Modelos Estadísticos muestreados en tiempos aleatorios

Consistencia del estimador de mínimos cuadrados en modelos muestreados en tiempos aleatorios y dirigidos por un ruido con larga memoria

Soledad Torres
soledad.torres@uv.cl

Universidad Central de Chile
CIMFAV- Universidad de Valparaíso.

November 10, 2024

Contenidos

Motivación

¿Por qué estudiar modelos con observaciones muestradas en tiempos aleatorios

- A lo largo de los años y en diferentes áreas de investigación, hemos notado que algunos eventos, como aquellos relacionados con fenómenos financieros, climatológicos y biológicos, entre otros; no ocurren de manera igualmente espaciada

¿Por qué estudiar modelos con observaciones muestradas en tiempos aleatorios

- A lo largo de los años y en diferentes áreas de investigación, hemos notado que algunos eventos, como aquellos relacionados con fenómenos financieros, climatológicos y biológicos, entre otros; no ocurren de manera igualmente espaciada
- Hay diferentes maneras de abordar este problema, como las series de tiempo irregularmente espaciadas o considerar que las observaciones son tomadas de manera aleatoria. En este trabajo, profundizaremos en el segundo enfoque.

¿Por qué estudiar modelos con observaciones muestradas en tiempos aleatorios

- A lo largo de los años y en diferentes áreas de investigación, hemos notado que algunos eventos, como aquellos relacionados con fenómenos financieros, climatológicos y biológicos, entre otros; no ocurren de manera igualmente espaciada
- Hay diferentes maneras de abordar este problema, como las series de tiempo irregularmente espaciadas o considerar que las observaciones son tomadas de manera aleatoria. En este trabajo, profundizaremos en el segundo enfoque.
- Estamos interesados en observar en un intervalo finito de tiempo, particularmente, nuestras observaciones estarán en el intervalo $[0, 1]$.

Otra característica interesante

- Una manera de explicar la fuerte correlación entre observaciones en un modelo, es representarla a través del exponente de Hurst, H, que toma valores entre $[0, 1]$.

Otra característica interesante

- Una manera de explicar la fuerte correlación entre observaciones en un modelo, es representarla a través del exponente de Hurst, H, que toma valores entre $[0, 1]$.
  - En particular, la dependencia de largo rango (*long-range dependence*), puede ser observada cuando $H \in (1/2, 1]$. Este fenómeno ha sido ampliamente estudiado a lo largo de los años.

Otra característica interesante

- Una manera de explicar la fuerte correlación entre observaciones en un modelo, es representarla a través del exponente de Hurst, H, que toma valores entre $[0, 1]$.
    - En particular, la dependencia de largo rango (*long-range dependence*), puede ser observada cuando $H \in (1/2, 1]$. Este fenómeno ha sido ampliamente estudiado a lo largo de los años.
    - El movimiento Browniano fraccionario (*fractional Brownian motion (fBm)*), es un proceso Gaussiano, con incrementos estacionarios, H-autosimilar y con la siguiente estructura de covarianza

$$R_H(t_i, t_j) := \mathbb{E}\left[B^H_{t_i} B^H_{t_j}\right] = \frac{\sigma^2}{2}\left[|t_i|^{2H} + |t_j|^{2H} - |t_i - t_j|^{2H}\right].$$

    - Existen otros procesos que son usados para representar el fenómeno de larga memoria y comparten la misma función de covarianza (a excepción de algunas constantes que dependen del valor de H) y que también pueden ser considerados para lo que sigue.

Trabajos previos

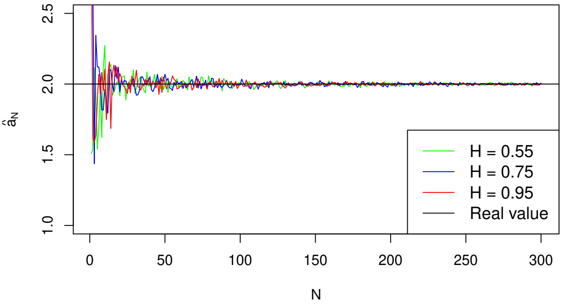
- Cuando las observaciones son muestreadas en tiempos aleatorios, podemos mencionar los trabajos de [2], [1] y [3], en cada uno de ellos el problema de estimación es abordado de un punto de vista no-paramétrico. Otros trabajos, también relacionados con los tiempos aleatorios, [4] y [5], ambos centrados en sistemas computacionales.
- Con respecto a trabajos donde el interés es la estimación paramétrica en series de tiempo que pueden ser representadas como una tendencia más un ruido con larga memoria podemos recordar los trabajos de [6], [7] y [8].

Conceptos previos

Modelo simple

Consideremos el siguiente modelo de regresión lineal simple

$$Y_{t_{i+1}} = at_{i+1} + \Delta B^H_{t_{i+1}}, \tag{1}$$



Modelo simple, con tiempos aleatorios

Consideremos el siguiente modelo de regresión lineal simple

$$Y_{\tau_{i+1}} = a\tau_{i+1} + \Delta B^H_{\tau_{i+1}}, \tag{2}$$

donde $a \in \mathbb{R}$ es el parámetro de drift (desconocido), $\Delta B^H_{\tau_{i+1}} = B^H_{\tau_{i+1}} - B^H_{\tau_i}$ y $\tau := \{\tau_i,\ 0 \leq i \leq N-1\}$ ó $\tau := \{\tau_i,\ 0 \leq i \leq N(1)\}$.

Tiempo aleatorio: Jittered case

Jittered sampling

Observaremos un proceso, en tiempos aleatorios τ,

$$\tau := \{\tau_i,\ 0 \leq i \leq N-1\},$$

con periodo $\delta = \frac{1}{N}$ pero contaminado por un ruido aditivo ν

$$\nu = \{\nu_{i,N};\ 0 \leq i \leq N-1\},$$

es una secuencia de variables aleatorias i.i.d. con función de densidad común $g_N(\cdot)$, dependiente de N y con soporte en $[0, 1/N]$. Entonces, definiremos

$$\tau_i = \frac{i}{N} + \nu_{i,N}, \quad i = 0, \ldots, N-1. \tag{3}$$

Algunas distribuciones que cumplen con la cualidad de ν son las distribuciones uniforme, triangular y raised cosine.

Tiempo aleatorio Renewal case

Renewal process

Observaremos un proceso, en tiempos aleatorios τ,

$$\tau := \{\tau_i,\ 0 \leq i \leq N(1)\},$$

con la diferencia que ahora consideraremos $N(1)$ (cantidad de observaciones antes de 1). Para este tiempo aleatorio, consideraremos que la sequencia τ satisface la propiedad de renovción, es decir

$$\tau_i = \sum_{j=1}^{i} t_j, \quad i = 1, 2, \dots, \tag{4}$$

donde t_j son variables aleatorias i.i.d con función de distribución $G_N(\cdot)$ con soporte en $[0, \infty)$ y que cumple con las siguientes propiedades

(HR1) $G_{N,i}$ es absolutamente continua con densidad $g_{N,i}(\cdot)$ en $[0, \infty)$,

(HR2) $G_{N,i}(0) = 0$, y

(HR3) $\int_0^\infty t dG_{N,i}(t) = N^{-1}$.

(HR4) $\int_0^\infty t^2 dG_{N,i}(t) \sim \frac{1}{N^\alpha}$ donde $\alpha > 1$.

(HR5) $\int_0^\infty t^4 dG_{N,i}(t) \sim \frac{1}{N^\beta}$ donde $\beta > 1$.

Comentarios sobre la cantidad de observaciones

- En el caso Jittered, consideramos una cantidad fija de observaciones, N, porque podemos asegurar que $\tau_{N-1} < 1$.
- En el caso Renewal, consideramos una cantidad aleatoria de observaciones, $N(1)$, porque no podemos asegurar que $\tau_N < 1$. Para este caso debemos considerar $\tau_{N(1)}$.

Resultados principales

Sobre el estimador de mínimos cuadrados, con una cantidad fija de observaciones

Para estimar el parámetro en el modelo (2), el estimador de mínimos cuadrados es determinado por

$$\hat{a}_N = \frac{\sum_{i=0}^{N-1} \tau_{i+1} Y_{\tau_{i+1}}}{\sum_{i=0}^{N-1} \tau_{i+1}^2}. \tag{5}$$

De (2) and (5), tenemos

$$\hat{a}_N - a = \frac{\sum_{i=0}^{N-1} \tau_{i+1} \Delta B^H_{\tau_{i+1}}}{\sum_{i=0}^{N-1} \tau_{i+1}^2}, \tag{6}$$

y agregando una normalización a (6), obtenemos

$$\hat{a}_N - a = \frac{\frac{1}{N} \sum_{i=0}^{N-1} \tau_{i+1} \Delta B^H_{\tau_{i+1}}}{\frac{1}{N} \sum_{i=0}^{N-1} \tau_{i+1}^2} = \frac{A_N}{D_N}. \tag{7}$$

Sobre el estimador de mínimos cuadrados, con una cantidad aleatoria de observaciones

Para estimar el parámetro en el modelo (2), el estimador de mínimos cuadrados es determinado por

$$\hat{a}_{N(1)} = \frac{\sum_{i=0}^{N(1)-1} \tau_{i+1} Y_{\tau_{i+1}}}{\sum_{i=0}^{N(1)-1} \tau_{i+1}^2}. \tag{8}$$

De (2) and (5), tenemos

$$\hat{a}_{N(1)} - a = \frac{\sum_{i=0}^{N(1)-1} \tau_{i+1} \Delta B^H_{\tau_{i+1}}}{\sum_{i=0}^{N(1)-1} \tau_{i+1}^2}, \tag{9}$$

y agregando una normalización a (6), obtenemos

$$\hat{a}_{N(1)} - a = \frac{\frac{1}{N} \sum_{i=0}^{N(1)-1} \tau_{i+1} \Delta B^H_{\tau_{i+1}}}{\frac{1}{N} \sum_{i=0}^{N(1)-1} \tau_{i+1}^2} = \frac{A_{N(1)}}{D_{N(1)}}. \tag{10}$$

Lema: convergencia del denominador

Lema, caso Jittered

Sea D_N definido por (7). Si τ_{i+1} satisface (3) , entonces

$$D_N \xrightarrow[N\to\infty]{a.s.} \frac{1}{3} \tag{11}$$

Lema, caso Renewal

Sea $D_{N(1)}$ definido por (10). Si τ_{i+1} satisface (4), entonces

$$D_{N(1)} \xrightarrow[N\to\infty]{a.s.} \frac{1}{3} \tag{12}$$

Esquema para la prueba

- Jittered case: notando que

$$D_N = \frac{1}{N}\sum_{i=0}^{N-1}\left(\frac{i^2}{N^2} + \frac{2i\nu_{i,N}}{N} + \nu_{i,N}^2\right)^2$$

  entonces, se estudia la convergencia con técnicas usuales. (convergencia en L^2 + Borel Cantelli).

- Renewal case: Reemplazamos $N(1)$ por N y transformamos D_N a una forma cuadrática en función de los valores de t_i, es decir

$$D_N = \frac{1}{N}\sum_{i=1}^{N}\sum_{j=1}^{N}(N-(i\vee j)+1)t_it_j,$$

  luego se propone una descomposición de acuerdo [**?**], para estudiar la convergencia de cada término.

Teoremas

Teorema, caso Jittered

Sea τ definidor por (3). Entonces el estimador (5) es fuertemente consistente, es decir,

$$\hat{a}_N \xrightarrow[N\to\infty]{a.s.} a.$$

Teorema, caso Renewal

Sea τ definidor por (4). Entonces el estimador (8) es fuertemente consistente, es decir,

$$\hat{a}_{N(1)} \xrightarrow[N\to\infty]{a.s.} a.$$

Esquema de la demostración

La base de la demostración es la descomposición de $\mathbb{E}\left[A_N^2\right]$.

$$\begin{aligned}
\mathbb{E}\left[A_N^2\right] &= \mathbb{E}\left[\frac{1}{N^2}\sum_{i=0}^{N-1}\tau_i^2\left(B^H_{\tau_{i+1}}-B^H_{\tau_i}\right)^2\right] \\
&+\mathbb{E}\left[\frac{1}{N^2}\sum_{0\le i,j\le N-1;|i-j|=1}\tau_i\tau_j\left(B^H_{\tau_{i+1}}-B^H_{\tau_i}\right)\left(B^H_{\tau_{j+1}}-B^H_{\tau_j}\right)\right] \\
&+\mathbb{E}\left[\frac{1}{N^2}\sum_{0\le i,j\le N-1;|i-j|\ge 2}\tau_i\tau_j\left(B^H_{\tau_{i+1}}-B^H_{\tau_i}\right)\left(B^H_{\tau_{j+1}}-B^H_{\tau_j}\right)\right] \\
&:=\mathbb{E}\left[A_N^{(1)}\right]+\mathbb{E}\left[A_N^{(2)}\right]+\mathbb{E}\left[A_N^{(3)}\right], \qquad (13)
\end{aligned}$$

además del condicionamiento adecuado.

Esquema de la demostración

Jittered: Probamos que el orden de convergencia, de los dos primeros términos, en (13) es de N^{1+2H} y del tercer término es de N^{2H}, por lo tanto obtenemos la convergencia casi segura.

Renewal:

1. Estudiamos la convergencia $|N(1) - N|$.
2. Estudiamos la convergencia de $\hat{a}_N$.
3. Estudiamos $|\hat{a}_{N(1)} - \hat{a}_N|$.

Simulación

Jittered case: Distribución uniforme

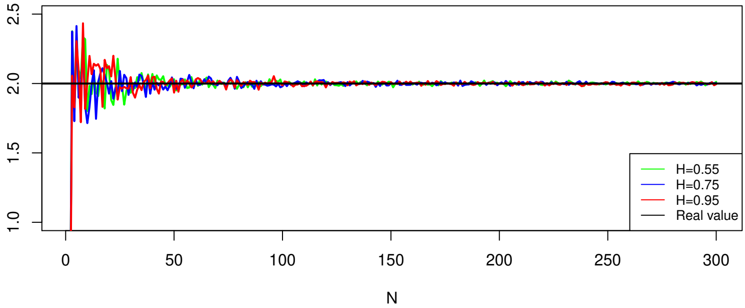


Figure: Convergencia del estimador $\hat{a}_N$ para diferentes valores de N

Jittered case: Distribución uniforme

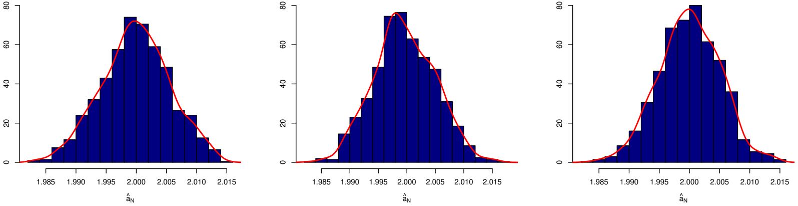


Figure: Distribución del valor del estimador $\hat{a}_N$ para distintos valores de $H = 0.55, 0.75, 0.95$

Jittered case: Distribución triangular

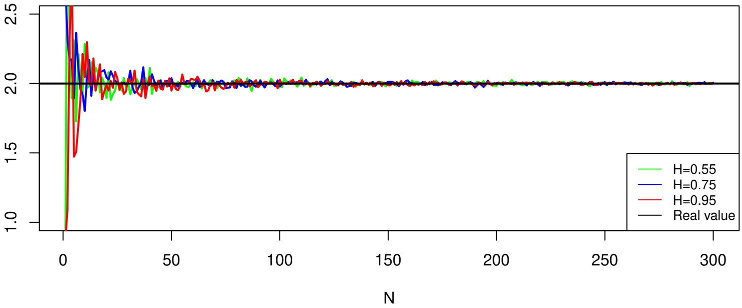


Figure: Convergencia del estimador $\hat{a}_N$ para diferentes valores de N

Jittered case: Distribución triangular

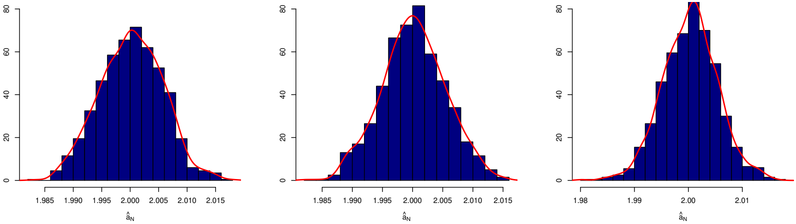


Figure: Distribución del valor del estimador $\hat{a}_N$ para distintos valores de $H = 0.55, 0.75, 0.95$

Renewal case: Distribución exponencial

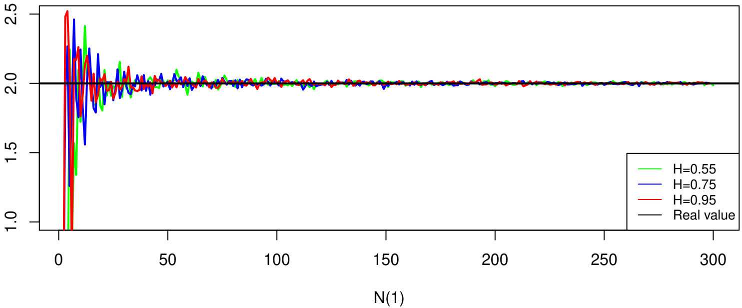


Figure: Convergencia del estimador $\hat{a}_N$ para diferentes valores de N

Renewal case: Distribución exponencial

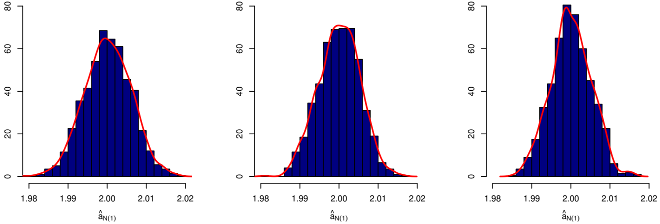


Figure: Distribución del valor del estimador $\hat{a}_N$ para distintos valores de $H = 0.55, 0.75, 0.95$

Renewal case: Distribución Beta prime

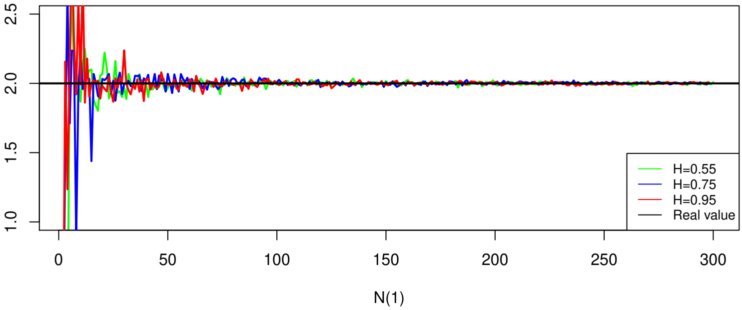


Figure: Convergencia del estimador $\hat{a}_N$ para diferentes valores de N

Renewal case: Distribución Beta prime

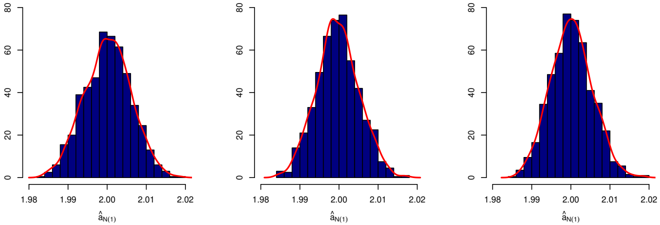


Figure: Distribución del valor del estimador $\hat{a}_N$ para distintos valores de $H = 0.55, 0.75, 0.95$

Generalizaciones

Caso "negatively superadditive dependent random observations"

- En [11] se generaliza el estudio de la consistencia
Un vector aleatorio $(X_1, X_2, \ldots, X_m)$ se dice que tiene una dependencia negativamente superaditiva si para toda función superaditiva ϕ tal que su esperanza $\mathbb{E}\phi(X_1, X_2, \ldots, X_m)$ existe,

$$\mathbb{E}\phi(X_1, X_2, \ldots, X_m) \leq \mathbb{E}\phi(X_1^*, X_2^*, \ldots, X_m^*),$$

donde $X_1^*, X_2^*, \ldots, X_m^*$ son independientes y para todo $1 \leq i \leq m$, X_i and X_i^* tienen la misma distribución (NSD).

Caso "negatively superadditive dependent random observations"

- En [11] se generaliza el estudio de la consistencia
  Un vector aleatorio $(X_1, X_2, \ldots, X_m)$ se dice que tiene una dependencia negativamente superaditiva si para toda función superaditiva ϕ tal que su esperanza $\mathbb{E}\phi(X_1, X_2, \ldots, X_m)$ existe,

$$\mathbb{E}\phi(X_1, X_2, \ldots, X_m) \leq \mathbb{E}\phi(X_1^*, X_2^*, \ldots, X_m^*),$$

  donde $X_1^*, X_2^*, \ldots, X_m^*$ son independientes y para todo $1 \leq i \leq m$, X_i and X_i^* tienen la misma distribución (NSD).
- Ejemplos

$$g(x) = |\Sigma|^{-1/2} g\left(x^t \Sigma x\right), \quad x \in \mathbb{R}^m$$

  son NSD, con $\Sigma = (\sigma_{i,j})$ satisface $\sigma_{i,j} < 0$ for $i \neq j$ y $\int_0^\infty g(t) t^{m/2-1} dt < \infty$. Como consecuencia un vector aleatorio gausisano con matriz de covarianza $\Sigma = (\sigma_{i,j})$ que satisface $\sigma_{i,j} < 0$ para $i \neq j$ es NSD.

Caso "negatively superadditive dependent random observations"

Nuestra hipótesis sobre los tiempos t_i es la siguiente.

Caso "negatively superadditive dependent random observations"

Nuestra hipótesis sobre los tiempos t_i es la siguiente.

(H_3) Para todo l $m \in \mathbb{N}$ (yN), el vector $(t_1, \ldots, t_m)$ es NSD.

- Note que (H_1) y (H_3) implican para $i \neq j$,

$$\mathbb{E}(t_i t_j) \leq \frac{1}{N^2}.$$

$$\mathbb{E}(t_i^2) \leq \frac{\tilde{C}^{2/(2+r)}}{N^2}.$$

- Para construir v.a. no-negativas NSD. tomamos $(X_1, \ldots, X_m)$ NSD y una función no decreciente $\varphi : \mathbb{R} \to \mathbb{R}_+$. Entonces $(\varphi(X_1), \ldots, \varphi(X_m))$ es no negativa y NSD.

Caso "negatively superadditive dependent random observations"

Nuestra hipótesis sobre los tiempos t_i es la siguiente.

(H_3) Para todo l $m \in \mathbb{N}$ (yN), el vector $(t_1, \ldots, t_m)$ es NSD.

- Note que (H_1) y (H_3) implican para $i \neq j$,

$$\mathbb{E}(t_i t_j) \leq \frac{1}{N^2}.$$

$$\mathbb{E}(t_i^2) \leq \frac{\tilde{C}^{2/(2+r)}}{N^2}.$$

- Para construir v.a. no-negativas NSD. tomamos $(X_1, \ldots, X_m)$ NSD y una función no decreciente $\varphi : \mathbb{R} \to \mathbb{R}_+$. Entonces $(\varphi(X_1), \ldots, \varphi(X_m))$ es no negativa y NSD.

Conclusiones

Conclusiones de este trabajo

- En un modelo de regresión lineal simple, tanto para tiempos no aleatorios como aleatorios, con un ruido con larga memoria, la consistencia del estimador es alcanzada.

Conclusiones de este trabajo

- En un modelo de regresión lineal simple, tanto para tiempos no aleatorios como aleatorios, con un ruido con larga memoria, la consistencia del estimador es alcanzada.
- Teórica y empíricamente, los estimadores son consistentes.

Conclusiones de este trabajo

- En un modelo de regresión lineal simple, tanto para tiempos no aleatorios como aleatorios, con un ruido con larga memoria, la consistencia del estimador es alcanzada.
- Teórica y empíricamente, los estimadores son consistentes.
- Considerar un esquema de muestreo aleatorio para la manera en que las observaciones es, en algunas ocasiones, un escenario más realista y las estimaciones propuestas en este trabajo tienen buenas propiedades .

References I

Vilar, J. A., & Vilar, J. M. (2000). Finite sample performance of density estimators from unequally spaced data. Statistics & probability letters, 50(1), 63-73.

Vilar, J. A. (1995). Kernel estimation of the regression function with random sampling times. Test, 4, 137-178.

Masry, E. (1983). Probability density estimation from sampled data. IEEE Transactions on Information Theory, 29(5), 696-709.

Chang, C. C. (2014). Optimum preventive maintenance policies for systems subject to random working times, replacement, and minimal repair. Computers & Industrial Engineering, 67, 185-194.

Zhao, X., Chen, M., & Nakagawa, T. (2014). Optimal time and random inspection policies for computer systems. Applied Mathematics & Information Sciences, 8(1L), 413-7.

Baillie, R. T., & Chung, S. K. (2002). Modeling and forecasting from trend-stationary long memory models with applications to climatology. International Journal of Forecasting, 18(2), 215-226.

References II

- Brockwell, A. E. (2007). Likelihood-based Analysis of a Class of Generalized Long-Memory Time Series Models. Journal of Time Series Analysis, 28(3), 386-407.
- Lobato, I. N., & Velasco, C. (2000). Long memory in stock-market trading volume. Journal of Business & Economic Statistics, 18(4), 410-427.
- Araya, H., Bahamonde, N., Fermn, L., Roa, T. and Torres, S. On the cinsistency of the least squares estimator in models sampled at random times driven by long memory noise: the renewal case Statistica Sinica 33 (2023), 1-26 doi:https://doi.org/10.5705/ss.202020.0457
- Araya, H., Bahamonde, N., Fermn, L., Roa, T. and Torres, S. On the cinsistency of the least squares estimator in models sampled at random times driven by long memory noise: the jittered case Statistica Sinica 33 (2023), 1-21 doi:https://doi.org/10.5705/ss.202020.0323
- Bertin, K., Torres, S., & Viitasaari, L. (2021). Least-square estimators in linear regression models under negatively superadditive dependent random observations. Statistics, 55(5), 1018–1034. https://doi.org/10.1080/02331888.2021.1993854