

## Introdução

Este trabalho apresenta o estudo de séries temporais envolvendo processos auto-regressivos de ordem  $p$ , com ruídos provenientes de uma distribuição  $\alpha$ -estável. O objetivo deste trabalho é melhor entender a distribuição e o seu impacto nos processos AR(p), além de estudar novas formas de analisar as séries obtidas. Para este fim, iniciamos os estudos com séries univariadas e verificamos seus comportamentos. Em seguida, estudamos o comportamento de séries bivariadas e verificamos a precisão dos estimadores obtidos. Por fim, iniciamos o estudo das séries quando o parâmetro de localização  $\delta$  varia em função do tempo.

## Análise Univariada

Iniciou-se com o estudo de rotina do grupo de pesquisa da orientadora, a fim de entender a lógica envolvida, além de adquirir maior conhecimento em R e para realizar análise de séries temporais. Foram realizadas simulações com séries temporais com tamanho amostral  $n = 1000$  e número de replicações  $re = 1$ , para testes.

A fim de melhor entender o comportamento da série, utilizamos as funções ACF e PACF, quando a variância é finita, ou a função Codiferença, quando a variância é infinita.

Foram estudadas as características de quatro casos específicos da distribuição  $\alpha$ -estável:

- ▶  $\alpha = 2$ , caso em que os processos são Gaussianos, com  $p \in \{1,2,3\}$  e  $\sigma_X \in \{0.1,1,2\}$ ;
- ▶  $\alpha = 1$ , processos correspondentes a distribuição Cauchy, com os mesmos valores de  $p$  e  $\sigma_X$ .
- ▶  $\alpha = 1.5$  e  $\alpha = 1.8$ , casos gerais  $\alpha$ -estável, novamente com os mesmos valores de  $p$  e  $\sigma_X$

A partir destes casos, obtivemos as seguintes figuras:

- ▶  $\alpha = 2, \sigma_X = 2, p=3$

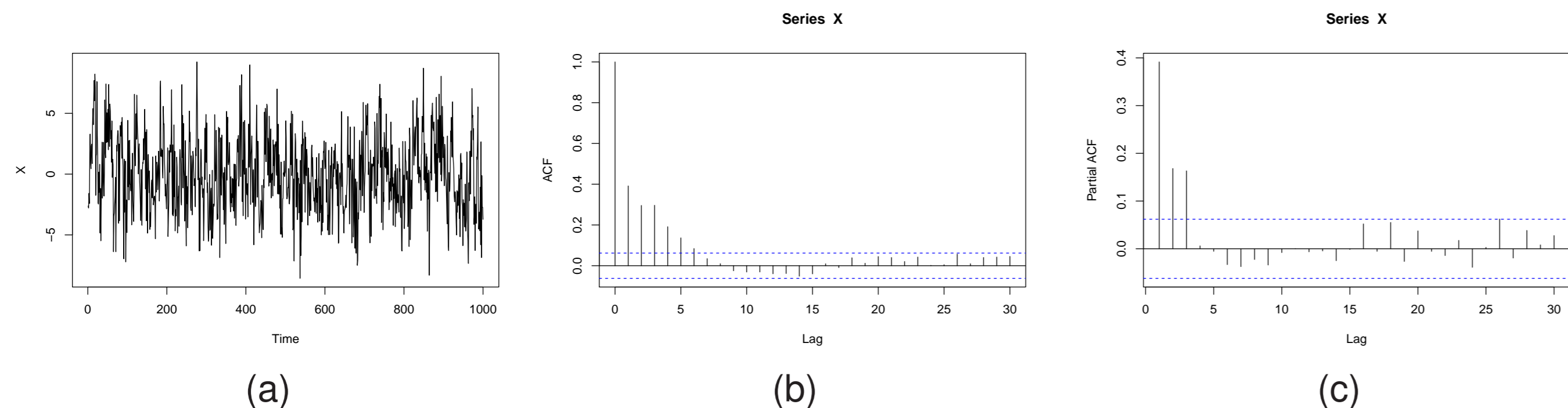


Figura 1: (a) Gráfico da série temporal; (b) ACF; (c) PACF.

Observando o gráficos da PACF, vemos que eles condizem com o fato de ser um processo auto-regressivo de ordem 3 (ver Morettin e Toloi, 2004). A ACF, no entanto, apresenta pequeno erro nos lags seguintes.

- ▶  $\alpha = 1, \sigma_X = 1, p=2$

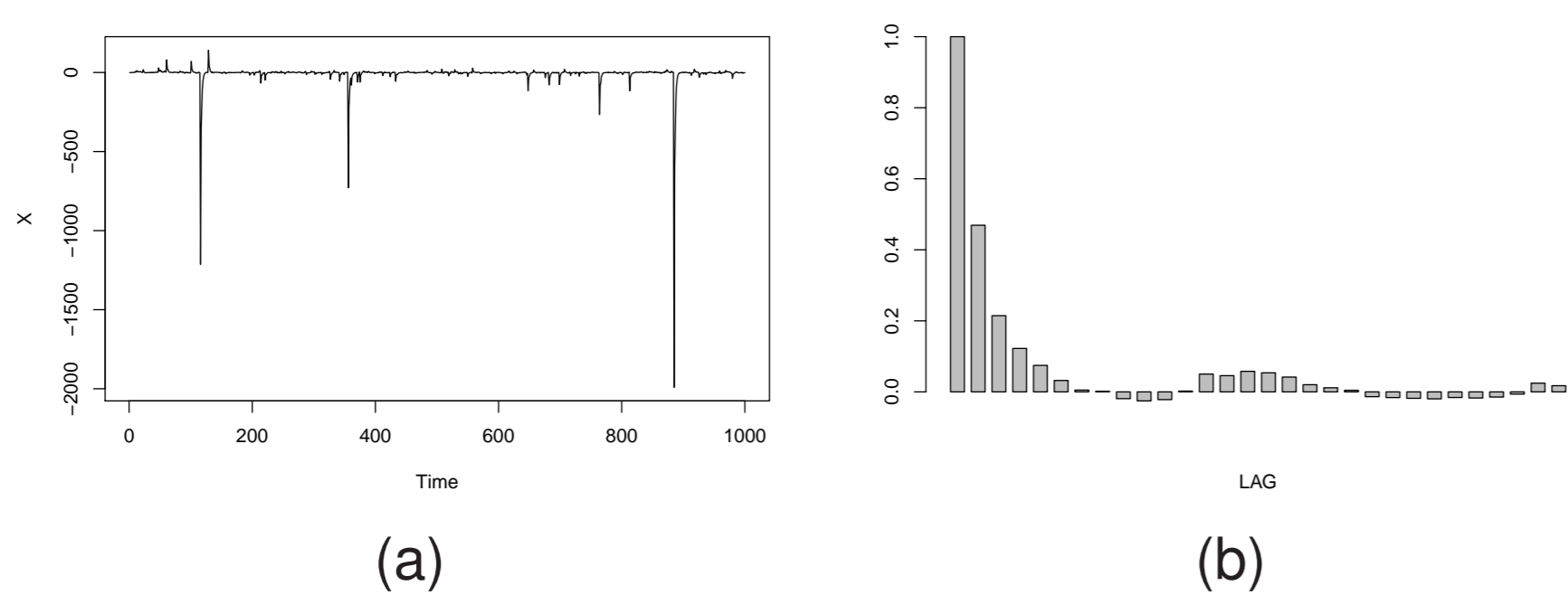


Figura 2: (a) Gráfico da série temporal; (b) Codiferença.

Através da Figura 2, podemos ver que a Codiferença se comporta como esperaríamos de um AR(2).

- ▶  $\alpha = 1.5, \sigma_X = 0.1, p=1$

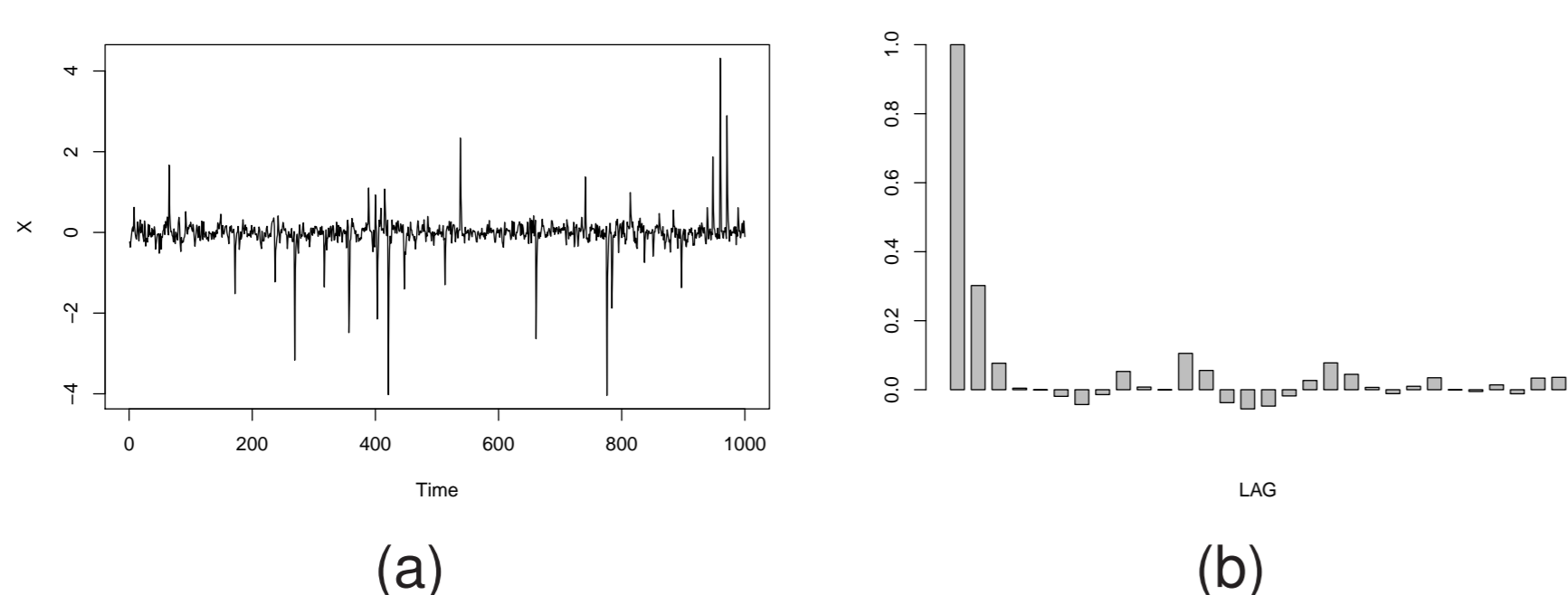


Figura 3: (a) Gráfico da série temporal; (b) Codiferença.

Na Figura 3, observa-se dois valores significantes na Codiferença, o que era esperado de um AR(1).

- ▶  $\alpha = 1.8, \sigma_X = 1, p=3$

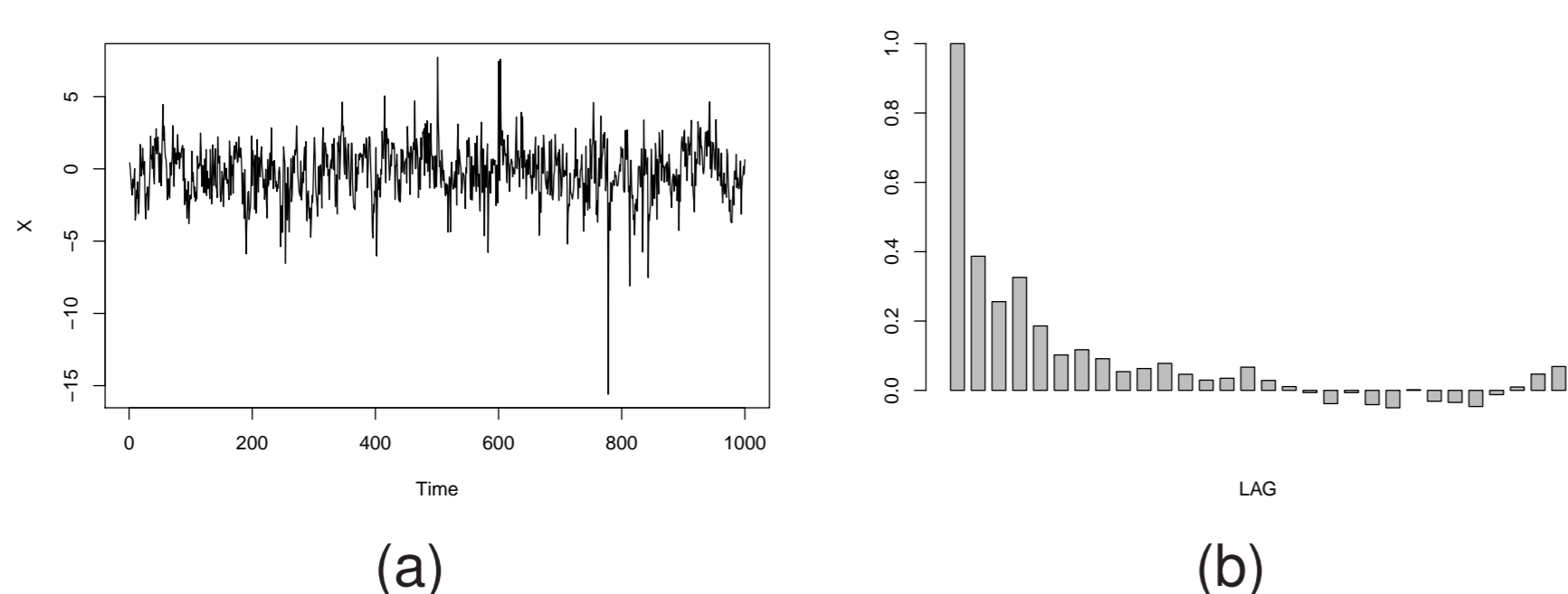


Figura 4: (a) Gráfico da série temporal; (b) Codiferença.

Observamos comportamento errôneo da Codiferença apresentada pela Figura 4 no quarto lag.

Comparando as demais figuras obtidas, concluímos que no caso  $p=3$  houve um grande número de incoerências quanto aos lags seguintes da ACF, PACF e da Codiferença. O caso  $p=1$ , no entanto, apresentou os melhores resultados de forma geral, como era esperado, uma vez que se trata de um processo bem mais simples.

## Análise Bivariada

Um processo auto-regressivo multivariado de ordem  $p$ , denotado por VAR(p), é um processo  $W_t$ , para todo  $t \in \mathbb{Z}$ , onde cada uma das suas componentes é um processo AR(p).

Neste trabalho, consideramos o caso bivariado, isto é,  $W_t \sim \text{VAR}(p)$ , onde  $W_t = (X_t, Y_t)$ , para todo  $t \in \mathbb{Z}$ . Os processos  $X_t$  e  $Y_t$  são auto-regressivos de ordem  $p$  com inovações  $\alpha$ -estáveis e codiferença dada por  $CD(X_t, X_t) = 1 = CD(Y_t, Y_t)$  e  $CD(X_t, Y_t) = \rho$ .

Para gerar os processos  $W_t$  utilizamos os seguintes parâmetros:

- ▶ tamanho amostral  $n = 10.000$ ,
- ▶  $\phi_0 = \varphi_0 = 0, \phi_1 = \varphi_1 = 0.3, \phi_2 = \varphi_2 = 0.1, \phi_3 = \varphi_3 = 0.2$ ,
- ▶ vetor de parâmetros  $\eta = (\alpha_1, \sigma_1, \alpha_2, \sigma_2)$ , onde  $X_t \sim \text{AR}(3)$  com  $\theta_1 = (\phi_0, \phi_1, \phi_2, \phi_3, \alpha_1, \sigma_1)$  e  $Y_t \sim \text{AR}(3)$  com  $\theta_2 = (\varphi_0, \varphi_1, \varphi_2, \varphi_3, \alpha_2, \sigma_2)$ .
- ▶  $\rho \in \{0.1, 0.3, 0.8\}$ .
- ▶  $re \in \{100, 500, 800\}$ .

Em seguida, foram feitas simulações com o intuito de decidir o valor mais adequado para as replicações.

Tabela 1: Resultado das estimativas para  $\eta$ , em um VAR(3), quando  $\rho = 0.3$  e  $re = 500$ .

Parâmetros		$\sigma_2 = 2$			
$\alpha_1$	$\alpha_2$	1.2	1.4	1.6	1.8
		$\sigma_1 = 1$	1.2	1.1852	1.3411
0.9049	1.0965			1.2714	1.5432
1.2178	1.6664			1.6137	1.9753
1.4	1.8632		2.3822	1.9960	2.1188
	1.3476		1.3449	1.3476	1.3476
	0.9753		0.9661	0.9753	0.9753
1.6	1.1961		1.9753	1.5967	1.9817
	2.1211		2.6833	2.0336	2.1564
	1.6086		1.6240	1.6125	1.5632
1.8	0.9542		1.0493	1.0383	0.9460
	1.9667		1.9735	1.9789	1.9740
	3.0595		2.6480	2.3432	2.1292
	1.9853	1.9831	1.9850	1.9847	
	1.0944	1.0822	1.0945	1.0783	
	1.2414	1.3958	1.6467	1.9778	
		2.0397	2.0987	2.1122	2.1320

Nota: No corpo da tabela os valores de  $\hat{\alpha}_1$  estão na primeira linha e de  $\hat{\alpha}_2$  na terceira linha enquanto que os de  $\hat{\sigma}_1$  estão na segunda linha e os de  $\hat{\sigma}_2$  na quarta linha. Os valores em negrito são aqueles que apresentaram uma diferença menor do que 10% em relação aos valores esperados.

Tabela 2: Resultado das estimativas para  $\eta$ , em um VAR(3), quando  $\rho = 0.3$  e  $re = 800$ .

Parâmetros		$\sigma_2 = 2$			
$\alpha_1$	$\alpha_2$	1.2	1.4	1.6	1.8
		$\sigma_1 = 1$	1.2	1.9845	1.9845
1.4622	1.4622			1.6306	1.5316
1.2639	1.4189			1.5620	1.8724
1.4	2.4075		2.0806	2.0627	1.8388
	1.4887		1.4439	1.6103	1.3831
	1.1252		1.0377	1.1106	0.9857
1.6	1.9754		1.9828	1.6405	1.9733
	3.0339		2.6699	2.0123	2.1585
	1.5779		1.9850	1.6022	1.6186
1.8	0.9835		1.1609	0.9927	1.0103
	1.9712		1.4248	1.9791	1.9725
	3.0363		1.9459	2.2822	2.1389
	1.9835	1.9851	1.9845	1.9859	
	1.0797	1.0684	1.0658	1.0890	
	1.2854	1.4162	1.6107	1.9782	
		2.0842	1.9568	2.0012	2.1334

Nota: No corpo da tabela os valores de  $\hat{\alpha}_1$  estão na primeira linha e de  $\hat{\alpha}_2$  na terceira linha enquanto que os de  $\hat{\sigma}_1$  estão na segunda linha e os de  $\hat{\sigma}_2$  na quarta linha. Os valores em negrito são aqueles que apresentaram uma diferença menor do que 10% em relação aos valores esperados.

Por fim, decidimos no uso de  $re=500$ , pois apesar de apresentar resultados comparáveis a  $re=800$ , o aumento no tempo computacional é desnecessário.

## Método com Parâmetro de Localização Dependendo do Tempo

Em seguida, iniciamos o estudo de séries temporais bivariadas  $W_t = (X_t, Y_t)$  onde cada componente é um processo auto-regressivo de ordem  $p$  com ruídos  $\alpha$ -estáveis, isto é,  $X_t \sim S_{\alpha_1}(\beta_1, \sigma_1, \delta_t)$  e  $Y_t \sim S_{\alpha_2}(\beta_2, \sigma_2, \delta_t)$  onde o parâmetro de localização  $\delta_t$  depende do tempo, isto é,  $\delta_{it}$  para  $i = 1, 2$ , e será dado por

$$\begin{aligned} \delta_{1t} &= \phi_0 + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \omega_t, \\ \delta_{2t} &= \varphi_0 + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \omega_t, \end{aligned}$$

onde o termo  $\omega_t$  é uma variável aleatória para cada  $t$  e terá função distribuição específica nas simulações. Utilizando o software estatístico R, realizamos simulações com os seguintes parâmetros:

- ▶ tamanho amostral  $n = 10000$ ,
- ▶  $re=1$ ,
- ▶  $p=1$ ,
- ▶  $\phi_0 = \varphi_0 = 0, \phi_1 = 0.4, \varphi_1 = 0.7$ ,
- ▶  $\alpha_1 = 2, \alpha_2 = 1.6, \sigma_1 = 1.5, \sigma_2 = 0.3, \beta_1 = \beta_2 = 0$ ,
- ▶  $\omega_t \sim \mathcal{N}(0, \sigma_\omega^2)$  e  $\sigma_\omega^2 \sim \mathcal{G}(0.1, 0.1)$ ,

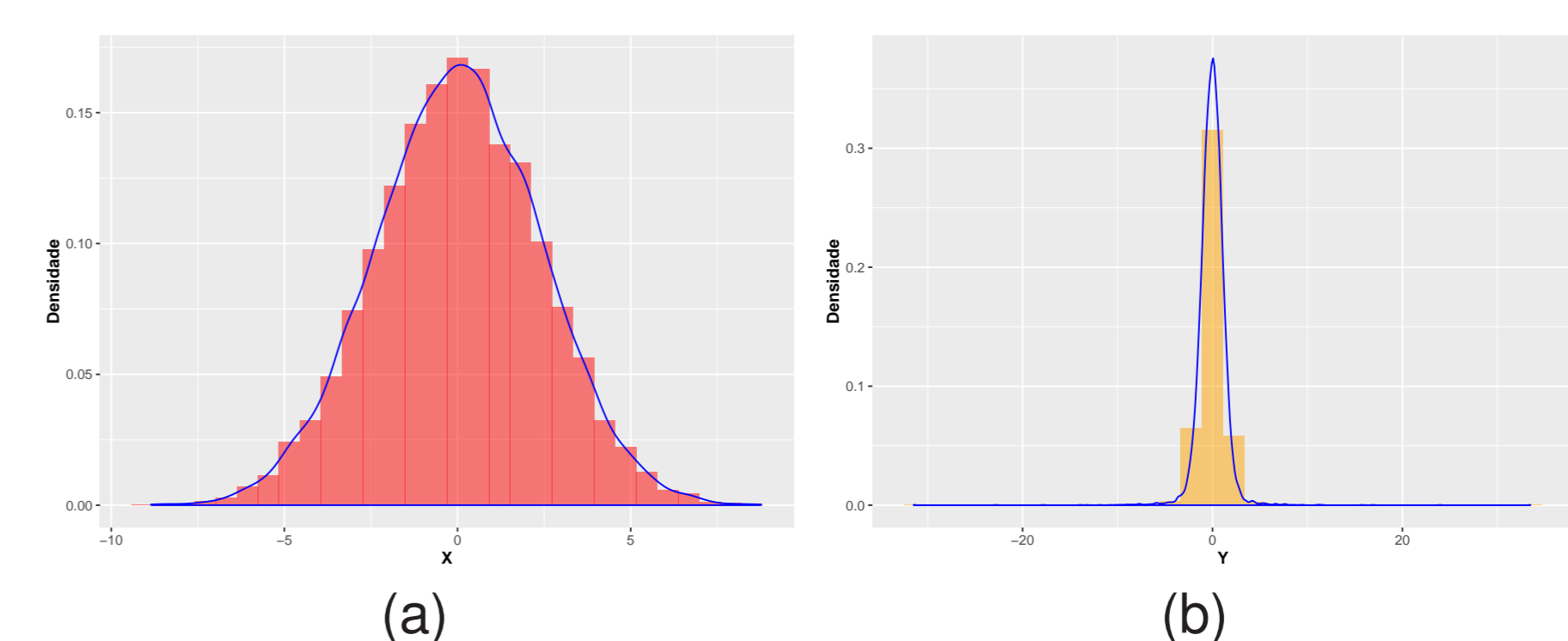


Figura 5: Gráfico da densidade de (a)  $X_t$  (b)  $Y_t$

## Resultados e Conclusão

Tendo em vista os estudos preliminares realizados até o momento, acreditamos ter avançado na análise de processos auto-regressivos bivariados com inovações  $\alpha$ -estáveis. Nosso objetivo é considerar o método com parâmetro de localização dependendo do tempo com mais detalhes e mais estudos de simulações. Outro objetivo é estimar o vetor de parâmetros  $\eta = (\alpha_1, \sigma_1, \delta_{1t}, \alpha_2, \sigma_2, \delta_{2t})'$  sob o enfoque da inferência Bayesiana, levando ainda em conta os parâmetros da componente  $\omega_t$ .

## Referências

- ▶ Brockwell, P.J. e R.A. Davis (1991). *Time Series: Theory and Methods*. New York: Springer-Verlag.
- ▶ Morettin, P.A. e C.M.C. Toloi (2004). *Análise de Séries Temporais*. ABE - Projeto Fisher. São Paulo: Edgard Blücher.
- ▶ Nolan, J.P. (2015). *Stable Distributions: Models for Heavy Tailed Data*. Available at [www.academic2.american.edu/~jpnolan](http://www.academic2.american.edu/~jpnolan).
- ▶ Rosadi, D. e M. Deistler (2009). "Estimating the codifference function of linear time series models with infinite variance". *Metrika*, Vol. 73, 395-429.
- ▶ Samorodnitsky, G. e M.S. Taqqu (1994). *Stable Non Gaussian Random Processes*. London: Chapman & Hall.
- ▶ Tsay, R.S. (2014). *Multivariate Time Series Analysis: With R and Financial Applications*. New Jersey: Wiley.