

Elaboración de redes analíticas interactivas con el paquete netCoin de R

Anónimo

Abstract En esta presentación se expone un paquete para producir visualizaciones dinámicas e interactivas mediante el análisis de estadístico de datos. Se cubrirán los fundamentos del análisis de coincidencias y el uso de netCoin, un paquete para crear gráficos de análisis de redes. Este tipo de gráficos se ha empleado no sólo para resolver problemas topográficos y representar estructuras de red, como las de las redes sociales, sino también para mostrar la correlación entre variables según modelos casuales. El análisis de trayectorias y los modelos de ecuaciones estructurales son, en efecto, bien conocidos por los científicos sociales, pero ambos estaban restringidos a las variables cuantitativas en sus primeras etapas. En esta presentación, propondremos una nueva forma de mostrar las redes sociales y las conexiones entre variables cualitativas de forma similar al análisis de correspondencias, pero utilizando otro conjunto de técnicas multivariantes, como la regresión lineal y logística, mezcladas con el análisis de redes. Por ejemplo, uno de los usos específicos de esta técnica de análisis consiste en la caracterización de diferentes perfiles de respuesta por diversas variables sociodemográficas. El ARC (análisis reticular de coincidencias) explora los vínculos o co-ocurrencias de personas, características o eventos en determinadas circunstancias. El paquete R netCoin implementa este análisis de coincidencias y genera atractivas visualizaciones interactivas de los datos. Este paquete permite analizar las relaciones en los datos de las encuestas, las conexiones a través de los datos de las redes sociales o cualquier otro tipo de datos de redes dibujando gráficos interactivos. Incluso, netCoin también puede utilizarse para visualizar modelos estadísticos como las regresiones lineales o los modelos de ecuaciones estructurales. En el breve tiempo disponible se incluirán ejemplos con datos de redes sociales y de encuestas.

Palabras clave: análisis reticular de coincidencias - paquete netCoin - análisis de redes - gráficos dinámicos e interactivos

Introducción

A la presente propuesta se la denominará Análisis Reticular de Coincidencias (ARC), puesto que su principal objetivo es descubrir una serie de fenómenos, opiniones o características que en un determinado campo suelen aparecer conjuntamente. Muchos estadísticos persiguen la ilusión de dar con las causas de los fenómenos a partir de la información. Sin embargo, hay que ser cautos pues, salvo que se aplique con rigor el método experimental, las herramientas estadísticas son muy limitadas en el estudio de causas y efectos. Por ello, se propone una serie de análisis que no tienen como meta el descubrimiento de las “verdaderas” causas de los fenómenos en estudio, sino sus pautas de concurrencia con el fin de proporcionar al investigador posibles sugerencias de cómo está estructurada la realidad. El primer uso que se le puede dar a este análisis proviene de la dificultad de trabajar en cuestionarios con preguntas multirespuesta. Un siguiente empleo del análisis reticular de coincidencias es el análisis de contenido. También este análisis ha sido empleado para analizar documentos fotográficos y para ver la evolución de los retuits de un determinado tema en la red de Twitter. El análisis reticular de coincidencias tiene como finalidad descubrir las pautas de concurrencia de una serie de sucesos en un conjunto de escenarios. Su objetivo es descubrir cómo se distribuyen conjuntamente una serie de características dispuestas en distintas unidades en las que pueden o no estar presentes. Pueden distinguirse diversos grados de coincidencias: la nula, la simple, la probable, la condicional, la estadísticamente condicional, la subtotal y la total. Además, para estimar qué grado de coincidencias presentan dos sucesos puede emplearse un amplio rango de estadísticos: desde la mera frecuencia, hasta medidas de distancias más complejas, como la de Haberman o la de Ruselll. Ahora bien, no solo es importante representar las coincidencias numéricamente con estadísticos. Una buena representación gráfica puede ayudar sobremedida a entender mejor la distribución de las coincidencias de un conjunto múltiple de sucesos.

El paquete netCoin para representar redes de coincidencias

netCoin es un paquete escrito en R que permite al usuario generar matrices de coincidencias con sus correspondientes grafos, así como crear una página web interactiva. En esta página interactiva pueden cambiarse un gran conjunto de elementos de los grafos, así como generar tablas y gráficos descargables.

Entre los elementos modificables se citan los siguientes:

- a. La etiqueta, el tamaño, el color y la forma de los sucesos o nodos, en función de sus propiedades. De igual modo se pueden representar áreas de nodos con las mismas características (conglomerados) e incluso reemplazar las formas geométricas de los nodos con imágenes.

- b. La etiqueta, el grosor y el color de las aristas que representan las coincidencias entre los sucesos, en función de las propiedades de los vínculos (frecuencias, grado de coincidencia, significación, ...)
- c. Se puede hacer una selección de nodos manual o en función de sus atributos.
- d. Se permite realizar una selección de las aristas en función de sus propiedades.
- e. De modos similar, puede cambiarse la disposición de los nodos del grafo según los diversos algoritmos que calculan otros paquetes de redes existentes en R.

Además de la representación gráfica, se obtienen dos tipos de tablas de atributos: la de los nodos y la de enlaces y en cada una de estas tablas aparecen las correspondientes propiedades de unos y otros. El paquete supone tiene dos características que lo distinguen de otras representaciones de redes entre los paquetes de redes. La primera es la elaboración de redes interactivas que pueden ubicarse en un servidor web a fin de que estén accesibles a cualquier persona con solo conocer su URL y la segunda es que permite una representación dinámica de las redes y el usuario puede ver cómo evolucionan los nodos y los enlaces a lo largo del paso del tiempo. En la presentación se empezará con una demostración con un par de ejemplos de resultados obtenidos con el paquete y se terminará con varios ejemplos de código para representar redes sencillas a través de las funciones del paquete.