Pessimistic Portfolio Allocation And Choquet Expected Utility

Roger Koenker
University of Illinois at Urbana-Champaign

La Roche-en-Ardennes: September, 2005

Joint work with Gib Bassett (UIC) and Gregory Kordas (U. Penn).

Outline

- Choquet Expected Utility

  Is there a useful role for pessimism in economic decision theory?

- Choquet Risk

  Is there a pessimistic theory of risk

- Choquet Portfolios

  How to be pessimistic?

St. Petersburg Paradox

What would you be willing to pay to play the game:

$$G = \{\text{pay: } p, \quad \text{win: } 2^n \text{ with probability } 2^{-n}, \quad n = 1, 2, ...\}$$

Bernoulli observed that even though the expected payoff was infinite, the gambler who maximized logarithmic utility would pay only a finite value to play. For example, given initial wealth 100,000 Roubles, our gambler would pay only 17 Roubles and 55 kopecks.

Expected Utility

To decide between two real valued gambles

$$X \sim F \quad \text{and} \quad Y \sim G$$

we choose X over Y if

$$Eu(X) = \int u(x)dF(x) \geq \int u(y)dG(y) = Eu(Y)$$

Bruno de Finetti 1906–1985

On Axiomatics

Suppose we have acts $P, Q, R, \ldots$ in a space $\mathcal{P}$, which admits enough convex structure to allow us to consider mixtures,

$$\alpha P + (1-\alpha)Q \in \mathcal{P} \quad \alpha \in (0,1)$$

Think of P, Q, R as probability measures on some underlying outcome/event space, $\mathcal{X}$.

Or better, view P, Q, R as acts mapping a space $\mathcal{S}$ of soon-to-be-revealed "states of nature" to the space of probability measures on the outcome space, $\mathcal{X}$.

Theorem(von-Neumann-Morgenstern) Suppose we have a preference relation $\{\succeq, \succ, \sim\}$ on $\mathcal{P}$ satisfying the axioms:

(A.1) (weak order) For all $P, Q, R \in \mathcal{P}$, $P \succeq Q$ or $Q \succeq P$, and $P \succeq Q$ and $Q \succeq P \Rightarrow P \succeq R$,

(A.2) (independence) For all $P, Q, R \in \mathcal{P}$ and $\alpha \in (0,1)$, then $P \succ Q \Rightarrow \alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)R$,

(A.3) (continuity) For all $P, Q, R \in \mathcal{P}$, if $P \succ Q$ and $Q \succ R$, then there exist α and $\beta \in (0,1)$, such that, $\alpha P + (1-\alpha)R \succ \beta Q(1-\beta)R$.

Then there exists a linear function u on $\mathcal{P}$ such that for all $P, Q \in \mathcal{P}$, $P \succ Q$ if and only if $u(P) > u(Q)$.

Weakening the Independence Axiom

The independence axiom seems quite innocuous, but it is extremely powerful. We will consider a weaker form of independence due to Schmeidler (1989).

(A.2') (comonotonic independence) For all pairwise comonotonic $P, Q, R \in \mathcal{P}$ and $\alpha \in (0,1)$ $P \succ Q \Rightarrow \alpha P + (1-\alpha)R \succ \alpha Q + (1-\alpha)R$,

Definition Two acts P and Q in $\mathcal{P}$ are comonotonic, or similarly ordered, if for no s and t in $\mathcal{S}$,

$$P(\{t\}) \succ P(\{s\}) \quad \text{and} \quad Q(\{s\}) \succ Q(\{t\}).$$

"If P is better in state t than state s, then Q is also better in t than s."

On Comonotonicity

Definition The two functions $X, Y : \Omega \to \mathbb{R}$ are comonotonic if there exists a third function $Z : \Omega \to \Re$ and increasing functions f and g such that $X = f(Z)$ and $Y = g(Z)$.

From our point of view the crucial property of comonotonic random variables is the behavior of quantile functions of their sums. For comonotonic random variables X, Y, we have

$$F_{X+Y}^{-1}(u) = F_X^{-1}(u) + F_Y^{-1}(u)$$

By comonotonicity we have a $U \sim U[0,1]$ such that $Z = g(U) = F_X^{-1}(U) + F_Y^{-1}(U)$ where g is left continuous and increasing, so by monotone invariance, $F_{g(U)}^{-1} = g \circ F_U^{-1} = F_X^{-1} + F_Y^{-1}$. Comonotonic random variables are maximally dependent $a\ la$ Fréchet

$$F_{X,Y}(x, y) = \min\{F_X(x), F_Y(y)\}.$$

Choquet Expected Utility

Among the many proposals offered to extend expected utility theory the most attractive (to us) replaces

$$E_F u(X) = \int_0^1 u(F^{-1}(t))dt \geq \int_0^1 u(G^{-1}(t))dt = E_G u(Y)$$

with

$$E_{\nu,F} u(X) = \int_0^1 u(F^{-1}(t))d\nu(t) \geq \int_0^1 u(G^{-1}(t))d\nu(t) = E_{\nu,G} u(Y)$$

The measure ν permits distortion of the probability assessments after ordering the outcomes. This rank dependent form of expected utility has been pioneered by Quiggin (1981), Schmeidler (1989), Wakker (1989) and Dennenberg (1990).

Pessimism

By relaxing the independence axiom we obtain a larger class of preferences representable as Choquet capacities and introducing pessimism.

The simplest form of Choquet expected utility is based on the "distortion"

$$\nu_\alpha(t) = \min\{t/\alpha, 1\}$$

so

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

This exaggerates the probability of the proportion α of least favorable events, and totally discounts the probability of the $1 - \alpha$ most favorable events. Expect the worst – and you won't be disappointed.

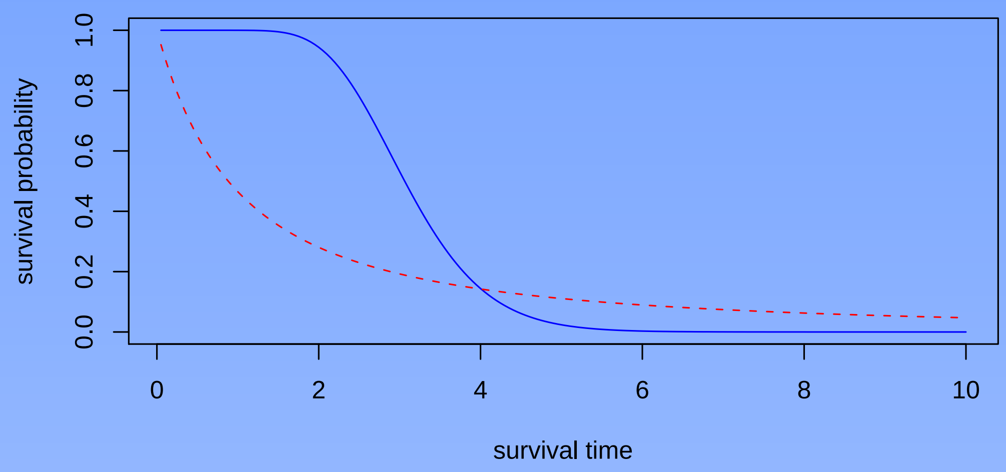
Savage on Pessimism

I have, at least once heard it objected against the personalistic view of probability that, according to that view, two people might be of different opinions, according as one is pessimistic and the other optimistic. I am not sure what position I would take in abstract discussion of whether that alleged property of personalistic views would be objectionable,

Savage on Pessimism

I have, at least once heard it objected against the personalistic view of probability that, according to that view, two people might be of different opinions, according as one is pessimistic and the other optimistic. I am not sure what position I would take in abstract discussion of whether that alleged property of personalistic views would be objectionable, but I think it is clear from the formal definition of qualitative probability that the particular personalistic view sponsored here does not leave room for optimism and pessimism, however these traits may be interpreted, to play any role in the person's judgement of probabilities. (Savage(1954), p. 68)

Pessimistic Medical Decision Making?



Survival Functions for a hypothetical medical treatment: The Lehmann quantile treatment effect is the horizontal distance between the survival curves. In this example consideration of the mean treatment effect would slightly favor the (dotted) treatment curve, but the pessimistic patient might favor the (solid) placebo curve.

What is Risk?

In the expected utility theory risk is in effect 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. The variance of X?

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:

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

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 Bicycle Messengers Risk Averse?

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

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

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 $\varrho : \mathcal{X} \to \mathcal{R}$ is called a coherent risk measure if,

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

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

Choquet α-Risk

The leading example of a coherent risk measure is

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

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

Pessimistic Risk Measures

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

$$\varrho(X) = \int_0^1 \varrho_{\nu_\alpha}(X) d\varphi(\alpha)$$

By Fubini

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

Approximating General Risk Measures

We can approximate general pessimistic risk measures by taking

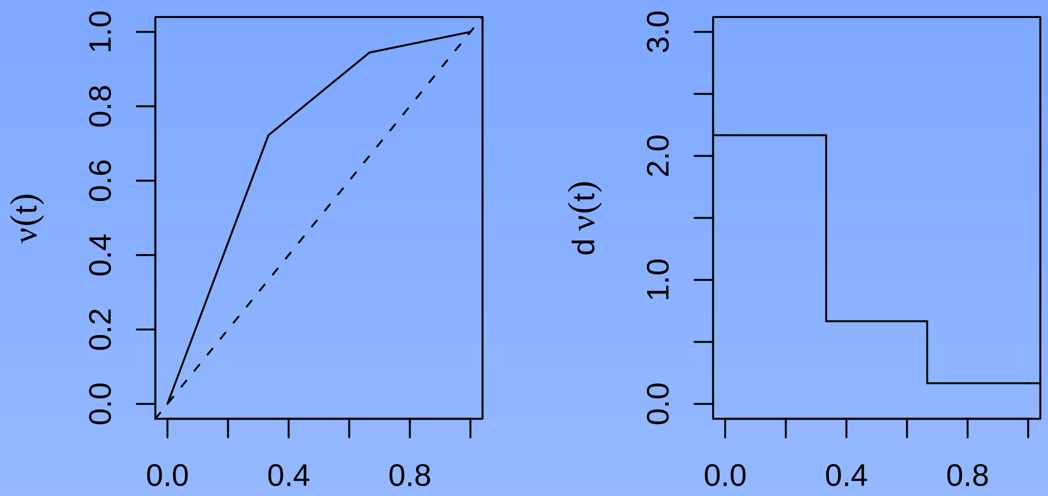
$$d\varphi(\alpha) = \sum \varphi_i \delta_{\tau_i}(\alpha) dt$$

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

$$\varrho(X) = -\int_0^1 F^{-1}(t) d\nu(t)$$

where $d\nu(t) = \sum \varphi_i \tau_i^{-1} I(t < \tau_i)$ and $\varphi_i > 0$, with $\sum \varphi_i = 1$, this is a piecewise constant density function.

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* 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

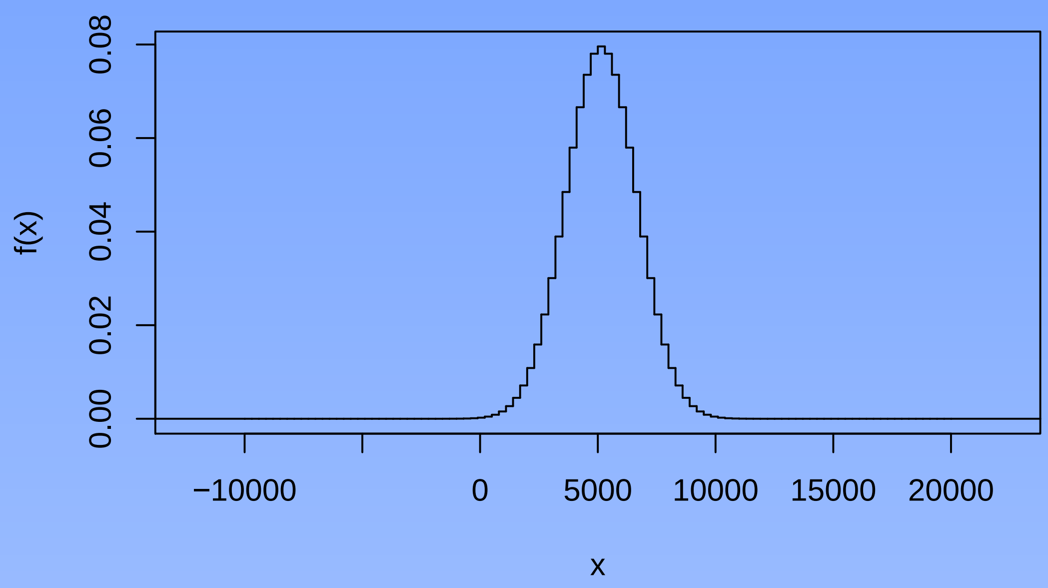
$$50 - 50 \text{ bet} \quad \left\langle \begin{array}{l} \text{win } 200 \\ \text{lose } 100 \end{array} \right.$$

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



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

Was Samuelson's colleague 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

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

So α risk can be estimated by the sample analogue

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


It was inevitable: 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

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

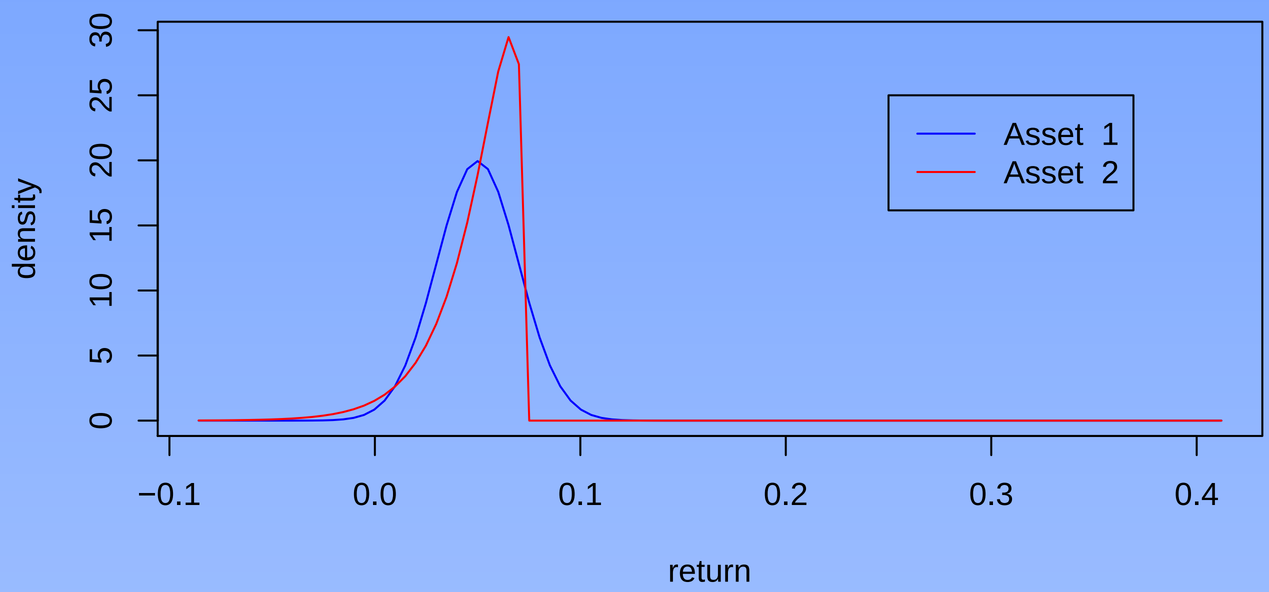
Minimize α-risk subject to a constraint on mean return.

This problem can be formulated as a quantile regression problem

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

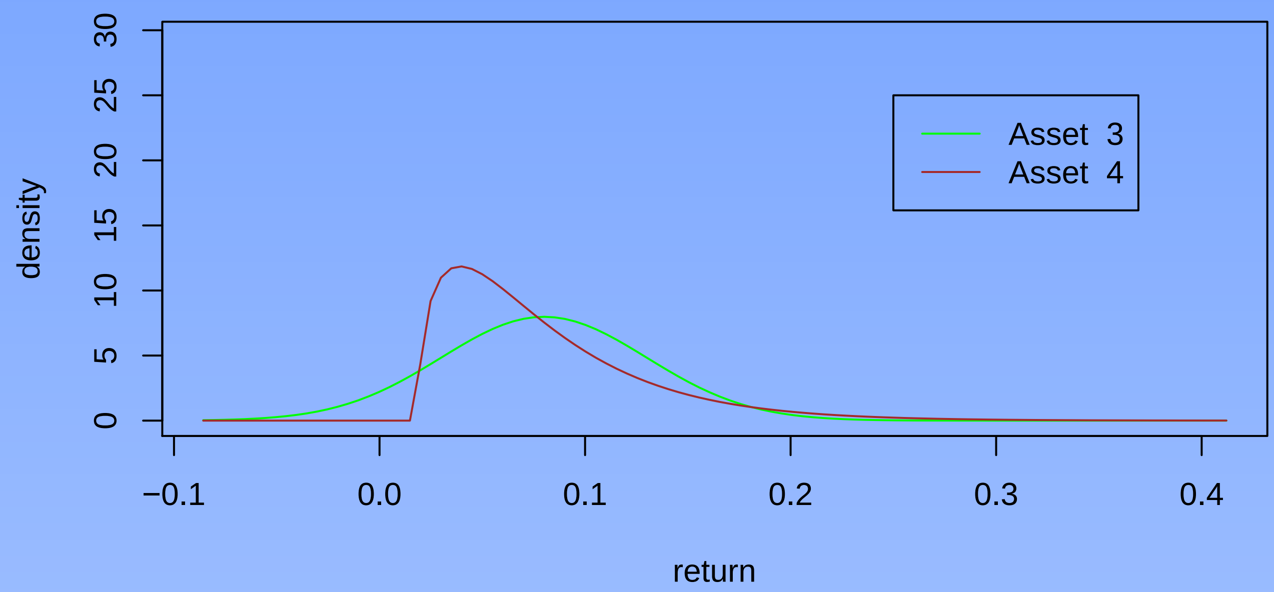
where $\pi(\beta) = (1 - \sum_{j=2}^p \beta_j, \beta^\top)^\top$.

An Example



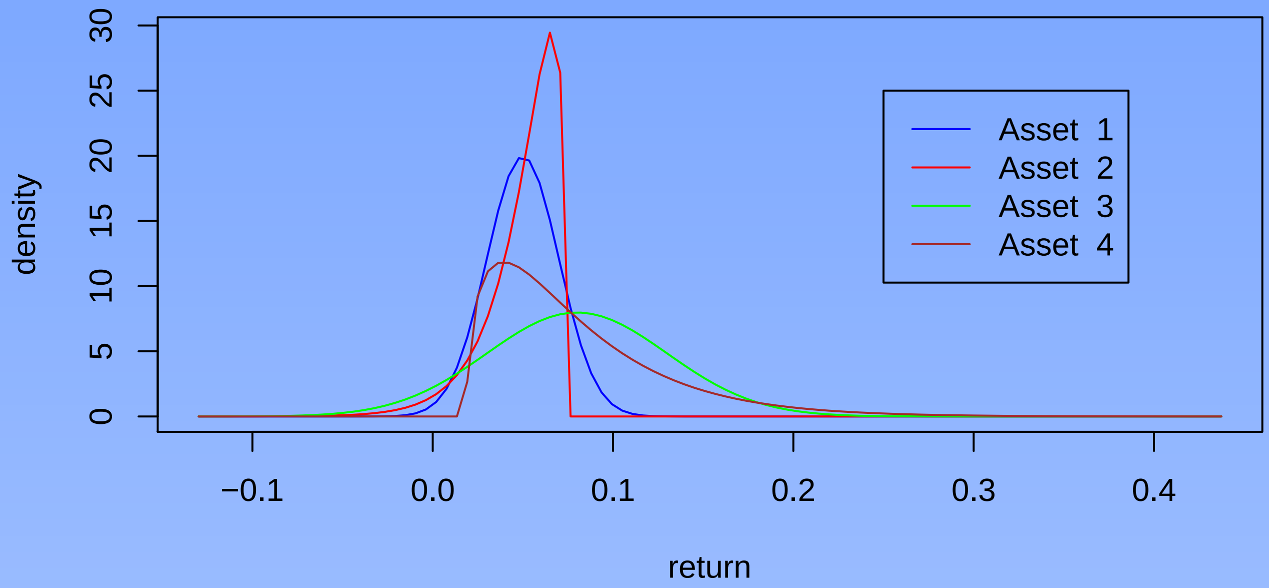
Two asset return densities with identical mean and variance.

An Example



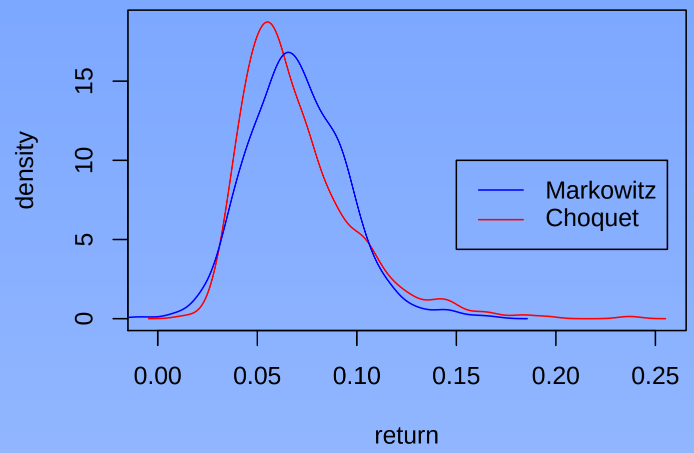
Two more asset return densities with identical mean and variance.

An Example



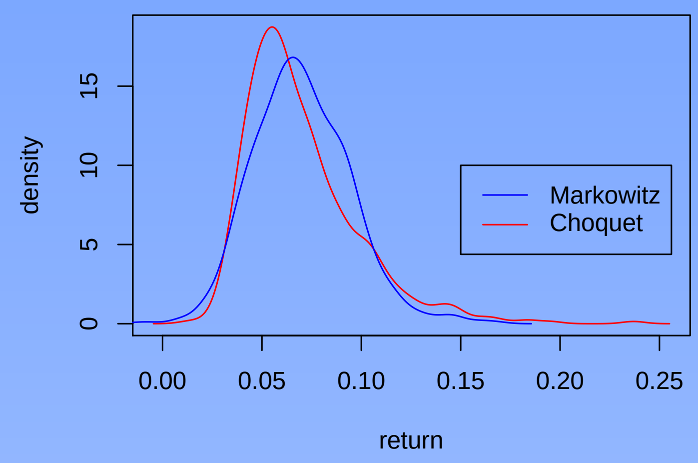
Two pairs of asset return densities with identical mean and variance.

Optimal Choquet and Markowitz Portfolio Returns



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



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 $\mu = .0768$, 68 basis points higher than the Markowitz portfolio.

A Unified Theory of Pessimism

Any pessimistic risk measure may be approximated by

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

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.

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

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.

- 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),