Introduction
000000000

Copula Density Estimator
0000000

Bandwidth Selection
0000

Numerical Experiments
0000000

Conclusion

Acknowledgements

Recursive Non-parametric Estimation of the Copula Density

Vladimir Yashin
supervised by Geoffrey Decrouez

Higher School of Economics

9 June 2018

Marginal and joint PDFs

What is the joint probability distribution of wage and age?

Let us present both random variables in terms of univariate CDFs

$$F_w(w) := \mathbb{P}(W \leqslant w), \quad \log(w) \sim \mathcal{N}(\mu_w, \sigma_w^2)$$
$$F_a(a) := \mathbb{P}(A \leqslant a), \quad a \sim \mathcal{N}(\mu_a, \sigma_a^2).$$

Then, the joint distribution has the following form

$$F(w, a) := \mathbb{P}(W \leqslant w, A \leqslant a).$$

Marginal and joint PDFs

What is the joint probability distribution of wage and age?

Let us present both random variables in terms of univariate CDFs

$$F_w(w) := \mathbb{P}(W \leqslant w), \quad \log(w) \sim \mathcal{N}(\mu_w, \sigma_w^2)$$
$$F_a(a) := \mathbb{P}(A \leqslant a), \quad a \sim \mathcal{N}(\mu_a, \sigma_a^2).$$

Then, the joint distribution has the following form

$$F(w, a) := \mathbb{P}(W \leqslant w, A \leqslant a).$$

Marginal and joint PDFs

What is the joint probability distribution of wage and age?

Let us present both random variables in terms of univariate CDFs

$$F_w(w) := \mathbb{P}(W \leqslant w), \quad \log(w) \sim \mathcal{N}\big(\mu_w, \sigma_w^2\big)$$
$$F_a(a) := \mathbb{P}(A \leqslant a), \quad a \sim \mathcal{N}\big(\mu_a, \sigma_a^2\big).$$

Then, the joint distribution has the following form

$$F(w, a) := \mathbb{P}(W \leqslant w, A \leqslant a).$$

Marginal and joint PDFs of independent samples
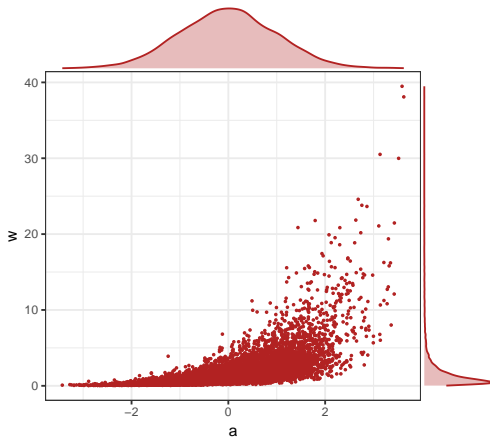

Joint PDF of random variables
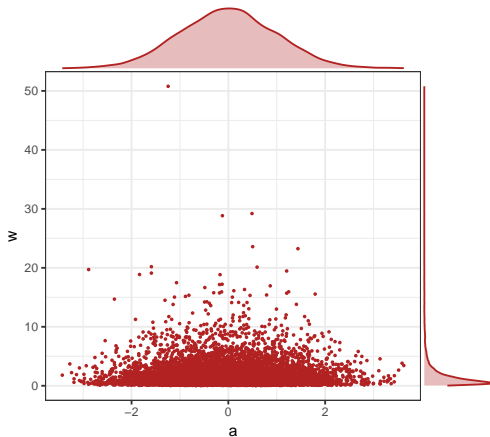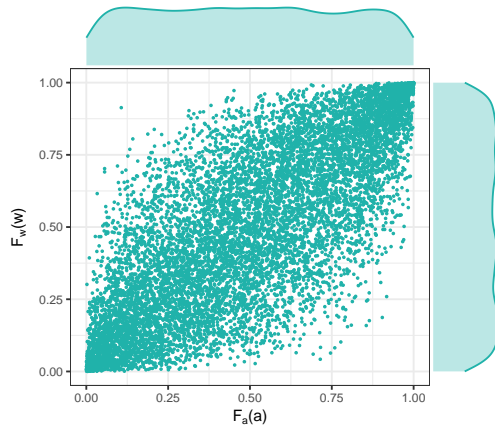
Joint PDF of marginal CDFs
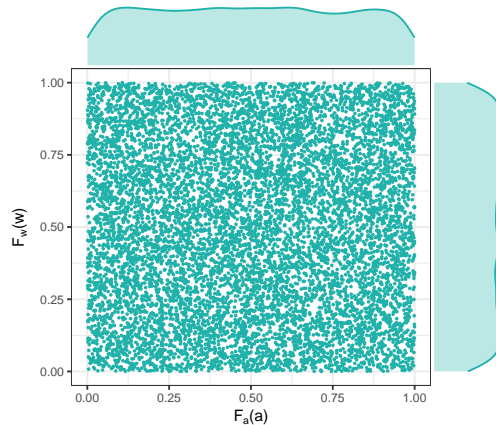
Marginal and joint PDFs of dependent samples ($\rho = 0.8$)

Joint PDF of random variables

Joint PDF of marginal CDFs

How to measure dependence between variables?

We wish to describe dependence structure between w and a. What to do?

1st idea: Pearson correlation coefficient (✗: non-linear dependence);

2nd idea: Spearman's rank correlation coefficient (✗: information is too general);

3rd idea: to consider joint PDF of marginals (✗: family of CDFs are rarely known).

Introduction
○○○●○○○○○

Copula Density Estimator
○○○○○○○

Bandwidth Selection
○○○○

Numerical Experiments
○○○○○○○

Conclusion

Acknowledgements

How to measure dependence between variables?

We wish to describe dependence structure between w and a. What to do?

1st idea: Pearson correlation coefficient (✗: non-linear dependence);

2nd idea: Spearman's rank correlation coefficient (✗: information is too general);

3rd idea: to consider joint PDF of marginals (✗: family of CDFs are rarely known).

How to measure dependence between variables?

We wish to describe dependence structure between w and a. What to do?

1st idea: Pearson correlation coefficient (✗: non-linear dependence);

2nd idea: Spearman's rank correlation coefficient (✗: information is too general);

3rd idea: to consider joint PDF of marginals (✗: family of CDFs are rarely known).

Introduction
○○○●○○○○○

Copula Density Estimator
○○○○○○○

Bandwidth Selection
○○○○

Numerical Experiments
○○○○○○○

Conclusion

Acknowledgements

How to measure dependence between variables?

We wish to describe dependence structure between w and a. What to do?

1st idea: Pearson correlation coefficient (✗: non-linear dependence);

2nd idea: Spearman's rank correlation coefficient (✗: information is too general);

3rd idea: to consider joint PDF of marginals (✗: family of CDFs are rarely known).

Let's use copulas

Copulas can address these problems

- ✓: can capture non-linear dependence;
- ✓: shows rich visual information of dependence structure;
- ✓: marginals might be of non-parametric families.

Let's use copulas

Copulas can address these problems

✓: can capture non-linear dependence;

✓: shows rich visual information of dependence structure;

✓: marginals might be of non-parametric families.

Let's use copulas

Copulas can address these problems

✓: can capture non-linear dependence;

✓: shows rich visual information of dependence structure;

✓: marginals might be of non-parametric families.

Let's use copulas

Copulas can address these problems

✓: can capture non-linear dependence;

✓: shows rich visual information of dependence structure;

✓: marginals might be of non-parametric families.

Let's use copulas

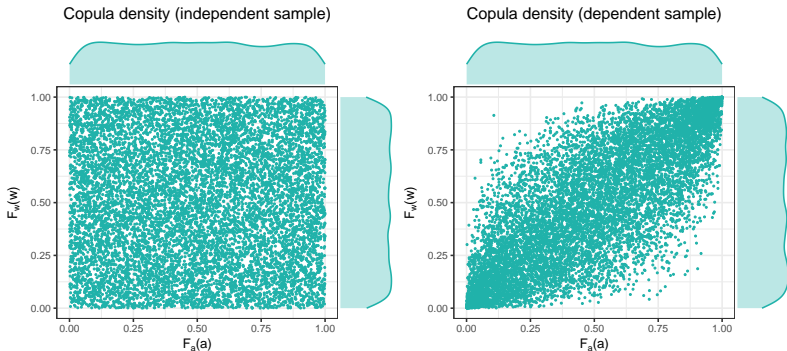Copulas can address these problems

- ✓: can capture non-linear dependence;
- ✓: shows rich visual information of dependence structure;
- ✓: marginals might be of non-parametric families.

What are copulas?

Copula is a multivariate CDF such that a marginal CDF for each variable is uniform.

What are copulas?

Copula is a multivariate CDF such that a marginal CDF for each variable is uniform.

Let's get technical

There exists[1] a unique d-variate copula $C : [0,1]^d \to [0,1]$ such that

$$F(X_1, \ldots, X_d) = C(F_1(X_1), \ldots, F_d(X_d)),$$

where $(X_1, \ldots, X_d)$ is a set of d random variables like age, wage, etc.

Therefore, the copula density

$$c(F_1(X_1), \ldots, F_d(X_d)) = \frac{f(X_1, \ldots, X_d)}{\prod_{j=d}^{d} f_j(X_j)}$$

---

[1]Sklar, 1959.

Let's get technical

There exists[1] a unique d-variate copula $C : [0, 1]^d \to [0, 1]$ such that

$$F(X_1, \ldots, X_d) = C\big(F_1(X_1), \ldots, F_d(X_d)\big),$$

where $(X_1, \ldots, X_d)$ is a set of d random variables like age, wage, etc.

Therefore, the copula density

$$c\big(F_1(X_1), \ldots, F_d(X_d)\big) = \frac{f(X_1, \ldots, X_d)}{\prod_{j=d}^d f_j(X_j)}$$

---

[1]Sklar, 1959.

Let's get technical

There exists[1] a unique d-variate copula $C : [0,1]^d \rightarrow [0,1]$ such that

$$F(X_1, \ldots, X_d) = C\big(F_1(X_1), \ldots, F_d(X_d)\big),$$

where $(X_1, \ldots, X_d)$ is a set of d random variables like age, wage, etc.

Therefore, the copula density

$$c\big(F_1(X_1), \ldots, F_d(X_d)\big) = \frac{f(X_1, \ldots, X_d)}{\prod_{j=d}^{d} f_j(X_j)}$$

---

[1]Sklar, 1959.

Introduction
○○○○○○○○●

Copula Density Estimator
○○○○○○○

Bandwidth Selection
○○○○

Numerical Experiments
○○○○○○○

Conclusion

Acknowledgements

Why old methods are bad?

The existing copula estimators are not on-line.

Therefore, it has to be entirely recalculated in case of persistent data flow.

Introduction
○○○○○○○○○

Copula Density Estimator
●○○○○○○

Bandwidth Selection
○○○○

Numerical Experiments
○○○○○○○

Conclusion

Acknowledgements

How to estimate copula density?

Recall that

$$c\big(F_1(X_1), \ldots, F_d(X_d)\big) = \frac{f(X_1, \ldots, X_d)}{\prod_{j=d}^{d} f_j(X_j)}.$$

First, we need to recursively estimate joint PDF f and each of marginal PDFs f_j.

Robbins-Monro procedure

A root of unknown function h using only noisy observations can be found as follows[2]

$$z_n = z_{n-1} - \alpha_n H(z_n, \varepsilon_n)$$

where $H(z_n, \varepsilon_n) \equiv (h(z_n) + \varepsilon_n)$ is a noisy observation and α_n is a sequence of positive numbers that converges to 0.

It was proved that z_n converges almost surely to z^* as $n \to \infty$, given $\mathbb{E}[\varepsilon] = 0$.

---

[2]Robbins and Monro, 1951.

Robbins-Monro procedure

A root of unknown function h using only noisy observations can be found as follows[2]

$$z_n = z_{n-1} - \alpha_n H(z_n, \varepsilon_n)$$

where $H(z_n, \varepsilon_n) \equiv (h(z_n) + \varepsilon_n)$ is a noisy observation and α_n is a sequence of positive numbers that converges to 0.

It was proved that z_n converges almost surely to $z^\star$ as $n \to \infty$, given $\mathbb{E}[\varepsilon] = 0$.

---

[2]Robbins and Monro, 1951.

Quantile estimation using Robbins-Monro procedure

To estimate a PDF we first need to estimate quantiles.

An estimator of u-th quantile can be derived from the Robbins-Monro procedure

$$x_n^u = x_{n-1}^u + \alpha_n\big(u - \mathbf{1}(X_{n-1} \leqslant x_{n-1}^u)\big),$$

where x_n^u is a u-th quantile estimator.

Introduction
000000000

Copula Density Estimator
0000000

Bandwidth Selection
0000

Numerical Experiments
0000000

Conclusion

Acknowledgements

Quantile estimation using Robbins-Monro procedure

To estimate a PDF we first need to estimate quantiles.

An estimator of u-th quantile can be derived from the Robbins-Monro procedure

$$x_n^u = x_{n-1}^u + \alpha_n\big(u - \mathbf{1}(X_{n-1} \leqslant x_{n-1}^u)\big),$$

where x_n^u is a u-th quantile estimator.

Introduction
000000000

Copula Density Estimator
0000●000

Bandwidth Selection
0000

Numerical Experiments
0000000

Conclusion

Acknowledgements

Smoothed recursive non-parametric estimation of marginal PDFs f_j

A smoothed adaptation of the quantile estimator[3]

$$f_{j,n}^{u_j} = (1 - 1/n) \cdot f_{j,n-1}^{u_j} + 1/n \cdot K_{h_{j,n}}^{u_j} \left(x_{n-1}^{u_j} - X_{n,j} \right)$$

$$a_{j,n}^{u_j} = \max \left[\mu_{j,n}, \min \left[f_{j,n}^{u_j}, \nu \ln(n+1) \right] \right]$$

$$x_n^{u_j} = x_{n-1}^{u_j} + \frac{1}{n a_{j,n}^{u_j}} \left[u_j - \mathbf{1} \left(X_{n,j} \leqslant x_{n-1}^{u_j} \right) \right],$$

where $K_{h_{j,n}}^{u_j} (\cdot)$ is a symmetric, univariate kernel with bandwidth $h_{j,n}$;
$\mu_{j,n}$ and ν are some positive constants.

---

[3]Definition and a proof of convergence in Amiri and Thiam, (2014)

Introduction
000000000

Copula Density Estimator
0000●00

Bandwidth Selection
0000

Numerical Experiments
0000000

Conclusion

Acknowledgements

Smoothed recursive non-parametric estimation of joint PDF f

A recursive estimator for the joint PDF defined as follows[4]

$$f_n^{\mathbf{u}} = (1 - 1/n) \cdot f_{n-1}^{\mathbf{u}} + 1/n \cdot \mathbf{K}_{b_n}^{\mathbf{u}} \left(x_{n-1}^{\mathbf{u}} - \mathbf{X}_n \right),$$

where now instead of u_j we have a grid of points $\mathbf{u}$ and multivariate data point $\mathbf{X}_n$; $\mathbf{K}_{b_n}^{\mathbf{u}}(\mathbf{X}) := \prod_{j=1}^d K_{b_n}^{u_j}(\mathbf{X}_j)$ where $K_{b_n}^{u_j}(\mathbf{X}_j)$ is a univariate kernel with bandwidth b_n.

---

[4]Definition and a proof of convergence in Mokkadem et al., (2009)

Some remarks on marginal and joint PDFs estimators

$$f_{j,n}^{u_j} = (1 - 1/n) \cdot f_{j,n-1}^{u_j} + 1/n \cdot K_{h_{j,n}}^{u_j}\left(x_{n-1}^{u_j} - X_{n,j}\right)$$
$$f_n^{\mathbf{u}} = (1 - 1/n) \cdot f_{n-1}^{\mathbf{u}} + 1/n \cdot \mathbf{K}_{b_n}^{\mathbf{u}}\left(x_{n-1}^{\mathbf{u}} - \mathbf{X}_n\right),$$

Remarks:

- The marginal estimator is just a special case of the joint estimator;

- The authors of the joint estimator didn't provide the estimation for quantiles $x^{\mathbf{u}}$;

- Therefore, the estimation of the quantile x^{u_j} has to be generalized ✓.

Introduction
000000000

Copula Density Estimator
0000000

Bandwidth Selection
0000

Numerical Experiments
0000000

Conclusion

Acknowledgements

Some remarks on marginal and joint PDFs estimators

$$f_{j,n}^{u_j} = (1 - 1/n) \cdot f_{j,n-1}^{u_j} + 1/n \cdot K_{h_{j,n}}^{u_j} \left(x_{n-1}^{u_j} - X_{n,j} \right)$$

$$f_n^{\mathbf{u}} = (1 - 1/n) \cdot f_{n-1}^{\mathbf{u}} + 1/n \cdot \mathbf{K}_{b_n}^{\mathbf{u}} \left(x_{n-1}^{\mathbf{u}} - \mathbf{X}_n \right),$$

Remarks:

- The marginal estimator is just a special case of the joint estimator;
- The authors of the joint estimator didn't provide the estimation for quantiles $x^{\mathbf{u}}$;
- Therefore, the estimation of the quantile x^{u_j} has to be generalized ✓.

Introduction
000000000

Copula Density Estimator
0000000

Bandwidth Selection
0000

Numerical Experiments
0000000

Conclusion

Acknowledgements

Some remarks on marginal and joint PDFs estimators

$$f_{j,n}^{u_j} = (1 - 1/n) \cdot f_{j,n-1}^{u_j} + 1/n \cdot K_{h_{j,n}}^{u_j} \left(x_{n-1}^{u_j} - X_{n,j} \right)$$
$$f_n^{\mathbf{u}} = (1 - 1/n) \cdot f_{n-1}^{\mathbf{u}} + 1/n \cdot \mathbf{K}_{b_n}^{\mathbf{u}} \left(x_{n-1}^{\mathbf{u}} - \mathbf{X}_n \right),$$

Remarks:

- The marginal estimator is just a special case of the joint estimator;
- The authors of the joint estimator didn't provide the estimation for quantiles $x^{\mathbf{u}}$;
- Therefore, the estimation of the quantile x^{u_j} has to be generalized ✓.

Introduction
○○○○○○○○○

Copula Density Estimator
○○○○○●○

Bandwidth Selection
○○○○

Numerical Experiments
○○○○○○○

Conclusion

Acknowledgements

Some remarks on marginal and joint PDFs estimators

$$f_{j,n}^{u_j} = (1 - 1/n) \cdot f_{j,n-1}^{u_j} + 1/n \cdot K_{h_{j,n}}^{u_j} \left(x_{n-1}^{u_j} - X_{n,j} \right)$$
$$f_n^{\mathbf{u}} = (1 - 1/n) \cdot f_{n-1}^{\mathbf{u}} + 1/n \cdot \mathbf{K}_{b_n}^{\mathbf{u}} \left(x_{n-1}^{\mathbf{u}} - \mathbf{X}_n \right),$$

Remarks:

- The marginal estimator is just a special case of the joint estimator;
- The authors of the joint estimator didn't provide the estimation for quantiles $x^{\mathbf{u}}$;
- Therefore, the estimation of the quantile x^{u_j} has to be generalized ✓.

Recursive non-parametric copula density estimator

Given

$$f_{j,n}^{u_j} = (1 - 1/n) \cdot f_{j,n-1}^{u_j} + 1/n \cdot K_{h_{j,n}}^{u_j}\left(x_{n-1}^{u_j} - X_{n,j}\right),$$
$$f_n^{\mathbf{u}} = (1 - 1/n) \cdot f_{n-1}^{\mathbf{u}} + 1/n \cdot \mathbf{K}_{b_n}^{\mathbf{u}}\left(x_{n-1}^{\mathbf{u}} - \mathbf{X}_n\right),$$

$c_n(\cdot)$ defined as follows

$$c_n = \frac{f_n^{\mathbf{u}}}{\prod_{j=1}^d f_{j,n}^{u_j}}$$

to the best of our knowdlenge, is the first recursive non-parametric estimator of the copula density $c(\cdot)$

$$c = \frac{f}{\prod_{j=d}^d f_j}.$$

How to select the bandwidths?

We have two sets of bandwidths h_j for each of marginals and b for the joint.

Two groups of bandwidths:

- Data-independent;
- Data-dependent.

How to select the bandwidths?

We have two sets of bandwidths h_j for each of marginals and b for the joint.

Two groups of bandwidths:

- Data-independent;
- Data-dependent.

Introduction
000000000

Copula Density Estimator
0000000

Bandwidth Selection
0●00

Numerical Experiments
0000000

Conclusion

Acknowledgements

Constant bandwidth

The simplest data-independent case: h_j and b are the same[5] and calculated as follows

$$b_n = h_{j,n} = n^{-\tau},$$

where $\tau \in (0, 1/2)$.

---

[5]As proposed in Robinson, (1975) and Holst, (1987) for non-smooth recursive algorithms.

Silverman's bandwidth

h_j and b are still the same but let's use aggregated information about data[6]

$$b_n = h_{j,n} = \left(\frac{4}{3}S_{j,n}^5\right)^{\frac{1}{5}} \cdot n^{-\frac{1}{5}},$$

where $S_{j,n}$ is the recursive estimator of the standard deviation.

---

[6]Adapted from the non-recursive estimator proposed in Silverman, (1986)

Introduction
000000000

Copula Density Estimator
0000000

Bandwidth Selection
000●

Numerical Experiments
0000000

Conclusion

Acknowledgements

"Specialized" bandwidths

Another idea is to use different bandwidths for marginal PDFs and the joint PDF.

An optimal bandwidth for the joint can be derived by optimizing $\text{MSE}(f_n^{\mathbf{u}})$. It results in

$$
\left[\frac{\widetilde{\eta}^d \cdot f^{\mathbf{u}} \cdot (d+2) \cdot d}{2 \cdot \eta^2 \cdot \left[\sum_{j=1}^d f_{(j,j)}^{\mathbf{u}} \right]^2 \cdot (d+4)} \right]^{\frac{1}{d+4}} \cdot n^{-\frac{1}{d+4}},
$$

where

$\widetilde{\eta} \approx 0.28$ and $\eta = 1$ if kernel is standard normal distribution;

$f_{(j,j)}^{\mathbf{u}}$ is the 2$^{\text{nd}}$ order derivative on j-th component;

In case of marginals $d = 1$.

"Specialized" bandwidths

Another idea is to use different bandwidths for marginal PDFs and the joint PDF.

An optimal bandwidth for the joint can be derived by optimizing $\text{MSE}(f_n^{\mathbf{u}})$. It results in

$$\left[\frac{\widetilde{\eta}^d \cdot f^{\mathbf{u}} \cdot (d+2) \cdot d}{2 \cdot \eta^2 \cdot \left[\sum_{j=1}^{d} f_{(j,j)}^{\mathbf{u}} \right]^2 \cdot (d+4)} \right]^{\frac{1}{d+4}} \cdot n^{-\frac{1}{d+4}},$$

where

$\widetilde{\eta} \approx 0.28$ and $\eta = 1$ if kernel is standard normal distribution;

$f_{(j,j)}^{\mathbf{u}}$ is the 2$^{\text{nd}}$ order derivative on j-th component;

In case of marginals $d = 1$.

Set-up ($d = 2$)

1. Select a copula which properties are known;
2. Sample data ($\mathbf{X}$) from the selected copula distribution;
3. Choose a grid of points $\mathbf{u}$ to estimate copula on.
   For example, $(0.1, \ldots, 0.9) \times (0.1, \ldots, 0.9)$;
4. Estimate values of the copula density at each point of the selected grid $\mathbf{u}$;
5. Calculate $\text{MSE}_n(c_n(\cdot), c(\cdot))$ at each n.

Here we use $n = 5000$ and number of Monte-Carlo simulations is 500.

Independent Copula

Independent Copula
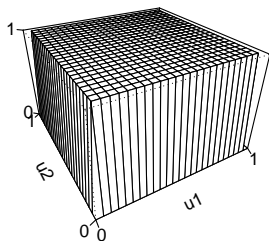
(Independence) Frank Copula, $\theta = 0.0$

Introduction
○○○○○○○○○

Copula Density Estimator
○○○○○○○

Bandwidth Selection
○○○○

Numerical Experiments
○○●○○○○

Conclusion

Acknowledgements
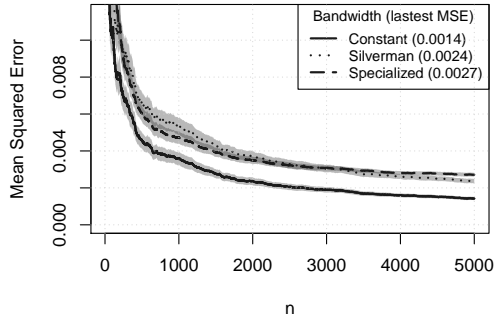
Student Copula ($\theta = 0.5$)
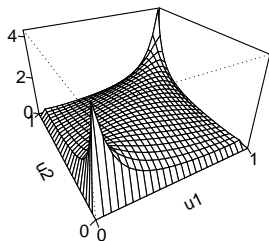
Student Copula, $\theta = 0.5$
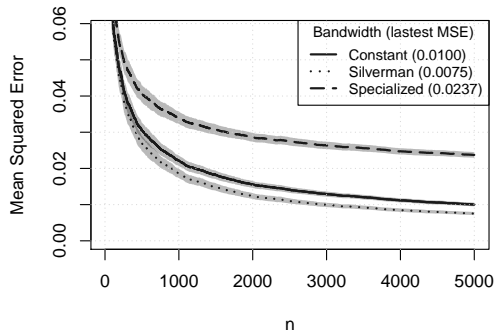
Student Copula, $\theta = 0.5$

Student Copula ($\theta = 0.0$)
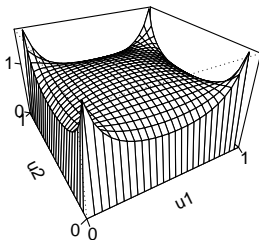
Student Copula, $\theta = 0.0$

Student Copula, $\theta = 0.0$

Clayton Copula ($\theta = -0.5$)

Clayton Copula, $\theta = -0.5$

Clayton Copula, $\theta = -0.5$

Bandwidth (lastest MSE)
—— Constant (0.0873)
······ Silverman (0.0738)
– – Specialized (0.1142)

Clayton Copula ($\theta = 10$)

Clayton Copula, $\theta = 10$

Clayton Copula, $\theta = 10.0$

Bandwidth (lastest MSE)
— Constant (7.4303)
⋯ Silverman (6.5356)
− − Specialized (4.3841)

Numerical results

1. The proposed recursive non-parametric estimator works of copula density works;

2. The usage of the constant bandwidths is a good choice;

3. Though, data-dependent bandwidths are more reliable in complicated cases.

Numerical results

1 The proposed recursive non-parametric estimator works of copula density works;

2 The usage of the constant bandwidths is a good choice;

3 Though, data-dependent bandwidths are more reliable in complicated cases.

Numerical results

1 The proposed recursive non-parametric estimator works of copula density works;

2 The usage of the constant bandwidths is a good choice;

3 Though, data-dependent bandwidths are more reliable in complicated cases.

Conclusion

Our contribution consists of the following

1. We introduced the kernel recursive non-parametric estimators of the joint PDF $f^{\mathbf{u}}$;

2. In addition, using this result we presented the recursive non-parametric estimator of the copula density $c(\cdot)$;

3. Github repository:
   https://github.com/vdyashin/CopulaDensityEstimator.

Conclusion

Our contribution consists of the following

1. We introduced the kernel recursive non-parametric estimators of the joint PDF $f^{\mathbf{u}}$;

2. In addition, using this result we presented the recursive non-parametric estimator of the copula density $c(\cdot)$;

3. Github repository:
   https://github.com/vdyashin/CopulaDensityEstimator.

Conclusion

Our contribution consists of the following

1. We introduced the kernel recursive non-parametric estimators of the joint PDF $f^{\mathbf{u}}$;

2. In addition, using this result we presented the recursive non-parametric estimator of the copula density $c(\cdot)$;

3. Github repository:
   https://github.com/vdyashin/CopulaDensityEstimator.

Introduction
000000000

Copula Density Estimator
0000000

Bandwidth Selection
0000

Numerical Experiments
0000000

Conclusion
0000000

Acknowledgements

The author sincerely acknowledges

Introduction
000000000

Copula Density Estimator
0000000

Bandwidth Selection
0000

Numerical Experiments
0000000

Conclusion

Acknowledgements

Recursive Non-parametric Estimation of the Copula Density

Vladimir Yashin
supervised by Geoffrey Decrouez

Higher School of Economics

9 June 2018

📄 Amiri, A. and B. Thiam (2014). "A smoothing stochastic algorithm for quantile estimation". In: *Statistics & Probability Letters* 93, pp. 116–125. ISSN: 0167-7152.

📄 Holst, U. (1987). "Recursive estimation of quantiles using recursive kernel density estimators". In: *Sequential Analysis* 6.3, pp. 219–237. ISSN: 0747-4946.

📄 Mokkadem, A., M. Pelletier, and Y. Slaoui (2009). "The stochastic approximation method for the estimation of a multivariate probability density". In: *Journal of Statistical Planning and Inference* 139.7, pp. 2459–2478. ISSN: 0378-3758.

📄 Robbins, H. and S. Monro (1951). "A stochastic approximation method". In: *The annals of mathematical statistics*, pp. 400–407. ISSN: 0003-4851.

📄 Robinson, D. W. (1975). "Non-parametric quantile estimation through stochastic approximation." PhD thesis.

📄 Silverman, B. W. (1986). *Density estimation for statistics and data analysis*. New-York: Routledge, p. 176. ISBN: 1351456172.

📄 Sklar, A. (1959). "Fonctions de repartition an dimensions et leurs marges". In: *Publ. Inst. Statist. Univ. Paris* 8, pp. 229–231.