Conditional Expected Shortfall

Nonparametric Estimation

Young-geun Kim
ygeunkim.github.io

Statistics, SKKU

13 Dec, 2019

1. Risk Measures
2. Nonparametric Estimation
3. Simulation
4. Data Analysis

Reviewed Paper

[1] Z. Cai and X. Wang. "Nonparametric estimation of conditional VaR and expected shortfall". In: *Journal of Econometrics* 147.1 (2008), pp. 120-130. ISSN: 0304-4076. DOI: 10.1016/j.jeconom.2008.09.005.

Keywords

Boundary effects, Empirical likelihood, Expected shortfall, Local linear estimation, Nonparametric smoothing, Value-at-risk, Weighted double kernel

Section 1

Risk Measures

Value at Risk

Given time horizon,

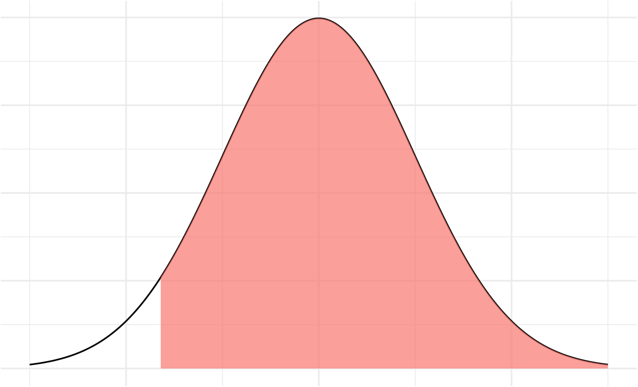


Figure 1: Loss Distribution - Can the financial institution still be in business after a catastrophic event?

Two Viewpoints

Tsay (2010) says that

- Financial institution: Maximal loss of a financial position during a given time period for a given probability
  - Measure of loss under *normal* market conditions
- Regulatory committe: Minimal loss under *extraordinary* market circumstances

Definition

- p: **Right** tail probability
- l: Time horizon
- $L(l)$: loss function of the asset from t to $t+l$
- F_l: CDF of $L(l)$

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

i.e. VaR can be computed by finding the p-th quantile.

Quantile Loss

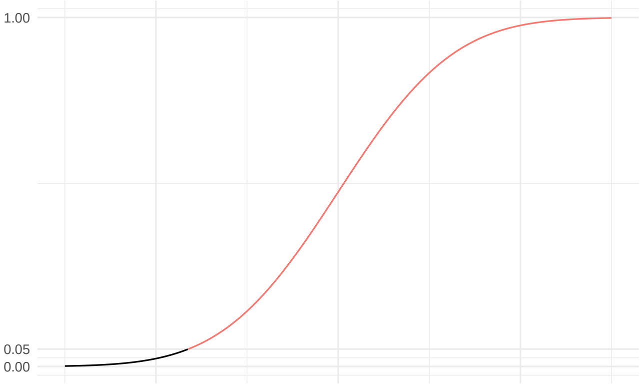


Figure 2: CDF of Loss

$$VaR = \inf\{x\colon F_l(x) \geq 1-p\}$$

Loss function and log Returns

Consider dollar $\{P_t\}$.

$$L(l) = P_{t+l} - P_t = R_{t+l} + \ldots + R_{t+1}$$

where $R_t = P_t - P_{t-1}$ is the return series.

- Loss occurs when the return R_t are *negative*
- Use *negative returns*

log Returns

Taylor expansion

For any $x_0 > 0$,

$$\ln x \approx \ln x_0 + \frac{1}{x_0}(x - x_0)$$

Write $x = x_2$, $x_0 = x_1$. Then

$$\ln \frac{x_2}{x_1} \approx \frac{x_2}{x_1} - 1 = \frac{x_2 - x_1}{x_1}$$

Percentage change

- Log returns $Y_t = \ln R_t$ correspond *approximately to percentage changes*
- Use *negative log returns*
- VaR by **Upper quantile of the distribution of log return**

VaR using log Return

log return

Cai and Wang (2008) used the following value in a real example part.

$$-100Y_{t+1} = -100\ln\frac{P_{t+1}}{P_t} \approx \text{percentage loss}$$

Dollar amount of VaR

From t to $t+1$,

$$VaR = P_t \times VaR(-100Y_{t+1})$$

Subadditivity

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

VaR and Expected Loss

- VaR: Quantile loss of the loss distribution
- ES: Expected value of loss function if the *VaR is exceeded*

$$ES := E\left[L(l) \mid L(l) \geq VaR\right]$$

Conditional information

- $\{X_t \colon t = 1, \dots n\}$
- Exogenous variable: economic, market variables
- or *past observed returns* e.g. $\{Y_{t-1}\}$

Conditional VaR

- Stationary log-return $\{Y_t \colon t = 1, \ldots n\}$
- Conditional information $\{X_t \colon t = 1, \ldots n\}$
- Conditional VaR (CVaR)

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

where

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

Conditional Expected Shortfall

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

Section 2

Nonparametric Estimation

Goal

Risk measures

- CVaR: $\hat{\nu}_p(x)$
- CES: $\hat{\mu}_p(x)$

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: $\hat{\nu}_p(x) = \hat{S}^{-1}(p \mid x)$ by inverting the conditional CDF

Weighted Kerel

To make $\hat{F}(y \mid x)$ satisfy its interval $[0, 1]$, Weighted Nadaraya Watson scheme (Cai 2001) is introduced.

Weight

$$W_{c,t}(x, h) = \frac{p_t(x) W_h(x - X_t)}{\sum_{i=1}^{n} p_i(x) W_h(x - X_i)}$$

Weights of kernel

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

Find λ which maximizes the log of the empirical likelihood,

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

Newton-Raphson

First order derivative

$$L_n'(\lambda) = \sum_{t=1}^{n} \frac{(X_t - x)W_h(x - X_i)}{1 + \lambda(X_t - x)W_h(x - X_i)}$$

Second order derivative

$$L_n''(\lambda) = -\sum_{t=1}^{n} \frac{\{(X_t - x)W_h(x - X_i)\}^2}{(1 + \lambda(X_t - x)W_h(x - X_i))^2} = -L_n'(\lambda)^2$$

Algorithm 1: Newton-Raphson method finding lambda

input : empirical log-likelihood, first and second order derivative

Initialize $\hat{\lambda}_{old}$;

while $|\hat{\lambda}_{old} - \hat{\lambda}_{new}| > \epsilon$ **do**

  Update λ by

$$\hat{\lambda}_{new} = \hat{\lambda}_{old} - \frac{L_n'(\hat{\lambda}_{old})}{L_n''(\hat{\lambda}_{old})}$$

end

output: $\hat{\lambda}_{new}$

Weighted Double Kernel Local Linear

Combining Double Kernel Local Liear (Cai and Wang 2008) and WNW (Cai 2001),

Conditional pdf

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

where initial estimate $Y_t^*(y) = K_{h_0}(y - Y_t)$ with symmetric kernel K_{h_0}.

Conditional cdf

$$\hat{F}_{cai}(y \mid x) = \sum_{t=1}^{n} W_{c,t}(x,h) G_{h_0}(y - Y_t)$$

where G_{h_0} is the cdf of K_{h_0}, i.e.

$$G_{h_0}(y - Y_t) = \int_{-\infty}^{y} K_{h_0}(y - Y_t) dy$$

WDKLL of CVaR

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

and **invert it**

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

Inverting method

Goal: find `y` value that has `1 - p` as cdf value.

1. Make grid of `y`
2. Make cdf value for each `y`
3. Take every value that is equal or larger than `1 - p`
4. Take minimum among them

WDKLL of CES

Plug-in the previous results

$$\hat{\mu}_p(x) = \frac{1}{p} \sum_{t=1}^{n} W_{c,t}(x,h) \left[Y_t \bar{G}_{h_0}(\hat{\nu}_p(x) - Y_t) + h G_{1,h_0}(\hat{\nu}_p(x) - Y_t) \right]$$

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

Asymptotic Normality

Bias of

- $\hat{f}_{cai}(y \mid x)$
- $\hat{S}_{cai}(y \mid x) = 1 - \hat{F}_{cai}(y \mid x)$
- $\hat{\nu}_p(x)$
- $\hat{\mu}_p(x)$

at both

Interior	Boundary
x	$x = ch$

AMSE

Bias

Note that

$$\hat{\mu}_p(x) - \mu(x) = O_p\left(h^2 + h_0^2 + (nh)^{-\frac{1}{2}}\right)$$

and hence, $\hat{\mu}_p(x)$ is a *consistent* with a convergent rate $\sqrt{nh}$

Optimal Bandwidth

$$n^{-\frac{4}{5}}$$

Bandwidth Selection

Criterion

- Nonparametric AIC (Cai and Tiwari 2000)

Two bandwidths

- Initial bandwidth h: insensitive to the final estimation
- WNW bandwith λ

Strategy

Use linear estimators

- WNW estimator: select one $\tilde{h}$
  - $h \leq 0.1\tilde{h}$: take small initial bandwidth
- Given h
  - Use $\hat{F}_{cai}$

Section 3

Simulation

AR(1)-GARCH(1, 0)

$$
\begin{cases}
X_t = Y_{t-1} \\
Y_t = 0.01 + 0.62 X_t + \sigma_t \epsilon_t \\
\sigma_t^2 = 0.15 + 0.65 \sigma_{t-1}^2 \\
\epsilon_t \sim N(0, 1)
\end{cases}
$$

True conditional distribution

Since $\epsilon_t \sim N(0,1)$,

$$\sigma_t \epsilon_t \sim N(0, \sigma_t^2)$$

$$Y_t \mid X_t \sim N\left(0.01 + 0.62 X_t, \sigma_t^2\right)$$

True CES

- For each X_t, `pnorm(x, mean, sd)` gives the conditional cdf value.
- Inverting $S(y \mid x) = 1 - F(y \mid x)$ gives $\nu_p(x)$.

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

- Plugging-in method gives $\mu_p(x)$.

$$\mu_p(x) = \frac{1}{p} \int_{\nu_p(x)}^{\infty} y f(y \mid x) dy$$

Goal of MC Simulation

- Compute the error between the true $\mu_p(x)$ and $\hat{\mu}_p(x)$
- Is the estimator of Cai and Wang (2008) good?

Process

Monte Carlo Samples:

1. For fixed x_t (pre-determined grid points)
2. Generate GARCH(1, 0): (σ_t, ϵ_t)
3. Generate Y_t using AR(1) for each $X_t = Y_{t-1}$
4. AR(1): $Y_t = 0.01 + 0.62 Y_{t-1} + \sigma_t \epsilon_t$

Expected Prediction Error

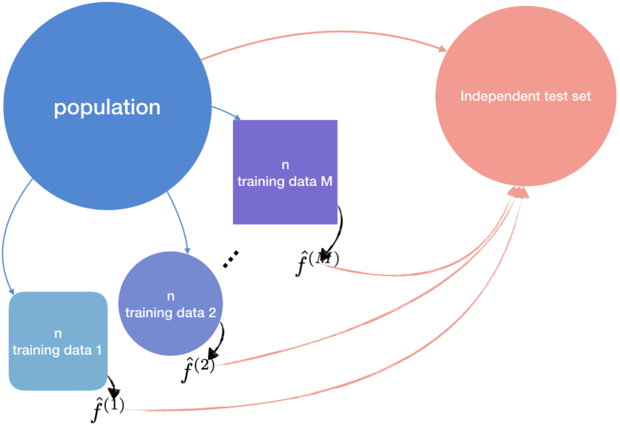
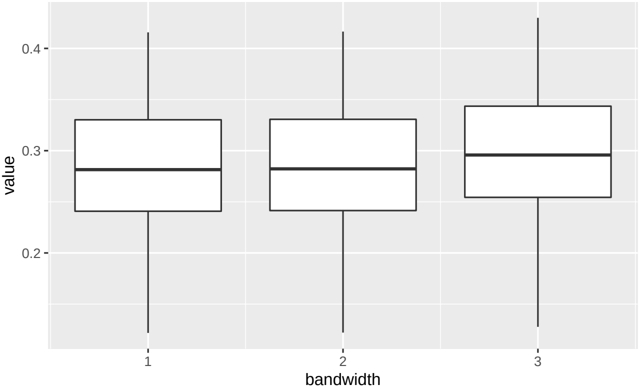


Figure 3: Simulating Expected prediction error

Monte Carlo Samples

```
#>        sigma    garch  mc    xt     yt
#>    1: 0.655 -0.41011  s1 0.005 -0.397
#>    2: 0.655  0.12022  s1 0.010  0.136
#>    3: 0.655 -0.54705  s1 0.015 -0.528
#>    4: 0.655  1.04436  s1 0.020  1.067
#>    5: 0.655  0.21571  s1 0.025  0.241
#>   ---
#> 1996: 0.655 -0.00527 s20 0.480  0.302
#> 1997: 0.655  0.67677 s20 0.485  0.987
#> 1998: 0.655 -0.52315 s20 0.490 -0.209
#> 1999: 0.655  0.65742 s20 0.495  0.974
#> 2000: 0.655 -0.20423 s20 0.500  0.116
```

Numerical Results



Section 4

Data Analysis

Bitcoin

Figure 4: Bitcoin price in USD

Daily Return

As mentioned, Cai and Wang (2008) used daily return defined by $y_t := -100 \ln \frac{P_t}{P_{t-1}}$

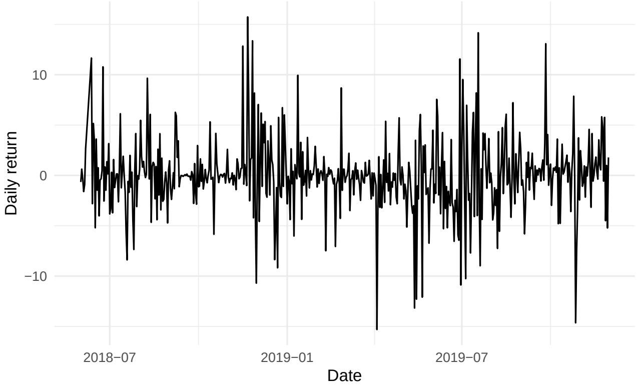


Figure 5: Daily return of bitcoin price

CES for DJI

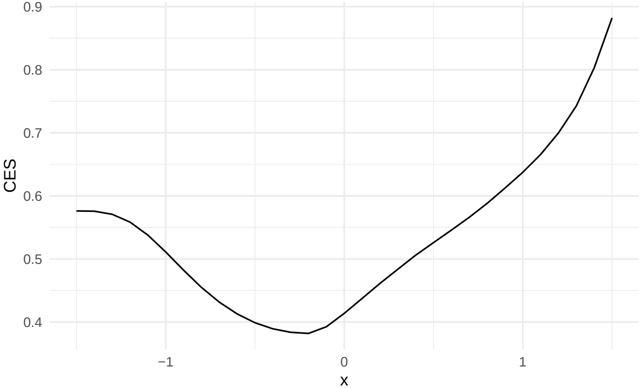


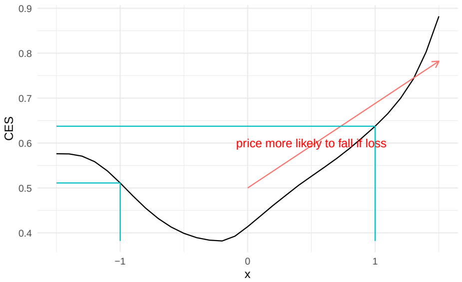
Figure 6: Conditional Expected Shortfall given each lagged variable value

- Volatility smile
- Conditional information x: positive y_{t-1} means loss

Interpretation

Following Cai and Wang (2008),

- Risk tends to be lower when *lagged log loss* is close to the emprirical average,
- and larger otherwise
- Bitcoin price is more likely to fall *if there were a loss* within the last day than if there was a same amount of positive return.



References

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

Cai, Zongwu, and Ram C Tiwari. 2000. "Application of a Local Linear Autoregressive Model to Bod Time Series." *Environmetrics: The Official Journal of the International Environmetrics Society* 11 (3): 341–50.

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.