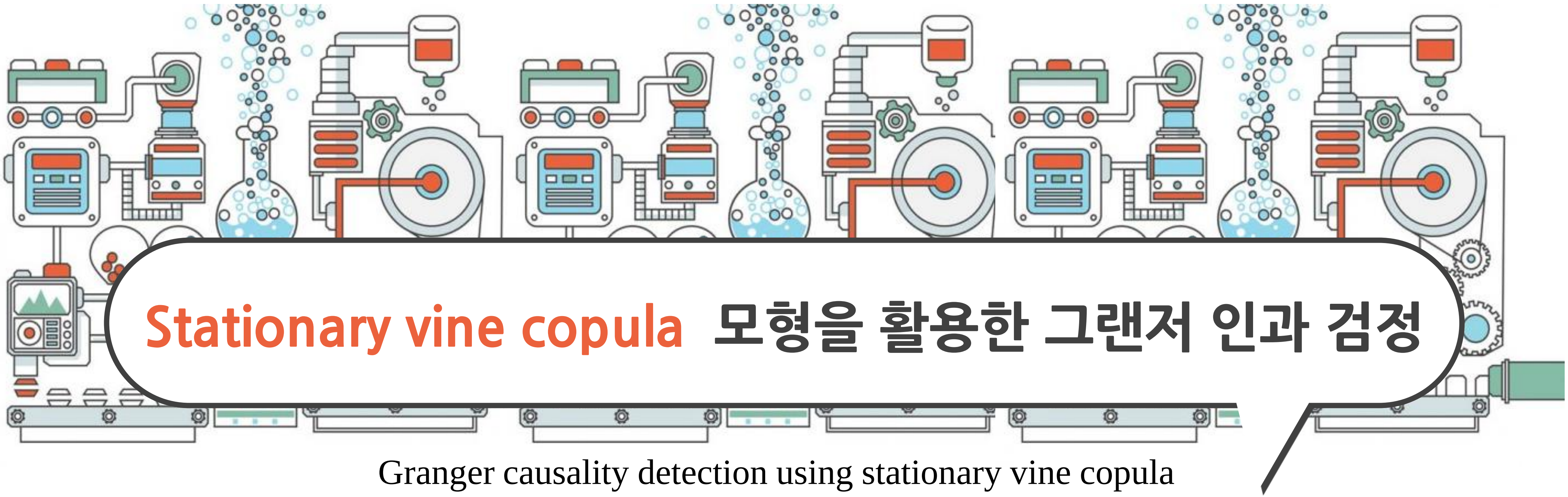
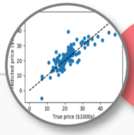


Stationary vine copula 모형을 활용한 그랜저 인과 검정

Granger causality detection using stationary vine copula

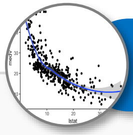
연구 목적

하나의 시계열 관측값이 다른 시계열의 과거 관측값에 영향을 받는다면
두 시계열 사이에는 그랜저 인과관계(Granger Causality)가 있다고 말함



VAR 모형 기반 GC검정

- **벡터자기회귀모형(VAR)활용**
  선형결합 구조를 가정
- **비선형 인과 관계 탐지 제약**



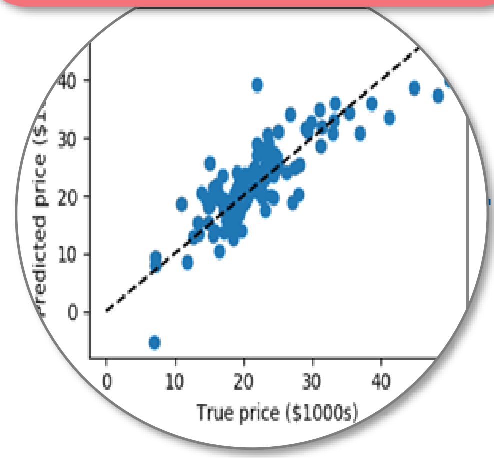
비모수회귀 기반 GC검정

- **Nonparametric regression 방법 활용**
  VAR 모형 기반 검정에 비해 nonlinear한 그랜저 인과를 탐지할 수 있는 장점이 있음
- **차원의 저주**
  모형화하는 함수의 차원이 증가하면 필요한 자료의 수가 기하급수적으로 증가
- **Smoothing parameter 추정의 어려움**
  모수의 개수 및 선택의 문제 등

연구 목적

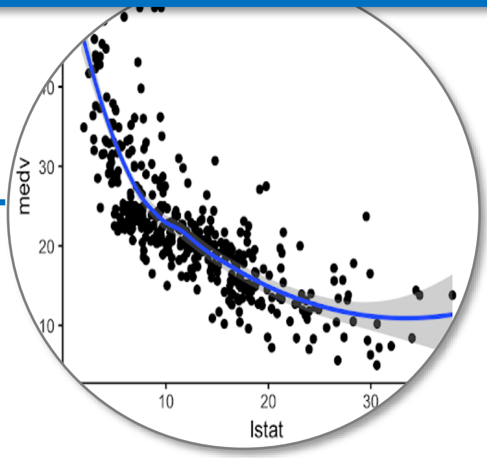
Parametric approach

VAR 모형 기반 GC검정



Non-parametric approach

비모수회귀 기반 GC검정

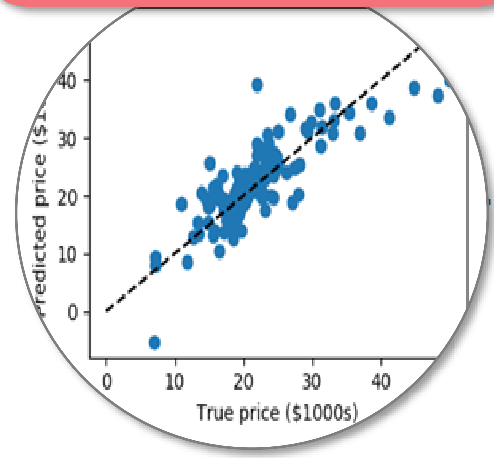


Flexible한 semi-parametric time series model을 활용한 그랜저 인과 검정 제안

연구 목적

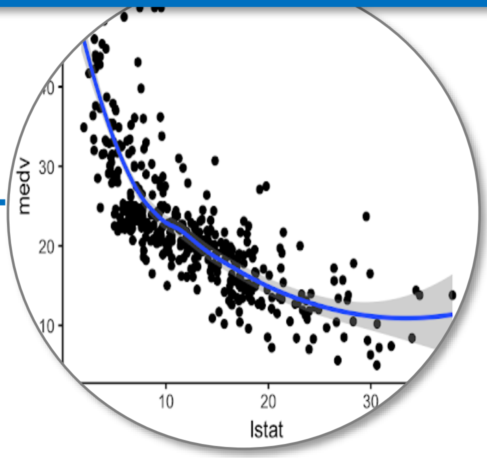
Parametric approach

VAR 모형 기반 GC검정



Non-parametric approach

비모수회귀 기반 GC검정



Stationary vine copula model을 활용한 그랜저 인과 검정 제안

배경지식

Copula

Sklar's theorem, X be a $d-$dimensional random vector and $F_i(x_i) \sim U(0,1)$ for $i = 1, \ldots, d$, function C called copula and c called copula density

$$F(x_1, x_2, \ldots, x_d) = \underbrace{C}_{ⓐ}\underbrace{(F_1(x_1), \ldots, F_1(x_1))}_{ⓑ}$$

$$C(F_1(x_1), \ldots, F_1(x_1)) \qquad C(F_1(x_1), \ldots, F_1(x_1))$$

ⓐ Copula component

ⓑ Marginal component

배경 지식

Copula

Sklar's theorem, X be a $d-$dimensional random vector and $F_i(x_i) \sim U(0,1)$ for $i = 1, \ldots, d$, function C called copula and c called copula density

$$F(x_1, x_2, \ldots, x_d) = C(\underbrace{F_1(x_1), \ldots, F_1(x_1)}_{\text{ⓑ Marginal component}})$$

Marginal component

Empirical Cumulative Distribution Function transform

$$\hat{F}(x) = \frac{1}{n+1} \sum_{i=1}^{n} 1_{\{x_i \leq x\}} \text{ , for all } x = (x_1, \ldots, x_n)$$

배경 지식

Copula

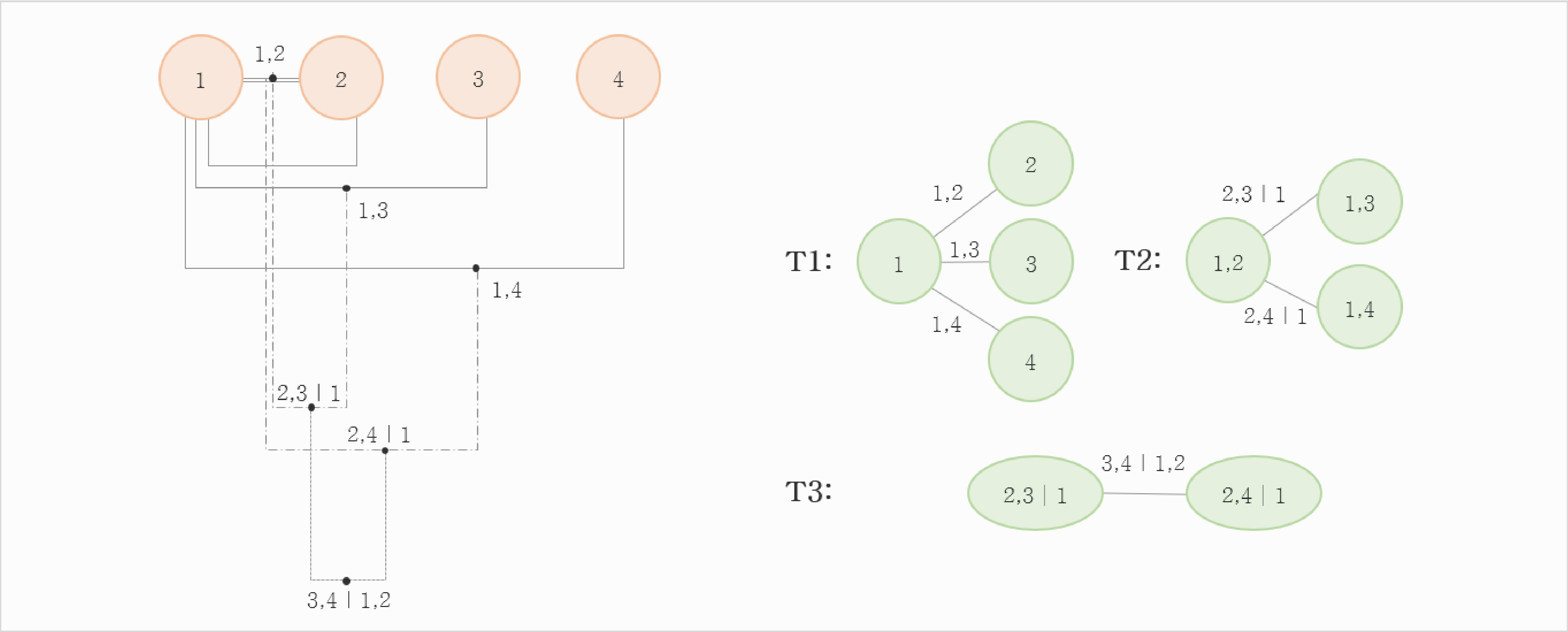
Sklar's theorem, X be a $\mathrm{d}-$ dimensional random vector and $\mathrm{F_i(x_i)} \sim \mathrm{U}(0,1)$ for $\mathrm{i} = 1, \dots, \mathrm{d}$, function C called copula and c called copula density

$$F(x_1, x_2, \dots, x_d) = \underbrace{C}_{ⓐ}(F_1(x_1), \dots, F_1(x_1))$$

Copula 함수의 추정만으로 다변량 결합 분포 함수의 모델링 가능

배경 지식

Copula ▶ 다변량 Copula ▶ Vine Copula Graph 이론을 활용한 다변량 copula 구조 분류



Vine copula structure 예시

배경 지식

Copula ▶ 다변량 Copula ▶ Vine Copula ▶ Stationary Vine Copula 시계열 특성을 반영

예) 정상성을 만족하는 Markov order p가 1인 이변량 시계열 $\{X_t, Y_t\}$에 대하여

1 $\{X_t, X_{t-1}, Y_t, Y_{t-1}\}$**의 4차원 copula 함수를 사용하여 모델링**

2 **정상성 조건**

ⓐ 시점에 상관없이 **주변분포는 동일** 함

$$X_t, \stackrel{d}{=} X_{t-1} \qquad Y_t, \stackrel{d}{=} Y_{t-1}$$

ⓑ 시점에 상관없이 관련된 **copula는 동일** 함

$$\begin{array}{cccc} & X_t & X_{t-1} & \\ c_t(X_t, Y_t) & | & | & c_{t-1}(X_{t-1}, Y_{t-1}) \\ & Y_t & Y_{t-1} & \end{array}$$

$$c_t(\cdot,\cdot) = c_{t-1}(\cdot,\cdot)$$

제안방법

제안 방법

GC 통계량

GC 통계량 추정

GC 검정

H_0: Y로부터의 X에 대한 그랜저 인과관계가 없다. H_1: not H_0

정상성을 만족하는 Markov order 1인 이변량 시계열 $\{X_t, Y_t\}$에 대한 GC 검정을 기준으로 설명

$$GC_{Y\to X}^{mean} = \log\left(\frac{E(X_t - E[X_t|X_{t-1}])^2}{E(X_t - E[X_t|X_{t-1}, Y_{t-1}])^2}\right)$$

가설	제곱합의 기댓값	통계량($GC_{Y\to X}^{mean}$)
귀무가설($Y \nrightarrow X$)하에서	분자, 분모의 제곱합이 동일	0
대립가설($Y \to X$)하에서	분모의 제곱합이 더 작아짐	0보다 큰 값

제안 방법

GC 통계량

GC 통계량 추정

GC 검정

H_0: Y로부터의 X에 대한 그랜저 인과관계가 없다. H_1: not H_0

$$GC_{Y\to X}^{mean} = \log\left(\frac{E(X_t - E[X_t|X_{t-1}])^2}{E(X_t - E[X_t|X_{t-1}, Y_{t-1}])^2}\right)$$

단계 1) 시점에 따라 자료 분할 - 학습 및 검증 자료

전체 자료의 길이 T $\Rightarrow$ 학습 자료의 시점 $1 \sim T^*$ + 검증 자료의 시점 $(T^*+1) \sim T$

- 학습자료 - 내부의 조건부기댓값을 추정하는 모형을 훈련
- 검증자료 - 외부의 기댓값을 추정

제안 방법

GC 통계량

GC 통계량 추정

GC 검정

H_0: Y로부터의 X에 대한 그랜저 인과관계가 없다.　　H_1: not H_0

단계 2) 학습 자료를 활용하여 모형 적합

$\{x_t\}$, $\{(x_t, y_t)\}$의 각 학습 자료를 사용하여 모형 적합 S_X, $S_{X,Y}$

단계 3) 검증 자료를 활용하여 표본 통계량 계산

$$\widehat{GC}^{mean}_{Y \to X} = \ln \left[\frac{(T - T^* + 1)^{-1} \sum_{t = T^* + 1}^{T} \left(x_t - \hat{E}[X_t \mid X_{t-1} = x_{t-1}] \right)^2}{(T - T^* + 1)^{-1} \sum_{t = T^* + 1}^{T} \left(x_t - \hat{E}[X_t \mid X_{t-1} = x_{t-1}, Y_{t-1} = y_{t-1}] \right)^2} \right]$$

제안 방법

GC 통계량

GC 통계량 추정

GC 검정

H_0: Y로부터의 X에 대한 그랜저 인과관계가 없다. H_1: not H_0

단계 3) 검증 자료를 활용하여 표본 통계량 계산

$$\hat{E}[X_t \mid X_{t-1} = x_{t-1}] = \frac{1}{N}\sum_{i=1}^{N} X_{t,S_X}^{(i)}$$

$$\hat{E}[X_t \mid X_{t-1} = x_{t-1},\ Y_{t-1} = y_{t-1}] = \frac{1}{N}\sum_{i=1}^{N} X_{t,S_{X,Y}}^{(i)}$$

$X_{t,S_X}^{(i)}, X_{t,S_{X,Y}}^{(i)}$: 적합 모형 S_X, $S_{X,Y}$을 통해 추정된 조건부 분포로부터 **시뮬레이션 예측값**

- $t = T^* + 1, \ldots, T$ 시점에서의 N개의 시뮬레이션 예측값을 생성
- 생성된 N개의 **시뮬레이션 예측값의 평균**을 분모, 분자의 조건부기댓값의 추정치로사용

제안 방법

GC 통계량

GC 통계량 추정

GC 검정

H_0: Y로부터의 X에 대한 그랜저 인과관계가 없다. H_1: not H_0

단계 3) 검증 자료를 활용하여 표본 통계량 계산

$$\widehat{GC}^{mean}_{Y\to X} = \ln\left[\frac{(T-T^*+1)^{-1}\sum_{t=T^*+1}^{T}\left(x_t - \boxed{\hat{E}[X_t \mid X_{t-1}=x_{t-1}]}\right.}{(T-T^*+1)^{-1}\sum_{t=T^*+1}^{T}\boxed{\hat{E}[X_t \mid X_{t-1}=x_{t-1}, Y_{t-1}=y_{t-1}]}}\right]$$

제안 방법

GC 통계량

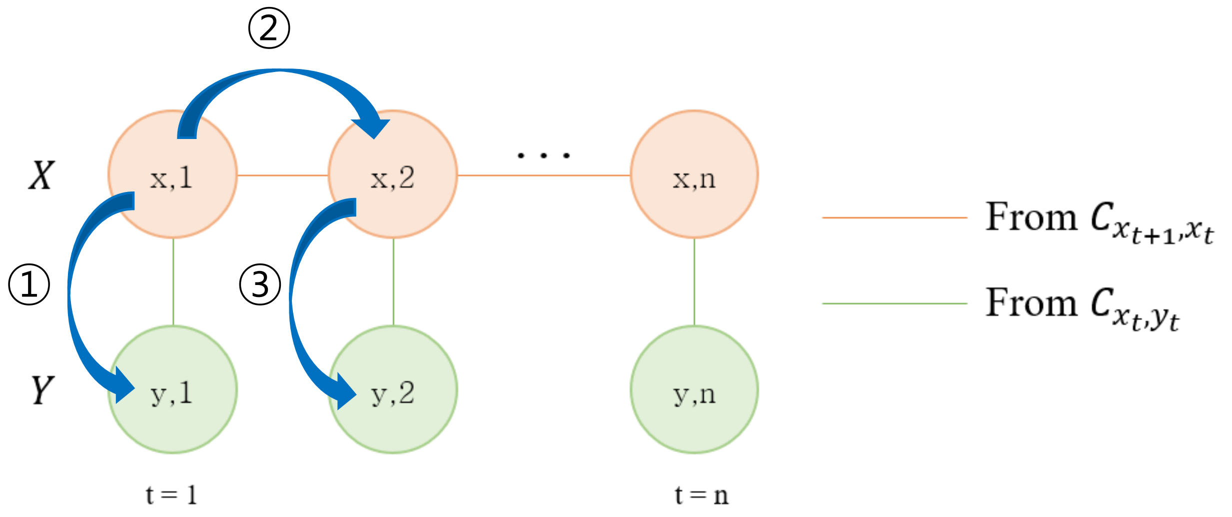
GC 통계량 추정

GC 검정

검정 통계량 분포 유도

검정 통계량의 귀무가설하의 이론적인 분포는 유도하기 매우 어려움
Smoothed local bootstrap 방법을 vine copula 기반으로 변형하여 활용

단계 1) 귀무가설($Y \nrightarrow X$)를 만족하는 B개의 bootstrap sample $\{(x_t^*, y_t^*)\}_{t=1}^{T}$ 생성



제안 방법

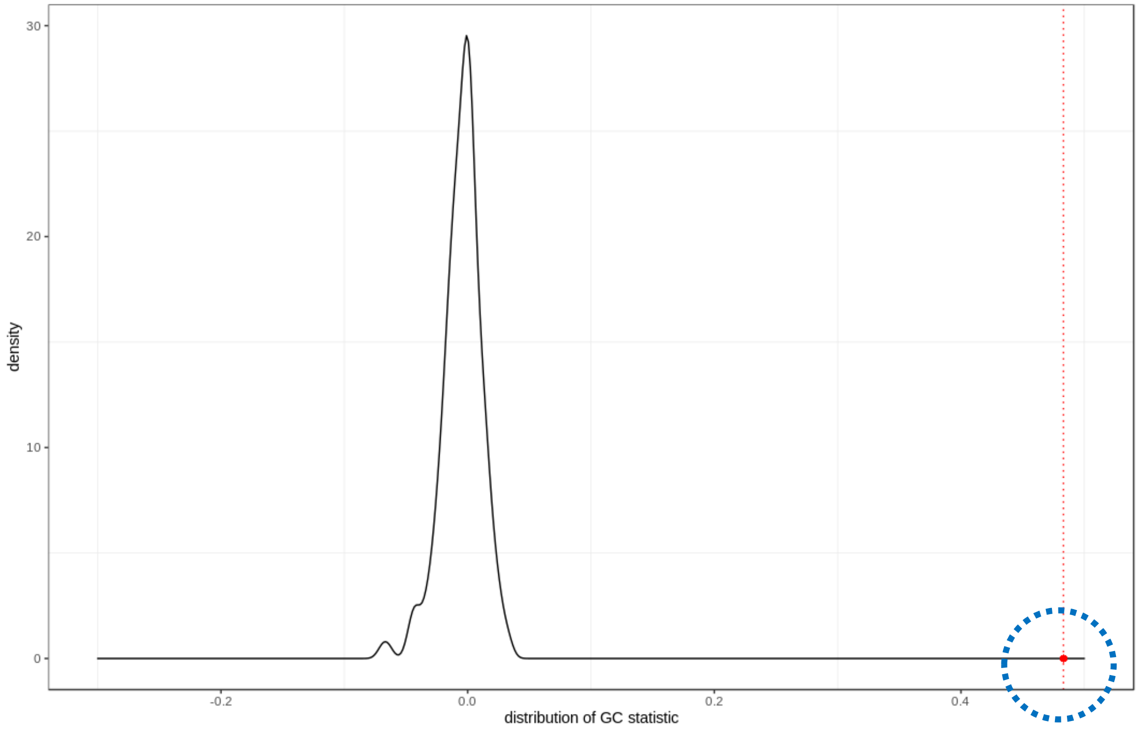
GC 통계량

GC 통계량 추정

GC 검정

검정 통계량 분포 유도

단계 2) 각각의 bootstrap sample에서 $\widehat{GC}^{*\,mean}_{j\;Y \to X}$ 계산



Bootstrap으로 얻어진 귀무가설하의 통계량의 분포(검은 실선)

대립가설하에서 생성된 하나의 자료에 대한 GC measure 추정치(빨간 점)와

제안 방법

GC 통계량

GC 통계량 추정

GC 검정

H_0: Y로부터의 X에 대한 그랜저 인과관계가 없다. H_1: not H_0

단계 3) 유의확률 계산 및 가설 검정

$$p^* = B^{-1}\sum_{j=1}^{B} I\left(\widehat{GC}_{j\,Y\to X}^{*\,mean} > \widehat{GC}_{Y\to X}^{mean}\right)$$

Reject H_0 when $p^* < \alpha$

시뮬레이션

시뮬레이션

정상성을 만족하는 **Markov order** 가 **1**인 이변량 시계열 $\{X_t, Y_t\}$에 대해 **유의수준 0.05**에서 **1000번** 반복 시뮬레이션 진행

Hyperparameter

01	02	03
N	B	*Split ratio*
# of i.i.d. replicates	# of Bootstrap sample	Split ratio
200	200	0.5

시뮬레이션

Size 측정 모형

- DGP S1 $Y_t = 0.5Y_{t-1} + \epsilon_t \quad X_t = 0.5X_{t-1} + \eta_t$
- DGP S2 $Y_t = 0.5Y_{t-1} + \epsilon_t \quad X_t = |X_{t-1}|^{0.8} + \eta_t$
- DGP S3 $Y_t = 0.5Y_{t-1} + 0.5X_{t-1}^2 + \epsilon_t \quad X_t = 0.5X_{t-1} + \eta_t$

Power 측정 모형

- DGP P1 $Y_t = 0.5Y_{t-1} + \epsilon_t \quad X_t = 0.5X_{t-1} + 0.5Y_{t-1} + \eta_t$
- DGP P2 $Y_t = 0.5Y_{t-1} + \epsilon_t \quad X_t = 0.5X_{t-1} + 0.5Y_{t-1} + 0.5\sin(-2Y_{t-1}) + \eta_t$
- DGP P3 $Y_t = 0.5Y_{t-1} + \epsilon_t \quad X_t = 0.5X_{t-1} + 0.5Y_{t-1}^2 + \eta_t$

시뮬레이션

비교대상방법

Svine	본 논문에서 제안한 방법
KLH	Kim, Lee and Hwang (2020)에서 제안한 방법
VAR	VAR 모형 기반의 GC 검정 방법
DP	Diks and Panchenko (2006)에서 제안한 방법
ST	Song and Taamouti (2018)에서 제안한 방법

※ ST 방법의 구현의 어려움으로, ST방법에 대한 시뮬레이션 결과는 논문에 제시된 결과를 그대로 가져왔음

결과

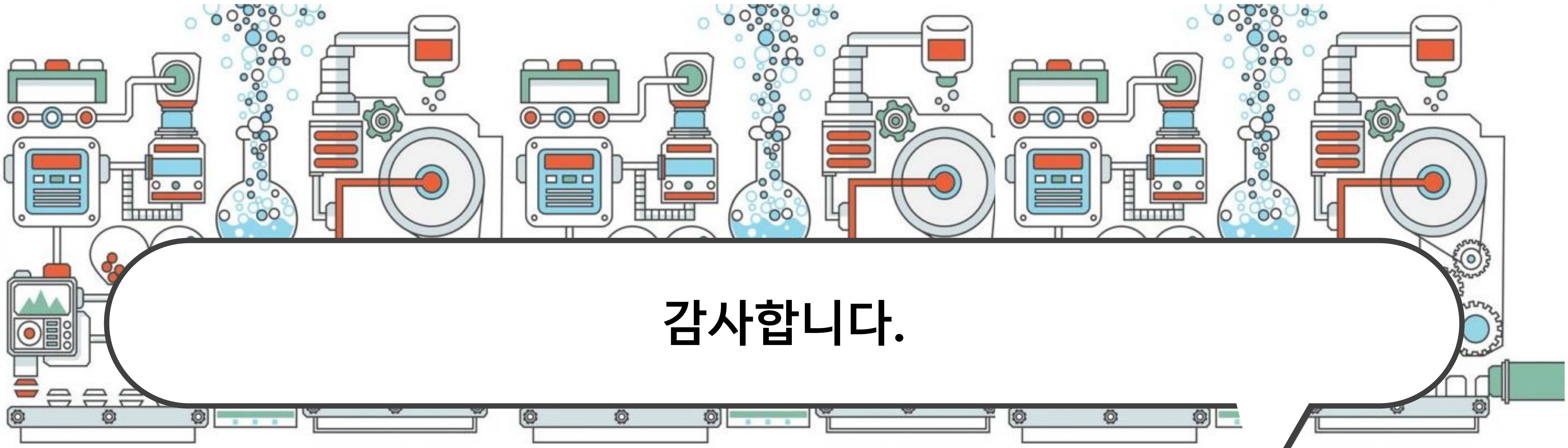
Size 측정 모형

Model	Method		T2 50	100	200	400
S1	Svine		.057	.042	.048	.060
	KLH		.064	.063	.046	.052
	VAR		.059	.047	.052	.051
	DP	$\epsilon=1$	.018	.014	.013	.019
		$\epsilon=1.5$	.017	.014	.020	.035
	ST	$\delta=0.6$	.046	–	.052	–
		$\delta=0.8$	.048	–	.050	–
S2	Svine		.066	.052	.046	.056
	KLH		.063	.057	.041	.050
	VAR		.054	.055	.042	.053
	DP	$\epsilon=1$	.013	.015	.015	.024
		$\epsilon=1.5$	.024	.027	.018	.038
	ST	$\delta=0.6$	.048	–	.048	–
		$\delta=0.8$	.052	–	.046	–
S3	Svine		.076	.091	.064	.059
	KLH		.057	.037	.058	.052
	VAR		.066	.044	.043	.045
	DP	$\epsilon=1$	.010	.014	.011	.020
		$\epsilon=1.5$	.014	.014	.026	.028
	ST	$\delta=0.6$	.064	–	.055	–
		$\delta=0.8$	.062	–	.054	–

결과

Power 측정 모형

Model	Method		T2			
			50	100	200	400
P1	Svine		.637	.897	.979	.999
	KLH		.935	.999	1.00	1.00
	VAR		.942	.999	1	1
	DP	$\epsilon=1$	.348	.727	.757	.773
		$\epsilon=1.5$	.524	.814	.628	.949
	ST	$\delta=0.6$	.232	–	.630	–
		$\delta=0.8$	.234	–	.664	–
P2	Svine		.492	.730	.936	.995
	KLH		.700	.914	.995	1.00
	VAR		.820	.987	.999	1
	DP	$\epsilon=1$	.210	.552	.858	.559
		$\epsilon=1.5$	.343	.649	.796	.588
	ST	$\delta=0.6$	.210	–	.752	–
		$\delta=0.8$	.212	–	.776	–
P3	Svine		.306	.495	.793	.965
	KLH		.234	.236	.234	.241
	VAR		.338	.367	.345	.348
	DP	$\epsilon=1$	.489	.818	.593	.978
		$\epsilon=1.5$	.547	.858	.554	.964
	ST	$\delta=0.6$	.398	–	.982	–
		$\delta=0.8$	.390	–	.950	–



감사합니다.

Granger causality detection using stationary vine copula