Expected Utility Theory

These are Alexander Wolitzky's MIT notes (14.121), slightly altered by Pedro Hemsley (IE-UFRJ)

Course so far introduced basic theory of choice and utility, extended to DM and producer theory.

Last topic extends in another direction: choice under uncertainty

Choice under Uncertainty

All choices made under some kind of uncertainty.

Sometimes useful to ignore uncertainty, focus on ultimate choices.

Other times, must model uncertainty explicitly.

Examples:

- Insurance markets.
- Financial markets.
- Game theory.

Overview

Impose extra assumptions on basic choice model of Lectures 1-2.

Rather than choosing outcome directly, decision-maker chooses uncertain prospect (or lottery).

A lottery is a probability distribution over outcomes.

Leads to von Neumann-Morgenstern expected utility model.

Consequences and Lotteries

Two basic elements of expected utility theory: consequences (or outcomes) and lotteries.

Consequences

Finite set C of consequences.

Consequences are what the decision-maker ultimately cares about.

Example: "I have a car accident, my insurance company covers most of the costs, but I have to pay a $500 deductible."

Decision-maker (DM) does not choose consequences directly.

Lotteries

DM chooses a lottery, p.

Lotteries are probability distributions over consequences:

$p: C \to [0,1]$ with $\sum_{c \in C} p(c) = 1$.

Set of all lotteries is denoted by P.

Example: "A gold-level health insurance plan, which covers all kinds of diseases, but has a $500 deductible."

Makes sense because DM assumed to rank health insurance plans only insofar as lead to different probability distributions over consequences.

Choice

Decision-maker makes choices from set of alternatives X.

What's set of alternatives here, C or P ?

Answer: P

DM does not choose consequences directly, but instead chooses lotteries.

Assume decision-maker has a rational preference relation $\succsim$ on P.

Natural to assume this?

Convex Combinations of Lotteries

Given two lotteries p and p', the convex combination $\alpha p + (1-\alpha)p'$ is the lottery defined by

$$(\alpha p + (1-\alpha)p')(c) = \alpha p(c) + (1-\alpha)p'(c) \text{ for all } c \in C.$$

One way to generate it:

- First, randomize between p and p' with weights α and $1-\alpha$.
- Second, choose a consequence according to whichever lottery came up.

Such a probability distribution over lotteries is called a compound lottery.

In expected utility theory, no distinction between simple and compound lotteries: simple lottery $\alpha p + (1-\alpha)p'$ and above compound lottery give same distribution over consequences, **so identified with same element of P**.

So, no problem if DM doesn't know exactly the distribution for something. We'll come back to this.

The Set P

As $\alpha p + (1-\alpha)p'$ is also a lottery, P is convex.

P is also closed and bounded (why?).

$\Longrightarrow P$ is a compact subset of $\mathbb{R}^n$, where $n = |C|$.

Whenever $\succsim$ is rational and continuous, can be represented by continuous utility function $U: P \to \mathbb{R}$:

$$p \succsim q \Leftrightarrow U(p) \geq U(q)$$

We're just applying it to lotteries because that's what the DM chooses now.

Intuitively, want more than this.

Want not only that DM has utility function over lotteries, but also that somehow related to "utility" over consequences.

Only care about lotteries insofar as affect distribution over consequences, so preferences over lotteries should have something to do with "preferences" over consequences.

Expected Utility

Best we could hope for is representation by utility function of following form:

Definition: a utility function $U: P \to \mathbb{R}$ has an expected utility form if there exists a function $u: C \to \mathbb{R}$ such that

$$U(p) = \sum_{c \in C} p(c)u(c) \text{ for all } p \in P.$$

In this case, the function U is called an **expected utility function**, and the function u is call a **von Neumann-Morgenstern utility function**.

If preferences over lotteries happen to have an expected utility representation, it's as if DM has a "utility function" over consequences (and chooses among lotteries so as to maximize *expected* "utility over consequences").

Remarks

$$U(p) = \sum_{c \in C} p(c)u(c)$$

Expected utility function $U: P \to \mathbb{R}$ represents preferences $\succsim$ on P just as we had before

$U: P \to \mathbb{R}$ is an example of a standard utility function.

von Neumann-Morgenstern utility function $u: C \to \mathbb{R}$ is **not** a standard utility function.

Can't have a "real" utility function on consequences, as DM never chooses among consequences.

If preferences over lotteries happen to have an expected utility representation, it's **as if** DM has a "utility function" over consequences.

This "utility function" over consequences is the von Neumann-Morgenstern utility function.

Example

Suppose hipster restaurant doesn't let you order steak or chicken, but only probability distributions over steak and chicken.

How should you assess menu item (p (steak), p (chicken)) ?

One way: ask yourself how much you'd like to eat steak, u(steak), and chicken, u(chicken), and evaluate according to

$$p(\text{ steak }) \cdot u(\text{ steak }) + p(\text{ chicken}) \cdot u(\text{ chicken })$$

If this is what you'd do, then your preferences have an expected utility representation.

Suppose instead you choose whichever menu item has p (steak) closest to $\frac{1}{2}$.

Your preferences are rational, so they have a utility representation.

But they do not have an expected utility representation – we'll see this.

Property of EU: Linearity in Probabilities

Theorem

If $U: P \rightarrow \mathbb{R}$ is an expected utility function, then

$$U(\alpha p + (1-\alpha)p') = \alpha U(p) + (1-\alpha)U(p')$$

In fact, a utility function $U: P \rightarrow \mathbb{R}$ has an expected utility form iff this equation holds for all p, p', and $\alpha \in [0,1]$.

Proof: appendix.

Property of EU: Invariant to Affine Transformations

Suppose $U: P \rightarrow \mathbb{R}$ is an expected utility function representing preferences $\succsim$.

Any increasing transformation of U also represents $\succsim$.

Not all increasing transformations of U have expected utility form.

Theorem

Suppose $U: P \rightarrow \mathbb{R}$ is an expected utility function representing preferences $\succsim$. Then $V: P \rightarrow \mathbb{R}$ is also an expected utility function representing $\succsim$ iff there exist $a, b > 0$ such that

$$V(p) = a + bU(p) \text{ for all } p \in P.$$

If this is so, we also have $V(p) = \sum_{c \in C} p(c)v(c)$ for all $p \in P$, where

$$v(c) = a + bu(c) \text{ for all } c \in C$$

Proof: appendix.

What Preferences have an Expected Utility Representation?

Preferences must be rational to have any kind of utility representation.

Preferences on a compact and convex set must be continuous to have a continuous utility representation.

Besides rationality and continuity, what's needed to ensure that preferences have an expected utility representation?

The Independence Axiom

Definition

A preference relation $\succsim$ satisfies independence if, for every

$p, q, r \in P$ and $\alpha \in (0,1)$,

$$p \succsim q \Leftrightarrow \alpha p + (1-\alpha) r \succsim \alpha q + (1-\alpha) r.$$

Can interpret as form of "dynamic consistency."

Doesn't need to hold for consequences.

Back to Example

Suppose choose lottery with $p(\text{steak})$ closest to $\frac{1}{2}$.

Let $p = \left(\frac{1}{2}, \frac{1}{2}\right)$, $q = (0,1)$, $r = (1,0)$, and $\alpha = \frac{1}{2}$.

Then

$$p = \left(\frac{1}{2}, \frac{1}{2}\right) \succ (0,1) = q$$

but

$$\alpha q + (1-\alpha) r = \left(\frac{1}{2}, \frac{1}{2}\right) \succ \left(\frac{3}{4}, \frac{1}{4}\right) = \alpha p + (1-\alpha) r$$

Does not satisfy independence.

Expected Utility: Characterization

Theorem (Expected Utility Theorem)

A preference relation $\succsim$ has an expected utility representation iff it satisfies **rationality**, **continuity**, and **independence**.

Intuition: both having expected utility form and satisfying independence boil down to having straight, parallel indifference curves.

Subjective Expected Utility Theory

So far, probabilities are objective.

In reality, uncertainty is usually subjective.

Subjective expected utility theory (Savage, 1954): under assumptions roughly similar to ones form this lecture, preferences have an expected utility representation where both the utilities over consequences and the subjective probabilities themselves are revealed by decision-maker's choices.

Thus, expected utility theory applies even when the probabilities are not objectively given.

(To learn more, a good starting point is Kreps (1988), "Notes on the Theory of Choice.")

Again, no problem if DM doesn't know the exact distribution.

The same holds in general equilibrium: allows for different individual priors.

One may go beyond and assume DM has some rule to deal with set of priors – e.g., DM may assume that nature will choose the worst possible prior, conditional on his optimal choice, leading to a mini-max structure that deals with fear of misspecification and relates to sub-rational behavior.

See nice discussion in Hansen and Sargent (2000) and a critique by Sims (AER 2001).

Attitudes toward Risk

Money Lotteries

Turn now to special case of choice under uncertainty where **outcomes are measured in dollars**.

Set of consequences C is subset of $\mathbb{R}$.

A lottery is a cumulative distribution function F on $\mathbb{R}$.

(Now we use F instead of p)

Assume preferences have expected utility representation:

$$U(F) = E_F[u(x)] = \int u(x)f(x)dx$$

More generally, we could write $\int u(x)dF(x)$.

This is useful if we do not know whether a density f exists.

We'll assume it does and make ${dF(x)}/{dx} = f(x)$, so that $dF(x) = f(x)dx$, leading to our representation above.

(But everything holds for a general $F(x)$.)

Assume u increasing, differentiable.

Question: how do properties of von Neumann-Morgenstern utility function u relate to decision-maker's attitude toward risk?

Expected Value vs. Expected Utility

Expected value of lottery F is

$$E_F[x] = \int xf(x)dx$$

Expected utility of lottery F is

$$E_F[u(x)] = \int u(x)f(x)dx$$

Can learn about DM's risk attitude by comparing $E_F[u(x)]$ and $u(E_F[x])$.

Risk Attitude: Definitions

Definition

A decision-maker is **risk-averse** if she always prefers the sure wealth level $E_F[x]$ to the lottery F : that is,

$$\int u(x)f(x)dx \leq u\left(\int xf(x)dx\right) \text{ for all } F.$$

A decision-maker is strictly risk-averse if the inequality is strict for all non-degenerate lotteries F.

A decision-maker is risk-neutral if she is always indifferent:

$$\int u(x)f(x)dx = u\left(\int xf(x)dx\right) \text{ for all } F.$$

A decision-maker is risk-loving if she always prefers the lottery:

$$\int u(x)f(x)dx \geq u\left(\int xf(x)dx\right) \text{ for all } F.$$

Risk Aversion and Concavity

Statement that $\int u(x)dF(x) \leq u\left(\int xdF(x)\right)$ for all F is called Jensen's inequality.

Fact: Jensen's inequality holds iff u is concave.

This implies:

Theorem

A decision-maker is (strictly) risk-averse if and only if u is (strictly) concave.

A decision-maker is risk-neutral if and only if u is linear.

A decision-maker is (strictly) risk-loving if and only if u is (strictly) convex.

Certainty Equivalents

Can also define risk-aversion using certainty equivalents.

Definition

The certainty equivalent of a lottery F is the sure wealth level that yields the same expected utility as F : that is,

$$u[CE(F,u)] = \int u(x)f(x)dx$$

That is,

$$CE(F,u) = u^{-1}\left(\int u(x)dF(x)\right).$$

Theorem

A decision-maker is risk-averse iff $CE(F,u) \leq E_F(x)$ for all F.

A decision-maker is risk-neutral iff $CE(F,u) = E_F(x)$ for all F.

A decision-maker is risk-loving iff $CE(F,u) \geq E_F(x)$ for all F.

Quantifying Risk Attitude

We know what it means for a DM to be risk-averse.

What does it mean for one DM to be **more risk-averse** than another?

Two possibilities:

1. u is more risk-averse than v if, for every F, $CE(F,u) \leq CE(F,v)$.
2. u is more risk-averse than v if u is "more concave" than v, in that $u = g \circ v$ for some increasing, concave g.

One more, based on local curvature of utility function: u is more-risk averse than v if, for every x,

$$-\frac{u''(x)}{u'(x)} \geq -\frac{v''(x)}{v'(x)}$$

$A(x,u) = -\frac{u''(x)}{u'(x)}$ is called the Arrow-Pratt coefficient of absolute risk-aversion.

An Equivalence

Theorem

The following are equivalent:

1. For every $F, CE(F,u) \leq CE(F,v)$.
2. There exists an increasing, concave function g such that $u = g \circ v$.
3. For every $x, A(x,u) \geq A(x,v)$.

Risk Attitude and Wealth Levels

How does risk attitude vary with wealth?

Natural to assume that a richer individual is more willing to bear risk: whenever a poorer individual is willing to accept a risky gamble, so is a richer individual.

Captured by decreasing absolute risk-aversion:

Definition

A von Neumann-Morenstern utility function u exhibits decreasing (constant, increasing) absolute risk-aversion if $A(x, u)$ is decreasing (constant, increasing) in x.

Risk Attitude and Wealth Levels

Theorem

Suppose u exhibits decreasing absolute risk-aversion.

If the decision-maker accepts some gamble at a lower wealth level, she also accepts it at any higher wealth level:

that is, for any lottery $F(x)$, if

$$E_F[u(w+x)] \geq u(w),$$

then, for any $w' > w$,

$$E_F[u(w'+x)] \geq u(w').$$

Multiplicative Gambles

What about gambles that multiply wealth, like choosing how risky a stock portfolio to hold? Are richer individuals also more willing to bear multiplicative risk? Depends on increasing/decreasing relative risk-aversion:

$$R(x, u) = -\frac{u''(x)}{u'(x)}x.$$

Theorem

Suppose u exhibits decreasing relative risk-aversion.

If the decision-maker accepts some multiplicative gamble at a lower wealth level, she also accepts it at any higher wealth level: that is, for any lottery $F(t)$, if

$$E_F[u(tw)] \geq u(w),$$

then, for any $w' > w$,

Relative Risk-Aversion vs. Absolute Risk-Aversion

$$R(x) = xA(x)$$

decreasing relative risk-aversion $\Longrightarrow$ decreasing absolute risk-aversion

increasing absolute risk-aversion $\Longrightarrow$ increasing relative risk-aversion

Ex. decreasing relative risk-aversion $\Longrightarrow$ more willing to gamble 1% of wealth as get richer.

So certainly more willing to gamble a fixed amount of money.

Application: Insurance

Risk-averse agent with wealth w, faces probability p of incurring monetary loss L.

Can insure against the loss by buying a policy that pays out a if the loss occurs.

Policy that pays out a costs qa.

How much insurance should she buy?

Agent's Problem

$$\max_a pu(w - qa - L + a) + (1 - p)u(w - qa)$$

u concave $\Longrightarrow$ concave problem, so FOC is necessary and sufficient.

FOC:

$$p(1 - q)u'(w - qa - L + a) = (1 - p)qu'(w - qa)$$

Equate marginal benefit of extra dollar in each state.

Actuarily Fair Prices

Insurance is actuarily fair if expected payout qa equals cost of insurance pa : that is, $p = q$.

With acturarily fair insurance, FOC becomes

$$u'(w - qa - L + a) = u'(w - qa)$$

Solution: $a = L$

A risk-averse DM facing actuarily fair prices will always fully insure.

Actuarily Unfair Prices

What if insurance company makes a profit, so $q > p$?

Rearrange FOC as

$$\frac{u'(w - qa - L + a)}{u'(w - qa)} = \frac{(1-p)q}{p(1-q)} > 1$$

Solution: $a < L$

A risk-averse DM facing actuarily unfair prices will never fully insure.

Intuition: u approximately linear for small risks, so not worth giving up expected value to insure away last little bit of variance.

Comparative Statics

$$\max_{a} pu(w - qa - L + a) + (1-p)u(w - qa)$$

Bigger loss $\Longrightarrow$ buy more insurance (a^* increasing in L) Follows from Topkis' theorem.

If agent has decreasing absolute risk-aversion, then she buys less insurance as she gets richer.

Prove it as an exercise!

Application: Portfolio Choice

Risk-averse agent with wealth w has to invest in a safe asset and a risky asset.

Safe asset pays certain return r.

Risky asset pays random return z, with cdf F.

Agent's problem

$$\max_{a \in [0,w]} \int u(az + (w - a)r) dF(z)$$

First-order condition

$$\int (z-r)u'(az+(w-a)r)dF(z)=0$$

Risk-Neutral Benchmark

Suppose $u'(x) = \alpha x$ for some $\alpha > 0$.

Then

$$U(a)=\int \alpha(az+(w-a)r)dF(z)$$

so

$$U'(a)=\alpha(E[z]-r)$$

Solution: set $a = w$ if $E[z] > r$, set $a = 0$ if $E[z] < r$.

Risk-neutral investor puts all wealth in the asset with the highest rate of return.

$r > E[z]$ Benchmark

$$U'(0)=\int (z-r)u'(w)dF=(E[z]-r)u'(w)$$

If safe asset has higher rate of return, then even risk-averse investor puts all wealth in the safe asset.

More Interesting Case

What if agent is risk-averse, but risky asset has higher expected return?

$$U'(0)=(E[z]-r)u'(w)>0$$

If risky asset has higher rate of return, then risk-averse investor always puts some wealth in the risky asset.

Comparative Statics

Does a less risk-averse agent always invest more in the risky asset?

Sufficient condition for agent v to invest more than agent u :

$$\int (z-r)u'(az+(w-a)r)dF = 0$$
$$\Rightarrow \int (z-r)v'(az+(w-a)r)dF \geq 0$$

u more risk-averse $\Rightarrow v = h \circ u$ for some increasing, convex h. Inequality equals

$$\int (z-r)h'(u(az+(w-a)r))u'(az+(w-a)r)dF \geq 0$$

$h'(\cdot)$ positive and increasing in z

$\Rightarrow$ multiplying by $h'(\cdot)$ puts more weight on positive $(z > r)$ terms, less weight on negative terms.

A less risk-averse agent always invests more in the risky=asset.

Comparing Risky Prospects

Risky Prospects

We've studied decision-maker's subjective attitude toward risk.

Now: study objective properties of risky prospects (lotteries, gambles) themselves, relate to individual decision-making.

Topics:

- First-Order Stochastic Dominance
- Second-Order Stochastic Dominance
- (Optional) Some recent research extending these concepts

First-Order Stochastic Dominance

When is one lottery unambiguously better than another?

Natural definition: F dominates G if, for every amount of money x, F is more likely to yield at least x dollars than G is.

Definition

For any lotteries F and G over $\mathbb{R}$, F first-order stochastically dominates (FOSD) G if

$$F(x) \leq G(x) \text{ for all } x.$$

FOSD and Choice

Main theorem relating FOSD to decision-making:

Theorem

F FOSD G iff every decision-maker with a non-decreasing utility function prefers F to G.

That is, the following are equivalent:

1. $F(x) \leq G(x)$ for all x.

2. $\int u(x)dF \geq \int u(x)dG$ for every non-decreasing function $u: \mathbb{R} \to \mathbb{R}$.

Proof:

Preferred by Everyone => FOSD

If F does not FOSD G, then there's some amount of money x^* such that G is more likely to give at least x^* than F is.

Consider a DM who only cares about getting at least x^* dollars.

She will prefer G.

FOSD => Preferred by Everyone

Main idea: F FOSD $G \Longrightarrow F$ gives more money "realization-by-realization."

Suppose draw x according to G, but then instead give decision-maker

$$y(x) = F^{-1}(G(x))$$

Then:

1. $y(x) \geq x$ for all x, and
2. y is distributed according to F.

$\Longrightarrow$ paying decision-maker according to F just like first paying according to G, then sometimes giving more money.

Any decision-maker who likes money likes this.

QED.

Second-Order Stochastic Dominance

Q : When is one lottery better than another for any decision-maker?

A: First-Order Stochastic Dominance.

Q: When is one lottery better than another for any risk-averse decision-maker?

A: Second-Order Stochastic Dominance.

Definition

F second-order stochastically dominates (SOSD) G iff every decision-maker with a non-decreasing and concave utility function prefers F to G : that is,

$$\int u(x)dF \geq \int u(x)dG$$

for every non-decreasing and concave function $u: \mathbb{R} \to \mathbb{R}$.

SOSD is a weaker property than FOSD.

SOSD for Distributions with Same Mean

If F and G have same mean, when will any risk-averse decision-maker prefer F ?

When is F "unambiguously less risky" than G ?

Mean-Preserving Spreads

G is a mean-preserving spread of F if G can be obtained by first drawing a realization from F and then adding noise.

Definition

G is a mean-preserving spread of F iff there exist random variables x, y, and ε such that

$$y = x + \varepsilon,$$

x is distributed according to F, y is distributed according to G, and $E[\varepsilon \mid x] = 0$ for all x.

Formulation in terms of cdfs:

$$\int_{-\infty}^{x} G(y)dy \geq \int_{-\infty}^{x} F(y)dy \text{ for all } x.$$

Characterization of SOSD for CDFs with Same Mean

Theorem

Assume that $\int xdF = \int xdG$. Then the following are equivalent:

1. F SOSD G.
2. G is a mean-preserving spread of F.

3. $\int_{-\infty}^{x} G(y)dy \geq \int_{-\infty}^{x} F(y)dy$ for all x.

General Characterization of SOSD

Theorem

The following are equivalent:

1. F $SOSD$ G.
2. $\int_{-\infty}^{x} G(y)dy \geq \int_{-\infty}^{x} F(y)dy$ for all x.
3. There exist random variables x, y, z, and ε such that

$$y = x + z + \varepsilon,$$

x is distributed according to F, y is distributed according to G, z is always non-positive, and $E[\varepsilon \mid x] = 0$ for all x.

4. There exists a cdf H such that F FOSD H and G is a mean-preserving spread of H.

Complete Dominance Orderings [Optional]

FOSD and SOSD are partial orders on lotteries:

"most distributions" are not ranked by FOSD or SOSD.

To some extent, nothing to be done:

If F doesn't FOSD G, some decision-maker prefers G.

If F doesn't SOSD G, some risk-averse decision-maker prefers G.

However, recent series of papers points out that if view F and G as lotteries over monetary gains and losses rather than final wealth levels, and only require that no decision-maker prefers G to F for all wealth levels, do get a complete order on lotteries (and index of lottery's "riskiness").

Acceptance Dominance

Consider decision-maker with wealth w, has to accept or reject a gamble F over gains/losses x.

Accept iff

$$E_F[u(w + x)] \geq u(w).$$

Definition

F acceptance dominates G if, whenever F is rejected by decision-maker with concave utility function u and wealth w, so is G.

That is, for all u concave and $w > 0$,

$$E_F[u(w+x)] \leq u(w)$$
$$E_G[u(w+x)] \leq u(w).$$

Acceptance Dominance and FOSD/SOSD

F SOSD G

$\Longrightarrow E_F[u(w+x)] \geq E_G[u(w+x)]$ for all concave u and wealth w

$\Longrightarrow F$ acceptance dominates G.

If $E_F[x] > 0$ but x can take on both positive and negative values, can show that F acceptance dominates lottery that doubles all gains and losses.

Acceptance dominance refines SOSD.

But still very incomplete.

Turns out can get complete order from something like: acceptance dominance at all wealth levels, or for all concave utility functions.

Wealth Uniform Dominance

Definition

F wealth-uniformly dominates G if, whenever F is rejected by decision-maker with concave utility function u at every wealth level w, so is G.

That is, for all $u \in \mathcal{U}^*$,

$$E_F[u(w+x)] \leq u(w) \text{ for all } w > 0$$
$$E_G[u(w+x)] \leq u(w) \text{ for all } w > 0.$$

Utility Uniform Dominance

Definition

F utility-uniformly dominates G if, whenever F is rejected at wealth level w by a decision-maker with any utility function $u \in \mathcal{U}^*$, so is G.

That is, for all $w > 0$,

$$E_F[u(w+x)] \leq u(w) \text{ for all } u \in \mathcal{U}^*$$
$$E_G[u(w+x)] \leq u(w) \text{ for all } u \in \mathcal{U}^*.$$

Uniform Dominance: Results

Hart (2011):

- Wealth-uniform dominance and utility-uniform dominance are complete orders.
- Comparison of two lotteries in these orders boils down to comparison of simple measures of the "riskiness" of the lotteries.
- Measure for wealth-uniform dominance: critical level of risk-aversion above which decision maker with constant absolute risk-aversion rejects the lottery.
- Measure for utility-uniform dominance: critical level of wealth below which decision-maker with log utility rejects the lottery.

Appendix: some proofs

U has expected utility form $\Leftrightarrow$ U linear in probabilities

Theorem

$U: P \rightarrow \mathbb{R}$ has an expected utility form if and only if

$$U(\alpha p + (1-\alpha)p') = \alpha U(p) + (1-\alpha)U(p')$$

holds for all p, p', and $\alpha \in [0,1]$.

Notice: this is MWG proposition 6B1. It uses the following notation: $U(\sum \alpha_k p_k) = \sum \alpha_k U(p_k)$, just substituting p for L (which stands for 'lottery').

Proof

Without loss of generality, we will assume only two consequences, c^1 and c^2.

Hence any lottery p may be written as $p = (p^1, p^2)$, in which $p^1 = Prob(c^1)$ and $p^2 = Prob(c^2)$.

All arguments below hold unchanged for $p = (p^1, \dots, p^n)$, that is, for n consequences $c^1, \dots, c^n$. This extension is shown in red below; you may simply ignore it in your first reading.

The arguments below also hold for $c \in [c^1, c^n] \in \mathbb{R}$, but the math is not exactly the same.

Necessity: U linear in probabilities $\Rightarrow$ U has expected utility form

Write lottery $p = (p^1, p^2)$ as a convex combination of degenerate lotteries (C^1, C^2):

$$p = p^1 C^1 + p^2 C^2 + \cdots + p^n C^n$$

That is, $C^1 = (1,0)$, meaning that consequence 1 (c^1) happens with probability 1, and $C^2 = (0,1)$, meaning that consequence 2 (c^2) happens with probability 1. The equation above is simply $p = (p^1, p^2) = (p^1, 0) + (0, p^2) = p^1 \cdot (1,0) + p^2 \cdot (0,1) = p^1C^1 + p^2C^2$.

Then:

$$U(p) = U(p^1C^1 + p^2C^2 + \cdots + p^nC^n) = p^1U(C^1) + p^2U(C^2) + \cdots + p^nU(C^n)$$

The second equality follows from our assumption: U is linear in probabilities.

But $U(C^1)$ is the utility from a degenerate lottery, that is, it's simply the vNM utility of consequence c_1: $U(C^1) = u(c^1)$.

Remember our notation: big U is for DM's actual utility; small u is form DM's vNM utility. Big C denotes a lottery; small c denotes a consequence.

The last equation may be rewritten as:

$$U(p) = U(p^1C^1 + p^2C^2 + \cdots + p^nC^n) = p^1U(C^1) + p^2U(C^2) + \cdots + p^nU(C^n)$$
$$= p^1u(c^1) + p^2u(c^2) + \cdots + p^nu(c^n)$$

In short:

$$U(p) = p^1u(c^1) + p^2u(c^2) + \cdots + p^nu(c^n)$$

This is exactly the expected utility property, concluding the proof.

Sufficiency: U has expected utility form $\Rightarrow$ U linear in probabilities

Consider a compound lottery:

$$(p_1, p_2, \dots, p_k\ ; \alpha_1, \alpha_2, \dots, \alpha_k)$$

Notice that now we have p_1 instead of p^1; and p_2 instead of p^2. **Superscripts** refer to consequences; **subscripts** refer to specific lotteries.

That is, p_1 and p_2 are different lotteries, and each one is a vector assigning probabilities for each of the two possible consequences c^1 and c^2:

$$p_i = (p_i^1, p_i^2, \dots, p_i^n)$$

For $i = 1, \dots, k$. That is, we have k lotteries, and each one is chosen with probability α_i in our compound lottery.

We will allow k to be generic. If you want, just take $k = 2$ in the following computations – again, it's without loss of generality, but be careful not to confuse the number of consequences with the number of lotteries.

Consider now the following (reduced) lottery:

$$\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k$$

Consider the utility of this lottery:

$$U(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k)$$

We may now use our assumption: U has the expected utility form. That is, one may rewrite this utility as:

$$U(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k) = u^1 \cdot Prob(u^1) + u^2 \cdot Prob(u^2) + \cdots + u^n \cdot Prob(u^n)$$

What are these u^i's? We just need to know that there are some u^i's that make this equation hold – our assumption guarantees this is the case. But we do have an interpretation for them: u^i is just the vNM utility of consequence c^i. Analogously, $Prob(u^i)$ is simply the probability of this consequence, computed from the compound lottery $\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k$.

This explains why we have subscripts on the LHS, but superscripts on the RHS. In the LHS, we have lotteries (that generate a compound lottery). On the RHS, we have consequences with vNM utilities u^1, u^2,...,u^n. If you want, you may think of the lottery on the LHS as any given lottery p.

Let's develop this equation:

$$U(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k) =$$

$$u^1 \cdot Prob(u^1) + u^2 \cdot Prob(u^2) =$$

$$u^1 \cdot \underbrace{(\alpha_1 p_1^1 + \alpha_2 p_2^1 + \cdots + \alpha_k p_k^1)}_{Prob(u^1)} + u^2 \cdot \underbrace{(\alpha_1 p_1^2 + \alpha_2 p_2^2 + \cdots + \alpha_k p_k^2)}_{Prob(u^2)} + \cdots + u^n \cdot \underbrace{(\alpha_1 p_1^n + \alpha_2 p_2^n + \cdots + \alpha_k p_k^n)}_{Prob(u^2)} =$$

$$\alpha_1 \cdot \underbrace{(u^1 \cdot p_1^1 + u^2 \cdot p_1^2 + \cdots + u^n \cdot p_1^n)}_{U(p_1)} + \alpha_2 \cdot \underbrace{(u^1 \cdot p_2^1 + u^2 \cdot p_2^2 + \cdots + u^n \cdot p_2^n)}_{U(p_2)} + \cdots + \alpha_k \cdot \underbrace{(u^1 \cdot p_k^1 + u^2 \cdot p_k^2 + \cdots + u^n \cdot p_k^n)}_{U(p_k)} =$$

$$\alpha_1 \cdot U(p_1) + \alpha_2 \cdot U(p_2) + \cdots + \alpha_k \cdot U(p_k)$$

From the second to the third line, we use the definition of $Prob(u^1)$: it is simply the first coordinate of the vector $\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k$, ie, the compound lottery. It is analogous for $Prob(u^2)$ to $Prob(u^n)$.

From the third to the fourth line: we use again our assumption: U has the expected utility form. Hence we may write $U(p_1) = u^1 \cdot p_1^1 + u^2 \cdot p_1^2 + \cdots + u^n \cdot p_1^n$. It is analogous for $U(p_2)$ to $U(p_k)$.

In short:

$$U(\alpha_1 p_1 + \alpha_2 p_2 + \cdots + \alpha_k p_k) = \alpha_1 \cdot U(p_1) + \alpha_2 \cdot U(p_2) + \cdots + \alpha_k \cdot U(p_k)$$

That is, U is linear in probabilities, concluding the proof.

QED.

Expected utility form is preserved under positive affine transformations

Theorem

$U, \tilde{U}$ have expected utility form (and represent the same preferences) $\Leftrightarrow$ there are $\beta > 0, \gamma$ such that for all p, $\tilde{U}(p) = \beta U(p) + \gamma$.

This is MWG proposition 6B2.

Proof

Choose $\overline{p}, \underline{p}$ such that for all lottery p, $\overline{p} \succcurlyeq p \succcurlyeq \underline{p}$.

If $\overline{p} \sim \underline{p}$, then all utility functions are constant, and the result follows immediately.

Assume now $\overline{p} \succ \underline{p}$.

Sufficiency: If U has expected utility form, then $\tilde{U}(p) = \beta U(p) + \gamma$ also has expected utility form.

Consider a compound lottery $\alpha_1 p_1 + \alpha_2 p_2$. That is, we have two lotteries (p_1 and p_2), and each is chosen with probability α_1 and α_2, respectively.

Without loss of generality, we consider only two lotteries, but the argument is unchanged for k lotteries.

Compute the utility of this compound lottery under $\tilde{U}$:

$$\tilde{U}(\alpha_1 p_1 + \alpha_2 p_2) =$$
$$\beta \cdot U(\alpha_1 p_1 + \alpha_2 p_2) + \gamma =$$
$$\beta \cdot [\alpha_1 U(p_1) + \alpha_2 U(p_2)] + \gamma =$$
$$\alpha_1 \beta \cdot U(p_1) + \alpha_2 \beta \cdot U(p_2) + \underbrace{[\alpha_1 \gamma + \alpha_2 \gamma]}_{=\gamma} =$$
$$\alpha_1 \cdot \underbrace{[\beta \cdot U(p_1) + \gamma]}_{\tilde{U}(p_1)} + \alpha_2 \cdot \underbrace{[\beta \cdot U(p_2) + \gamma]}_{\tilde{U}(p_2)} =$$
$$\alpha_1 \cdot \tilde{U}(p_1) + \alpha_2 \cdot \tilde{U}(p_2)$$

From the first to the second line: we use the definition $\tilde{U}(p) = \beta U(p) + \gamma$.

From the second to the third line: we use the assumption that U has the expected utility form: hence, $U(\alpha_1 p_1 + \alpha_2 p_2) = \alpha_1 U(p_1) + \alpha_2 U(p_2)$.

From the third to the fourth line: we simply write $\gamma = \alpha_1 \gamma + \alpha_2 \gamma$, which is true because $\alpha_1 + \alpha_2 = 1$ (it's a probability distribution, so it must sum up to one).

From the fourth to the fifth line: we factor out α_1 and α_2.

In short:

$$\widetilde{U}(\alpha_1 p_1 + \alpha_2 p_2) = \alpha_1 \cdot \widetilde{U}(p_1) + \alpha_2 \cdot \widetilde{U}(p_2)$$

That is, $\widetilde{U}$ has the expected utility form, as we wanted to show.

Necessity: U and $\widetilde{U}$ have the expected utility form (and represent the same preferences) implies that for some $\beta > 0, \gamma$, one has $\widetilde{U}(p) = \beta U(p) + \gamma$

Fix a lottery p.

Choose $\lambda_p \in [0,1]$ such that:

$$U(p) = \lambda_p \cdot U(\overline{p}) + (1 - \lambda_p) \cdot U(\underline{p})$$

This equation has two implications.

First implication: $p \sim \lambda_p \overline{p} + (1 - \lambda_p)\underline{p}$

This holds because we're assuming U has the expected utility form. The previous theorem states that if this is the case, then U is linear in probabilities. Hence the RHS of this last equation may be rewritten as:

$$\lambda_p \cdot U(\overline{p}) + (1 - \lambda_p) \cdot U(\underline{p}) = U(\lambda_p \cdot \overline{p} + (1 - \lambda_p) \cdot \underline{p})$$

Hence $U(p) = U\left(\lambda_p \cdot \overline{p} + (1-\lambda_p) \cdot \underline{p}\right)$. By definition of a utility function, the arguments on each side must be indifferent for the DM.

Second implication:

$$\lambda_p = \frac{U(p) - U\left(\underline{p}\right)}{U(\overline{p}) - U\left(\underline{p}\right)}$$

This is just a rearrangement of the equation above.

We know that $\widetilde{U}$ is linear in probabilities (previous theorems) and represents the same preferences. Hence:

$$\widetilde{U}(p) = \lambda_p \cdot \widetilde{U}(\overline{p}) + (1-\lambda_p) \cdot \widetilde{U}\left(\underline{p}\right) =$$

$$\lambda_p \cdot \left[\widetilde{U}(\overline{p}) - \widetilde{U}\left(\underline{p}\right)\right] + \widetilde{U}\left(\underline{p}\right) =$$

$$\underbrace{\left(\frac{U(p) - U\left(\underline{p}\right)}{U(\overline{p}) - U\left(\underline{p}\right)}\right)}_{\lambda_p} \cdot \left[\widetilde{U}(\overline{p}) - \widetilde{U}\left(\underline{p}\right)\right] + \widetilde{U}\left(\underline{p}\right)$$

In short:

$$\widetilde{U}(p) = \underbrace{\left(\frac{U(p) - U\left(\underline{p}\right)}{U(\overline{p}) - U\left(\underline{p}\right)}\right)}_{\lambda_p} \cdot \left[\widetilde{U}(\overline{p}) - \widetilde{U}\left(\underline{p}\right)\right] + \widetilde{U}\left(\underline{p}\right)$$

In this last expression, only $U(p)$ depends on p. All other terms are parameters. Rearrange this expression to get the following:

$$\tilde{U}(p) = \left[\frac{\tilde{U}(\overline{p}) - \tilde{U}\left(\underline{p}\right)}{U(\overline{p}) - U\left(\underline{p}\right)}\right] \cdot U(p) + \tilde{U}\left(\underline{p}\right) - U\left(\underline{p}\right) \cdot \left[\frac{\tilde{U}(\overline{p}) - \tilde{U}\left(\underline{p}\right)}{U(\overline{p}) - U\left(\underline{p}\right)}\right]$$

Again: except $U(p)$, everything in this expression is a parameter, built from functions (U or $\tilde{U}$) evaluated at specific arguments ($\overline{p}$ or $\underline{p}$). We can label them as we want. Let's choose:

$$\beta = \left[\frac{\tilde{U}(\overline{p}) - \tilde{U}\left(\underline{p}\right)}{U(\overline{p}) - U\left(\underline{p}\right)}\right]$$

$$\gamma = \tilde{U}\left(\underline{p}\right) - U\left(\underline{p}\right) \cdot \underbrace{\left[\frac{\tilde{U}(\overline{p}) - \tilde{U}\left(\underline{p}\right)}{U(\overline{p}) - U\left(\underline{p}\right)}\right]}_{\beta}$$

Then one has:

$$\tilde{U}(p) = \beta U(p) + \gamma$$

This is what we wanted to show, concluding the proof.

QED.

Expected utility theorem

Theorem

(Rational and continuous) Preferences may be represented by an utility function with the expected utility form if and only if it respects the axiom of independence.

Proof of Necessity: if $\succcurlyeq$ respect the axiom of independence, then it may be represented by a utility function with the expected utility form.

Assume there are lotteries $\overline{p}$ and $\underline{p}$ such that for all p, one has $\overline{p} \succcurlyeq p \succcurlyeq \underline{p}$.

If $\overline{p} \sim \underline{p}$, the result follows immediately: use a constant utility function.

Assume from now on $\overline{p} \succ \underline{p}$.

Step 1

Take α and β such that $1 > \beta > \alpha > 0$.
Write:

$$\overline{p} =$$
$$\beta\overline{p} + (1-\beta)\overline{p} \succ$$
$$\beta\overline{p} + (1-\beta)\underline{p} =$$
$$(\beta - \alpha)\overline{p} + \alpha\overline{p} + (1-\beta)\underline{p} \succ$$
$$(\beta - \alpha)\underline{p} + \alpha\overline{p} + (1-\beta)\underline{p} =$$
$$\alpha\overline{p} + (1-\alpha)\underline{p} \succ$$
$$\alpha\underline{p} + (1-\alpha)\underline{p} =$$
$$\underline{p}$$

From the first to the second line: $\overline{p}$ is the average of $\overline{p}$ and $\overline{p}$!

From the second to the third line: we apply the axiom of independence. Observe that we keep $\overline{p}$ in the first term of the sum, but substitute $\overline{p}$ for $\underline{p}$ in the second term. Since $\overline{p} \succ \underline{p}$, the axiom of independence implies the strict preference.

From the third to the fourth line: add and subtract $\alpha\overline{p}$.

From the fourth to the fifth line: again, we just substitute $\overline{p}$ for $\underline{p}$ in one term of the sum, and leave the rest unchanged. The axiom of independence applies again.

From the fifth to the sixth line: we cancel out $\beta\underline{p}$.

From the sixth to the seventh line: we repeat the argument of the 2nd to 3rd line, in reverse order.

From the seventh to the eighth line: we repeat the argument of the 1st to 2nd line, in reverse order.

Step 2: for all p, there is only one λ_p such that $\lambda_p\overline{p} + (1-\lambda_p)\underline{p} \sim p$

Existence follows from continuity. For any lottery p, define the sets:

$$\left\{\lambda \in [0,1]: \lambda\overline{p} + (1-\lambda)\underline{p} \succcurlyeq p\right\}$$

$$\left\{\lambda \in [0,1]: \lambda\overline{p} + (1-\lambda)\underline{p} \preccurlyeq p\right\}$$

Continuity and completeness of $\succcurlyeq$ imply that both sets are closed (why?). Moreover, any λ belongs to at least one of these sets. Since both sets are non-empty and $[0,1]$ is connected, there must by some λ belonging to both (again: why?). Define it as λ_p.

Uniqueness follows from the previous step. If we were to slightly increase the value of λ_p (from α to β in the notation of the previous step), we would get a new lottery strictly preferred to the DM, breaking indifference.

Step 3: $U(p) = \lambda_p$ is a utility function that represents $\succcurlyeq$

Consider two lotteries p and q.

From steps 1 and 2, we may write:

$$p \succcurlyeq q \Leftrightarrow \lambda_p\overline{p} + (1-\lambda_p)\underline{p} \succcurlyeq \lambda_q\overline{p} + (1-\lambda_q)\underline{p} \Leftrightarrow \lambda_p \geqslant \lambda_q$$

The first $\Leftrightarrow$ comes from step 2: use $p \sim \lambda_p\overline{p} + (1-\lambda_p)\underline{p}$, and analogously $q \sim \lambda_q\overline{p} + (1-\lambda_q)\underline{p}$.

The second $\Leftrightarrow$ comes from step 1, taking $\lambda_p = \beta$ and $\lambda_q = \alpha$.

In short, $p \succcurlyeq q \Leftrightarrow \lambda_p \succcurlyeq \lambda_q$. This is the definition of an utility function representing $\succcurlyeq$.

Step 4: $U(p) = \lambda_p$ has the expected utility form.

We have to show that for all $\alpha \in [0,1]$, and for any lotteries p, p', one has:

$$U[\alpha p + (1-\alpha)p'] = \alpha U(p) + (1-\alpha)U(p')$$

From step 2, we have:

$$p \sim \lambda_p \overline{p} + (1-\lambda_p)\underline{p}$$

$$p' \sim \lambda_{p'} \overline{p} + (1-\lambda_{p'})\underline{p}$$

From step 3, we have : $U(p) = \lambda_p$. These two relations become:

$$p \sim U(p)\overline{p} + (1-U(p))\underline{p}$$

$$p' \sim U(p')\overline{p} + (1-U(p'))\underline{p}$$

Take a convex combination $\alpha p + (1-\alpha)p'$. Given the two relations above, we have:

$$\alpha p + (1-\alpha)p' \sim \alpha\left[U(p)\overline{p} + (1-U(p))\underline{p}\right] + (1-\alpha)\left[U(p')\overline{p} + (1-U(p'))\underline{p}\right]$$

Factor out $\overline{p}$ and $\underline{p}$

$$\sim [\alpha U(p) + (1-\alpha)U(p')] \cdot \overline{p} + \left[\alpha(1-U(p)) + (1-\alpha)(1-U(p'))\right] \cdot \underline{p}$$

$$\sim [\alpha U(p) + (1-\alpha)U(p')] \cdot \overline{p} + \left[1 - \left(\alpha U(p) + (1-\alpha)U(p')\right)\right] \cdot \underline{p}$$

Notice now that the terms in red are the same. We may denote it by any letter – for example, λ (without the subscript to distinguish it from both λ_p and $\lambda_{p'}$): $\lambda = \alpha U(p) + (1-\alpha)U(p')$.

The last line becomes:

$$\lambda \cdot \overline{p} + [1-\lambda] \cdot \underline{p}$$

In short, and without all the colors:

$$\alpha p + (1-\alpha)p' \sim \lambda \cdot \overline{p} + [1-\lambda] \cdot \underline{p}$$

But this is the very definition of λ_p as defined in step 2, only applied to $\alpha p + (1-\alpha)p'$. (If you want, you may write $\lambda_{\alpha p+(1-\alpha)p'}$ instead of λ.)

Or, using step 4, we have the more intuitive notation $\lambda = U[\alpha p + (1-\alpha)p']$.

But we just defined $\lambda = \alpha U(p) + (1-\alpha)U(p')$.

Since these are the same λ, one has:

$$U[\alpha p + (1-\alpha)p'] = \alpha U(p) + (1-\alpha)U(p')$$

That is, U has the expected utility property, concluding the proof.

QED.