Ficha Metodológica para la Construcción de un Índice Compuesto de Riesgo de Depresión

1. Selección y Preparación de Variables

Las variables fueron agrupadas según las dimensiones teóricas y clasificadas según su naturaleza:

- **Numéricas/Ordinales:** Pueden tratarse con Análisis de Componentes Principales (ACP/PCA).
- **Categóricas/Nominales:** Se analizaron mediante Análisis de Correspondencias Múltiples (ACM/MCA).

Antes de su inclusión en el análisis:

- Se revisó la completitud de los datos, eliminando registros con valores faltantes en las variables clave.
- Se aplicaron transformaciones necesarias, incluyendo recodificación y creación de variables dummy para los análisis requeridos.

2. Reducción de Dimensionalidad y Asignación de Pesos

2.1. Análisis por Dimensión

a) Dimensiones con variables continuas/ordinales (Ejemplo: Pobreza, Puntaje Disfuncionalidad Familiar)

- **Método:** Análisis de Componentes Principales (ACP/PCA).
- **Justificación:** Se evaluó la adecuación mediante el índice KMO (Kaiser-Meyer-Olkin) y la prueba de esfericidad de Bartlett.
- **Criterios de calidad:**
    - **KMO > 0.6:** Adecuación aceptable.
    - **Bartlett p < 0.05:** Correlación suficiente.
    - **Varianza explicada por el primer componente (PC1) > 40%:** Suficiente capacidad de síntesis.
- **Resultado:** Se extrajo el primer componente principal (PC1) y se usaron sus cargas factoriales como pesos internos para cada variable.

b) Dimensiones con variables categóricas (Ejemplo: Depresión, cuando hay muchas variables tipo sí/no)

- **Método:** Análisis de Correspondencias Múltiples (ACM/MCA).
- **Justificación:** El ACM es la técnica adecuada para reducción de dimensionalidad con variables categóricas.
- **Criterios de calidad:**
    - **Varianza explicada (inercia) del primer eje > 15%:** Suficiente capacidad de síntesis para fenómenos sociales.
- **Resultado:** Se usaron las coordenadas factoriales del primer eje como pesos internos para cada modalidad de variable.

2.2. Construcción de Subíndices Normalizados

Para cada dimensión, se calculó el subíndice individual como el score (puntaje) del primer componente (PC1 para ACP, Dim1 para ACM), y se normalizó a media cero y desviación estándar uno (z-score):

$$score_{normalizado} = \frac{Score - \mu}{\sigma}$$

3.3. Integración y Ponderación de Subíndices: Pesos Integrados

Para construir el índice global:

1. Se calculó la varianza explicada por el primer componente/dimensión en cada análisis (w_j).
2. Se ponderó cada subíndice normalizado por la proporción de varianza explicada en su dimensión.
3. Se integró el índice compuesto como la suma ponderada de los subíndices normalizados:

$$Indice_{Final} = \frac{\sum_{j=1}^{D}(score_j \; X \; w_j)}{\sum_{j=1}^{D} w_j}$$

Donde:

- $score_j$: Puntaje normalizado de la dimensión j.
- w_j: Varianza explicada por el primer componente de la dimensión j.
- *D*: Número de dimensiones.

Los pesos integrados de cada variable/modalidad se calcularon multiplicando el peso interno en la dimensión por la importancia relativa de la dimensión en el índice global:

$$Peso\ Integrado\ Variable = \frac{Peso\ en\ Dimensión\ X\ w_j}{\sum_{j=1}^{D} w_j}$$

Estos pesos son los que deben emplearse para el cálculo, interpretación y pronóstico del índice en nuevas bases de datos.

4. Validación y Evaluación de la Calidad del Modelo

Para cada dimensión, se calcularon y reportaron las siguientes métricas:

- **KMO** y **Bartlett** (para ACP/PCA): aseguran la pertinencia del análisis.
- **Varianza explicada** (en PCA o inercia en MCA): indica la capacidad de síntesis del modelo.
- **Número de variables/modalidades**: describe la complejidad y riqueza de cada dimensión.

Criterios de aceptabilidad: Se consideró aceptable una varianza explicada por el primer componente superior al 40% en PCA y superior al 15% en MCA, junto con KMO > 0.6 y Bartlett significativo.

5. Resultados y Aplicación

El resultado del proceso es un índice compuesto, interpretable y robusto, que permite:

- La identificación y priorización de hogares, individuos o comunidades en mayor riesgo.
- La comparabilidad entre unidades de análisis y entre distintos momentos temporales.
- El análisis de sensibilidad y simulación para evaluar el impacto de cambios en las variables.

Los pesos integrados de cada variable/modalidad están documentados y disponibles en un archivo Excel, junto con las métricas de calidad. (enviados anoche)

Referencias

- Filmer, D., & Pritchett, L. H. (2001). Estimating wealth effects without expenditure data—or tears: An application to educational enrollments in states of India. *Demography*, 38(1), 115-132.
- Vyas, S., & Kumaranayake, L. (2006). Constructing socio-economic status indices: how to use principal components analysis. *Health Policy and Planning*, 21(6), 459-468.