Notes on hedging cryptos with spectral risk measures

March 15, 2021

Abstract

We investigate different methods of hedging cryptocurrencies with Bitcoin futures. A useful generalisation of variance-based hedging uses spectral risk measures and copulas.

Contents

1 TODO **4**

2 Optimal hedge ratio **4**

3 Spectral risk measures **7**

3.1 Representation of spectral risk measures 7

3.2 Exponential spectral risk measures 8

4 Hedge effectiveness **9**

5 D_1 Operator **9**

6 Special copulas 10
6.1 Gaussian Copula . . . 11
6.2 Archimedean copulas . . . 14
6.3 Elliptical Copulas and normal variance mixtures . . . 15
6.4 Normal mean-variance mixtures . . . 18
6.5 Normal inverse Gaussian factor copula model . . . 20
6.5.1 Normal inverse Gaussian distribution . . . 20
6.5.2 NIG factor copula . . . 23
6.5.3 Fitting the NIG factor model . . . 25
6.5.4 Fitting the NIG factor copula . . . 26
6.5.4.1 Bitcoin data . . . 28
6.5.5 Hedge calculation in the NIG factor model . . . 29
6.6 t-copulae . . . 34
6.6.1 Vanilla t-copula . . . 35
6.6.2 Skewed t copula . . . 35
6.6.3 Double-t copula . . . 35
6.6.4 Normal Inverse Gaussian Copula . . . 36

7 Estimation 38

8 Estimation 38
8.1 Simulated Method of Moments . . . 38
8.2 Maximum Likelihood Estimation . . . 40
8.3 Comparison . . . 41

9 Backtesting 42

10 Backtesting value-at-risk 42

1. TODO

- Discussion on calibration using method of moments. Make clear that creating a hedge with one instrument is linear, so only co-movement can be hedged. With BTC we see quite a few counter-movement events, so this might be idiosyncratic risk, but it might also point to another risk factor that could be hedged separately
- Basis risk: the risk that prices of financial instruments in a hedging strategy are imperfectly correlated, reducing the effectiveness of the hedging strategy (see `Minimum-capital-requirements-for-market-risk-BIS-2019.pdf`).

2. Optimal hedge ratio

Following (Barbi and Romagnoli, 2014), we consider the problem of optimal hedge ratios by extending the commmonly known minimum variance hedge ratio to more general risk measures and dependence structures.

Hedge portfolio: $R_t^h = R_t^S - hR_t^F$, involving returns of spot and future contract and where h is the hedge ratio
Optimal hedge ratio: $h^* = \operatorname{argmin}_h \rho_\phi(s, h)$, for given confidence level $1 - s$ (if applicable, e.g. in the case of VaR, ES), where ρ_ϕ is a spectral risk measure with weighting function ϕ (see below).
The distribution function of R^h can be expressed in terms of the copula and the marginal distributions as Proposition 1 result shows (this is a corrected version of Corollary 2.1 of (Barbi and Romagnoli, 2014)). For practical applications, it is numerically faster and more stable to use additional information about the specific copula and marginal distributions. We therefore derive semi-analytic formulas for a number of special cases, such as the Gaussian-, Student t-, normal inverse Gaussian (NIG) and Archimedean copulas in Section 6.

Proposition 1. *Let R^S and R^F be two real-valued random variables on the same probability space $(\Omega, \mathcal{A}, \mathbf{P})$ with corresponding absolutely continuous copula $C_{R^S,R^F}s$ and continuous marginals F_{R^S} and F_{R^F}. Then, the distribution of of R^h is given by*

$$F_{R^h}(x) = 1 - \int_0^1 D_1 C_{R^S,R^F}\left(u, F_{R^F}\left(\frac{F_{R^S}^{-1}(u) - x}{h}\right)\right) \mathrm{d}u. \tag{1}$$

Here $D_1C(u,v) = \dfrac{\partial}{\partial u}C(u,v)$, which is easily shown to fulfil, see e.g. Equation (5.15) of (McNeil *et al.*, 2005):[1]

$$D_1C_{X,Y}(F_X(x), F_Y(y)) = \mathbf{P}(Y \leq y | X = x). \tag{2}$$

Proof. Using the identity (2) gives

$$F_{R^h}(x) = \mathbf{P}(R^s - hR^F \leq x) = \mathbb{E}\left[\mathbf{P}\left(R^F \geq \frac{R^S - x}{h} \Big| R^S\right)\right]$$

$$= 1 - \mathbb{E}\left[\mathbf{P}\left(R^F \leq \frac{R^S - x}{h} \Big| R^S\right)\right] = 1 - \int_0^1 D_1C_{R^S,R^F}\left(u, F_{R^F}\left(\frac{F_{R^S}^{(-1)}(u) - x}{h}\right)\right) \mathrm{d}u.$$

□

In addition to Barbi and Romagnoli (2014) we propose a more handy expression for the density of r^h

Proposition 2.1. With the same setting of the above proposition, the density of r^h can be written as

$$f_{r^h}(y) = \left|\frac{1}{h}\right| \int_0^1 c_{r^S,r^F}\left[u, F_{r^F}\left\{\frac{F_{r^S}^{-1}(u) - y}{h}\right\}\right] \cdot f_{r^F}\left\{\frac{F_{r^S}^{-1}(u) - y}{h}\right\} du \tag{3}$$

[1] Let $F_X(x) = u$, $F_Y(y) = v$. Then, formally,

$$\begin{aligned}
\frac{\partial}{\partial F_X(x)}C(F_X(x), F_Y(y)) &= \frac{\partial}{\partial F_X(x)}\mathbf{P}(U \leq F_X(x), V \leq F_Y(y)) = \mathbf{P}(U \in \mathrm{d}F_X(x), V \leq F_Y(y)) \\
&= \mathbf{P}(V \leq F_Y(y) | U = F_X(x)) \cdot \mathbf{P}(U \in \mathrm{d}F_X(x)) = \mathbf{P}(Y \leq y | X = x) \cdot \mathbf{P}(U \in \mathrm{d}u) \\
&= \mathbf{P}(Y \leq y | X = x).
\end{aligned}$$

, or

$$f_{r^h}(y) = \int_0^1 c_{r^S,r^F}\left[u, F_{r^S}\left\{y + hF^{-1}_{r^F}(u)\right\}\right] \cdot f_{r^S}\left\{y + hF^{-1}_{r^F}(u).\right\} du \tag{4}$$

The two expression are equivalent. One can use any of them to get the density of r^h. Notice that the density of r^h in the above proposition is readily accessible as long as we have the copula density and the marginal densities. A generic expression can be found in the appendix (TODO).

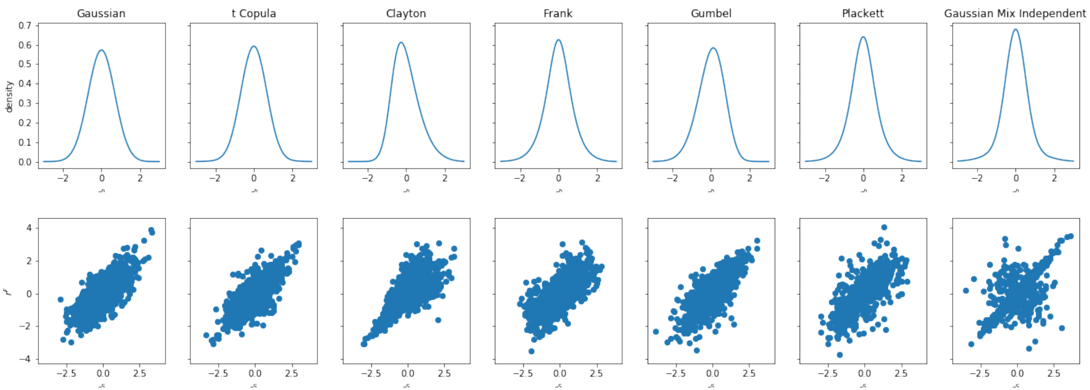


Figure 1: Upper Panel: Density of $R^h = R^S - hR^F$ of different copulas with $R^S, R^F \sim N(0,1)$, 0.75 Spearman's rho between R^S and R^F, and $h = 0.5$; Lower Panel: Scatter plot of samples from copulas. This illustration shows how dependence structure modelled by copulas affects the density of the linear combination of margins. Notice that the r^h modelled by the asymmetric copulas namely the Clayton and Gumbel copulas are skewed to right and left respectively.

3. Spectral risk measures

Spectral risk measure (Acerbi, 2002; Cotter and Dowd, 2006):

$$\rho_\phi = -\int_0^1 \phi(p)\, q_p \,\mathrm{d}p,$$

where q_p is the p-quantile of the return distribution and $\phi(s)$, $s \in [0,1]$, is the so-called *risk aversion function*, a weighting function such that[2]

(i) $\phi(p) \geq 0$,

(ii) $\int_0^1 \phi(p)\,\mathrm{d}p = 1$,

(iii) $\phi'(p) \leq 0$.

Examples: VaR, ES
Replacing the last property with $\phi'(p) > 0$ rules out risk-neutral behaviour.
Spectral risk measures are coherent (Acerbi, 2002).

3.1. Representation of spectral risk measures

To prevent numerical instabilities involving the quantile function, re-write spectral risk measures as follows:

- Integration by substitution: $\displaystyle\int_a^b g(\varphi(x))\,\varphi'(x)\,\mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} g(u)\,\mathrm{d}u.$
- Spectral risk measures: $\displaystyle -\int_0^1 \phi(p)\, F^{(-1)}(p)\,\mathrm{d}p$

[2] Note that the treatment in (Acerbi, 2002) is measure-based and therefore slightly different.

- Set $\varphi(x) = F(x)$, $g(p) = \phi(p)\, F^{(-1)}(p)$.
- Then:
$$-\int_0^1 \phi(p)\, F^{(-1)}(p)\, \mathrm{d}p = -\int_{-\infty}^{\infty} \phi(F(x))\, x\, f(x)\, \mathrm{d}x.$$

3.2. Exponential spectral risk measures

- Choose exponential utiliy function: $U(x) = -\mathbf{e}^{-kx}$, where $k > 0$ is the Arrow-Pratt coefficient of absolute risk aversion (ARA).
- Coefficient of absolute risk aversion: $R_A(x) = -\dfrac{U''(x)}{U'(x)} = k$
- Coefficient of relative risk aversion: $R_R(x) = -\dfrac{xU''(x)}{U'(x)} = xk$
- Weighting function $\phi(p) = \lambda \mathbf{e}^{-k(1-p)}$, where λ is an unknown positive constant.
- Set $\lambda = \dfrac{k}{1-\mathbf{e}^{-k}}$ to satisfy normalisation.
- Exponential spectral risk measure:
$$\rho_\phi = \int_0^1 \phi(p)\, F^{(-1)}(p)\, \mathrm{d}p = \frac{k}{1-\mathbf{e}^{-k}} \int_0^1 \mathbf{e}^{-k(1-p)}\, F^{(-1)}(p)\, \mathrm{d}p.$$
(If calculation of quantiles is a problem use change of variables above.)
- What exactly is the link between risk measure and utility? I think there is no direct link: the exponential risk measure is *inspired* by ARA utility.

4. Hedge effectiveness

(Ederington, 1979): hedge effectiveness similar to R^2 (Barbi R.): similar, but with actual risk measures (e.g. VaR) This uses the optimisation criterion to check out-of-sample performance of each method. It'll be interesting to compare across methods in which respect methods meet their target.
Also use P&L to compare across methods, this might give insights to what extent methods achieve their objective and how this compares across methods (e.g. hedging general risk vs. hedging tail risk).

5. D_1 Operator

Introduce copulas first
The D_1 operator is given as

$$D_1 C_{X,Y}(F_X(x), F_Y(y)) = \mathbf{P}(Y \leq y | X = x).$$

In the context of the above notation, we obtain

$$\begin{aligned} D_1 C_{R^s,R^F}\{w, g(w)\} &= \mathbf{P}[R_F \leq F_F^{(-1)}\{g(w)\} | R_s = F_S^{(-1)}(w)] = \mathbf{P}\{V \leq g(w) | U = w\} \\ &= \frac{\mathbf{P}\{U \in \mathrm{d}w, V \leq g(w)\}}{\mathbf{P}(U \in \mathrm{d}w)} = \mathbf{P}\{U \in \mathrm{d}w | V \leq g(w)\} = \int_0^{g(w)} c(w, v) \, \mathrm{d}v. \end{aligned}$$

The last line can also be written as Which statement in the last line does the refer to?

$$\frac{\partial}{\partial w} C\{w, g(w')\}\Big|_{w'=w}.$$

We give an explicit equation of the D_1 operator for Archimedean copulae.
The D_1 operator is defined as the partial derivatives of the first input to the copula function, so we fix the second argument while taking derivative with respect to the first, and then evaluate the function. we have

Function	Gumbel	Frank	Clayton	Independence
$\phi(t)$	$\{-\log(t)\}^{\theta}$	$-\ln\left\{\frac{\exp(-\theta t)-1}{\exp(-\theta)-1}\right\}$	$\frac{1}{\theta}(t^{-\theta}-1)$	Same to Gumbel where $\theta=1$
$\phi^{-1}(t)$	$\exp(-t^{1/\theta})$	$\frac{-1}{\theta}\log[1+\exp(-t)\{\exp(-\theta)-1\}]$	$(1+\theta t)^{-\frac{1}{\theta}}$	
$\partial\phi(t)/\partial t$	$\theta\frac{\phi(t)}{t\log(t)}$	$\frac{\theta\exp(-\theta t)}{\exp(-\theta t)-1}$	$-t^{-(\theta+1)}$	
$\partial\phi^{-1}(t)/\partial t$	$\frac{-1}{\theta}t^{\frac{1}{\theta}-1}\phi^{-1}(t)$	$\frac{1}{\theta}\frac{\exp(-t)\{\exp(-\theta)-1\}}{1+\exp(-t)\{\exp(-\theta)-1\}}$	$\theta(1+\theta t)^{-\frac{1}{\theta}-1}$	

Table 1: Archemdean Copulae's Generator, Generator Inverse, and their derivative.

$$\left.\frac{\partial C\{v,g(w)\}}{\partial v}\right|_{v=w} = \frac{\partial\phi^{-1}[\phi(v)+\phi\{g(w)\}]}{\partial[\phi(v)+\phi\{g(w)\}]}\left.\frac{\partial[\phi(v)+\phi\{g(w)\}]}{\partial v}\right|_{v=w} \tag{5}$$

$$= \frac{\partial\phi^{-1}[\phi(v)+\phi\{g(w)\}]}{\partial[\phi(v)+\phi\{g(w)\}]}\left.\frac{\partial\phi(v)}{\partial v}\right|_{v=w} \tag{6}$$

$$= \frac{\partial\phi^{-1}[\phi(w)+\phi\{g(w)\}]}{\partial[\phi(w)+\phi\{g(w)\}]}\frac{\partial\phi(w)}{\partial w} \tag{7}$$

$$, \text{where } g(w) = F_{R^F}\left\{\frac{F_{R^S}^{-1}(w)-r^h}{h}\right\} \tag{8}$$

$$\tag{9}$$

6. Special copulas

Add basics about copulas, Sklar's Theorem

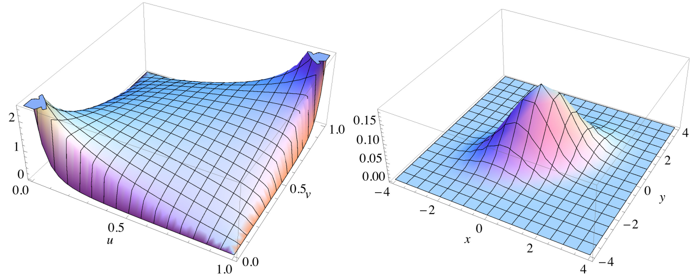


Figure 2: Left: Density of the Gaussian (Normal) copula. Right: Density of the bivariate Normal distribution ($\rho = 0.5$ in both cases).

Currently: Archimedean (Clayton, Gumbel), Gaussian, NVM / elliptical (Student t, NIG), NIG factor
Possibly: double t, Frank, skewed-t
For CF-expansion of Student t, see Abramowitz Stegun 26.7.5, see also `https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.152.7540&rep=rep1&type=pdf`

6.1. Gaussian Copula

The **Gaussian copula**, also called **Normal copula**, is the copula generated by jointly normally distributed random variables (given here in bivariate form):

$$C_{\rho,\mathrm{N}}(u,v) := \mathrm{N}_2(\mathrm{N}^{(-1)}(u), \mathrm{N}^{(-1)}(v); \rho),$$

where N_2 and N are the bivariate and univariate normal distribution functions, respectively, and ρ is the correlation parameter.

The Gaussian or Normal copula is

$$C_{\Sigma}^{Ga}(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \int_{-\infty}^{\Phi^{-1}(x_1)} \cdots \int_{-\infty}^{\Phi^{-1}(x_d)} \exp\left\{-\frac{1}{2} y^{\top}\Sigma^{-1} y\right\} dy_1 \dots dy_d, \quad x \in \mathbb{R}^d. \tag{10}$$

The copula density is

$$c_{\Sigma}^{Ga}(x) = \frac{1}{(2\pi)^{d/2}|\Sigma|^{1/2}} \exp\left\{-\frac{1}{2}\begin{pmatrix}\Phi^{-1}(x_1)\\ \vdots \\ \Phi^{-1}(x_d)\end{pmatrix}^{\top} \Sigma^{-1} \begin{pmatrix}\Phi^{-1}(x_1)\\ \vdots \\ \Phi^{-1}(x_d)\end{pmatrix}\right\} \tag{11}$$

Simplified notation bivariate Gaussian copula

$$C_{\rho}^{Ga}\{w, g(w)\} = \Phi_{\rho}[\Phi^{-1}(w), \Phi^{-1}\{g(w)\}], \tag{12}$$

where $g(w) : [0,1] \mapsto \mathbb{R}$ is defined above, ρ is the dependency parameter of a bivariate Gaussian copula, Φ_{ρ} is bivariate normal distribution with mean 0 and covariance $\begin{bmatrix}1 & \rho\\ \rho & 1\end{bmatrix}$, $\Phi(\cdot)$ is CDF of standard normal, $\phi(\cdot)$ is PDF of standard normal, $\Phi^{-1}(\cdot)$ is quantile function of standard normal.

[I suggest to move the part below to the section where the formula for R^h is discussed. Also, there is a simpler version that requires only univariate function evluations.]

The bivariate $D_1 C^{Ga}\{w, g(w)\}$ is

$$D_1 C_{\rho}^{Ga}\{w, g(w)\} = \int_{-\infty}^{\Phi^{-1}\{g(w)\}} \phi_{\rho}\{\Phi^{-1}(w), u\} du \cdot \frac{1}{\phi\{\Phi^{-1}(w)\}} \tag{13}$$

Proof.

$$D_1 C_\rho\{w, g(w)\} = \left.\frac{\partial C_\rho\{w, g(w')\}}{\partial w}\right|_{w'=w} \tag{14}$$

$$= \left.\frac{\partial \Phi_\rho[\Phi^{-1}(w), \Phi^{-1}\{g(w)\}]}{\partial \Phi^{-1}(w)} \frac{\partial \Phi^{-1}(w)}{\partial w}\right|_{w'=w} \tag{15}$$

$$= \frac{1}{2\pi\rho} \int_{-\infty}^{\Phi^{-1}\{g(w)\}} \exp\left\{-\frac{1}{2(1-\rho^2)} \Phi^{-1}(w)^2 - 2\rho\Phi^{-1}(w)u + u^2\right\} du \cdot \frac{1}{\phi\{\Phi^{-1}(w)\}} \tag{16}$$

□

Hedge distribution

Let $R^S \sim \mathrm{N}(\mu_S, \sigma_S^2)$ and $R^F \sim \mathrm{N}(\mu_F, \sigma_F^2)$ and assume further that they are jointly normally distributed with correlation ρ. Then,

$$R^h = R^S - hR^F \sim \mathrm{N}(\mu_S - h\mu_F, \sigma_s^2 + h^2\sigma_F^2 - 2\rho h \sigma_S \sigma_F).$$

More generally, if $R^k \sim F^k$, $k \in \{S, F\}$, then the distribution of R^h can be expressed with univariate expressions:

$$\begin{aligned}
\mathbf{P}(R^S - hR^F \leq x) &= 1 - \mathbb{E}\left[\mathbf{P}\left(R^F \leq \frac{R^S - x}{h} \middle| R^S\right)\right] \\
&= 1 - \mathbb{E}\left[\mathbf{P}\left(\mathrm{N}^{(-1)}(F_F(R^F)) \leq \mathrm{N}^{(-1)}\left(F_F\left(\frac{R^S - x}{h}\right)\right) \middle| R^S\right)\right] \\
&= 1 - \mathbb{E}\left[\mathbf{P}\left(\rho \mathrm{N}^{(-1)}(F_S(R^S)) + \sqrt{1-\rho^2}\varepsilon \leq \mathrm{N}^{(-1)}\left(F_F\left(\frac{R^S - x}{h}\right)\right) \middle| R^S\right)\right] \\
&= 1 - \mathbb{E}\left[\mathrm{N}\left(\frac{\mathrm{N}^{(-1)}\left(F_F\left(\frac{R^S-x}{h}\right)\right) - \rho \mathrm{N}^{(-1)}(F_S(R^S))}{\sqrt{1-\rho^2}}\right)\right] \\
&= 1 - \int_0^1 \mathrm{N}\left(\frac{\mathrm{N}^{(-1)}\left(F_F\left(\frac{F_S^{(-1)}(u)-x}{h}\right)\right) - \rho \mathrm{N}^{(-1)}(u)}{\sqrt{1-\rho^2}}\right) \mathrm{d}u.
\end{aligned}$$

6.2. Archimedean copulas

- A well-studied one-parameter family of copulas are the **Archimedean copulas**.
- Let $\phi : [0,1] \to [0,\infty]$ be a continuous and strictly decreasing function with $\phi(1) = 0$ and $\phi(0) \leq \infty$.
- We define the **pseudo-inverse** of ϕ as

$$\phi^{(-1)}(t) = \begin{cases} \phi^{-1}(t), & 0 \leq t \leq \phi(0), \\ 0, & \phi(0) < t \leq \infty. \end{cases}$$

- If, in addition, ϕ is convex, then the following function is a copula:

$$C(u,v) = \phi^{(-1)}(\phi(u) + \phi(v)).$$

- Such copulas are called **Archimedean copulas**, and the function ϕ is called an **Archimedean copula generator**.
- Examples of Archimedean copulas are the **Gumbel** and the **Clayton** copulas:

$$C_{\theta,\mathrm{Gu}}(u,v) = \exp\left\{-((-\ln u)^{\theta} + (-\ln v)^{\theta})^{1/\theta}\right\}, \qquad 1 \leq \theta < \infty,$$
$$C_{\theta,\mathrm{Cl}}(u,v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \qquad 0 < \theta < \infty.$$

- In the case of the Gumbel copula, the independence copula is attained when $\theta = 1$ and the comonotonicity copula is attained as $\theta \to \infty$.
- Thus, the Gumbel copula interpolates between independence and perfect dependence.
- In the case of the Clayton copula, the independence copula is attained as $\theta \to 0$, whereas the comonotonicity copula is attained as $\theta \to \infty$.

6.3. Elliptical Copulas and normal variance mixtures

See e.g. Theorem 3.22, Definition 3.26 and Theorem 3.28 of (McNeil *et al.*, 2005):

Definition 1. A random vector $\boldsymbol{Z} = (Z_0, \ldots, Z_d)^T$ follows an elliptical distribution if it has a representation

$$\boldsymbol{Z} \stackrel{\mathcal{L}}{=} GA\boldsymbol{U},$$

where $G > 0$ is a scalar random variable, the so-called *mixing variable*, A is a deterministic $(d+1) \times (d+1)$ matrix with $AA^T := \Sigma$, which in turn is a $(d+1) \times (d+1)$ nonnegative definite symmetric matrix of rank $d+1$, and $\boldsymbol{U}$ is a $(d+1)$-dimensional random vector uniformly distributed on the unit sphere $\mathcal{S}_{d+1} := \{\boldsymbol{z} \in \mathbb{R}^{d+1} : \boldsymbol{z}^T\boldsymbol{z} = 1\}$, and $\boldsymbol{U}$ is independent of G.

A subclass of elliptical distributions are the so-called *normal variance mixtures (NVM)*, see Section 3.3 of (McNeil *et al.*, 2005). For the connection between NVM and elliptical distributions, see also Theorem 3.25 of (McNeil *et al.*, 2005).

Definition 2 (Normal variance mixture (NVM))**.** The random vector $\mathbf{X} = (X_1, \ldots, X_k)^T$ follows a multivariate *normal variance mixture (NVM) distribution* if

$$\mathbf{X} \overset{\mathcal{L}}{=} \mu + \sqrt{W} A \mathbf{Z},$$

where

(i) $\mathbf{Z} \sim \mathrm{N}_k(\mathbf{0}, I_k)$, i.e., $\mathbf{Z}$ are independent, standard normally distributed,

(ii) $W \geq 0$ is a random variable independent of $\mathbf{Z}$,

(iii) $A \in \mathbb{R}^{d \times k}$ and $\mu \in R^d$ are a matrix and vector of constants, respectively.

It is easily observed that $\mathbf{X}|W = w \sim \mathrm{N}_d(\mu, w\Sigma)$, where $\sigma = AA'$.
In general, we will assume that Σ is positive definite and that $W > 0$ $\mathbf{P}$–a.s.. Then, the density of $\mathbf{X}$ is given by

$$\begin{aligned} f(\mathbf{x}) &= \int f_{\mathbf{X}|W}(\mathbf{x}|w) \,\mathrm{d}H(w) \\ &= \int \frac{w^{-d/2}}{(2\pi)^{d/2} |\Sigma|^{1/2}} \exp\left(- \frac{(\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)}{2w} \right) \mathrm{d}H(w), \end{aligned}$$

where H is the distribution function of W.
Special cases:

- Normal distribution: W constant
- Student t distribution: $W \sim Ig(1/2\nu, 1/2\nu)$, where Ig is an inverse gamma distribution

- Symmetric generalised hyperbolic distribution: $W \sim N^{-}(\lambda, \chi, \psi)$ where N^{-} refers to the generalised inverse Gaussian (GIG) distribution;
- Normal inverse Gaussian (NIG): W follows a GIG distribution with $\lambda = -0.5$.

Copulas are obtained from elliptical distributions via Sklar's theorem by transforming the margins to uniforms.

Hedge distribution

Let (R^S, R^F) follow a normal variance mixture, i.e., there exists a decomposition such that

$$\begin{aligned} R^S &= \mu_S + \sqrt{W}\sigma_S Z_1 \\ R^F &= \mu_F + \sqrt{W}\sigma_F(\rho Z_1 + \sqrt{1-\rho^2} Z_2), \end{aligned}$$

where W is the mixing variable and Z_1, Z_2 are independent standard normal variables. Then, R^h follows a NVM distribution with

$$R^h = R^S - hR^F = \mu_S - h\mu_F + \sqrt{W}\left((\sigma_S - h\sigma_F\rho)Z_1 - h\sqrt{1-\rho^2}\sigma_F Z_2\right) = \mu_S - h\mu_F + \sqrt{W}Z_3,$$

where $Z_3 \sim \mathrm{N}(0, \sigma_S^2 + h^2\sigma_F^2 - 2\rho h\sigma_S\sigma_F)$.
More generally, let $R^k \sim F^k$, $k \in \{S, F\}$ and write V as the marginal distribution functions of the NVM distribution components. Let $V^{(-1)}(F_S(R^S)) \overset{\mathcal{L}}{=} \sqrt{W}Z_1$ and $V^{(-1)}(F_F(R^F)) \overset{\mathcal{L}}{=} \sqrt{W}\rho Z_1 + \sqrt{W}\sqrt{1-\rho^2}Z_2$, were Z_1, Z_2 are

independent standard normals. Then

$$
\begin{aligned}
\mathbf{P}(R^S - hR^F \leq x) &= 1 - \mathbb{E}\left[\mathbf{P}\left(R^F \leq \frac{R^S - x}{h} \middle| R^S\right)\right] \\
&= 1 - \mathbb{E}\left[\mathbf{P}\left(V^{(-1)}(F_F(R^F)) \leq V^{(-1)}\left(F_F\left(\frac{R^S - x}{h}\right)\right) \middle| W, Z_1\right)\right] \\
&= 1 - \mathbb{E}\left[\mathbf{P}\left(\sqrt{W}\rho Z_1 + \sqrt{W}\sqrt{1-\rho^2}Z_2 \leq V^{(-1)}\left(F_F\left(\frac{F_S^{(-1)}(V(\sqrt{W}Z_1)) - x}{h}\right)\right) \middle| W, Z_1\right)\right] \\
&= 1 - \mathbb{E}\left[\mathrm{N}\left(\frac{V^{(-1)}\left(F_F\left(\frac{F_S^{(-1)}(V(\sqrt{W}Z_1)) - x}{h}\right)\right) - \rho\sqrt{W}Z_1}{\sqrt{W}\sqrt{1-\rho^2}}\right)\right] \\
&= 1 - \int_0^\infty \int_{-\infty}^\infty \mathrm{N}\left(\frac{V^{(-1)}\left(F_F\left(\frac{F_S^{(-1)}(V(\sqrt{w}z_1)) - x}{h}\right)\right) - \rho\sqrt{w}z_1}{\sqrt{w}\sqrt{1-\rho^2}}\right) \varphi(z_1)\, f_W(w)\, \mathrm{d}z_1\, \mathrm{d}w.
\end{aligned}
$$

As in the Gaussian copula case, this expression contains evaluations of only univariate distribution function.

6.4. Normal mean-variance mixtures

Elliptical distributions are constrained to be symmetric. Extending NVM's to normal mean-variance mixtures (NMVM) introduces skewness, which may be more realistic for financial returns. We refer to Section 3.2.2 of (McNeil *et al.*, 2005) for more details.

Definition 3. The random vector $\mathbf{X}$ is said to have a (multivariate) normal mean-variance mixture distribution if

$$
\mathbf{X} \overset{\mathcal{L}}{=} \mathbf{m}(W) + \sqrt{W} A \mathbf{Z},
$$

where

(i) $\mathbf{Z} \sim \mathrm{N}_k(\mathbf{0}, I_k)$;

(ii) $W \geq 0$ is a non-negative, scalar-valued random variable independent of $\mathbf{Z}$;

(iii) $A \in \mathbb{R}^{d \times k}$ is a matrix;

(iv) $\mathbf{m} : [0, \infty) \to \mathbb{R}^d$ is a measurable function.

We have

$$\mathbf{X}|W = w \sim \mathrm{N}_d(\mathbf{m}(w), w\Sigma),$$

where $\Sigma = AA'$. A possible concrete specification of $\mathbf{m}(W)$ is

$$\mathbf{m}(W) = \mu + W\gamma, \tag{17}$$

where μ and γ are vectors in $\mathbb{R}^d$. If $\gamma = 0$, then the distribution is a NVM.
A special case are the *generalized hyperbolic (GH) distributions*, which a NMVM's with mean specification (17) and mixing distribution $W \sim N^-(\lambda, \chi, \psi)$, a generalised inverse Gaussian (GIG) distribution. We write $\mathbf{X} \sim \mathrm{GH}_d(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$. The specification is not unique in the sense that scaled versions of the parameters described the same distribution.
A different parameterization is obtained using the parameters $\alpha, \delta, \Delta, \beta$:

$$\Delta = |\Sigma|^{-1/3}\Sigma, \quad \beta = \Sigma^{-1}\gamma, \quad \delta = \sqrt{\chi|\Sigma|^{1/d}}, \quad \alpha = \sqrt{|\Sigma|^{-1/d}(\psi + \gamma'\Sigma^{-1}\gamma)}.$$

It is not uncommon to restrict $|\Delta| = 1$. The parameters have the following interpretation: μ is the location parameter, δ the scale parameter, β the skewness parameter and α and λ are shape parameters.
If $\alpha, \delta, \Delta, \beta$ is used, then

$$\Sigma = \Delta, \quad \gamma = \Delta\beta, \quad \chi = \delta^2, \quad \psi = (\alpha^2 - \beta'\Delta\beta).$$

Amongst several special cases, we consider the NIG and the skewed t distributions:

- $\lambda = -1/2$ gives rise to the NIG distribution
- $\lambda = -1/2\nu$, $\chi = \nu$, $\psi = 0$ gives rise to an asymmetric or skewed t distribution. The multivariate t distribution is obtained as $\gamma \to 0$.

A different parameterization is obtained by setting

$$\beta = \Sigma^{-1}\gamma, \quad \delta = \sqrt{\chi}, \quad \alpha = \sqrt{\psi + \gamma'\Sigma^{-1}\gamma},$$

satisfying the constraints $\delta \geq 0$, $\alpha^2 > \beta'\Sigma\beta$ if $\lambda > 0$; $\delta > 0$, $\alpha^2 > \beta'\Sigma\beta$ if $\lambda = 0$; and $\delta > 0$, $\alpha^2 \geq \beta'\Sigma\beta$ if $\lambda < 0$.

6.5. Normal inverse Gaussian factor copula model

Literature on NIG, factor copulas and fitting of copulas:

- (Kalemanova *et al.*, 2007): NIG factor model, calibrated to CDO's
- (Genest, 1987; Genest and Rivest, 1993): "Method of moments" for nonparametric estimation of copula parameters; mainly refers to Kendall's tau or Spearman's rho
- (Patton and Oh, 2012; Oh and Patton, 2013): Factor copula models (in particular skewed t), estimated with "method of moments" involving Spearman's rank correlation, and $0.05, 0.1, 0.9, 0.95$ quantile dependence measures (there is a weight matrix, but they use the identity matrix, see ASA paper, page 694, right column, top paragraph)

6.5.1. Normal inverse Gaussian distribution

It was established in the previous section that a multivariate normal inverse Gaussian (NIG) distribution is a normal variance mixture. Here the random components have a joint scalar mixing variable. An additional copula model can be derived as a factor model from the NIG distribution: under certain conditions, its distribution type is preserved

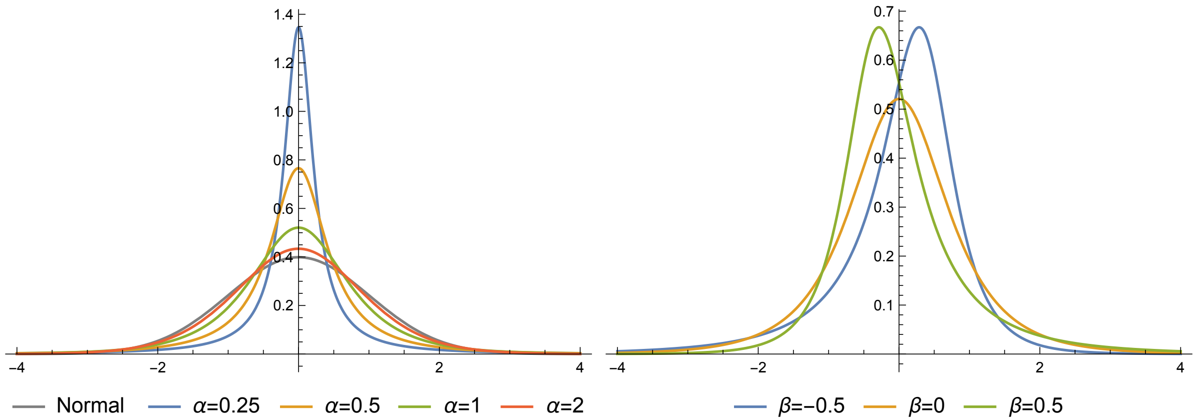


Figure 3: Standardised NIG's density functions, i.e., μ and δ yield zero expectation and unit variance. Left: $\beta = 0$; right: $\alpha = 1$. With increasing α, the standardised NIG converges to a normal distribution.

under linear combinations. This property, together with its infinite divisibility, allow for the construction of Lévy processes from NIG distributions.
Following (Barndorff-Nielsen, 1997), a normal inverse Gaussian (NIG) distribution has density function

$$g(x; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \mathbf{e}^{\delta\sqrt{\alpha^2-\beta^2}-\beta\mu} \frac{1}{q((x-\mu)/\delta)} K_1 \left[\delta\alpha q \left(\frac{x-\mu}{\delta} \right) \right] \mathbf{e}^{\beta x}, \quad x > 0,$$

where $q(x) = \sqrt{1+x^2}$ and where K_1 is the modified Bessel function of third order and index 1. The parameters satisfy $0 \le |\beta| \le \alpha$, $\mu \in \mathbb{R}$ and $\delta > 0$. The parameters have the following interpretation: μ and δ are location and scale parameters, respectively, α determines the heaviness of the tails and β determines the degree of asymmetry. If $\beta = 0$, then the distribution is symmetric around μ. Figure 3 shows examples of NIG densities.

The moment-generating function of the NIG distribution is given by

$$M(u;\alpha,\beta,\mu,\delta)=\exp\left(\delta\left(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2}\right)+\mu u\right).$$

As a direct consequence, moments are easily calculated with the expectation and variance of the NIG distribution being

$$\mathbb{E}X=\mu+\frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}} \tag{18}$$

$$\mathrm{Var}(X)=\frac{\alpha^2\delta}{(\alpha^2-\beta^2)^{3/2}}. \tag{19}$$

Let $\mathrm{IG}(\delta,\gamma)$ denote the *inverse Gaussian distribution* with density function[3]

$$d(w;\delta,\gamma)=\frac{1}{\sqrt{2\pi}}\exp(\delta\gamma)w^{-3/2}\exp(-\frac{\delta^2/w+\gamma^2 z}{2}). \tag{20}$$

The $\mathrm{NIG}(\alpha,\beta,\mu\,\delta)$ distribution is a normal variance-mean mixture: X follows an $\mathrm{NIG}(\alpha,\beta,\mu,\delta)$ distribution if X conditional on W follows a normal distribution with mean $\mu+\beta W$ and variance W, i.e.,

$$X|W\overset{\mathcal{L}}{\sim}\mathrm{N}(\mu+\beta W,W),$$

where W follows an $\mathrm{IG}(\delta,\sqrt{\alpha^2-\beta^2})$ distribution.

[3] The density of the IG distribution in `Mathematica` is given as

$$f(x)=\sqrt{\frac{\lambda}{x^3}}\frac{1}{\sqrt{2\pi}}\mathbf{e}^{-\frac{\lambda(x-\mu)^2}{2x\mu^2}},\quad x>0,$$

with parameters $\mu=\delta/\gamma$ and $\lambda=\delta^2$.

It is easily seen from the moment-generating function that linear combinations of NIG random variables are again NIG-distributed provided they share the parameters α and β. Let $X_i \sim \mathrm{NIG}(\alpha, \beta, \mu_i, \delta_i)$, $i = 1, 2$, be independent NIG variables. Then,

$$\mathbb{E}\left[\mathbf{e}^{u(X_1+X_2)}\right] = \mathbb{E}\left[\mathbf{e}^{uX_1}\right]\mathbb{E}\left[\mathbf{e}^{uX_2}\right] = \exp\left((\delta_1+\delta_2)\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+u)^2}\right) + (\mu_1+\mu_2)u\right),$$

hence $X_1 + X_2 \sim \mathrm{NIG}(\alpha, \beta, \mu_1 + \mu_2, \delta_1 + \delta_2)$. (This is also a direct consequence from the properties of the normal inverse Gaussian Lévy process X_t, which may be represented as Brownian motion with a random time change,

$$X_t = B_{W_t} + \mu t,$$

where $B = (B_t)_{t\geq 0}$ is a Brownian motion and $W = (W_t)_{t\geq 0}$ is a Lévy process with density given by (20). The random variable W_t can be interpreted as a first-passage time of an independent Brownian motion $\overline{B}$, i.e., $W_t = \inf\{s > 0 : \overline{B}_s + \sqrt{\alpha^2 - \beta^2} s = \delta t\}$.)
As a consequence, the NIG distribution gives rise to two copulas:

- a copula determined from a linear combination of independent NIG random variables with identical parameters α, β (essentially a factor model);
- a copula determined from the multivariate normal-mean-variance mixture, which is a linear combination of normal random variables scaled by one scalar inverse Gaussian random variable.

6.5.2. NIG factor copula

We consider a simple factor model consisting of NIG-distributed random variables.

Proposition 2. *Let $Z \sim NIG(\alpha, \beta, \mu, \delta)$ and $Z_i \sim NIG(\alpha, \beta, \mu_i, \delta_i)$, $i = 1, \ldots, n$ be independent NIG-distributed random variables. Then (i) $X_i = Z + Z_i \sim NIG(\alpha, \beta, \mu + \mu_i, \delta + \delta_i)$ and (ii)*

$$\begin{aligned} Cov(X_i, X_j) &= Var(Z), \\ Corr(X_i, X_j) &= \frac{\delta}{\sqrt{(\delta+\delta_i)(\delta+\delta_j)}}. \end{aligned} \tag{21}$$

Proof. (i) This follows directly from the moment-generating function.

(ii) For the covariance,

$$\begin{aligned}\mathrm{Cov}(X_i, X_j) &= \mathbb{E}[(Z+Z_i)(Z+Z_j)] - \mathbb{E}[Z+Z_i]\mathbb{E}[Z+Z_j] \\ &= \mathbb{E}[Z^2] - (\mathbb{E}Z)^2.\end{aligned}$$

The correlation is determined directly from (19).

□

The NIG factor copula is obtained by transforming the margins to uniforms (see Sklar's Theorem). The NIG factor copula function can be written as (Krupskii and Joe, 2013):

$$C_{R^S,R^F}\{F_{R^S}(r^S), F_{R^F}(r^F)\} = \int_{\mathbb{R}} F_{Z_1}\{F_{X_1}^{-1} \circ F_{R^S}(r^S) - z\} \cdot F_{Z_2}\{F_{X_2}^{-1} \circ F_{R^F}(r^F) - z\} \cdot f_Z(z)dz \tag{22}$$

The Spearman rho of NIG factor copula is

$$\rho_S = 12 \int\int\int_{\mathbb{R}^3} F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot f_{Z_1}(x_1 - z) \cdot f_{Z_2}(x_2 - z) \cdot f_Z(z)dx_1dx_2dz - \frac{1}{48} \tag{23}$$

Proof.

$$\rho_S(R^S, R^F) = \rho\{F_{R^S}(R^S), F_{R^F}(R^F)\} \tag{24}$$

$$= \rho\{F_{X_1}(X_1), F_{X_2}(X_2)\} \tag{25}$$

$$= 12 \cdot \mathbb{E}\{F_{X_1}(X_1) \cdot F_{X_2}(X_2)\} - \frac{1}{48} \tag{26}$$

$$= 12 \cdot \int\int_{\mathbb{R}^2} F_{X_1}(X_1) \cdot F_{X_2}(X_2)dF_{X_1,X_2}(x_1, x_2) \tag{27}$$

$$\tag{28}$$

Because

$$F_{X_1,X_2}(x_1,x_2) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2) \tag{29}$$
$$= \mathbb{P}(Z_1 \leq x_1 - Z, Z_2 \leq x_2 - Z) \tag{30}$$
$$= \int_{\mathbb{R}} \mathbb{P}(Z_1 \leq x_1 - z) \cdot \mathbb{P}(Z_2 \leq x_2 - z) \cdot f_Z(z) dz, \tag{31}$$

so,

$$\rho_S(R^S, R^F) = 12 \cdot \int\int\int_{\mathbb{R}^3} F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot f_{Z_1}(x_1 - z) \cdot f_{Z_2}(x_2 - z) \cdot f_Z(z) dx_1 dx_2 dz - \frac{1}{48} \tag{32}$$

□

6.5.3. Fitting the NIG factor model

In this section we assume that the marginal distributions are NIG.
Given that the parameters $\alpha, \beta, \mu, \delta$ denote tail heaviness, skewness, location and scale, one can fit the NIG distribution by matching the first four moments. Let $\hat{\mu}_r = \frac{1}{n}\sum_{i=1}^n x_i^r$ denote the r-th moment of the sample $x_1, \ldots, x_n$. Then moment-matching the NIG distribution corresponds to solving for $\alpha, \beta, \mu, \delta$ the system of equations

$$\begin{pmatrix} \frac{\partial}{\partial u} M(u; \alpha, \beta, \mu, \delta)|_{u=0} \\ \frac{\partial^2}{\partial u^2} M(u; \alpha, \beta, \mu, \delta)|_{u=0} \\ \frac{\partial^3}{\partial u^3} M(u; \alpha, \beta, \mu, \delta)|_{u=0} \\ \frac{\partial^4}{\partial u^4} M(u; \alpha, \beta, \mu, \delta)|_{u=0} \end{pmatrix} = \begin{pmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \hat{\mu}_3 \\ \hat{\mu}_4 \end{pmatrix}$$

Bivariate data, as present in our case, can be fit by moment-matching involving the parameters $(\alpha, \beta, \mu_1, \mu_2, \delta_1, \delta_2)$ and then fitting δ of the joint factor via the empirical correlation. Without loss of generality, we set $\mu = 0$. Then,

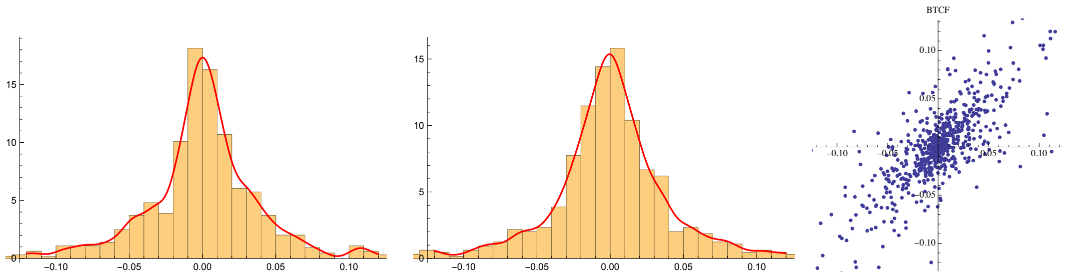


Figure 4: BRR and BTCF return distributions (empirical and fitted to NIG) as well as scatter plot.

$\alpha, \beta, \mu_i, \tilde{\delta}_i$, $i = 1, 2$, are determined from

$$\min_{\alpha,\beta,\mu_1,\mu_2,\tilde{\delta}_1,\tilde{\delta}_2} \sqrt{\sum_{k=1}^{2}\sum_{r=1}^{4}\left(\hat{\mu}_{r,k} - \frac{\partial^r}{\partial u^r} M(u;\alpha,\beta,\mu_k,\tilde{\delta}_k)\Big|_{u=0}\right)^2}.$$

Here, $\tilde{\delta}_k$ refers to the scale parameter of X_k. The scale parameters of the independent NIG components Z_k, $k = 1, 2$, are obtained as $\delta_k = \delta - \tilde{\delta}_k$ after solving δ from (21), which fixes the dependence between the margins. Figure 4 shows histograms and the fitted NIG densities of the Bitcoin reference rate (BRR) returns and the Bitcoin Futures contract returns (BTCF).

6.5.4. Fitting the NIG factor copula

If the margins are not NIG-distributed, then the NIG factor copula model is calibrated by the so-called "method of moments" as described by (Oh and Patton, 2013). In this setting, Spearmans' rho (rank correlation) and quantile dependence measures at the $0.05, 0.1, 0.9, 0.95$ quantiles are calibrated against the empirical counterparts. Spearman's rho and quantile dependence of a pair (X, Y) with copula C are defined as

$$\rho_S = 12 \int\int_{I^2} C(u, v)\, \mathrm{d}u\, \mathrm{d}v - 3 \tag{33}$$

$$\lambda_q = \begin{cases} \mathbf{P}(F_X(X) \leq q | F_Y(Y) \leq q) = \dfrac{C(q,q)}{q}, & \text{if } q \in (0, 0.5], \\ \mathbf{P}(F_X(X) > q | F_Y(Y) > q) = \dfrac{1 - 2q + C(q,q)}{1-q}, & \text{if } q \in (0.5, 1). \end{cases} \tag{34}$$

The empirical counterparts are

$$\hat{\rho}_S = \frac{12}{n} \sum_{k=1}^{n} \hat{F}_X(x_k) \hat{F}_Y(y_k) - 3,$$

$$\hat{\lambda}_q = \begin{cases} \dfrac{1}{n} \displaystyle\sum_{k=1}^{n} \dfrac{\mathbf{1}_{\{\hat{F}_X(x_k) \leq q, \hat{F}_Y(y_k) \leq q\}}}{q}, & \text{if } q \in (0, 0.5], \\ \dfrac{1}{n} \displaystyle\sum_{k=1}^{n} \dfrac{\mathbf{1}_{\{\hat{F}_X(x_k) > q, \hat{F}_Y(y_k) > q\}}}{1-q}, & \text{if } q \in (0.5, 1). \end{cases}$$

Here, the empirical distribution function is defined such that the data lie strictly in the interior of the unit cube (see (Oh and Patton, 2013) and (McNeil *et al.*, 2005, p. 232)):

$$\hat{F}(x) = \frac{1}{n+1} \sum_{k=1}^{n} \mathbf{1}_{\{x_k \leq x\}}.$$

Convergence of the sample measures to the population counterparts is established by (Fermanian *et al.*, 2004). Denote the *standardised* NIG distribution function (i.e., expectation 0 and variance 1) by $F(\cdot; \alpha, \beta)$. In this case $\delta(\alpha, \beta) = \dfrac{(\alpha^2 - \beta^2)^{3/2}}{\alpha^2}$ and $\mu(\alpha, \beta) = -\dfrac{\delta\beta}{\sqrt{\alpha^2 - \beta^2}} = \dfrac{\beta^3}{\alpha^2} - \beta$. In the factor model setting, the parameters α, β and δ are calibrated, where δ refers to the scaling parameter of the latent variable Z.

Denote the margins by $U_i \sim U(0,1)$, $i = 1,2$. The factor model of Proposition 2 is obtained by transforming the uniform margins to standardised NIG distributions by setting $X_i := F^{(-1)}(U_i; \alpha, \beta, \mu, \delta + \delta_i)$ with $\delta + \delta_i = \frac{(\alpha^2 - \beta^2)^{3/2}}{\alpha^2}$ and $\mu = -\frac{(\delta+\delta_i)\beta}{\sqrt{\alpha^2 - \beta^2}} = -\frac{(\alpha^2 - \beta^2)\beta}{\alpha^2} = \frac{\beta^3}{\alpha^2} - \beta$. Here, δ refers to the scaling factor of the latent variable Z. In this case, $\mathbb{E}X_i = 0$ and $\mathrm{Var}(X_i) = 1$.

Calibrating the NIG factor copula corresponds to minimising the root mean square error of the sample and distribution copula measures:

$$\min_{\alpha,\beta,\delta} \sqrt{(\rho_S - \hat{\rho}_S)^2 + (\lambda_{0.05} - \hat{\lambda}_{0.05})^2 + (\lambda_{0.1} - \hat{\lambda}_{0.1})^2 + (\lambda_{0.9} - \hat{\lambda}_{0.9})^2 + (\lambda_{0.95} - \hat{\lambda}_{0.95})^2}$$

To avoid numerical calculation of ρ_S in Equation (33), which requires integrating the copula function, we approximate ρ_S by the rank correlation for the Gaussian copula (see (McNeil *et al.*, 2005, p. 215)),

$$\rho_S(X_1, X_2) = \frac{6}{\pi} \arcsin \frac{1}{2}\rho, \tag{35}$$

where ρ refers to the correlation coefficient, (21). This is justified if β is close to zero and for sufficiently large values of α, as the NIG converges to the normal distribution as $\alpha \to \infty$.

In addition, as ρ_S in (35) depends only on δ (through ρ), we transform the optimisation of δ into a constraint, giving

$$\min_{\alpha,\beta} \sqrt{(\lambda_{0.05} - \hat{\lambda}_{0.05})^2 + (\lambda_{0.1} - \hat{\lambda}_{0.1})^2 + (\lambda_{0.9} - \hat{\lambda}_{0.9})^2 + (\lambda_{0.95} - \hat{\lambda}_{0.95})^2},$$

subject to $\hat{\rho}_S = \frac{6}{\pi} \arcsin \frac{1}{2} \frac{\delta}{\delta + \delta_i}$.

[Possibly compute AIC in order to compare the calibration performance across models?]

6.5.4.1. Bitcoin data The copula parameters obtained in the calibration of the Bitcoin data are $\alpha = 0.773$, $\beta = 0.02933$, $\delta = 0.5782$. The rank correlation $\hat{\rho}_S = 0.7337$, whereas the calibrated rank correlation (subject to the

Measure	Empirical	Calibrated
ρ_s	0.7337	0.7470
$\lambda_{0.05}$	0.5581	0.5872
$\lambda_{0.1}$	0.6357	0.6100
$\lambda_{0.9}$	0.6202	0.6156
$\lambda_{0.95}$	0.5891	0.5954

Table 2: Empirical versus calibrated dependence measures. The calibrated parameters are $\alpha = 0.773$, $\beta = 0.02933$, $\delta = 0.5782$. The RMSE is 0.0395.

error from the the Gaussian approximation of (35)) gives 0.7470. Figure 5 shows scatter plots of the NIG copula applied to the BTC data and to simulated data.

6.5.5. Hedge calculation in the NIG factor model

Optimising for the hedge quantity h requires fast calculations of the hedge distribution function (1). The position obtained from hedging Bitcoin with h units of the Bitcoin future returns

$$R^h = X_1 - hX_2 = Z + Z_1 - hZ - hZ_2 = (1-h)Z + Z_1 - hZ_2,$$

with Z, Z_1, Z_2 independent NIG-distributed random variables. Without loss of generally, set $\mu = 0$. Because of the scaling with $1-h$ and h, R^h will not follow an NIG distribution, unless $h = 0$. A direct calculation of the distribution function is achieved by numerically computing the integral

$$\mathbf{P}(R^h \leq x) = \int_0^\infty \int_0^\infty \int_0^\infty \mathrm{N}(y; \mu_1 - h\mu_2 + \beta w_1 - hw_2, (1-h)^2 w + w_1 + h^2 w_2)\, f_W(w)\, f_{W_1}(w_1)\, f_{W_2}(w_2)\, \mathrm{d}w\, \mathrm{d}w_1\, \mathrm{d}w_2,$$

where $\mathrm{N}(x; \mu, \sigma^2)$ denotes the normal distribution function with expectation μ and variance σ^2, and f_W, f_{W_1}, f_{W_2} are the inverse Gaussian density functions of the scaling variables W, W_1 and W_2 with parameters $\delta, \delta_1, \delta_2$ and $\sqrt{\alpha^2 - \beta^2}$.

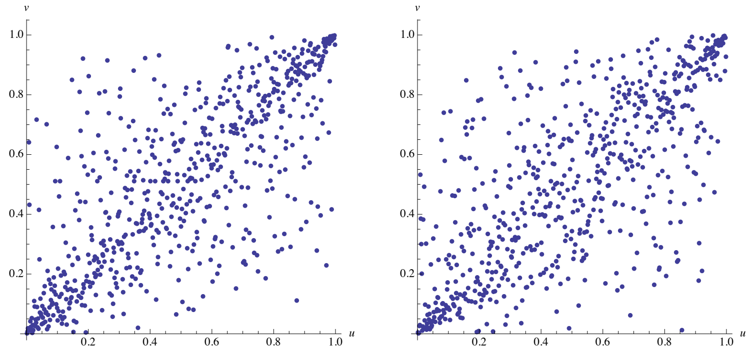


Figure 5: Left: Scatter plot of empirical copula; right: scatter plot of simulated sample with calibrated parameters. The sample size is 645 in both cases.

However, for determining the optimal hedge parameter h, numerical computation of the three-fold integral may be slow, compared to other methods that make explicit use of the simple form of the moment-generating function. The following two methods achieve a much faster computation for the NIG factor model. The first method uses Fourier inversion to calculate the density from the characteristic function. The second method approximates the distribution of R^h by an NIG distribution. Both methods use the moment-generating function of R^h, which is given by

$$\begin{aligned}\varphi_h(u) := \mathbb{E}\left[\mathbf{e}^{uR^h}\right] &= \mathbb{E}\left[\mathbf{e}^{u[(1-h)Z+Z_1-hZ_2]}\right] = \mathbb{E}\left[\mathbf{e}^{u(1-h)Z}\right]\,\mathbb{E}\left[\mathbf{e}^{uZ_1}\right]\,\mathbb{E}\left[\mathbf{e}^{-uhZ_2}\right] \\ &= M\left(u;\frac{\alpha}{1-h},\frac{\beta}{1-h},(1-h)\mu,(1-h)\delta\right)\,M\left(u;\alpha,\beta,\mu_1,\delta_1\right)\,M\left(-u;\frac{\alpha}{h},\frac{\beta}{h},h\mu_2,h\delta_2\right). \end{aligned} \tag{36}$$

Add copula stuff; here the assumption is still that R_S and R_F are NIG.

Calculation via Fourier inversion Because of the simple form of the moment-generating function (36), the density f of R^h can be calculated numerically using the inverse Fourier transform of the characteristic function $\varphi(u)$:

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathbf{e}^{-iux} \varphi_h(u) \, \mathrm{d}u,$$

and the distribution function of R^h can be calculated as

$$\mathbf{P}(R^h \le x) = \lim_{b \to -\infty} \frac{1}{2\pi} \int_{\mathbb{R}} \frac{\mathbf{e}^{-iux} - \mathbf{e}^{-iub}}{iu} \varphi_h(u) \, \mathrm{d}u,$$

see e.g. II.§12, Theorem 3 of (Shiryaev, 1996).

NIG approximation of hedged position If $h = 1$ and $\beta = 0$, then R^h is NIG, hence, if h is close to 1, it may be feasible to approximate R^h by an NIG. The parameters can be determined either by moment-matching or by a first-order Taylor approximation of the moment-generating function's exponent.
For the moment-matching procedure, set $R^h \approx \mathrm{NIG}(\alpha_h, \beta_h, \mu_h, \delta_h)$ and solve for $\alpha_h, \beta_h, \mu_h, \delta_h$,

$$\begin{pmatrix} \frac{\partial}{\partial u} \varphi_h(u)\big|_{u=0} \\ \frac{\partial^2}{\partial u^2} \varphi_h(u)\big|_{u=0} \\ \frac{\partial^3}{\partial u^3} \varphi_h(u)\big|_{u=0} \\ \frac{\partial^4}{\partial u^4} \varphi_h(u)\big|_{u=0} \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial u} M(u; \alpha_h, \beta_h, \mu_h, \delta_h)_{u=0} \\ \frac{\partial^1}{\partial u^2} M(u; \alpha_h, \beta_h, \mu_h, \delta_h)_{u=0} \\ \frac{\partial^2}{\partial u^3} M(u; \alpha_h, \beta_h, \mu_h, \delta_h)_{u=0} \\ \frac{\partial^3}{\partial u^4} M(u; \alpha_h, \beta_h, \mu_h, \delta_h)_{u=0} \end{pmatrix}.$$

An alternative to determine the parameters is to assume that $h = 1$ so that $R^1 = Z_1 - Z_2$ and using the following first-order Taylor approximation around zero:

$$\sqrt{\alpha^2 - (\beta + u)^2} - \sqrt{\alpha^2 - (\beta - u)^2} \approx -\frac{2\beta}{\sqrt{\alpha^2 - \beta^2}} u.$$

This gives

$$
\begin{aligned}
\varphi_1(u) &= M(u;\alpha,\beta,\mu_1,\delta_1)M(-u;\alpha,\beta,\mu_2,\delta_2)\\
&= \exp\left(\delta_1(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2})+\mu_1\, u+\delta_2(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta-u)^2})-\mu_2\, u\right)\\
&= \exp\left(\delta_1(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2})+\mu_1\, u+\delta_2(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2})-\mu_2\, u\right)\\
&\cdot \exp\left(\delta_2(\sqrt{\alpha^2-(\beta+u)^2}-\sqrt{\alpha^2-(\beta-u)^2})\right)\\
&\approx \exp\left((\delta_1+\delta_2)(\sqrt{\alpha^2-\beta^2}-\sqrt{\alpha^2-(\beta+u)^2})+\left(\mu_1-\mu_2-\frac{2\delta_2\beta}{\sqrt{\alpha^2-\beta^2}}\right)u\right)\\
&= M\left(u;\alpha,\beta,\mu_1-\mu_2-\frac{2\delta_2\beta}{\sqrt{\alpha^2-\beta^2}},\delta_1+\delta_2\right).
\end{aligned}
$$

Figure 6 shows an example of the distribution of $R^{0.95}$ for the different approximations as well as their error relative to direct integration. It turns out that the NIG moment-matching approximation performs best in terms of the error, while the Taylor approximation performs worst. Taking into account CPU times, the Fourier inversion technique performs best in terms of balancing error and CPU time.

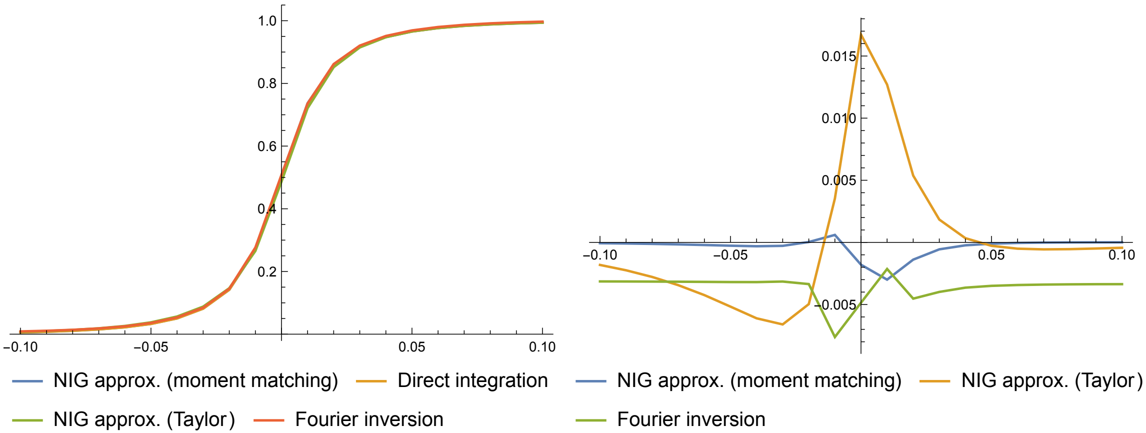


Figure 6: Left: CDF of $R^{0.95}$ using Bitcoin data. Right: Difference of different methods to direct integration. CPU times: NIG approximation (moment matching): 0.2 seconds plus 18.6 seconds for parameter calibration; Direct integration: 82 seconds; NIG approximation (Taylor): 0.18 seconds; Fourier inversion: 0.034 seconds plus 1.55 seconds to generate smooth density function.

[The definition below is one way of introducing the elliptical copula, but not the most practical one for our purposes.]

Definition 4. Elliptical Distribution. The d-dimensional random vector $\boldsymbol{y}$ has an elliptical distribution if and only if the characteristic function $\boldsymbol{t} \mapsto \mathbb{E}\{\exp(i\boldsymbol{t}^\top \boldsymbol{y})\}$ with $\boldsymbol{t} \in \mathbb{R}^d$ has the representation

$$\phi_g(\boldsymbol{t}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\nu}) = \exp(i\boldsymbol{t}^\top \boldsymbol{\mu}) g(\boldsymbol{t}^\top \boldsymbol{\Sigma} \boldsymbol{t}; \boldsymbol{\nu}) \tag{37}$$

where $g(\cdot; \nu) : [0, \infty[\mapsto \mathbb{R}$, $\nu \in \mathbb{R}^d$, and Σ is a symmetric positive semidefinite $d \times d$-matrix.

If r has a density, then the density of $\boldsymbol{y}$ is of the form

$$|\Sigma|^{\frac{1}{2}} g\{(\boldsymbol{y} - \boldsymbol{\mu})^\top \Sigma^{-1} (\boldsymbol{y} - \boldsymbol{\mu})\}. \tag{38}$$

The function $g(\cdot; \nu)$ is known as characteristic generator, whereas $\boldsymbol{\nu}$ is parameter that determines the shape, in particular the tai index of the distribution.

Corollary 3. *(?, equation 2.12) If $\boldsymbol{y}$ follows an elliptical distribution, then $\boldsymbol{y}$ has a stochastic representation*

$$\boldsymbol{y} = \boldsymbol{\mu} + r\boldsymbol{A}^\top \boldsymbol{u}, \tag{39}$$

where $r \in \mathbb{R}_+$ is independent of $\boldsymbol{u}$, and $\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{\Sigma}$.

Distribution	$r \sim$	$g(\boldsymbol{t})$
Gaussian	χ_n	

Table 3: Generators of Elliptical Distributions summarised from (?, Chapter 2)

6.6. t-copulae

The t copula is to represent the dependency structure by t distribution (?Embrechts *et al.*, 2002). ? extend this idea to skewed t copula and grouped t copula to allow more flexibility in the modelling of dependency structure.

6.6.1. Vanilla t-copula

The t-copula is

$$C^t_{\nu,\Sigma}(x) = \int_{-\infty}^{t_\nu^{-1}(x_1)} \cdots \int_{-\infty}^{t_\nu^{-1}(x_n)} \frac{\Gamma\left\{\frac{\nu+i}{2}\right\}}{\Gamma\left\{\frac{\nu}{2}\right\}(\pi\nu)^{i/2}|\Sigma|^{1/2}} \left(1 + \frac{y^\top \Sigma^{-1} y}{\nu}\right)^{-\frac{\nu+i}{2}} dy_1 \dots dy_n, \tag{40}$$

where t_ν^{-1} is the quantile function of a univariate student-t distribution with degree of freedom ν.

6.6.2. Skewed t copula

Mean variance mixture

6.6.3. Double-t copula

[It is OK to introduce the copula without any reference to CDO's.] **?** present an alternative way to the Gaussian copula for valuing CDO tranches. The double-t copula model is a weighted sum of a common (or market) variable M and a idiosyncratic variable Z_i. *[They are t-distributed, right?]* The double-t copula is

$$X_i = w_i M + \sqrt{1 - w_i^2} Z_i \tag{41}$$

where M and Z_i are independent random variables with zero mean and unit variance, and X_i is an indicator variable for i^{th} asset. The authors map the time to default of the i^{th} obligor, t_i, to X_i,

$$F_{X_i}(x) = F_{t_i}(t). \tag{42}$$

In our case, we map X_i to log-returns of portfolio constituents,

$$F_{X_1}(x) = F_{r^S}(s) \text{ and } F_{X_2}(x) = F_{r^F}(t). \tag{43}$$

This is also known as percentile-to-percentile mapping(**?**).

[The percentile-to-percentile mapping is just the property that applying the cdf to a random variable yields a $U(0,1)$ variable:

$$F(x) = \mathbf{P}(X \leq x) = \mathbf{P}(F(X) \leq F(x)) = \mathbf{P}(U \leq F(x)),$$

since by definition, $U \sim U(0,1)$ fulfills $\mathbf{P}(U \leq u) = u$, $0 \leq u \leq 1$. This could be introduced when copulas and Sklar's Theorem are introduced.] The reason for this mapping is to turn incomprehensible dependency structures into known structure.

6.6.4. Normal Inverse Gaussian Copula

Normal Inverse Gaussian (NIG) distribution is a flexible 4-parameter distribution that can produce fat tails and skewness, unlike student-t distribution, NIG's convolution is stable under certain conditions and the CDF, PDF and quantile function can still be computed sufficiently fast (**?**, chapter 5). NIG distribution is a mixture of normal and inverse Gaussian distribution.

Definition 5. Inverse Gaussian Distribution. A non-negative random variable Y has an Inverse Gaussian (IG) distribution with parameters $\alpha > 0$ and $\beta > 0$ if its density funcion is of the form

$$f_{\mathrm{IG}}(y; \alpha, \beta) = \frac{\alpha}{\sqrt{2\pi\beta}} y^{-1.5} \exp\left\{-\frac{(\alpha - \beta z)^2}{2\beta z}\right\} \tag{44}$$

The corresponding distribution function is:

$$F_{\mathrm{IG}}(y; \alpha, \beta) = \frac{\alpha}{\sqrt{2\pi\beta}} \int_0^y z^{-1.5} \exp\left\{-\frac{(\alpha - \beta z)^2}{2\beta z}\right\} dz. \tag{45}$$

We write $Y \sim \mathrm{IG}(\alpha, \beta)$.

Definition 6. Normal Inverse Gaussian Distribution. A random variable X has an Normal Inverse Gaussian (NIG) distribution with parameters α, β, μ and δ if its density funcion is of the form

$$X|Y = y \sim \Phi(\mu + \beta y, y) \tag{46}$$

$$Y \sim \mathrm{IG}(\delta\gamma, \gamma^2) \text{ with } \gamma \stackrel{\text{def}}{=} \sqrt{\alpha^2 - \beta^2} \tag{47}$$

The corresponding distribution function is:

$$F_{\mathrm{NIG}}(y;\alpha,\beta) = \frac{\alpha}{\sqrt{2\pi\beta}} \int_0^y z^{-1.5} \exp\left\{-\frac{(\alpha-\beta z)^2}{2\beta z}\right\} dz. \tag{48}$$

7. Estimation

8. Estimation

8.1. Simulated Method of Moments

This method is suggested by Oh and Patton (2013). In our setting, rank correlation e.g. Spearman's ρ or Kendall's τ, and quantile dependence measures at different levels λ_q are calibrated against their empirical counterparts.

Spearman's rho, Kendall's tau, and quantile dependence of a pair (X, Y) with copula C are defined as

$$\rho_S = 12 \int\int_{I^2} C_{\boldsymbol{\theta}}(u, v)\,\mathrm{d}u\,\mathrm{d}v - 3 \tag{49}$$

$$\tau_K = 4\,\mathsf{E}[C_{\boldsymbol{\theta}}\{F_X(x), F_Y(y)\}] - 1, \tag{50}$$

$$\lambda_q = \begin{cases} \mathbf{P}(F_X(X) \le q | F_Y(Y) \le q) = \dfrac{C_{\boldsymbol{\theta}}(q,q)}{q}, & \text{if } q \in (0, 0.5], \\ \mathbf{P}(F_X(X) > q | F_Y(Y) > q) = \dfrac{1 - 2q + C_{\boldsymbol{\theta}}(q,q)}{1-q}, & \text{if } q \in (0.5, 1). \end{cases} \tag{51}$$

The empirical counterparts are

$$
\begin{aligned}
\hat{\rho}_S &= \frac{12}{n}\sum_{k=1}^{n}\hat{F}_X(x_k)\hat{F}_Y(y_k) - 3,\\
\hat{\tau}_K &= \frac{4}{n}\sum_{k=1}^{n}\hat{C}\{\hat{F}_X(x_i),\hat{F}_X(y_i)\} - 1,\\
\hat{\lambda}_q &= \begin{cases} \frac{1}{n}\sum_{k=1}^{n}\frac{\mathbf{1}_{\{\hat{F}_X(x_k)\le q,\hat{F}_Y(y_k)\le q\}}}{q}, & \text{if } q\in(0,0.5],\\ \frac{1}{n}\sum_{k=1}^{n}\frac{\mathbf{1}_{\{\hat{F}_X(x_k)> q,\hat{F}_Y(y_k)> q\}}}{1-q}, & \text{if } q\in(0.5,1). \end{cases},
\end{aligned}
$$

where $\hat{F}(x) := \frac{1}{n}\sum_{k=1}^{n} 1_{\{x_i\le x\}}$ and $\hat{C}(u,v) := \frac{1}{n}\sum_{k=1}^{n} 1_{\{u_i\le u, v_i\le v\}}$.

We denote $\tilde{\boldsymbol{m}}(\boldsymbol{\theta})$ be a m-dimensional vector of dependence measures according the the dependence parameters $\boldsymbol{\theta}$,and $\hat{\boldsymbol{m}}$ be the corresponding empirical counterpart. The difference between dependence measures and their counterpart is denoted by

$$
\boldsymbol{g}(\boldsymbol{\theta}) = \hat{\boldsymbol{m}} - \tilde{\boldsymbol{m}}(\boldsymbol{\theta}).
$$

The SMM estimator is

$$
\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\operatorname{argmin}}\, \boldsymbol{g}(\boldsymbol{\theta})^\intercal \hat{\boldsymbol{W}} \boldsymbol{g}(\boldsymbol{\theta}),
$$

where $\hat{W}$ is some positive definite weigh matrix.

In this work, we use $\tilde{\boldsymbol{m}}(\boldsymbol{\theta}) = (\rho_S, \lambda_{0.05}, \lambda_{0.1}, \lambda_{0.9}, \lambda_{0.95})^\intercal$ for calibration of Bitcoin price and CME Bitcoin future.

8.2. Maximum Likelihood Estimation

By Sklar's theorem, the joint density of a d-dimensional random variable $\boldsymbol{X}$ with sample size n can be written as

$$\boldsymbol{f}_{\boldsymbol{X}}(x_1, ..., x_d) = \boldsymbol{c}\{F_{X_1}(x_1), ..., F_{X_d}(x_d)\} \prod_{j=1}^{d} f_{X_i}(x_i). \tag{52}$$

We follow the treatment of MLE documented in section 10.1 of Joe (1997), namely the inference functions for margins or IFM method. The log-likelihood $\sum_{i=1}^{n} f_{\boldsymbol{X}}(X_{i,1}, ..., X_{i,d})$ can be decomposed into dependence part and marginal part,

$$\begin{aligned} L(\boldsymbol{\theta}) &= \sum_{i=1}^{n} \boldsymbol{c}\{F_{X_1}(x_{i,1}; \boldsymbol{\delta}_1), ..., F_{X_d}(x_{i,d}; \boldsymbol{\delta}_d); \boldsymbol{\gamma}\} + \sum_{i=1}^{n}\sum_{j=1}^{d} f_{X_j}(x_{i,j}; \boldsymbol{\delta}_j) && (53) \\ &= L_C(\boldsymbol{\delta}_1, ..., \boldsymbol{\delta}_d, \boldsymbol{\gamma}) + \sum_{j=1}^{d} L_j(\boldsymbol{\delta}_j) && (54) \end{aligned}$$

where $\boldsymbol{\delta}_j$ is the parameter of the j-th margin, $\boldsymbol{\gamma}$ is the parameter of the parametric copula, and $\boldsymbol{\theta} = (\boldsymbol{\delta}_1, ..., \boldsymbol{\delta}_d, \boldsymbol{\gamma})$. Instead of searching the $\boldsymbol{\theta}$ is a high dimensional space, Joe (1997) suggests to search for $\hat{\boldsymbol{\delta}}_1, ..., \hat{\boldsymbol{\delta}}_d$ that maximize $L_1(\boldsymbol{\delta}_1), ..., L_d(\boldsymbol{\delta}_d)$, then search for $\hat{\boldsymbol{\gamma}}$ that maximize $L_C(\hat{\boldsymbol{\delta}}_1, ..., \hat{\boldsymbol{\delta}}_d, \boldsymbol{\gamma})$.
That is, under regularity conditions, $(\hat{\boldsymbol{\delta}}_1, ..., \hat{\boldsymbol{\delta}}_d, \hat{\boldsymbol{\gamma}})$ is the solution of

$$\left(\frac{\partial L_1}{\partial \boldsymbol{\delta}_1}, ..., \frac{\partial L_d}{\partial \boldsymbol{\delta}_d}, \frac{\partial L_C}{\partial \boldsymbol{\gamma}}\right) = \mathbf{0}. \tag{55}$$

However, the IFM requires making assumption to the distribution of of the margins. Genest *et al.* (1995) suggests to replace the estimation of marginals parameters estimation by non-parameteric estimation. Given non-parametric estimator $\hat{F}_i$ of the margins F_i, the estimator of the dependence parameters $\boldsymbol{\gamma}$ is

$$\hat{\boldsymbol{\gamma}} = \underset{\boldsymbol{\gamma}}{\operatorname{argmax}} \sum_{i=1}^{n} \boldsymbol{c}\{\hat{F}_{X_1}(x_{i,1}), ..., \hat{F}_{X_d}(x_{i,d}); \boldsymbol{\gamma}\}. \tag{56}$$

8.3. Comparison

Both the simulated method of moments and the maximum likelihood estimation are unbiased and proven to give good fits. The problem remain is which procedure is more suitable for hedging.
Figure 7 shows the empirical quantile dependence of Bitcoin and CME future and the copula implied quantile dependence from MLE and MM calibration procedures. Although the MLE is a better fit to a range of quantile dependence in the middle, it fails to address the situation in the tails. Our data empirically has weaker quantile dependence in the ends, and those points generate PnL to the hedged portfolio. MM is preferred visually as it produces a better fit to the dependence structure in the two extremes.

9. Backtesting

Especially for VaR and ES, at the 99% level, but also at the 95% level, out-of-sample testing the hedge effectiveness by means of a 30-day window may lead to unstable results. To evaluate these risk measures we therefore employ backtesting techniques from Basel regulation (references here).

10. Backtesting value-at-risk

The standard procedure for backtesting VaR, the so-called "traffic light approach", dates back to the beginnings of VaR and Basel regulation BIS (1996). The method is described in e.g. Section 2.3 of McNeil *et al.* (2005).
At time t, let VaR_α^t denote the true one-period value-at-risk (VaR) at the confidence level α. The realised P&L of the position, $\Delta V_t = V_{t+1} - V_t$, is assumed to follow a continuous distribution. A violation (also called exception or outlier) of VaR is expressed by the Bernoulli variable $\mathbf{1}_{t+1} := \mathbf{1}_{\{\Delta V_t > \mathrm{VaR}_\alpha^t\}}$. If the sequence $(\mathbf{1}_{t+1})_{t\in\mathbb{Z}}$ is adapted to the filtration $(\mathcal{F}_t)_{t\in\mathbb{Z}}$ and $\mathbb{E}(\mathbf{1}_{t+1}|\mathcal{F}_t) = 1-p$, for all t, then $(\mathbf{1}_t)_{t\in\mathbb{Z}}$ is a process of identically distributed Bernoulli variables, see Lemma 4.29 McNeil *et al.* (2005).
In order to backtest a VaR model, one-period estimates are compared against realised P&L of the position, giving a sequence $\hat{\mathbf{1}}_{t+1} := \hat{\mathbf{1}}_{\{\Delta V_t > \widehat{\mathrm{VaR}}_\alpha^t\}}$, $t+1 = 1, \ldots, n$. If indeed the model measures VaR at the given level α, then the number of violations X follows a binomial distribution, i.e., $X \sim B(n, 1-p)$. The "traffic light approach" in the Basel accord specifies the acceptable number of violations such that falsely rejecting the null hypothesis of a vaild model at most 5%. In our setting, because $n \cdot p$ is sufficiently large, we approximate the number of violations by a normal distribution, i.e. $X \overset{a}{\sim} \mathrm{N}\,(n\,p, n\,p(1-p))$ and hence reject the null hypothesis if

$$\frac{\hat{X} - n\,p}{\sqrt{n\,p(1-p)}} > \mathrm{N}_{1-\delta},$$

where δ denotes the level of statistical significance.

References

Acerbi, C. Spectral measures of risk: A coherent representation of subjective risk aversion. *Journal of Banking & Finance*, 26(7):1505–1518, 2002.

Barbi, M. and S. Romagnoli. A copula-based quantile risk measure approach to estimate the optimal hedge ratio. *Journal of Futures Markets*, 34(7):658–675, 2014.

Barndorff-Nielsen, O. E. Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of statistics*, 24(1):1–13, 1997.

BIS. Supervisory framework for the use of "backtesting"' in conjunction with the internal models approach to market risk capital requirements. Basel Committee on Banking Supervision, Bank for International Settlements, 1996.

Cotter, J. and K. Dowd. Extreme spectral risk measures: an application to futures clearinghouse margin requirements. *Journal of Banking & Finance*, 30(12):3469–3485, 2006.

Embrechts, P., A. McNeil, and D. Straumann. Correlation and dependence in risk management: properties and pitfalls. *Risk management: value at risk and beyond*, 1:176–223, 2002.

Fermanian, J.-D., D. Radulovic, and M. Wegkamp. Weak convergence of empirical copula processes. *Bernoulli*, 10(5):847–860, October 2004.

Genest, C. and L.-P. Rivest. Statistical inference procedures for bivariate archimedean copulas. *Journal of the American statistical Association*, 88(423):1034–1043, 1993.

Genest, C., K. Ghoudi, and L.-P. Rivest. A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika*, 82(3):543–552, 1995.

Genest, C. Frank's family of bivariate distributions. *Biometrika*, 74(3):549–555, 1987.

Joe, H. *Multivariate models and multivariate dependence concepts*. CRC Press, 1997.

Kalemanova, A., B. Schmid, and R. Werner. The normal inverse Gaussian distribution for synthetic CDO pricing. *The Journal of Derivatives*, 14(3):80–94, 2007.

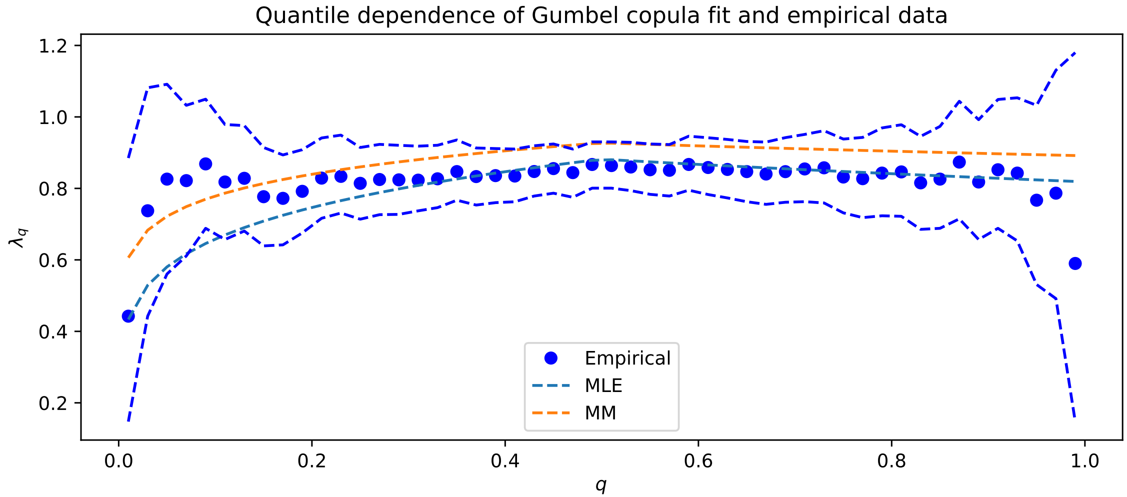
Krupskii, P. and H. Joe. Factor copula models for multivariate data. *Journal of Multivariate Analysis*, 120:85–101, 2013.

McNeil, A., R. Frey, and P. Embrechts. *Quantitative Risk Management*. Princeton University Press, Princeton, NJ, 2005.

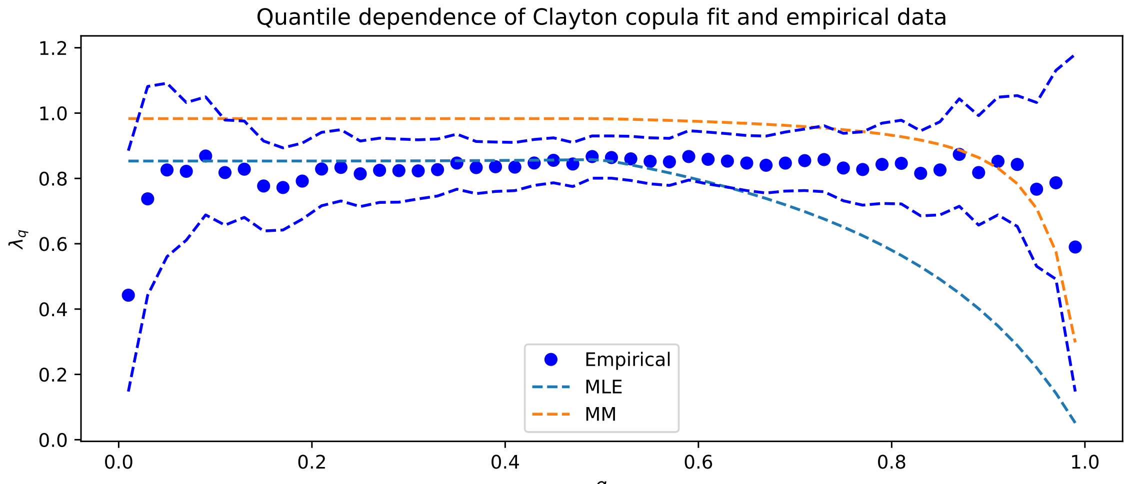
Oh, D. H. and A. J. Patton. Simulated method of moments estimation for copula-based multivariate models. *Journal of the American Statistical Association*, 108(502):689–700, 2013.

Patton, A. and D. H. Oh. Modelling dependence in high dimensions with factor copulas. Working Paper, Duke University, 2012.

Shiryaev, A. N. *Probability*. Springer, Berlin, 2nd edition, 1996.



Quantile dependence of Gumbel copula fit and empirical data
1.2
1.0
0.8
0.6
0.4
0.2
λ_q
0.0
0.2
0.4
0.6
0.8
1.0
q
Empirical
MLE
MM



Quantile dependence of Clayton copula fit and empirical data
1.2
1.0
0.8
0.6
0.4
0.2
0.0
λ_q
0.0
0.2
0.4
0.6
0.8
1.0
Empirical
MLE
MM