Generalizing expectation to general risk measures

Yuxi Liu

Australian National University

Monday 28th October, 2019

Overview

An investment problem (4 min)

If you have some cash, and you'd like to purchase some stocks, hold for a year, then sell them. How do you maximize gain? That is, minimize loss?

An investment problem (4 min)

If you have some cash, and you'd like to purchase some stocks, hold for a year, then sell them. How do you maximize gain? That is, minimize loss?

Let there be N stocks to pick from, and your strategy is to invest a portion x_i of the cash into stock i. Then your strategies are defined by a vector $x = (x_1, ...x_N)$

An investment problem (4 min)

If you have some cash, and you'd like to purchase some stocks, hold for a year, then sell them. How do you maximize gain? That is, minimize loss?

Let there be N stocks to pick from, and your strategy is to invest a portion x_i of the cash into stock i. Then your strategies are defined by a vector $x = (x_1, ...x_N)$

Let the loss from the investment be $L = L(x)$. Since investment outcome is uncertain, $L(x)$ is a random variable. We would like to find some x that "minimizes" $L(x)$.

An investment problem (4 min)

If you have some cash, and you'd like to purchase some stocks, hold for a year, then sell them. How do you maximize gain? That is, minimize loss?

Let there be N stocks to pick from, and your strategy is to invest a portion x_i of the cash into stock i. Then your strategies are defined by a vector $x = (x_1, ...x_N)$

Let the loss from the investment be $L = L(x)$. Since investment outcome is uncertain, $L(x)$ is a random variable. We would like to find some x that "minimizes" $L(x)$.

$L(x)$ is random, so it can't be minimized directly. The usual solution is to minimize its expectation:

$$x^* = \arg\min_x \mathbb{E}(L(x))$$

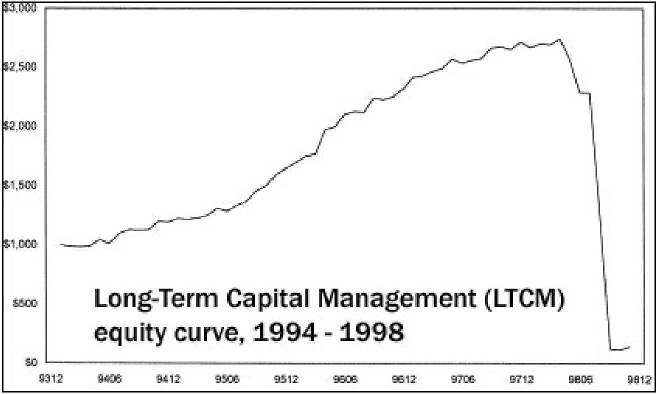
Expectation can be dangerous (7 min)

Minimizing expectation could be, however, dangerous when there is a small chance of catastrophe.

Consider an example from finance.

Expectation can be dangerous (7 min)

Collapse of Long-Term Capital Management cost $4 billion.



Controlling the tail (13 min)

To deal with the dangerous long tail, we can control its tail directly.

Controlling the tail (13 min)

To deal with the dangerous long tail, we can control its tail directly.

Consider a normally distributed random variable $X \sim \mathcal{N}(\mu, \sigma^2)$.
$\mathcal{N}$ denotes a normal distribution with mean μ and variance σ^2.

Controlling the tail (13 min)

To deal with the dangerous long tail, we can control its tail directly.

Consider a normally distributed random variable $X \sim \mathcal{N}(\mu, \sigma^2)$.
$\mathcal{N}$ denotes a normal distribution with mean μ and variance σ^2.

Then, instead of the expectation, we consider the tail expectation:

$$\mathbb{E}(X|X > q_\alpha(X))$$

where $q_\alpha(X)$ is the α-quantile of X.

CVaR: Conditional value at risk (18 min)

Definition of CVaR

For any random variable X, and $0 \leq \alpha < 1$,

$$\mathrm{CVaR}_\alpha(X) = \mathbb{E}(X | X > q_\alpha(X))$$

The $\alpha = 1$ case is special: $\mathrm{CVaR}_1(X) = \sup(X)$

CVaR: Conditional value at risk (18 min)

Definition of CVaR

For any random variable X, and $0 \leq \alpha < 1$,

$$\mathrm{CVaR}_\alpha(X) = \mathbb{E}(X | X > q_\alpha(X))$$

The $\alpha = 1$ case is special: $\mathrm{CVaR}_1(X) = \sup(X)$

Example

Let X be uniform over $[0, 1]$, then,

$$q_\alpha(X) = \alpha, \mathrm{CVaR}_\alpha(X) = \frac{1}{2}(1 + \alpha)$$

Discrete approximation (23 min)

Given random variable X, we can construct an approximation of X by sampling the first n terms of its IID process $X_1, X_2, \ldots X_n$. Then let L_n be a random variable that is equal to X_i with probability $1/n$.

Discrete approximation (23 min)

Given random variable X, we can construct an approximation of X by sampling the first n terms of its IID process $X_1, X_2, ...X_n$. Then let L_n be a random variable that is equal to X_i with probability $1/n$.

If n is big, and X is "nice", then L_n should be "similar" to X. For example, we should have

$$\mathbb{E}(L_n) = \frac{1}{n}\sum_{i=1}^{n} X_i \approx \mathbb{E}(X)$$

Side remark: This is essentially "bootstrapping" from statistics.

Central limit theorem (33 min)

Intuitively, the central limit theorem states that for any X with variance σ^2, we have

$$\mathbb{E}(L_n) = \frac{1}{n}\sum_{i=1}^{n} X_i \approx \mathbb{E}(X) + \frac{1}{\sqrt{n}}\mathcal{N}(0,\sigma^2) + o\left(n^{-1/2}\right)$$

That is, $\sqrt{n}(\mathbb{E}(L_n) - \mathbb{E}(X))$ converges to $\mathcal{N}(0,\sigma^2)$ in distribution. This suggests the generalization

Central limit theorem for CVaR

For any X with finite variance, there exists some function $\sigma : [0,1) \to [0,\infty)$, such that

$$\mathrm{CVaR}_\alpha(L_n) \approx \mathrm{CVaR}_\alpha(X) + \frac{1}{\sqrt{n}}\mathcal{N}(0,\sigma(\alpha)^2) + o\left(n^{-1/2}\right)$$

Central limit theorem (33 min)

Central limit theorem for CVaR

Since L_n is a mixture of $X_1, \dots X_n$, we have

$$\mathrm{CVaR}_\alpha(L_n) \approx \frac{1}{(1-\alpha)n} \sum_{i=1}^{(1-\alpha)n} X_{(i)}$$

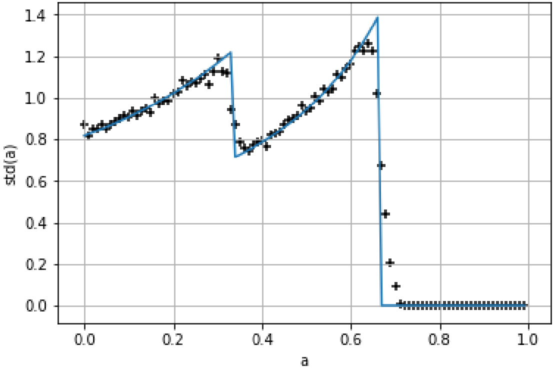
where $X_{(i)}$ is the i-th greatest among all $X_1, \dots X_n$.

We proved that $\sigma(\alpha)^2$ equals

$$\mathbb{V}\left(\frac{1}{1-\alpha} (X - q_\alpha(X))^+ \right)$$

Numerical experiments (38 min)

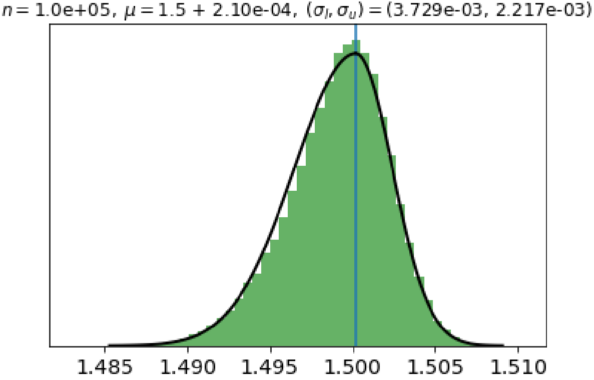
For X uniform over $\{0, 1, 2\}$, we generated trials of $\text{CVaR}_\alpha(L_{1000})$, and graphed theoretical vs actual $\sigma(\alpha)^2$:



What happens at the "jumps", like $\alpha = 1/3$?

Numerical experiments (38 min)

For X uniform over $\{0, 1, 2\}$, we generated trials of $\mathrm{CVaR}_\alpha(L_{1000})$,



The distribution of $\mathrm{CVaR}_{1/3}(L_n)$ becomes "mixed Gaussian"!

Proof of central limit theorem (43 min)

The proof of the theorem proceeded in 4 steps:

1. Use the Gärtner–Ellis theorem to calculate the result as an integral equation.

Proof of central limit theorem (43 min)

The proof of the theorem proceeded in 4 steps:

1. Use the Gärtner–Ellis theorem to calculate the result as an integral equation.
2. Solve the equation when X is a mixture of uniform distributions over intervals on the real line.

Proof of central limit theorem (43 min)

The proof of the theorem proceeded in 4 steps:

1. Use the Gärtner–Ellis theorem to calculate the result as an integral equation.
2. Solve the equation when X is a mixture of uniform distributions over intervals on the real line.
3. Take the limit so that X has discrete distribution.

Proof of central limit theorem (43 min)

The proof of the theorem proceeded in 4 steps:

1. Use the Gärtner–Ellis theorem to calculate the result as an integral equation.
2. Solve the equation when X is a mixture of uniform distributions over intervals on the real line.
3. Take the limit so that X has discrete distribution.
4. A general X distribution is takes as the limit of a sequence of discrete distributions.

Strong law of large numbers (48 min)

Finally, we have the remarkable generalization

Uniform strong law of large numbers for general risk measures

Let X be a random variable with bounded range.
With probability 1, for any m probability distribution over $[0,1]$, the risk measure defined by

$$\mathcal{F}(X) = \int_0^1 \mathrm{CVaR}_\alpha(X) dm(\alpha)$$

gives

$$\lim_n \mathcal{F}(L_n) = \mathcal{F}(X)$$

Prove the base case with Gärtner-Ellis, then "bootstrap" from it.

Conclusion (50 min)

1. Expectation is not necessarily the best for describing the relevant behaviors of a random variable.

Conclusion (50 min)

1. Expectation is not necessarily the best for describing the relevant behaviors of a random variable.
2. Minimizing the CVaR of risk, instead of the expectation, allows more prudent planning.

Conclusion (50 min)

1. Expectation is not necessarily the best for describing the relevant behaviors of a random variable.
2. Minimizing the CVaR of risk, instead of the expectation, allows more prudent planning.
3. There are remarkable generalizations to basic theorems of probability theory, once expectation is replaced with CVaR.