Chapter 5

Uncertainty

Example.

Consider an investor who must decide how much of their initial wealth w to put into a risky asset. The risky asset can have any of the positive or negative rates of return r_i with probabilities p_i, $i = 1, 2, \cdots, n$. Suppose the investor is an expected utility maximizer and their utility for x amount of money for sure can be represented by a twice continuously differentiable, strictly increasing and strictly concave utility function $u(x)$. Let a^* be the investor's optimal amount of money to put in the risky asset. Give a necessary and sufficient condition for the investor to have strict incentives to invest in the risky asset, that is, $a^* > 0$ and a^* is strictly preferred to $a = 0$.

Solution.

Since the investor is an expected utility maximizer, optimization problem is given by:

$$\max_{0 \leq a \leq w} U(a) = \sum_{i=1}^{n} p_i \cdot u(w + ar_i)$$

A sufficient condition for $a^* > 0$ is $U'(0) = u'(w) \cdot \sum_{i=1}^{n} p_i r_i > 0$, that is, the expected rate of return $\sum_{i=1}^{n} p_i r_i > 0$.

Next suppose $\sum_{i=1}^{n} p_i r_i \leq 0$. Since $u(\cdot)$ is strictly concave, $U''(a) = \sum_{i=1}^{n} p_i \cdot r_i^2 \cdot u''(w + ar_i) \leq 0$, which implies $U(\cdot)$ is also strictly concave. Consequently, $U'(a) \leq U'(0) \leq 0$ for $a \geq 0$. It follows that the condition is also necessary.

The expected utility representation is convenient, but as written it looks ad hoc — why *expected* utility specifically, rather than some other functional of the probability distribution? The agenda of this chapter mirrors what we did under certainty: identify the axioms on the agent's preference relation that force an expected utility representation.

5.1 Setups

Definition 5.1.1: Simple Lottery; Compound Lottery

Let $X = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$ denote a finite set of *sure* outcomes (i.e., without uncertainty).

- A *simple lottery* $\boldsymbol{p} = p_1 \circ \boldsymbol{x}_1 + p_2 \circ \boldsymbol{x}_2 + \cdots + p_n \circ \boldsymbol{x}_n$ $(p_1 + p_2 + \cdots + p_n = 1)$ is a probability distribution over a finite number of sure outcomes $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n\}$. When there is no confusion, we will also use $\boldsymbol{p} = \{p_1, p_2, \cdots, p_n\}$ to denote such a simple lottery.
- A *compound lottery* $\sum_{j=1}^{k} \alpha_j \boldsymbol{p}^j$ $(\alpha_1 + \alpha_2 + \cdots + \alpha_k = 1)$ is a probability distribution over a finite number of simple lotteries $\{\boldsymbol{p}^1, \boldsymbol{p}^2, \cdots, \boldsymbol{p}^k\}$.

We make two simplifying assumptions about how the decision-maker perceives lotteries:

Assumption 5.1.2

1. The probabilities are **objective**: the probabilities are regarded as objective or exogenously given facts by the decision-maker (in contrast to subjective evaluation).
2. The decision-maker is a **consequentialist**: decision-makers only care about the outcomes, instead of how lotteries are mixed.

   Under this assumption, the compound lottery $\sum_{j=1}^{k} \alpha_j \boldsymbol{p}^j$ is identical to the simple lottery it induces, i.e.,

$$\sum_{j=1}^{k} \alpha_j \boldsymbol{p}^j \iff \left(\sum_{j=1}^{k} \alpha_j \boldsymbol{p}_1^j, \sum_{j=1}^{k} \alpha_j \boldsymbol{p}_2^j, \cdots, \sum_{j=1}^{k} \alpha_j \boldsymbol{p}_n^j \right)$$

The first assumption can be relaxed.

Given the two assumptions, we can restrict our focus on the space of simple lotteries $\mathcal{P} = \Delta(X)$.

Definition 5.1.3: Space of Simple Lotteries

The space of simple lotteries, $\Delta(x)$, is defined as

$$\Delta(X) = \{(p_1, \cdots, p_n) : p_i \geq 0 \;\; \forall i, p_1 + \cdots + p_n = 1\}.$$

Consider the agent's preference relation $\succsim$ over $\mathcal{P}$. To ensure a utility representation $U : \mathcal{P} \to \mathbb{R}$, we maintain the three basic assumptions as in the certainty benchmark.

- Completeness: For any $\boldsymbol{p}^1, \boldsymbol{p}^2 \in \mathcal{P}$, either $\boldsymbol{p}^1 \succsim \boldsymbol{p}^2$ or $\boldsymbol{p}^2 \succsim \boldsymbol{p}^1$ (or both).
- Transitivity: For any $\boldsymbol{p}^1, \boldsymbol{p}^2, \boldsymbol{p}^3 \in \mathcal{P}$, if $\boldsymbol{p}^1 \succsim \boldsymbol{p}^2$ and $\boldsymbol{p}^2 \succsim \boldsymbol{p}^3$, then $\boldsymbol{p}^1 \succsim \boldsymbol{p}^3$.

- Continuity: For any two sequences $\left(\boldsymbol{p}^i\right)_{i=1}^n, \left(\boldsymbol{q}^i\right)_{i=1}^n \in \mathcal{P}$, if $\boldsymbol{p}^i \succsim \boldsymbol{q}^i$ $(\forall i = 1, 2, \cdots, n)$ and $\lim_{n\to\infty} \boldsymbol{p}^i = \boldsymbol{p}^*$, $\lim_{n\to\infty} \boldsymbol{q}^i = \boldsymbol{q}^*$, then $\boldsymbol{p}^* \succsim \boldsymbol{q}^*$.

Remark.

> There is an alternative definition of continuity. Let $\succsim$ be a complete, transitive preference relation on $\mathcal{P}$. $\succsim$ is *continuous* if for any $\boldsymbol{p}^1, \boldsymbol{p}^2, \boldsymbol{p}^3 \in \mathcal{P}$ such that $\boldsymbol{p}^1 \succsim \boldsymbol{p}^2 \succsim \boldsymbol{p}^3$, there exists $t \in [0,1]$ such that $\boldsymbol{p}^2 \sim t\boldsymbol{p}^1 + (1-t)\,\boldsymbol{p}^3$.

To ensure an expected utility representation, we need an additional restriction.

Definition 5.1.4: Independence

A preference relation $\succsim$ on $\mathcal{P}$ satisfies *independence* if for any $\boldsymbol{p}^1, \boldsymbol{p}^2, \boldsymbol{p}^3 \in \mathcal{P}$ and $t \in (0,1)$,

$$\boldsymbol{p}^1 \succsim \boldsymbol{p}^2 \iff t\boldsymbol{p}^1 + (1-t)\,\boldsymbol{p}^3 \succsim t\boldsymbol{p}^2 + (1-t)\,\boldsymbol{p}^3$$

Independence says that mixing both lotteries with a common third lottery $\boldsymbol{p}^3$ does not flip the ranking. Note that $t\boldsymbol{p}^1 + (1-t)\boldsymbol{p}^3$ and $t\boldsymbol{p}^2 + (1-t)\boldsymbol{p}^3$ are formally *compound* lotteries; under the consequentialist assumption, we identify each with the simple lottery it induces.

Moreover, if the preference relation $\succsim$ on $\mathcal{P}$ satisfies independence, then for any $\boldsymbol{p}^1, \boldsymbol{p}^2, \boldsymbol{p}^3 \in \mathcal{P}$ and $t \in (0,1)$, we have

- $\boldsymbol{p}^1 \succ \boldsymbol{p}^2 \iff t\boldsymbol{p}^1 + (1-t)\,\boldsymbol{p}^3 \succ t\boldsymbol{p}^2 + (1-t)\,\boldsymbol{p}^3$
- $\boldsymbol{p}^1 \sim \boldsymbol{p}^2 \iff t\boldsymbol{p}^1 + (1-t)\,\boldsymbol{p}^3 \sim t\boldsymbol{p}^2 + (1-t)\,\boldsymbol{p}^3$
- If $\boldsymbol{p}^1 \succsim \boldsymbol{p}^2$ and $\boldsymbol{p}^3 \succsim \boldsymbol{p}^4$, then $t\boldsymbol{p}^1 + (1-t)\,\boldsymbol{p}^3 \succsim t\boldsymbol{p}^2 + (1-t)\,\boldsymbol{p}^4$.

In summary, the four assumptions we maintain throughout this chapter — completeness, transitivity, continuity, and independence — are exactly what is needed for an expected utility representation.

Assumption 5.1.5

- **Completeness**: For any $\boldsymbol{p}^1, \boldsymbol{p}^2 \in \mathcal{P}$, either $\boldsymbol{p}^1 \succsim \boldsymbol{p}^2$ or $\boldsymbol{p}^2 \succsim \boldsymbol{p}^1$ (or both).
- **Transitivity**: For any $\boldsymbol{p}^1, \boldsymbol{p}^2, \boldsymbol{p}^3 \in \mathcal{P}$, if $\boldsymbol{p}^1 \succsim \boldsymbol{p}^2$ and $\boldsymbol{p}^2 \succsim \boldsymbol{p}^3$, then $\boldsymbol{p}^1 \succsim \boldsymbol{p}^3$.
- **Continuity**: For any two sequences $\left(\boldsymbol{p}^i\right)_{i=1}^n, \left(\boldsymbol{q}^i\right)_{i=1}^n \in \mathcal{P}$, if $\boldsymbol{p}^i \succsim \boldsymbol{q}^i$ for all i and $\lim_{n\to\infty} \boldsymbol{p}^i = \boldsymbol{p}^*$, $\lim_{n\to\infty} \boldsymbol{q}^i = \boldsymbol{q}^*$, then $\boldsymbol{p}^* \succsim \boldsymbol{q}^*$.
- **Independence**: For any $\boldsymbol{p}^1, \boldsymbol{p}^2, \boldsymbol{p}^3 \in \mathcal{P}$ and any $t \in (0,1)$,

$$\boldsymbol{p}^1 \succsim \boldsymbol{p}^2 \iff t\boldsymbol{p}^1 + (1-t)\boldsymbol{p}^3 \succsim t\boldsymbol{p}^2 + (1-t)\boldsymbol{p}^3.$$

5.2 Expected Utility Representation

Before starting the representation theorem, we first formalize what we mean by an expected utility function.

Definition 5.2.1: Expected Utility Form

A utility function $U : \mathcal{P} \to \mathbb{R}$ has an *expected utility form* (or a *von Neumann-Morgenstern expected utility function*) if there is an assignment of numbers $(u_1, u_2, \cdots, u_n)$ to each of the n sure outcomes $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_n$ such that for any $\boldsymbol{p} \in \mathcal{P}$,

$$U(\boldsymbol{p}) = \sum_{i=1}^{n} p_i u_i.$$

A sure outcome $\boldsymbol{x}_i$ can be represented by a simple lottery $\mathbf{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0)$ and we have $U(\mathbf{e}_i) = u(\boldsymbol{x}_i) = u_i$. Notice that an expected utility function can be written as

$$U(\boldsymbol{p}) = \sum_{i=1}^{n} p_i U(\mathbf{e}_i).$$

Thus, $U(\boldsymbol{p})$ is linear in the probabilities. The following proposition reveals that this observation rings true more generally.

Proposition 5.2.2

A linear function $U : \mathcal{P} \to \mathbb{R}$ has an expected utility form if and only if it is *linear*, that is, for any k (simple) lotteries $\boldsymbol{p}^j \in \mathcal{P}, (j = 1, 2, \cdots, k)$ and associated probabilities $\alpha_j \geq 0$ with $\sum_{j=1}^{k} \alpha_j = 1$, we have

$$U\left(\sum_{j=1}^{k} \alpha_j \boldsymbol{p}^j\right) = \sum_{j=1}^{k} \alpha_j U\left(\boldsymbol{p}^j\right).$$

Proof for Proposition.

- "If" part

  Let $\boldsymbol{p}^j = \boldsymbol{e}_j$ for all $j = 1, 2, \cdots, k$. Thus, $\boldsymbol{p} = \sum_{j=1}^{k} \alpha_j \boldsymbol{p}^j = (\alpha_1, \alpha_2, \cdots, \alpha_k)$. The condition of U being linear is then transformed to

  $$U(\boldsymbol{p}) = \sum_{j=1}^{k} p_j U(\mathbf{e}_j) = \sum_{j=1}^{k} \alpha_j u_j.$$

  This directly implies that U has an expected utility form.

- "Only if" part

  - Since the decision-maker is a consequentialist, the compound lottery $\sum_{j=1}^{k} \alpha_j \circ \boldsymbol{p}^j$

is identical to the simple lottery it induces,

$$\alpha_1 \circ \boldsymbol{p}^1 + \alpha_2 \circ \boldsymbol{p}^2 + \cdots + \alpha_k \circ \boldsymbol{p}^k \iff \left(\sum_{j=1}^k \alpha_j p_1^j, \sum_{j=1}^k \alpha_j p_2^j, \cdots, \sum_{j=1}^k \alpha_j p_n^j \right)$$

– Since the utility function has the expected utility form, we have

$$U\left(\sum_{j=1}^k \alpha_j p_1^j, \sum_{j=1}^k \alpha_j p_2^j, \cdots, \sum_{j=1}^k \alpha_j p_n^j \right) = \sum_{i=1}^n \left(\sum_{j=1}^k \alpha_j p_i^j \right) u_i$$

– On the other hand, $U\left(\boldsymbol{p}^j\right) = \sum_{i=1}^n p_i u_i$, so

$$\sum_{j=1}^k \alpha_j U\left(\boldsymbol{p}^j\right) = \sum_{j=1}^k \alpha_j \left(\sum_{i=1}^n p_i^j u_i \right)$$

– Jointly, we have proved the result:

$$\begin{aligned} U\left(\sum_{j=1}^k \alpha_j \boldsymbol{p}^j \right) &= U\left(\sum_{j=1}^k \alpha_j p_1^j, \sum_{j=1}^k \alpha_j p_2^j, \cdots, \sum_{j=1}^k \alpha_j p_n^j \right) \\ &= \sum_{i=1}^n \left(\sum_{j=1}^k \alpha_j p_i^j \right) u_i \\ &= \sum_{j=1}^k \alpha_j U\left(\boldsymbol{p}^j\right) \end{aligned}$$

Theorem 5.2.3

A complete and transitive preference relation $\succsim$ on $\mathcal{P}$ satisfies continuity and independence if and only if it admits an expected utility representation $U : \mathcal{P} \to \mathbb{R}$.

Proof for Theorem

The "if" part is easy to verify. We mainly focus on the "only if" part.

Suppose the complete and transitive preference relation $\succsim$ satisfies continuity and independence. We construct an expected utility function $U\left(\boldsymbol{p}\right) = \sum_{i=1}^n p_i u_i$ that represents $\succsim$ as follows:

- First, with a slight abuse of notation, interpret the consequences $x_1, \cdots, x_n$ as degenerate lotteries, so that x_i puts probability 1 on x_i and probability 0 on all other consequences. Then any lottery p can be viewed as a mixture of those degenerate lotteries: $p = \sum_i p_i x_i$.
- Label the consequences so that (when viewed as degenerate lotteries) x_1 is the best and x_n is the worst, i.e., $x_1 \succsim x_i \succsim x_n$ for all i.
- By continuity, for each consequence x_i, there is some $\lambda_i \in [0, 1]$ such that $x_i \sim \lambda_i x_1 +$

$(1-\lambda_i)\, x_n$.

- By independence (iterated n times for each consequence), for any lottery $p = \sum_i p_i x_i$,

$$p \sim \sum_i p_i \left(\lambda_i x_1 + (1-\lambda_i)\, x_n\right) = \left(\sum_i p_i \lambda_i\right) x_1 + \left(1 - \sum_i p_i \lambda_i\right) x_n$$

  – If $x_1 \sim x_n$ then by independence, for all p we have

$$p \sim \left(\sum_i p_i \lambda_i\right) x_1 + \left(1 - \sum_i p_i \lambda_i\right) x_1 = x_1$$

  Hence, letting $u_i = 0$ for all i gives an expected utility representation of $\succsim$.

  – If $x_1 \succ x_n$, then for $1 \geq \lambda \geq \lambda' \geq 0$, letting $\delta = \lambda - \lambda' \in [0,1]$ and $\hat{p} = \frac{\lambda'}{1-\delta} x_1 + \frac{1-\lambda}{1-\delta} x_n$ and using the Independence Axiom,

$$\begin{aligned} \lambda x_1 + (1-\lambda)\, x_n &= \delta x_1 + (1-\delta)\, \hat{p} \\ &\succ \delta x_1 + (1-\delta)\, \hat{p} \\ &= \lambda' x_n + (1-\lambda')\, x_n \end{aligned}$$

  Thus, letting $u_i = \lambda_i$ yields an expected utility representation of $\succsim$.

5.3 Measures of Risk

5.3.1 Risk Attitudes

Let the set of sure outcomes $X = \mathbb{R}$. For any sure outcome $x \in X$, its utility is determined by utility function $u : X \to \mathbb{R}$. A lottery is fully characterized by a corresponding distribution function F, whose expected utility is then $U(F) = \int_X u(x)\, \mathrm{d}F(x)$. For simplicity, suppose $u(\cdot)$ is strictly increasing and continuous and $U(F) = \mathbb{E}_F[u(x)] < +\infty$.

Here we naturally extend the notion of lottery to $X = \mathbb{R}$, the infinite set, and avoid any technical issues related to the definition of lotteries on finite-many sure outcomes.

Definition 5.3.1: Risk Averse

For any non-degenerate lottery F, define its average payoff δ_{E_F} as

$$\delta_{\mathrm{E}_F} = \int_X x \mathrm{d}F(x)\,.$$

A decision-maker is strictly *risk-averse* if for any non-degenerate lottery F, the sure outcome δ_{E_F} is strictly preferred to the lottery F, i.e., $\delta_{\mathrm{E}_F} \succ F$. Or in utility terms,

$$u(\delta_{\mathrm{E}_F}) = u\left(\int_X x \mathrm{d}F(x)\right) > U(F) = \int_X u(x)\, \mathrm{d}F(x)$$

Corollary 5.3.2

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

Intuitively, a risk-averse individual has *decreasing* marginal utility — each additional dollar is worth less than the previous one. *Risk-loving* (u convex, $F \succ \delta_{\mathbb{E}[F]}$) and *risk-neutral* ($u$ linear, $F \sim \delta_{\mathbb{E}[F]}$) attitudes are defined symmetrically.

Risk attitudes contrast the utility of an average payoff with the average utility of the payoffs. The *certainty equivalent* quantifies how large the gap between the two is — in monetary units.

Definition 5.3.3: Certainty Equivalent

The *certainty equivalent* $c(F, u)$ of a money lottery is defined as

$$u(c(F, u)) = U(F).$$

Intuitively, $c(F, u)$ is the sure amount of money the agent considers equivalent to facing the risky lottery F.

Certainty equivalent depends on initial wealth. For simplicity, we assume $w_0 = 0$, i.e., the agent does not have initial wealth; then $c(F, u) \leq \mathbb{E}[F]$ if and only if the agent is risk-averse.

5.3.2 Measures of Risk Aversion

So far "risk-averse" is a yes/no property, controlled by concavity of u. The vNM expected utility is cardinally meaningful (only up to positive affine transformations), which raises the natural follow-up question: can we use u to *quantify* how risk-averse an agent is, and rank two agents accordingly? The two standard measures below do exactly that.

Consider a risk-averse agent who has an initial wealth w and faces a (small) fair gamble.

- Scenario 1: The gamble $\tilde{\varepsilon}$ is measured in terms of monetary unit. Then consider how large a has to be to make
$$u(w - a) = U(w + \tilde{\varepsilon}).$$
- Scenario 2: The gamble $\tilde{\delta}$ is measured in terms of percentage of the initial wealth. Then consider how large r has to be to ensure
$$u((1 - r) w) = U\left(\left(1 + \tilde{\delta}\right) w\right).$$

Since the agent is risk-averse, $a, r > 0$. Intuitively, a measures how much money the agent is willing to give up to avoid the gamble; r measures the percentage of initial wealth the agent is willing to give up to avoid the gamble. If the agent were risk-neutral, $a = r = 0$. As a result, given initial wealth w and the small fair gamble, both a and r measure the agent's level of risk aversion.

Coefficient of Absolute Risk Aversion When $\tilde{\varepsilon}$ is small, a is small. By Taylor expansion and definition of expected utility:

$$\begin{aligned} u(w-a) &\approx u(w) - u'(w)\cdot a \\ U(w+\tilde{\varepsilon}) &= \mathbb{E}_{\varepsilon}[u(w+\varepsilon)] = \int_{\varepsilon} u(w+\varepsilon)\,\mathrm{d}\varepsilon \\ &\approx \int_{\varepsilon}\left[u(w) + u'(w)\cdot\varepsilon + \frac{1}{2}u''(w)\cdot\varepsilon^2\right]\mathrm{d}\varepsilon \\ &= u(w)\int_{\varepsilon} 1\mathrm{d}\varepsilon + u'(w)\int_{\varepsilon}\varepsilon\mathrm{d}\varepsilon + \frac{1}{2}u''(w)\int_{\varepsilon}\varepsilon^2\mathrm{d}\varepsilon \\ &= u(w) + \frac{1}{2}u''(w)\cdot\mathrm{Var}[\tilde{\varepsilon}] \end{aligned}$$

It follows that

$$\begin{aligned} u(w) - u'(w)\cdot a &= u(w) + \frac{1}{2}u''(w)\cdot\mathrm{Var}[\tilde{\varepsilon}] \\ \Longrightarrow a &\approx -\frac{1}{2}\cdot\frac{u''(w)}{u'(w)}\cdot\mathrm{Var}[\tilde{\varepsilon}] \end{aligned}$$

Definition 5.3.4: Coefficient of Absolute Risk Aversion

Suppose the agent has initial wealth $x > 0$, and the utility function $u:\mathbb{R}\to\mathbb{R}$ is twice continuously differentiable. Then the *coefficient of absolute risk aversion*, $A(x,u)$, is defined as

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

Coefficient of Relative Risk Aversion Similarly, we use Taylor expansion and definition of expected utility to compute r:

$$\begin{aligned} u(w(1-r)) &\approx u(w) - u'(w)\cdot(wr) \\ U\left(w+w\tilde{\delta}\right) &= \mathbb{E}_G[u(w+w\delta)] = \int_{\delta} u(w+w\delta)\,\mathrm{d}G(\delta) \\ &\approx \int\left[u(w) + u'(w)\cdot(w\delta) + \frac{1}{2}u''(w)\cdot(w\delta)^2\right]\mathrm{d}G(\delta) \\ &= u(w)\int 1\mathrm{d}G(\delta) + u'(w)\int_{\delta} w\delta\mathrm{d}G(\delta) + \frac{1}{2}u''(w)\int_{\delta}(w\delta)^2\,\mathrm{d}G(\delta) \\ &= u(w) + \frac{1}{2}u''(w)\cdot w^2\mathrm{Var}\left[\tilde{\delta}\right] \end{aligned}$$

It follows that

$$\begin{aligned} u((1-r)w) &= U\left(\left(1+\tilde{\delta}\right)w\right) \\ \Longrightarrow r &\approx -\frac{wu''(w)}{2u'(w)}\mathrm{Var}\left[\tilde{\delta}\right] \end{aligned}$$

Definition 5.3.5: Coefficient of Relative Risk Aversion

Suppose the agent has initial wealth $x > 0$, and the utility function $u : \mathbb{R} \to \mathbb{R}$ is twice continuously differentiable. Then the *coefficient of relative risk aversion*, $R(x, u)$, is defined as

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

Proposition 5.3.6: Equivalence of Different Measures of Risk Aversion

Suppose utility functions u and v are strictly increasing and twice differentiable. The following definitions of an agent characterized by u being more risk averse than another agent characterized by v are equivalent:

1. For any lottery F and sure outcome δ_X. if $F \succsim_u \delta_X$, then $F \succsim_v \delta_X$.
2. For any lottery F, $c(F, u) \leq c(F, v)$.
3. The function u is "more concave" than v, that is, there exists some increasing and concave function g such that $u = g \circ v$.
4. $r(x) = \dfrac{u'(x)}{v'(x)}$ is non-increasing in x.
5. For any x, $A(x, u) = -\dfrac{u''(x)}{u'(x)} \geq -\dfrac{v''(x)}{v'(x)} = A(x, v)$.

Proof for Proposition.

- (1) $\Longleftrightarrow$ (2)
  - (1) $\Longrightarrow$ (2): By definition of certainty equivalent, $\delta_{c(F,u)} \sim_u F$. By condition (1), we have $\delta_{c(F,v)} \sim_v F \succsim_v \delta_{c(F,u)}$. Recall that $v(\cdot)$ is strictly increasing, so we have $c(F, u) \leq c(F, v)$.
  - (2) $\Longrightarrow$ (1): If $\delta_{c(F,u)} \sim F \succsim_u \delta_x$, then by strictly-increasing property of u, it follows that $c(F, u) \geq x$. So $c(F, v) \geq c(F, u) \geq x \Longrightarrow c(F, v) \geq x$. Again by v being strictly increasing, we have $c(F, v) \sim F \succsim_v \delta_x$.
- (2) $\Longleftrightarrow$ (3)
  - Since u and v are strictly increasing, v^{-1} is well defined and $g = u \circ v^{-1}$ is strictly increasing.
  - Again by strict monotonicity of u, we have $u(c(F, u)) \leq u(c(F, v))$ for any lottery F.
  $$u(c(F, u)) = U(F) = \int_X u(x)\, \mathrm{d}F(x) = \int_X g(v(x))\, \mathrm{d}F(x)$$
  $$u(c(F, v)) = g(v(c(F, v))) = g\left(\int_X v(x)\, \mathrm{d}F(x)\right)$$
  - By Jensen's inequality, $\int_X g(v(x))\, \mathrm{d}F(x) \leq g\left(\int_X v(x)\, \mathrm{d}F(x)\right)$ if and only if g is concave.

- (3) $\Longleftrightarrow$ (4)
  - Since u and v are strictly increasing, v^{-1} is well defined and $g = u \circ v^{-1}$ is strictly increasing.
  - When u and v are both differentiable, $u'(x) = g'(v(x)) \cdot v'(x) \Longrightarrow \frac{u'(x)}{v'(x)} = g'(v(x))$.
  - Since v is strictly increasing, $\frac{u'(x)}{v'(x)}$ is non-increasing in x if and only if g' is non-increasing, that is, g is concave.
- (4) $\Longleftrightarrow$ (5)
  - When $r(x) > 0$, $r(x)$ is non-increasing in x if and only if $\ln r(x) = \ln u'(x) - \ln v'(x)$ is non-increasing in x.
  - When u and v are twice differentiable,

$$(\ln r(x))' = \frac{u''(x)}{u'(x)} - \frac{v''(x)}{v'(x)}$$

  - It follows that $r(x) = \frac{u'(x)}{v'(x)}$ is non-increasing in x if and only if $A(x, u) \geq A(x, v)$.

Example.

Consider an investor who must decide how much of their initial wealth w to put into a risky asset. The risky asset can have any of the positive or negative rates of return r_i with probabilities p_i, $i = 1, 2, \cdots, n$. Suppose the investor is an expected utility maximizer and their utility for x amount of money for sure can be represented by a twice continuously differentiable, strictly increasing and strictly concave utility function $u(x)$. Let a^* be the investor's optimal amount of money to put in the risky asset. Suppose $\sum_{i=1}^{n} p_i r_i > 0$.

1. Give a sufficient condition for $a^* < w$.
2. Suppose $a^* < w$ and that $A(x, u)$ is strictly decreasing in x, how would $a^*(w)$ change with w?

5.4 Comparison of Risky Prospects

The coefficient of absolute risk aversion $A(x, u)$ lets us compare agents (or the same agent at different wealth levels). The natural next question is about comparing *lotteries*: when can we say $F \succsim G$ for *every* (risk-averse) decision-maker, regardless of the specific shape of u?

A monetary lottery is summarized by two features that any agent will care about: its expected payoff and its dispersion around that expectation. This suggests two unambiguous cases in which F should be ranked above G:

1. F gives a (weakly) higher payoff than G at every probability level — F "shifts the distribution to the right."

2. F and G have the same expected payoff, but F is less dispersed.

These two cases correspond to *first-order* and *second-order* stochastic dominance respectively.

5.4.1 First-Order Stochastic Dominance

Proposition 5.4.1: First-Order Stochastic Dominance

The following two conditions are equivalent:

1. For any non-decreasing $u : \mathbb{R} \to \mathbb{R}$,

$$\int_X u(t)\,\mathrm{d}F(t) \geq \int_X u(t)\,\mathrm{d}G(t)\,.$$

2. $F(x) \leq G(x)$ for almost every x.

In this case, F *first-order stochastically dominates* G.

Proof for Proposition.

Assume that u is continuously differentiable, and $u'(\cdot) > 0$. Let $X = [\underline{x}, \bar{x}]$.

$$\begin{aligned}
&\int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}F(t) - \int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}G(t) \\
&= \left[u(t)\,F(t) \Big|_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} u'(t)\,F(t)\,\mathrm{d}t \right] - \left[u(t)\,G(t) \Big|_{\underline{x}}^{\bar{x}} - \int_{\underline{x}}^{\bar{x}} u'(t)\,G(t)\,\mathrm{d}t \right] \\
&= \left[(u(\underline{x}) - u(\bar{x})) - \int_{\underline{x}}^{\bar{x}} u'(t)\,F(t)\,\mathrm{d}t \right] - \left[(u(\underline{x}) - u(\bar{x})) - \int_{\underline{x}}^{\bar{x}} u'(t)\,G(t)\,\mathrm{d}t \right] \\
&= \int_{\underline{x}}^{\bar{x}} u'(t)\,(G(t) - F(t))\,\mathrm{d}t
\end{aligned}$$

It follows that $\int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}F(t) \geq \int_{\underline{x}}^{\bar{x}} u(t)\,\mathrm{d}G(t)$ if and only if $G(x) \geq F(x)$ almost everywhere. ∎

The first condition is the universal characterization: every decision-maker who prefers more money to less ranks F above G, regardless of their risk attitude or the specific shape of u.

5.4.2 Second-Order Stochastic Dominance

Definition 5.4.2: Mean-Preserving Spread

Let X, Y be two random variables with distribution functions F and G, respectively. Then G is a *mean-preserving spread* of F if $Y = X + \tilde{\varepsilon}$, where $\mathbb{E}[\tilde{\varepsilon}|X] = 0$.

A mean-preserving spread keeps the expected payoff fixed but adds noise to the realization — G is obtained from F by spreading out the probability mass without shifting its mean.

Proposition 5.4.3: Second-Order Stochastic Dominance

Suppose $\mathbb{E}[F] = \mathbb{E}[G]$. Then the following three conditions are equivalent:

1. For any (non-decreasing) concave $u : \mathbb{R} \to \mathbb{R}$,

$$\int_X u(t)\, \mathrm{d}F(t) \geq \int_X u(t)\, \mathrm{d}G(t)$$

2. For almost every $x \in X$,

$$\int_{-\infty}^{x} F(t)\, \mathrm{d}t \leq \int_{-\infty}^{x} G(t)\, \mathrm{d}t$$

3. G is a mean-preserving spread of F.

In this case, F *second-order stochastically dominates* G.

Proof for Proposition.

For simplicity, suppose $X = [\underline{x}, \bar{x}]$ and u is twice continuously differentiable.

- (1) $\iff$ (2)
  - By earlier calculation, $\int_{\underline{x}}^{\bar{x}} u(t)\, \mathrm{d}F(t) - \int_{\underline{x}}^{\bar{x}} u(t)\, \mathrm{d}G(t) = \int_{\underline{x}}^{\bar{x}} u'(t)\,(G(t) - F(t))\, \mathrm{d}t$.
  - Again by integration by parts and noticing $\mathbb{E}[F] = \mathbb{E}[G]$, we have:

$$\int_{\underline{x}}^{\bar{x}} u(t)\, \mathrm{d}F(t) - \int_{\underline{x}}^{\bar{x}} u(t)\, \mathrm{d}G(t) = \int_{\underline{x}}^{\bar{x}} u''(x) \left[\int_{\underline{x}}^{x} G(t)\, \mathrm{d}t - \int_{\underline{x}}^{x} F(t)\, \mathrm{d}t\right] \mathrm{d}x$$

  - Because u is concave, $u'' \leq 0$. It follows that $\int_{\underline{x}}^{\bar{x}} u(t)\, \mathrm{d}F(t) \geq \int_{\underline{x}}^{\bar{x}} u(t)\, \mathrm{d}G(t)$ if and only if $\int_{\underline{x}}^{x} F(t)\, \mathrm{d}t \geq \int_{\underline{x}}^{x} G(t)\, \mathrm{d}t$ almost everywhere.
- (1) $\iff$ (3)
  - We will only show (3) $\implies$ (1).
  - By the law of iterated expectation,

$$\mathbb{E}_G[u(Y)] = \mathbb{E}_F[\mathbb{E}_G[u(Y)\,|X]]$$

  - Since $u(\cdot)$ is concave, by Jensen's inequality,

$$\mathbb{E}_G[u(Y)] = \mathbb{E}_F[\mathbb{E}_G[u(Y)\,|X]] \leq \mathbb{E}_F[u(\mathbb{E}_G[Y|X])] = \mathbb{E}_F[u(X)]$$

∎

The first condition says that any risk-averse decision-maker would prefer lottery F over G. Additionally, when $\mathbb{E}[F] = \mathbb{E}[G]$, we actually do not need $u(\cdot)$ to be non-decreasing. If $\mathbb{E}[F] \neq \mathbb{E}[G]$, then the second and the third conditions are still equivalent, but we actually need $u(\cdot)$ to be non-decreasing and concave.

5.4.3 Likelihood Ratio Dominance

Definition 5.4.4: Likelihood Ratio Dominance

Let F and G be two distribution functions with common support $[\underline{x}, \bar{x}]$. Suppose the density functions exist and are given by f and g, respectively. F *dominates* G *in the likelihood ratio order* if $\frac{f(x)}{g(x)}$ is non-decreasing in x.

Intuitively, when F dominates G in the likelihood ratio order, F puts higher probabilities on higher returns compared with G.

Proposition 5.4.5: Likelihood Ratio Dominance Implies First-Order Stochastic Dominance

If F dominates G in the likelihood ratio order, then F first-order stochastically dominates G.

Proof for Proposition.

Since $F(\cdot)$ and $G(\cdot)$ are both non-decreasing, in order to show that $F(x) \leq G(x)$, it suffices to show that

$$\frac{F(x)}{1-F(x)} \leq \frac{G(x)}{1-G(x)}, \forall x \in (\underline{x}, \bar{x}).$$

We have

$$\begin{aligned}\frac{F(x)}{1-F(x)} &= \frac{\int_{\underline{x}}^{x} f(t)\,\mathrm{d}t}{\int_{x}^{\bar{x}} f(t)\,\mathrm{d}t} \\ &= \frac{\int_{\underline{x}}^{x} \frac{f(t)}{g(t)} \cdot g(t)\,\mathrm{d}t}{\int_{x}^{\bar{x}} \frac{f(t)}{g(t)} \cdot g(t)\,\mathrm{d}t}\end{aligned}$$

Since $\frac{f(x)}{g(x)}$ is non-decreasing,

$$\begin{aligned}\forall t \in [\underline{x}, x], \frac{f(t)}{g(t)} &\leq \frac{f(x)}{g(x)} \\ \forall t \in [x, \bar{x}], \frac{f(t)}{g(t)} &\geq \frac{f(x)}{g(x)}\end{aligned}$$

Therefore, we have

$$\frac{F(x)}{1-F(x)} = \frac{\int_{\underline{x}}^{x} \frac{f(t)}{g(t)} \cdot g(t)\,\mathrm{d}t}{\int_{x}^{\bar{x}} \frac{f(t)}{g(t)} \cdot g(t)\,\mathrm{d}t} \leq \frac{\frac{f(x)}{g(x)} \cdot \int_{\underline{x}}^{x} g(t)\,\mathrm{d}t}{\frac{f(x)}{g(x)} \cdot \int_{x}^{\bar{x}} g(t)\,\mathrm{d}t} = \frac{\int_{\underline{x}}^{x} g(t)\,\mathrm{d}t}{\int_{x}^{\bar{x}} g(t)\,\mathrm{d}t} = \frac{G(x)}{1-G(x)}$$

5.5 Comparative Statics Under Risk

Our earlier discussions on comparative statics did not specifically take uncertainty into consideration. Whether and how would the comparative statics results carry over to choices under uncertainty?

The first result is that, if a random variable is "introduced" into a supermodular function, the expected utility function still preserves supermodularity.

Lemma 5.5.1: Expected Utility Function Preserves Supermodularity

Let $X \subset \mathbb{R}^n$ be a sublattice and $T \subset \mathbb{R}$. Suppose $u : X \times T \to \mathbb{R}$ is supermodular in $\boldsymbol{x}$. Then for any distribution function F on T, the function $U : X \to \mathbb{R}$ defined by

$$U(\boldsymbol{x}) = \int_t u(\boldsymbol{x}, t)\,\mathrm{d}F(t)$$

is supermodular in $\boldsymbol{x}$.

Proof for Lemma

Take any $\boldsymbol{x}, \boldsymbol{x}' \in X$. we have

$$\begin{aligned} U(\boldsymbol{x} \wedge \boldsymbol{x}') + U(\boldsymbol{x} \vee \boldsymbol{x}') &= \int_t [u(\boldsymbol{x} \wedge \boldsymbol{x}', t) + u(\boldsymbol{x} \vee \boldsymbol{x}', t)]\,\mathrm{d}F(t) \\ &\geq \int_t [u(\boldsymbol{x}, t) + u(\boldsymbol{x}', t)]\,\mathrm{d}F(t) \\ &= U(\boldsymbol{x}) + U(\boldsymbol{x}') \end{aligned}$$

where the inequality in the second line follows from the supermodularity of $u(\boldsymbol{x}, t)$ in $\boldsymbol{x}$. Consequently by definition of supermodularity, $U(\cdot)$ is supermodular in $\boldsymbol{x}$. ■

The second result states that, first-order stochastic dominance preserves increasing differences.

Lemma 5.5.2: First-Order Stochastic Dominance Preserves Increasing Differences

Let $X \subset \mathbb{R}^n$ be a sublattice and $T, \Theta \subset \mathbb{R}$. Suppose $u : X \times T \to \mathbb{R}$ has increasing differences in $(\boldsymbol{x}, t)$, and $\{F_\theta\}_{\theta \in \Theta}$ is a family of distribution function on T such that $F_\theta \geq_{FOSD} F_{\theta'}$ if $\theta > \theta'$. Then the function $U : X \times \Theta \to \mathbb{R}$ defined by

$$U(\boldsymbol{x}, \theta) = \int_t u(\boldsymbol{x}, t)\,\mathrm{d}F_\theta(t)$$

has increasing differences in $(\boldsymbol{x}, \theta)$.

Proof for Lemma

Take any $\boldsymbol{x} > \boldsymbol{x}'$ and $\theta > \theta'$, we have

$$\begin{aligned} U(\boldsymbol{x}, \theta) - U(\boldsymbol{x}', \theta) &= \int_t [u(\boldsymbol{x}, t) - u(\boldsymbol{x}', t)]\,\mathrm{d}F_\theta(t) \\ &\geq \int_t [u(\boldsymbol{x}, t) - u(\boldsymbol{x}', t)]\,\mathrm{d}F_{\theta'}(t) \\ &= U(\boldsymbol{x}, \theta') - U(\boldsymbol{x}', \theta') \end{aligned}$$

where the inequality holds because, $\delta(t) := u(\boldsymbol{x}, t) - u(\boldsymbol{x}', t)$ is non-decreasing in t by increasing differences of $u(\cdot, \cdot)$ in $(\boldsymbol{x}, t)$ and $F_\theta \geq_{FOSD} F_{\theta'}$. Consequently, $U(x, \theta)$ has increasing differences in $(\boldsymbol{x}, \theta)$. ■

Intuitively, θ can be interpreted as a signal of t, indicating the distribution of t.

Proposition 5.5.3: Comparative Statics Under Risk

Let $X \subset \mathbb{R}^n$ be a sublattice and $T, \Theta \subset \mathbb{R}$. Suppose $u : X \times T \to \mathbb{R}$ is supermodular in $\boldsymbol{x}$ and has increasing differences in $(\boldsymbol{x}, t)$. Further suppose $\{F_\theta\}_{\theta \in \Theta}$ is a family of distribution function on T such that $F_\theta \geq_{FOSD} F_{\theta'}$ if $\theta > \theta'$. Then for the function $U : X \times \Theta \to \mathbb{R}$ defined by

$$U(\boldsymbol{x}, \theta) = \int_t u(\boldsymbol{x}, t) \, \mathrm{d}F_\theta(t),$$

the set of maximizers, $\arg\max_{\boldsymbol{x} \in X} U(\boldsymbol{x}, \theta)$, is non-decreasing in θ in the strong set order.

Proof for Proposition.

- Fixing any θ, by the lemma of expected utility function preserving supermodularity, $U(\boldsymbol{x}, \theta)$ is supermodular in $\boldsymbol{x}$.
- By the lemma of first-order stochastic dominance preserving increasing differences, $U(\boldsymbol{x}, \theta)$ has increasing differences in $(\boldsymbol{x}, \theta)$.
- Since $X \times \Theta$ forms a product set, $U(\boldsymbol{x}, \theta)$ is supermodular in $(\boldsymbol{x}, \theta)$.
- By the multivariate Topkis' theorem, the set of maximizers, $\arg\max_{\boldsymbol{x} \in X} U(\boldsymbol{x}, \theta)$, is non-decreasing in θ in the strong set order.

■

Example.

Suppose a monopolist faces an uncertain demand and must make a production decision prior to learning the realized demand for its product. The monopolist does learn some information about demand prior to choosing its output. Formally, suppose the inverse demand function is $p(q, t) = \hat{p}(q) + t$, where $\hat{p}(q)$ is the estimated inverse demand, and t is a random noise. The ex-post profit of the firm is therefore

$$\pi(q, t) = p(q, t) \cdot q - c(q)$$

with $c(q)$ being the monopolist's cost function. The monopolist observes a signal θ that is informative about the parameter t. Specifically, suppose the distribution of t conditional on θ is $F_\theta(\cdot)$, and $F_\theta \geq_{FOSD} F_{\theta'}$ for $\theta > \theta'$. Determine how would the monopolist's optimal output $q^*(\theta)$ change with the observed signal θ.

Solution.

The monopolist solves the following maximization problem:

$$\max_{q \geq 0} \Pi(q, \theta) = \int_t \pi(q, t) \, \mathrm{d}F_\theta(t)$$

Notice that $\frac{\partial \pi(q,t)}{\partial t} = q$, increasing in q. It follows that $\pi(q, t)$ has increasing differences in (q,t). Since $F_\theta \geq_{FOSD} F_{\theta'}$ for $\theta > \theta'$, $\Pi(q,\theta)$ has increasing differences in (q,θ) and $q^*(\theta)$ is non-decreasing in θ in the strong set order.

Example.

Consider an agent with an initial wealth w, who faces a random loss $\tilde{l} \in [0,w]$. To counter the potential loss, the agent may purchase any fraction of insurance $y \in [0,1]$. Specifically, if the agent purchases y unit of insurance, then they pay a total price of yp upfront, and get paid yl if they incur a loss of l. Suppose the agent is an expected utility maximizer and has a utility function $u(x)$ for x amount of money for sure. Further suppose $u(\cdot)$ is twice continuously differentiable with $u'(\cdot) > 0$ and $u''(\cdot) < 0$.

1. First suppose $p \leq \mathbb{E}\left[\tilde{l}\right]$. Show the optimal amount of insurance $y^* = 1$.
2. Next suppose $p > \mathbb{E}\left[\tilde{l}\right]$. Show the optimal amount of insurance $y^* < 1$.
3. Let $A(x,u)$ be the coefficient of absolute risk aversion. Write down its expression and show that if $A(x,u) = c_0$ (a constant), then the optimal amount of insurance y^* is independent of w.
4. Now suppose $A(x,u)$ strictly decreases with x. Show the optimal amount of insurance y^* is non-increasing in the agent's initial wealth w.

Solution.

1. Since the agent is an expected utility maximizer, their choice problem is given by

$$\max_{y\in[0,1]} U(y) = \mathbb{E}\left[u\left(w - yp - (1-y)l\right)\right]$$

Simple calculation shows

$$\begin{aligned} U'(y) &= \mathbb{E}\left[(l-p)\cdot u'\left(w - l + (l-p)y\right)\right] \\ U''(y) &= \mathbb{E}\left[(l-p)^2 \cdot u''\left(w - l + (l-p)y\right)\right] \leq 0 \end{aligned}$$

It follows that $U'(y) \geq U'(1) = u'(w-p)\left(\mathbb{E}\left[\tilde{l}\right] - p\right) \geq 0$, with strict inequality as long as $p < \mathbb{E}\left[\tilde{l}\right]$. Consequently, $y^* = 1$.

2. When $p > \mathbb{E}\left[\tilde{l}\right]$, $U'(1) = u'(w-p)\left(\mathbb{E}\left[\tilde{l}\right] - p\right) < 0$, and $U''(y) \leq 0$, we have $y^* < 1$.

3. By definition, $A(x,u) = -\dfrac{u''(x)}{u'(x)}$.

   Take any $w_1 < w_2$, and let

$$\begin{aligned} v_1(x) &= u(w_1 + x) \\ v_2(x) &= u(w_2 + x) \end{aligned}$$

Then at initial wealth w_i $(i = 1, 2)$, the agent's maximization problem is given by

$$\max_{y\in[0,1]} V_i(y) = \mathbb{E}\left[v_i(-l + (l-p)y)\right]$$

From the previous two questions, it suffices to focus on the case of $p > \mathbb{E}\left[\tilde{l}\right]$.

By earlier characterization, $A(x,u) = c_0$ implies $\dfrac{v_1'(x)}{v_2'(x)} = c_1$, which is also a constant.

Simple calculation shows

$$\begin{aligned} V_i'(y) &= \mathbb{E}\left[(l-p)\, v_i'(-l + (l-p)y)\right] \\ V_i''(y) &= \mathbb{E}\left[(l-p)^2\, v_i''(-l + (l-p)y)\right] \le 0 \end{aligned}$$

If $V_1'(0) \le 0$, then $V_2'(0) \le 0$, so $y_1^* = y_2^* = 0$. Otherwise, if $V_1'(y^*) = 0$, then $V_2'(y^*) = 0$; by $V_i''(\cdot) < 0$ for $i = 1, 2$, $V_i'(y)$ strictly decreases with y and achieves maximum at y, so $y_1^* = y_2^* = y^*$.

4. Suppose by contradiction that $y_2^* > y_1^* \ge 0$. By optimality of y_2^*,

$$\int_0^w v_2'(-l + (l-p)y_2^*)(l-p)\,\mathrm{d}F(l) = 0.$$

By our earlier characterization, $A(x,u)$ strictly decreases with x implies $\dfrac{v_1'(x)}{v_2'(x)}$ is non-increasing in x.

Consequently, we have

$$\begin{aligned} V_1'(y_2^*) &= \int_0^p v_1'(-l + (l-p)y_2^*)(l-p)\,\mathrm{d}F(l) \\ &\quad + \int_p^w v_1'(-l + (l-p)y_2^*)(l-p)\,\mathrm{d}F(l) \\ &= \int_0^p \frac{v_1'(-l + (l-p)y_2^*)}{v_2'(-l + (l-p)y_2^*)} v_2'(-l + (l-p)y_2^*)(l-p)\,\mathrm{d}F(l) \\ &\quad + \int_p^w \frac{v_1'(-l + (l-p)y_2^*)}{v_2'(-l + (l-p)y_2^*)} v_2'(-l + (l-p)y_2^*)(l-p)\,\mathrm{d}F(l) \\ &\ge \frac{v_1'(-p)}{v_2'(-p)} \int_0^w v_2'(-l + (l-p)y_2^*)(l-p)\,\mathrm{d}F(l) \\ &= 0 \end{aligned}$$

Given $V_i''(\cdot) < 0$ and our assumption of $y_2^* > y_1^* \ge 0$, we have $V_1'(y_1^*) > V_1'(y_2^*) \ge 0$, which is a contradiction.