Risk, Choquet Portfolios and Quantile Regression

Roger Koenker

University of Illinois, Urbana-Champaign

5th RMetrics Workshop, Meielisalp: 28 June 2011

Joint work with Gib Bassett (UIC) and Gregory Kordas (Athens)

Risk as Pessimism?

In expected utility theory risk is entirely an attribute of the utility function:

Risk Neutrality	$\Rightarrow$	$u(x) \sim$ affine
Risk Aversion	$\Rightarrow$	$u(x) \sim$ concave
Risk Attraction	$\Rightarrow$	$u(x) \sim$ convex

Risk as Pessimism?

In expected utility theory risk is entirely an attribute of the utility function:

Risk Neutrality	$\Rightarrow$	$u(x) \sim$ affine
Risk Aversion	$\Rightarrow$	$u(x) \sim$ concave
Risk Attraction	$\Rightarrow$	$u(x) \sim$ convex

Locally, the risk premium, i.e. the amount one is willing to pay to accept a zero mean risk, X, is

$$\pi(w, X) = \frac{1}{2} A(w) V(X)$$

where $A(w) = -u''(w)/u'(w)$ is the Arrow-Pratt coefficient of absolute risk aversion and $V(X)$ is the variance of X.

Risk as Pessimism?

In expected utility theory risk is entirely an attribute of the utility function:

Risk Neutrality	$\Rightarrow$	$u(x) \sim$ affine
Risk Aversion	$\Rightarrow$	$u(x) \sim$ concave
Risk Attraction	$\Rightarrow$	$u(x) \sim$ convex

Locally, the risk premium, i.e. the amount one is willing to pay to accept a zero mean risk, X, is

$$\pi(w, X) = \tfrac{1}{2}A(w)V(X)$$

where $A(w) = -u''(w)/u'(w)$ is the Arrow-Pratt coefficient of absolute risk aversion and $V(X)$ is the variance of X. Why is variance a reasonable measure of risk?

A Little Risk Aversion is a Dangerous Thing

Would you accept the gamble:

$$G_1 \qquad 50-50 \left\langle \begin{array}{l} \text{win } \$110 \\ \text{lose } \$100 \end{array} \right.$$

A Little Risk Aversion is a Dangerous Thing

Would you accept the gamble:

$$G_1 \qquad 50-50 \left\langle \begin{array}{l} \text{win } \$110 \\ \text{lose } \$100 \end{array} \right.$$

Suppose you say "no", then what about the gamble:

A Little Risk Aversion is a Dangerous Thing

Would you accept the gamble:

$$G_1 \qquad 50 - 50 \left\langle \begin{array}{c} \text{win } \$110 \\ \text{lose } \$100 \end{array} \right.$$

Suppose you say "no", then what about the gamble:

$$G_2 \qquad 50 - 50 \left\langle \begin{array}{c} \text{win } \$700,000 \\ \text{lose } \$1,000 \end{array} \right.$$

A Little Risk Aversion is a Dangerous Thing

Would you accept the gamble:

G_1 $\quad 50 - 50 \left\langle \begin{array}{c} \text{win } \$110 \\ \text{lose } \$100 \end{array} \right.$

Suppose you say "no", then what about the gamble:

G_2 $\quad 50 - 50 \left\langle \begin{array}{c} \text{win } \$700,000 \\ \text{lose } \$1,000 \end{array} \right.$

If you say "no" to G_1 for any initial wealth up to \$300,000, then you must also say "no" to G_2.

A Little Risk Aversion is a Dangerous Thing

Would you accept the gamble:

$$G_1 \qquad 50-50 \quad \left\langle \begin{array}{c} \text{win } \$110 \\ \text{lose } \$100 \end{array} \right.$$

Suppose you say "no", then what about the gamble:

$$G_2 \qquad 50-50 \quad \left\langle \begin{array}{c} \text{win } \$700,000 \\ \text{lose } \$1,000 \end{array} \right.$$

If you say "no" to G_1 for any initial wealth up to $300,000, then you must also say "no" to G_2.
Moral: A little local risk aversion over small gambles implies implausibly large risk aversion over large gambles. Reference: Rabin (2000)

Are Swiss Bicycle Messengers Risk Averse?

When Veloblitz and Flash bicycle messengers from Zurich were confronted with the bet:

$$50 - 50 \left\langle \begin{array}{l} \text{win 8 CHF} \\ \text{lose 5 CHF} \end{array} \right.$$

Are Swiss Bicycle Messengers Risk Averse?

When Veloblitz and Flash bicycle messengers from Zurich were confronted with the bet:

$$50 - 50 \left\langle \begin{array}{l} \text{win 8 CHF} \\ \text{lose 5 CHF} \end{array} \right.$$

More than half (54%) rejected the bet.
Reference: Fehr and Götte (2002)

Coherent Risk

Definition (Artzner, Delbaen, Eber and Heath (1999)) For real valued random variables $X \in \mathcal{X}$ on $(\Omega, \mathcal{A})$ a mapping $\rho : \mathcal{X} \rightarrow \mathcal{R}$ is called a coherent risk measure if,

1. Monotone: $X, Y \in \mathcal{X}$, with $X \leqslant Y \Rightarrow \rho(X) \geqslant \rho(Y)$.
2. Subadditive: $X, Y, X+Y \in \mathcal{X}, \Rightarrow \rho(X+Y) \leqslant \rho(X) + \rho(Y)$.
3. Linearly Homogeneous: For all $\lambda \geqslant 0$ and $X \in \mathcal{X}$, $\rho(\lambda X) = \lambda\rho(X)$.
4. Translation Invariant: For all $\lambda \in \mathcal{R}$ and $X \in \mathcal{X}$, $\rho(\lambda + X) = \rho(X) - \lambda$.

Coherent Risk

Definition (Artzner, Delbaen, Eber and Heath (1999)) For real valued random variables $X \in \mathcal{X}$ on $(\Omega, \mathcal{A})$ a mapping $\rho : \mathcal{X} \rightarrow \mathcal{R}$ is called a coherent risk measure if,

1. Monotone: $X, Y \in \mathcal{X}$, with $X \leqslant Y \Rightarrow \rho(X) \geqslant \rho(Y)$.
2. Subadditive: $X, Y, X + Y \in \mathcal{X}, \Rightarrow \rho(X + Y) \leqslant \rho(X) + \rho(Y)$.
3. Linearly Homogeneous: For all $\lambda \geqslant 0$ and $X \in \mathcal{X}$, $\rho(\lambda X) = \lambda\rho(X)$.
4. Translation Invariant: For all $\lambda \in \mathcal{R}$ and $X \in \mathcal{X}$, $\rho(\lambda + X) = \rho(X) - \lambda$.

Many conventional measures of risks including those based on standard deviation are ruled out by these requirements. So are quantile based measures like "value at risk."

Choquet α-Risk

The leading example of a coherent risk measure is

$$\rho_{\nu_\alpha}(X) = -\alpha^{-1} \int_0^\alpha F^{-1}(t)dt$$

Variants of this risk measure have been introduced under several names

Choquet α-Risk

The leading example of a coherent risk measure is

$$\rho_{\nu_\alpha}(X) = -\alpha^{-1} \int_0^\alpha F^{-1}(t)dt$$

Variants of this risk measure have been introduced under several names

- Expected shortfall (Acerbi and Tasche (2002))
- Conditional VaR (Rockafellar and Uryasev (2000))
- Tail conditional expectation (Artzner, et al (1999)).

Choquet α-Risk

The leading example of a coherent risk measure is

$$\rho_{\nu_\alpha}(X) = -\alpha^{-1} \int_0^\alpha F^{-1}(t)dt$$

Variants of this risk measure have been introduced under several names

- Expected shortfall (Acerbi and Tasche (2002))
- Conditional VaR (Rockafellar and Uryasev (2000))
- Tail conditional expectation (Artzner, et al (1999)).

Note that $\rho_{\nu_\alpha}(X) = -E_{\nu_\alpha,F}(X)$, so Choquet α-risk is just negative Choquet expected utility with the distortion function ν_α.

Pessimistic Risk Measures

Definition A risk measure ρ will be called pessimistic if, for some probability measure φ on [0, 1]

$$\rho(X) = \int_0^1 \rho_{\nu_\alpha}(X) d\varphi(\alpha)$$

Pessimistic Risk Measures

Definition A risk measure ρ will be called pessimistic if, for some probability measure φ on [0, 1]

$$\rho(X) = \int_0^1 \rho_{\nu_\alpha}(X) d\varphi(\alpha)$$

By Fubini

$$\begin{aligned}
\rho(X) &= -\int_0^1 \alpha^{-1} \int_0^\alpha F^{-1}(t) dt d\varphi(\alpha) \\
&= -\int_0^1 F^{-1}(t) \int_t^1 \alpha^{-1} d\varphi(\alpha) dt \\
&\equiv -\int_0^1 F^{-1}(t) d\nu(t)
\end{aligned}$$

Approximating General Pessimistic Risk Measures

We can approximate any pessimistic risk measure by taking

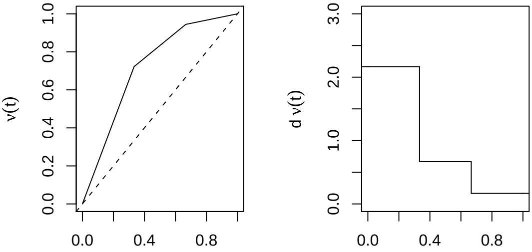
$$d\varphi(t) = \sum \varphi_i \delta_{\tau_i}(t)$$

where δ_τ denotes (Dirac) point mass 1 at τ. Then

$$\rho(X) = -\varphi_0 F^{-1}(0) - \int_0^1 F^{-1}(t)\gamma(t)dt$$

where $\gamma(t) = \sum \varphi_i \tau_i^{-1} I(t < \tau_i)$ and $\varphi_i > 0$, with $\sum \varphi_i = 1$.

An Example



$$d\varphi(t) = \frac{1}{2}\delta_{1/3}(t) + \frac{1}{3}\delta_{2/3}(t) + \frac{1}{6}\delta_1(t)$$

A Theorem

Theorem (Kusuoka (2001)) A regular risk measure is *coherent* in the sense of Artzner *et. al.* if and only if it is *pessimistic*.

A Theorem

Theorem (Kusuoka (2001)) A regular risk measure is *coherent* in the sense of Artzner *et. al.* if and only if it is *pessimistic*.

- Pessimistic Choquet risk measures correspond to *concave* ν, i.e., *monotone decreasing* $d\nu$.

A Theorem

Theorem (Kusuoka (2001)) A regular risk measure is *coherent* in the sense of Artzner *et. al.* if and only if it is *pessimistic*.

- Pessimistic Choquet risk measures correspond to *concave* ν, i.e., *monotone decreasing* $d\nu$.
- Probability assessments are distorted to accentuate the probability of the least favorable events.

A Theorem

Theorem (Kusuoka (2001)) A regular risk measure is *coherent* in the sense of Artzner *et. al.* if and only if it is *pessimistic*.

- Pessimistic Choquet risk measures correspond to *concave* ν, i.e., *monotone decreasing* $d\nu$.
- Probability assessments are distorted to accentuate the probability of the least favorable events.
- The crucial coherence requirement is subadditivity, or submodularity, or 2-alternatingness in the terminology of Choquet capacities.

An Example

Samuelson (1963) describes asking a colleague at lunch whether he would be willing to make a

$$50 - 50 \text{ bet} \quad \left\langle \begin{array}{l} \text{win } 200 \\ \text{lose } 100 \end{array} \right.$$

An Example

Samuelson (1963) describes asking a colleague at lunch whether he would be willing to make a

$$50 - 50 \text{ bet} \quad \left\langle \begin{array}{l} \text{win } 200 \\ \text{lose } 100 \end{array} \right.$$

The colleague (later revealed to be E. Cary Brown) responded

"no, but I *would* be willing to make 100 such bets."

An Example

Samuelson (1963) describes asking a colleague at lunch whether he would be willing to make a

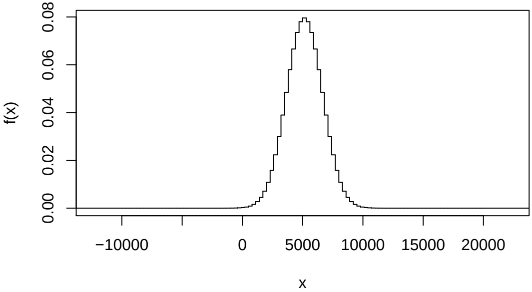
$$50 - 50 \text{ bet} \quad \left\langle \begin{array}{l} \text{win } 200 \\ \text{lose } 100 \end{array} \right.$$

The colleague (later revealed to be E. Cary Brown) responded

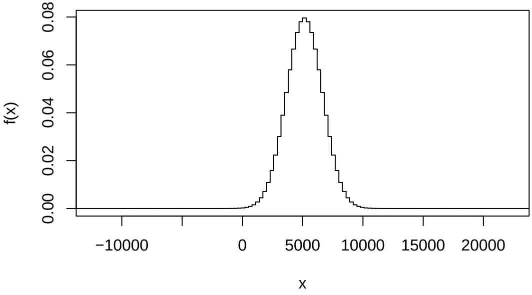
"no, but I *would* be willing to make 100 such bets."

This response has been interpreted not only as reflecting a basic confusion about how to maximize expected utility but also as a fundamental misunderstanding of the law of large numbers.

Payoff Density of 100 Samuelson trials



Payoff Density of 100 Samuelson trials



Odds of losing money on the 100 trial bet is 1 chance in 2300.

Was Brown really irrational?

Suppose, for the sake of simplicity that

$$d\varphi(t) = \tfrac{1}{2}\delta_{1/2}(t) + \tfrac{1}{2}\delta_1(t)$$

so for one Samuelson coin flip we have the unfavorable evaluation,

$$E_{\nu,F}(X) = \tfrac{1}{2}(-100) + \tfrac{1}{2}(50) = -25$$

but for $S = \sum_{i=1}^{100} X_i \sim \mathcal{B}in(.5, 100)$ we have the favorable evaluation,

$$\begin{aligned} E_{\nu,F}(S) &= \tfrac{1}{2}2\int_0^{1/2} F_S^{-1}(t)dt + \tfrac{1}{2}(5000) \\ &= 1704.11 + 2500 \\ &= 4204.11 \end{aligned}$$

How to be Pessimistic

Theorem Let X be a real-valued random variable with $EX = \mu < \infty$, and $\rho_\alpha(u) = u(\alpha - I(u < 0))$. Then

$$\min_{\xi \in \mathcal{R}} E\rho_\alpha(X - \xi) = \alpha\mu + \rho_{\nu_\alpha}(X)$$

So α risk can be estimated by the sample analogue

$$\hat{\rho}_{\nu_\alpha}(x) = (n\alpha)^{-1} \min_\xi \sum \rho_\alpha(x_i - \xi) - \hat{\mu}_n$$

How to be Pessimistic

Theorem Let X be a real-valued random variable with $EX = \mu < \infty$, and $\rho_\alpha(u) = u(\alpha - I(u < 0))$. Then

$$\min_{\xi \in \mathcal{R}} E\rho_\alpha(X - \xi) = \alpha\mu + \rho_{\nu_\alpha}(X)$$

So α risk can be estimated by the sample analogue

$$\hat{\rho}_{\nu_\alpha}(x) = (n\alpha)^{-1} \min_{\xi} \sum \rho_\alpha(x_i - \xi) - \hat{\mu}_n$$

 I knew it! Eventually everything looks like quantile regression to this guy!

Pessimistic Portfolios

Now let $X = (X_1, \ldots, X_p)$ denote a vector of potential portfolio asset returns and $Y = X^\top \pi$, the returns on the portfolio with weights π. Consider

$$\min_{\pi} \rho_{\nu_\alpha}(Y) - \lambda \mu(Y)$$

Minimize α-risk subject to a constraint on mean return.

Pessimistic Portfolios

Now let $X = (X_1, \ldots, X_p)$ denote a vector of potential portfolio asset returns and $Y = X^\top \pi$, the returns on the portfolio with weights π. Consider

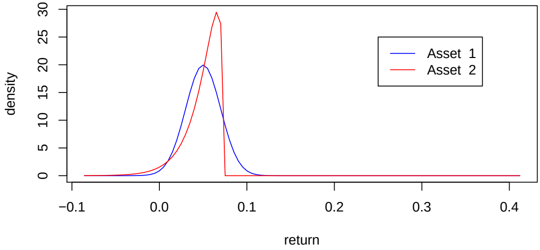
$$\min_{\pi} \rho_{\nu_\alpha}(Y) - \lambda \mu(Y)$$

Minimize α-risk subject to a constraint on mean return.
This problem can be formulated as a linear quantile regression problem

$$\min_{(\beta,\xi)\in\mathcal{R}^p} \sum_{i=1}^{n} \rho_\alpha(x_{i1} - \sum_{j=2}^{p}(x_{i1} - x_{ij})\beta_j - \xi) \quad \text{s.t.} \quad \bar{x}^\top \pi(\beta) = \mu_0,$$

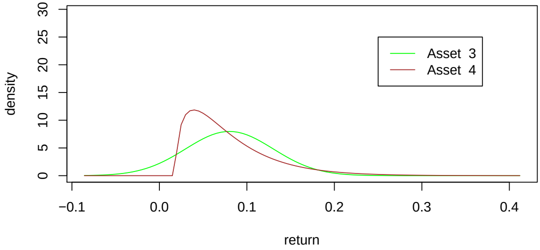
where $\pi(\beta) = (1 - \sum_{j=2}^{p} \beta_j, \beta^\top)^\top$.

An Example



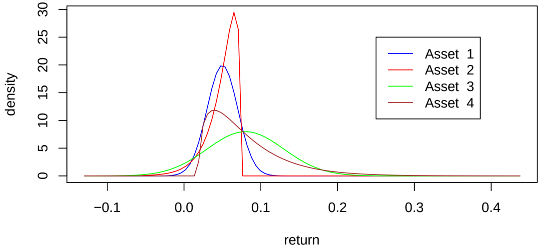
Two asset return densities with identical mean and variance.

An Example



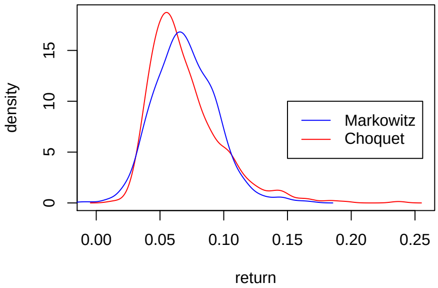
Two more asset return densities with identical mean and variance.

An Example

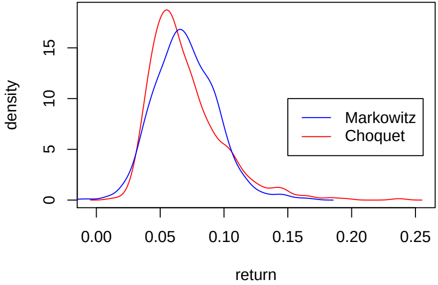


Two pairs of asset return densities with identical mean and variance.

Optimal Choquet and Markowitz Portfolio Returns

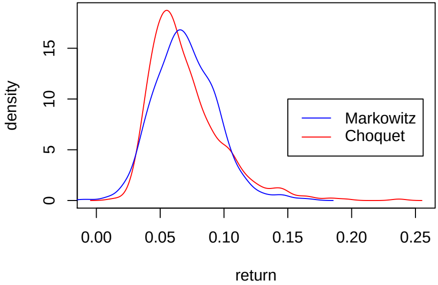


Optimal Choquet and Markowitz Portfolio Returns

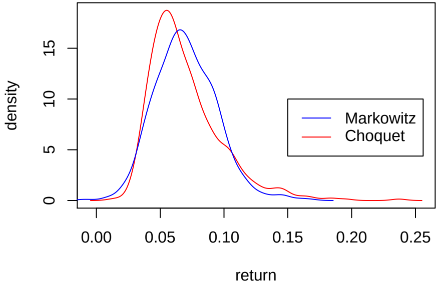


Markowitz portfolio minimizes the standard deviation of returns subject to mean return $\mu = .07$. The Choquet portfolio minimizes Choquet risk (for $\alpha = .10$) subject to earning the same mean return. The Choquet portfolio has better performance in both tails than mean-variance Markowitz portfolio.

Optimal Choquet and Markowitz Portfolio Returns



Optimal Choquet and Markowitz Portfolio Returns



Now, the Markowitz portfolio minimizes the standard deviation of returns subject to mean return $\mu = .07$. The Choquet portfolio maximizes expected return subject to achieving the same Choquet risk (for $\alpha = .10$) as the Markowitz portfolio. Choquet portfolio has expected return $\mu = .08$ a full percentage point higher than the Markowitz portfolio.

A Unified Theory of Pessimism

Any pessimistic risk measure may be approximated by

$$\rho_\nu(X) = \sum_{k=1}^{m} \varphi_k \rho_{\nu_{\alpha_k}}(X)$$

where $\varphi_k > 0$ for $k = 1, 2, \ldots, m$ and $\sum \varphi_k = 1$.
Portfolio weights can be estimated for these risk measures by solving linear programs that are weighted sums of quantile regression problems:

$$\min_{(\beta,\xi)\in\mathcal{R}^p} \sum_{k=1}^{m} \sum_{i=1}^{n} \nu_k \rho_{\alpha_k}(x_{i1} - \sum_{j=2}^{p} (x_{i1} - x_{ij})\beta_j - \xi_k) \quad s.t. \quad \bar{x}^\top \pi(\beta) = \mu_0,$$

A Unified Theory of Pessimism

Any pessimistic risk measure may be approximated by

$$\rho_\nu(X) = \sum_{k=1}^{m} \varphi_k \rho_{\nu_{\alpha_k}}(X)$$

where $\varphi_k > 0$ for $k = 1, 2, ..., m$ and $\sum \varphi_k = 1$.
Portfolio weights can be estimated for these risk measures by solving linear programs that are weighted sums of quantile regression problems:

$$\min_{(\beta,\xi)\in\mathcal{R}^p} \sum_{k=1}^{m} \sum_{i=1}^{n} \nu_k \rho_{\alpha_k}(x_{i1} - \sum_{j=2}^{p} (x_{i1} - x_{ij})\beta_j - \xi_k) \quad \text{s.t.} \quad \bar{x}^\top \pi(\beta) = \mu_0,$$

Software in R is available on from my web pages.

Conclusions

- Expected Utility is unsatisfactory both as a positive, i.e., descriptive, theory of behavior and as a normative guide to behavior.

Conclusions

- Expected Utility is unsatisfactory both as a positive, i.e., descriptive, theory of behavior and as a normative guide to behavior.
- Choquet (non-additive, rank dependent) expected utility provides a simple, tractable alternative.

Conclusions

- Expected Utility is unsatisfactory both as a positive, i.e., descriptive, theory of behavior and as a normative guide to behavior.
- Choquet (non-additive, rank dependent) expected utility provides a simple, tractable alternative.
- Mean-variance Portfolio allocation is also unsatisfactory since it relies on unpalatable assumptions of Gaussian returns, or quadratic utility.

Conclusions

- Expected Utility is unsatisfactory both as a positive, i.e., descriptive, theory of behavior and as a normative guide to behavior.
- Choquet (non-additive, rank dependent) expected utility provides a simple, tractable alternative.
- Mean-variance Portfolio allocation is also unsatisfactory since it relies on unpalatable assumptions of Gaussian returns, or quadratic utility.
- Choquet portfolio optimization can be formulated as a quantile regression problem thus providing an attractive practical alternative to the dominant mean-variance approach of Markowitz (1952).