**Familia: Barrera o facilitador en el proceso de inclusión laboral de 3 poblaciones en edad productiva**

Marco teórico

# Escalamiento óptimo: Análisis de componentes principales categórico

El objetivo del escalamiento óptimo es asignar cuantificaciones numéricas a las categorías de una variable cualitativa, esto para aplicar los métodos estadísticos tradicionales en la solución de un problema. En este sentido el análisis de componentes principales categórico conocido por el acrónimo CATPCA, del inglés CATegorical Principal Components Analysis, busca la reducción de un conjunto original de variables en un conjunto más pequeño de componentes independientes que representen la mayor parte de la información encontrada en las variables originales, esto a través de la cuantificación óptima de las variables cualitativas bien sean estas de tipo: nominal u ordinal (IBM SPSS, 2011).

El procedimiento opera de la siguiente forma:

* En primera medida se asignan las restricciones correspondientes, dadas por la escala de medición de cada una de las variables en la tabla de datos.
* A continuación, el método determina dos tipos de parámetros: parámetros de estimación del modelo del análisis de componentes principales y parámetros de cuantificación óptima. De este modo los parámetros de cuantificación óptima se inicializan con valores aleatorios.
* Acto seguido el algoritmo ajusta un Análisis de componentes principales clásico (PCA) sobre los parámetros de cuantificación óptima iniciales, los cuales se van actualizando durante el proceso, reemplazando los parámetros viejos por los nuevos, hasta conseguir la convergencia.

El fin de efectuar un CATPCA para la presente investigación consiste en cuantificar de manera apropiada las variables categóricas que componen la base de datos con el propósito de identificar la influencia de las componentes de entorno del núcleo familiar sobre la inclusión laboral, luego posterior a la cuantificación apropiada de las variables se procede a utilizar estas con el método PLS-PM explicado en el siguiente apartado.

# Partial Least Squares – Path Modeling (Mínimos cuadrados parciales – Modelación de rutas)

Esta técnica estadística de modelación tiene como finalidad estudiar relaciones entre *bloques de variables* a través de *métodos de regresión lineal*, permitiendo explorar la asociación entre variables latentes (no observables). Cada bloque o grupo de variables define una variable latente (Sánchez, 2013). El modelo se compone de dos partes:

***Modelo estructural o modelo interno***. Esta es la parte del modelo que explora las relaciones entre las variables latentes y se ve expresada de un modelo de regresión lineal múltiple de la siguiente manera:

Donde,

: Variable latente de interés o respuesta

: Variables latentes explicativas o predictoras

: Coeficientes de regresión que miden el aporte de cada variable latente predictora sobre la variable latente de respuesta.

: Error aleatorio que cuantifica la variación no explicada por las variables latentes explicativas incluidas en el modelo.

***Modelo de medida o modelo externo***. Corresponde a la parte del modelo que se encarga de medir las relaciones entre cada variable latente con su propio bloque de variables observadas. En este punto se distinguen dos casos de interés:

* El primero denominado modelo reflectivo, donde la variable latente se considera una causa de las variables observadas.
* Y el modelo formativo, en el cual las variables observadas se consideran como variables predictoras de la variable latente.

Para el presente estudio, se consideró un modelo formativo, cuya expresión matemática se presenta a continuación:

Donde,

: Variable latente (puede ser explicativa o de respuesta)

: Coeficientes o carga que cuantifica la relación entre cada variable observada y su correspondiente variable latente.

: Variable observada que se encuentra asociada a la variable latente

: Error aleatorio que mide la variabilidad no representada por las variables observadas que se asociaron a la variable latente .