1. CHOICE UNDER UNCERTAINTY

You have studied choices that yield certain outcomes. We now consider choice under uncertainty. We shall first discuss how to represent alternatives that have uncertain outcomes and our basic assumptions. We next establish the expected utility theorem, which plays a central role in the theory of decision under certainty. We then introduce the concepts of risk and its measurements, and discuss some applications of the theory we have developed. We will conclude with a discussion of comparing payoff distributions in terms of returns and risk.

1.1. Representation of and Preferences over Risky Alternatives

Let the set of all possible outcomes, indexed by \( n = 1, 2, \ldots, N \), be \( C \). A decision maker needs to choose among a number of risky alternatives that may result in outcome \( n \) in \( C \) with some known probabilities. We call a risky alternative a lottery, \( L \).

**Definition 1** A simple lottery \( L \) is a list \( L = (p_1, \ldots, p_N) \) with \( p_n \) being the probability of outcome \( n \) occurring, \( p_n \geq 0 \) for \( n = 1, \ldots, N \) and \( \sum_n p_n = 1 \).

We can consider \( L \) as a point in the \((N - 1)\) dimensional simplex, \( \Delta = \{ p \geq 0 : p_1 + \ldots + p_N = 1 \} \).

**Definition 2** Given \( k \) simple lotteries \( L_k = (p^k_1, \ldots, p^k_N) \), \( k = 1, \ldots, K \), and probabilities \( \alpha_k \), the compound lottery \( (L_1, \ldots, L_K; \alpha_1, \ldots, \alpha_K) \) is the risky alternative that yields simple lottery \( L_k \) with probability \( \alpha_k \).

The reduced lottery of any compound lottery is

\[
L = \alpha_1 L_1 + \ldots + \alpha_K L_K,
\]
where
\[ p_n = \alpha_1 p_{n1} + \ldots + \alpha_K p_{nK}. \]

The consequentialist Premise: For any risky alternative, only the reduced lottery over final outcomes is of relevance to the decision maker. With this assumption, we can restrict our analysis of lotteries to the set of all simple lotteries, \( \mathcal{L} \). We next assume that the decision maker has a rational preference relation \( \succeq \) on \( \mathcal{L} \), and \( \succeq \) is complete, transitive, and continuous. We further assume that preference satisfies the independence axiom:

**Definition 3** The preference relation on \( \mathcal{L} \) satisfies the independence axiom if for all \( L, L', L'' \in \mathcal{L} \), and \( \alpha \in (0,1) \), we have
\[ L \succeq L' \text{ iff } \alpha L + (1 - \alpha) L'' \succeq \alpha L' + (1 - \alpha) L''. \]

The independence axiom says that the preference ordering of two lotteries is not changed if each of them is mixed with a third lottery in the same way. The I.A. plays a central role in the theory of choice under uncertainty. Intuitively, we can think of an experiment where event 1 happens with probability \( \alpha \) and event 2 happens with probability \( 1 - \alpha \), \( \alpha L + (1 - \alpha) L'' \) as the compound lottery where \( L \) is chosen if event 1 happens and \( L'' \) is chosen if event 2 happens, and \( \alpha L' + (1 - \alpha) L'' \) as the compound lottery where \( L' \) is chosen if event 1 happens and \( L'' \) is chosen if event 2 happens. Now conditional on event 1 happens, \( \alpha L + (1 - \alpha) L'' \) is at least as good as \( \alpha L' + (1 - \alpha) L'' \); and conditional on event 2 happens, these two lotteries yield the same result. The I.A. requires that \( \alpha L + (1 - \alpha) L'' \succeq \alpha L' + (1 - \alpha) L'' \). Notice that the decision maker does not actually consumer both \( L \) and \( L'' \) for the compound lottery \( \alpha L + (1 - \alpha) L'' \); he obtains either \( L \) or \( L'' \). This is different from a consumer who mixes a consumption bundle \( x \) with \( x'' \), in which case he then consumers both \( x \) and \( x'' \), and \( x \succeq x' \) need not imply \( xUx'' \succeq x'Ux'' \). The I.A. explores the unique feature of the decision problem over lotteries.
Definition 4 The utility function $U : \mathcal{L} \rightarrow \mathbb{R}$ has an expected utility form if there is an assignment of numbers $(u_1, ..., u_N)$ to the $N$ outcomes such that for every $L \in \mathcal{L}$ we have

$$U(L) = u_1p_1 + ... + u_Np_N.$$

A utility function $U : \mathcal{L} \rightarrow \mathbb{R}$ with the expected utility form is called a von Neumann-Morgenstern (v.N-M) expected utility function.

Let $L^n$ be the lottery that yields outcome $n$ with probability 1. Then $U(L^n) = u_n$, and the expected utility form can be written as $U(L) = \sum_n p_n U(L^n)$. An important property of the expected utility form is its lineality. We have:

Proposition 1 A utility function $U : \mathcal{L} \rightarrow \mathbb{R}$ has an expected utility form if and only if it is linear, that is, if and only if it satisfies

$$U\left(\sum_k \alpha_k L_k\right) = \sum_k \alpha_k U(L_k) \quad (1)$$

for any $K$ lotteries $L_k \in \mathcal{L}$ and probabilities $\alpha_k$.

Proof. If $U(\cdot)$ satisfies (1), then for any $L = (p_1, ..., p_N) \in \mathcal{L}$,

$$U(L) = U\left(\sum_n p_n L^n\right) = \sum_n p_n U(L^n) = \sum_n p_n u_n,$$

and thus $U(\cdot)$ has the expected utility form.

In the other direction, suppose $U(\cdot)$ has the expected utility form. Consider any compound lottery $(L_1, ..., L_K; \alpha_1, ..., \alpha_K)$, and its reduced lottery $L = \sum_k \alpha_k L_k$. Notice that outcome $n$ occurs with probability $p_n^k$ in $L_k$, and $\sum_k \alpha_k p_n^k$ in $L$. Thus

$$U\left(\sum_k \alpha_k L_k\right) = \sum_n u_n \left(\sum_k \alpha_k p_n^k\right) = \sum_k \alpha_k \left(\sum_n u_n p_n^k\right) = \sum_k \alpha_k U(L_k).$$
Another important property of the expected utility form is that it is preserved by increasing linear transformations. That is, if $U : \mathcal{L} \rightarrow R$ is a v.N-M expected utility function, then $\tilde{U} : \mathcal{L} \rightarrow R$ is another v.N-M expected utility function if $\tilde{U}(L) = \beta U(L) + \gamma$ for every $L \in \mathcal{L}$, where $\beta > 0$ and $\gamma$ are both scalars.

A consequence of the lineality of the expected utility form is that if the preference relation is represented by a utility function that has the expected utility form, then the preference relation satisfies the independence axiom. The proof is straightforward and you are asked to check this as an exercise. The expected utility theorem says that the converse is also true.

1.2. The Expected Utility Theorem

**Proposition 2** Suppose that $\succeq$ on $\mathcal{L}$ is continuous and satisfies the independence axiom. Then $\succeq$ admits a utility representation of the expected utility form. That is, we can assign a number $u_n$ to each outcome $n = 1, \ldots, N$ in such a manner that for any two lotteries $L = (p_1, \ldots, p_N)$ and $L' = (p'_1, \ldots, p'_N)$, we have

$$L \succeq L' \text{ if and only if } \sum u_n p_n \geq \sum u_n p'_n.$$ 

**Proof.** We proceed in six steps.

Step 1. One can verify that there are $\bar{L}$ and $\underline{L}$ in $\mathcal{L}$ such that $\bar{L} \succeq L \succeq \underline{L}$ for all $L \in \mathcal{L}$. This is left as your exercise. Now if $\bar{L} \sim \underline{L}$, the conclusion holds trivially. So assume $\bar{L} \succ \underline{L}$.

Step 2. If $L \succ L'$ and $\alpha \in (0, 1)$, then $L \succ \alpha L + (1 - \alpha)L'$. This is because

$$L = (1 - \alpha)L + \alpha L \succ (1 - \alpha)L' + \alpha L \succ (1 - \alpha)L' + \alpha L' = L',$$

where the $\succ$ relations follow from the independence axiom.

Step 3. Let $\alpha, \beta \in [0, 1]$. Then $\beta \bar{L} + (1 - \beta)\underline{L} \succ \alpha \bar{L} + (1 - \alpha)\underline{L}$ if and only if $\beta > \alpha$.
If $\beta > \alpha$, 

$$
\beta L + (1 - \beta)L = \frac{\beta - \alpha}{1 - \alpha} L + \left( 1 - \frac{\beta - \alpha}{1 - \alpha} \right) (\alpha L + (1 - \alpha)L) > \alpha L + (1 - \alpha)L
$$

by repeatedly applying Step 2.

For the other direction, suppose that $\beta \leq \alpha$. A similar argument as above implies that $\beta L + (1 - \beta)L \gtrless \alpha L + (1 - \alpha)L$. This establishes that $\beta L + (1 - \beta)L > \alpha L + (1 - \alpha)L$ only if $\beta > \alpha$.

Step 4. For any $L \in \mathcal{L}$, there is a unique $\alpha_L$ such that $\alpha_L L + (1 - \alpha_L)L \sim L$.

This follows from the continuity of $\gtrless$ and the fact that $L$ and $L$ are the best and worst lottery, respectively.

Step 5. The function $U : \mathcal{L} \to R$ that assigns $U(L) = \alpha_L$ for all $L \in \mathcal{L}$ represents the preference relation $\gtrless$.

From Step 4, we know that for any $L, L' \in \mathcal{L}$, $L \gtrless L'$ if and only if 

$$
\alpha_L L + (1 - \alpha_L)L \gtrless \alpha_{L'} L + (1 - \alpha_{L'})L.
$$

Thus, from step 3, $L \gtrless L'$ if and only if $U(L) = \alpha_L \geq U(L') = \alpha_{L'}$.

Step 6. The utility function $U(\cdot)$ that assigns $U(L) = \alpha_L$ is linear and therefore has the expected utility form.

We need to show that for any $L, L' \in \mathcal{L}$ and $\beta \in [0, 1]$, we have

$$
U(\beta L + (1 - \beta)L') = \beta U(L) + (1 - \beta)U(L').
$$

By definition, $L \sim U(L) L + (1 - U(L)) L$, and $L' \sim U(L') L + (1 - U(L')) L$.

By the independence axiom, we then have:

$$
\beta L + (1 - \beta)L' \sim \beta (U(L) L + (1 - U(L)) L) + (1 - \beta)L'
$$

$$
\sim \beta (U(L) L + (1 - U(L)) L) + (1 - \beta) (U(L') L + (1 - U(L')) L).
$$

But the last lottery above has the same reduced form as

$$
[\beta U(L) + (1 - \beta)U(L')] L + [1 - \beta U(L) - (1 - \beta)U(L')] L.
$$
Thus
\[ \beta L + (1 - \beta)L' \sim [\beta U(L) + (1 - \beta)U(L')] \bar{L} + [1 - \beta U(L) - (1 - \beta)U(L')] \bar{L}. \]

We therefore have, from the construction of \( U(\cdot) \), that
\[ \beta L + (1 - \beta)L' = \beta U(L) + (1 - \beta)U(L'). \]

The expected utility theorem is extremely useful in applications. Its underlying assumptions also seem reasonable for most economic applications. But there has also been difficulties about it. A famous example is the Allais Paradox. Suppose that there are three possible prices (outcomes): 2.5; 0.5; and 0 in million dollars. Consider two choice tests:

First: choose between \( L_1 = (0, 1, 0) \) and \( L_1' = (.10, .89, .01) \).
Second: choose between \( L_2 = (0, .11, .89) \) and \( L_2' = (.10, 0, .90) \).

It is not uncommon for people to have \( L_1 \succ L_1' \) but \( L_0 \succ L_2 \). But this is not consistent with the expected utility theorem. If there is a v.N-M utility function, denote by \( u_{25} \), \( u_{05} \), and \( u_0 \) the utility values of the three outcomes. Then \( L_1 \succ L_1' \) implies
\[ u_{05} > .10u_{25} + .89u_{05} + .01u_0. \]

Adding \( .89u_0 - .89u_{05} \) to both sides, we have
\[ .11u_{05} + .89u_0 > .10u_{25} + .9u_0, \]
and therefore any individual with v.N-M utility function must have \( L_2 \succ L_2' \).

1.3. Money Lotteries and Risk Aversion

So far, we have treated the outcome set as any abstract set. We now turn to outcomes that are amounts of money. It is more convenient to treat money as a continuous variable, and we can describe a money lottery by c.d.f. \( F : R \to [0, 1] \). That
is, for any $x$, $F(x)$ is the probability that the realized payoff is less than or equal to $x$. If $F$ admits a p.d.d. $f(x)$, then $F(x) = \int_{-\infty}^{x} f(x)dx$. The set $\mathcal{L}$ is then the set of all possible c.d.f.s. The expected utility theorem then says that there is an assignment of utility values $u(x)$ to nonnegative amount of money with the property that any $F$ can be evaluated by a utility function of the form

$$U(F) = \int u(x)dF(x).$$

Notice that the v.N-M expected utility function $U(F)$ is defined over lotteries, while $u(x)$, which is called Bernoulli utility function, is defined over sure amounts of money.

We can now discuss risk aversion and its measurements.

**Definition 5** A decision maker is risk averse if, for any lottery $F(\cdot)$, he prefers the certain amount $\int xdF(x)$ to the lottery $F(\cdot)$ itself. He is strictly risk averse if the preference holds strictly. He is risk neutral if he is always indifferent between these two lotteries.

That is, a person is risk averse if

$$\int u(x)dF(x) \leq u\left(\int xdF(x)\right) \text{ for all } F(\cdot).$$

But this is the familiar Jensen’s inequality that is the defining property of a concave function. Thus, risk aversion is equivalent to the concavity of $u(x)$, and strictly risk aversion is equivalent to the strict concavity of $u(x)$. Further, risk neutrality is equivalent to $u(\cdot)$ being linear. In what follows we shall assume that the decision maker is risk averse.

**Definition 6** The certainty equivalent of $F(\cdot)$, denoted $c(F,u)$, is the amount of money for which the individual is indifferent between the lottery $F(\cdot)$ and the certain amount of $c(F,u)$; that is,

$$u(c(F,u)) = \int u(x)dF(x).$$
For a risk averter, we have $c(F, u) \leq \int x dF(x)$ since

$$u(c(F, u)) = \int u(x) dF(x) \leq u \left( \int x dF(x) \right)$$

and $u(\cdot)$ is non-decreasing.
Example 1 *Demand for Insurance.*

(1) How much are you willing to pay for complete insurance?

Suppose that \( u(x) = \sqrt{x} \); your wealth without an accident is 100, with an accident is 64. The probability of having an accident is \( \beta \). Let \( m \) be the maximum amount of insurance premium that you are willing to pay to buy the insurance so that you will always receive 100, then

\[
\begin{align*}
  u(\beta 64 + (1 - \beta)100) &= \sqrt{64\beta + 100(1 - \beta)} \\
  U(L) &= \beta \sqrt{64} + (1 - \beta)\sqrt{100} = 8\beta + 10(1 - \beta) = 10 - 2\beta
\end{align*}
\]

\[
\begin{align*}
  \beta^2 64 + (1 - \beta)^2 100 + 2\beta(1 - \beta)80 - (64\beta + 100(1 - \beta)) &= -4\beta(1 - \beta) \\
  \sqrt{100 - m} &= 10 - 2\beta \\
  m &= 100 - (10 - 2\beta)^2 = 4\beta (10 - \beta)
\end{align*}
\]

(2) How much insurance to purchase?

A strictly risk averse person who has wealth \( w \) may suffer a loss of \( D \) with probability \( \pi \). One unit of insurance costs \( q \) dollars and pays 1 dollar if the loss occurs. If \( \alpha \) units insurance are bought, the wealth of the person will be \( w - \alpha q \) with prob. \( (1 - \pi) \) and \( w - D + \alpha - \alpha q \) with prob. \( \pi \). The person’s expected wealth is \( w - \pi D + \alpha(\pi - q) \). The person solves

\[
\max_{\alpha \geq 0} (1 - \pi) u(w - \alpha q) + \pi u(w - D + \alpha - \alpha q).
\]

In the optimum,

\[
-q(1 - \pi)u'(w - \alpha q) + (1 - q)\pi u'(w - D + \alpha - \alpha q) \leq 0 \text{ with " =" if } \alpha > 0.
\]
Now if the price \( q \) of one unit of insurance is fair, in the sense of it being equal to the expected cost of insurance, \( \pi \). The 1st order condition becomes

\[-u'(w - \alpha q) + u'(w - D + \alpha - \alpha q) \leq 0 \quad \text{with “” if } \alpha > 0.\]

Now if \( \alpha = 0 \), we would have \( u'(w - D) > u'(w) \). Thus we must have \( \alpha > 0 \). Thus

\[u'(w - \alpha q) = u'(w - D + \alpha - \alpha q),\]

which implies

\[w - \alpha q = w - D + \alpha - \alpha q,\]

or

\[\alpha^* = D.\]

Thus if insurance is fair, the person will insure completely.

**Example 2 Demand for a Risky Asset**

There are two assets: a safe asset that returns \(1 + r\) dollars for a dollar invested, and a risky asset that has a random return \(z\) for each dollar invested, with p.d.f. \(F(z)\) and \(\int z dF(z) > 1 + r\). A person needs to invest his total wealth \(w\) between the two assets, \(\alpha\) for risky asset and \(\beta\) for safe asset, with \(\alpha + \beta = w\). The problem

\[
\max_{\alpha, \beta \geq 0} \int u(\alpha z + \beta (1 + r)) \, dF(z) \quad \text{s.t. } \alpha + \beta = w
\]

or

\[
\max_{0 \leq \alpha \leq w} \int u(\alpha z + (w - \alpha) (1 + r)) \, dF(z). \]

In the optimum, \(\phi(\alpha^*) = \int u'(\alpha^* z + (w - \alpha^*) (1 + r))(z - (1 + r)) \, dF(z)\) \(\leq 0 \quad \text{if } \alpha^* < w\)

\[
\geq 0 \quad \text{if } \alpha^* > 0
\]
Since

\[ \phi(0) = \int u'(w(1+r))(z - (1 + r)) \, dF(z) = u'(w(1 + r)) \int (z - (1 + r)) \, dF(z) > 0, \]

we must have \( \alpha^* > 0 \). That is, a strictly risk averse person will always invest some portion of his wealth in risky asset if the expected return from a risky asset is higher than that of a safe asset.
There are two commonly-used measures of risk aversion.

The Arrow-Pratt coefficient of absolute risk aversion is defined as

\[
r_A(x) = -\frac{u''(x)}{u'(x)}.
\]

**Example 3** The utility function that has constant coefficient of absolute risk aversion: \( u(x) = -e^{-ax} \) for \( a > 0 \).

The notion of decreasing absolute risk aversion. This is an assumption often made in applications.

Comparisons across individuals.

Another useful concept is:

**Definition 7** The coefficient of relative risk aversion at \( x \) is \( r_R(x, u) = -x \frac{u''(x)}{u'(x)} \).

Notice that \( r_R(x, u) = x \cdot r_A(x, u) \). Thus

\[
\frac{dr_R(x, u)}{dx} = r_A(x, u) + x \frac{dr_A(x, u)}{dx}.
\]

Therefore, for a risk averter, if he has decreasing relative risk aversion, he must also have decreasing absolute risk aversion; but the converse need not be true.
1.4. Comparing Payoff Distributions in Terms of Return and Risk

When outcomes are stochastic, how can we say one distribution has higher return or lower risk than another? The concepts are first-order stochastic dominance and second-order stochastic dominance. Assume the distributions have support on $[0, \bar{x}]$ for some $\bar{x} < \infty$.

**Definition 8** The distribution $F(\cdot)$ first-order stochastically dominates $G(\cdot)$ if, for every nondecreasing function $u : R \to R$, we have

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

In applications, we often use an alternative definition, which says $F$ first-order stochastically dominates $G$ if $F(x) \leq G(x)$ for every $x$. These two definitions can be shown to be equivalent.

**Proposition 3** $F(\cdot)$ first-order stochastically dominates $G(\cdot)$ if and only if $F(x) \leq G(x)$ for every $x$.

**Proof.** Let $H(x) = F(x) - G(x)$. To show the “only if” part, suppose that there is some $\hat{x}$ such that $F(\hat{x}) > G(\hat{x})$, or $H(\hat{x}) > 0$. Consider

$$u(x) = \tilde{u}(x) = \begin{cases} 0 & \text{if } x \leq \hat{x} \\ 1 & \text{if } x > \hat{x} \end{cases}.$$  

Then clearly $\tilde{u}(x)$ is nondecreasing. But

$$\int \tilde{u}(x)dH(x) = \int_{\hat{x}}^{\infty} dH(x) = -H(\hat{x}) < 0,$$

which implies

$$\int \tilde{u}(x)dF(x) < \int \tilde{u}(x)dG(x).$$
To show next the “if” part, we have,

\[
\int u(x) dH(x) = [u(x)H(x)]_0^\infty - \int u'(x) H(x) dx = - \int u'(x) H(x) dx \geq 0
\]

if \(-\int H(x) dx \geq 0\), because \(u'(x) \geq 0\). If \(F(x) \leq G(x)\) for every \(x\), then indeed \(-\int H(x) dx \geq 0\), and thus \(\int u(x) dH(x) \geq 0\).

\[
\Box
\]

Example. Consider the class of distributions \(F_\alpha(x) = x\alpha, 0 < x < 1, 1 \leq \alpha\). If \(\alpha_1 > \alpha_2\), we have

\[x^{\alpha_1} < x^{\alpha_2},\]

and then \(F_{\alpha_1}(x)\) first-order stochastically dominates \(F_{\alpha_2}(x)\).

Example. In the literature, \(\frac{f(x)}{1-F(x)}\) is called the hazard rate. Show that if

\[
\frac{f(x)}{1-F(x)} \geq \frac{g(x)}{1-G(x)} \text{ for all } x, \text{ then } F \text{ first-order stochastically dominates } G.
\]

Since \(\frac{f(x)}{1-F(x)} \geq \frac{g(x)}{1-G(x)}\),

\[
\int_{-\infty}^x \frac{f(x)}{1-F(x)} dx \geq \int_{-\infty}^x \frac{f(x)}{1-F(x)} dx.
\]

or

\[
[\ln (1 - F(x))]_x^{-\infty} \geq [\ln (1 - G(x))]_x^{-\infty}.
\]

That is,

\[
0 - \ln (1 - F(x)) \geq 0 - \ln (1 - G(x))
\]

or

\[
\ln (1 - F(x)) \leq \ln (1 - G(x))
\]

Since \(\ln(t)\) decreases in \(t\) for \(t < 1\), we have

\[1 - F(x) \geq 1 - G(x),\]

14
or

\[ F(x) \leq G(x) \text{ for any } x. \]

The converse is not true.

The concept of first-order stochastic dominance has applications beyond comparing returns. For instance, when firms prices are stochastic, one firm’s price may be higher than another firm in the sense of first-order stochastic dominance. In the study of principal-agent problems, we often assume that the output of a firm increases in the effort of the manager in the sense of first-order stochastic dominance.

We now compare distributions with the same mean to focus on the idea of which distribution involves more risk.

**Definition 9** For any two distributions with the same mean, \( F(\cdot) \) second-order stochastically dominates (or is less risky than) \( G(\cdot) \) if for every nondecreasing concave function \( u : \mathbb{R}_+ \to \mathbb{R} \), we have

\[
\int u(x)dF(x) \geq \int u(x)dG(x).
\]

An alternative way to say that \( F \) second-order stochastically dominates \( G \) is to say that \( G \) is a mean-preserving spread of \( F \).

Consider the following lottery: in the first stage, we have a lottery over \( x \) with distribution \( F \). In the second stage, we randomize each possible outcome \( x \) further so that the final payoff is \( x + z \), where \( z \) has a distribution function \( H_x(z) \) with a mean of zero. Thus the mean of \( x + z \) is \( x \). The resulting reduced lottery, denoted by \( G \), is called a mean-preserving spread of \( F \). If \( u(\cdot) \) is concave,

\[
\int u(x)dG(x) = \int \left( \int u(x + z)dH_x(z) \right)dF(x) \leq \int u \left( \int (x + z)dH_x(z) \right)dF(x) = \int u(x)dF(x).
\]

Thus if \( G \) is a mean-preserving spread of \( F \), than \( F \) second-order stochastically dominates \( G \).