Dependence sampling and quasi-random numbers

Christiane Lemieux

University of Waterloo, Canada

MCQMC 2016

Stanford University, August 16th, 2016

Two sides to the story

- New results on the use of quasi-random numbers to sample from dependence models based on copulas
  $\rightarrow$ joint work with M. Cambou (Lausanne) and M. Hofert (Waterloo))

  including ways of using importance sampling
  $\rightarrow$ joint work with Yoshihiro Taniguchi (Waterloo)
- Viewing quasi-random numbers as a form of dependence sampling; use this idea to study the variance of the corresponding estimators

Outline

(1) QRN to sample dependence models

1. Problem setup and background
2. QMC-based simulation of copula models
3. Error behavior
4. Importance sampling for copula models
5. Numerical examples

(2) QRN as dependence sampling

1. Problem setup and background
2. Generalized Hoeffding's Lemma and covariance result for negatively upper orthant dependent (NUOD) points
3. Dependence structure of scrambled nets

1- Problem setup and background

- In many applications, in particular in finance and insurance, quantities of interest can be written as $\mathbb{E}[\Psi_0(\boldsymbol{X})] = \int \Psi_0(\boldsymbol{x})dH(\boldsymbol{x})$, where

$$\boldsymbol{X} = (X_1, \ldots, X_d) \in \mathbb{R}^d \text{ is a random vector with distribution function } H : \mathbb{R}^d \to [0,1]$$

$$\Psi_0 : \mathbb{R}^d \to \mathbb{R} \text{ is a function.}$$

- Examples: X_j (or e^{X_j} or $e^{q(X_j, X_{d/2+j})}$) might be the price of a financial asset at time j; X_j (or e^{X_j}) might the price at time T of the jth financial asset/risk.
- MC and QMC methods have been used widely on such problems, with d sometimes very large. In most cases, the dependence among the X_j's has been modeled by a multivariate normal distribution (MVN)

- When $\boldsymbol{X} \sim \text{MVN}(\mu, \Sigma)$, can generate $\boldsymbol{X}$ as follows: let $\boldsymbol{Z} = (Z_1, \ldots, Z_d)$ with Z_j's independent $\mathcal{N}(0,1)$ and then let

$$\boldsymbol{X} = A\boldsymbol{Z} + \boldsymbol{\mu},$$

  where A is such that $AA^T = \Sigma$.
- Includes methods such as Brownian Bridge and PCA.
- Can rewrite our problem in terms of independent $Z_1, \ldots, Z_d$:

$$\Psi_0(\boldsymbol{X}) = \Psi_0(g_1(\boldsymbol{Z})) \quad \text{where } \boldsymbol{X} = g_1(\boldsymbol{Z}) = A\boldsymbol{Z} + \boldsymbol{\mu} \qquad \text{(linear).}$$

- Can go one step further and rewrite the problem as a function of $\boldsymbol{U} \in [0,1)^d$ by writing $Z_j = \Phi^{-1}(U_j)$ where $U_1, \ldots, U_d$ are iid $U(0,1)$ and $\Phi(z) = P(\mathcal{N}(0,1) \leq z)$.
- Can then write

$$\mathbb{E}_X[\Psi_0(\boldsymbol{X})] = \mathbb{E}_Z[\Psi_0(g_1(\boldsymbol{Z}))] = \mathbb{E}_U[\Psi_0(g_1(g_2(\boldsymbol{U})))]$$

  where $g_2 : [0,1)^d \to \mathbb{R}^d$ is given by

$$g_2(\boldsymbol{U}) = (\Phi^{-1}(U_1), \ldots, \Phi^{-1}(U_d))^T.$$

- Idea of QMC is then to use a low-discrepancy point set $P_n = \{\mathbf{u}_i, i = 1, \ldots, n\}$ and approximate $\mathbb{E}_X[\Psi_0(\boldsymbol{X})]$ by

$$\frac{1}{n} \sum_{i=1}^{n} \Psi_0(g_1(g_2(\mathbf{u}_i))).$$

- There are alternative transformations that can be used to generate $\boldsymbol{Z}$, but g_2 is often preferred for QMC because it is monotone and is a bijection from $[0,1)^d$ to $\mathbb{R}^d$, while other alternatives might be from $[0,1)^k$ to $\mathbb{R}^d$ with $k > d$.

Beyond the MVN distribution

- MVN is too simple; especially not suitable for tail dependence, which tend to produce the largest values of $\Psi_0(\boldsymbol{X})$.
- Copulas offer a much more flexible tool for modeling, and are becoming increasingly important for risk management, especially for capital requirements due to *operational risk* (Basel Committee on Banking Supervision).

Background on copula models

- Let $H(\boldsymbol{x}) = P(X_1 \le x_1, \ldots, X_d \le x_d)$ be the joint CDF of $\boldsymbol{X}$.
  Then we can write *(Sklar's Theorem)*

$$H(\boldsymbol{x}) = C(F_1(x_1), \ldots, F_d(x_d)), \quad \boldsymbol{x} \in \mathbb{R}^d,$$

  where $F_j(x) = P(X_j \le x)$, are the marginal distribution functions of H and $C : [0,1]^d \to [0,1]$ is a *copula*, which is a distribution function, i.e., $C(u_1, \ldots, u_d) = P(U_1 \le u_1, \ldots, U_d \le u_d))$ with $U_j \sim U(0,1)$
- A copula model allows one to separate the dependence structure from the marginal distributions.
- For example, if $X_1, \ldots, X_d$ are independent, then

$$H(x_1, \ldots, x_d) = \prod_{j=1}^{d} F_j(x_j)$$

  so it means $C(u_1, \ldots, u_d) = \prod_{j=1}^{d} u_j$.

Examples of copulas

- Gaussian copula: is given by

$$C_P^{\mathrm{Ga}}(\boldsymbol{u}) = \Phi_P(\underbrace{\Phi^{-1}(u_1)}_{X_1}, \dots, \underbrace{\Phi^{-1}(u_d)}_{X_d}),$$

where $\Phi_P : \mathbb{R}^d \to [0,1]$ denotes the d-variate standard normal distribution function with correlation matrix P.

- A d-dimensional copula C is called *Archimedean* if it permits the representation

$$C(\boldsymbol{u}) = \psi(\psi^{-1}(u_1) + \cdots + \psi^{-1}(u_d)),$$

for some generator ψ: $[0,\infty] \to [0,1]$, which is a continuous, decreasing function ψ satisfying $\psi(0) = 1$, $\psi(\infty) = \lim_{t\to\infty} \psi(t) = 0$, and is strictly decreasing on $[0, \inf\{t : \psi(t) = 0\}]$.

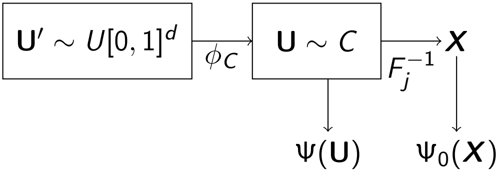
Example of an Archimedean copula: the Clayton copula

- Uses generator $\psi(t) = (1+t)^{-1/\theta}$, $t \geq 0$, $\theta \in [-1, \infty)\backslash\{0\}$.
- Means $C(u_1, u_2) = \max(u_1^{-\theta} + u_2^{-\theta} - 1, 0)^{-1/\theta}$.
- Most negative dependence: $\theta = -1$, so $C(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$ (corresponds to having $U_1 = 1 - U_2$.)
- Most positive dependence: $\theta \to \infty$, so that $C(u_1, u_2) = \min(u_1, u_2)$ (correponds to having $U_1 = U_2$.)

In terms of a copula model, we may write

$$\mathbb{E}[\Psi_0(\boldsymbol{X})] = \mathbb{E}_C[\Psi(\boldsymbol{U})] = \int_{[0,1)^d} \Psi(\mathbf{u}) dC(\mathbf{u}),$$

where $\boldsymbol{U} \sim C$, $\Psi(u_1, \ldots, u_d) = \Psi_0(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d))$, and $F_j^{-1}(p) = \inf\{x \in \mathbb{R} : F_j(x) \geq p\}$, $j \in \{1, \ldots, d\}$.



Hence to estimate $\mathbb{E}_C[\Psi(\boldsymbol{U})]$, proceed as follows:

1) sample $\boldsymbol{U}_i = \phi_C(\mathbf{U}'_i) \sim C$
2) transform each U_{ij} according to $x_{ij} = F_j^{-1}(U_{ij}), j = 1, \ldots, d$, and use

$$\frac{1}{n}\sum_{i=1}^{n} \Psi_0(\boldsymbol{x}_i) = \frac{1}{n}\sum_{i=1}^{n} \Psi(\boldsymbol{U}_i)$$

$\Rightarrow$ Main challenge is the sampling of the copula via ϕ_C

2- QMC-based simulation of copula models

- To sample a d-dimensional copula C, use transformation $\phi_C : [0,1)^k \to [0,1)^d$, $k \geq d$, such that

$$\boldsymbol{U} = \phi_C(\boldsymbol{U}') \sim C \text{ when } \boldsymbol{U}' \sim U(0,1)^d.$$

- In other words, ϕ_C transforms $U(0,1)^k$ random variables to random variables with distribution function C.
- With QMC-based sampling, prefer one-to-one $\phi_C : [0,1)^d \to [0,1)^d$. Would also like $\phi_C(\cdot) = (\phi_{C,1}(\cdot), \ldots, \phi_{C,d}(\cdot))$ such that *each* $\phi_{C,j} : [0,1]^d \to [0,1]$ *is monotone in each argument, and smooth*.
- Can use the conditional distribution method (CDM), which is the only known algorithm for sampling an arbitrary copula without further knowledge about its structure.

Conditional Distribution Method

For $j \in \{2, \ldots, d\}$, let

$$C(u_j \mid u_1, \ldots, u_{j-1}) = \mathbb{P}(U_j \leq u_j \mid U_1 = u_1, \ldots, U_{j-1} = u_{j-1})$$

denote the *conditional copula of U_j at u_j given $U_1 = u_1, \ldots, U_{j-1} = u_{j-1}$*. If $C^{-1}(u_j \mid u_1, \ldots, u_{j-1})$ denotes the corresponding quantile function, the CDM is given as follows (Schmitz, 2003):

Theorem (Conditional distribution method)

Let C be a d-dimensional copula, $\boldsymbol{U}' \sim U[0,1]^d$, and ϕ_C^{CDM} be given by

$$\begin{aligned} U_1 &= U_1' &&= \phi_{C,1}^{CDM}(U_1') \\ U_2 &= C^{-1}(U_2' \mid U_1) &&= \phi_{C,2}^{CDM}(U_1', U_2') \\ &\vdots \\ U_d &= C^{-1}(U_d' \mid U_1, \ldots, U_{d-1}) &&= \phi_{C,d}^{CDM}(U_1', \ldots, U_d'). \end{aligned}$$

Then $\boldsymbol{U} = (U_1, \ldots, U_d) = \phi_C^{CDM}(\boldsymbol{U}') \sim C$.

CDM for Archimedean copulas

Assuming ψ to be sufficiently often differentiable, the conditional Archimedean copulas follow from above theorem and are given by

$$C(u_j \mid u_1, \dots, u_{j-1}) = \frac{\psi^{(j-1)}(W_j)}{\psi^{(j-1)}(W_{j-1})},$$

where $W_j = \sum_{k=1}^{j} \psi^{-1}(u_k)$ and thus

$$C^{-1}(u_j \mid u_1, \dots, u_{j-1}) = \psi\left(\psi^{(j-1)^{-1}}\left(u_j \psi^{(j-1)}(W_{j-1})\right) - W_{j-1}\right).$$

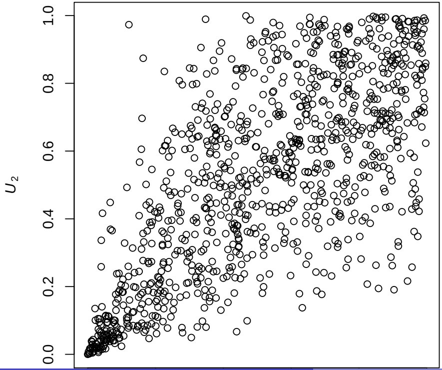
The generator derivatives $\psi^{(j-1)}$ and their inverses $\psi^{(j-1)^{-1}}$ can be challenging to compute.

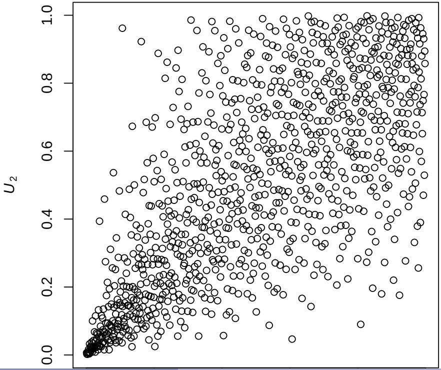
CDM for Clayton copula

For $\psi(t) = (1+t)^{-1/\theta}, \theta > 0$, then $\phi_{C,j}^{CDM}(\cdot)$ is given by

$$C^{-1}(u_j \mid u_1, \ldots, u_{j-1}) = \left(1 + \left(1 - (j-1) + \sum_{k=1}^{j-1} u_k^{-\theta}\right)\left(u_j^{-\frac{1}{j-1+1/\theta}} - 1\right)\right)^{-\frac{1}{\theta}}.$$

1000 realizations of a Clayton copula with $\theta = 2$ generated by MC (left) and 2-d Halton (right).





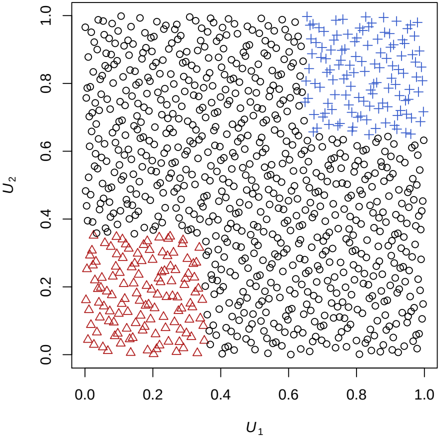
Alternative methods

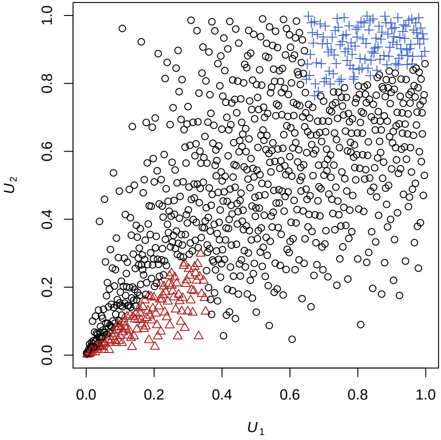
- Approaches other than the CDM are often used to sample copulas.
- Based on transformations $\phi_C : [0,1)^k \to [0,1)^d$ with $k > d$ making use of well-chosen *stochastic representation*
- For example, the Marshall–Olkin algorithm uses a *conditional independence approach*: based on the fact that for an Archimedean copula C with completely monotone generator ψ,

$$\boldsymbol{U} = (\psi(E_1/V), \ldots, \psi(E_d/V)) \sim C,$$

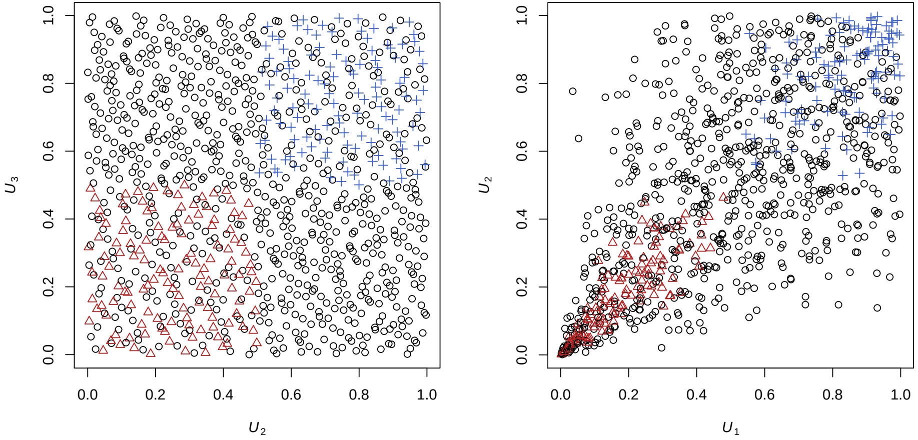
where $V \sim F = \mathcal{LS}^{-1}[\psi]$ (i.e., F is the distribution whose Laplace transform is ψ), independent of $E_1, \ldots, E_d \sim \mathrm{Exp}(1)$ (iid).
Here $k = d + 1$ if we use inversion to generate E_j's and V.

1000 realizations of the 1st two components of a 3d-Halton sequence with marked points ($\triangle$ and $+$) in the respective regions $[0, \sqrt{1/8}]^2$ and $[1 - \sqrt{1/8}, 1]^2$ (left): corresponding ϕ_C^{CDM}-transformed points (right) to a Clayton copula with $\theta = 2$





1000 realizations of 2nd & 3rd components of a 3d-Halton sequence with marked points ($\triangle$ and $+$) corresponding to respective regions $[0, 0.5]^3$ and $[0.5, 1]^3$ (left): corresponding ϕ_C^{MO}-transformed points (right) to a Clayton copula with $\theta = 2$.



3- Error behavior

When $\phi_C = \phi_C^{\text{CDM}}$, one can easily show that

$$\mathbb{E}_C\left[\Psi(\boldsymbol{U})\right] = \int_{[0,1]^d} \Psi(\phi_C(\boldsymbol{v}))d\boldsymbol{v}.$$

Natural to then work with $f = \Psi \circ \phi_C$ and usual star-discrepancy, i.e.,

Result: *Let $\boldsymbol{u}_i = \phi_C^{CDM}(\boldsymbol{v}_i)$ and $P_n = \{\boldsymbol{v}_i, i = 1, \ldots, n\}$ be a point set in $[0,1)^d$. Assume Ψ_0, C, and $F_1, \ldots, F_d$ are such that $V(\Psi \circ \phi_C^{CDM}) < \infty$, where $V(\cdot)$ is the variation in the sense of Hardy and Krause. Then*

$$\left| \frac{1}{n} \sum_{i=1}^{n} \Psi(\boldsymbol{u}_i) - \mathbb{E}_C[\Psi(\boldsymbol{U})] \right| \leq V(\Psi \circ \phi_C^{CDM}) D^*(P_n).$$

where $D^*(P_n)$ is the usual star-discrepancy.

But $V(\Psi) < \infty \not\Rightarrow V(\Psi \circ \phi_C^{CDM}) < \infty$: conditions under which it holds?

Conditions to have bounded variation with CDM

Proposition: Assume that Ψ has continuous mixed partial derivatives up to total order d and there exist $m, M, K > 0$ such that for all $\boldsymbol{u} \in (0,1)^d$, $c(\boldsymbol{u}) \geq m > 0$, $C_{i|1\ldots i-1} = C(u_i \mid u_1, \ldots, u_{i-1})$ and

$$\left| \frac{\partial^k C_{i|1\ldots i-1}}{\partial u_{\alpha_1} \cdots \partial u_{\alpha_k}} \right| \leq M, \ \alpha_1, \ldots, \alpha_k \in \{1, \ldots, i\}, \tag{1}$$

for each $1 \leq k \leq i \leq d$, $1 \leq k \leq l \leq d$ and $\{\alpha_1, \ldots, \alpha_l\} \subseteq \{1, \ldots, d\}$, we have

$$\left| \frac{\partial^k \Psi(u_1, \ldots, u_d)}{\partial u_{\beta_1} \cdots \partial u_{\beta_k}} \right| \leq K, \quad \beta_j \in \{\alpha_1, \ldots, \alpha_l\}, \ 1 \leq j \leq k. \tag{2}$$

Then there exists a constant $C^{(d)}$ s.t. for $\boldsymbol{u}_i = \phi_C^{\text{CDM}}(\boldsymbol{v}_i)$, $i = 1, \ldots, n$,

$$\left| \frac{1}{n} \sum_{i=1}^{n} \Psi(\boldsymbol{u}_i) - \mathbb{E}[\Psi(\boldsymbol{U})] \right| \leq D^*(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n) K C^{(d)} (M^d/m)^{2d-1}.$$

Remarks on this result

- Proof is a "translation" of Hlawka and Mück (1972)
- This approach avoids having to deal with ϕ_C and its derivatives
- Unfortunately (1) does not often hold for typical copulas
- Conditions in (2) are stronger than those required to have

$$\|\Psi\|_{d,1} = \sum_{l=1}^{d} \sum_{\boldsymbol{\alpha}} \int_{[0,1)^l} \left| \frac{\partial^l \Psi(u_{\alpha_1}, \ldots, u_{\alpha_l}, \mathbf{1})}{\partial u_{\alpha_1} \cdots \partial u_{\alpha_l}} \right| du_{\alpha_1} \ldots du_{\alpha_l}$$

bounded

Results for MO

- Here we need to decompose mixed partial derivatives of the form

$$\frac{\partial^l (\Psi \circ \phi_C)(v_{\alpha_1}, \ldots, v_{\alpha_l}, \mathbf{1})}{\partial v_{\alpha_1} \cdots \partial v_{\alpha_l}}$$

  in terms of Ψ and $\phi_C(\cdot) = (\phi_{C,1}(\cdot), \ldots, \phi_{C,d}(\cdot))$ separately.
- We use Theorem 2.1 from Constantine and Savits (1996) and get

$$\frac{\partial^l \Psi \circ \phi_C(v_{\alpha_1}, \ldots, v_{\alpha_l}, \mathbf{1})}{\partial v_{\alpha_1} \cdots \partial v_{\alpha_l}} = \sum_{1 \le |\boldsymbol{\beta}| \le l} \frac{\partial^{|\boldsymbol{\beta}|} \Psi}{\partial^{\beta_1} u_1 \ldots \partial^{\beta_d} u_d} \sum_{s=1}^{l} \sum_{\gamma, \boldsymbol{k}} c_\gamma \prod_{j=1}^{s} \frac{\partial^{|\gamma_j|} \phi_{C,k_j}(v_{\alpha_1}, \ldots, v_{\alpha_l}, \mathbf{1})}{\partial^{\gamma_{j,1}} v_{\alpha_1} \ldots \partial^{\gamma_{j,l}} v_{\alpha_l}} \quad (3)$$

where $\boldsymbol{\beta} \in \mathbb{N}_0^d$ and $|\boldsymbol{\beta}| = \sum_{j=1}^{d} \beta_j$.

- From (3), we see that a sufficient condition to show that $\|\Psi \circ \Phi_C\|_{d,1} < \infty$ is to establish that all products of the form

$$\frac{\partial^{|\boldsymbol{\beta}|}\Psi}{\partial^{\beta_1} u_1 \ldots \partial^{\beta_d} u_d} \prod_{j=1}^{s} \frac{\partial^{|\gamma_j|}\phi_{C,k_j}(v_{\alpha_1}, \ldots, v_{\alpha_l}, \mathbf{1})}{\partial^{\gamma_{j,1}} v_{\alpha_1} \ldots \partial^{\gamma_{j,l}} v_{\alpha_l}}, \; s \in \{1, \ldots, l\}, \quad (4)$$

  are in L_1.
- In MO algorithm, if
  - v_1 is used to generate V
  - v_{j+1} is used to generate E_j

  then $\phi_{C,j}$ is a function of v_1 and v_{j+1} only, for $j \in \{1, \ldots, d\}$.
  $\Rightarrow$ the only partial derivatives of $\phi_{C,j}$ that are nonzero are those with respect to variables in $\{v_1, v_{j+1}\}$.

Error behaviour for MO for continuous V

Proposition: Let $P_n = \{\boldsymbol{v}_i, i = 1, \ldots, n\} \subseteq [0,1)^{d+1}$ and $\boldsymbol{u}_i = \phi_C^{\mathrm{MO}}(\boldsymbol{v}_i)$.
Assume $V \sim F$ is continuously distributed and that:

1. the point set P_n excludes the origin and there exists some $p \geq 1$ such that $\min_{1 \leq i \leq n} v_{i,1} \geq 1/pn$;
2. the function Ψ satisfies $|\Psi(\boldsymbol{u})| < \infty$ for all $\boldsymbol{u} \in [0,1)^{d+1}$ and
$$\frac{\partial^{|\boldsymbol{\beta}|}\Psi}{\partial^{\beta_1} u_1 \ldots \partial^{\beta_d} u_d} < \infty \text{ for all } \boldsymbol{\beta} = (\beta_1, \ldots, \beta_d), \tag{5}$$
with $\beta_l \in \{0, \ldots, d\}$ and $|\boldsymbol{\beta}| \leq d$;
3. the generator $\psi(\cdot)$ is such that $\psi'(t) + t\psi''(t)$ has at most one zero t^* in $(0, \infty)$ and it satisfies $-t^*\psi'(t^*) < \infty$; and
4. $F^{-1}(1 - 1/pn)$ is in $O(n^a)$ for some constant $a > 0$;

then
$$\left| \frac{1}{n} \sum_{i=1}^{n} \Psi(\boldsymbol{u}_i) - \mathbb{E}[\Psi(\boldsymbol{U})] \right| \leq C^{(d)}(\log n) D^*(P_n).$$

Error behaviour for MO algorithm for discrete V

Proposition Let $P_n = \{\boldsymbol{v}_i, i = 1, \ldots, n\} \subseteq [0,1)^{d+1}$ and $\boldsymbol{u}_i = \phi_C^{\text{MO}}(\boldsymbol{v}_i)$. Assume $V \sim F$ is discrete. If (5) holds and

1. there exists some $p \geq 1$ such that the point set P_n satisfies $\max_{1 \leq i \leq n} v_{i,1} \leq 1 - 1/pn$;
2. there exist constants $c > 0$ and $q \in (0,1)$ such that $1 - F(l) \leq cq^l$ for $l \geq 1$;

then there exists a constant $C^{(d)}$ such that

$$\left| \frac{1}{n} \sum_{i=1}^{n} \Psi(\boldsymbol{u}_i) - \mathbb{E}[\Psi(\boldsymbol{U})] \right| \leq C^{(d)} (\log n) D^*(P_n).$$

Copula-based discrepancy

- Similarly to the usual Lebesgue discrepancy used with KH, one can define the discrepancy with respect to a copula measure C as

$$D_C^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n) = \sup_{B \in \mathcal{B}} \left| \frac{\#\{\boldsymbol{u}_i \in B\}}{n} - C(B) \right|, \tag{6}$$

  where $\mathcal{B}$ is again restricted to hyperrectangles with a vertex at $\mathbf{0}$ *(e.g., Hlawka and Mück (1972), Aistleitner and Dick (2014))*.
- Could then use result from Aistleitner and Dick (2014):

Theorem

Let $\{\boldsymbol{u}_i, i = 1, \ldots, n\}$ be a set of points in $[0,1]^d$. Assume Ψ_0 and $F_1, \ldots, F_d$ are such that $V(\Psi) < \infty$. Then we have

$$\left| \frac{1}{n} \sum_{i=1}^{n} \Psi(\boldsymbol{u}_i) - \mathbb{E}_C[\Psi(\boldsymbol{U})] \right| \leq V(\Psi) D_C^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n).$$

Remarks on this approach

- General copulas C are not in a product form. So if $B \in \mathcal{B}$, then $\phi_C^{-1}(B)$ may not be in $\mathcal{B}$, so cannot easily infer that good behaviour of $D^*(\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ implies good behaviour of $D_C^*(\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$ when using $\mathbf{u}_i = \phi_C(\mathbf{v}_i)$.
- Very conservative bounds based on (uniform) discrepancy of original low-discrepancy point set exists
- Could we directly construct point sets that minimize copula-based discrepancy?

4- Importance sampling for copula models

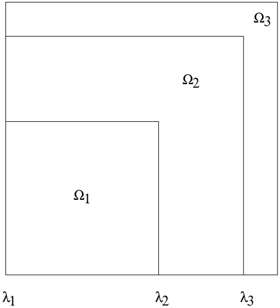
- Idea proposed in Arbenz, Cambou, and Hofert (2015)
- Large losses typically occur when at least one asset is large
- In terms of copula, want to change measure so that $M(\mathbf{U}) := \max_{1 \le i \le d} U_i$ close to 1 occurs more often
- How? Define threshold rv Λ with distribution π_Λ; generate $\Lambda = \lambda$ then generate $\mathbf{U}$ from C given $\max_{1 \le i \le d} U_i > \lambda$
- ACH propose to sample using CDM approach or rejection sampling; propose heuristic to choose discrete π_Λ characterized by probabilities $q_k = P(\Lambda = \lambda_k)$, for fixed threshold $\lambda_1, \ldots, \lambda_L$.
- Corresponding IS estimator is of the form

$$\hat{\mu}_{imp} = \frac{1}{n} \sum_{i=1}^{n} \Psi(\mathbf{U}_i) \omega(\mathbf{U}_i)$$

where $\omega(\mathbf{u})^{-1} = \sum_{k=1}^{L} \mathbf{1}(\lambda_k \le M(\mathbf{u})) q_k / (1 - C(\lambda_k, \ldots, \lambda_k))$.

4- IS for copula models – continued

- We proposed to use this idea within MO algorithm
- Requires to sample from $(E_1, \ldots, E_d, V)|M(\mathbf{U}) \geq \lambda_k$, where E_i's are iid $\mathrm{Exp}(1)$ and $V \sim F$
- Analyze the variance by making connection with stratified sampling

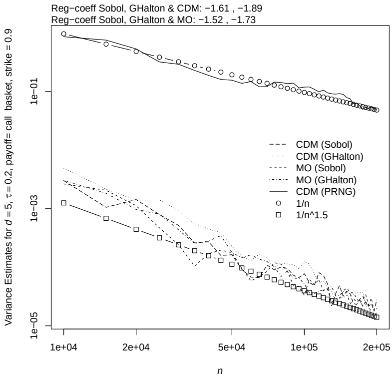


4- IS for copula models – continued

- Can find the q_k's that minimize variance of this IS estimator (requires estimates of $\mu_k^{(2)} = \mathrm{E}(\Psi^2(\mathbf{U})|\mathbf{U} \in \Omega_k)$, $k = 1, \ldots, L$).
- **Result:** With those optimal q_k's and under the condition $\mu_1^{(2)} \leq \ldots \leq \mu_L^{(2)}$ we have that $\mathrm{Var}(\hat{\mu}_{imp}) \leq \mathrm{Var}(\hat{\mu}_{mc})$.
- Can also directly define a stratified sampling estimator using the above stratas
- Requires to sample from $(E_1, \ldots, E_d, V)|\lambda_k \leq M(\mathbf{U}) < \lambda_{k+1}$;
- Corresponding estimator with optimal strata allocation has variance no larger than IS estimator

5- Numerical Examples

- Asset prices X_j for jth asset at time T, for $j = 1, \ldots, d$.
- Same marginal distribution for each j but dependence across assets induced by copula.
- Consider functionals of the aggregate sum $S = X_1 + \ldots + X_d$:
- Basket call option: $\Psi_0(S) = \max(0, S/d - K)$.
- Value-at-Risk: $\mathrm{VaR}_{0.99}(S) = F_S^{-1}(0.99) = \inf\{x \in \mathbb{R} : F_S(x) \geq 0.99\}$
- Expected shortfall: $\mathrm{ES}_{0.99}(S) = \frac{1}{1-0.99} \int_{0.99}^{1} F_S^{-1}(u) du$



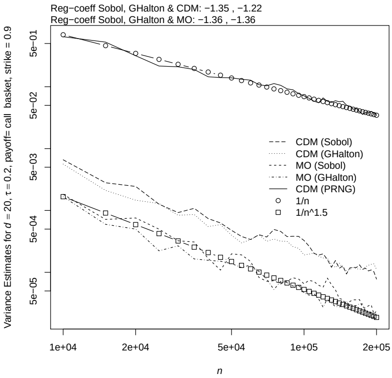


Figure : Variance estimates for the functional Basket Call with lognormal margins based on $B = 25$ repetitions for a Clayton copula with parameter such that Kendall's tau equals 0.2 for $d = 5$ (left) and $d = 20$ (right).

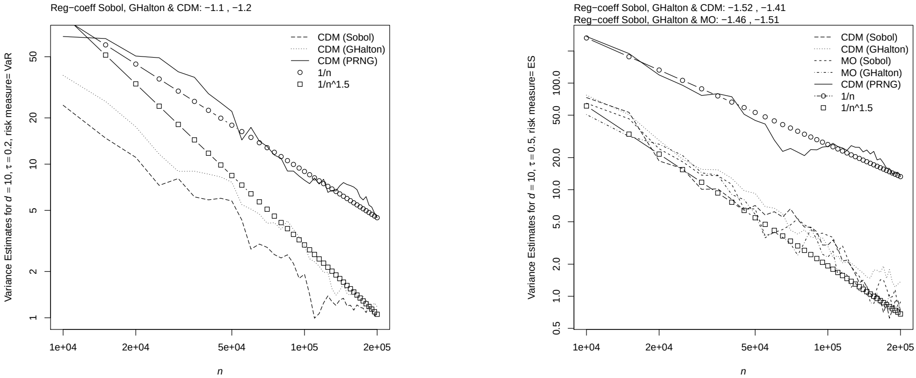


Figure : Variance estimates for the functional $\mathrm{Var}_{0.99}$ with lognormal margins and exchangeable t copula (left); and for $\mathrm{ES}_{0.99}$ with Pareto margins for a Clayton copula (right) for $d = 10$

Table : Var. Red. Factor for Clayton Copula, $d = 5$, $n = 20000$

fct	IS	Str
$\mathrm{E}[(S-K)^+]$	14	34
$\mathrm{VaR}_{0.995}(S)$	25	56
$\mathrm{CVaR}_{0.99}(S)$	36	57

Table : VRF for QMC Clayton Copula, $d = 5$, $n = 20000$

fct	Plain QMC	IS-min	Str
$\mathrm{E}[(S-K)^+]$	13.3	250	927
$\mathrm{VaR}_{0.995}(S)$	5.7	225	142
$\mathrm{CVaR}_{0.99}(S)$	6.2	756	880

QRN as dependence sampling

Motivation:

- Methods such as Antithetic Variates and Latin Hypercube Sampling can be shown to never be worse than MC for functions monotone in each coordinate: what about (R)QMC?
- It is of interest to have this type of upper bounds with easily verifiable conditions on integrands.

Problem setup and background

- Start with randomized QMC point set $\tilde{P}_n$ where each $\mathbf{U}_i \sim U(0,1)^d$ but the $\mathbf{U}_i$'s are dependent
- Assess its dependence via distribution function

$$H(\mathbf{u},\mathbf{v};\tilde{P}_n) := \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j>i} P(\mathbf{U}_i \leq \mathbf{u}, \mathbf{U}_j \leq \mathbf{v}). \tag{7}$$

- Useful to study variance of $\hat{\mu}_n = \sum_{i=1}^{n} f(\mathbf{U}_i)/n$ because

$$\mathrm{Var}(\hat{\mu}_n) = \frac{\sigma^2}{n} + \frac{n-1}{n}\sigma_{I,J}, \tag{8}$$

where $\sigma^2 = \mathrm{Var}(f(\mathbf{U}_i))$ and

$$\sigma_{I,J} = \mathrm{Cov}(f(\mathbf{U}_I), f(\mathbf{U}_J)) = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j>i} \mathrm{Cov}(f(\mathbf{U}_i), f(\mathbf{U}_j)).$$

- Negative Quadrant Dependence (NQD): We say X and Y are NQD if

$$P(X \leq x, Y \leq y) \leq P(X \leq x)P(Y \leq y).$$

- Hoeffding's Lemma: If X and Y are two random variables with joint cdf $H(x,y)$ and respective marginal distributions $F(x)$ and $G(y)$, then

$$\mathrm{Cov}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[H(x,y) - F(x)G(y)]dxdy.$$

- ($\approx$ Lehmann's result:) if $\tilde{P}_n$ has independent pairwise NQD coordinates, i.e.,

$$P(\mathbf{U}_i \leq \mathbf{u}, \mathbf{U}_j \leq \mathbf{v}) = \prod_{l=1}^{d} P(U_{i,l} \leq u_l, U_{j,l} \leq v_l) \leq \prod_{l=1}^{d} u_l v_l$$

and f is monotone wrt each coordinate, then $\mathrm{Cov}(f(\mathbf{U}_i), f(\mathbf{U}_j)) \leq 0$ and thus $\mathrm{Var}(\hat{\mu}_n) \leq \mathrm{Var}(\hat{\mu}_{MC})$.

Beyond dimension 2 and pairwise NQD

- A vector $\mathbf{X}$ of rv's is Negative Upper Orthant Dependent (NUOD) if

$$P(X_1 \geq x_1, \ldots, X_d \geq x_d) \leq \prod_{l=1}^{d} P(X_l \geq x_l).$$

- NUOD Sampling scheme:

$$T(\mathbf{u}, \mathbf{v}; \tilde{P}_n) := \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{j>i} P(\mathbf{U}_i \geq \mathbf{u}, \mathbf{U}_j \geq \mathbf{v}) \geq \prod_{l=1}^{s} (1-u_l)(1-v_l)$$

Generalized Hoeffding's Lemma

- Multivariate integration by parts (Mardia and Thompson 1972):
- Let μ_F be a finite signed measure over $[0,1]^s$. Assume that f is integrable with respect to μ_F over $[0,1)^s$ and that it admits the representation

$$f(\mathbf{u}) = \int_{\mathbf{v} \leq \mathbf{u}} g(\mathbf{v}) d\eta_f(\mathbf{v}), \tag{9}$$

for every $\mathbf{u} \in [0,1)^s$, where η_f is a finite signed Borel measure over $[0,1]^s$, and g is a Borel-measurable integrable function with respect to $\mu_F \times \eta$. Then we have

$$\int_{[0,1)^s} f(\mathbf{u}) d\mu_F(\mathbf{u}) = \int_{[0,1)^s} g(\mathbf{u}) \mu_F([\mathbf{u}, \mathbf{1}]) d\eta_f(\mathbf{u}).$$

Remarks on this result

- Other variants exist (e.g., papers by Young (1917))
- Condition on f is weaker than asking for bounded HK variation

Generalization of Hoeffding's lemma

Lemma

Let $\mathbf{U}$ *and* $\mathbf{V}$ *be random vectors over* $[0,1]^d$ *with joint survival function* $T(\mathbf{u},\mathbf{v})$ *and marginal survival functions* $R(\mathbf{u})$ *and* $S(\mathbf{v})$*, respectively. Let* f *be a function defined over* $[0,1]^s$ *satisfying the representation* (9) *with function* g *and measure* η_f*. Then*

$$\mathrm{Cov}(f(\mathbf{U}), f(\mathbf{V})) = \int_{[0,1]^{2d}} (T(\mathbf{u},\mathbf{v}) - R(\mathbf{u})S(\mathbf{v}))g(\mathbf{u})g(\mathbf{v})d\eta_f(\mathbf{u})d\eta_f(\mathbf{v}),$$

assuming f *satisfies all integrability conditions required for this covariance term to be well defined.*

Result: Let f be a function defined over $[0,1]^d$ satisfying the representation (9) with each of g and η_f not changing sign over their domain.
If $\tilde{P}_n$ is an NUOD sampling scheme with corresponding estimator $\hat{\mu}_n$, then

$$\mathrm{Var}(\hat{\mu}_n) = \mathrm{Var}(\hat{\mu}_{mc,n}) + (n-1)\sigma_{I,J}/n \leq \mathrm{Var}(\hat{\mu}_{mc,n}),$$

where the covariance term $\sigma_{I,J} = \mathrm{Cov}(f(\mathbf{U}_I), f(\mathbf{U}_J))$ is given by

$$\sigma_{I,J} = \int_{[0,1]^{2d}} [T(\mathbf{u},\mathbf{v};\tilde{P}_n) - \prod_{i=1}^{s}(1-u_i)\prod_{j=1}^{s}(1-v_j)]g(\mathbf{u})g(\mathbf{v})d\eta_f(\mathbf{u})d\eta_f(\mathbf{v}).$$

Variance reduction with Conditional NQD property

Result:

- Let $(X_1, \ldots, X_s)$ and $(Y_1, \ldots, Y_s)$ be vectors of random variables with the property that the distribution of (X_1, Y_1) and the conditional distribution of (X_j, Y_j) given $(X_1, Y_1), \ldots, (X_{j-1}, Y_{j-1})$ for each $j = 2, \ldots, s$ are NQD. (Conditional NQD property)
- Let f and g be functions of s variables and let $X = f(X_1, \ldots, X_s)$, $Y = g(Y_1, \ldots, Y_s)$.
- Then $\mathrm{Cov}(X, Y) \leq 0$ if f and g are concordant in the jth coordinate, for $j = 1, \ldots, s$.

Dependence structure of scrambled nets

- Can prove that a scrambled $(0, m, 2)$-net is an NUOD sampling scheme.
- Can also prove pairs are conditionally NQD as in previous result.
- Coordinates of pairs of points are not independent.
- Using previous result we get:

Result: Let f be a function defined over $[0, 1]^2$ that is monotone in each coordinate. Let $\hat{\mu}_n$ be the estimator for $I(f)$ based on a scrambled $(0, m, 2)$-net in base b, and let $\hat{\mu}_{mc,n}$ be the MC estimator for $I(f)$ based on $n = b^m$ points. Then

$$\mathrm{Var}(\hat{\mu}_n) \leq \mathrm{Var}(\hat{\mu}_{mc,n}).$$