Conditional Expected Shortfall

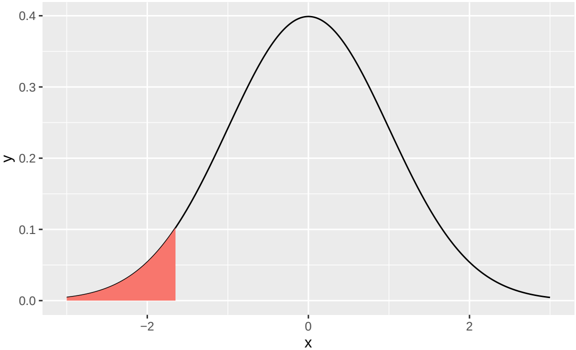
Nonparametric Estimation

Young-geun Kim
ygeunkim.github.io
2019711358, Department of Statistics

27 Nov, 2019

Expected Shortfall

Value at Risk



Tsay (2010) says that

> Measure of loss under *normal* market conditions
> Minimal loss under *extraordinary* market circumstances

Value at Risk

p: **Right** tail probability
l: Time horizon
$L(l)$: loss function of the asset
F: CDF of the loss

$$p = P\left[L(l) \geq VaR\right]$$

Subadditivity

Coherent risk measure

- Homogeneity
- Monotonicity
- Translation invariance (risk-free condition)
- Subadditivity

VaR

does not satisfy subadditivity

- When two portfolios are merged, the risk measure should not be greater than the sum of each.
- VaR *underestimates* the actual loss.

Conditional VaR

Stationary log-return $\{Y_t \colon t = 1, \ldots n\}$
Exogenous variable $\{X_t \colon t = 1, \ldots n\}$
Conditional VaR (CVaR) or Expected Shortfall (ES)

$$\nu_p(x) = S^{-1}(p \mid x)$$

where

$S(y \mid x) := 1 - F(y \mid x)$
F: conditional CDF of Y_t given $X_t = x$.

Conditional Expected Shortfall

We are interested in *Expected Shortfall given exogenous variable values*
Conditional Expected Shortfall (CES)

$$\mu_p(x) = E\left[Y_t \mid Y_t \geq \nu_p(x), X_t = x\right]$$

Formulating CES

Let $B \equiv \{\omega\colon Y_t \geq \nu_p(x)\} \in \mathcal{B}$. Then

$$
\begin{aligned}
\mu_p(x) &= E\left[Y_t \mid Y_t \geq \nu_p(x), X_t = x\right] \\
&= \frac{1}{P(B)} \int_B Y_t dP \\
&= \frac{1}{P\left(Y_t \geq \nu_p(x) \mid X_t = x\right)} \int_{\nu_p(x)}^{\infty} y f(y \mid x) dy \\
&= \frac{1}{p} \int_{\nu_p(x)}^{\infty} y f(y \mid x) dy
\end{aligned}
$$

Nonparametric Estimation

Workflow of Estimation

Plugging-in Method

$$\hat{\mu}_p(x) = \frac{1}{p}\int_{\hat{\nu}_p(x)}^{\infty} y\hat{f}(y \mid x)dy$$

What to estimate

Conditional PDF: $\hat{f}(y \mid x)$
CVaR: $\hat{\nu}_p(x) = \hat{S}^{-1}(p \mid x)$

Conditional Disribution

Taylor expansion

Consider any symmetric kernel $K_h(\cdot)$. Then

$$\begin{aligned} E[K_h(y - Y_t) \mid X_t = x] &= K_h * f_{y|x}(y) \\ &= f(y \mid x) + \frac{h^2}{2}\mu_2(K) f^{(2)}(y \mid x) + o(h^2) \end{aligned}$$

where $\mu_j(K) = \int_{\mathbb{R}} u^j K(u) du$.

Smoothing

$$f(y \mid x) \approx E\left[K_h(y - Y_t) \mid X_t = x\right]$$

Methods

Local Linear
Weighted Nadaraya Watson
WDKLL (Cai and Wang 2008)

Local Linear

Denote $Y_t^*(y) \equiv K_h(y - Y_t)$.

$$\hat{f}(y \mid x) = \underset{\alpha(x),\beta(x)}{\operatorname{argmin}} \sum_{t=1}^{n} W_\lambda(x - X_t) \left[Y_t^*(y) - \alpha(x) - \beta(x)(X_t - x) \right]^2$$

Since this is involved in the two kernel ($K_h(\cdot)$, $W_\lambda(\cdot)$), Cai and Wang (2008) names this as *double kernel*.

Local Linear Solution

Note that the local linear estimate is equivalent to WLS.

$$\mathbf{Y}_y^* = (Y_1(y), \dots, Y_n(y))^T \in \mathbb{R}^n$$
$$\mathbf{b}_x(x_t) := (1, x_t - x)^T \in \mathbb{R}^2 \text{ and } \mathbf{b}_x(x) = \mathbf{e}_1 := (1, 0)^T$$
$$X_x := \left(\mathbf{b}_x(x_i)^T\right) \in \mathbb{R}^{n \times 2}$$
$$W_x := diag(W_\lambda(x - X_j)) \in \mathbb{R}^{n \times n}$$

Then $\hat{f}_{ll} = \hat{\alpha}$:

$$\begin{aligned} \hat{f}_{ll}(y \mid x) &= \mathbf{e}_1^T (X_x^T W_x X_x)^{-1} X_x^T W_x \mathbf{Y}_y^* \\ &= \mathbf{l}(x)^T \mathbf{Y}_y^* \\ &\equiv \sum_{t=1}^{n} l_t(x) Y_t^*(y) \end{aligned}$$

Linear Smoother

$$\mathbf{l}(x)^T = \mathbf{e}_1^T (X_x^T W_x X_x)^{-1} X_x^T W_x$$

By annoying arithmetic,

$$l_t(x) = \frac{S_2(x) - (X_t - x) S_1(x)}{S_0(x) S_2(x) - [S_1(x)]^2} W_\lambda(x - X_t)$$

where $S_j(x) := \sum_{t=1}^{n} W_\lambda(x - X_t)(X_t - x)^j$.

Matrix computations

Let $w_t \equiv W_\lambda(x - X_t)$

$$(X_x^T W_x X_x) = \left[\begin{array}{cc} \sum_t w_t & \sum_t w_t(x_t - x) \\ \sum_t w_t(x_t - x) & \sum_t w_t(x_t - x)^2 \end{array} \right] \equiv \left[\begin{array}{cc} S_0 & S_1 \\ S_1 & S_2 \end{array} \right]$$

$$X_x^T W_x = \left[\begin{array}{ccc} w_1 & \cdots & w_n \\ w_1(x_1 - x) & \cdots & w_n(x_n - x) \end{array} \right]$$

Thus,

$$\mathbf{l}(x)^T = \frac{1}{S_0 S_2 - S_1^2} \left[\begin{array}{ccc} S_2 w_1 - S_1 w_1(x_1 - x) & \cdots & S_2 w_n - S_1 w_n(x_n - x) \end{array} \right]$$

Discrete Moments Conditions

$$S_j(x) := \sum_{t=1}^{n} W_\lambda(x - X_t)(X_t - x)^j = \delta_{0,j} = \begin{cases} 1 & j = 0 \\ 0 & \text{o/w} \end{cases}$$

CVaR

Invert $\hat{F}_{ll}(y \mid x)$

Conditional CDF

$$\hat{F}_{ll}(y \mid x) = \int_{\infty}^{y} \hat{f}_{ll}(y \mid x) dy$$
$$= \sum_{t=1}^{n} l_t(x) G_h(y - Y_t)$$

where $G(\cdot)$ is the cdf of $K(\cdot)$.

Problem

It must be $\hat{F}_{ll} \in [0, 1]$ and monotone increasing
However, LL does not guarantee these properties.

Weighted Nadaraya Watson

To get the right shape of CDF

$$\hat{F}_{NW}(y \mid x) = \sum_{t=1}^{n} H_t(x,\lambda) I(Y_t \leq y)$$

where

$$H_t(x,\lambda) = \frac{p_t(x) W_\lambda(x - X_t)}{\sum_{i=1}^{n} p_i(x) W_\lambda(x - X_i)}$$

$p_t(x)$ is *weighted* for each NW weight.
Cai (2001) finds the best weights $\{p_t\}_1^n$ by maximizing the *empirical likelihood*.

Choosing weights

Constraints

$p_t(x) \geq 0$
$\sum_t p_t(x) = 1$
Discrete moments conditions $\sum_{t=1}^{n} H_t(x, \lambda)(X_t - x)^j = \delta_{0,j},\ 0 \leq j \leq 1$

Empirical likelihood

Maximize $\sum_t \ln p_t(x)$. Lagrangian multiplier gives that

$$p_t(x) = \frac{1}{n\left[1 + \gamma(X_t - x) W_\lambda(x - X_i)\right]} \geq 0$$

and $\gamma(\cdot)$ uniquely maximizing the log of the empirical likelihood

$$L_n(\gamma) = -\sum_{t=1}^{n} \ln\left[1 + \gamma(X_t - x) W_\lambda(x - X_i)\right]$$

Weighted Double Kernel Local Linear

In a local linear scheme,
replace linear smoother with WNW weight

$$\hat{f}_{cai}(y \mid x) = \sum_{t=1}^{n} H_t(x, \lambda) Y_t^*(y)$$

and hence,

$$\begin{aligned}\hat{F}_{cai}(y \mid x) &= \int_{\infty}^{y} \hat{f}_{cai}(y \mid x) dy \\ &= \sum_{t=1}^{n} H_t(x, \lambda) G_h(y - Y_t)\end{aligned}$$

Inverting and Plugging-in

CVaR

$$\hat{\nu}_p^{(cai)}(x) = \hat{S}_{cai}^{-1}(p \mid x)$$

where $\hat{S}_{cai}(y \mid x) = 1 - \hat{F}_{cai}(y \mid x)$

CES

$$\hat{\mu}_p(x) = \frac{1}{p} \sum_{t=1}^{n} H_t(x, \lambda) \left[Y_t \bar{G}_h(\hat{\nu}_p(x) - Y_t) + h G_{1,h}(\hat{\nu}_p(x) - Y_t) \right]$$

where $\bar{G}(u) = 1 - G(u)$ and $G_1(u) = \int_u^{\infty} vK(v)dv$.

Statistical Properties

References

Cai, Zongwu. 2001. "Weighted Nadaraya–Watson Regression Estimation." *Statistics & Probability Letters* 51 (3): 307–18.

Cai, Zongwu, and Xian Wang. 2008. "Nonparametric estimation of conditional VaR and expected shortfall." *Journal of Econometrics* 147 (1): 120–30. https://doi.org/10.1016/j.jeconom.2008.09.005.

Tsay, Ruey S. 2010. *Analysis of Financial Time Series*. John Wiley & Sons.