**Anexo**

**A.1 Fórmulas**

**Media contra armónica (MCA) e Índice Mazziotta Pareto (MPI) para datos ponderados**

a) *Media contra armónica* *ponderada* (*MCAw*):

Ec. (A.1)

b) *MCAw*, *MPIw* *y sus parámetros*. Si el índice tiene un mensaje negativo (e.g., rezago o marginación):

Si el mensaje es positivo (e.g., bienestar):

Cuando la suma de los pesos es igual a la unidad, la media de los valores ponderados, la desviación estándar y el coeficiente de variación son[[1]](#footnote-1):

Ec. (A.2)

Ec. (A.3)

Ec. (A.4)

Ec. (A.5)

Cuando la suma de los pesos no es igual a la unidad,

; Ec. (A.6)

Ec. (A.7)

Donde, *w* es el peso asignado a cada variable zj estandarizada en el caso i; *N’* es el número de variables (igual al número de ponderaciones distintas de cero); son la media y desviación estándar de los valores estandarizados y ponderados, respectivamente.

**Media Geométrica (MG)**

La MG es resultado de la multiplicación de valores suavizado por la raíz a la potencia *n*. Esta suavización impide que los valores altos predominen sobre los valores bajos, como en la media aritmética. La media geométrica es menor a la media contra-armónica y la media aritmética, pero superior a la media armónica. La fórmula para valores no ponderados es:

Ec. (A.8

Donde: z son los valores estandarizados por un procedimiento balanceado, como , y *n* es el número de variables o dimensiones.

La MG para valores ponderados es:

Ec. (A.9)

Cuando la suma de las ponderaciones es igual a la unidad:

Ec. (A.10)

Donde:

Π= Letra griega mayúscula pi que indica el cálculo de un producto resultante de una multiplicación.

*zi*= Valor de la variable estandarizada.

*wi*= Ponderador de *wi*.

=suma de los ponderadores *w1*, *w2*, . . ., *wn* desde la primera variable hasta la variable *j*, para cada observación *i*.

Otra versión de la *MGw* con el mismo resultado, utilizando logaritmos:

Ec. (A.11)

Cuando la suma de las ponderaciones es igual a la unidad:

Ec. (A.12)

En esta segunda versión de *MGw* puede apreciarse que el log (o ln) ‘endereza’ o corrige cualquier asimetría que pudiera estar presente en *zi*. El resultado es el mismo con log o ln.

**Método de la distancia de Pena (DP2)**

El índice DP2, dada la distancia de cada observación de una referencia, que corresponde al área con el mayor valor en el caso del rezago social (peor escenario), se obtiene de la siguiente manera:

Ec. (A.13)

*i*= 1. . . .n (áreas); *j*=1,2, . . . .m (variables).

= diferencia entre el valor observado *xij* y el valor de referencia x\* (máximo en el caso del rezago) de la variable *j*.

*sj* es la desviación estándar de la variable *j*.

*R2* Es el cuadrado del coeficiente de correlación múltiple de la regresión de *xj* con *xj-1*, *xj-2*,. . . .*x1,* incluidas sucesivamente, una a la vez.

es un factor de corrección o ponderación. La primera variable recibe la ponderación de la unidad (no hay otra variable previa, ); la segunda tiene un peso de ; la tercera, ; etc.

Distancia Frechet (DF) = es la suma de las variables *j* estandarizadas, para cada observación *i*.

= Ec. (A.14)

Pasos para calcular DP2 (Montero, Chasco y Larraz 2010):

1. Calcular la Distancia Frechet (DF) para cada observación (suma horizontal) con la Ec. (15). Para cada unidad espacial *j*, DF es el valor máximo que DP2 puede alcanzar, definido por la Ec. (14).

2. Calcular el coeficiente de correlación entre cada indicador y DF para identificar el orden descendiente de las variables a fin de estimar su peso .

3. Calcular DP-1 con la Ec. (14) considerando el orden de las variables definido en el paso previo.

4. Obtener un nuevo orden descendiente siguiendo la correlación entre cada variable y DP-1.

5. Calcular DP-2 con la Ec. (14) considerando el orden de las variables definido en el paso 4.

6. Repetir el proceso de manera iterativa hasta que se alcance un criterio de convergencia; por ejemplo, que la diferencia entre dos DP contiguos sea cero. Si no se alcanzara el criterio de convergencia, puede tomarse el DP-1 o el promedio de los dos últimos.

**Estandarización z clásica re-escalada a 100 ± 10**

Ec. (A.15)

Donde y son la media y desviación estándar, respectivamente de la variable *j* y ± es el signo es la polaridad de la variable (+ si es positiva y – si la polaridad es negativa).

**Estandarización de z mín-máx de 0 a 100**

Ec. (A.16)

Donde *mín* y *máx* son el valor mínimo y máximo, respectivamente de la variable *j*.

**Estandarización zMP de 100 ± 30 (aproximado)**

Cutillo, Mazziotta y Pareto (2021) sugieren la siguiente estandarización para variables con polaridad positiva ():

Ec. (A.17)

Donde, *xj0* es el dato de referencia, usualmente el promedio nacional o del universo estudiado. El valor de es *aproximadamente* 100 ± 30; *xij* es el valor original de la variable *j* para el caso *i*; *mínj* y *máx*j son los valores mínimo y máximo de la variable *j*, respectivamente. Cuando *xij* es igual a *xj0*, el numerador en la fórmula es cero y es 100. Cuando *xij* es igual al máximo, el numerador es la mitad superior de los datos (aproximadamente +0.5), el término entre paréntesis es aproximadamente +30 y el resultado es *aproximadamente* 130. El procedimiento es similar para la mitad inferior de la distribución. En este caso, el término entre paréntesis es aproximadamente -30 y el resultado de la fórmula es *aproximadamente* 70. La Ec. (A.17) es formulada para variables con polaridad positiva porque un incremento en la variable original (*xij*) se refleja en un incremento en el valor estandarizado () y en el índice compuesto.

La versión para variables con polaridad negativa, no reportada por los autores, es la siguiente:

Ec. (A.18)

O, más sencillo:

Ec. (A.19)

La Ec. (A.18) o Ec. (A.19) estandariza y revierte la polaridad negativa de una variable.

La Ec. (A.17) es más sencilla de calcular y arroja exactamente los *mismos resultados obtenidos con la expresión utilizada por los autores en trabajos previos* (v.g., Mazziotta y Pareto 2016 y 2017), retomada por algunos seguidores (v.g., Ivaldi, Parra Saiani, Primosich, y Bruzzi 2020). La versión 2021 y anteriores, generan resultados centrados en 100, pero los valores en los extremos (70 y 130) no son exactos y el desbalance persiste porque la asimetría no cambia. Mazziotta y Pareto (2022a y 2022b) presentan también una versión para series de tiempo.

**A.2 Cuadros estadísticos**

Cuadro A.1 Matriz de correlación para las variables de rezago social, 2020

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 |
| x1 | 1.000 |  |  |  |  |  |  |  |  |  |  |
| x2 | 0.491 | 1.000 |  |  |  |  |  |  |  |  |  |
| x3 | 0.928 | 0.577 | 1.000 |  |  |  |  |  |  |  |  |
| x4 | 0.491 | 0.410 | 0.509 | 1.000 |  |  |  |  |  |  |  |
| x5 | 0.892 | 0.404 | 0.811 | 0.374 | 1.000 |  |  |  |  |  |  |
| x6 | 0.624 | 0.051 | 0.515 | 0.140 | 0.546 | 1.000 |  |  |  |  |  |
| x7 | 0.860 | 0.316 | 0.746 | 0.354 | 0.904 | 0.481 | 1.000 |  |  |  |  |
| x8 | 0.787 | 0.153 | 0.728 | 0.214 | 0.824 | 0.527 | 0.785 | 1.000 |  |  |  |
| x9 | 0.617 | 0.227 | 0.571 | 0.037 | 0.705 | 0.478 | 0.667 | 0.774 | 1.000 |  |  |
| x10 | 0.842 | 0.280 | 0.723 | 0.501 | 0.847 | 0.519 | 0.787 | 0.674 | 0.554 | 1.000 |  |
| x11 | 0.828 | 0.373 | 0.736 | 0.570 | 0.714 | 0.460 | 0.703 | 0.651 | 0.540 | 0.889 | 1.000 |

Cuadro A.2 Correlación anti-imagen

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 |
| x1 | .814a |  |  |  |  |  |  |  |  |  |  |
| x2 | -0.283 | .505a |  |  |  |  |  |  |  |  |  |
| x3 | -0.665 | -0.185 | .877a |  |  |  |  |  |  |  |  |
| x4 | 0.095 | -0.083 | -0.214 | .859a |  |  |  |  |  |  |  |
| x5 | -0.060 | -0.436 | -0.008 | -0.047 | .785a |  |  |  |  |  |  |
| x6 | -0.546 | 0.395 | 0.177 | 0.033 | -0.091 | .745a |  |  |  |  |  |
| x7 | -0.466 | 0.289 | 0.271 | -0.046 | -0.456 | 0.344 | .859a |  |  |  |  |
| x8 | -0.215 | 0.613 | -0.145 | 0.063 | -0.451 | 0.209 | 0.119 | .789a |  |  |  |
| x9 | 0.328 | -0.252 | -0.113 | 0.336 | -0.131 | -0.284 | -0.189 | -0.416 | .825a |  |  |
| x10 | -0.105 | 0.464 | 0.060 | -0.017 | -0.694 | 0.060 | 0.161 | 0.415 | 0.047 | .762a |  |
| x11 | -0.277 | -0.259 | 0.131 | -0.274 | 0.574 | 0.097 | -0.039 | -0.291 | -0.242 | -0.746 | .776a |
| a Medidas de adecuación de muestreo (MSA) | | | | | | | | | | | | |

Fuente: Cálculos del autor en base a la información del CONEVAL disponible en <https://www.coneval.org.mx/Medicion/IRS/Paginas/Indice_de_Rezago_Social_2020_anexos.aspx>

Cuadro A.3 Ajuste en las ponderaciones asignadas por el criterio de razón insuficiente al omitir una variable del rezago social, 2020\*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable que se omite** | | **Peso w para** | | | | | | | | | | |
| **z1** | **z2** | **z3** | **z4** | **z5** | **z6** | **z7** | **z8** | **z9** | **z10** | **z11** |
| **Ninguna** |  | 0.067 | 0.067 | 0.067 | 0.200 | 0.200 | 0.050 | 0.050 | 0.050 | 0.050 | 0.100 | 0.100 |
| **Analf** | **sin z1** | X | 0.100 | 0.100 | 0.200 | 0.200 | 0.050 | 0.050 | 0.050 | 0.050 | 0.100 | 0.100 |
| **No\_asiste** | **sin z2** | 0.100 | X | 0.100 | 0.200 | 0.200 | 0.050 | 0.050 | 0.050 | 0.050 | 0.100 | 0.100 |
| **Edu\_incom** | **sin z3** | 0.100 | 0.100 | X | 0.200 | 0.200 | 0.050 | 0.050 | 0.050 | 0.050 | 0.100 | 0.100 |
| **Salud** | **sin z4** | 0.083 | 0.083 | 0.083 | X | 0.250 | 0.063 | 0.063 | 0.063 | 0.063 | 0.125 | 0.125 |
| **Tierra** | **sin z5** | 0.083 | 0.083 | 0.083 | 0.250 | X | 0.063 | 0.063 | 0.063 | 0.063 | 0.125 | 0.125 |
| **Excusa** | **sin z6** | 0.067 | 0.067 | 0.067 | 0.200 | 0.200 | X | 0.067 | 0.067 | 0.067 | 0.100 | 0.100 |
| **Agua** | **sin z7** | 0.067 | 0.067 | 0.067 | 0.200 | 0.200 | 0.067 | X | 0.067 | 0.067 | 0.100 | 0.100 |
| **Dren** | **sin z8** | 0.067 | 0.067 | 0.067 | 0.200 | 0.200 | 0.067 | 0.067 | X | 0.067 | 0.100 | 0.100 |
| **Elec** | **sin z9** | 0.067 | 0.067 | 0.067 | 0.200 | 0.200 | 0.067 | 0.067 | 0.067 | X | 0.100 | 0.100 |
| **Lava** | **sin z10** | 0.067 | 0.067 | 0.067 | 0.200 | 0.200 | 0.050 | 0.050 | 0.050 | 0.050 | X | 0.200 |
| **Refri** | **sin z11** | 0.067 | 0.067 | 0.067 | 0.200 | 0.200 | 0.050 | 0.050 | 0.050 | 0.050 | 0.200 | X |

\* La suma horizontal es igual a la unidad. Cada vez que se omite una variable, los pesos se ajustan considerando el número de indicadores y sus componentes.

Cuadro A.4 Ajuste en las ponderaciones obtenidas por el Anáisis de Componentes Principales (ACP) al omitir una variable del rezago social, 2020\*

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable que se omite** | | **Peso w para** | | | | | | | | | | |
| **z1** | **z2** | **z3** | **z4** | **z5** | **z6** | **z7** | **z8** | **z9** | **z10** | **z11** |
| **Ninguna** |  | 0.137 | 0.064 | 0.128 | 0.071 | 0.133 | 0.088 | 0.127 | 0.120 | 0.103 | 0.126 | 0.122 |
| **Analf** | **sin z1** | X | 0.140 | 0.153 | 0.141 | 0.157 | 0.144 | 0.154 | 0.152 | 0.148 | 0.154 | 0.153 |
| **No\_asiste** | **sin z2** | 0.073 | X | 0.068 | 0.064 | 0.074 | 0.071 | 0.075 | 0.077 | 0.071 | 0.075 | 0.072 |
| **Edu\_incom** | **sin z3** | 0.146 | 0.129 | X | 0.131 | 0.146 | 0.135 | 0.145 | 0.142 | 0.138 | 0.145 | 0.143 |
| **Salud** | **sin z4** | 0.081 | 0.070 | 0.077 | X | 0.083 | 0.077 | 0.082 | 0.083 | 0.082 | 0.077 | 0.075 |
| **Tierra** | **sin z5** | 0.154 | 0.136 | 0.150 | 0.138 | X | 0.140 | 0.147 | 0.146 | 0.142 | 0.149 | 0.150 |
| **Excusa** | **sin z6** | 0.100 | 0.092 | 0.100 | 0.092 | 0.101 | X | 0.100 | 0.096 | 0.093 | 0.099 | 0.099 |
| **Agua** | **sin z7** | 0.146 | 0.131 | 0.144 | 0.132 | 0.143 | 0.134 | X | 0.139 | 0.135 | 0.142 | 0.142 |
| **Dren** | **sin z8** | 0.139 | 0.125 | 0.136 | 0.126 | 0.135 | 0.125 | 0.133 | X | 0.125 | 0.136 | 0.135 |
| **Elec** | **sin z9** | 0.121 | 0.107 | 0.118 | 0.110 | 0.116 | 0.107 | 0.115 | 0.109 | X | 0.117 | 0.116 |
| **Lava** | **sin z10** | 0.146 | 0.131 | 0.144 | 0.129 | 0.143 | 0.133 | 0.142 | 0.141 | 0.137 | X | 0.137 |
| **Refri** | **sin z11** | 0.141 | 0.125 | 0.138 | 0.124 | 0.142 | 0.129 | 0.139 | 0.137 | 0.133 | 0.134 | X |

\*La suma horizontal de los pesos w (Coeficiente de puntuación del componente F1 sin rotar, en el ACP) es distinta a la unidad. Cada línea (renglón) de pesos implica una corrida del ACP sin la variable que se indica.

Cuadro A.5 Amplitud de rango. Efecto de la igual ponderación en las estandarizaciones zMP y zEB

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Amplitud de rango en estandarizaciones | | | |
|  | Sin ponderar | | Ponderadas | |
|  | zMP | zEB | wzMP | wzEB |
| Analf | 60 | 60 | 4 | 4 |
| No\_asiste | 60 | 60 | 4 | 4 |
| Edu\_incom | 60 | 60 | 4 | 4 |
| Salud | 60 | 60 | 12 | 12 |
| Tierra | 60 | 60 | 12 | 12 |
| Excusa | 60 | 60 | 3 | 3 |
| Agua | 60 | 60 | 3 | 3 |
| Dren | 60 | 60 | 3 | 3 |
| Elec | 60 | 60 | 3 | 3 |
| Lava | 60 | 60 | 6 | 6 |
| Refri | 60 | 60 | 6 | 6 |

Fuente: Cálculos del autor en base a las ponderaciones en el Cuadro 1.

Cuadro A.6 Correlación de rangos, excepto donde se indica.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | DP2 | MCAw | MGw | ACP |
| DP2 | 1.000 |  |  |  |
| MCAw | 0.944 | 1.000 |  |  |
| MGw | 0.951 | 0.991 (0.997)a | 1.000 |  |
| ACP | 0.981 | 0.952 | 0.952 | 1.000 |

Fuente: Cálculos del autor en base al Cuadro 3 del texto.

a Correlación de Pearson entre paréntesis.

Cuadro A.7 Estabilidad al omitir una variable endistintos procedimientos de agregación con distintos criterios de estandarización y sin ponderación.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Variable que se omite | MCA | MG | MCA | MG | MCA | MG |
| zEB sin ponderar | | zMP sin ponderar | | z sin ponderar | |
| Analf | 0.44 | 0.38 | 0.50 | 0.38 | 0.50 | 0.44 |
| No\_asiste | 0.81 | 0.75 | 0.75 | 0.81 | 1.13 | 1.06 |
| Edu\_incom | 0.56 | 0.56 | 0.38 | 0.44 | 0.44 | 0.44 |
| Salud | 1.13 | 0.75 | 1.44 | 1.38 | 1.31 | 1.13 |
| Tierra | 0.38 | 0.56 | 0.06 | 0.25 | 0.38 | 0.38 |
| Excusa | 0.63 | 0.50 | 0.63 | 0.56 | 0.69 | 0.94 |
| Agua | 0.63 | 0.94 | 0.75 | 0.88 | 0.81 | 0.75 |
| Dren | 0.94 | 0.88 | 0.56 | 0.81 | 0.44 | 0.69 |
| Elec | 1.00 | 0.94 | 1.19 | 1.19 | 1.31 | 1.38 |
| Lava | 0.69 | 0.81 | 0.63 | 0.75 | 0.69 | 0.56 |
| Refri | 0.31 | 0.44 | 0.63 | 0.56 | 0.38 | 0.50 |
|  |  |  |  |  |  |  |
| Media | 0.68 | **0.68** | 0.68 | 0.73 | 0.73 | 0.75 |
| DS | 0.25 | **0.19** | 0.35 | 0.32 | 0.35 | 0.32 |

Fuente. Cálculos del autor.

**Referencias en el Anexo**

Ivaldi, E., Parra Saiani, P., Primosich, J. J., y Bruzzi, C. (2020). Health and deprivation: A new approach applied to 32 Argentinian urban areas. Social Indicators Research, 151(1), 155-179. <https://doi.org/10.1007/s11205-020-02369-w>

Mazziotta, M., y Pareto, A. (2022a). Normalization methods for spatio‐temporal analysis of environmental performance: Revisiting the Min–Max method. Environmetrics.

<https://doi.org/10.1002/env.2730>

Mazziotta, M., y Pareto, A. (2022b). Normalization methods for aggregating time series: the constrained Min–Max method. Rivista Italiana di Economia Demografia e Statistica, 75(4), 101-108. <http://www.sieds.it/wp-content/uploads/2022/01/9_01441RV_Mazziotta.pdf>

Mazziotta, M., y Pareto, A. (2021). Weighting in non-compensatory composite indices: The weighted Mazziotta-Pareto index [Conference session]. Statistical Society of Slovenia and University of Ljubljana/17th Applied Statistics 2021 [Virtual conference], September 20–22, 2021, Eslovenia

<https://akastrin.si/as2021/files/contributed/38.pdf>

Mazziotta, M., y Pareto, A. (2017). Measuring well-being over time: The adjusted Mazziotta–Pareto index versus other non-compensatory indices. *Social Indicators Research*, 136(3), 967-976. <https://doi.org/10.1007/s11205-017-1577-5>

Mazziotta, M., y Pareto, A. (2016a). On a generalized non-compensatory composite index for measuring socio-economic phenomena. Social Indicators Research, 127(3), 983-1003. <https://doi.org/10.1007/s11205-015-0998-2>

Montero, J., Chasco, C., & Larraz, B. (2010). Building an environmental quality index for a big city: A spatial interpolation approach combined with a distance indicator. Journal of Geographical Systems, 12(4), 435-459.

<https://doi.org/10.1007/s10109-010-0108-6>

1. La fórmula de la media ponderada es muy conocida y puede inferirse fácilmente la versión para datos ponderados cuando la suma de las ponderaciones es igual a la unidad. La fórmula de la desviación estándar correspondiente es menos conocida. La versión que se reporta aquí proviene de la presentación virtual de Matteo Mazziotta en la *17th Applied Statistics 2021* (Mazziotta y Pareto 2021). [↑](#footnote-ref-1)