Using Autocopulas for the Testing of Linearity Against Markov-switching Type of Nonlinearity

Jana Lenčuchová

Department of Mathematics, Faculty of Civil Engineering, STU Bratislava

ISCAMI 2012, Malenovice, May 10-13, 2012

Content

Inspiration

- Rakonczai, P., Márkus, L. and Zempléni, A., Autocopulas: investigating the interdependence structure of stationary time series, in *Methodology and Computing in Applied Probability*, 14(1), 149167 (2012).
- Rakonczai, P., Márkus, L. and Zempléni, A.,Goodness of Fit for Auto-Copulas: Testing the Adequacy of Time Series Models, in *Proceedings of the 4th International Workshop in Applied Probability*, Compiegne, France, CD-ROM, paper No.73, (2008).

Copulas

Definition

A copula is a function $C : [0,1]x[0,1] \to [0,1]$ with the following properties:

1. for every u, v in $[0,1]$

$$C(u,0) = C(0,v) = 0 \tag{1}$$

   and

$$C(u,1) = u \quad \textit{and} \quad C(1,v) = v; \tag{2}$$

2. 2-increasing property - for every u_1, u_2, v_1, v_2 in $[0,1]$ such that $u_1 \leqq u_2$ and $v_1 \leqq v_2$,

$$C(u_2,v_2) - C(u_1,v_2) - C(u_2,v_1) + C(u_1,v_1) \geqq 0. \tag{3}$$

Sklar's theorem (1959)

For any pair (X, Y) of random variables, the joint distribution function H can be represented as

Theorem

$$H(x, y) = C(F(x), G(y)), \tag{4}$$

where $C(.)$ is a copula function and $F(x), G(y)$ are marginal distributions. Moreover, if the marginal distributions are continuous, the copula C is unique.

AutoCopulas

Definition

Given a strictly stationary time series Y_t and $l \in \mathbb{Z}^+$, the l-lag autocopula $C_{Y,l}$ is the copula of the bivariate random vector (Y_t, Y_{t-l}).The l-lag autocopulas as the function of the lag l give the autocopula function.

Some families of copulas

Name	$C_\theta(u,v)$	$\phi_\theta(t)$	$\theta \in$
Gumbel	$e^{-\left((-\ln u)^\theta + (-\ln v)^\theta\right)^{\frac{1}{\theta}}}$	$(-\log t)^\theta$	$[1,\infty)$
Frank	$-\frac{1}{\theta}\ln\left(1+\frac{\left(e^{-\theta u}-1\right)\left(e^{-\theta v}-1\right)}{\left(e^{-\theta}-1\right)}\right)$	$-\log\frac{e^{-\theta t}-1}{e^{-\theta}-1}$	$(-\infty,0)\bigcup(0,\infty)$
Clayton	$\left(u^{-\theta}+v^{-\theta}-1\right)^{\frac{-1}{\theta}}$	$t^{-\theta}-1$	$(0,\infty)$
Joe	$1-\left((1-u)^\theta+(1-v)^\theta-(1-x)^\theta(1-v)^\theta\right)^{1/\theta}$	$-\log\left(1-(1-t)^\theta\right)$	$[1,\infty)$

Copula GoF test

Focused Regions	Test statistics
Global	$S_1 = \frac{1}{m}\sum_{t_i \in [0+\epsilon, 1-\epsilon]} (\mathcal{K}(\theta_n, t_i) - \mathcal{K}_n(t_i))^2$
Upper Tail	$S_2 = \frac{1}{m}\sum_{t_i \in [0+\epsilon, 1-\epsilon]} \frac{(\mathcal{K}(\theta_n, t_i) - \mathcal{K}_n(t_i))^2}{1-\mathcal{K}(\theta_n, t_i)}$
Lower Tail	$S_3 = \frac{1}{m}\sum_{t_i \in [0+\epsilon, 1-\epsilon]} \frac{(\mathcal{K}(\theta_n, t_i) - \mathcal{K}_n(t_i))^2}{\mathcal{K}(\theta_n, t_i)}$
Lower and Upper Tail	$S_4 = \frac{1}{m}\sum_{t_i \in [0+\epsilon, 1-\epsilon]} \frac{(\mathcal{K}(\theta_n, t_i) - \mathcal{K}_n(t_i))^2}{\mathcal{K}(\theta_n, t_i)(1-\mathcal{K}(\theta_n, t_i))}$

Table: Numerically approximated Cramér-von Mises type test statistics on the Kendalls process

- $\mathcal{K}(t) = P(H(X,Y) \leq t) = P(C(F(X), G(Y)) \leq t) = P(C(u,v) \leq t)$
- $\mathcal{K}_n(t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(E_{in} \leq t)$, where $E_{in} = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}(U_j \leq U_i, V_j \leq V_i)$ and $\{(U_i, V_i), i \in \mathcal{T})\}$ is i.i.d. sample
- $(t_i)_{i=1}^{m}$ is an appropriately fine division of the interval $(0, 1))$

Markov-switching models

- random variable s_t in case of N possible states, can attain values from set $\{1, 2, 3, \ldots, N\}$
- stochastic process $\{s_t\}$ - a first-order ergodic Markov process (Hamilton 1989)
- $P(s_t = j | s_{t-1} = i, s_{t-2} = k, ...) = P(s_t = j | s_{t-1} = i) = p_{ij}$
- $p_{ij} > 0,\ i, j = 1, ..., N$
- $p_{i1} + p_{i2} + ... + p_{iN} = 1, \qquad i = 1, ..., N$
- Complete probability distribution of Markov chain is defined by the initial distribution $\pi_i = P(s_1 = i)$ and the state transition probability matrix $M = (p_{ij})_{i,j=1,...,N}$

Markov-switching models

- observable time series $y_1, \ldots, y_T$
- $y_t = \phi_{0,s_t} + \phi_{1,s_t} y_{t-1} + \ldots + \phi_{q,s_t} y_{t-q} + \epsilon_t, \qquad s_t = 1, \ldots, N$
  - $\epsilon_t \sim N(0, \sigma^2)$
  - ϕ_{j,s_t} are autoregressive coefficients of an appropriate regime $s_t = 1, \ldots, N$, $j = 0, 1, \ldots, q$
  - q is model order

Classical approach - Likelihood ratio test

Testing of a linear model against a 2-regime model

$H_0 : \boldsymbol{\varphi_1} = \boldsymbol{\varphi_2}$

against

$H_1 : \phi_{i,1} \neq \phi_{i,2}$ for at least one $i \in \{0, 1, 2, ..., q\}$

$\boldsymbol{\varphi_1}$, $\boldsymbol{\varphi_2}$ represents AR coefficients of a Markov-switching model in both regimes

Likelihood ratio test

$L = L_{MSW} - L_{AR}$

- L_{MSW} and L_{AR} are loglikelihood functions for the corresponding Markov-switching model and AR model
- this test statistic has non-standard distribution (Hansen 1992)
- simulation has to be carried out

New test using autocopula - testing procedure

1. Simulate AR time series with the size $n = 10000$ for 500 times.
2. Obtain their l-lag autocopula sample from the thinned by $s = 10$ series.
3. Calculate the test statistics $S_{i,j}$, $i = 1, \ldots, 4$ and $j = 1, \ldots, 500$ and choose the 0.95 quantiles as critical values $Q_{i,0.95}$, $i = 1, \ldots, 4$.
4. Simulate MSW model time series with the size $n = 10000$ for 500 times.
5. Obtain their l-lag autocopula sample from the thinned by $s = 10$ series.
6. Calculate the test statistics $S_{i,j}$, $i = 1, \ldots, 4, j = 1, \ldots, 500$ and reject H_0 whenever $S_{i,j} > Q_{i,0.95}$.

Testing linearity against MSW-type of nonlinearity using autocopulas

Copula	TS	l=1	l=2	l=3	Copula	TS	l=1	l=2	l=3
GUMBEL	S1	6,4%	7,6%	6,4%	NORMAL	S1	20,6%	4,8%	13,8%
	S2	5,8%	5,6%	6,0%		S2	29,0%	3,6%	15,8%
	S3	6,0%	9,2%	5,2%		S3	22,4%	5,2%	18,0%
	S4	6,4%	7,0%	5,8%		S4	28,8%	4,2%	16,4%
FRANK	S1	32,0%	36,0%	25.8%	JOE	S1	11,8%	16.0%	12.4%
	S2	44,4%	32,8%	27.2%		S2	7,2%	13.2%	13.4%
	S3	47,2%	37,2%	24.0%		S3	16,4%	25.6%	15.8%
	S4	47,2%	36,2%	25.8%		S4	12,0%	18.6%	16.0%
CLAYTON	S1	0%	0%	0%	FGM	S1	33,2%	6,8%	7,5%
	S2	0%	0%	0%		S2	33,4%	9,0%	13,0%
	S3	0%	0%	0%		S3	29,6%	9,6%	10,5%
	S4	0%	0%	0%		S4	39,4%	10,2%	10,0%

Table: Rejection rate of the null hypothesis of the AR copula if the true model is a 2-regime MSW model, significance level $\alpha = 0.05$

Thank you for your attention!