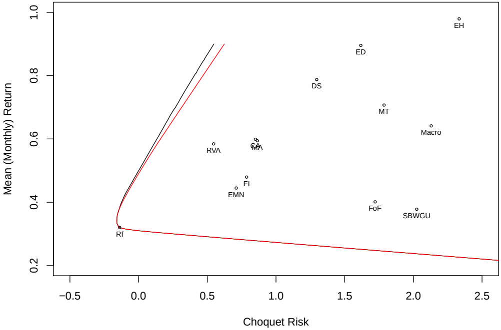
Risk, Uncertainty and the Pessimistic Portfolio

Roger Koenker

University of Illinois, Urbana-Champaign

Knight Lecture: Cornell 24 March 2015



A Knightian Principle

"The fundamental fact of organized activity is the tendency to transform the uncertainties of human opinion and action into measurable probabilities by forming an approximate evaluation of the judgement and capacity of the man."

F.H. Knight, (1921) *Risk, Uncertainty and Profit*, p. 311.

An Introductory Anecdote

Paul Samuelson describes in a 1963 paper asking a colleague at lunch whether he would be willing to make a

$$50 - 50 \text{ bet} \quad \Big\langle \begin{array}{c} \text{win } \$200 \\ \text{lose } \$100 \end{array}$$

An Introductory Anecdote

Paul Samuelson describes in a 1963 paper 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, the chair of the MIT department at the time) responded "no, but I *would* be willing to make 100 such bets."

An Introductory Anecdote

Paul Samuelson describes in a 1963 paper asking a colleague at lunch whether he would be willing to make a

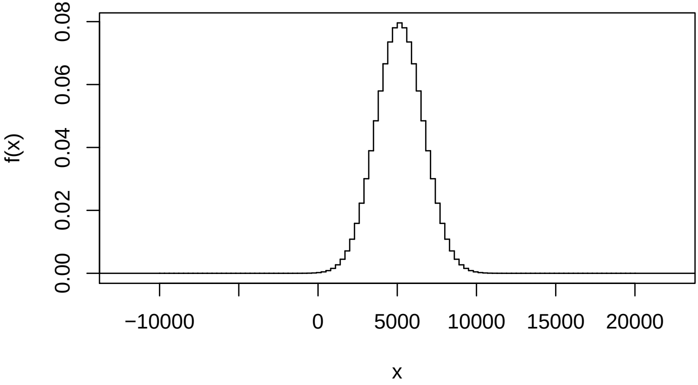
$$50 - 50 \text{ bet} \quad \left\langle \begin{array}{l} \text{win } \$200 \\ \text{lose } \$100 \end{array} \right.$$

The colleague (later revealed to be E. Cary Brown, the chair of the MIT department at the time) responded "no, but I *would* be willing to make 100 such bets."

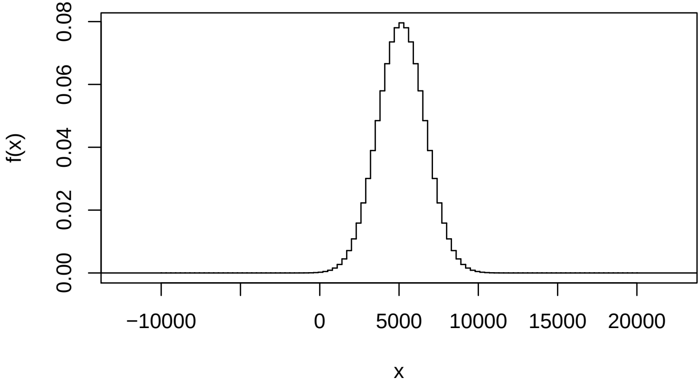
This response was interpreted by Samuelson not only as reflecting:

- A basic confusion about maximizing expected utility, and
- A fundamental misunderstanding of the law of large numbers.

Payoff Density of 100 Samuelson coin flips

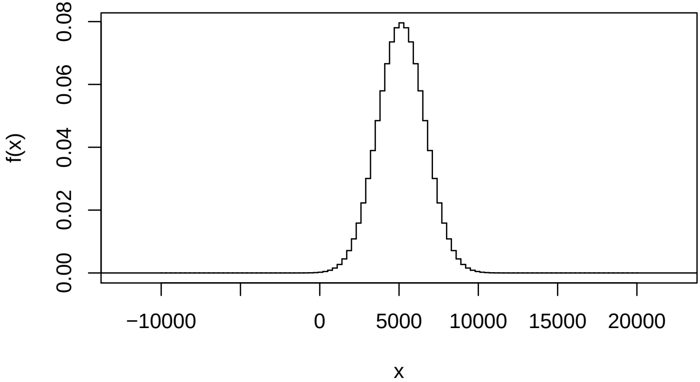


Payoff Density of 100 Samuelson coin flips



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

Payoff Density of 100 Samuelson coin flips



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

Given that he didn't like the one-shot bet, was Brown being irrational?

Samuelson's Argument

Samuelson (1963) begins his diatribe about Brown with an rather *ad hominem* definition he attributes to Stanislaw Ulam:

"I define a coward as someone who will not bet when you offer him two-to-one odds and let him choose *his* side."

Samuelson's Argument

Samuelson (1963) begins his diatribe about Brown with an rather *ad hominem* definition he attributes to Stanislaw Ulam:

"I define a coward as someone who will not bet when you offer him two-to-one odds and let him choose *his* side."

Samuelson then asserts that for any expected utility maximizer:

"No sequence [of gambles] is acceptable, if each of its single plays is not acceptable."

Samuelson's Argument

Samuelson (1963) begins his diatribe about Brown with an rather *ad hominem* definition he attributes to Stanislaw Ulam:

"I define a coward as someone who will not bet when you offer him two-to-one odds and let him choose *his* side."

Samuelson then asserts that for any expected utility maximizer:

"No sequence [of gambles] is acceptable, if each of its single plays is not acceptable."

Pratt (1989) qualifies Samuelson's "Basic Theorem" by noting that it only applies to those with proper utility functions. Maybe Brown was merely improper.

Are Swiss Bicycle Messengers Improper?

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

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

Are Swiss Bicycle Messengers Improper?

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

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

More than half (54%) rejected the bet, just like Brown!
Reference: Fehr and Götte (2002)

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

$$E_F u(X) = \int u(x)dF(x) \geqslant \int u(y)dG(y) = E_G u(Y)$$

or, after a change of variable,

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

A Brief Axiomatic Interlude

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

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

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

For example, P, Q, R might be portfolios consisting of various underlying asset returns.

von Neumann-Morgenstern (Trivialized)

Theorem Suppose we have a preference relation $\{\succeq, \succ, \sim\}$ on $\mathcal{P}$ satisfying the axioms:

von Neumann-Morgenstern (Trivialized)

Theorem 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 $Q \succeq R \Rightarrow P \succeq R$,

von Neumann-Morgenstern (Trivialized)

Theorem 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 $Q \succeq R \Rightarrow P \succeq R$,

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

von Neumann-Morgenstern (Trivialized)

Theorem 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 $Q \succeq R \Rightarrow P \succeq R$,

(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, $\alpha P + (1 - \alpha)R \succ \beta Q + (1 - \beta)R$.

von Neumann-Morgenstern (Trivialized)

Theorem 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 $Q \succeq R \Rightarrow P \succeq R$,

(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. Suppose we consider a weaker form of independence due to Schmeidler (1989).

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

Weakening the Independence Axiom

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

1. 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 $Z : \Omega \to \mathbb{R}$ and increasing functions f and g such that $X = f(Z)$ and $Y = g(Z)$.

On Comonotonicity

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

From my 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

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

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

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

Choquet Expected Utility

Among the many proposals offered to extend expected utility theory the most attractive (to me) replaces the linear (additive)

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

with the nonlinear (non-additive)

Choquet Expected Utility

Among the many proposals offered to extend expected utility theory the most attractive (to me) replaces the linear (additive)

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

with the nonlinear (non-additive)

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

Choquet Expected Utility

Among the many proposals offered to extend expected utility theory the most attractive (to me) replaces the linear (additive)

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

with the nonlinear (non-additive)

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

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), Yaari (1987), Schmeidler (1989), Wakker (1989) and Dennenberg (1990), among others.

Choquet 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.

Choquet 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.

A Smoother example

Another simple, yet intriguing, one-parameter family of pessimistic Choquet distortions is the measure:

$$\nu_\theta(t) = 1 - (1-t)^\theta \qquad \theta \geqslant 1$$

Note that, again changing variables, $t \to F_X(x)$, we have,

$$E_{\nu_\theta} X = \int_0^1 F_X^{-1}(t) d\nu(t) = \int_{-\infty}^{\infty} x d(1 - (1 - F_X(x))^\theta)$$

The pessimist imagines that he gets not a single draw from X but instead gets θ draws, and from these he always gets the worst. The parameter θ is a natural "measure of pessimism," and need not be an integer.

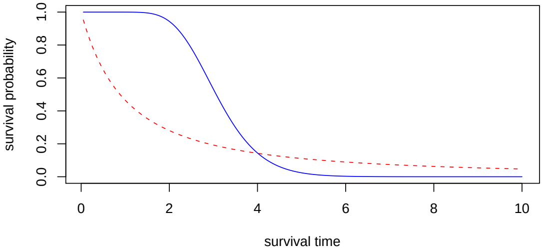
Savage on Pessimism

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 (QTE) 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. Only the luckiest 15% actually do better under the treatment.

How Should We Measure Risk?

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

How Should We Measure Risk?

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.

How Should We Measure Risk?

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. Why is variance a reasonable measure of risk?

Coherent Risk

Definition (Artzner, Delbaen, Eber and Heath (1999)) For real valued random variables $X \in \mathcal{X}$ on $(\Omega, \mathcal{A})$ a mapping $\rho : \mathcal{X} \to \mathbb{R}$ 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 $\lambda \in \mathbb{R}$ 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 $\rho : \mathcal{X} \to \mathbb{R}$ 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 $\lambda \in \mathbb{R}$ and $X \in \mathcal{X}$, $\rho(\lambda + X) = \rho(X) - \lambda$.

Many conventional measures of risks including those based on variance and 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 linear utility and 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

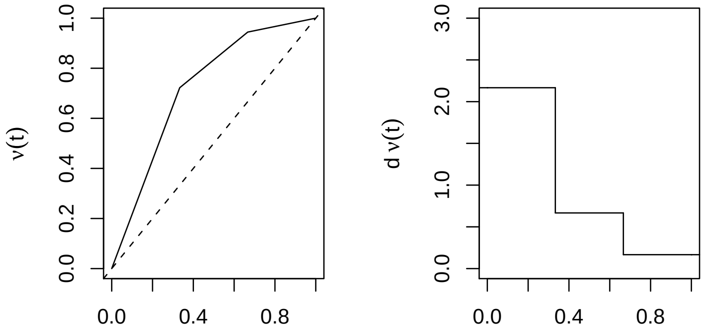
$$d\varphi(t) = \sum_{i=1}^{m} \varphi_i \delta_{\tau_i}(t)$$

where δ_τ denotes (Dirac) point mass 1 at τ, and $\varphi_i \geqslant 0$, with $\sum_{i=0}^{m} \varphi_i = 1$. Then

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

where $\gamma(t) = \sum_{i=1}^{m} \varphi_i \tau_i^{-1} I(t < \tau_i)$.

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* $\mathrm{d}\nu$.

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.

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

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

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

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

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

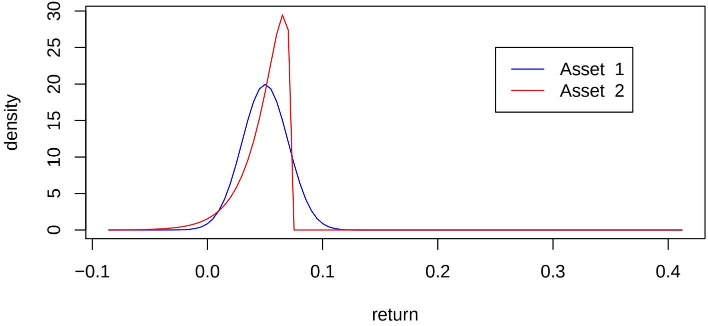
$$\min_{\pi} \rho_{\nu_\alpha}(Y) - \lambda\mu(Y)$$

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

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

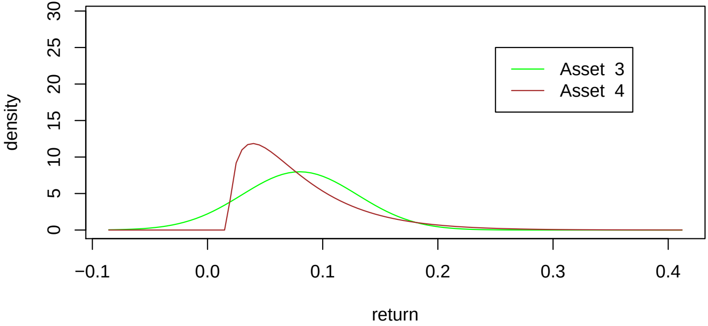
where $\pi(\beta) = (1 - \sum_{j=2}^{p} \beta_j, \beta^\top)^\top$.

A Toy Example



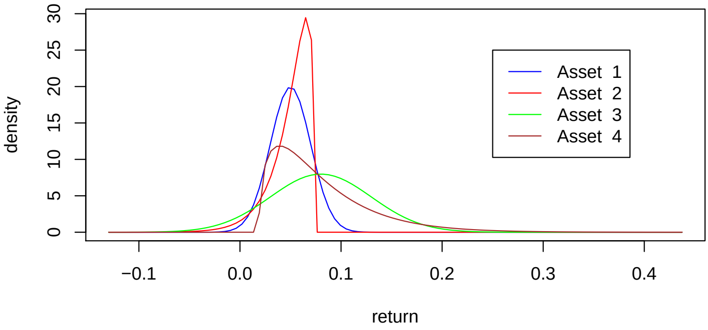
Two asset return densities with identical mean and variance.

A Toy Example



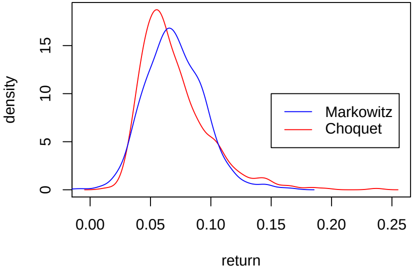
Two more asset return densities with identical mean and variance.

A Toy Example

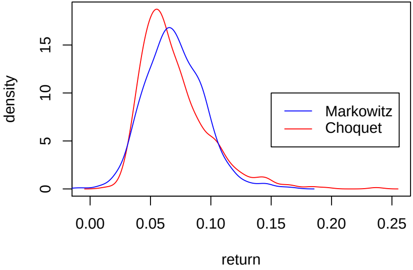


Two pairs of asset return densities with identical mean and variance.

Optimal Choquet and Markowitz Portfolio Returns

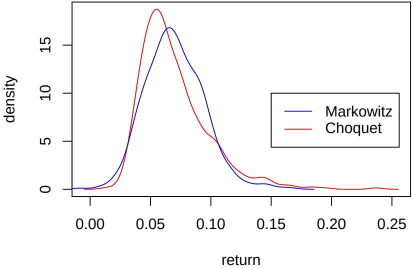


Optimal Choquet and Markowitz Portfolio Returns

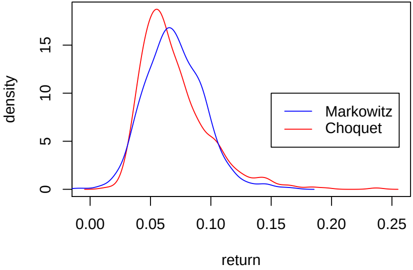


Markowitz portfolio minimizes the standard deviation of returns subject to mean return $\mu = .07$. The Choquet portfolio minimizes Choquet risk (for $\alpha = .10$) subject to earning the same mean return. The Choquet portfolio has better performance in both tails than mean-variance Markowitz portfolio.

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 Pessimistic Portfolios

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:

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

A Unified Theory of Pessimistic Portfolios

Any pessimistic risk measure may be approximated by

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

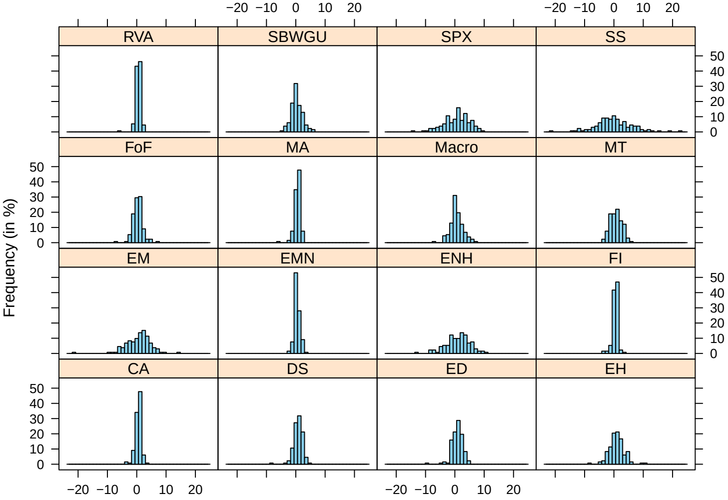
where $\varphi_k > 0$ for $k = 1, 2, ..., m$ and $\sum \varphi_k = 1$.
Portfolio weights can be estimated for these risk measures by solving linear programs that are weighted sums of quantile regression problems:

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

Software in R is available in my package `quantreg`.

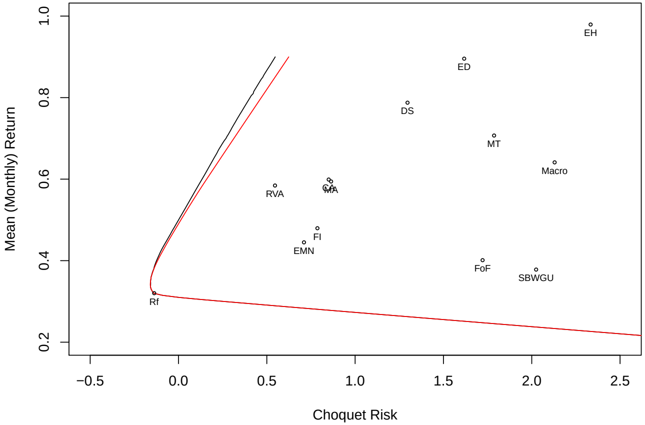
Proof of the Pudding

Lucas and Siegmann (2008) consider portfolios comprised of several hedge funds and some conventional index funds. Marginal (monthly) returns 1994-2004 are illustrated below.



Risk-Return Frontier

Minimizing Choquet Risk with the distortion measure: $d\nu = \frac{1}{2}\delta_{0.1} + \frac{1}{2}\delta_{0.5}$ for various levels of mean return yields this risk-return frontier.



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.

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.
- Large sample theory of Choquet portfolios is an on-going research project.

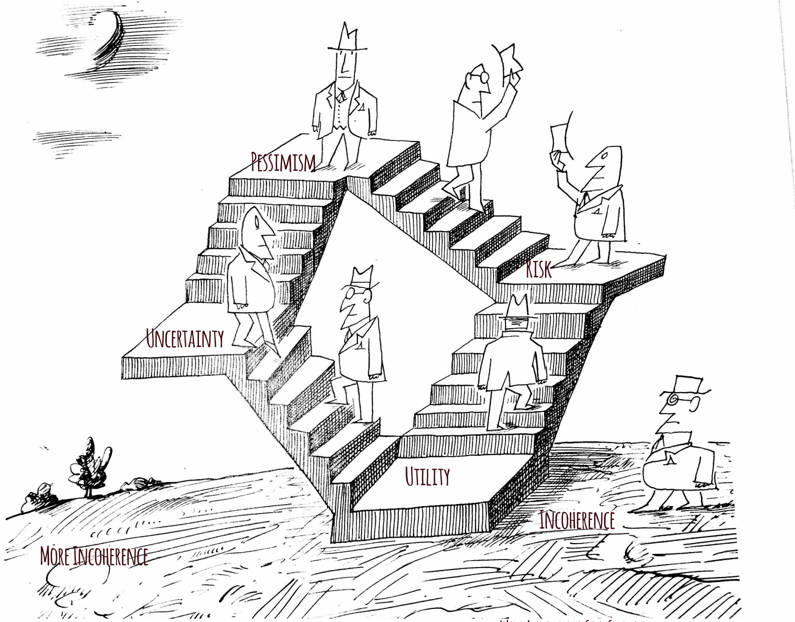
Conclusions

- 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.
- Large sample theory of Choquet portfolios is an on-going research project.
- The real problem is finding a reliable model of asset returns.

Rhetorical Review: Dyads of Risk

Knight	Organized – Disorganized
Ulam, Samuelson	Courageous – Cowardly
Pratt, Samuelson	Proper – Improper
Savage, Artzner et al	Coherent – Incoherent
Choquet, Schmeidler	Pessimistic – Optimistic

Organized Activity



With Apologies to Saul Steinberg