

Informe IX: Birnbaum–Saunders autoregressive conditional duration models applied to high-frequency financial data

Econometric Modeling and Solving Social Problems

Katy Pacheco Manchego

24 julio 2020

Introducción

En la actualidad se tiene que los mercados financieros son muy variables a través que pasa el tiempo, para tener un control sobre los eventos y analizar la estructura de los tiempos entre eventos comerciales se registran el tiempo preciso de cada operación bursátil, junto con el precio y el volumen. Para hacer unas buenas predicciones de los eventos comerciales. En Saulo et al. (2019) Propusieron dos modelos de duración condicional autorregresiva de Birnbaum-Saunders especificados en términos de mediana condicional variable de tiempo y duraciones medias y compararon el rendimiento entre ellos. Es por ello que el doctor leiva, dada la flexibilidad y facilidad de estimación de los modelos propuestos por Saulo et al. (2019), ellos proporcionan una alternativa novedosa a los modelos de duración condicional autorregresivos existentes para hacer predicciones de los eventos comerciales. En resumen, el artículo Saulo et al. (2019) se enfoca en el modelo Birnbaum – Saunders ACD (BSACD) propuesto por Bhatti (2010), el cual se basa en la distribución BS conocida por que tiene propiedades interesantes y ser ampliamente estudiada. En ese contexto, consideraron y propusieron dos modelos:

- El primero, es un nuevo modelo basado en la media (BSACD1) especificado en términos de una TD media condicional variable en el tiempo, como es habitual en los modelos ACD, utilizando una versión reparameterizada de la distribución BS
- El segundo es un modelo basado en la mediana (BSACD2) especificado en términos de una TD media condicional variable en el tiempo.

Por lo tanto, el objetivo principal del artículo Saulo et al. (2019) es comparar los modelos BSACD1 y BSACD2. Además, obtener los estimadores de máxima verosimilitud (ML) de los parámetros del modelo BSACD1 y BSACD2 y evaluar su rendimiento utilizando un estudio de simulación Monte Carlo (MC), derivar herramientas de diagnóstico para los modelos BSACD1 y BSACD2 y evaluar la solidez de cada modelo en casos atípicos, ajustar los modelos BSACD1 y BSACD2 a un conjunto de datos del mundo real para evaluar, establecer la capacidad de pronóstico de los modelos BSACD1 y BSACD2 para la evaluación (C2) y comparar los rendimientos de ajuste y pronóstico de los modelos BSACD1 y BSACD2 con los modelos existentes en la literatura.

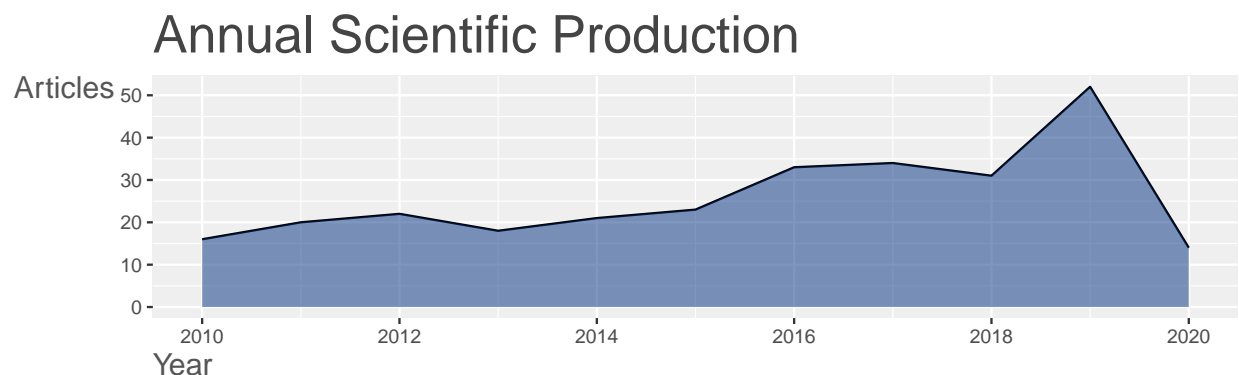
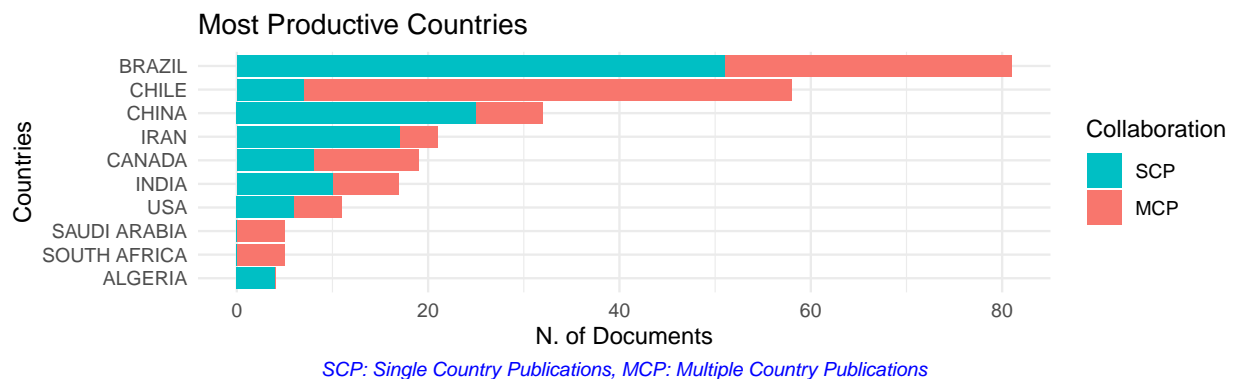
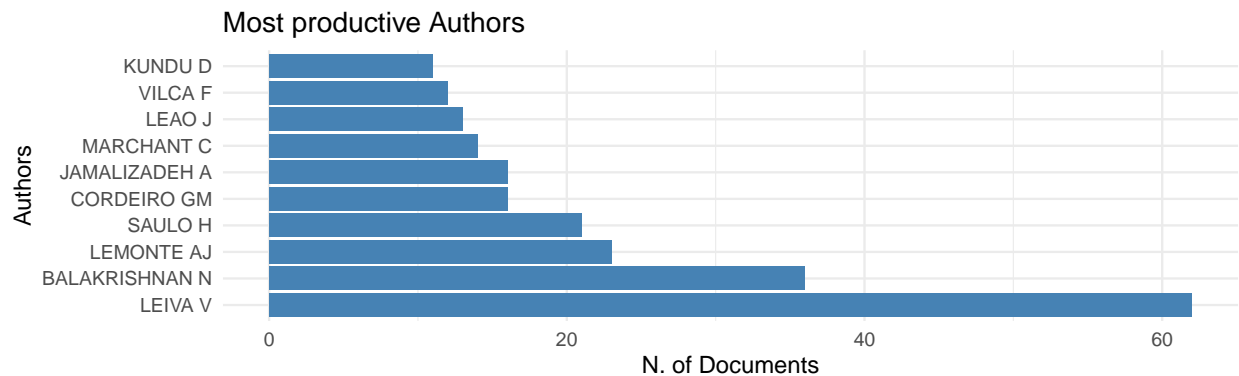
Análisis de la literatura Relacionada

En esta sección se identifican los artículos relacionados con Birnbaum–Saunders autoregressive conditional duration models applied to high-frequency financial data , durante el periodo 2010-2020, para ello se realizó un análisis de la literatura comprendido en los últimos 10 años, utilizando como palabra clave: Birnbaum–Saunders , y autoregressive conditional. Esta búsqueda se realizó en las bases de datos Web of Science donde encontramos 294 artículos relacionadas en el periodo de tiempo analizado

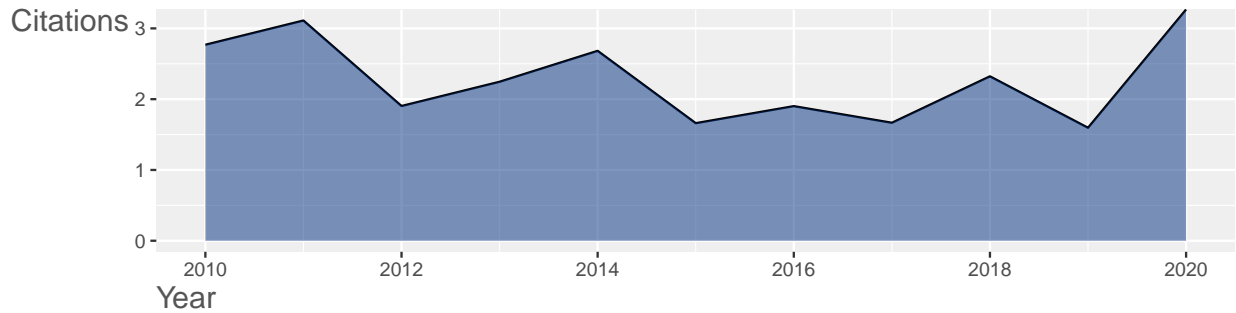
Se observa en las figuras que la producción científica por año ha venido aumentando de manera exponencial, además notamos que los años con mayor contribución son 2018 y 2019, en el año 2020 vemos que ha disminuido la producción científica, y que en los años de 2010-2014 fue casi que constante la producción científica.

En las Figuras, analizamos la áreas donde mas se realizan contribuciones y encontramos que estos temas son implementados e investigados en áreas como Matemáticas, ingenieras y economía.

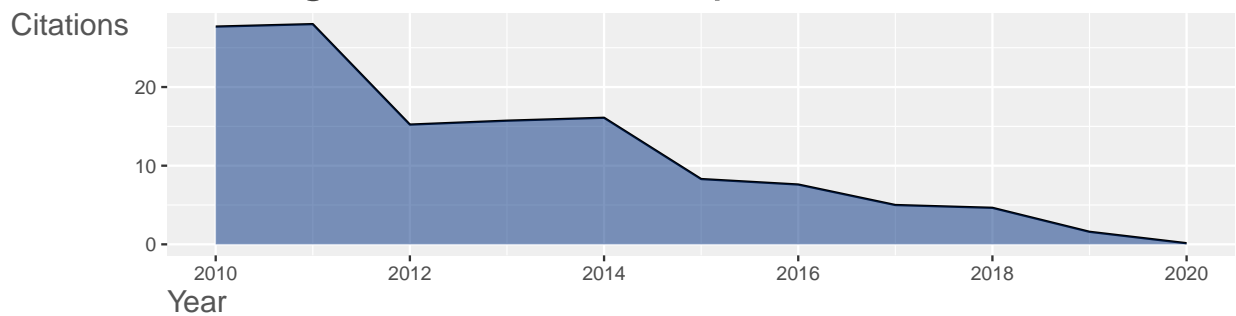
Por último analizamos los países con mayor contribución en estas área de investigación y encontramos que Brazil, chile, china, Iran, etc, son los países mas contribuyentes a estas áreas, destacando también que con respecto a los autores el **Phd Victor Leiva** es el mayor contribuyente en los últimos años en los temas de la distribución Birnbaum–Saunders



Average Article Citations per Year



Average Total Citations per Year



Marco Teórico (Descripción General)

En esta sección se hace una revisión de algunos conceptos teóricos relacionados con el artículo [1] y una breve mención de otros aspectos relevantes de ésta.

Big data

Según Wu et al. (n.d.) Big Data se refiere a conjuntos de datos crecientes y complejos de gran volumen con múltiples fuentes autónomas. Con el rápido desarrollo de las redes, el almacenamiento de datos y la capacidad de recopilación de datos, Big Data ahora se está expandiendo rápidamente en todos los dominios de la ciencia y la ingeniería, incluidas las ciencias físicas, biológicas y biomédicas.

Aunque el tamaño utilizado para determinar si un conjunto de datos determinado se considera Big Data no está firmemente definido y sigue cambiando con el tiempo, la mayoría de los analistas y profesionales actualmente se refieren a conjuntos de datos que van desde 30-50 Terabytes a varios Petabytes.

La naturaleza compleja del Big Data se debe principalmente a la naturaleza no estructurada de gran parte de los datos generados por las tecnologías modernas, como los web logs, la identificación por radiofrecuencia (RFID), los sensores incorporados en dispositivos, la maquinaria, los vehículos, las búsquedas en Internet, las redes sociales como Facebook, computadoras portátiles, teléfonos inteligentes y otros teléfonos móviles, dispositivos GPS y registros de centros de llamadas.

En la mayoría de los casos, con el fin de utilizar eficazmente el Big Data, debe combinarse con datos estructurados (normalmente de una base de datos relacional) de una aplicación comercial más convencional, como un ERP (Enterprise Resource Planning) o un CRM (Customer Relationship Management).

Distribución de Birnbaum-Saunders

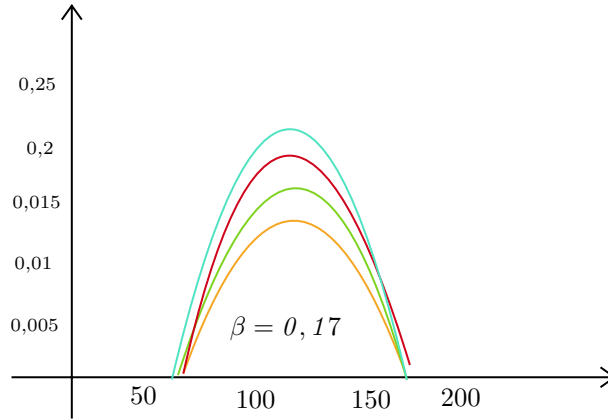
Según Bhatti (2010) una variable aleatoria X se distribuye BS si se puede representar mediante la transformación de una variable aleatoria con distribución normal estándar, $Z \sim N(0, 1)$, dado por

$$X = \sigma \left[kZ/2 + \{(kZ/2)^2 + 1\}^{1/2} \right]^2, \quad (1)$$

dónde $k < 0$ y $\sigma > 0$ son parámetros de forma y escala respectivamente. En este caso, se utiliza la notación $X \sim BS(k, \sigma)$ y la función de densidad correspondiente se obtiene como

$$f(x; k; \sigma) = \frac{1}{\sqrt{2\pi}} \exp \left(-\frac{1}{2k^2} \left[\frac{x}{\sigma} + \frac{\sigma}{x} - 2 \right] \right) \frac{x^{-3/2} [x + \sigma]}{2k\sigma^{1/2}}, \quad x > 0. \quad (2)$$

Entonces, se puede demostrar que $E[X] = \sigma[1 + k^2/2]$, $Var[X] = [k\sigma]^2[1 + 5k^2/4]$ y La mediana de la distribución es el parámetro de escala σ . Tenga en cuenta que la distribución BS está cerrado bajo escala y transformaciones recíprocas, es decir, $bX \sim BS(k, b\sigma)$, con $b > 0$ y $1/X \sim BS(k, 1/\sigma)$, respectivamente.



simulación del Monte Carlo

La simulación de Montecarlo según (“Simulación de Montecarlo - Qué es, definición y concepto Economipedia,” n.d.) es un método estadístico utilizado para resolver problemas matemáticos complejos a través de la generación de variables aleatorias. La simulación de Montecarlo o método de Montecarlo, le debe el nombre al famoso casino del principado de Mónaco. La ruleta es el juego de casino más famoso y también el ejemplo más sencillo de mecanismo que permite generar números aleatorios.

La clave de este método está en entender el término *simulación*. Realizar una simulación consiste en repetir o duplicar las características y comportamientos de un sistema real. Así pues, el objetivo principal de la simulación de Montecarlo es intentar imitar el comportamiento de variables reales para, en la medida de lo posible, analizar o predecir cómo van a evolucionar. A través de la simulación se pueden resolver desde problemas muy sencillos, hasta problemas muy complejos. Algunos problemas pueden solucionarse con papel y bolígrafo. Sin embargo, la mayoría requieren el uso de programas informáticos como Excel, R Studio o Matlab. Sin estos programas, resolver determinados problemas llevaría muchísimo tiempo.

Contribución del Trabajo

En el artículo (Saulo et al., 2019) propusieron dos modelos de duración condicional autorregresiva de Birnbaum-Saunders, El primero es un nuevo modelo basado en la media (BSACD1) y el segundo es un modelo basado en la mediana (BSACD2). Al realizar un estudio de diagnóstico influenciado determinaron

que el modelo basado en la mediana es más robusto para casos atípicos que el modelo basado en la media, por lo que es una opción más confiable usar ese modelo para hacer predicciones en condiciones de mercado altamente impredecibles. En general, los resultados que se obtuvieron en la investigación presentados en el artículo (Saulo et al., 2019) han demostrado que los dos modelos de duración condicional autorregresiva de Birnbaum-Saunders tienen rendimientos similares en términos de ajuste del modelo y capacidad de pronóstico. Por lo tanto, en (Saulo et al., 2019) hicieron una gran contribución con dos modelos predictivos de eventos comerciales que superan en todos los casos a los modelos existentes en la literatura para hacer predicciones en estos casos. ## Comentario Adicional Se tiene que los eventos comerciales son muy variantes y a veces poco predecibles y esto muchas veces se da por que hay muchos factores que influyen a resultados diversos. Seria de mucha mas ayuda realizar estos dos modelos predictivos que se proponen en (Saulo et al., 2019) de forma multivariada, es decir, donde hayan mas variables explicativas que ayuden a predecir con mas exactitud la variable respuesta.

References

- Bhatti, C.R., 2010. The Birnbaum–Saunders autoregressive conditional duration model. *Mathematics and Computers in Simulation (MATCOM)* 80, 2062–2078.
- Saulo, H., Leão, J., Leiva, V., Aykroyd, R.G., 2019. Birnbaum–Saunders autoregressive conditional duration models applied to high-frequency financial data. *Statistical Papers* 60, 1605–1629. <https://doi.org/10.1007/s00362-017-0888-6>
- Simulación de Montecarlo - Qué es, definición y concepto Economipedia, n.d.
- Wu, X., Zhu, X., Wu, G.-Q., Ding, W., n.d. *Data Mining with Big Data* 26.