EBERHARD KARLS UNIVERSITÄT TÜBINGEN

WIRTSCHAFTS- UND SOZIALWISSENSCHAFTLICHE FAKULTÄT

Transfer Entropy Estimation Based on Smoothed Quantile Regressions

Johannes Bleher and Thomas Dimpfl

1 Motivation

In order to estimate relative entropy measures like mutual information or transfer entropy, estimates for joint densities are necessary.

Mutual Information

$$I(X,Y,Z) = \iiint_{\mathbb{R}^3} f_{X,Y,Z}(x,y,z) \log\left(\frac{f_{X,Y,Z}(x,y,z)}{f_X(x)f_Y(y)f_Z(z)}\right) dz dx dy$$
$$= \mathbb{E}\left[\log\left(\frac{f_{X,Y,Z}(x,y,z)}{f_X(x)f_Y(y)f_Z(z)}\right)\right]$$

Transfer Entropy Information

$$\hat{T}_{X\to Y} = \mathbb{E}\left[\log\left(\frac{f_{Y_t|X_{t-1},Y_{t-1}}(y_t \mid x_{t-1}, y_{t-1})}{f_{Y_t|Y_{t-1}}(y_t \mid y_{t-1})}\right)\right]$$

Usually, these measures are calculated by discretizing continuous variables.

Ideas of this paper

Calculate Transfer Entropy without throwing away information i.e. without discretization.

Measure information flow from data by ...

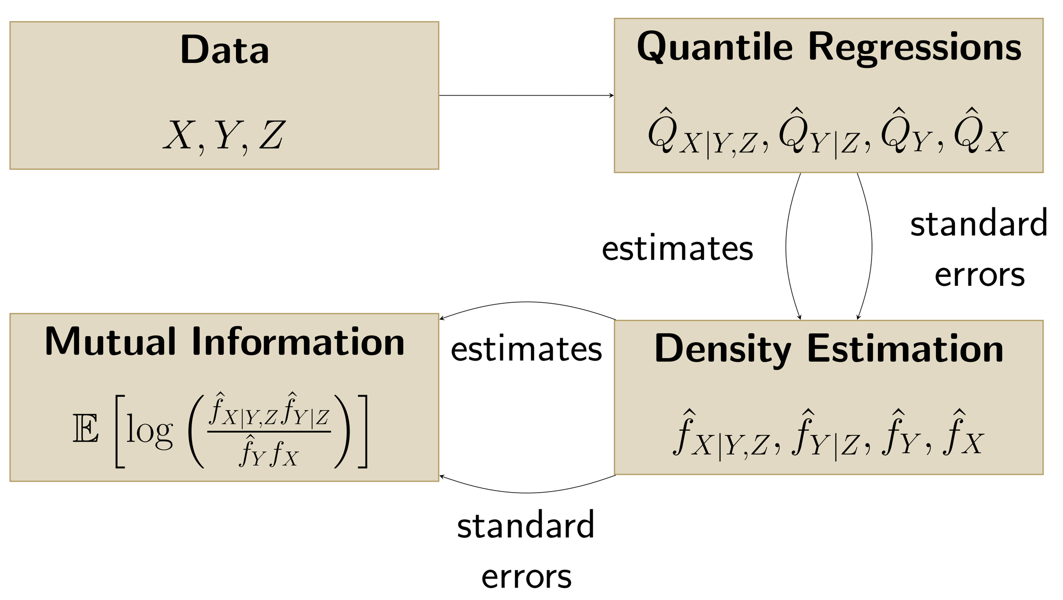
- ... using as much information as possible.
- ... using as few assumptions as necessary.
- ... keeping computational cost low.
- ... utilizing a testable framework.

Interesting for...

- ... variable selection.
- ... causality analysis.

2 The Theoretical Framework

In order to estimate relative entropy measures, we use quantile regressions on multidimensional (time) series data. We smooth the resulting estimated conditional quantile functions to calculate conditional densities. In a last step, the relative entropy measure is calculated.



The same process can be applied to time series data and transfer entropy.

3 Quantile Regression

$$F_j = \min_{\boldsymbol{\theta}\in\mathbb{R}} \sum_{i=1}^{N} \rho_{\tau_j}(y_i - \boldsymbol{x}_i'\boldsymbol{\theta}_j)$$

where $\rho_{\tau_j}(u) = u\,(\tau_j - \mathbb{1}(u<0))$ with $\mathbb{1}(\cdot)$ as the indicator function (essentially a Heaviside step function). Fitted quantiles can then be calculated by

$$\boldsymbol{Q}_{\boldsymbol{y}}(\tau_j \mid \boldsymbol{X}) = \boldsymbol{X}'\hat{\boldsymbol{\theta}}(\tau_j)$$

Problem: Joint standard error of several QR on same data? **Solution:** GMM asymptotics

---

Reformulate Quantile Regression as GMM problem by substituting the Heaviside step function $\mathbb{1}(u<0)$ with the sigmoid function $\mathbf{1}(u) = (1+e^u)^{-1}$:

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmin}}\, \mathbf{g}_N(\boldsymbol{\theta})'\mathbf{W}\mathbf{g}_N(\boldsymbol{\theta})$$

Collect empirical sampling errors of moment conditions

$$g_l(\tau_j, \boldsymbol{\theta}_j) = \mathbb{E}[X_l\,(\tau_j - \mathbf{1}(U<0)) - X_l\,U\,\mathbf{1}(U=0)]$$

$$\mathbf{g}_N(\boldsymbol{\theta}_j) = \begin{pmatrix} \frac{1}{N}\sum_{i=1}^N g_{1i}(\tau_j, \hat{\boldsymbol{\theta}}_j) \\ \vdots \\ \frac{1}{N}\sum_{i=1}^N g_{Ki}(\tau_j, \hat{\boldsymbol{\theta}}_j) \end{pmatrix}$$

Joint asymptotic normality of quantile regression estimates

$$\sqrt{N}(\mathrm{vec}(\hat{\boldsymbol{\theta}}) - \mathrm{vec}(\boldsymbol{\theta})) \sim \mathcal{N}\left(0, \mathrm{Avar}(\hat{\boldsymbol{\theta}})\right)$$

with

$$\widehat{\mathrm{Avar}}(\hat{\boldsymbol{\theta}}) = (\hat{\boldsymbol{D}}'\boldsymbol{W}\hat{\boldsymbol{D}})^{-1}\,\hat{\boldsymbol{D}}'\boldsymbol{W}\ \hat{\boldsymbol{S}}\ \boldsymbol{W}\hat{\boldsymbol{D}}\,(\hat{\boldsymbol{D}}'\boldsymbol{W}\hat{\boldsymbol{D}})^{-1}$$

4 Density Estimation by Smoothing Conditional Quantile Functions

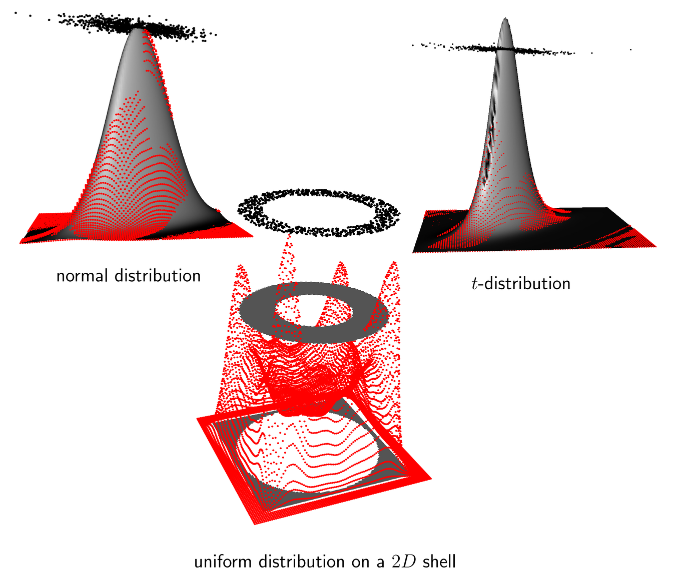
Given a point $P = (\hat{\boldsymbol{\theta}}\mathbf{x}_0, y_0)$

$$\hat{\gamma} = (\boldsymbol{Z}_P'\boldsymbol{W}_P\boldsymbol{Z}_P)^{-1}\boldsymbol{Z}_P'\boldsymbol{W}_P\boldsymbol{\tau}$$

where

- $\underset{Q\times p}{\boldsymbol{Z}_P} = \left(\boldsymbol{\iota}, \hat{\boldsymbol{\theta}}\mathbf{x}_0 - y_0, (\hat{\boldsymbol{\theta}}\mathbf{x}_0 - y_0)^2, \ldots, (\hat{\boldsymbol{\theta}}\mathbf{x}_0 - y_0)^p\right)$
- $\boldsymbol{\iota}$ as the column vector of ones
- $\boldsymbol{W}_P = \mathrm{diag}\left(h^{-1}K(\frac{\hat{\boldsymbol{\theta}}\mathbf{x}_0 - y_0}{h})\right)$
- $K(\cdot)$ denotes the weight function
- h is the approximately optimal bandwidth

$\Rightarrow \hat{\gamma}_1$ is a biased density estimate at point P.



Standard Errors of Densities via the Delta-Method

$$\lim_{N\to\infty} \sqrt{Qh^3}\frac{\left(\hat{f}(\hat{\boldsymbol{\theta}}) - f\right)}{\mathrm{vec}(\hat{\boldsymbol{\theta}}) - \mathrm{vec}(\boldsymbol{\theta})} \approx \left.\frac{\partial \hat{f}}{\partial \bar{\boldsymbol{\theta}}}\right|_{\bar{\boldsymbol{\theta}}=\hat{\boldsymbol{\theta}}}.$$

follows

$$\mathrm{Avar}\left(\hat{\gamma}_1\left(\boldsymbol{\theta}\right)\right) \approx \frac{1}{NQh^3}\mathrm{vec}\left(\mathbf{H}\right)'\mathrm{Avar}(\hat{\boldsymbol{\theta}})\,\mathrm{vec}\left(\mathbf{H}\right)$$

- $\mathbf{H}$ derivative of density estimate wrt $\hat{\boldsymbol{\theta}}$
  - ✓ estimation uncertainty from QR
  - ✓ estimation of bandwidth
  - ✓ ghost points

5 Estimating and Testing Mutual Information and Transfer Entropy

Standard Errors of Mutual Information and Transfer Entropy via the Delta Method for each contribution C_i:

$$\lim_{N\to\infty}\sqrt{Q}\frac{\hat{C}_i(\hat{\boldsymbol{\theta}}) - C_i - C_i^*}{\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}} = \lim_{N\to\infty}\frac{1}{\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}}\sqrt{Q}\log\left(\frac{h_{X|Y,Z}^{-3/2}(\hat{f}_{X|Y,Z} - f_{X|Y,Z})h_{Y|Z}^{-3/2}(\hat{f}_{Y|Z} - f_{Y|Z})}{h_X^{-3/2}(\hat{f}_X - f_X)h_Y^{-3/2}(\hat{f}_Y - f_Y)}\right) = \left.\frac{\partial \hat{C}_{x_i,y_i,z_i}}{\partial\bar{\theta}_{lm}}\right|_{\bar{\theta}_{lm}=\hat{\theta}_{lm}},$$

For each contribution (approximately)

$$\hat{C}_i\left(\hat{\boldsymbol{\theta}}\right) + C_i^* \sim \mathcal{N}\left(C_{\mathbf{i}}, \frac{1}{QN}\mathrm{vec}\left(\boldsymbol{\Upsilon}_{\mathbf{i}}\right)'\mathrm{Avar}(\hat{\boldsymbol{\theta}})\,\mathrm{vec}\left(\boldsymbol{\Upsilon}_{\mathbf{i}}\right)\right).$$

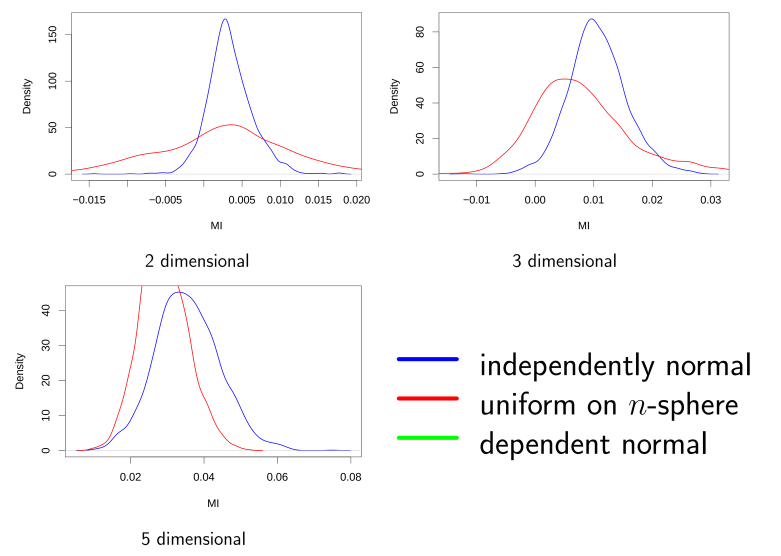
This yields the variance estimator for MI

$$\mathrm{var}\left(\hat{I}_{X,Y,Z}\right) = \frac{1}{QN}\left[\frac{1}{N}\sum_{i=1}^{N}\mathrm{vec}\left(\boldsymbol{\Upsilon}_i\right)\right]'\mathrm{Avar}(\hat{\boldsymbol{\theta}})\left[\frac{1}{N}\sum_{j=1}^{N}\mathrm{vec}\left(\boldsymbol{\Upsilon}_j\right)\right].$$

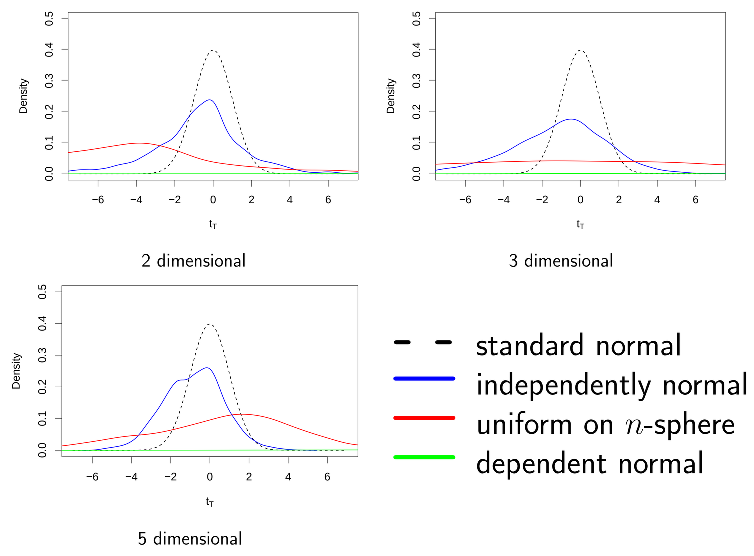
Analogously, the variance estimator for TE can be worked out. With this a z-score test statistic can be formulated

$$t_I = \frac{\hat{I} - I_0}{\sqrt{\mathrm{var}(\hat{I})}} \sim \mathcal{N}\left(0,1\right).$$

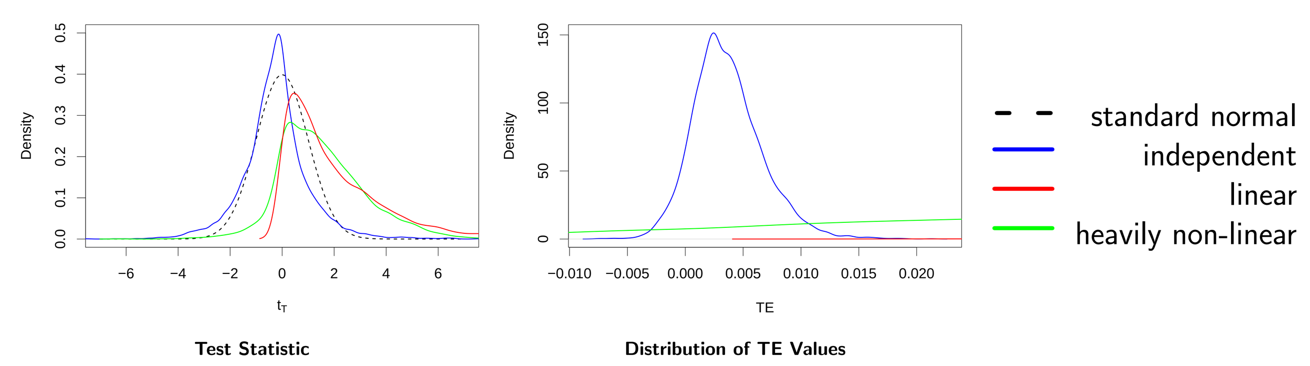
Simulation Results MI values



Simulation Results Test Statistics



Simulation Results – $TE_{X\to Z}$



6 Conclusions

- We derived the joint distribution of multiple quantile regressions on the same sample.
- Worked out the conditional density estimates using smoothed quantile regressions.
- Estimated MI and Transfer Entropy.
- Explored test statistics.
- Two empirical applications.

Johannes Bleher
University of Tübingen
johannes.bleher@uni-tuebingen.de

Thomas Dimpfl
University of Tübingen
thomas.dimpfl@uni-tuebingen.de