 2001 corresponde a Jinotega, y en 2007, a RAAN.⁴⁴ Esto implica que el cálculo de tasas extremas no incluye en el denominador a las mismas unidades de análisis.

Lo más importante, sin embargo, es que el ordenamiento según el indicador socioeconómico difiere claramente del ordenamiento según las TMI, aunque haya un alto grado de covariación entre la dimensión socioeconómica y la dimensión de salud. La TMI más alta no es la del departamento con la peor situación según la métrica socioeconómica F1.⁴⁵

El examen de los índices muestra grandes incrementos para los que se basan en rangos, y ello fundamentalmente debido a la inversión de los departamentos con el peor valor de F1 y a la reducción de la tasa de Managua. Las dos versiones de la varianza entre grupos (VEG y VEG*) disminuyen y también lo hace el IDP. El resto de los índices se incrementa. El incremento relativo es notablemente alto para los índices basados en comparaciones dos a dos, para las dos versiones del IPK (la ponderada y la no ponderada) y moderadamente alta para la DLM, todo ello, por supuesto, de acuerdo a una apreciación comparativa, necesariamente subjetiva. Hay una diferencia notable entre los valores ponderados y no ponderados de los índices basados en rangos. La ponderación achica las brechas porque éstas se multiplican por un factor menor que 1, debido a que la población con las tasas menores es mucho mayor.

Los índices de desigualdad que toman en cuenta toda la jerarquía poblacional y no sólo los valores extremos, experimentan incrementos inferiores a 10%. El índice de concentración tiene también un pequeño incremento de 1.1%.

⁴⁴ En el centro de la distribución hay otras inversiones en el ordenamiento de los departamentos

⁴⁵ Otro tanto acontece si en lugar de F1 se construye el ordenamiento a partir de cualquiera de los cuatro indicadores socioeconómicos por separado.

Tabla IV.7 Índices de desigualdad social en salud. Nicaragua 2001 y 2007 (sin incluir el departamento de Masaya)

Indices	2001	2007	Incremento relativo
CT	3.8	8.0	110.5
CPT	1.2	1.8	50.0
DT	9.8	16.9	72.4
DPT	3.0	3.8	26.7
RAP	10.0	10.2	2.0
RAP(%)	74.1%	81.0%	9.3
IDP	-15.36	-14.5	-5.6
IRD(Pamuk)	-1.14	-1.20	5.3
IRD (K-M)	3.63	3.72	2.5
IRD (corregido)	2.67	2.69	0.7
IPK	3.34	5.38	61.1
VEG	50.81	41.90	-17.5
IPK*	2.87	4.26	48.4
VEG*	3.76	3.32	-11.7
DLM	0.20	0.26	30.0
THEIL	0.16	0.17	6.3
I_{int}	0.36	0.42	16.7
Concentración	-0.186	-0.188	1.1

Supongamos ahora que en el lapso de 2001 a 2007, todo hubiese ocurrido exactamente igual a como ocurrió, con la única diferencia de que Managua, la capital, no hubiese sido tan exitosa en la reducción de su ya baja tasa de mortalidad infantil y que ésta se mantuviera en el mismo nivel de 3.5 x mil nacidos vivos, que ya exhibía 6 años antes.

La medición de la desigualdad se habría comportado del siguiente modo:

Tabla IV.8 Índices de desigualdad social en salud. Nicaragua 2001 y 2007 (asumiendo TMI invariante en Managua y sin incluir el valor atípico de Masaya)

Indices	2001	2007	Incremento relativo
CT	3.8	5.5	44.7
CPT	1.2	1.2	0
DT	9.8	15.8	61.2
DPT	3.0	3.6	20
RAP	10.0	9.4	-6
RAP(%)	74.1%	72.9%	-1.6
IDP	-15.4	-13.5	-12.3
IRD(Pamuk)	-1.14	-1.03	-9.6
IRD (K-M)	3.63	3.11	-14.3
IRD (corregido)	2.67	2.39	-10.5
IPK	3.34	3.39	1.5
VEG	50.8	36.3	-28.6
IPK*	2.87	2.68	-6.6
VEG*	3.76	2.81	-25.3
DLM	0.20	0.18	-10
THEIL	0.16	0.13	-18.8
I_{int}	0.36	0.31	-13.9
Concentración	-0.186	-0.167	-10.2

Es notable, que ahora el panorama cambia completamente. Casi todos los índices (salvo los que se basan en las comparaciones dos a dos y el IPK) habrían disminuido. La diametral diferencia se debe sólo al hecho de que nos hemos basado en el supuesto de que Managua no alcanzó los avances que sí experimentó en la realidad. El impacto que esto tiene sobre la apreciación de la realidad es radical. En gran parte, debido a que Managua redujo su tasa de mortalidad de 3.5 a 2.4 x 1000 nacidos vivos (lo cual representa un total de 34 muertes menos de niños menores de un año que las que se habrían producido de mantener las ya bajas tasas), las desigualdades se incrementaron.

Si volvemos a la pregunta inicial (¿han disminuido o han aumentado las desigualdades?) se repara de inmediato en la necesidad de matizarla. Las desigualdades sanitarias (desigualdades en las tasas de mortalidad infantil no asociadas a la dimensión socioeconómica) se han incrementado si se toma como referente a la tasa menor (correspondiente a Managua) para el cálculo de las brechas, pero han disminuido si se toma como referente al promedio del país. Esta elección no es, claramente, de naturaleza meramente técnica. ¿Es razonable esperar que todo el país pueda alcanzar los indicadores de la capital? ¿O es más natural medir la desigualdad con respecto a los indicadores promedio del país? De esta disyuntiva depende la elección entre el IPK o la

VEG, el IPK* o la VEG*, y por consiguiente la respuesta en cuanto al aumento o la disminución de la desigualdad.

De acuerdo a la racionalidad implícita en los indicadores basados en entropía, la desigualdad se ha incrementado, de acuerdo a los resultados que muestra la tabla IV.7. El índice integrado -que se recomienda en este texto- revela un incremento de 0.36 a 0.42 (16.7%). La diferencia entre la cantidad real de muertes de niños menores de un año y la que correspondería a la cifra esperada según la cantidad de nacimientos vivos, fue mayor en 2007, y si esas diferencias son ponderadas de acuerdo al logaritmo de las tasas (para que tengan mayor contribución en departamentos con tasas altas) se tiene el incremento relativo ya referido de 16.7 por ciento.

Si lo que interesa, por el contrario es la desigualdad asociada a la condición socioeconómica (es decir, en su covariación con dicha condición), la desigualdad ha experimentado un ligero incremento entre 2001 y 2007, que en términos relativos es menor del 5% para todos los índices. En particular para el índice relativo de desigualdad corregido es de apenas 0.7%. En términos de concentración de la carga de mortalidad, el incremento fue de 1.1%

Nuevamente, es importante notar el impacto que tiene la capital Managua, que concentra más del 20% de los nacimientos del país. La gran disminución relativa de la TMI de casi un tercio de su valor inicial, ponderada por su tamaño poblacional (en particular por su número de nacimientos) tiene un enorme peso en la tendencia de la desigualdad en el país. Si Managua hubiese mantenido su baja tasa inicial de mortalidad infantil (3.5 por mil nacidos vivos), tanto la desigualdad sanitaria per se, como la desigualdad asociada a la condición socioeconómica, habrían experimentado una disminución relativa de entre 10 y 20% para casi todos los índices, con la excepción de los que se basan en comparaciones entre los extremos de la jerarquía socioeconómica.

Ejemplo No. 2

En este ejemplo no se aplica ningún índice en particular, sino que se desarrolla un simple ejercicio crítico de evaluación del cambio que toma en cuenta la relación entre dos estratos socioeconómicos dados por la escolaridad, y la concentración de un indicador de salud. El ejemplo ilustra un problema relacionado con la interpretación de la medición de la desigualdad cuando en el mecanismo que la produce se mezclan causas injustas con factores estructurales irreductibles. En tales casos puede ocurrir que al actuar sobre las raíces moralmente inaceptables de la desigualdad, la resultante sea un incremento relativo de esta que se produce al disminuir la carga total de enfermedad en la población y quedar confinada dicha carga a los segmentos poblacionales con una vulnerabilidad esencial que no responde a las acciones políticas o programáticas para corregirla. Supongamos que la relación que vincula la escolaridad materna con la desnutrición crónica en niños menores de 5 años es la que muestran los datos ficticios del cuadro siguiente en dos tiempos t_0 y t_1 .

Analfabetismo materno y desnutrición crónica en el menor de 5 años (tiempo t_0)

Escolaridad	Desnutridos	No desnutridos	Total
Analfabetas	20	80	100
No analfabetas	20	880	900
Total	40	960	1000

Analfabetismo materno y desnutrición crónica en el menor de 5 años (tiempo t_1)

Escolaridad	Desnutridos	No desnutridos	Total
Analfabetas	18	32	50
No analfabetas	10	940	950
Total	28	972	1000

La asociación entre la desnutrición en el menor de 5 años y la escolaridad materna es bien conocida. La literatura recoge varias experiencias acerca de la reducción de la desnutrición crónica mediante el incremento de la escolaridad de la madre. Mejorar la provisión de alimentos tiene un impacto reducido en la nutrición del niño pequeño, si no se desarrollan programas que promuevan la educación materna. (27)

Supongamos entonces que se trabaja sobre la población ficticia del primero de los dos cuadros anteriores y se reduce el analfabetismo materno en 50% (de 10% a 5% en el lapso entre t_0 y t_1). La desnutrición crónica se ha reducido en ese tiempo, de 4% a 2.8%, en gran medida, debido a los programas de alfabetización. Sin embargo, dada la covariación entre educación materna y desnutrición, gran parte de la desnutrición crónica que prevalece, se concentra entre los hijos de madres analfabetas. En efecto, en el tiempo t_0 , 10% de la población concentraba al 50% (20/40) de la desnutrición crónica del menor de 5 años, mientras que en el tiempo t_1 , 5% de la población concentra 64% de la desnutrición crónica (18/28).

No hay dudas de que la situación que describe el segundo de los cuadros anteriores es más favorable que la que describe el primero (el analfabetismo se redujo a la mitad y la desnutrición crónica a menos de la mitad) aunque se concentra cada vez más en la población analfabeta y aunque el cociente de las tasas se ha incrementado notablemente: 9 en t_0 ((20/100) / (20/900)) y 34.2 en t_1 ((18/50) / (10/950)).

Referencias

1. Wagstaff A. Inequality aversion, health inequality and health achievement. The World Bank Development Research Group Public Services and Human Development Network Health, Nutrition and Population Team. Policy Research Working Paper 2765. 2002.
2. World Health Organization. Regional Office for Europe. Environmental health inequalities in Europe. Assessment Report. Copenhagen. 2012.
3. Northington V, Stone D, Ladenheim K, Gibbs BK. Comparative perspectives on health disparities. *J Health Politics, Policy and Law* 2006; 31: 1-2. Disponible en: <http://jhppl.yale.edu/current/abstracts311.html>.
4. Spinakis A, Anastasiou G, Panousis V, Spiliopoulus K, Palaiologou S, Yfantopoulos J. Expert Review and Proposals for Measurement of Health Inequalities in the European Union. Full Report. European Commission Directorate General for Health and Consumers. Luxembourg 2011.
5. Andreopoulos B, An A, Wang X, Schroeder M. A roadmap of clustering algorithms: finding a match for a biomedical application. *Briefing in Bioinformatics* 2009; 10: 297-314.
6. Bacallao J. Cluster analysis. In Sarah Boslaugh (ed). *Encyclopedia of Epidemiology*. New York. Sage. 2007.
7. Galimberti G, Soffritti G. Finite mixture models for clustering multilevel data with multiple cluster structures. *Statistical Modelling* 2010; 10: 265-290.
8. Vermunt JK, Magidson J. Latent class cluster analysis. En: Hagenaars JA, McCutcheon AN. *Applied Latent Class Analysis*. Cambridge. Cambridge University Press. 2002).
9. Vermunt JK, Magidson J. Latent class models for classification. *Comput Stat and Data Analysis* 2003; 41: 531-537.
10. Vermunt JK, Magidson J. *LG-syntax User's Guide*. Manual for Latent Gold 4.5 Syntax Module. Belmont MA. Statistical Innovations. 2008.
11. Qian G, Wu Y, Shao Q. A procedure for estimating the number of clusters in logistic regression clustering. *Journal of Classification* 2009; 26: 183-199.
12. De Boeck P, Wilson M, Acton DS. A conceptual and psychometric framework for distinguishing categories and dimensions. *Psychological Rev* 2005; 112: 129-158.
13. Lubke GH, Neale MC Distinguishing between latent classes and continuous factors: resolution by maximum likelihood? *Multivariate Behav Res* 2006; 41: 499-532.
14. Vermunt JK, Tran B, Magidson J. Latent class models in longitudinal research. En: Menard S (Ed). *Handbook of longitudinal research. Design, measurement and analysis*. Burlington, Ma. Elsevier. 2008: 373-385.
15. Yang MS, Yu NY. Estimation of parameters in latent class models using fuzzy clustering algorithms. *Eur J Oper Res* 2005; 160: 515-531.;
16. Steinley D, Brusco MJ. Evaluating mixture modeling for clustering: recommendations and cautions. *Psychol Methods* 2011; 16: 63-79.

17. Vermunt JK. K-Means may perform as well as mixture model clustering but may also be much worse. Comment on Steinley and Brusco (2011). *Psychol Methods* 2011; 16: 82-88.
18. Dempster AP, Laird NM, Rubin DB. Maximum likelihood for incomplete data via the EM algorithm. (with discussion). *J Royal Statistical Soc 1977 Series B*; 39: 1-38.
19. McLachlan G, Krishnan T. *The EM algorithm and extensions*. New York. Wiley. 1997.
20. Fraley C, Raftery AE. MCLUST: Software for model-based cluster analysis. *J Classif* 1999; 16:297-306)
21. Fraley C, Raftery AE. Model-based clustering, discriminant analysis and density estimation. *JASA* 2002; 97: 611-612.
22. Pourahmadi M, Daniels M, Park T. Simultaneous modeling of the Cholesky decomposition of several covariance matrices. *J Multivariate Anal* 2007; 98: 568-587).
23. Raftery AE, Dean N. Variable selection for model clustering. Technical Report. Department of Statistics. University of Washington. 2004).
24. Qian G, Wu Y. Estimation and selection in regression clustering. *Eur J of Pure and Applied Math* 2011; 4: 455-466.
25. Datos estadísticos básicos. INDE, MINSA: DGDP-ONE, Nicaragua, 2008.
26. Díaz-Perera G. Determinantes contextuales de las enfermedades asociadas a la aterosclerosis y sus factores de riesgo. Tesis presentada en opción a la categoría de Doctor en Ciencias de la Salud. La Habana. 2011.
27. Smith LC, Hadad LJ. Explaining child malnutrition in developing countries. A cross country analysis. International Food Policy Research Institute. Research Report No. 111. New York. 2000.

CAPÍTULO V

DESIGUALDADES EN SALUD EN RELACIÓN CON DIMENSIONES SOCIOECONÓMICAS NO EXPLÍCITAS. UN EJEMPLO EN CIUDAD DE LA HABANA

*El ojo que ves
No es ojo porque tú lo veas
Es ojo porque te ve.*

Antonio Machado

Resumen

Este capítulo se dedica básicamente a exponer aplicaciones de los aportes metodológicos contenidos en el apartado 4.8 del capítulo precedente, que por su trascendencia conceptual son bastante más que contribuciones métricas en el ámbito de las desigualdades sociales en salud. Una razón adicional y no menos importante para dedicar un capítulo separado a estas aplicaciones es que ellas se realizan a partir de los datos recogidos en un proyecto de investigación sobre determinantes contextuales de la aterosclerosis incluido en el plan temático del Centro de Investigaciones y Referencia de Aterosclerosis de La Habana, del cual el tema de las desigualdades sociales es un componente importante. Los resultados ya fueron parcialmente expuestos en una tesis de doctorado tutelada por el autor de esta tesis y en dos publicaciones recientes. Los resultados que se exponen en este capítulo tienen el propósito de mostrar la conveniencia de explorar modos alternativos de estratificar la población para la medición de las desigualdades sociales en salud. Se sugiere que, a falta de la posibilidad de fundamentar hipótesis concretas en relación con los determinantes de las desigualdades y sus mecanismos de acción, la búsqueda inductiva de dimensiones latentes, es un buen derrotero analítico para satisfacer este propósito.

Ya se ha argumentado en este propio texto que si los criterios de estratificación social no se eligen adecuadamente, grandes desigualdades sociales podrían quedar ocultas tras una baja desigualdad entre grupos. Las propuestas del índice de concentración ajustado y del análisis de conglomerados basados en modelos, que se expusieron en los apartados 4.6 y 4.8 del capítulo precedente, parten de esa premisa. Ciertos criterios de estratificación normativos como el sexo, la raza, el ingreso o la educación, que en muchos países exhiben grandes -y a veces crecientes- diferencias con respecto a muchos indicadores de salud, podrían no ser útiles en otros países que han avanzado en

la reducción de las desigualdades en esos ejes, pero que no por ello han conseguido erradicarlas.

Desde hace años se viene prestando especial atención al aumento de las desigualdades sociales en Cuba y a los fenómenos económicos, políticos, sociales y culturales que las generan y fomentan. (1-7) Algunos estudios han abordado, directa o indirectamente, las manifestaciones de estas desigualdades en el campo de la salud. (8-11) Más recientemente, y como parte de un proyecto de investigación institucional bajo la dirección del autor de esta tesis, se han estudiado las desigualdades sociales asociadas a los factores de riesgo y a los determinantes de la aterosclerosis. (12,13).

Es un hecho reconocido que en Cuba las crisis económica a escala nacional e internacional han tenido un efecto mucho menor sobre los principales logros en materia de salud, que en otros países. (14) Se han conjugado para ello diversos factores entre los que se cuentan el principio constitucional del derecho universal a la salud, la estructura organizativa del sistema de salud que garantiza el acceso sin exclusiones, un activismo centralizado del Estado cubano, que ha sido exitoso en garantizar la participación comunitaria, y una reconocida cultura sanitaria de la población.

El propósito de este trabajo no es adentrarse en el estudio y análisis de las desigualdades sanitarias asociadas a la condición socioeconómica en Cuba, y mucho menos en sus causas, pero sí pretende aportar indicios cuantitativos objetivos de la existencia de dimensiones latentes en relación con las cuales la población cubana exhibe heterogeneidades que se asocian a diferencias en los perfiles epidemiológicos y de riesgo.

El objeto de estudio de los resultados que se exponen en este capítulo son los factores de riesgo y las enfermedades asociadas a la aterosclerosis, cuyos determinantes individuales y contextuales son el centro de interés de un proyecto de investigación en el plan temático del Centro de Investigaciones y Referencia de Aterosclerosis de La Habana (CIRAH).

Los detalles del diseño, el proceso de recogida de la información y las variables utilizadas pueden encontrarse en tres publicaciones recientes (12, 14, 15) y se resumen brevemente a continuación. Una conclusión común de estas referencias, que se toma como punto de partida para las aplicaciones que contiene este apartado, es la relativamente baja covariación entre la dimensión económica y la dimensión salud en Cuba.

Se estudiaron 840 familias pertenecientes a 12 consultorios de La Habana, agrupados en 7 áreas de salud. Las áreas de salud y los consultorios fueron seleccionados a discreción, a partir de criterios prácticos de factibilidad relacionados con la presencia de residentes de MGI en todas esas áreas con intereses en la temática y con proyectos de investigación en la línea de los determinantes sociales de las enfermedades asociadas

a la aterosclerosis. En cada consultorio las familias fueron seleccionadas mediante muestreo simple aleatorio, hasta completar la cuota de 70 familias por consultorio.

Se midieron tres variables de salud a nivel de la familia: las densidades de hipertensión, diabetes y fumadores entre personas de 20 años o más. Las densidades son variables cuya unidad de anclaje es la vivienda, y se definen, respectivamente, como el cociente entre hipertensos, diabéticos y fumadores entre el total de susceptibles. Los sujetos se definieron como positivos por declaración en el caso de las tres variables, por declaración o dispensarización en el caso de la diabetes, y por declaración, dispensarización o toma de la tensión arterial (reiterada si resultó ser alta en la primera medición) en el caso de la hipertensión. Las variables socioeconómicas fueron:

- El número de personas por vivienda.
- La máxima escolaridad en el hogar (el mayor de los grados de escolaridad alcanzados por los miembros del hogar: primaria, secundaria, preuniversitaria o equivalente y universitaria)
- El número de equipos de primera necesidad (radio, plancha, ventilador, televisor, refrigerador y olla de presión)
- El número de equipos de segunda necesidad (grabadora, batidora, aire acondicionado, tocadiscos, horno microonda, video, máquina de coser, congelador, calentador de agua y otros).
- La percepción subjetiva de la condición económica de la familia (excelente, buena, regular, mala o muy mala)

Dentro de un marco de opciones prácticas limitadas por los objetivos y el alcance de un proyecto de investigación mayor que le dio cobertura⁴⁶ la elección de las variables anteriores se justifica con los siguientes argumentos:

El número de personas por vivienda: esta variable, particularmente cuando se expresa en términos relativos al número de habitaciones de la vivienda, suele ser un buen marcador, en otros países, de las condiciones de higiene y saneamiento básico que se asocian con la pobreza y con gran cantidad de indicadores de salud. En Cuba, dada la precariedad del fondo habitacional que padecen, sobre todo, las áreas urbanas, se ha generado una gran heterogeneidad en la composición de la familia: de un lado "familias extendidas" compuestas por personas de varias generaciones, con nexos muy variados de consanguinidad, y con modos de vida y perfiles de salud muy diferentes. El tamaño y la composición familiar han sido objeto de estudio de diversos autores con hallazgos no siempre coincidentes en relación con la pobreza y las desventajas sociales. (16-19)

La máxima escolaridad en el hogar: la escolaridad es una de las primeras opciones como indicador de condición socioeconómica, debido fundamentalmente a que en etapas

⁴⁶ Determinantes individuales y contextuales de los factores de riesgo y las enfermedades consecuentes de la aterosclerosis. Centro de Investigaciones y Referencia de Aterosclerosis de La Habana (CIRAH): 2009 - 2012

relativamente tempranas de la vida se alcanzan valores invariantes de este indicador, lo que lo pone a salvo del fenómeno de causalidad reversa que padecen otros indicadores como el empleo, que influye y a su vez es influido por la salud. (20) Como marcador de la condición socioeconómica, el indicador "escolaridad" tiene en Cuba dos problemas: su escasa variabilidad y la también escasa representatividad en los grupos de escolaridad más baja, que son a su vez, los más vulnerables en términos de riesgo y de conciencia de riesgo. Se optó por elegir la escolaridad máxima dentro de la familia, bajo el presupuesto de que ese miembro familiar con mayor escolaridad sería capaz de ejercer por su cuenta todos los efectos potencialmente atribuibles a la variable.

Los equipos de primera y de segunda necesidad: las condiciones del hogar han sido también tradicionalmente un indicador socioeconómico preferido en las investigaciones sanitarias. (21) Los habituales son el agua, la electricidad y el saneamiento básico, pero también se han empleado las posesiones del hogar (propiedad del inmueble, posesión de automóvil etc.). El acceso o no a artículos de primera necesidad en el hogar es un marcador específico de pobreza muy útil en sociedades pobres, que exhiben grandes brechas en el acceso a los servicios de salud. En Cuba, y particularmente en La Habana, la posesión de ciertos artículos como computadoras, hornos microondas y congeladores, podrían marcar la diferencia entre familias con y sin acceso a los mercados en divisa, lo cual a su vez podría asociarse con la calidad de la alimentación y la conciencia de riesgo, en relación, por ejemplo, con ciertas conductas como el tabaquismo.

La percepción subjetiva de la condición económica: la psicobiología ha hecho relevantes aportes recientes a la comprensión de los mecanismos que vinculan la condición socioeconómica y la salud. Uno de esos aportes se relaciona con el significado de la condición socioeconómica relativa y las reacciones que ella desencadena en el sujeto a lo largo del curso de la vida. (22) En tal sentido, el modo en que el individuo se contempla a sí mismo, en relación con las otras personas de su comunidad o de su entorno inmediato podría tener una asociación mucho mayor con su salud, que su propia condición socioeconómica objetiva. La inclusión de esta variable responde a la intención de capturar de algún modo esa posición relativa y no a la de disponer de un marcador de la condición socioeconómica objetiva. Claramente, la percepción subjetiva lleva implícitos los sesgos de la personalidad y no es, por tanto, un buen marcador de la condición socioeconómica real. Dentro de la familia, se resolvió trabajar con la percepción más favorable y con la menos favorable. Los resultados que se exponen en este capítulo corresponden a esta última debido a que no hubo diferencias perceptibles. No es posible descartar que en relación con esta variable pueda presentarse un fenómeno de doble causalidad o incluso de causalidad reversa. En efecto, la condición económica puede influir sobre la salud, pero su percepción subjetiva puede depender de la salud individual.

Procedimientos

Se basan íntegramente en los procedimientos ya descritos en 4.8.1 y 4.8.2

Resultados

a. Aplicación del análisis de conglomerados basado en modelos (ver 4.8.1)

La Tabla V.1 muestra la mezcla de subpoblaciones y sus proporciones respectivas en 794 de las 840 viviendas estudiadas. Las 46 restantes no fueron consideradas por contener alguna variable faltante. Uno de los conglomerados que contiene una proporción estimada entre 10.7 y 11.9 de casos, presenta, como se verá oportunamente, un perfil de riesgo diferente al resto de los conglomerados.

Tabla V.1

Distribución de las familias por conglomerados: proporción estimada de los conglomerados en la mezcla poblacional.

Conglomerados	No. de familias	%	Estimación
C1 ^a	90	11,3%	(10,7-11,9)
C2	430	54,2%	(53,9-54,6)
C3	274	34,5%	(34,1-34,9)
Total	794	100 %	
Casos faltantes	46 ^b		
Total	840		

a Conglomerado con perfil singular de riesgo

b Casos faltantes (por ausencia de alguna de las variables)

Tanto en las medias como en las medianas⁴⁷ se aprecian las diferencias sistemáticas pero muy discretas de C1 con respecto a los otros dos conglomerados de hogares. C1 tiene:

- Un número menor de habitantes por vivienda
- Tres años menos en la mediana de la escolaridad máxima por hogar.
- Valores discretamente mayores (peores)⁴⁸ en la percepción subjetiva de la condición económica.
- Valores casi idénticos en los artículos de primera necesidad.
- Valores idénticos al segundo conglomerado, pero menores que el tercero en los equipos de segunda necesidad.

⁴⁷ Se incluye también la mediana porque las densidades tienen una distribución considerablemente asimétrica.

⁴⁸ Notar que la percepción tiene una polaridad invertida, desde 1. Excelente hasta 5. Muy mala.

Tabla V.2 Caracterización de los conglomerados según variables socioeconómicas (medias, medianas y desviaciones estándar)

Conglomerados		Personas por vivienda	Máxima escolaridad	Situación económica subjetiva ^b	Equipos de primera necesidad	Equipos de segunda necesidad
C1^a (N=90)	Media	2.9	7.5	3.3	5.6	1.3
	Mediana	2.5	9.0	3.0	6.0	1.0
	DE	1.2	3.7	1.0	0.8	1.9
C2 (N=430)	Media	3.2	10.2	3.0	5.7	1.2
	Mediana	3.0	12.0	3.0	6.0	1.0
	DE	1.3	3.0	0.2	0.6	1.4
C3 (N=274)	Media	3.1	9.6	2.3	5.8	2.1
	Mediana	3.0	12.0	2.0	6.0	2.0
	DE	1.2	4.3	0.8	0.6	2.0

a Conglomerado con perfil singular de riesgo

b Esta variable tiene la polaridad invertida con respecto a las demás (los valores más bajos representan una percepción más favorable)

A estas diferencias discretas corresponden, sin embargo, notables diferencias con respecto a las densidades de hipertensión, diabetes y tabaquismo. Tanto la media como la mediana de la densidad de hipertensión son casi dos veces superiores a las de los otros conglomerados. Algo parecido ocurre con respecto a la densidad de fumadores, y también con respecto a la densidad de diabetes, aunque en este último caso sólo en relación con la media. Estas circunstancias explican que se haya identificado a C1 como un conglomerado con un perfil singular de riesgo.

Tabla V.3 Caracterización de los conglomerados según variables de salud (medias, medianas y desviaciones estándar)

Conglomerados		Densidad de hipertensos	Densidad de diabéticos	Densidad de fumadores
C1^a (N=90)	Media	0.45	0.20	0.36
	Mediana	0.42	0.00	0.33
	DE	0.37	0.30	0.40
C2 (N=430)	Media	0.26	0.07	0.23
	Mediana	0.25	0.00	0.20
	DE	0.27	0.15	0.26
C3 (N=274)	Media	0.21	0.06	0.16
	Mediana	0.20	0.00	0.00
	DE	0.23	0.14	0.23

La distribución de los conglomerados en los 12 consultorios estudiados de La Habana se presenta en la Tabla V.4. El conglomerado atípico aparece en todos los consultorios estudiados y tiene una presencia en un rango de 2.9 en el Policlínico Primero de Enero en Playa, a 23.3 en el Policlínico Enrique Betancourt Neninger de Habana del Este. Esta distribución geográfica más o menos homogénea en toda la ciudad, es típica de una arquitectura social que ha ido perfilándose en Cuba después de 1959, y que desdibujó las fronteras socioeconómicas que sí estaban bien demarcadas en la geografía urbana, física y social, de la Cuba prerrevolucionaria.

Tabla V.4 Distribución de los conglomerados según consultorios médicos.

No	Consultorio	Número de viviendas					
		C1		C2		C3	
		No	%	No	%	No	%
1	CMF No 10.1 Policlínico 1ero de enero. Playa	2	2,9%	32	47,1%	34	50,0%
2	CMF No 31 Policlínico Dr. Rafael Valdés. Cotorro	7	12,5%	38	67,9%	11	19,6%
3	CMF No 22 Policlínico Dr. Rafael Valdés. Cotorro	11	15,7%	43	61,4%	16	22,9%
4	CMF No 1 Policlínico Héroes de Girón. Cerro	8	13,3%	12	20,0%	40	66,7%
5	CMF No 6 Policlínico Héroes de Girón. Cerro	4	5,7%	50	71,4%	16	22,9%
6	CMF No 29 Policlínico René Bedia Morales. Boyero	12	19,4%	25	40,3%	25	40,3%
7	CMF No 31 Policlínico 19 de abril. Plaza de la Revolución	14	20,0%	35	50,0%	21	30,0%
8	CMF No 38 Policlínico 19 de abril. Plaza de la Revolución	3	4,3%	34	49,3%	32	46,4%
9	CMF No 13-22 Policlínico 19 de abril. Plaza de la Revolución	5	7,2%	34	49,3%	30	43,5%
10	CMF No 7 Policlínico 19 de abril. Plaza de la Revolución	5	7,1%	38	54,3%	27	38,6%
11	CMF No 6 Policlínico Cristóbal Labra. La Lisa	5	7,1%	48	68,6%	17	24,3%
12	CMF No 1 Policlínico Enríquez Betancourt Neninger. La Habana del Este	14	23,3%	41	68,3%	5	8,3%

La edad no se utilizó como variable para la búsqueda exploratoria de los conglomerados, pero exhibe algunas diferencias. La población de C1 es en promedio entre 5 y 7 años más vieja que la de los otros dos conglomerados.

Tabla V.4 Medias e intervalos de confianza para la edad según conglomerados

	Conglomerados	Media	Intervalo de confianza 95%	
			Inferior	Superior
Edad promedio	C1 ^a (N=90)	48,9	44,9	52,4
	C2 (N=430)	43,8	42,4	45,2
	C3 (N=274)	41,8	40,2	43,6

Sin embargo, las diferencias en el perfil de densidades de los conglomerados no se deben sólo a la edad, ni tampoco al resto de los marcadores socioeconómicos. Es lo que puede claramente inferirse de los resultados del ajuste de sendos modelos lineales, en los que las densidades se han utilizado como variables dependientes, las variables socioeconómicas que dieron lugar a los conglomerados como independientes, y el conglomerado propiamente como efecto principal. En vista de los resultados de la Tabla V.4, la edad promedio de las personas de la vivienda se incluyó también en los modelos como variable de control.⁴⁹ Los resultados del ajuste de estos modelos se exponen en la Tabla V.5.

Tabla V.5. Resultados del modelo lineal (estadístico F de Fisher y valor de p asociado) para evaluar las diferencias entre conglomerados, ajustadas para variables individuales.

Variable	Densidad de diabéticos		Densidad de hipertensos		Densidad de fumadores	
	F	"p"	F	"p"	F	"p"
Edad	38.19	0.00	165.17	0.00	0.87	0.34
Personas x vivienda	0.02	0.88	3.11	0.07	0.00	0.99
Máxima escolaridad del hogar	0.04	0.82	2.64	0.10	4.57	0.03
Situación económica	1.08	0.29	1.99	0.15	1.22	0.26
Equipos ^a	0.79	0.37	0.12	0.72	1.04	0.30
Conglomerado	18.47	0.00	16.71	0.00	11.73	0.00

a Incluye la suma de los equipos de primera y de segunda necesidad

Un examen detenido de la tabla anterior muestra que aún después de ajustar para la edad y para los factores socioeconómicos que permitieron la identificación de los conglomerados, hay un "efecto conglomerado" altamente significativo que ocurre a expensas de las diferencias del conglomerado C1 con respecto a los otros dos.⁵⁰ Este es un resultado de una gran trascendencia porque apunta hacia la existencia de factores no identificados que definen a este segmento de la población con un perfil singular de

⁴⁹ En los tres modelos el nivel de anclaje es la vivienda.

⁵⁰ Se realizaron contrastes "post hoc" no incluidos en el texto que así lo demuestran.

riesgo, como un grupo de gran interés, particularmente para la focalización de las políticas y programas de salud. Dada la naturaleza exploratoria de este estudio y el consiguiente diseño, sólo es posible conjeturar que en este grupo poblacional confluyen aspectos de naturaleza cultural o psicológica que podrían estar mediando el efecto de otros indicadores socioeconómicos, o tener simplemente un efecto independiente que queda ahora no explícito u oculto tras el efecto de conglomerado. Fue una buena decisión controlar para la edad que según muestra el cuadro estaría ejerciendo un efecto de confusión entre el conglomerado y las densidades, con la excepción de la densidad de fumadores, que no se modifica con la edad. Las variables incluidas exhiben, en general, bastante poco impacto sobre las densidades, en armonía con el presupuesto inicial de una escasa covariación entre la dimensión económica y la dimensión salud, al menos en nuestro medio social urbano.

Un análisis cualitativo no comparativo de este conglomerado singular reveló además las siguientes características: (12,13)

- Viviendas con problemas constructivos.
- Núcleos familiares con pocos miembros (2 a 3).
- Familias monoparentales con predominio de mujeres que tienen a su cargo menores de edad o ancianos.
- Mayoría de integrantes del núcleo familiar mayores de 60 años.
- Predominio de mestizos y negros que en conjunto arrojan un porcentaje superior al de los blancos.
- Escolaridad en entorno del 9no. grado pero con un número relativamente alto de casos por debajo de ese nivel.
- Trabajadores estatales jubilados.
- Frecuente percepción de regular a muy mala de la situación económica.

b. Aplicación del análisis de conglomerados basado en regresión (ver 4.8.2)

Como ya se expuso en el capítulo anterior, el análisis de conglomerados basado en regresión opera bajo la lógica de identificar una variable latente cualitativa que modifica el efecto de un conjunto de regresores sobre una variable de respuesta y de generar con ella una partición que minimice la varianza residual. (23,24) Los valores de dicha variable latente corresponden a segmentos o estratos poblacionales. Esta lógica es bastante armónica con el principio de que el contexto, independientemente de que pueda él mismo influir o no sobre la salud, es capaz de modificar radicalmente el efecto de otros factores, por ejemplo, de los factores de riesgo individual o de otros factores contextuales. (12)

Los resultados de esta línea de búsqueda inductiva de conglomerados con un significado desde el punto de vista de las desigualdades sociales en salud se exponen en lo que sigue. La Tabla V.6 muestra tres modelos radicalmente diferentes en la estructura de sus parámetros para explicar un constructo de salud que se obtuvo como el primer factor en un análisis factorial clásico sobre las tres densidades (de diabetes, hipertensión y tabaquismo). Este eje factorial

explica más de 80% de la variabilidad común de las tres variables de salud.⁵¹ Es obvio que de un conglomerado a otro cambian la magnitud y el signo, prácticamente de todos los parámetros. Por ejemplo, la estimación del parámetro que corresponde al intercepto se mueve en un rango de -3.72 en el conglomerado 3 a 1.78 en el 2; y el de la condición económica subjetiva, de -0.42 en el conglomerado 3 hasta -0.10 en el 2. Hay una clara modificación del patrón de asociación entre la dimensión socioeconómica y la dimensión salud. Se omiten los errores estándar de los estimadores y los valores de p asociados que son irrelevantes en relación con el argumento básico de esta exposición.

Tabla V.6 Parámetros estimados de los modelos de regresión para un indicador sintético de las variables de salud según conglomerados

Variable	Conglomerado 1	Conglomerado 2	Conglomerado 3
Intercepto	-1.61	1.78	-3.72
Hab x viv.	0.45	-0.11	0.38
Escolaridad	0.03	-0.006	0.05
Equipos de 1ra	-0.09	0.008	-0.09
Equipos de 2da	0.16	-0.002	0.19
Cond_econ	-0.36	-0.10	-0.42

Los resultados que se exponen en la Tabla V.7 permiten constatar que prácticamente no hay diferencias entre los grupos con respecto a los valores promedio de los indicadores socioeconómicos. Por ejemplo, las medianas de todas las variables son idénticas en los tres conglomerados.

⁵¹ En el texto se omiten los resultados de este análisis factorial, que tienen una utilidad meramente instrumental en la exposición.

Tabla V.7 Caracterización de los conglomerados según variables socioeconómicas (medias, medianas y desviaciones estándar).

Conglomerados		Personas por vivienda	Máxima escolaridad	Situación económica subjetiva	Equipos de primera necesidad	Equipos de segunda necesidad
C1 (N=294)	Media	2.96	9.81	2.75	5.71	1.52
	Mediana	3.00	12.00	3.00	6.00	1.00
	DE	1.13	3.39	0.73	0.65	1.55
C2 (N=358)	Media	2.96	9.32	2.76	5.64	1.53
	Mediana	3.00	12.00	3.00	6.00	1.00
	DE	1.36	4.01	0.70	0.72	1.84
C3 (N=142)	Media	3.09	9.44	2.78	5.72	1.42
	Mediana	3.00	12.00	3.00	6.00	1.00
	DE	1.50	3.89	0.70	0.52	1.76

Sin embargo, si se examinan los indicadores de salud, se advierten de inmediato grandes diferencias: el conglomerado 3 (C3) tiene medias y medianas mucho más altas de las densidades de diabetes e hipertensión y moderadamente más alta de tabaquismo. El conglomerado 1 es un conglomerado intermedio en relación con el perfil de los tres factores de riesgo elegidos, y el conglomerado 2 (el que contiene mayor cantidad de viviendas) un conglomerado con densidades notablemente inferiores al promedio.

Tabla V.8 Caracterización de los conglomerados según variables de salud (medias, medianas y desviaciones estándar)

Conglomerados		Densidad de diabéticos	Densidad de hipertensos	Densidad de fumadores
C1 (N=294)	Media	0.06	0.35	0.22
	Mediana	0.00	0.33	0.17
	DE	0.13	0.18	0.26
C2 (N=358)	Media	0.003	0.04	0.18
	Mediana	0.00	0.00	0.00
	DE	0.02	0.09	0.26
C3 (N=142)	Media	0.31	0.63	0.30
	Mediana	0.33	0.50	0.33
	DE	0.25	0.27	0.30

Carece de sentido intentar una validación cruzada a partir de los resultados de dos enfoques que se basan en presupuestos y filosofías analíticas completamente diferentes: el primero en la identificación de los componentes de mezclas de subpoblaciones con distribuciones de probabilidad diferentes; el segundo en la búsqueda de una variable latente categórica que modifica la asociación entre un conjunto de regresores y una variable de respuesta.

Cabe sin embargo, hacer notar, que ambas apuntan hacia el hecho particularmente interesante de la influencia del contexto sobre la salud, que se hace sentir de dos modos diferentes: mediante un efecto directo, y mediante una modificación del efecto de otros factores. Lo primero se puso de manifiesto en el muy significativo efecto del conglomerado sobre las tres variables de salud, aun después de remover la influencia de la edad y de los factores socioeconómicos considerados. Lo segundo, las drásticas diferencias entre conglomerados con respecto al patrón de asociación entre dichos factores socioeconómicos y las variables de salud.

En un orden más práctico, ambos enfoques permiten identificar grupos con respecto a los cuales emergen claras desigualdades en los indicadores de salud, que en este estudio se restringen a tres conocidos factores de riesgo para las enfermedades consecuentes de la aterosclerosis. En el ambiente urbano de La Habana, las estratificaciones de acuerdo a estas variables latentes, permiten desvelar desigualdades que no se manifiestan, o lo hacen en una medida mucho menor, en los factores de clasificación clásicos para la medición de desigualdades sociales en salud.

Es imperativo avanzar en la identificación y caracterización de estas magnitudes latentes de desigualdad y en criterios e indicadores objetivos para medirlas, y con ello estar en condiciones de formular mejores criterios de focalización de los programas y políticas de salud.

Referencias

1. Zabala MC. Familia y pobreza en Cuba. Estudio de casos. Publicaciones Acuario, Centro Félix Varela. La Habana, 2010.
2. Espina M. Desigualdad y política social en Cuba hoy. Conference "The Future of Social Justice in Cuba". Panel "Social Justice in Cuba: Assessment of Current Situation". Bellagio Conference Center May 27-29, 2008.
3. Espina M. Desarrollo, desigualdad y políticas sociales. Acercamientos desde una perspectiva compleja. Publicaciones Acuario. Centro Félix Varela. La Habana, 2010
4. Espina M. Políticas de atención a la pobreza y la desigualdad: examinando el rol del estado en la experiencia cubana. CLACSO, Consejo Latinoamericano de Ciencias Sociales, Buenos Aires. 2008.
5. Castro, F. Discurso pronunciado en la graduación de los alumnos de la Escuela de Trabajadores Sociales efectuada en el teatro Carlos Marx, el 15 de febrero de 2001.

6. Castro, F. Discurso pronunciado en el acto de graduación del segundo curso trabajadores sociales, efectuado en el teatro Astral, Ciudad de La Habana, el 31 de diciembre de 2001.
7. Dilla H. La reestructuración espacial en Cuba. Cuadernos del CENDES. Año 25. N° 68, mayo-agosto 2008.
8. Corral A, Castañeda I, Bárzaga E, Torres M, Santana MC. Determinantes sociales en la diferencial de mortalidad entre mujeres y hombres en Cuba. Rev Cub Salud Pública 2010; 36: 66-77.
9. Iñiguez L. Territorio y contextos en la salud de la población. Rev Cub Salud Pública 2008; 34. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-34662008000100006&lng=es.
10. Méndez E, Lloret MC. Índice de Desarrollo Humano a nivel territorial en Cuba. Período 1985- 2001. Rev Cub Salud Pública 2005; 31. Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-34662005000200003&lng=es.
11. Álvarez AG, García A, Bonet M. Pautas conceptuales y metodológicas para explicar los determinantes de los niveles de salud en Cuba. Rev Cub Salud Pública 2007; 33: (2). Disponible en: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S0864-34662007000200013&lng=es.
12. Díaz-Perera G. Determinantes contextuales de las enfermedades asociadas a la aterosclerosis y sus factores de riesgo. Tesis presentada en opción a la categoría de Doctora en Ciencias de la Salud. La Habana. 2011.
13. Díaz-Perera G, Bacallao J, Alemañy E. Subpoblaciones con perfiles epidemiológicos y de riesgo singulares en La Habana, Cuba: diabetes, hipertensión y tabaquismo. Rev Panam Salud Pública 2012; 32: 9-14.
14. Borowy I. Similar but different: health and economic crisis in 1990's Cuba and Russia. Soc Sci Med 2011; 72: 1489-1498.
15. Díaz-Perera G, Bacallao J, Alemañy E. Relación entre la dimensión socioeconómica y la dimensión salud en familias cubanas. Rev Cub Salud Pública 2012, 38: 403-413.
16. Ferriol, A "Política social cubana: situación y transformaciones" en Temas (No. 11) . La Habana. 1998.
17. Zabala, MC. "Situación de la pobreza en el Caribe: actualidad y perspectivas. Cuba en el contexto caribeño", Seminario Internacional Estrategias de Reducción de la Pobreza en el Caribe: los Actores Externos y su Impacto, CLACSO-CROP, La Habana. 2002.
18. Zabala, MC. "Los estudios cualitativos de la pobreza en Cuba", Taller XX Aniversario del Centro de Investigaciones Psicológicas y Sociológicas, La Habana. 2003
19. Rodríguez, P et al. "¿Pobreza, marginalidad o exclusión?: un estudio sobre el barrio Alturas del Mirador", La Habana, Centro de Antropología, Informe Preliminar de Investigación. 2004.
20. Grundy G, Holt E. The socioeconomic status of older adults. How should we measure it in studies of health inequalities. J Epidemiol Community Health 2001; 55: 895-904.
21. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position. J Epidemiol Community Health 2006; 60: 7-12.

22. Mathews KA, Gallo LC. Psychological perspectives on pathways linking socioeconomic status and physical health. *Annu Rev Psychol* 2011; 62: 501-530.
23. Zhong S, Ghosh J. A unified framework for model-based clustering. *J Machine Learning Research* 2003; 4: 1001-1037.
24. Zhu X, Goldberg A. Kernel regression with ordered preferences. In. *Proceedings of the Twenty-Second AAAI Conference on Artificial Intelligence*. 2007

CAPÍTULO VI

EPIDAT 4.1. EJEMPLOS RESUELTOS. UNA APLICACIÓN A LA MEDICIÓN DE DESIGUALDADES SANITARIAS EN CUBA

Resumen:

En este capítulo se presenta la versión 4.1 (en desarrollo) del programa Epidat, se exponen las facilidades que brinda el módulo de desigualdades y se ilustran con la solución de algunos problemas de medición. Casi todos los ejemplos se basan en datos ficticios de prueba (algunos de los cuales pueden encontrarse en el expediente informático del programa). Se incluye también un ejemplo con estadísticas sanitarias oficiales de Cuba, con lo cual se satisface el doble propósito de presentar el recurso de cómputo y de incursionar preliminarmente en la medición de desigualdades en el ámbito nacional con el empleo de esta herramienta. Este último ejemplo se comenta y discute.

Epidat (1) es un programa de libre distribución para el análisis epidemiológico de datos. Su versión 4.0, y la más reciente al momento de redactar este documento (la 4.1), contienen un módulo dedicado a la medición de desigualdades. Este ha sido concebido especialmente para el cálculo de los índices en estratos constituidos por unidades geodemográficas (países, regiones, provincias, departamentos, municipios, comunidades y otras unidades político-administrativas), con el propósito de complementar de ese modo la generación de evidencia epidemiológica que hacen los países como parte de sus estadísticas sanitarias. El recuadro que sigue es la pantalla inicial de presentación, que incluye el modo de citarlo y la vía de acceso

Acerca de

E Epidat: Programa para análisis epidemiológico de datos - Versión 4.1 - Marzo 2013

Historial

Instituciones auspiciadoras

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Dirección Xeral de Innovación e Xestión da Saúde Pública

Organización Panamericana de la Salud
Organización Mundial de la Salud

UNIVERSIDAD CES

Créditos

Créditos

Contacto

Dirección Xeral de Innovación e Xestión da Saúde Pública
Edificio administrativo San Lázaro, s/n
15703, Santiago de Compostela
A Coruña, España

Tt: (34) 881-54 67 59
Fax: (34) 881-54 29 70
E-mail: suporte_epidat@sergas.es
Web: <http://dxsp.sergas.es>

Cómo citar Epidat

Epidat: programa para análisis epidemiológico de datos Versión 4.1, marzo 2013. Consellería de Sanidade, Xunta de Galicia, España; Organización Panamericana de la Salud (OPS-OMS); Universidad CES, Colombia. Disponible en: <http://dxsp.sergas.es>

Licencia

Reconocimiento - NoComercial - SinObraDerivada (by-nc-nd): No se permite un uso comercial de la obra original ni la generación de obras derivadas.

Aceptar

En los tres pequeños recuadros siguientes se exponen tres ventanas funcionales del programa que muestran respectivamente el menú vigente a la fecha, el submenú correspondiente a la medición de las desigualdades sociales en salud, y para información del lector, un listado de los módulos aún pendientes, pero en vías de ser incorporados a la versión corriente.

Módulos	Herramientas	Ventana	Ayuda
Análisis descriptivo			→
Muestreo			→
Concordancia y consistencia			→
Ajuste de tasas			→
Demografía			→
Estimación de la mortalidad atribuida			→
Distribuciones de probabilidad			→
Análisis Bayesiano			→
Medición de desigualdades en salud			→

Índices basados en rangos	→
Índices basados en el concepto de disparidad	→
Índices basados en el concepto de desproporcionalidad	→
Índices basados en el modelo de regresión	→
Índices basados en el concepto de entropía	→

Enero 2012
El contenido de Epidat 4.0 se estructuró en 19 módulos, 9 de los cuales ya están implementados en el programa. Los módulos que están pendientes y que se irán incorporando progresivamente son:
2. Depuración e imputación de datos
4. Inferencia sobre parámetros
5. Tablas de contingencia
7. Pruebas diagnósticas
11. Regresión logística
12. Análisis de supervivencia
15. Metanálisis
16. Vigilancia en Salud Pública
17. Índices de desarrollo o privación
19. Métodos de evaluación económica

A continuación se presentan y resuelven varios ejemplos, casi todos ellos disponibles como parte de la documentación de libre acceso del programa.

Ejemplo 1. (Indicadores basados en comparaciones dos a dos⁵²)

Datos: (10 provincias, cantidad de nacidos vivos, orden de las provincias según necesidades básicas insatisfechas (NBI) y tasas de mortalidad infantil (TMI) como variable de salud.

Tabla VI.1

PROVINCIA	NV	NBI	TMI
A	2 300	14.1	26
B	2 500	14.8	39
C	1 000	15.4	50
D	12 000	15.5	57
E	550	15.6	58
F	10 000	15.8	59
G	2 400	17	69
H	1 200	17.4	73
I	250	17.6	78
J	450	27.9	109
TOTALES	32 650		1834950
	TMI TOTAL=	56.20	
	Cociente de tasas=	4.19	
	Diferencia de tasas=	83	
	RAP=	30.20	
	RAP relativo (%)=	53.74	

Sombreados y en negrita se incluyen los cálculos realizados en una hoja EXCEL que pueden verificarse con las salidas de Epidat que siguen.

Salidas de Epidat

[1] Índices basados en rangos:

Entrada automática:

Archivo de trabajo: C:\Instalaciones\Epidat Abr 2011\Epidat 4.0\Epidat 4.0\Ejemplos\Medición de desigualdades en salud\RANGOS.xls⁵³

Tabla: Tabla3

VARIABLES:

Identificadores unidades geodemográficas: PROVINCIA

Poblaciones: NV

Variable socioeconómica: NBI

VARIABLES DE SALUD: TMI

Datos:

⁵² Indicadores basados en rango, de acuerdo a la terminología de Epidat)

⁵³ Se refiere a la ubicación de la base de datos en el ordenador en el que se realiza el cómputo

Número de unidades geodemográficas: 10
 Número de variables de salud: 1
 Sentido de la variable socioeconómica: Negativa⁵⁴
 Sentido de la variable de salud: Negativa
 Ordenar por: Variable socioeconómica

Resultados:

Índice	TMI
Cociente de tasas extremas	4,19
Diferencia de tasas extremas	83
Riesgo atribuible poblacional	30,20
Riesgo atribuible poblacional relativo (%)	53,73

La tasa de mortalidad infantil en el grupo con más alto porcentaje de NBI es 4.19 veces mayor que la del grupo con el porcentaje más bajo de NBI. La diferencia absoluta entre dichos grupos es de 83 niños fallecidos por cada 1000 nacidos vivos. Si toda la población se comportase como el grupo mejor en términos del indicador socioeconómico habría 30,2 muertes menos por cada mil nacidos vivos, lo que representa una reducción de 53,73%.

Ejemplo 2: Índices basados en el concepto de disparidad

Datos: Cuatro variantes de las tasas mortalidad infantil (TMI, TMI_1, TMI_2 y TMI_3) en 10 provincias. Población en miles y tamaños relativos de las poblaciones.

Tabla VI.2

PROVINCIA	NV	W_i	TMI	TMI_1	TMI_2	TMI_3
A	2 300	0.07	26	16	23.4	24.7
B	2 500	0.08	39	29	35.1	37.05
C	1 000	0.03	50	40	45	42.5
D	12 000	0.37	26	16	23.4	24.7
E	550	0.02	27	17	24.3	25.65
F	10 000	0.31	28	18	25.2	26.6
G	2 400	0.07	61	51	54.9	51.85
H	1 200	0.04	63	53	56.7	53.55
I	250	0.01	68	58	61.2	57.8
J	450	0.01	69	59	62.1	58.65
	32 650					
	t_{total}		33.21	23.21	29.89	30.57
	$t_{ref}=tasa\ mínima$		26	16	23.4	24.7

⁵⁴ En casi todos los módulos el usuario debe especificar la polaridad de las variables socioeconómica y de salud, es decir, si se trata de una variable con significado negativo (como la mortalidad y la morbilidad) o positivo (la cobertura de inmunizaciones)

Antes de analizar las salidas, es importante notar que la mayor reducción de las tasas se consigue en el caso 1 (TMI_1)

Epidat produce las salidas siguientes:

Archivo de trabajo: C:\Instalaciones\Epidat Abr 2011\Epidat 4.0\Epidat 4.0\Ejemplos\Medición de desigualdades en salud\DISPARIDAD.xls

Tabla: Datos

Variables:

Identificadores unidades geodemográficas: PROVINCIA

Poblaciones: NV

Variables de salud: TMI, TMI_1, TMI_2, TMI_3

Datos:

Número de unidades geodemográficas: 10

Número de variables de salud: 4

Valor de referencia: Mínimo (26; 16; 23,4; 24,7)

Resultados:

Tabla VI.3

Índice	TMI	TMI_1	TMI_2	TMI_3
Índice de Pearcy-Keppel	0,76	1,23	0,76	0,63
Variante del índice de Pearcy-Keppel ^(a)	0.28	0.45	0.28	0.24
Varianza entre grupos	159,24	159,24	128,98	97,14
Variante de la varianza entre grupos ^(b)	4,80	6,86	4,33	3,18

(a) Índice de Pearcy-Keppel ponderado. (b) Varianza entre grupos relativa.

Es interesante constatar que el escenario que representa una mayor reducción de la tasa poblacional, implica también un incremento de la desigualdad con respecto a todos los índices con la excepción de la varianza entre grupos que se mantiene constante, ya que la varianza no es sensible a cambios de posición, pero sí a cambios de escala. Sólo en la situación 3 (TMI_3) disminuyen todos los índices de desigualdad; sin embargo, la reducción en la tasa global es inferior a 3 puntos porcentuales.

Minujin y Delamónica, 2003 (2) califican el desempeño de las políticas en términos de los dos ejes que la OMS considera esenciales en la formulación de las metas programáticas en

salud: la salud promedio de la población y las brechas sociales relativas. El resultado ideal es aquel que combina una mejoría de la salud poblacional con una reducción de las brechas. No es esa, sin embargo, la tendencia que exhibe la mayoría de los países del mundo. Las inequidades en salud se hacen sentir en todos los países, más allá de su nivel de desarrollo. Aun en países que exhiben mejoría de sus indicadores poblacionales de salud, las brechas sanitarias entre algunos subgrupos poblacionales se han incrementado o se han mantenido, pese a lo cual importantes iniciativas de desarrollo en salud como los Objetivos de Desarrollo del Milenio no incluyen mediciones y menos aun análisis de la equidad sanitaria.

Ejemplo 3: Índices basados en el modelo de regresión

Datos: Población (en miles de habitantes), porcentaje poblacional con necesidades básicas insatisfechas (NBI) y tasa de mortalidad infantil (TMI) por mil nacidos vivos en provincias argentinas.

Tabla VI.4

PROVINCIA	POBLACIÓN	NBI	TMI
Formosa	13 273	33.6	25.5
Chaco	24 215	33	26.7
Salta	27 492	31.6	18.9
Santiago del Estero	16 254	31.3	12.4
Jujuy	13 768	28.8	20.6
Corrientes	22 094	28.5	23.8
Misiones	23 846	27.1	22.3
Tucumán	29 921	23.9	24.3
Catamarca	8 040	21.5	20.6
La Rioja	6 303	20.4	20.8
Río Negro	11 144	17.9	14.8
Entre Ríos	22 088	17.6	16.5
San Juan	13 070	17.4	20.2
Neuquén	9 712	17	11.7
Buenos Aires	247 738	15.8	15.8
San Luis	8 365	15.6	16.5
Chubut	8 000	15.5	17.8
Mendoza	27 804	15.4	12.3
Santa Fe	52 690	14.8	14
Tierra del Fuego	2 199	14.1	9.1
Córdoba	52 453	13	14.8
Santa Cruz	4183	10.4	17.2
La Pampa	5 272	10.3	13.1
Ciudad de B. Aires	41 509	7.8	10

Las salidas analíticas con Epidat 4.0 son las siguientes:

Índices de desigualdad:

Archivo de trabajo: D:\trabajo\Tesis\Tesis_active\carpeta de trabajo\working version\ejemplos en EXCEL para EPIDAT\regression.xls

Tabla: Hoja1

VARIABLES:

Identificadores unidades geodemográficas: provincia

Poblaciones: popnv

Variable socioeconómica: NBI

VARIABLES DE SALUD: tmi

DATOS:

Número de unidades geodemográficas: 24
 Número de variables de salud: 1
 Sentido de la variable socioeconómica: Negativa
 Sentido de la variable de salud: Negativa
 Transformación logarítmica de la variable de salud: No

Resultados:

Tabla VI.5

Índice	tmi
Índice de desigualdad de la pendiente	-10,17
Índice relativo de desigualdad de Pamuk	0,61
Índice relativo de desigualdad de Kunst y Mackenbach	1,86
Índice relativo de desigualdad acotado ⁵⁵	1,80

Entre los dos extremos teóricos de la jerarquía social hay una diferencia de 10.17 fallecimientos por cada 1000 nacidos vivos. El extremo con mayor privación socioeconómica en términos del porcentaje de NBI tiene una TMI 1.86 veces más alta. El cociente esperado de estas tasas extremas, de acuerdo al modelo teórico basado en la regresión es de 1.80.

Ejemplo 4. Índices basados en el concepto de entropía

Datos: Población y cantidad de médicos en los 25 departamentos peruanos. La cuarta columna es la tasa de médicos por cada 100 mil habitantes.

Tabla VI.6

DEPARTAMENTO	POBLACIÓN	MÉDICOS	T_MEDICOS
Amazonas	391 078	124	31.71

⁵⁵ Se llama así en Epidat 4.0 al índice relativo de desigualdad corregido basado en el cociente de tasas

Ancash	1 045 921	653	62.43
Apurímac	417 280	119	28.52
Arequipa	1 035 773	1451	140.09
Ayacucho	519 315	222	42.75
Cajamarca	1 377 297	412	29.91
Callao	736 243	1488	202.11
Cusco	1 131 061	559	49.42
Huancavelica	421 638	115	27.27
Huánuco	747 263	266	35.60
Ica	628 684	738	117.39
Junín	1 161 581	660	56.82
La Libertad	1 415 512	1128	79.69
Lambayeque	1 050 280	791	75.31
Lima	7 204 000	13141	182.41
Loreto	839 748	331	39.42
Madre de Dios	79 172	73	92.20
Moquegua	142 475	160	112.30
Pasco	245 651	138	56.18
Piura	1 506 716	939	62.32
Puno	1 171 838	387	33.03
San Martín	692 408	272	39.28
Tacna	261 336	263	100.64
Tumbes	83 609	122	145.92
Ucayali	394 889	156	39.50

Epdat 4.0 arroja los siguientes resultados:

Índices basados en el concepto de entropía:

Entrada automática:

Archivo de trabajo: D:\trabajo\Tesis\Tesis_active\carpeta de trajín\working version\ejemplos en EXCEL para EPIDAT\ENTROPIA.xls

Variables:

Identificadores unidades geodemográficas: DEPARTAMENTO

Poblaciones: POBLACIÓN

Variables de salud: MEDICOS

Datos:

Número de unidades geodemográficas: 25
 Número de variables de salud: 1

Resultados:

Índice	MEDICOS
Índice de Hoover (o de disimilitudes)	0,29
Índice de Theil	0,20

Índices estandarizados a escala [0,1]

Índice	MEDICOS
Índice de Hoover (o de disimilitudes)	0,25
Índice de Theil	0,18

Transformación de equivalencia de entropía

Índice	MEDICOS
Índice de Hoover (o de disimilitudes)	0,31
Índice de Theil	0,34

El índice de Hoover, sin estandarización, ni conversión a equivalencia de entropía, tiene una interpretación directa: para que haya una distribución completamente equitativa de los médicos en el país, habría que redistribuir 29% de ese recurso humano.

El índice de Theil, en su equivalencia de entropía, implica que la desigualdad entre provincias peruanas en cuanto a la distribución de los médicos, equivale a la de un país con dos territorios, uno de los cuales, con el 34% de la población posee el 66% de los médicos, mientras que el otro, con 66% de la población, tiene sólo el 34% de los médicos. Esto implica un potencial redistribución de 16%.

Ejemplo 5. Índices basados en el concepto de desproporcionalidad

Datos: Necesidades básicas insatisfechas y tasas de mortalidad infantil en 10 provincias

Tabla VI.7

PROVINCIA	NV	NBI	TMI
A	2 300	60.3	69
B	2 500	60.1	68
C	1 000	51.9	63
D	12 000	51.5	61
E	550	46.6	58
F	10 000	41.2	57
G	2 400	30	56
H	1 200	21.8	50
I	250	15.7	39
J	450	13.4	26
	32 650		59.5

Los resultados que arroja Epidat son los siguientes:

Índice de concentración:

Archivo de trabajo: D:\trabajo\Tesis\Tesis_active\carpeta de trabajo\working version\ejemplos en EXCEL para EPIDAT\DESproporcionalidad.xls

Tabla: Tabla8

Variables:

Identificadores unidades geodemográficas: PROVINCIA

Poblaciones: NV

Variable socioeconómica: NBI

Variables de salud: TMI

Datos:

Número de unidades geodemográficas: 10
Número de variables de salud: 1
Sentido de la variable socioeconómica: Negativa
Sentido de la variable de salud: Negativa
Suavizar el índice y curva de concentración: No

El usuario del sistema debe especificar la polaridad de la variable socioeconómica y la variable de salud, porque de ello depende el ordenamiento que se realiza para construir la curva de concentración.

Resultados:

	TMI
Índice de concentración	-0,0475

El Índice de Concentración negativo implica que la curva de concentración se encuentra por encima de la recta de plena equidad, o lo que es lo mismo, que las muertes infantiles tienden a concentrarse en los más pobres, en el sentido de las necesidades básicas insatisfechas. Sin embargo, a partir de la relación (ver Anexo)

$$I_C = \frac{2IDP \text{ var}(r)}{t_{total}} = 2IRD_{PAMUK} \text{ var}(r)$$

... se obtiene un índice de desigualdad de la pendiente (IDP) de -18.42 cuya interpretación es inmediata, de modo similar a como ya se discutió en el ejemplo 3.

Ejemplo 6: Epidat aplicado a la medición de la desigualdad con estadísticas sanitarias del ámbito nacional.

Esta sexta y última aplica de Epidat tiene especial interés porque se realiza con estadísticas sanitarias cubanas de recursos humanos, servicios y mortalidad específica. Los datos íntegros para este análisis están contenidos en las dos tablas EXCEL que se insertan a continuación:

EJEMPLO 6

TABLA VI.8

Provincias	Población total	Médicos	Médicos de familia	Estomatólogos	Enfermeras	Camas médicas	Camas sociales
P. del Río	591987	3852	705	642	6022	2168	282
Artemisa	505836	2324	451	407	3788	1197	345
La Habana	2132965	20540	2355	3032	16224	13760	2875
Mayabeque	380725	1686	394	374	2404	773	505
Matanzas	694532	4079	725	784	4685	1965	732
Villa Clara	799028	5551	465	831	7023	2841	1370
Cienfuegos	408007	2720	1036	456	2904	1116	385
S. Spíritus	465890	3661	647	527	3852	1838	652
Ciego	425492	2913	499	434	3553	1223	635
Camagüey	779622	5114	826	849	6487	3056	1087
Las Tunas	539039	3042	597	503	4312	2007	638
Holguín	1037833	6133	1053	927	8223	3680	774
Granma	837285	5311	1311	934	10300	2817	422
Santiago	1048416	7987	1429	1403	9580	4845	952
Guantánamo	511447	3662	774	551	6264	1955	359
I Juventud	86439	547	100	139	803	365	64
Cuba	11244543	79122	13367	12793	96424	45606	12077

Fuente: Oficina Nacional de Estadísticas e Información (ONEI). Anuario Estadístico de Cuba 2011. Salud Pública y Asistencia Social. Edición 2012.

EJEMPLO 6

TABLA VI.9

Provincias	Población total	m materna	u5_mr	enf_cor	neoplas	ecvasc	IDH
P. del Río	591987	28.9	4.8	106.1	98.5	44.4	0.358
Habana Interior	886561	79.5	6.6	98.2	111.9	38.4	0.565
La Habana	2132965	59.2	5.6	121.9	121	43.1	0.677
Matanzas	694532	25.5	6.2	122.5	116.3	38.1	0.536
Villa Clara	799028	11.9	6.7	90.6	99.6	30.7	0.501
Cienfuegos	408007	43.2	6.1	102.6	115.2	34.6	0.638
S. Spíritus	465890	19	7.2	87.3	109.6	37.3	0.465
Ciego	425492	55.9	5.5	94.2	105.4	39	0.553
Camagüey	779622	51.9	6.6	96.4	121.5	41.3	0.453
Las Tunas	539039	58.7	4.7	97.4	117.6	30	0.303
Holguín	1037833	31	4.7	89.8	106.3	35.2	0.371
Granma	837285	17.6	5	103.7	112.5	40.9	0.347
Santiago	1048416	61.4	7	98.4	117.1	51.5	0.347
Guantánamo	511447	25.2	7.4	95.4	116.2	44.4	0.273
I Juventud	86439	8.7	9.6	108.4	112	34.9	0.589

Fuentes:

- Oficina Nacional de Estadísticas e Información (ONEI). Anuario Estadístico de Cuba 2011. Salud Pública y Asistencia Social. Edición 2012.
- Oficina Nacional de Estadísticas e Información (ONEI). Anuario Estadístico de Cuba 2011. Población. Edición 2012.
- Ministerio de Salud Pública. Dirección Nacional de Registros Médicos y Estadísticas de Salud. Anuario Estadístico de Salud (DNRMES). 2011.

La población total estimada⁵⁶ al 30 de junio de 2011 se obtuvo del Anuario Estadístico de Salud. Los datos sobre recursos humanos en salud, y las tasas de mortalidad materna, en el menor de 5 años y por causas seleccionadas, del Anuario Estadístico Nacional de la ONEI. El índice de desarrollo humano (IDH) provincial, de un artículo publicado en la Revista Cubana de Salud Pública. (3) Los datos de la ONEI y de la DNRMES del MINSAP temporalmente compatibles, pero las estimaciones del IDH se refieren al promedio de una serie temporal de valores estimados entre 1985 y 2001. En esas fechas no existían aún las provincias de Artemisa y Mayabeque. Por otro lado, las fuentes consultadas contenían datos faltantes de las tasas de mortalidad materna en la provincia de Ciego de Ávila y en el municipio especial de Isla de la Juventud. Fue necesario tomar las siguientes providencias:

- Se estimaron los datos faltantes utilizando un procedimiento de imputación simple refinado basado en el empleo de los vecinos más próximos. (4)
- Para el análisis de los datos de mortalidad se obtuvo una estimación ponderada de las provincias de Artemisa y Mayabeque que debe proporcionar una aproximación razonable a la antigua provincia de La Habana (Habana Interior en la tabla). Esto se hizo con el fin de poder utilizar el valor del IDH obtenido del estudio ya referido. (3)
- Por último, se calculó el coeficiente de correlación de Spearman entre las estimaciones del IDH con el fin de constatar si sería posible utilizar el valor del IDH

⁵⁶ Se trata de proyecciones poblacionales anteriores a la realización del Censo Nacional de Población y Viviendas de 2012.

promedio como un criterio fiable de estratificación para el cálculo de los índices de desigualdad.

Las dos tablas que siguen muestran las altas autocorrelaciones del IDH en el tiempo, suficientes para justificar el empleo del promedio como insumo para los análisis posteriores.

Tablas VI.10 y VI.11. Autocorrelaciones temporales del IDH
Matriz de correlación (Spearman):

Variables	1998	1999	2000	2001	Promedio
1998	1	0.611	0.625	0.435	0.654
1999	0.611	1	0.768	0.411	0.652
2000	0.625	0.768	1	0.460	0.847
2001	0.435	0.411	0.460	1	0.675
Promedio	0.654	0.652	0.847	0.675	1

Valores de "p":

Variables	1998	1999	2000	2001	Promedio
1998	0	0.017	0.015	0.106	0.010
1999	0.017	0	0.001	0.127	0.010
2000	0.015	0.001	0	0.085	< 0.0001
2001	0.106	0.127	0.085	0	0.007
Promedio	0.010	0.010	< 0.0001	0.007	0

A continuación se resume todo el análisis de desigualdad, mediante la aplicación de la batería de índices contenidos en Epidat.

Tabla VI.12

Índices de Hoover para los indicadores de recursos humanos y servicios. Valor máximo de la diferencia entre valor real y valor esperado bajo la hipótesis de plena igualdad.

Indicador	Médicos	Médicos de familia	Estomatólogos	Enfermeras	Camas médicas	Camas sociales
Índice de Hoover	0.083	0.098	0.067	0.070	0.125	0.152
Diferencia máxima	7%	4.1%	4.7%	3.2%	11.2%	4.8%
Provincia	La Habana	Cienfuegos	La Habana	Granma	La Habana	La Habana

Lo más interesante de la tabla anterior es, en primer lugar, el valor relativamente bajo del Índice de Hoover para todos los indicadores. El más alto -el que corresponde a las camas sociales- implica que se necesitaría redistribuir el 15% de este recurso en el país para conseguir la completa equidad (idéntica proporción de camas que de habitantes). En

segundo lugar, que la máxima contribución a la desigualdad (máxima diferencia entre el valor real y el valor esperado bajo el supuesto de plena equidad) corresponde en casi todos los casos a la capital, con la excepción del número de médicos de familia y del número de personal de enfermería: La Habana dispone de más médicos, más estomatólogos y más camas (para atención médica y para atención social) que lo que cabría esperar bajo la hipótesis de equidad distributiva de los recursos humanos y de los servicios.

En las tablas que siguen, se aplican los índices de desigualdad a las provincias del país ordenadas según el IDH promedio. También se aplica el Índice de Theil en su transformación de equivalencia de entropía.

Tabla VI.13 Índices de desigualdad basados en regresión para 5 indicadores específicos de mortalidad

Índices de desigualdad basados en regresión	Indicadores de mortalidad específica				
	Mortalidad materna	Mortalidad en el menor de 5 años	Enfermedades del corazón	Neoplasias	Enfermedades cerebro-vasculares
IDP	7.82	1.15	14.35	0.04	-5.41
IRD	0.18	0.19	0.14	0.00	-0.14
IRD _{K-M}	1.18	1.17	1.13	1.00	1.15
IRD _{CORREGIDO}	0.84	0.85	0.88	1.00	1.13

Tabla VI.14 Índices basados en rangos para 5 indicadores específicos de mortalidad

Índices de desigualdad basados en rangos	Indicadores de mortalidad específica				
	Mortalidad materna	Mortalidad en el menor de 5 años	Enfermedades del corazón	Neoplasias	Enfermedades cerebro-vasculares
CT	0.43	1.32	0.78	0.96	1.03
DT	-34	1.8	-26.5	-4.8	1.3
RAP	-15.77	0.37	-18.83	-7.76	-3.07
RAP relativo	-36.31	6.27	-18.27	-6.85	-7.66

Tabla VI.15 Índice y curva de concentración, Índice de efecto y equivalencia de entropía del Índice de Theil para 5 indicadores de mortalidad específica

Índices de desigualdad	Indicadores de mortalidad específica				
	Mortalidad materna	Mortalidad en el menor de 5 años	Enfermedades del corazón	Neoplasias	Enfermedades cerebro-vasculares
Concentración	0.09	0.00	0.04	0.01	-0.01
Theil (equivalencia de entropía)	0.38	0.46	0.47	0.48	0.42
Índice de efecto	57.51	0.41	56.74	17.25	-2.39

El análisis en bloque de las Tablas V1.12 a V1.14 revela lo siguiente: sólo las enfermedades cerebro-vasculares exhiben un IDP -y por consiguiente- un Índice de Concentración negativo, aunque muy pequeño. Consistentemente con esos resultados, sólo el CT que corresponde a ese indicador es -aunque ligeramente- mayor que la unidad. El IRD_{CORREGIDO} que como se ha expuesto es un cociente corregido de tasas extremas, también es

ligeramente mayor que la unidad. Tres de las 5 DT son negativas, justamente las que corresponden a los CT inferiores a la unidad. La relación entre la dimensión económica y la dimensión salud tiene un signo contrario al que exhiben la mayoría de los países, lo cual se pone de manifiesto, tanto en el IDP como en el Índice de Efecto. El RAP es negativo (salvo la mortalidad para el menor de 5 años), porque las tasas promedio son inferiores a las que exhibe la capital del país, que es la provincia con el IDH más alto. Los valores de equivalencia de entropía son muy bajos, lo cual es expresión del bajo potencial redistributivo de la carga de enfermedad. Sólo en el caso de la mortalidad materna alcanza un valor de 0.38 lo que implica un 12% de redistribución potencial de los casos totales para alcanzar la completa igualdad. Este resultado se explica perfectamente como consecuencia de las altas tasas de mortalidad materna de La Habana interior en comparación con otras provincias, que, no obstante, no se asocia con los valores relativos del IDH de esa provincia.⁵⁷

Referencias:

1. Santiago MI, Hervada, X, Naviera G, Silva LC, Fariñas H, Vázquez E, Bacallao J, Mujica OJ. El programa EPIDAT: usos y perspectivas. *Rev Panam Salud Pública* 2010; 27: 80-82.
2. Minujin A, Delamónica E. Equality matters for a world fit for children: lessons from the 90's. UNICEF Staff Working Papers. Division of Policy and Planning Series (Draft. Number DPP-03). New York. UNICEF, 2003.
3. Méndez E, Lloret MC. Índice de desarrollo humano a nivel territorial en Cuba. Período 1985-2001. *Revista Cubana de Salud Pública* 2005; 31. Versión On-line ISSN 0864-3466.
4. Bacallao Guerra J. Método de imputación múltiple basado en sustitución hot-deck usando un criterio refinado de vecino y un criterio de cercanía métrica. Tesis de Maestría. Facultad de Matemática. UH. 2008

⁵⁷ En los datos originales, para 2011, otro tanto ocurre con las provincias de Mayabeque y Artemisa (71.5 y 85.5, respectivamente).

CONCLUSIONES Y RECOMENDACIONES

Conclusiones

En relación con el primer objetivo (Exponer y discutir tres aspectos esenciales que determinan la necesidad de medir y el modo de hacerlo)

- La relación no lineal entre la condición socioeconómica y la salud es insuficiente para explicar las diferencias sanitarias en relación con la magnitud de las desigualdades sociales. El balance de la revisión de la literatura y algunos indicios preliminares de investigaciones realizadas en el país apuntan a que las desigualdades sociales tienen un efecto negativo sobre la salud.
- En la práctica, la polémica en relación con la medición de desigualdades sanitarias totales o con respecto a grupos se ha decantado a favor de esta última opción, que implica estratificar la medición de la salud según indicadores socioeconómicos y añadir a dicha medición indicadores que muestran las desigualdades. No obstante, los pro y los contra de las dos opciones se han examinado detenidamente y se ha formulado y desarrollado una propuesta que combina los elementos positivos de ambas.
- La medición multidimensional de las desigualdades dentro de la cual la salud sería una dimensión más junto a un conjunto de indicadores socioeconómicos, es una opción relativamente poco explorada y con considerables complicaciones metodológicas derivadas de la dificultad de definir relaciones de orden en espacios multidimensionales. La lógica que la sustenta es aceptable, pero se desvía de la práctica común en la medición de las desigualdades sociales en salud, en la que las variables sanitarias tienen un rol bien definido como variables dependientes o de respuesta en la acción de medir.

En relación con el segundo objetivo (Examinar los criterios y condiciones a los que se subordina la elección de una métrica)

- Se examinaron seis de estos criterios y condiciones y se argumentó que en todos los casos consideraciones de naturaleza política, ética o axiológica, antes o además que técnica, son determinantes de la elección. Estas consideraciones son básicamente subjetivas, y aunque no suelen hacerse explícitas, están presentes en el momento de la elección. No hay, por consiguiente, una métrica óptima desde el punto de vista metodológico, ni tampoco una que responda a un juicio de valor que sea universalmente preferible a todos los demás.

En relación con el tercer objetivo (Evaluar críticamente los índices que se emplean con mayor frecuencia)

Del examen crítico se han derivado juicios técnicos sobre todos los índices, sus condiciones de uso, su alcance y sus limitaciones. De manera general:

- Los índices que se basan en comparaciones dos a dos desconocen otros grupos de la clasificación o ignoran lo que ocurre en el centro de la jerarquía, si se trata de una clasificación ordenada. Por otra parte, no toman en cuenta el tamaño de los grupos.
- Los índices basados en modelos de regresión incorporan la dimensión socioeconómica de modo explícito, utilizan la información de todos los grupos y toman en cuenta el tamaño de éstos si el modelo se ajusta mediante el empleo de estimaciones ponderadas. Su principal limitación es que no pueden emplearse con criterios nominales de estratificación socioeconómica y que pueden transmitir una falsa apreciación de causalidad que desorientaría las decisiones programáticas que se deriven de la medición.
- Los índices basados en el concepto de disparidad deben usarse sólo para criterios nominales de estratificación, por lo que no incorporan explícitamente la condición socioeconómica. Son fáciles de calcular e interpretar y se asemejan a índices usados tradicionalmente por los estadísticos.
- Los índices basados en la noción de entropía no tienen un vínculo explícito con la condición socioeconómica. Un vínculo indirecto puede conseguirse si se estratifica de acuerdo a una variable socioeconómica, pero aún en ese caso no toman en cuenta el orden intrínseco de las clases, si éste existiera. Sin embargo, ambos tienen la interesante propiedad de que describen el potencial redistributivo de un sistema de clases y su distribución de probabilidad correspondiente para alcanzar la completa igualdad. El índice de Theil, además, es invariante a cambios de escala, toma en cuenta todas las clases y puede descomponerse en un componente dentro de clases y otro entre clases.
- El índice de concentración, que es el más importante de los índices basados en la noción de desproporcionalidad aplicado al campo de la salud, cumple tres de las propiedades consideradas como muy deseables para los índices: asociación explícita con la dimensión socioeconómica, uso de la información contenida en toda la estratificación, y sensibilidad a los cambios en la variable de salud y en la composición de los grupos. Sin embargo, son asimétricos al empleo de indicadores que tienen tanto una versión positiva como negativa.⁵⁸ Por otra parte, la noción de concentración de la carga de enfermedad es inseparable de los indicadores globales de salud, lo cual hace muy problemática y subjetiva la interpretación de los cambios en la concentración.

⁵⁸ Cobertura o falta de ella, analfabetos o no analfabetos, tasa de bajo peso o tasa de peso normal

En relación con el cuarto objetivo (Exponer y fundamentar novedades metodológicas de carácter instrumental y conceptual)

- La mayoría de los índices se ha enriquecido con algún refinamiento metodológico que mejora sus propiedades, aunque sin ninguna aspiración de optimizarlos:
 - En dos de los más comunes índices basados en comparaciones binarias se introdujo el sencillo expediente de la ponderación por el cociente de los tamaños de las clases.
 - En los índices basados en el concepto de disparidad, se introdujeron pequeñas variantes que pueden mejorar sus propiedades o complementar su información, en un caso añadiendo la ponderación por el tamaño relativo de los grupos (IPK), y en otro construyendo una versión relativa del índice (VEG).
 - En los índices basados en regresión se propuso un nuevo índice llamado "índice de desigualdad de la pendiente corregido" que constituye una extensión natural de los índices basados en comparaciones dos a dos de estratos extremos, corrigiendo la estimación en función de la posición de los grupos intermedios.
- Se propuso un "índice de concentración ajustado" que corrige el índice de concentración para la presencia de un factor confusor.
- Se introdujo un nuevo índice basado en la noción de entropía al que se llamó "índice integrado" que es una divergencia entre distribuciones de probabilidad que pondera las diferencias en función del tamaño de las tasas log-transformadas.
- En relación con el tema crucial del empleo de medidas absolutas o relativas de desigualdad se formularon expresiones simples que relacionan los cambios en los indicadores globales de salud con el incremento o la disminución de la desigualdad.
- Se propone la novedad conceptual de medir la desigualdad en dimensiones latentes bajo dos metodologías con un fundamento analítico diferente: el análisis de conglomerados basados en modelos y el análisis de conglomerados basados en regresión. Se sostiene el punto de vista de que cuando los criterios normativos habituales no exhiben gradientes de salud, las desigualdades sanitarias podrían estar ocultas en dimensiones latentes que hay que desvelar e interpretar. Se considera que ese es el caso de la sociedad cubana actual.

En relación con el quinto objetivo (Exponer y discutir los resultados de la aplicación de procedimientos basados en el análisis de clases latentes con datos ámbito nacional.)

- Los dos enfoques analíticos empleados revelan la existencia de grupos con perfiles de riesgo singulares en relación con el resto de la población a partir de indicadores calculados a escala familiar. En estos grupos, la densidad de diabéticos, hipertensos y fumadores es notablemente más alta que en el resto de los grupos.
- Ninguna de las variables socioeconómicas exhibe diferencias apreciables entre los grupos, lo que hace pensar en la existencia de un constructo

socioeconómico complejo que define subpoblaciones vulnerables o en riesgo para las variables de salud estudiadas.

- Las evidencias preliminares sugieren que en Cuba y en otros escenarios que no exhiban altos gradientes de salud en relación con los criterios normativos de estratificación para el estudio de las desigualdades sanitarias, debe explorarse la existencia de dimensiones latentes de desigualdad.

En relación con Epidat y su aplicación

- Los recursos simples de Epidat aplicados al cálculo de las desigualdades sanitarias, asociadas o no a estratificaciones sociales en las estadísticas sanitarias nacionales, confirman que en Cuba hay también niveles bajos de desigualdad territorial, lo cual refuerza la última conclusión del apartado anterior.

Recomendaciones

Se formulan recomendaciones de dos tipos: (a) las líneas de investigación que abre este trabajo y que le darían continuidad y (b) sus aplicaciones inmediatas.

(a)

- ◆ Caracterizar las poblaciones con perfiles singulares de riesgo mediante la búsqueda de los indicadores socioeconómicos y socioculturales apropiados.
- ◆ Extender el estudio y análisis de las desigualdades a otras áreas de interés en el campo de la salud (eg. el cáncer y las enfermedades emergentes y re-emergentes)
- ◆ Evaluar otros índices aplicables a la medición de las desigualdades sociales y de las desigualdades sociales en salud, en especial los que se basan en la noción de entropía generalizada.
- ◆ Profundizar en el estudio de los métodos de clasificación basados en el análisis de clases latentes y en su generalización a otras posibles aplicaciones en el campo de la salud.

(b)

- ◆ Revitalizar el análisis de la situación de salud, incorporando de rutina el estudio de las desigualdades sociales en salud.
- ◆ Además de las vías académicas, divulgar los resultados de este estudio para la generación de evidencia epidemiológica que acompañe a las estadísticas sanitarias del país.
- ◆ Capacitar a los gestores de estadísticas sanitarias en el uso y la interpretación de los índices, en particular de los más sencillos y que se emplean con mayor frecuencia internacionalmente.
- ◆ Incorporar normativamente la medición de las desigualdades sociales a las estadísticas sanitarias a escala nacional y local.

ANEXO

A. Sobre la relación entre el índice de concentración y los índices de desigualdad de la pendiente (absoluto y relativo)

De la página 48:

Wagstaff y colaboradores⁵⁹ demostraron que

$$I_C = \frac{2}{t_{total}} \text{cov}(r, t)$$

en donde $\text{cov}(r, t)$ es la covarianza entre los ridits (el rango relativo de los grupos ordenados de acuerdo a una variable socioeconómica) y la variable de salud, y t_{total} es la variable de salud en la población.

Pero la estimación mínimo-cuadrática de β en

$$t_i = \alpha + \beta r_i$$

viene dada por:

$$\hat{\beta} = \frac{\text{cov}(r, t)}{\text{Var}(r)}$$

Por consiguiente:

$$I_C = \frac{2}{t_{total}} \hat{\beta} * \text{Var}(r)$$

Pero

$$\hat{\beta} \equiv IDP$$

Por tanto

⁵⁹ Wagstaff A, Paci P, van Doorslaer E. On the measurement of inequalities in health. Soc Sci Med 1991; 33: 545 - 557.

$$I_C = \frac{2IDP \text{ var}(r)}{t_{total}} = 2IRD_{PAMUK} \text{ var}(r)$$

ya que

$$IRD_{PAMUK} = \frac{IDP}{t_{total}}$$

B. Sobre los cambios en las tasas y su efecto sobre las desigualdades según CT y DT (prueba de la propiedad (iii))

De la página 53:

Si se realiza una intervención que produce una disminución relativa de k_x en el grupo con la tasa más alta y de k_y en el grupo con la tasa más baja, y si $k_x \geq k_y$. Las nuevas tasas serán, respectivamente, $x - xk_x$, y $y - yk_y$. Entonces:

- La diferencia entre las tasas se mantiene constante sólo si $xk_x = yk_y$ ⁶⁰, y disminuye sólo si $xk_x \geq yk_y$.

$$DT = (x - xk_x) - (y - yk_y) = (x - y) - (xk_x - yk_y)$$

$$DT = x - y \Leftrightarrow xk_x - yk_y = 0$$

- c. El cociente entre las tasas se mantiene constante si $k_x = k_y$ y disminuye sólo si $k_x > k_y$.

$$CT = \frac{x - xk_x}{y - yk_y} = \frac{x(1 - k_x)}{y(1 - k_y)} \leq \frac{x}{y} \Leftrightarrow (1 - k_x) \leq (1 - k_y) \Leftrightarrow k_x \geq k_y$$

puesto que

$$k_x, k_y < 1$$

⁶⁰ Esta condición es equivalente a que la reducción absoluta en las tasas sea idéntica en ambos grupos.

C. Sobre el índice de concentración ajustado para un confusor

De la página 57:

Para datos no agrupados (ver III.25 y IV.12), el índice de concentración puede calcularse según la expresión

$$I_C = \frac{2}{nS} \sum_{i=1}^n s_i r_i - 1 \quad (\text{I})$$

Por otra parte, la variable de salud s_i depende de una variable y que genera la estratificación para el cálculo de la desigualdad, y de una variable z , que confunde el efecto de y sobre s . Se supone además que la relación puede modelarse como:

$$s_i = \alpha + \alpha_y y_i + \alpha_z z_i + \varepsilon_i \quad (\text{II})$$

Entonces es posible descomponer el índice de concentración del modo que resume la expresión siguiente:

$$I_C = IC_y \frac{\alpha_y \mu_y}{S} + IC_z \frac{\alpha_z \mu_z}{S} + \frac{IC_\varepsilon}{S} \quad (\text{III})$$

en donde

μ_y, μ_z : medias de las variables y, z respectivamente

S : media de la variable de salud en la población

IC_y, IC_z : índices de concentración de las variables y, z , respectivamente.

IC_ε : índice de concentración generalizado⁶¹ para ε

Debe notarse primeramente que: $S = \alpha + \alpha_y \mu_y + \alpha_z \mu_z$ ya que $\sum_i \varepsilon_i = 0$ (IV)

Sustituyendo (II) en (I) resulta

⁶¹ El índice de concentración generalizado es el índice de concentración multiplicado por la media poblacional, y equivale, como ya se expuso, al índice de desigualdad de la pendiente

$$I_C = \frac{2}{nS} \sum_i r_i (\alpha + \alpha_Y y_i + \alpha_Z z_i + \varepsilon_i) - 1 \quad (\text{V})$$

$$= \frac{2}{nS} \left(\alpha \sum_i r_i + \alpha_Y \sum_i r_i y_i + \alpha_Z \sum_i r_i z_i + \sum_i r_i \varepsilon_i \right) - 1 \quad (\text{VI})$$

Aplicando (I) a las variables y y z , y tomando en cuenta que:

$$\sum_i r_i = \frac{n}{2}$$

Resulta:

$$\sum_y y_i r_i = \frac{n\mu_Y}{2} (IC_Y + 1) \quad \text{y} \quad \sum_z z_i r_i = \frac{n\mu_Z}{2} (IC_Z + 1) \quad (\text{VII})$$

Por lo que

$$I_C = \frac{2}{nS} \left(\alpha \frac{n}{2} + \alpha_Y \frac{n\mu_Y}{2} (IC_Y + 1) + \alpha_Z \frac{n\mu_Z}{2} (IC_Z + 1) + \frac{n}{2} IC_\varepsilon \right) - 1 \quad (\text{VIII})$$

$$I_C = \left(\frac{\alpha}{S} + \frac{\alpha_Y \mu_Y}{S} + \frac{\alpha_Z \mu_Z}{S} + \frac{\alpha_Y \mu_Y}{S} IC_Y + \frac{\alpha_Z \mu_Z}{S} IC_Z + \frac{IC_\varepsilon}{S} \right) - 1 \quad (\text{IX})$$

Tomando en cuenta (IV) se demuestra (III).

D. Sobre un índice integrado basado en la noción de entropía

De la página 61:

$$\begin{aligned} DLM &= -\sum_{i=1}^K \pi_i [\ln(t_i) - \ln(t_{total})] = -\sum_i \pi_i \ln \left[\frac{s_i/n_i}{S/N} \right] = -\sum_i \pi_i [\ln(s_i) - \ln(n_i) - \ln(S) + \ln(N)] \\ &= -\sum_i \pi_i \left[\ln \frac{s_i}{S} - \ln \frac{n_i}{N} \right] = \sum_i \pi_i \left[\ln \frac{n_i}{N} - \ln \frac{s_i}{S} \right] = \sum_i \pi_i \ln \left(\frac{\pi_i}{p_i} \right) \quad (\text{X}) \end{aligned}$$

Por otra parte, el índice de Theil se calcula como:

$$I_T = \sum_{i=1}^K p_i \ln \left(\frac{p_i}{\pi_i} \right) \quad (\text{XI}) \text{ ver (III.18)}$$

Se ha definido el "índice de entropía integrado" (I_{int}) como el promedio de I_T y DLM.

$$I_{\text{int}} = \frac{1}{2}(I_T + DLM) = \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) * \ln \left(\frac{p_i}{\pi_i} \right) \quad (\text{XII}) \quad (\text{IV.18})$$

Considerado como distancia entre distribuciones de probabilidad I_{int} cumple dos de los axiomas de distancia. En efecto:

- i. $d(\pi, p) = 0$ si y sólo si $p_i = \pi_i$ para todo i .
- ii. $d(\pi, p) = d(p, \pi)$

La "descomponibilidad" de I_{int} es una consecuencia inmediata de la descomponibilidad de I_T y de la simetría entre I_T y DLM.

A partir de (XII) se tiene que:

$$I_{\text{int}} = \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) * \ln \left(\frac{p_i}{\pi_i} \right) = \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) * \ln \left(\frac{s_i/S}{n_i/N} \right) = \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) \ln \left(\frac{s_i/n_i}{S/N} \right) = \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) \ln \left(\frac{s_i}{n_i} \right) + \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) \ln \left(\frac{S}{N} \right)$$

Pero el último sumando de la expresión anterior es cero porque p_i y π_i son dos distribuciones de probabilidad y $\ln(S/N)$ es una constante. Luego

$$I_{\text{int}} = \frac{1}{2}(I_T + DLM) = \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) * \ln \left(\frac{p_i}{\pi_i} \right) = \frac{1}{2} \sum_{i=1}^K (p_i - \pi_i) * \ln \left(\frac{s_i}{n_i} \right)$$

y se comprueba que I_{int} es una divergencia entre distribuciones ponderada por las tasas log-transformadas.