Quantitative Investments

Dale W.R. Rosenthal\textsuperscript{1}

1 June 2018

\textsuperscript{1}info@q36llc.com
Recall

Last lecture we discussed risk & returns.

- Randomness;
- Statistical metrics for returns;
- Risk premia; and,
- How actual returns differ from normality.

Today we will talk about risk measures.
Warning

⚠️ 注意！ The material in this chapter is more challenging than most chapters. Proceed carefully.
Risk Measures

Chapter 8, *A Quantitative Primer on Investments with R*
This part discusses risk measures, specifically:

- Value-at-risk;
- Coherent risk measures; and,
- Risk-based performance measures.
If we are going to measure risk, we need to be humble.

Rule #1: models are generalizations; plot summaries.
  - “All models are wrong, some models are useful.” — George Box

Every model has assumptions; ask “what if those don’t hold?”

Many models approximate; ask “when is that inaccurate?”

Some claim modeling is pointless: models ≠ reality.
  - Nouveau Ludditism. Useful models should not be perfect.
  - Don’t let the perfect be the enemy of the good.
Seems reasonable we should consider $\sigma, \gamma, \kappa$.

But really, what is risk? Can we get a simple definition?

One idea: 5-th (or $N$-th) percentile of returns.

- This (scaled by notional) is the 5% value at risk (VaR).
- 5% of the time, we expect a loss of at least 5%-VaR.

For a normal distribution, 5%-VaR is $1.64 \sigma$ below the mean.

In general: 5%-VaR is the best of worst 5% of days.
Example: Suppose we can invest $1 in two bonds: A and B.
- Each defaults independently with probability 0.04.

Default probabilities for $\frac{A+B}{2}$ portfolio (half A, half B):
- $P(\text{no defaults}) = 0.9216$;
- $P(\text{A&B default}) = 0.0016$.
- Thus $P(\text{one bond defaults}) = 0.0768$. 
Value at Risk?

- \( E(\text{value}|A \text{ or } B) = E(\text{value}|\frac{A+B}{2}) = 0.96. \)
- \( \text{Var}(\text{value}|A \text{ or } B) = 0.0384; \text{Var}(\text{value}|\frac{A+B}{2}) = 0.0192. \)
- \( P(\text{total loss}|A \text{ or } B) = 0.04 \Rightarrow P(\text{total loss}|\frac{A+B}{2}) = 0.0016 \)
  - So payoff variability, \( P(\text{total loss}) \) is lower for \( \frac{A+B}{2} \) portfolio.
- 5%-VaR for $1 in A or B = $0; for \( \frac{A+B}{2} \) portfolio = $0.50.
- VaR is deeply flawed!
  - VaR tells us investing in one bond is less risky.
  - VaR leads us into more concentrated portfolios than is wise.
- Basel II accords set international bank capital standards. . .
  - . . .capital adequacy levels were set based on VaR.
A better risk measure: *semideviation*.

Idea: “standard deviation” using only returns below the mean.

\[
\theta_p := \sqrt{E((r_t - \mu)^2 | r_t < \mu)} \tag{1}
\]

\[
\hat{\theta}_p = \sqrt{\frac{1}{n_-} \sum_{t=1}^{n} (r_t - \mu)^2 I(r_t < \mu)} \tag{2}
\]

where \(n_- = \sum_{t=1}^{n} I(r_t < \mu)\) \tag{3}

Similar and informative: *maximum drawdown*.

Largest peak–trough decline of portfolio over a time period.

While these are better, they are not ideal.
Artzner, Delbaen, Eber, and Heath (1997) rethought risk.

- Proposed properties of sensible (coherent) risk measures.
- Cash, diversification, dominance help; risk \( \geq 0 \); linearity.

The “Gang of Four” also proposed a coherent risk measure.

- **Expected shortfall (ES)** aka **Conditional value at risk (CVaR)**.

ES: expected return when at/below a given %ile.

Thus 5%-ES := \( E(r \cdot \text{Not}l | r \cdot \text{Not}l \leq 5\%-\text{VaR}) \).

Idea: Don’t just give the best of the worst days;

Rather: Give the average returns on the “worst” days.
A picture makes the difference clear.

Best of the worst days is not a good risk measure.
Now we can concoct metrics for risk-adjusted return.

The *Sharpe ratio* was discovered by Sharpe (1966):

\[
S_p := \frac{E(r_p - r_f)}{\sigma_r}; \quad \hat{S}_p = \frac{\bar{r}_p - \bar{r}_f}{\hat{\sigma}_p} = \frac{\hat{\mu}_p - \hat{\mu}_f}{\hat{\sigma}_p}. \tag{4}
\]

Sortino and Price (1994) discovered the *Sortino ratio*:

\[
S_{op} := \frac{E(r - r_f)}{\theta_p}; \quad \hat{S}_{op} = \frac{\bar{r}_p - \bar{r}_f}{\hat{\theta}_p} = \frac{\hat{\mu}_p - \hat{\mu}_f}{\hat{\theta}_p}. \tag{5}
\]

*Conditional Sharpe ratio* uses ES in denominator.

\[
CS_p := \frac{E(r - r_f)}{ES_p}; \quad \hat{CS}_p = \frac{\bar{r}_p - \bar{r}_f}{\hat{ES}_p} = \frac{\hat{\mu}_p - \hat{\mu}_f}{\hat{ES}_p}. \tag{6}
\]

“\(S_P \approx S_{op}\) so who cares?” Nonsense!
- Rarely run out of cash... but really care when you do.
- We care about the *unusual* risky times.
Helpful-yet-difficult approach: estimate the return density.
The benefit of this is it lets us simulate unseen scenarios.
The difficulties: shape? tail behavior? conditional stationarity?
A few common approaches:
1. Use the empirical (historical) distribution;
2. Estimate a parametric distribution;
3. Approximate the density with a kernel density estimator; or,
4. Use measured moments in an asymptotic expansion.
Which is best? Not clear; we will look at all of them.
Common approaches use one of two simple methods:

- use empirical distribution (aka *histogram*); or,
- estimate parameters of a pre-specified distribution.

Histogram estimator yields no insight on unseen scenarios.

Parametric approach assumes a distribution (and thus tail behavior).

Advantage: These approaches are easy and transparent.
Kernel Density Estimation

- A *kernel* function integrates to 1 and is even.
- Use kernel to find weighted average of nearby data.
- Typically parametrize a kernel $K$ with some bandwidth $h$:
  \[
  K_h(x) = \frac{1}{h} K(x/h). \tag{7}
  \]

- Nadaraya-Watson kernel density estimator is most common:
  \[
  \hat{Y} = E(Y|X = x) = \frac{\sum_{t=1}^{n} Y_t K_h(x - X_t)}{\sum_{t=1}^{n} K_h(x - X_t)}, \tag{8}
  \]
  \[
  \hat{f} = E(f|X = x) = \frac{1}{nh} \sum_{t=1}^{n} K \left( \frac{x - X_t}{h} \right), \tag{9}
  \]
  for $n$ observations of data $X_t$.
- Typical kernels: uniform (⇒ histogram), triangle, Gaussian.
Kernel Density Estimation: Examples

- Gaussian kernel density estimate for various bandwidths $h$.
- Generating density (– –). Sheather-Jones suggested $h$: lower right
- Problem: Tail behavior is still only Gaussian.
Asymptotic expansions: approximating series using moments.

Typically use either Edgeworth or Cornish-Fisher expansion.

Edgeworth expansion: Taylor series of cumulant generating fcn.

Typically done using a base normal distribution.

\[
\hat{f}_r(z) = \frac{\phi(z)}{\sigma} \left[ 1 + \frac{\gamma(z^3 - 3z)}{6} + \frac{(\kappa - 3)(z^4 - 6z^2 + 3)}{24} \right.
\]
\[
\left. + \frac{\gamma^2(z^6 - 15z^4 + 45z^2 - 15)}{72} \right] + O(n^{-3/2})
\]

where \( \phi(z) \) = standard normal pdf, \( z = (r - \mu)/\sigma \).
Example Edgeworth expansions for data from unimodal, bimodal pdfs:
Asymptotic Expansions: Cornish-Fisher

- **Cornish-Fisher expansion**: series for inverse cdf.
- Also often done using a base normal distribution.

\[
\hat{F}^{-1}_r(p) = \mu + \sigma \left[ q_N + \gamma \frac{q_N^2 - 1}{6} + (\kappa - 3) \frac{q_N^3 - 3q_N}{24} 
- \gamma^2 \frac{2q_N^3 - 5q_N}{36} \right] + O(n^{-3/2}),
\]

(11)

where \(q_N = \Phi^{-1}(p)\), inverse of standard normal cdf.
We can also estimate risk using *extreme value theory*. Idea: find density for largest point \( X_{(n)} \), \( f(n) \):

\[
f(n)(x) = nP(x_j \leq x_i = x \forall j \neq i) f(x_i = x) = nF(x)^{n-1}f(x).
\]  

Fisher-Tippett: \( F_n \) converges in distribution to \( H \).

\( H \) (maximal domain of attraction) may have three forms:

1. Thin-tailed returns \( F \): \( H \sim \text{Weibull} \); \( H(x) = e^{-(−x)^\alpha} I_{x<0} \);
2. Medium-tailed returns \( F \): \( H \sim \text{Gumbel} \); \( H(x) = e^{-e^{-x}} \); and,
3. Heavy-tailed returns \( F \): \( H \sim \text{Fréchet} \); \( H(x) = e^{-x^{-\alpha}} I_{x>0} \).

Can then use appropriate form to efficiently simulate maximum loss.
Often work with generalized extreme value distribution:

\[
H_\xi(x) = \begin{cases} 
  e^{-(1+\xi x)^{-1/\xi}} & \xi \neq 0, 1+\xi x > 0 \\
  e^{-e^{-x}} & \xi = 0,
\end{cases}
\]

(13)

Sometime use similar generalized Pareto distribution:

\[
G_\xi(x) = \begin{cases} 
  1 - (1 + \xi x)^{-1/\xi} & \xi \neq 0, 1+\xi x > 0 \\
  1 - e^{-x} & \xi = 0.
\end{cases}
\]

(14)

\[G_\xi = 1 + \log(H_\xi(x)) \approx H(x)\] (but easier to use).

\[\xi \mapsto H: \text{Fréchet (}\xi < 0\text{), Gumbel (}\xi = 0\text{), Weibull (}\xi > 0\text{).}\]

How to estimate \(\xi\)? ML or points over threshold.

But game theory still matters: hard to protect against malice.
Density estimates then yield risk measures (e.g. VaR, ES).

- **Empirical**: sort for $\alpha - \text{VaR}$, average for $\alpha - \text{ES}$.
- **Parametric**: find quantile (VaR) or tail integral (ES).
- **Kernel Density**: Numerically integrate to get VaR, ES.
- **Edgeworth/Cornish-Fisher**: closed-form, numeric, or use package.
- **EVT**: Use R packages to estimate GEV or GPD distributions.

Results will vary. Which to use? Can average or take worst.
Why Convergence Matters

- It now makes sense to relate convergence to risk management.
- Some strategies have different convergence due to economics.
- Compare convergence in distribution, probability, almost surely:

![Graph showing convergence in probability and almost surely](image)

- Convergence in probability: worst-loss days become less frequent.
- Convergence almost surely: worst-loss days also get better.
Finally, must consider other risks which are not easily quantified.
- Model error: risk that model is wrong/mis-estimated.
- Also possible: risk varies across time, with state of economy.
- Operational risks: failures in business processes, fraud.
- Physical risks: bugs, threats, break-ins, unrest.
- Regulatory risks: politicians may regulate, threaten, expropriate.
- Risks may even be impossible to quantify/measure.
  - *Knightian uncertainty*: may not even know form.
  - Unknown unknowns; differs from known unknowns.
The Road Ahead

We have covered risk measures; on to diversification next time!

- Measuring: Diversification and Statistical Modeling;
- Valuation: Fixed Income, Yield Curves, Equity Valuation;
- Valuation II: Factor Models, Microfoundations, Global Investing, FX;
- Risk Alleviation: Futures, Options, Credit, Structured Products; and,
- All Together Now: Active Portfolios, Investment Firms, Crises.