Notes on hedging cryptos with spectral risk measures

November 6, 2020

Abstract

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

1. Optimal hedge ratio

Following (Barbi and Romagnoli, 2014), we consider the problem of the optimal hedge ratios by extending 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: $h^* = \text{argmin}_h\, \rho_\phi(s, h)$, 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).
Corollary 2.1 of (Barbi and Romagnoli, 2014), corrected: 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 $C^t_{R^S,R^F}(w, \lambda)$ and continuous

marginals F_{R^S} and F_{R^F}. Then, the s-quantile of R^h solves the following:

$$F_{R^h}(r^h) = 1 - \int_0^1 D_1 C_{R^S,R^F} \left\{ w, F_{R^F} \left[\frac{F_{R^S}^{-1}(w) - r^h}{h} \right] \right\} dw.$$

[..]

Here $D_1 C(u,v) = \dfrac{\partial}{\partial u} C(u,v)$, which can be shown to fulfil (Cherubini *et al.*, 2011)

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

2. 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[1]

(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).

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

2.1. Representation of spectral risk measures

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

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

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

2.2. 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.

3. D_1 Operator

The D_1 operator is given as

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

In the context of the above notation, we obtain

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

The last line can also be written as

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

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.

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

$$= \left.\frac{\partial\phi^{-1}[\phi(v)+\phi\{g(w)\}]}{\partial[\phi(v)+\phi\{g(w)\}]}\frac{\partial\phi(v)}{\partial v}\right|_{v=w} \tag{2}$$

$$= \frac{\partial\phi^{-1}[\phi(w)+\phi\{g(w)\}]}{\partial[\phi(w)+\phi\{g(w)\}]}\frac{\partial\phi(w)}{\partial w} \tag{3}$$

$$\text{, where } g(w) = F_{R^F}\left\{\frac{F_{R^S}^{-1}(w)-r^h}{h}\right\} \tag{4}$$

$$\tag{5}$$

4. Dependence

Dependence through copula (e.g. Student t, Clayton or Gumbel)

4.1. 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:

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

- 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 $\theta \to \infty$.

4.2. Elliptical Copulae

Definition 1. 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

$$\phi_g(\boldsymbol{t}; \boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\nu}) = \exp(i\boldsymbol{t}^\top \boldsymbol{\mu}) g(\boldsymbol{t}^\top \boldsymbol{\Sigma} \boldsymbol{t}; \boldsymbol{\nu}) \tag{6}$$

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.

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 distribtion.

Corollary 1. *(Fang, 2018, equation 2.12) If $\boldsymbol{y}$ follows an elliptical distribution, then $\boldsymbol{y}$ has a stochastic representation*

$$\boldsymbol{y} = \boldsymbol{\mu} + r\boldsymbol{A}^\top \boldsymbol{u}, \tag{7}$$

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 2: Generators of Elliptical Distributions summarised from (Fang, 2018, Chapter 2)

4.3. Extreme-value copulae

5. Estimation

5.1. Two-Stage Estimation

Joe (2005) study the efficiency of a two-stage estimation procedure of copula estimation. The authors also call this method inference function for margins IFM.

Pros

1. Almost as efficient as MLE methods but easier to be implemented
2. Yields an asymptotically Gaussian, unbiased estimate

Cons

1. Subject to specification of marginals Kim *et al.* (2007)

Our data

$$\boldsymbol{y} = \begin{bmatrix} y_{11} & \cdots & y_{1i} \\ \vdots & \ddots & \vdots \\ y_{n1} & \cdots & y_{ni} \end{bmatrix} \tag{8}$$

Let F and f be the joint cdf and joint density of $\boldsymbol{y}$ with parameters $\boldsymbol{\delta}$, and let F_i and f_i be the marginal cdf and marginal density for the i^{th} random variable with parameters $\boldsymbol{\theta}_i$, we have

$$f(\boldsymbol{y};\boldsymbol{\theta}_1,\boldsymbol{\theta}_2,\ldots\boldsymbol{\theta}_i,\boldsymbol{\delta}) = c\{F_1(\boldsymbol{y}_1;\boldsymbol{\theta}_1),F_2(\boldsymbol{y}_2;\boldsymbol{\theta}_2),\ldots,F_i(\boldsymbol{y}_1;\boldsymbol{\theta}_i);\boldsymbol{\delta}\}\prod_{j=1}^{i} f_i(\boldsymbol{y}_j;\boldsymbol{\theta}_j) \tag{9}$$

For a sample of size n, the log-likelihood of functions of the i^{th} univariate margin is

$$L_i(\theta_i) = \sum_{m=1}^{n} \log f_i(y_{mi};\theta_i), \tag{10}$$

and the log-likelihood function for the joint distribution is

$$L(\delta,\theta_1,\theta_2,\ldots,\theta_i) = \sum_{m=1}^{n}\sum_{j=1}^{i} \log f(y_{mj};\delta,\theta_1,\theta_2,...,\theta_i) \tag{11}$$

In most cases, one does not have closed form estimators and numerical techniques are needed. Numerical ML estimation difficulty increase when the total number of parameters increases. The two-stage estimation is designed to overcome this problem.

The two-stage procedure is

1. estimate the univariate parameters from separate univariate likelihoods to get $\tilde{\boldsymbol{\theta}}_1, ..., \tilde{\boldsymbol{\theta}}_i$
2. maximize $L(\boldsymbol{\delta}, \tilde{\boldsymbol{\theta}}_1, \ldots, \tilde{\boldsymbol{\theta}}_i)$ over $\boldsymbol{\delta}$ to get $\tilde{\boldsymbol{\delta}}$

Under regularity conditions [2] , $(\tilde{\boldsymbol{\theta}}_1, \ldots \tilde{\boldsymbol{\theta}}_i, \tilde{\boldsymbol{\delta}})$ is the solution of

$$(\partial L_1/\partial \boldsymbol{\theta}_1^{\intercal}, \ldots, \partial L_i/\partial \boldsymbol{\theta}_i^{\intercal}, \partial L/\partial \boldsymbol{\delta}_1^{\intercal}) = \mathbf{0} \tag{12}$$

For comparison, if we optimize L directly without the two-stage procedure (i.e. MLE), we solve for

$$(\partial L/\partial \boldsymbol{\theta}_1^{\intercal}, \ldots, \partial L/\partial \boldsymbol{\theta}_i^{\intercal}, \partial L/\partial \boldsymbol{\delta}_1^{\intercal}) = \mathbf{0} \tag{13}$$

We denote the two solutions as $\tilde{\boldsymbol{\eta}} = (\tilde{\boldsymbol{\theta}}_1, \ldots \tilde{\boldsymbol{\theta}}_i, \tilde{\boldsymbol{\delta}})$ for two-stage procedure; $\hat{\boldsymbol{\eta}} = (\hat{\boldsymbol{\theta}}_1, \ldots \hat{\boldsymbol{\theta}}_i, \hat{\boldsymbol{\delta}})$ for MLE procedure. and compare the asymptotic relative efficiency of $\tilde{\boldsymbol{\eta}}$ and $\hat{\boldsymbol{\eta}}$.

Asymptotics: yet to be done. Kim *et al.* (2007) show the estimation of $\boldsymbol{\theta}$ may be seriously affected. They compare the two-stage approach and Canonical Maximum Likelihood Method by simulation and conclude that Canonical Maximum Likelihood is prefered from a computational statistics and data analysis point of view.

[2]Regularity conditions include 1. $\exists \frac{\partial \log f(x;\theta)}{\partial \theta}, \frac{\partial^2 \log f(x;\theta)}{\partial \theta^2}, \frac{\partial^3 \log f(x;\theta)}{\partial \theta^3}$ for all x; 2. $\exists g(x), h(x) and H(x)$ such that for θ in a neighborhood $N(\theta_0)$ the relations $\left|\frac{\partial f(x;\theta)}{\partial theta}\right| \leq g(x)$, $\left|\frac{\partial^2 f(x;\theta)}{\partial \theta^2}\right| \leq h(x)$, $\left|\frac{\partial^3 f(x;\theta)}{\partial \theta^3}\right| \leq H(x)$ hold for all x, and $\int g(x)dx < \infty$, $\int h(x)dx < \infty$, $\mathbb{E}_\theta\{H(X)\} < \infty$ for $\theta \in N(\theta_0)$; 3. For each $\theta \in \Theta$, $0 < \mathbb{E}_\theta \left\{ \left(\frac{\partial \log f(X;\theta)}{\partial \theta} \right)^2 \right\}$. For detail see section 4.2.2 of Serfling (2009)

5.2. Canonical Maximum Likelihood Method

This approach was studied by Genest *et al.* (1995) and Shih and Louis (1995). also by Wang and Ding, 2000; Tsukahara, 2005 This is also known as pseudo maximum likelihood (PML) and as canonical maximum likelihood (see Cherubini et al., 2004)

Genest and Werker (2002) obtained conditions under which the PMLE is asymptotically efficient.

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.

Cherubini, U., S. Mulinacci, and S. Romagnoli. A copula-based model of speculative price dynamics in discrete time. *Journal of Multivariate Analysis*, 102(6):1047–1063, 2011.

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.

Fang, K. W. *Symmetric multivariate and related distributions*. CRC Press, 2018.

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.

Joe, H. Asymptotic efficiency of the two-stage estimation method for copula-based models. *Journal of multivariate Analysis*, 94(2):401–419, 2005.

Kim, G., M. J. Silvapulle, and P. Silvapulle. Comparison of semiparametric and parametric methods for estimating copulas. *Computational Statistics & Data Analysis*, 51(6):2836–2850, 2007.

Serfling, R. J. *Approximation theorems of mathematical statistics*, volume 162. John Wiley & Sons, 2009.

Shih, J. H. and T. A. Louis. Inferences on the association parameter in copula models for bivariate survival data. *Biometrics*, pages 1384–1399, 1995.