Versión 1.0.1 | Estudio

Autor: Vojtech Royce

Traducción del original en inglés. English original

Cómo cambia el ganador según las ponderaciones

Una prueba reproducible de las puntuaciones compuestas de igualdad

Documento de trabajo 1.0.1. Revisión del autor pendiente. Todos los perfiles numéricos de este artículo son ejemplos didácticos artificiales.

Resumen

Un índice compuesto puede invertir el orden de dos perfiles sin que cambie una sola observación subyacente. Este artículo demuestra el mecanismo mediante dos perfiles expresamente artificiales medidos en dos dimensiones. El perfil A tiene valores de 90 y 40; el perfil B tiene valores de 60 y 80. Con una media aritmética ponderada, su orden se invierte cuando la ponderación de la primera dimensión cruza el umbral de cuatro séptimos. El cálculo es exacto y reproducible. No constituye evidencia sobre ningún país ni grupo demográfico. Utilizamos el ejemplo para especificar una norma práctica de presentación para World Equality Index: publicar los valores de los componentes, definir las ponderaciones admisibles, mostrar los umbrales de cambio, distinguir la sensibilidad del modelo de la incertidumbre muestral y explicar si las fortalezas pueden compensar las debilidades. Un panel de dimensiones diferenciadas sigue siendo el objeto empírico principal. Una puntuación agregada, cuando sea defendible, es una interpretación declarada de esas dimensiones.

Por qué importa para la investigación sobre igualdad

Los derechos, las oportunidades, los resultados, la contribución, la carga y la resiliencia responden a preguntas distintas. Una jurisdicción puede obtener buenos resultados en una dimensión y malos en otra. Sumar esas dimensiones introduce una afirmación adicional: que las compensaciones entre ellas tienen sentido en una escala común y que las relaciones de intercambio elegidas son adecuadas para el propósito declarado. Una clasificación puede parecer precisa incluso cuando esa interpretación es la parte más controvertida del cálculo.

El manual de la OCDE y del Centro Común de Investigación de la Comisión Europea trata la selección de indicadores, la normalización, la ponderación y la agregación como decisiones que requieren una justificación transparente. Su análisis de la robustez recomienda examinar cómo afectan a un índice la incertidumbre y los supuestos alternativos. Este artículo adopta esa orientación metodológica general y aporta un ejemplo analítico propio y sencillo y una propuesta de reglas de presentación para WEI. Manual de la OCDE, la UE y el JRC de la CE, 2008, especialmente los pasos 6 y 7

En esta demostración no se utilizan observaciones reales de países. Las denominaciones perfil A y perfil B son deliberadamente artificiales. Las etiquetas de las dimensiones, Derechos y Resultados, son nombres ilustrativos, no códigos jurídicos ni resultados demográficos medidos. Sus valores numéricos se eligieron para crear una compensación transparente que pueda resolverse a mano. Nunca deben mostrarse como puntuaciones nacionales, incorporarse a un explorador empírico como observaciones reales ni describirse como hallazgos sobre grupos sociales.

El modelo

Supongamos que las puntuaciones de dos dimensiones se sitúan en una escala fija de 0 a 100 y que, según la convención ilustrativa de puntuación, los valores mayores se consideran mejores. El perfil A tiene Derechos = 90 y Resultados = 40. El perfil B tiene Derechos = 60 y Resultados = 80. Sea w la ponderación de Derechos (Rights) y 1 menos w la ponderación de Resultados (Outcomes). El intervalo admisible va de 0 a 1.

La puntuación aritmética es:

$S(w) = w \times \text{Rights} + (1 - w) \times \text{Outcomes}$

Al sustituir los valores de cada perfil se obtiene:

$S_A(w) = 90w + 40(1 - w) = 40 + 50w$

$S_B(w) = 60w + 80(1 - w) = 80 - 20w$

Al restar las dos puntuaciones se obtiene:

$S_A(w) - S_B(w) = 70w - 40$

Los dos perfiles empatan cuando 70w es igual a 40, es decir, w = 4/7, aproximadamente 0,571429. Por debajo de ese umbral, B ocupa una posición superior. Por encima, A ocupa una posición superior. En el umbral empatan. Las observaciones y la aritmética permanecen fijas; el cambio de orden se debe enteramente a la elección de ponderación permitida.

Resultados de la demostración

Ponderación de Derechos	Ponderación de Resultados	Puntuación del perfil A	Puntuación del perfil B	Orden
0,00	1,00	40,00	80,00	B por encima de A
0,25	0,75	52,50	75,00	B por encima de A
0,50	0,50	65,00	70,00	B por encima de A
4/7	3/7	68,571429	68,571429	Empate
0,75	0,25	77,50	65,00	A por encima de B
0,80	0,20	80,00	64,00	A por encima de B
1,00	0,00	90,00	60,00	A por encima de B

Cada valor de esta tabla se calcula a partir de los dos perfiles artificiales. No se obtuvo ninguna estimación de un conjunto de datos sociales. Con ponderaciones iguales, B aventaja a A en cinco puntos. Con una ponderación de Derechos de 0,80, A aventaja a B en dieciséis puntos. Calificar cualquiera de los dos resultados como la clasificación global objetivamente correcta ocultaría la decisión sustantiva incorporada en la ponderación.

El hallazgo relevante es más informativo que afirmar que las ponderaciones importan. Identifica exactamente dónde cambia la respuesta. Si un informe de investigación hubiera justificado una ponderación de Derechos entre 0,20 y 0,50, B se situaría por delante en todo ese intervalo especificado. Si hubiera justificado un intervalo de 0,65 a 0,90, A se situaría por delante en todo él. Si el intervalo admisible abarca valores a ambos lados de cuatro séptimos, el orden es contingente dentro de la familia de modelos. La justificación de un intervalo debe evaluarse antes de seleccionar la clasificación preferida.

La robustez se define para un conjunto de supuestos

Una clasificación solo es robusta respecto de un conjunto explícito de supuestos admisibles. Presentar una clasificación como robusta después de variar una ponderación de 0,49 a 0,51 dice poco sobre un intervalo plausible de 0,10 a 0,90. El intervalo, el método de normalización, la regla de agregación y el tratamiento de los datos faltantes deben figurar junto al resultado de robustez. El resultado no puede separarse del experimento que lo produjo.

Un caso especial útil es la dominancia por componentes. Si un perfil alcanza al menos el mismo valor en todas las dimensiones y un valor estrictamente mayor en una dimensión que recibe una ponderación positiva, una puntuación aritmética con ponderaciones no negativas lo sitúa por encima. Ninguna compensación invierte ese orden dentro de esos supuestos.

Nuestro ejemplo carece intencionalmente de dominancia: A aventaja a B en un componente y B aventaja a A en el otro. El cruce es una propiedad de ese conflicto, no un defecto de la aritmética.

Con más de dos dimensiones, un único control deslizante ya no abarca todas las combinaciones de ponderaciones. Una futura implementación empírica debería definir una región factible de ponderaciones y examinar las clasificaciones en toda ella. Una malla de evaluación poco densa puede mostrar ejemplos, pero no permite establecer automáticamente que una posición extrema sea imposible entre los puntos muestreados. Las afirmaciones sobre los límites de las posiciones alcanzables deberían sustentarse en una optimización exacta o en un procedimiento de aproximación documentado.

Qué significa y qué no significa un control deslizante

Un control interactivo de ponderaciones puede ayudar al lector a explorar supuestos. El control debería mostrar el vector completo de ponderaciones y normalizarlo de forma coherente, en lugar de modificar valores complementarios ocultos. Debería mantener los componentes originales en pantalla y permitir exportar un escenario. Un escenario guardado expresa el cálculo elegido por un lector, no sustituye al método de referencia publicado.

Si una herramienta muestrea ponderaciones uniformemente, la proporción de muestras en las que A se sitúa por delante es una propiedad de ese esquema de muestreo. No es la probabilidad de que A sea realmente más igualitario, la probabilidad de acuerdo público ni un nivel de confianza estadística. En este ejemplo bidimensional, una distribución uniforme de w en el intervalo de 0 a 1 sitúa a A por delante en los tres séptimos superiores del intervalo. Eso describe una proporción geométrica bajo una distribución arbitraria de ponderaciones. Otra distribución cambia la proporción sin modificar ningún valor de los perfiles.

Los intervalos de sensibilidad tampoco son intervalos de incertidumbre muestral. Un intervalo de confianza aborda la incertidumbre de una cantidad estimada mediante un procedimiento estadístico. Un intervalo de puntuaciones obtenido al variar las ponderaciones aborda las consecuencias de cambiar la regla de puntuación. Ambos pueden importar en el trabajo empírico, pero combinarlos en una banda de incertidumbre sin distinguir ambas fuentes dificultaría la interpretación del gráfico.

Compensación y escala

La fórmula aritmética permite la compensación. Una mejora en una dimensión puede compensar una pérdida en la otra. Algunos propósitos de investigación pueden aceptar ese intercambio; otros pueden exigir un umbral mínimo para un derecho o un servicio esencial. Una regla con umbrales obligatorios responde a una pregunta distinta de la que responde una media sin restricciones. WEI debería especificar la compensación permitida en lugar de tratar una fórmula conocida como una regla neutral por defecto.

La escala elegida también importa. Pasar una dimensión de un porcentaje a un recuento sin reajustar su escala puede cambiar drásticamente su influencia numérica. Incluso una ponderación nominalmente igual no garantiza la misma influencia práctica cuando las distribuciones de los componentes difieren. Antes de asignar ponderaciones, un artículo debe definir el constructo, el sentido, las unidades, la transformación y la población de

referencia de cada componente. Una puntuación alta debe tener una interpretación precisa en ese contexto.

En las medidas de igualdad, incluso el sentido puede ser objeto de controversia. Una diferencia menor entre grupos en un resultado observado no establece que exista igualdad de trato, y unos resultados desiguales pueden surgir de más de un proceso institucional. La decisión de premiar la paridad de resultados debería documentarse como el constructo que se está puntuando. Una medida de derechos legales debe proceder de pruebas sobre los derechos, en lugar de inferirse silenciosamente de una distribución del empleo.

Propuesta de norma de publicación para WEI

Cada índice compuesto empírico debería tener un propósito explícito y un inventario de componentes. El inventario debería incluir las versiones de las fuentes, la cobertura geográfica, los períodos, los denominadores, los valores faltantes, las transformaciones y el significado sustantivo de los valores más altos. Los lectores deberían poder examinar un componente sin aceptar la interpretación agregada.

El método debería identificar después un vector de ponderaciones de referencia y justificar su uso. Las ponderaciones y reglas de agregación alternativas admisibles deberían especificarse antes de examinar qué entidades se benefician. Los resultados deberían presentar comparaciones por componentes, puntuaciones de referencia, umbrales de cambio cuando puedan calcularse e intervalos de posiciones o puntuaciones bajo las alternativas declaradas. Los casos en los que cambia el orden merecen una indicación visible, no una nota al pie que describa el índice como generalmente robusto.

Los componentes faltantes necesitan una regla que preserve la comparación declarada. Renormalizar las ponderaciones para una entidad mientras se mantienen las originales para otra puede llevar a comparar constructos distintos. Un cálculo debería indicar cuándo se excluyen entidades, cuándo se imputan datos y si un resultado depende de esa decisión. Un valor faltante no debe convertirse silenciosamente en un cero ni en la media muestral.

Por último, la publicación debería incluir cálculos ejecutables y un registro de las decisiones editoriales. Las correcciones deberían identificar si cambió la evidencia subyacente, se corrigió un error de programación o se modificó una decisión normativa de puntuación. Una nueva política de ponderación constituye una revisión del método aunque el conjunto de datos siga siendo idéntico. Las descargas versionadas proporcionan al lector un objeto estable que citar y permiten separar los desacuerdos sobre la interpretación de los desacuerdos sobre la aritmética.

Límites y uso

Esta demostración prueba una inversión de posiciones dentro de un modelo aritmético deliberadamente sencillo. No evalúa los índices de igualdad existentes, no establece ponderaciones preferibles, no mide ninguna jurisdicción real ni aporta pruebas para una conclusión política. Su reproducibilidad es exacta porque los datos de entrada son ejemplos definidos, no porque la medición social esté exenta de incertidumbre.

El script adjunto reproduce todas las puntuaciones mostradas y verifica el empate en cuatro séptimos. La implicación práctica es una obligación de publicación: cuando el orden cambie

bajo supuestos admisibles, mostrar el cambio y explicar los supuestos. Cuando los componentes respondan a preguntas distintas, mantener visibles sus significados. Un índice compuesto puede resumir una perspectiva declarada, mientras que una plataforma transparente permite al lector ver cómo esa perspectiva dio forma al resultado.