



Chapitre de livre

2007

Accepted version

Open Access

This is an author manuscript post-peer-reviewing (accepted version) of the original publication. The layout of the published version may differ .

The estimation of copulas : theory and practice

Charpentier, Arthur; Fermanian, Jean-David; Scaillet, Olivier

How to cite

CHARPENTIER, Arthur, FERMANIAN, Jean-David, SCAILLET, Olivier. The estimation of copulas : theory and practice. In: Copulas: from theory to application in finance. Jörn Rank (Ed.). London : Risk Books, 2007. p. 35–64.

This publication URL: <https://archive-ouverte.unige.ch/unige:41917>

The Estimation of Copulas: Theory and Practice

**Arthur Charpentier; Jean-David Fermanian;
Olivier Scaillet¹**

Ensaie-Crest and Katholieke Universiteit Leuven; BNP-Paribas and Crest;
HEC Genève and Swiss Finance Institute

INTRODUCTION

Copulas are a way of formalising dependence structures of random vectors. Although they have been known about for a long time (Sklar (1959)), they have been rediscovered relatively recently in applied sciences (biostatistics, reliability, biology, etc). In finance, they have become a standard tool with broad applications: multi-asset pricing (especially complex credit derivatives), credit portfolio modelling, risk management, etc. For example, see Li (1999), Patton (2001) and Longin and Solnik (1995).

Although the concept of copulas is well understood, it is now recognised that their empirical estimation is a harder and trickier task. Many traps and technical difficulties are present, and these are, most of the time, ignored or underestimated by practitioners. The problem is that the estimation of copulas implies usually that every marginal distribution of the underlying random vectors must be evaluated and plugged into an estimated multivariate distribution. Such a procedure produces unexpected and unusual effects with respect to the usual statistical procedures: non-standard limiting behaviours, noisy estimations, etc (eg, see the discussion in Fermanian and Scaillet, 2005).

In this chapter, we focus on the practical issues practitioners are faced with, in particular concerning estimation and visualisation.

01 In the first section, we give a general setting for the estimation
02 of copulas. Such a framework embraces most of the available
03 techniques. In the second section, we deal with the estimation of
04 the copula density itself, with a particular focus on estimation near
05 the boundaries of the unit square.

07 A GENERAL APPROACH FOR THE ESTIMATION OF 08 COPULA FUNCTIONS

09 Copulas involve several underlying functions: the marginal cumu-
10 lative distribution functions (CDF) and a joint CDF. To estimate
11 copula functions, the first issue consists in specifying how to esti-
12 mate separately the margins and the joint law. Moreover, some of
13 these functions can be fully known. Depending on the assumptions
14 made, some quantities have to be estimated parametrically, or semi-
15 or even non-parametrically. In the latter case, the practitioner has to
16 choose between the usual methodology of using “empirical coun-
17 terparts” and invoking smoothing methods well-known in statis-
18 tics: kernels, wavelets, orthogonal polynomials, nearest neighbours,
19 etc.

20 Obviously, the estimation precision and the graphical results are
21 functions of all these choices. A true known marginal can greatly
22 improve the results under well-specification, but the reverse is
23 true under misspecification (even under a light one). Without any
24 valuable prior information, non-parametric estimation should be
25 favoured, especially for marginal estimation.

26 To illustrate this point Figure 2.1 shows the graphical behaviour
27 of the exceeding probability function

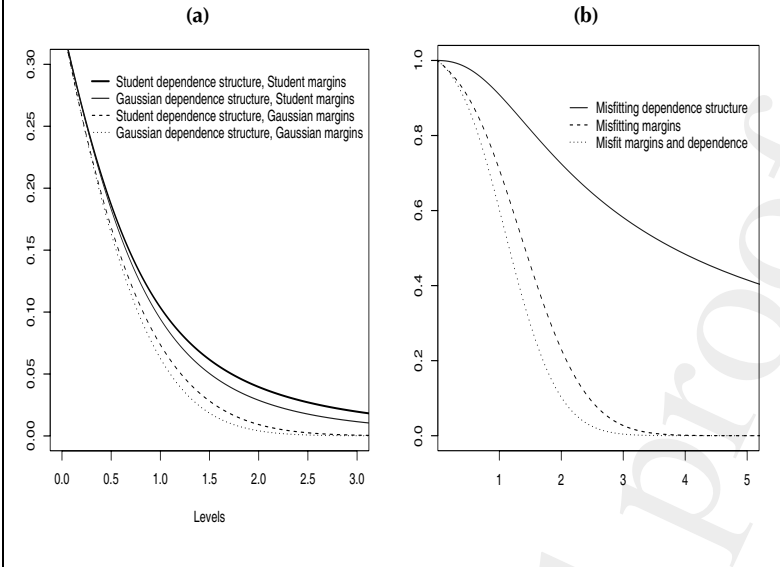
$$28 \quad \chi : p \mapsto P(X > F_X^{-1}(p), Y > F_Y^{-1}(p))$$

29
30 If the true underlying model is a multivariate Student vector (X, Y) ,
31 the associated probability is the upper line. If either marginal dis-
32 tributions are misspecified (eg, Gaussian marginal distributions),
33 or the dependence structure is misspecified (eg, joint Gaussian dis-
34 tribution), these probabilities are always underestimated, especially
35 in the tails.

36 Now, let us introduce our framework formally. Consider the
37 estimation of a d -dimensional copula C , that can be written

$$38 \quad C(\mathbf{u}) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$$

Figure 2.1 (a) The function χ when (X, Y) is a Student random vector, and when either margins or the dependence structure are misspecified. (b) The associated ratios of exceeding probability corresponding to the χ function obtained for the misspecified model versus the true χ (for the true Student model).



Obviously, all the marginal CDFs have been denoted by F_k , $k = 1, \dots, d$, when the joint CDF is F . Throughout this chapter, the inverse operator $^{-1}$ should be understood to be a generalised inverse; namely that for every function G ,

$$G^{-1}(x) = \inf\{y \mid G(y) \geq x\}$$

Assume we have observed a T -sample $(X_i)_{i=1, \dots, T}$. These are some realisations of the d -random vector $\mathbf{X} = (X_1, \dots, X_d)$. Note that we do not assume that $X_i = (X_{1i}, \dots, X_{di})$ are mutually independent (at least for the moment).

Every marginal CDF, say the k th, can be estimated empirically by

$$F_k^{(1)}(x) = \frac{1}{T} \sum_{i=1}^T \mathbb{1}(X_{ki} \leq x)$$

01 and $[F_k^{(1)}]^{-1}(u_k)$ is simply the empirical quantile corresponding to
 02 $u_k \in [0, 1]$. Another means of estimation is to smooth such CDFs,
 03 and the simplest way is to invoke the kernel method (eg, see Härdle
 04 and Linton (1994) or Pagan and Ullah (1999) for an introduction):
 05 consider a univariate kernel function $K: \mathbb{R} \rightarrow \mathbb{R}$, $\int K = 1$, and a
 06 bandwidth sequence h_T (or simply h hereafter), $h_T > 0$ and $h_T \rightarrow 0$
 07 when $T \rightarrow \infty$. Then, $F_k(x)$ can be estimated by

$$08$$

$$09 \quad F_k^{(2)}(x) = \frac{1}{T} \sum_{i=1}^T \mathbb{K} \left(\frac{x - X_{ki}}{h} \right)$$

$$10$$

$$11$$

12 for every real number x , by denoting \mathbb{K} the primitive function of K :
 13 $\mathbb{K}(x) = \int_{-\infty}^x K$.

14 There exists another common case: assume that an underlying
 15 parametric model has been fitted previously for the k th margin.
 16 Then, the natural estimator for $F_k(x)$ is some CDF $F_k^{(3)}(x, \hat{\theta}_k)$ that
 17 depends on the relevant estimated parameter $\hat{\theta}_k$. When such a
 18 model is well-specified, $\hat{\theta}_k$ is tending almost surely to a value θ_k
 19 such that $F_k(\cdot) = F_k^{(3)}(\cdot, \theta_k)$. The last limiting case is the knowledge
 20 of the true CDF F_k . Formally, we will set $F_k^{(0)} = F_k$.

21 Similarly, the joint CDF F can be estimated empirically by

$$22$$

$$23 \quad F^{(1)}(\mathbf{x}) = \frac{1}{T} \sum_{i=1}^T \mathbb{1}(\mathbf{X}_i \leq \mathbf{x})$$

$$24$$

$$25$$

26 or by the kernel method

$$27$$

$$28 \quad F^{(2)}(\mathbf{x}) = \frac{1}{T} \sum_{i=1}^T \mathbb{K} \left(\frac{\mathbf{x} - \mathbf{X}_i}{h} \right)$$

$$29$$

$$30$$

31 with a d -dimensional kernel K , so that

$$32 \quad \mathbb{K}(\mathbf{x}) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} K$$

$$33$$

$$34$$

35 for every $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$. Besides, there may exist an under-
 36 lying parametric model for \mathbf{X} : F is assumed to belong to a set of
 37 multivariate CDFs indexed by a parameter τ . A consistent estima-
 38 tion $\hat{\tau}$ for the “true” value τ allows setting $F^{(3)}(\cdot) = F(\cdot, \hat{\tau})$. Finally,
 39 we can denote $F^{(0)} = F$.

Therefore, generally speaking, a d -dimensional copula C can be estimated by

$$\hat{C}(\mathbf{u}) = F^{(j)}\left([F_1^{(j_1)}]^{-1}(u_1), \dots, [F_d^{(j_d)}]^{-1}(u_d)\right) \quad (2.1)$$

for each of the indexes j, j_1, j_2, \dots, j_d that belong to $\{0, 1, 2, 3\}$. Thus, it is not so obvious to discriminate between all these competitors, especially without any parametric assumption.

Every estimation method has its own advantages and drawbacks. The full empirical method ($j = j_1 = \dots = j_d = 1$ with the notations of Equation (2.1)) has been introduced in Deheuvels (1979, 1981a, 1981b) and studied more recently by Fermanian *et al* (2004), in the independent setting, and by Doukhan *et al* (2004) in a dependent framework. It provides a robust and universal way for estimation purposes. Nonetheless, its discontinuous feature induces some difficulties: the graphical representations of the copula can be not very nice from a visual point of view and not intuitive. Moreover, there is no unique choice for building the inverse function of $F_k^{(1)}$. In particular, if $X_{k1} \leq \dots \leq X_{kT}$ is the ordered sample on the k -axis, the inverse function of $F_k^{(1)}$ at some point i/T may be chosen arbitrarily between X_{ki} and $X_{k(i+1)}$. Finally, since the copula estimator is not differentiable when only one empirical CDF is involved in Equation (2.1), it cannot, for example, be used straightforwardly to derive an estimate of the associated copula density (by differentiation of $\hat{C}(\mathbf{u})$ with respect to all its arguments) or for optimisation purposes.

Smooth estimators are better suited to graphical usage, and can provide more easily the intuition to achieve the “true” underlying parametric distribution. However, they depend on an auxiliary smoothing parameter (eg, h in the case of the kernel method), and suffer from the well-known “curse of dimensionality”: the higher the dimension (d with our notations), the worse the performance in terms of convergence rates. In other words, as the dimension increases, the complexity of the problem increases exponentially.² Such methods can be invoked safely in practice when $d \leq 3$ and for sample sizes larger than, say, two hundred observations (which is usual in finance). The theory of fully smoothed copulas ($j = j_1 = \dots = j_d = 2$ with the notation of Equation (2.1)) can be found in Fermanian and Scaillet (2003) in a strongly dependent framework.

01 A more comfortable situation exists when “good” parametric
 02 assumptions are put into (2.1) for the marginal CDFs and/or the
 03 joint CDF F . The former case is relatively usual because there exist
 04 a great many univariate models for financial variables (eg, see
 05 Alexander (2002)). Nevertheless, for a lot of dynamic models (eg,
 06 stochastic volatility models), their (unconditional) marginal CDFs
 07 cannot be written explicitly. Obviously, we are under the threat of a
 08 misspecification, which can have disastrous effects (see Fermanian
 09 and Scaillet (2005)). Concerning a parametric assumption for F
 10 itself, our opinion is balanced. At first glance, we are absolutely free
 11 to choose an “interesting” parametric family \mathcal{F} of d -dimensional
 12 CDFs that would contain the true law F . But, by setting for every
 13 real number x and every $k = 1, \dots, d$

$$14 \quad F_k^{(3)}(x) = F(+\infty, \dots, +\infty, x, +\infty, \dots | \hat{\tau})$$

15
 16
 17 where x is the k th argument of F , we should have found the “right”
 18 marginal distributions too, to be self-coherent. Indeed, the joint law
 19 contains the marginal ones. Then, the estimated copula should be

$$20 \quad \hat{C}(\mathbf{u}) = F^{(3)}\left([F_1^{(3)}]^{-1}(u_1), [F_2^{(3)}]^{-1}(u_2), \dots, [F_d^{(3)}]^{-1}(u_d)\right)$$

21
 22 In reality, the problem is finding a sufficiently rich family \mathcal{F} *ex ante*
 23 that might generate all empirical features. What people do is more
 24 clever. They choose a parametric family \mathcal{F}^* and other marginal
 25 parametric families \mathcal{F}_k^* , $k = 1, \dots, d$, and set

$$26 \quad C(\mathbf{u}) = \hat{F}^*\left([\hat{F}_1^*]^{-1}(u_1), \dots, [\hat{F}_d^*]^{-1}(u_d)\right)$$

27
 28
 29 for some $\hat{F}^* \in \mathcal{F}$, and $\hat{F}_k^* \in \mathcal{F}_k$ for every $k = 1, \dots, d$. Note that the
 30 choice of all the parametric families is absolutely free of constraints,
 31 and that these families are not related to each other (they can
 32 be arbitrary and independently chosen). This is the usual way of
 33 generating new copula families. The price to be paid is that the true
 34 joint law F does not belong to \mathcal{F}^* generally speaking. Similarly, the
 35 true marginal laws F_k do not belong to the sets \mathcal{F}_k^* in general.

36
 37 If a parametric assumption is made in such a case, the standard
 38 estimation procedure is semi-parametric: the copula is a function of
 39 some parameter $\theta = (\tau, \theta_1, \dots, \theta_d)$. Recall that the copula density c

is the derivative of C with respect to each of its arguments:

$$c_{\theta}(\mathbf{u}) = \frac{\partial^d}{\partial_1 \dots \partial_d} C(\mathbf{u})$$

Here, the copula density c_{θ} itself can be calculated under a full parametric assumption. Thus, we get an estimator of θ by maximizing the log-likelihood

$$\sum_{i=1}^T \log c_{\theta}(\hat{F}_1(X_{1i}), \dots, \hat{F}_d(X_{di}))$$

for some \sqrt{T} -convergent estimates $\hat{F}_k(X_{ki})$ of the marginal CDFs. Obviously, we may choose $\hat{F}_k = F_k^{(1)}$ or $F_k^{(2)}$.

Note that such an estimator is called an “omnibus estimator”, and it can be seen as a maximum-likelihood estimator of θ after replacing the unobservable ranks $F_k(X_{ki})$ by the pseudo-observations. The asymptotic distribution of the estimator has been studied in Genest *et al* (1995) and Shi and Louis (1995). The main aim of semi-parametric estimation is to avoid possible misspecification of marginal distributions, which may overestimate the degree of dependence in the data (eg, see Silvapulle *et al* (2004)). Note finally that Chen and Fan (2004a, 2004b) have developed the theory of this semi-parametric estimator in a time-series context.

Thus, depending on the degree of assumptions about the joint and marginal models, there exists a wide range of possibilities for estimating copula functions as provided by Equation (2.1). The only trap to avoid is to be sure that the assumptions made for marginals are consistent with those drawn for the joint law. The statistical properties of all these estimators are the usual ones, namely consistency and asymptotic normality.

THE ESTIMATION OF COPULA DENSITIES

After the estimation of C by \hat{C} as in Equation (2.1), it is tempting to define an estimate of the copula density c at every $\mathbf{u} \in [0, 1]^d$ by

$$\hat{c}(\mathbf{u}) = \frac{\partial^d}{\partial_1 \dots \partial_d} \hat{C}(\mathbf{u})$$

Unfortunately, this works only when \hat{C} is differentiable. Most of the time, this is the case when the marginal and joint CDFs are

01 parametric or nonparametrically smoothed (by the kernel method,
 02 for instance). In the latter case and when d is “large” (more than 3),
 03 the estimation of c can be relatively poor because of the curse of
 04 dimensionality.

05 Nonparametric estimation procedures for the density of a copula
 06 function have already been proposed by Behnen *et al* (1985) and
 07 Gijbels and Mielniczuk (1990). These procedures rely on symmetric
 08 kernels, and have been detailed in the context of uncensored data.
 09 Unfortunately, such techniques are not consistent on the bound-
 10 aries of $[0, 1]^d$. They suffer from the so-called boundary bias. Such
 11 bias can be significant in the neighbourhood of the boundaries
 12 too, depending on the size of the bandwidth. Hereafter, we will
 13 propose some solutions to cope with such issues. To ease notation
 14 and without a lack of generality, we will restrict ourselves to the
 15 bivariate case ($d = 2$). Thus, our random vector will be denoted by
 16 (X, Y) instead of (X_1, X_2) .

17 In the following sections, we will study some properties of some
 18 kernel-based estimators, and illustrate some of these by simula-
 19 tions. The benchmark will be a simulated sample, whose size is
 20 $T = 1,000$ and that will be generated by a Frank copula with copula
 21 density

$$22 \quad c^{Fr}(u, v, \theta) = \frac{\theta[1 - e^{-\theta}]e^{-\theta(u+v)}}{23 \quad ([1 - e^{-\theta}] - (1 - e^{-\theta u})(1 - e^{-\theta v}))^2}$$

24 and Kendall’s tau equal to 0.5. Hence, the copula parameter is $\theta =$
 25 5.74. This density can be seen in Figure 2.2 together with its contour
 26 plot on the right.
 27

28
 29 **Nonparametric density estimation for distributions with finite**
 30 **support**

31 An initial approach relies on a kernel-based estimation of the den-
 32 sity based on the pseudo-observations $(F_{X,T}(X_i), F_{Y,T}(Y_i))$, where
 33 $F_{X,T}$ and $F_{Y,T}$ are the empirical distribution functions
 34

$$35 \quad F_{X,T}(x) = \frac{1}{T+1} \sum_{i=1}^T \mathbb{1}(X_i \leq x) \quad \text{and} \quad F_{Y,T}(y) = \frac{1}{T+1} \sum_{i=1}^T \mathbb{1}(Y_i \leq y)$$

36
 37
 38 where the factor $T + 1$ (instead of standard T , as in Deheuvels
 39 (1979) for instance) allows the avoidance of boundary problems: the

01 quantities $F_{X,T}(X_i)$ and $F_{Y,T}(Y_i)$ are the ranks of the X_i 's and the Y_i 's
 02 divided by $T + 1$, and therefore take values

$$03 \quad \left\{ \frac{1}{T+1}, \frac{2}{T+1}, \dots, \frac{T}{T+1} \right\}$$

06 Standard kernel-based estimators of the density of pseudo-
 07 observations yield, using diagonal bandwidth (see Wand and Jones
 08 (1995))

$$09 \quad \hat{c}_h(u, v) = \frac{1}{Th^2} \sum_{i=1}^T K\left(\frac{u - F_{X,T}(X_i)}{h}, \frac{v - F_{Y,T}(Y_i)}{h}\right)$$

12 for a bivariate kernel $K : \mathbb{R}^2 \rightarrow \mathbb{R}$, $\int K = 1$.

14 The variance of the estimator can be derived, and is $\mathcal{O}((Th^2)^{-1})$.
 15 Moreover, it is asymptotically normal at every point $(u, v) \in (0, 1)$:

$$16 \quad \frac{\hat{c}_h(u, v) - E(\hat{c}_h(u, v))}{\sqrt{\text{Var}(\hat{c}_h(u, v))}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

19 As a benchmark, Figure 2.2 shows the theoretical density of a
 20 Frank copula. In Figure 2.3 we plot the standard Gaussian kernel
 21 estimator based on the sample of pseudo-observations $(\hat{U}_i, \hat{V}_i) \equiv$
 22 $(F_{X,T}(X_i), F_{Y,T}(Y_i))$.

23 Recall that even if kernel estimates are consistent for distribu-
 24 tions with unbounded support and the support is bounded, the
 25 boundary bias can yield some “ill” underestimation (even if the
 26 distribution is twice differentiable in the interior of its support).

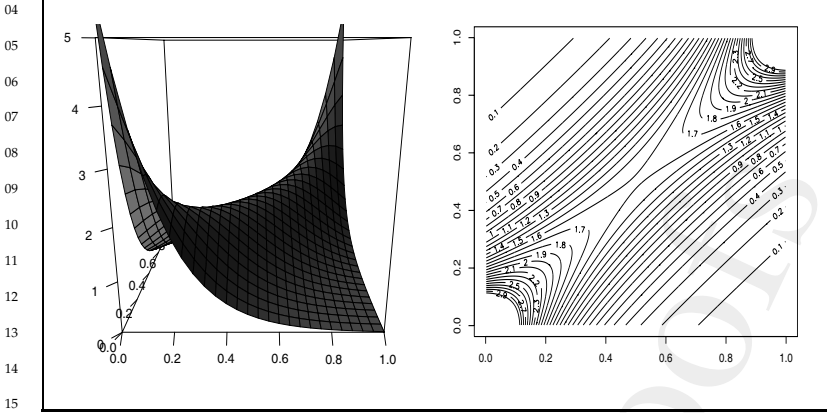
27 We can explain this phenomenon easily in the univariate case.
 28 Consider a T sample X_1, \dots, X_T of a positive random variable with
 29 density f . The support of their density is then \mathbb{R}^+ . Let K denote
 30 a symmetric kernel, whose support is $[-1, +1]$. Then, for all $x \geq 0$,
 31 using a Taylor expansion, we get

$$32 \quad E(\hat{f}_h(x)) = \int_{-1}^{x/h} K(y) f(x - hy) dy$$

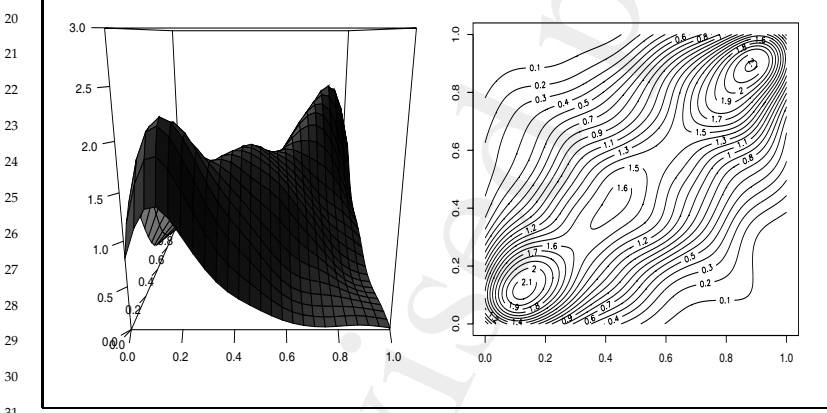
$$33 \quad = f(x) \cdot \int_{-1}^{x/h} K(y) dy$$

$$34 \quad - h \cdot f'(x) \cdot \int_{-1}^{x/h} yK(y) dy + \mathcal{O}(h^2)$$

01 **Figure 2.2** Density of the Frank copula with a Kendall tau equal
 02 to 0.5.
 03



15
 16
 17 **Figure 2.3** Estimation of the copula density using a Gaussian kernel
 18 based on 1,000 observations drawn from a Frank copula.
 19

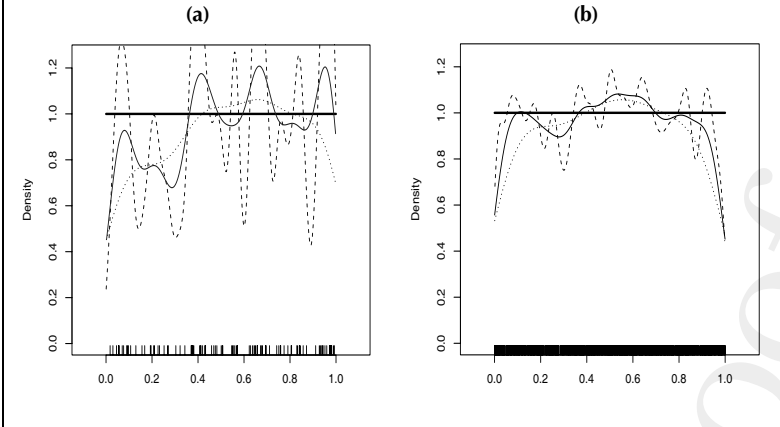


32 Hence, since the kernel is symmetric, $\int_{-1}^{x/h} K(y) dy \xrightarrow{h \rightarrow 0} 1/2$ when
 33 $x = 0$, and therefore
 34

35
$$E(\hat{f}_h(0)) = \frac{1}{2}f(0) + \mathcal{O}(h)$$

36
 37 Note that, if $x > 0$, the expression $\int_{-1}^{x/h} K(y) dy$ is 1 when h is
 38 sufficiently small (when $x > h$ to be specific). Thus, this integral
 39 cannot be one, uniformly, with respect to every $x \in (0, 1]$. And for

Figure 2.4 Estimation of the uniform density on $[0, 1]$ using a Gaussian kernel and different bandwidth with (a) $n = 100$ and (b) $n = 1,000$ observations.



more general kernels, it has no reason to be equal to 1. In the latter case, since this expression can be calculated, normalising $\hat{f}_h(x)$ by dividing by $\int_{-1}^{x/h} K(z) dz$ (at each x) achieves consistency. Nonetheless, it remains a bias that is of the order of $\mathcal{O}(h)$. Using some boundary kernels (see Gasser and Müller (1979)), it is possible to achieve $\mathcal{O}(h^2)$ everywhere in the interior of the support.

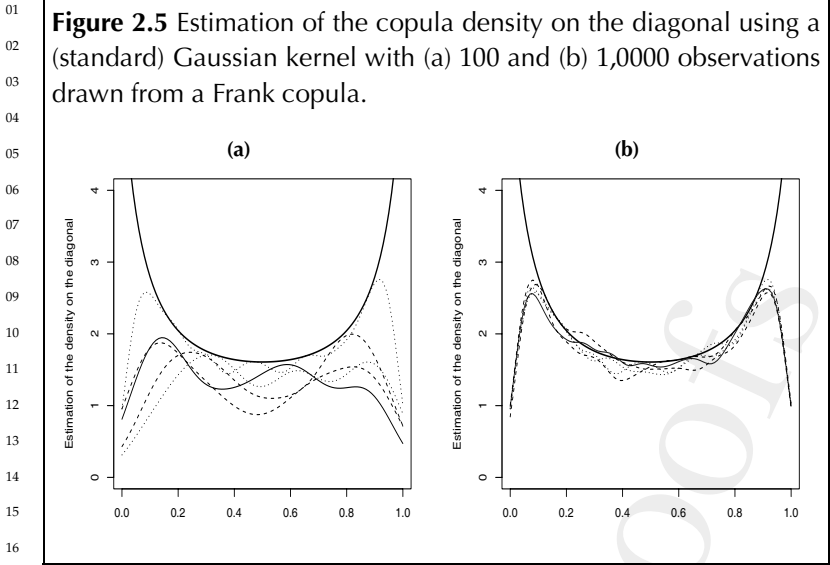
Consider the case of variables uniformly distributed on $[0, 1]$, U_1, \dots, U_n . Figure 2.4 shows kernel-based estimators of the uniform density, with Gaussian kernel and different bandwidths, with $n = 100$ and 1,000 simulated variables. In that case, for any $h > 0$,

$$E(\hat{f}_h(0)) = \int_0^1 K_h(y) dy = \frac{1}{h\sqrt{2\pi}} \int_0^1 \exp\left(-\frac{y^2}{2h^2}\right) dy \xrightarrow{h \rightarrow 0} \frac{1}{2} = \frac{f(0)}{2}$$

and in the interior, ie, $x \in (0, 1)$,

$$\begin{aligned} E(\hat{f}_h(x)) &= \int_0^1 K_h(y-x) dy \\ &= \frac{1}{h\sqrt{2\pi}} \int_0^1 \exp\left(-\frac{(y-x)^2}{2h^2}\right) dy \xrightarrow{h \rightarrow 0} 1 = f(x) \end{aligned}$$

Dealing with bivariate copula densities, we observe the same phenomenon. On boundaries, we obtain some “multiplicative



17
 18 bias", $1/4$ in corners and $1/2$ in the interior of borders. The addi-
 19 tional bias is of the order of $\mathcal{O}(h)$ on the frontier, and standard $\mathcal{O}(h^2)$
 20 in the interior. More precisely, in any corners (eg, $(0, 0)$)
 21

$$22 \quad E(\hat{c}_h(0, 0)) = \frac{1}{4}c(u, v) + \mathcal{O}(h)$$

23
 24 on the interior of the borders (eg, $u = 0$ and $v \in (0, 1)$)

$$25 \quad E(\hat{c}_h(0, v)) = \frac{1}{2}c(u, v) + \mathcal{O}(h)$$

26
 27 and in the interior ($(u, v) \in (0, 1) \times (0, 1)$)

$$28 \quad E(\hat{c}_h(u, v)) = c(u, v) + \mathcal{O}(h^2)$$

29
 30 The bias can be observed in Figure 2.5, which represents the diago-
 31 nal of the estimated density for several samples.

32 Several techniques have been introduced to obtain a better esti-
 33 mation on the borders for univariate densities:
 34

- 35 • mirror image modification (Schuster (1985); Deheuvels and
 36 Hominal (1979)), where artificial data are obtained, using sym-
 37 metric (mirror) transformations on the borders;
- 38 • transformed kernels (Devroye and Györfi (1985); Wand *et al*
 39 (1991)), where the idea is to transform the data X_i using a

bijjective mapping ϕ so that the $\phi(X_i)$ have support \mathbb{R} . Efficient kernel-based estimation of the density of the $\phi(X_i)$ can be derived, and, by the inverse transformation, we get back the density estimation of the X_i themselves;

- boundary kernels (Gasser and Müller (1979); Rice (1984); Müller (1991)), where a smooth distortion is considered near the border, so that the bandwidth and the kernel shape can be modified (the closer to the border, the smaller).

Finally, the last section will briefly mention the impact of pseudo-observations, ie, working on samples

$$\{(F_{X,T}(X_1), F_{Y,T}(Y_1)), \dots, (F_{X,T}(X_T), F_{Y,T}(Y_T))\}$$

instead of

$$\{(F_X(X_1), F_Y(Y_1)), \dots, (F_X(X_T), F_Y(Y_T))\}$$

as if we know the true marginal distributions.

Mirror image

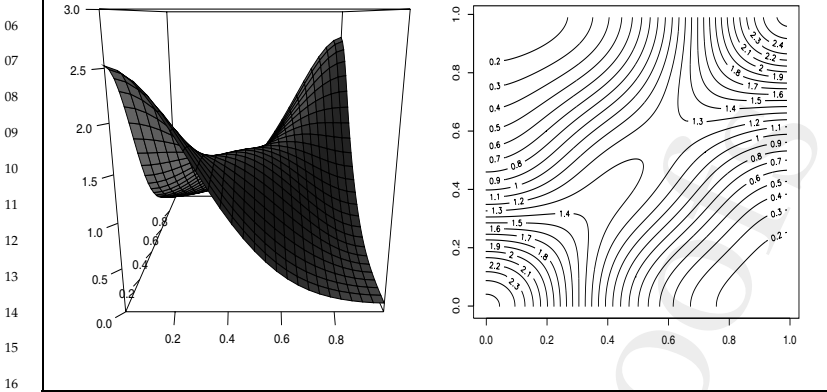
The idea of this method, developed by Deheuvels and Hominal (1979) and Schuster (1985), is to add some “missing mass” by reflecting the sample with respect to the boundaries. They focus on the case where variables are positive, ie, whose support is $[0, \infty)$. Formally and in its simplest form, it means replacing $K_h(x - X_i)$ by $K_h(x - X_i) + K_h(x + X_i)$. The estimator of the density is then

$$\hat{f}_h(x) = \frac{1}{Th} \sum_{i=1}^T \left\{ K\left(\frac{x - X_i}{h}\right) + K\left(\frac{x + X_i}{h}\right) \right\}$$

In the case of densities whose support is $[0, 1] \times [0, 1]$, the non-consistency can be corrected on the boundaries, but the convergence rate of the bias will remain $\mathcal{O}(h)$ on the boundaries, which is larger than the usual rate $\mathcal{O}(h^2)$ obtained in the interior if $h \rightarrow 0$. The only case where the usual rate of convergence is obtained on boundaries is when the derivative of the density is zero on such subsets. Note that the variance is 4 times higher in corners and 2 times higher in the interior of borders.

For copulas, instead of using only the “pseudo-observations” $(\hat{U}_i, \hat{V}_i) \equiv (F_{X,T}(X_i), F_{Y,T}(Y_i))$, the mirror image consists in reflecting each data point with respect to all edges and corners of the

01 **Figure 2.6** Estimation of the copula density using a Gaussian kernel
 02 and the mirror reflection principle with 1,000 observations from the
 03 Frank copula.
 04



17 unit square $[0, 1] \times [0, 1]$. Hence, additional observations can be
 18 considered; ie, the $(\pm \hat{U}_i, \pm \hat{V}_i)$, the $(\pm \hat{U}_i, 2 - \hat{V}_i)$, the $(2 - \hat{U}_i, \pm \hat{V}_i)$
 19 and the $(2 - \hat{U}_i, 2 - \hat{V}_i)$, so that one can consider
 20

21 $\hat{c}_h(u, v)$
 22
$$= \frac{1}{Th^2} \sum_{i=1}^T \left\{ K\left(\frac{u - \hat{U}_i}{h}\right) K\left(\frac{v - \hat{V}_i}{h}\right) + K\left(\frac{u + \hat{U}_i}{h}\right) K\left(\frac{v - \hat{V}_i}{h}\right) \right.$$

 23
$$+ K\left(\frac{u - \hat{U}_i}{h}\right) K\left(\frac{v + \hat{V}_i}{h}\right) + K\left(\frac{u + \hat{U}_i}{h}\right) K\left(\frac{v + \hat{V}_i}{h}\right)$$

 24
$$+ K\left(\frac{u - \hat{U}_i}{h}\right) K\left(\frac{v - 2 + \hat{V}_i}{h}\right) + K\left(\frac{u + \hat{U}_i}{h}\right) K\left(\frac{v - 2 + \hat{V}_i}{h}\right)$$

 25
$$+ K\left(\frac{u - 2 + \hat{U}_i}{h}\right) K\left(\frac{v - \hat{V}_i}{h}\right) + K\left(\frac{u - 2 + \hat{U}_i}{h}\right) K\left(\frac{v + \hat{V}_i}{h}\right)$$

 26
$$\left. + K\left(\frac{u - 2 + \hat{U}_i}{h}\right) K\left(\frac{v - 2 + \hat{V}_i}{h}\right) \right\}$$

 27
 28
 29
 30
 31
 32
 33

34 Figure 2.6 has been obtained using the reflection principle.
 35 We can check that the fit is far better than in Figure 2.3.

36
 37 **Transformed kernels**

38 Recall that c is the density of (U, V) , $U = F_X(X)$ and $V = F_Y(Y)$.
 39 The two latter random variables (RVs) follow uniform distributions

(marginally). Consider a distribution function G of a continuous distribution on \mathbb{R} , with differentiable strictly positive density g . We build new RVs $\tilde{X} = G^{-1}(U)$ and $\tilde{Y} = G^{-1}(V)$. Then, the density of (\tilde{X}, \tilde{Y}) is

$$f(x, y) = g(x)g(y)c[G(x), G(y)] \quad (2.2)$$

This density is twice continuously differentiable on \mathbb{R}^2 , and the standard kernel approach applies.

Since we do not observe a sample of (U, V) but instead make pseudo-observations (\hat{U}_i, \hat{V}_i) , we build an “approximated sample” of the transformed variables $(\tilde{X}_1, \tilde{Y}_1), \dots, (\tilde{X}_T, \tilde{Y}_T)$ by setting $\tilde{X}_i = G^{-1}(\hat{U}_i)$ and $\tilde{Y}_i = G^{-1}(\hat{V}_i)$. Thus, the kernel estimator of f is

$$\hat{f}(x, y) = \frac{1}{Th^2} \sum_{i=1}^T K\left(\frac{x - \tilde{X}_i}{h}, \frac{y - \tilde{Y}_i}{h}\right) \quad (2.3)$$

The associated estimator of c is then deduced by inverting (2.2),

$$c(u, v) = \frac{f(G^{-1}(u), G^{-1}(v))}{g(G^{-1}(u))g(G^{-1}(v))}, \quad (u, v) \in [0, 1] \times [0, 1]$$

and therefore we get

$$\begin{aligned} \hat{c}_h(u, v) &= \frac{1}{Th^2 g(G^{-1}(u)) \cdot g(G^{-1}(v))} \\ &\times \sum_{i=1}^T K\left(\frac{G^{-1}(u) - G^{-1}(\hat{U}_i)}{h}, \frac{G^{-1}(v) - G^{-1}(\hat{V}_i)}{h}\right) \end{aligned}$$

Note that this approach can be extended by considering different transformations G_X and G_Y , different kernels K_X and K_Y , or different bandwidths h_X and h_Y , for the two marginal random variables.

Figure 2.7 was obtained using the transformed kernel, where K was a Gaussian kernel and G was respectively the CDF of the $\mathcal{N}(0, 1)$ distribution.

The absence of a multiplicative bias on the borders can be observed in Figure 2.8, where the diagonal of the copula density is plotted, based on several samples. The copula density estimator obtained with transformed samples has no bias, is asymptotically normal, etc. Actually, we get all the usual properties of the multivariate kernel density estimators.

01
02
03
04
05
06
07
08
09
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39

Figure 2.7 Estimation of the copula density using a Gaussian kernel and Gaussian transformations with 1,000 observations drawn from the Frank copula.

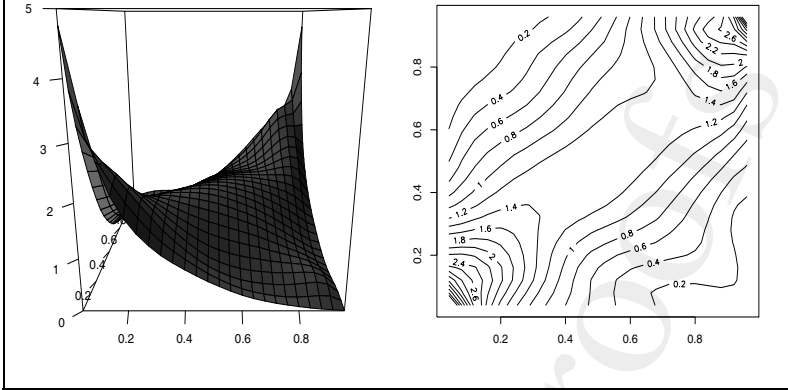
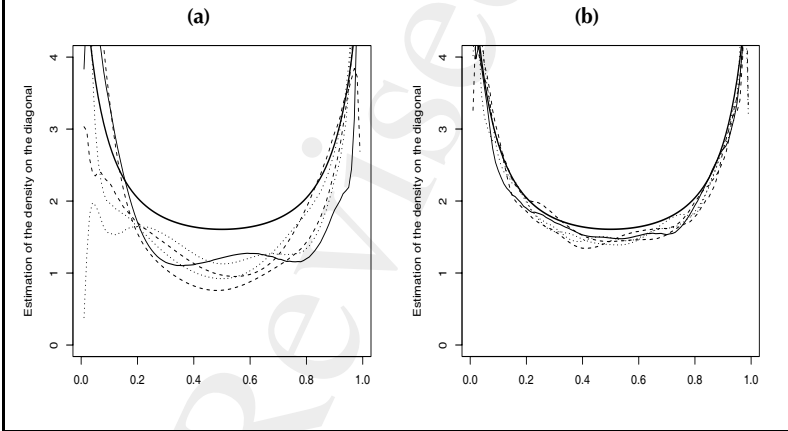


Figure 2.8 Estimation of the copula density on the diagonal using a Gaussian kernel and Gaussian transformations with (a) 100 and (b) 1,000 observations drawn from a Frank copula.



Beta kernels

In this section we examine the use of the beta kernel introduced by Brown and Chen (1999), and Chen (1999, 2000) for nonparametric estimation of regression curves and univariate densities with compact support, respectively.

Following an idea by Harrell and Davis (1982), Chen (1999, 2000) introduced the beta kernel estimator as an estimator of a density function with known compact support $[0, 1]$, to remove the boundary bias of the standard kernel estimator:

$$\hat{f}_h(x) = \frac{1}{T} \sum_{i=1}^T K\left(X_i, \frac{x}{h} + 1, \frac{1-x}{h} + 1\right)$$

where $K(\cdot, \alpha, \beta)$ denotes the density of the beta distribution with parameters α and β ,

$$K(x, \alpha, \beta) = \frac{x^\alpha (1-x)^\beta}{B(\alpha, \beta)}, \quad x \in [0, 1]$$

where

$$B(\alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)}$$

The main difficulty when working with this estimator is the lack of a simple “rule of thumb” for choosing the smoothing parameter h .

The beta kernel has two leading advantages. First it can match the compact support of the object to be estimated. Secondly it has a flexible form and changes the smoothness in a natural way as we move away from the boundaries. As a consequence, beta kernel estimators are naturally free of boundary bias and can produce estimates with a smaller variance. Indeed we can benefit from a larger effective sample size since we can pool more data. Monte Carlo results available in these papers show that they have better performance compared to other estimators which are free of boundary bias, such as local linear (Jones (1993)) or boundary kernel (Müller (1991)) estimators. Renault and Scaillet (2004) also report better performance compared to transformation kernel estimators (Silverman (1986)). In addition, Bouezmarni and Rolin (2001, 2003) show that the beta kernel density estimator is consistent even if the true density is unbounded at the boundaries. This feature may also arise in our situation. For example the density of a

01 bivariate Gaussian copula is unbounded at the corners $(0, 0)$ and
 02 $(1, 1)$. Therefore beta kernels are appropriate candidates to build
 03 well-behaved nonparametric estimators of the density of a copula
 04 function.

05 The beta-kernel based estimator of the copula density at point
 06 (u, v) is obtained using product beta kernels, which yields

$$\begin{aligned} \hat{c}_h(u, v) &= \frac{1}{Th^2} \sum_{i=1}^T K\left(X_i, \frac{u}{h} + 1, \frac{1-u}{h} + 1\right) \\ &\quad \times K\left(Y_i, \frac{v}{h} + 1, \frac{1-v}{h} + 1\right) \end{aligned}$$

12 Figure 2.9 shows that the shape of the product beta kernels for
 13 different values of u and v is clearly adaptive.

14 For convenience, the bandwidths are here assumed to be equal,
 15 but, more generally, one can consider one bandwidth per compo-
 16 nent. See Figure 2.10 for an example of an estimation based on beta
 17 kernels and a bandwidth $h = 0.05$.

18 Let $(u, v) \in [0, 1] \times [0, 1]$. The bias of $\hat{c}_h(u, v)$ is of the order of h ,
 19 $\hat{c}_h(u, v) = c(u, v) + \mathcal{O}(h)$. The absence of a multiplicative bias on the
 20 boundaries can be observed on Figure 2.11, where the diagonal of the
 21 copula density is plotted, based on several samples.

22 On the other hand, note that the variance depends on the loca-
 23 tion. More precisely, $\text{Var}(\hat{c}_h(u, v))$ is $\mathcal{O}((Th^\kappa)^{-1})$, where $\kappa = 2$ in
 24 corners, $\kappa = 3/2$ in borders, and $\kappa = 1$ in the interior of $[0, 1] \times$
 25 $[0, 1]$. Moreover, as well as “standard” kernel estimates, $\hat{c}_h(u, v)$ is
 26 asymptotically normally distributed:

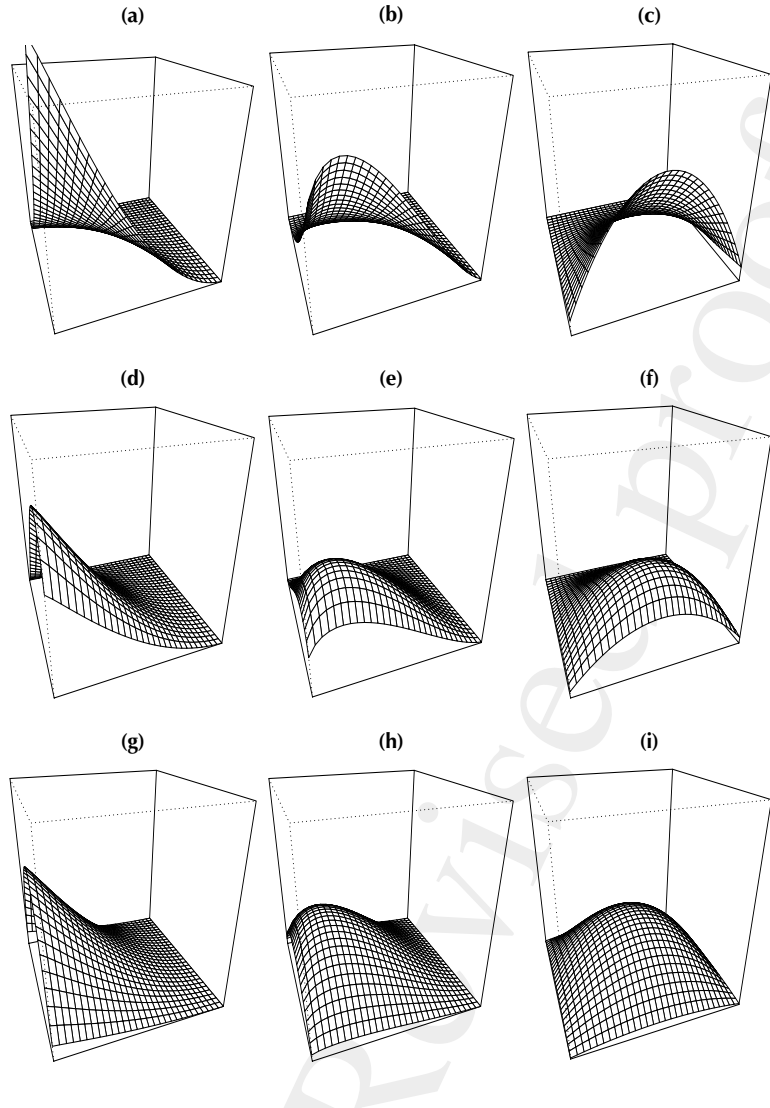
$$\begin{aligned} \sqrt{Th^{\kappa'}} [\hat{c}_h(u, v) - c(u, v)] &\xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma(u, v)^2) \\ \text{as } Th^{\kappa'} &\rightarrow \infty \text{ and } h \rightarrow 0 \end{aligned}$$

27 where κ' depends on the location, and where $\sigma(u, v)^2$ is propor-
 28 tional to $c(u, v)$.

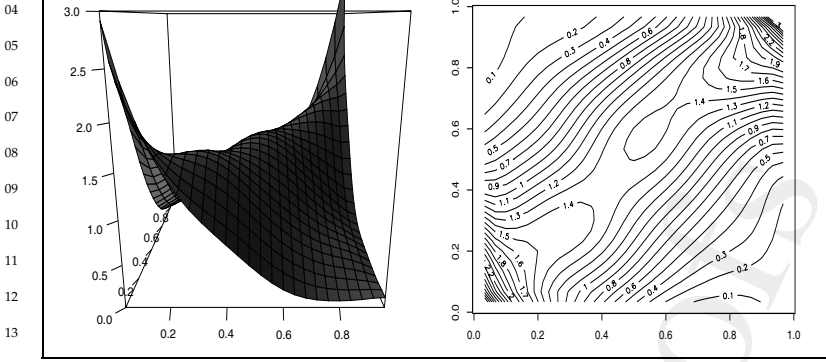
29 **Working with pseudo-observations**

30 As we know, most of the time the marginal distributions of random
 31 vectors are unknown, as recalled in the first section. Hence, the
 32 associated copula density should be estimated not on samples
 33 $(F_X(X_i), F_Y(Y_i))$ but on pseudo-samples $(F_{X,T}(X_i), F_{Y,T}(Y_i))$.

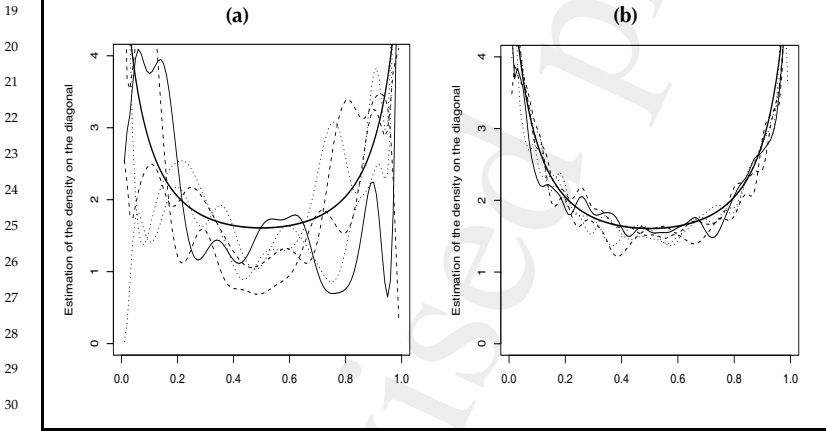
Figure 2.9 Shape of bivariate beta kernels for different values of u and v . (a) $u = 0.0, v = 0.0$; (b) $u = 0.2, v = 0.0$; (c) $u = 0.5, v = 0.0$; (d) $u = 0.0, v = 0.2$; (e) $u = 0.2, v = 0.2$; (f) $u = 0.5, v = 0.2$; (g) $u = 0.0, v = 0.5$; (h) $u = 0.2, v = 0.5$; (i) $u = 0.5, v = 0.5$.



01 **Figure 2.10** Estimation of the copula density using beta kernels ($u =$
 02 0.05) with 1,000 observations drawn from a Frank copula.
 03

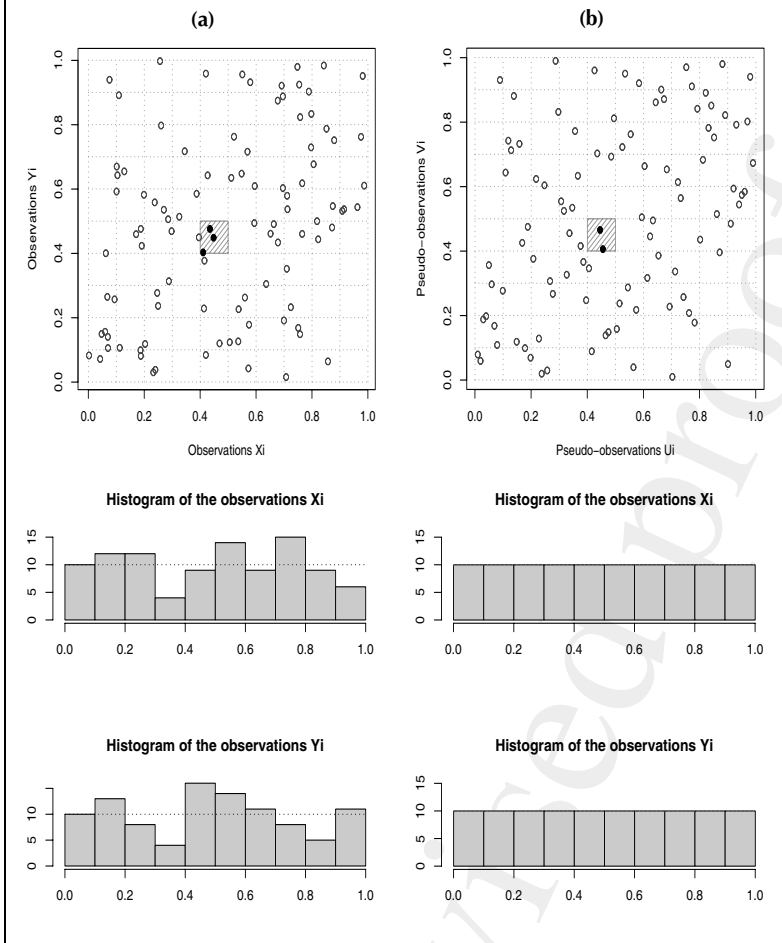


15 **Figure 2.11** Estimation of the copula density on the diagonal using
 16 beta kernels with (a) 100 and (b) 1,000 observations drawn from a
 17 Frank copula.
 18



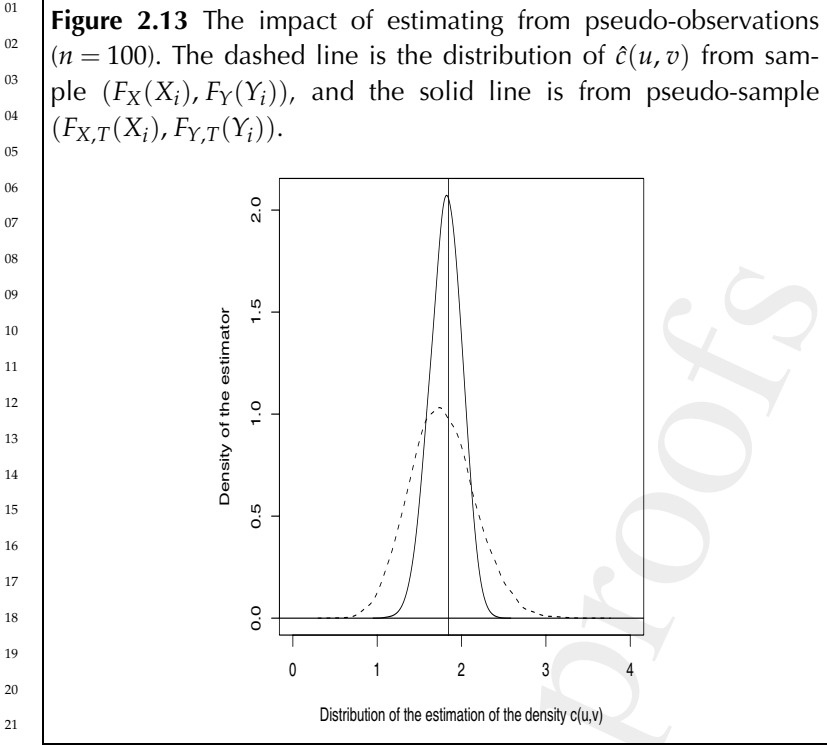
34 Figure 2.12 shows some scatterplots when the margins are
 35 known (ie, we know $(F_X(X_i), F_Y(Y_i))$), and when margins are esti-
 36 mated (ie, $(F_{X,T}(X_i), F_{Y,T}(Y_i))$). Note that the pseudo-sample is more
 37 “uniform”, in the sense of a lower discrepancy (as in quasi Monte
 38 Carlo techniques; eg, see Niederreiter (1992)). Here, by mapping
 39 every point of the sample on the marginal axis, we get uniform

Figure 2.12 Scatterplots of observations and pseudo-observations and histograms of margins. (a) 100 observations (X_i, Y_i) drawn from a Frank copula and (b) associated pseudo-sample $(U_i, V_i) = (\hat{F}_X(X_i), \hat{F}_Y(Y_i))$.



grids, which is a type of “Latin hypercube” property (eg, see Jäckel (2002)).

Because samples are more “uniform” using ranks and pseudo-observations, the variance of the estimator of the density, at some given point $(u, v) \in (0, 1) \times (0, 1)$, is usually smaller. For instance,



22
 23
 24 Figure 2.13 shows the impact of considering pseudo-observations,
 25 ie, substituting $F_{X,T}$ and $F_{Y,T}$ into unknown marginal distributions
 26 F_X and F_Y . The dashed line shows the density of $\hat{c}(u, v)$ from 100
 27 observations (U_i, V_i) (drawn from the same Frank copula), and the
 28 solid line shows the density of $\hat{c}(u, v)$ from the sample of pseudo-
 29 observations (ie, the ranks of the observations).
 30

31 A heuristic interpretation can be obtained from Figure 2.12.
 32 Consider the standard kernel-based estimator of the density, with
 33 a rectangular kernel. Consider a point (u, v) in the interior, and a
 34 bandwidth h such that the square $[u - h, u + h] \times [v - h, v + h]$ lies
 35 in the interior of the unit square. Given a T sample, an estimation
 36 of the density at point (u, v) involves the number of points located
 37 in the small square around (u, v) . Such a number will be denoted
 38 by N , and it is a random variable. Larger N provides more precise
 39 estimations.

01 Assume that the margins are known, or equivalently, let
 02 $(U_1, V_1), \dots, (U_T, V_T)$ denote a sample with distribution func-
 03 tion C . The number of points in the small square, say N_1 , is random
 04 and follows a binomial law with size T and some parameter p_1 .
 05 Thus, we have $N_1 \sim \mathcal{B}(T, p_1)$ with

$$\begin{aligned} p_1 &= P((U, V) \in [u - h, u + h] \times [v - h, v + h]) \\ &= C(u + h, v + h) + C(u - h, v - h) \\ &\quad - C(u - h, v + h) - C(u + h, v - h) \end{aligned}$$

10 and therefore

$$\text{Var}(N_1) = Tp_1(1 - p_1)$$

12 On the other hand, assume that margins are unknown, or equiv-
 13 alently that we are dealing with a sample of pseudo-observations
 14 $(\hat{U}_1, \hat{V}_1), \dots, (\hat{U}_T, \hat{V}_T)$. By construction of pseudo-observations, we
 15 have

$$\#\{\hat{U}_i \in [u - h, u + h]\} = \lfloor 2hT \rfloor$$

18 where $\lfloor \cdot \rfloor$ denotes the integer part. As previously, the number of
 19 points in the small square N_2 satisfies $N_2 \sim \mathcal{B}(\lfloor 2hT \rfloor, p_2)$ where

$$\begin{aligned} p_2 &= P((\hat{U}, \hat{V}) \in [u - h, u + h] \times [v - h, v + h] \mid \hat{U} \in [u - h, u + h]) \\ &= \frac{P((\hat{U}, \hat{V}) \in [u - h, u + h] \times [v - h, v + h])}{P(\hat{U} \in [u - h, u + h])} \\ &\approx \{C(u + h, v + h) + C(u - h, v - h) \\ &\quad - C(u - h, v + h) - C(u + h, v - h)\} / 2h \\ &= \frac{p_1}{2h} \end{aligned}$$

28 Therefore the expected number of observations is the same for both
 29 methods ($E[N_1] \simeq E[N_2] \simeq Tp_1$), but

$$\text{Var}(N_2) \approx 2hTp_2(1 - p_2) = 2hT \frac{p_1}{2h} \left(1 - \frac{p_1}{2h}\right) = \frac{T}{2h} p_1(2h - p_1)$$

33 Thus

$$\frac{\text{Var}(N_2)}{\text{Var}(N_1)} = \frac{Tp_1(2h - p_1)}{2hTp_1(1 - p_1)} = \frac{2h - p_1}{2h - 2hp_1} \leq 1$$

36 since $h \leq 1/2$ and thus $2hp_1 \leq p_1$.

37 So finally, the variance of the number of observations in the
 38 small square around (u, v) is larger than the variance of the num-
 39 ber of pseudo-observations in the same square. Therefore, this

01 larger uncertainty concerning the relevant sub-sample used in the
02 neighbourhood of (u, v) in the former case implies a loss of effi-
03 ciency. The consequence of this result is largely counterintuitive.
04 By working with pseudo-observations instead of “true” ones, we
05 would expect an additional noise, which should induce more noisy
06 estimated copula densities. This is not in fact the case as we have
07 just shown.

08

09

CONCLUDING REMARKS

10

11 We have discussed how various estimation procedures impact the
12 estimation of tail probabilities in a copula framework. Parametric
13 estimation may lead to severe underestimation when the paramet-
14 ric model of the margins and/or the copula is misspecified. Non-
15 parametric estimation may also lead to severe underestimation
16 when the smoothing method does not take into account potential
17 boundary biases in the corner of the density support. Since the
18 primary focus of most risk management procedures is to gauge
19 these tail probabilities, we think that the methods analysed above
20 might help to better understand the occurrence of extreme risks in
21 stand-alone positions (single asset) or inside a portfolio (multiple
22 assets). In particular we have shown that nonparametric methods
23 are simple, powerful visualisation tools that enable the detection
24 of dependencies among various risks. A clear assessment of these
25 dependencies should help in the design of better risk measurement
26 tools within a VAR or an expected shortfall framework.

27

28 **1** The third author acknowledges financial support by the National Centre of Competence in
29 Research “Financial Valuation and Risk Management” (NCCR FINRISK).

30 **2** For example, the number of histogram grid cells increases exponentially. This effect cannot
31 be avoided, even using other estimation methods. Under smoothness assumptions on
32 the density, the amount of training data required for nonparametric estimators increases
33 exponentially with the dimension (eg, see Stone (1980)).

33

34

REFERENCES

35

Alexander, C., 2002, *Market Models* (New York: John Wiley and Sons).

36

Behnen, K., M. Huskova, and G. Neuhaus, 1985, “Rank Estimators of Scores for Testing
37 Independence”, *Statistics and Decision*, 3, pp 239–262.

37

38

Bouezmarni, T. and J. M. Rolin, 2001, “Consistency of Beta Kernel Density Function Estima-
39 tor”, *The Canadian Journal of Statistics*, 31, pp 89–98.

39

- 01 **Bouezmarni, T. and J. M. Rolin**, 2003, "Bernstein Estimator for Unbounded Density Function",
02 Université Catholique de Louvain-la-Neuve.
- 03 **Brown, B. M. and S. X. Chen**, 1999, "Beta-Bernstein Smoothing for Regression Curves with
04 Compact Supports", *Scandinavian Journal of Statistics*, **26**, pp 47–59.
- 05 **Chen, S. X.**, 1999, "A Beta Kernel Estimation for Density Functions", *Computational Statistics
06 and Data analysis*, **31**, pp 131–135.
- 07 **Chen, S. X.**, 2000, "Beta Kernel for Regression Curve", *Statistica Sinica*, **10**, pp 73–92.
- 08 **Chen, S. X. and Y. Fan**, 2004a, "Efficient Semiparametric Estimation of Copulas", Working
09 Paper.
- 10 **Chen, S. X. and Y. Fan**, 2004b, "Copula-based Tests for Dynamic Models", Working Paper.
- 11 **Deheuvels, P.**, 1979, "La fonction de dépendance empirique et ses propriétés", *Acad. Roy. Belg.,
12 Bull. C1 Sci. 5ième sér.*, **65**, pp 274–292.
- 13 **Deheuvels, P.**, 1981a, "A Kolmogorov–Smirnov Type Test for Independence and Multivariate
14 Samples", *Rev. Roum. Math. Pures et Appl.*, **26**(2), pp 213–226.
- 15 **Deheuvels, P.**, 1981b, "A Nonparametric Test for Independence", *Pub. Inst. Stat. Univ. Paris*,
16 **26**(2), pp 29–50.
- 17 **Deheuvels, P. and P. Hominal**, 1979, "Estimation non paramétrique de la densité compte tenu
18 d'informations sur le support", *Revue de Statistique Appliquée*, **27**, pp 47–68.
- 19 **Devroye, L. and L. Györfi**, 1985, *Nonparametric Density Estimation: The L_1 View* (New York: John
20 Wiley and Sons).
- 21 **Doukhan, P., J.-D. Fermanian and G. Lang**, 2004, "Copulas of a Vector-valued Stationary
22 Weakly Dependent Process", Working Paper.
- 23 **Fermanian, J.-D., D. Radulovic and M. Wegkamp**, 2004, "Weak Convergence of Empirical
24 Copula Processes", *Bernoulli*, **10**, pp 847–860.
- 25 **Fermanian, J.-D. and O. Scaillet**, 2003, "Nonparametric Estimation of Copulas for Time
26 Series", *Journal of Risk*, **95**, pp 25–54.
- 27 **Fermanian, J.-D. and O. Scaillet**, 2005, "Some Statistical Pitfalls in Copula Modelling for
28 Financial Applications", in Klein, E. (ed), *Capital Formation, Governance and Banking* (New York:
29 Nova Science Publishers), pp 59–74.
- 30 **Gasser, T. and H. G. Müller**, 1979, "Kernel Estimation of Regression Functions", in Gasser, T.
31 and M. Rosenblatt (eds), *Smoothing Techniques for Curve Estimation*, Lecture Notes in Mathemat-
32 ics, **757** (Springer), pp 23–68.
- 33 **Genest, C., K. Ghoudi and L. P. Rivest**, 1995, "A Semiparametric Estimation Procedure of
34 Dependence Parameters in Multivariate Families of Distributions", *Biometrika*, **82**, pp 543–552.
- 35 **Gijbels, I. and J. Mielniczuk**, 1990, "Estimating the Density of a Copula Function", *Communi-
36 cations in Statistics: Theory and Methods*, **19**, pp 445–464.
- 37 **Härdle, W. and O. Linton**, 1994, "Applied Nonparametric Methods", in Engle, R. and
38 D. McFadden (eds), *Handbook of Econometrics*, **IV** (Amsterdam: North-Holland).
- 39 **Harrell, F. E. and C. E. Davis**, 1982, "A New Distribution-free Quantile Estimator", *Biometrika*,
69, pp 635–640.
- Jäckel, P.**, 2002, *Monte Carlo Methods in Finance* (New York: John Wiley and Sons).
- Jones, M. C.**, 1993, "Simple Boundary Correction for Kernel Density Estimation", *Statistics and
Computing*, **3**, pp 135–146.

- 01 **Li, D. X.**, 1999, "On Default Correlation: a Copula Function Approach", RiskMetrics Group,
02 Working Paper.
- 03 **Longin, F. and B. Solnik**, 1995, "Is the Correlation in International Equity Returns Constant:
04 1960–1990?", *Journal of International Money and Finance*, **14**, pp 3–26.
- 05 **Müller, H. G.**, 1991, "Smooth Optimum Kernel Estimators near Endpoints", *Biometrika*, **78**,
06 pp 521–530.
- 07 **Niederreiter, H.**, 1992, "Random Number Generation and Quasi-Monte Carlo Methods",
08 *CBMS-SIAM*, **63**, pp 86–112.
- 09 **Pagan, A. and A. Ullah**, 1999, *Nonparametric Econometrics* (Cambridge: Cambridge University
10 Press).
- 11 **Patton, A.**, 2001, "Modelling Time-Varying Exchange Rate Dependence Using the Conditional
12 Copula", University of California, San Diego, Discussion Paper, **01-09**.
- 13 **Renault, O. and O. Scaillet**, 2004, "On the Way to Recovery: a Nonparametric Bias Free
14 Estimation of Recovery Rate Densities", *Journal of Banking & Finance*, **28**, pp 2915–2931.
- 15 **Rice, J. A.**, 1984, "Boundary Modification for Kernel Regression", *Communication in Statistics*,
16 **12**, pp 899–900.
- 17 **Schuster, E.**, 1985, "Incorporating Support Constraints into Nonparametric Estimators of
18 Densities", *Communications in Statistics: Theory and Methods*, **14**, pp 1123–1136.
- 19 **Shih, J. H. and T. A. Louis**, 1995, "Inferences on the Association Parameter in Copula Models
20 for Bivariate Survival Data", *Biometrics*, **55**, pp 1384–1399.
- 21 **Silvapulle, P., G. Kim and M. J. Silvapulle**, 2004, "Robustness of a Semiparametric Estimator
22 of a Copula", *Econometric Society 2004 Australasian Meetings*, **317**.
- 23 **Silverman, B.**, 1986, *Density Estimation for Statistics and Data Analysis* (New York: Chapman
24 and Hall).
- 25 **Sklar, A.**, 1959, "Fonctions de répartition à n dimensions et leurs marges", *Publ. Inst. Statist.*
26 *Univ. Paris*, **8**, pp 229–231.
- 27 **Stone, C. J.**, 1980, "Optimal Rates of Convergence for Nonparametric Estimators", *Annals of*
28 *Probability*, **12**, pp 361–379.
- 29 **Wand, M. P. and M. C. Jones**, 1995, *Kernel Smoothing* (New York: Chapman and Hall).
- 30 **Wand, M. P., J. S. Marron and D. Ruppert**, 1991, "Transformations in Density Estima-
31 tion: Rejoinder (in Theory and Methods)", *Journal of the American Statistical Association*, **86**,
32 pp 360–361.
- 33
- 34
- 35
- 36
- 37
- 38
- 39