Chapter 5

Decision Making under Uncertainty

In previous lectures, we considered decision problems in which the decision maker does not know the consequences of his choices but he is given the probability of each consequence under each choice. In most economic applications, such a probability is not given. For example, in a given game, a player cares not only about what he plays but also about what other players play. Hence, the description of consequences include the strategy profiles. In that case, in order to fit in that framework, we would need to give other players' mixed strategy profiles in the description of the game, making Game Theoretical analysis moot. Likewise in a market, the price is formed according to the collective actions of all market participants, and hence the price distribution is not given.

In all these problems, the decision makers hold subjective beliefs about the unknown aspects of the problem and use these beliefs in making their decisions. For example, a player chooses his strategy according to his beliefs about what other players may play, and he may reach these beliefs through a combination of reasoning and the knowledge of past behavior. This is called decision making under *uncertainty*.

As established by Savage and the others, under some reasonable assumptions, such subjective beliefs can be represented by a probability distribution, in the sense that the decision maker finds an event more likely than another if and only if the probability distribution assigns higher probability to the former event than latter. In that case, using the probability distribution, one can convert a decision problem under uncertainty to a decision problem under risk, and apply the analysis of the previous lecture. In this lecture, I will describe this program in detail. In particular, I will describe

- the conditions such consistent beliefs impose on the preferences,
- the elicitation of the beliefs from the preferences, and
- the representation of the beliefs by a probability distribution.

5.1 Acts, States, Consequences, and Expected Utility Representation

Consider a finite set C of consequences. Let S be the set of all states of the world. Take a set F of acts $f : S \to C$ as the set of alternatives (i.e., set $X = F$). Each state $s \in S$ describes all the relevant aspects of the world, hence the states are mutually exclusive. Moreover, the consequence $f(s)$ of act f depends on the true state of the world. Hence, the decision maker may be uncertain about the consequences of his acts. Recall that the decision maker cares only about the consequences, but he needs to choose an act.

Example 5.1 (Game as a Decision Problem) *Consider a complete information game with set $N = \{1, \ldots, n\}$ of players in which each player $i \in N$ has a strategy space S_i. The decision problem of a player i can be described as follows. Since he cares about the strategy profiles, the set of consequences is $C = S_1 \times \cdots \times S_n$. Since he does not know what the other players play, the set of states is $S = S_{-i} \equiv \prod_{j \neq i} S_j$. Since he chooses among his strategies, the set of acts is $F = S_i$, where each strategy s_i is represented as a function $s_{-i} \mapsto (s_i, s_{-i})$. (Here, (s_i, s_{-i}) is the strategy profile in which i plays s_i and the others play s_{-i}.) Traditionally, a complete-information game is defined by also including the VNM utility function $u_i : S_1 \times \cdots \times S_n \to \mathbb{R}$ for each player. Fixing such a utility function is equivalent to fixing the preferences on all lotteries on $S_1 \times \cdots \times S_n$.*

Note that above example is only a way to model a player's uncertainty in a game, although it seems to be most direct way to model a player's uncertainty about the others' strategies. Depending on the richness of the player's theories in his decision making, one may consider richer state spaces. For example, the player may think that the other players react to whether it is sunny or rainy in their decisions. In that case, one would include the state of the weather in the space space, e.g., by taking $S =$

$S_{-i} \times \{\text{sunny, rainy}\}$. Sometimes, it may also be useful to consider a state space that does not directly refer to the others' strategies.

We would like to represent the decision maker's preference relation $\succeq$ on F by some $U : F \to \mathbb{R}$ such that

$$U(f) \equiv E[u \circ f]$$

(in the sense of (OR)) where $u : C \to \mathbb{R}$ is a "utility function" on C and E is an expectation operator on S. That is, we want

$$f \succeq g \iff U(f) \equiv E[u \circ f] \geq E[u \circ g] \equiv U(g). \tag{EUR}$$

In the formulation of Von Neumann and Morgenstern, the probability distribution (and hence the expectation operator E) is objectively given. In fact, acts are formulated as lotteries, i.e., probability distributions on C. In such a world, as we have seen in the last lecture, $\succeq$ is representable in the sense of (EUR) if and only if it is a continuous preference relation and satisfies the Independence Axiom.

For the cases of our concern in this lecture, there is no objectively given probability distribution on S. We therefore need to determine the decision maker's (subjective) probability assessment on S. This is done in two important formulations. First, Savage carefully elicits the beliefs and represents them by a probability distribution in a world with no objective probability is given. Second, Anscombe and Aumann simply uses indifference between some lotteries and acts to elicit preferences. I will first describe Anscombe and Aumann's tractable model, and then present Savage's deeper and more useful analysis.

5.2 Anscombe-Aumann Model

Anscombe and Aumann consider a tractable model in which the decision maker's subjective probability assessments are determined using his attitudes towards the lotteries (with objectively given probabilities) as well as towards the acts with uncertain consequences. To do this, they consider the decision maker's preferences on the set P^S of all "acts" whose outcomes are lotteries on C, where P is the set of all lotteries (probability distributions on C). In the language defined above, they assume that the consequences and the decision maker's preferences on the set of consequences have the special structure of Von-Neumann and Morgenstern model.

Note that an act f assigns a probability $f(x|s)$ on any consequence $x \in C$ at any state $s \in S$. The expected utility representation in this set up is given by

$$f \succeq g \iff \sum_{s \in S} \sum_{x \in C} u(x) f(x|s) p(s).$$

In this set up, it is straightforward to determine the decision maker's probability assessments. Consider a subset A of S and any two consequences $x, y \in C$ with $x \succ y$. Consider the act f_A that yields the sure lottery of x on A,[1] and the sure lottery of y on $S \backslash A$. That is, $f_A(x|s) = 1$ for any $s \in A$ and $f_A(y|s) = 1$ for any $s \notin A$. (See Figure 5.1.) Under some continuity assumptions (which are also necessary for representability), there exists some $\pi_A \in [0,1]$ such that the decision maker is indifferent between f_A and the act g_A with $g_A(x|s) = \pi_A$ and $g_A(y|s) = 1 - \pi_A$ at each $s \in S$. That is, regardless of the state, g_A yields the lottery p_A that gives x with probability π_A and y with probability $1 - \pi_A$. Then, π_A is the (subjective) probability the decision maker assigns to the event A — under the assumption that π_A does not depend on which alternatives x and y are used. In this way, one obtains a probability distribution on S. Using the theory of Von Neumann and Morgenstern, one then obtains a representation theorem in this extended space where we have both subjective uncertainty and objectively given risk.

While this is a tractable model, it has two major limitation. First, the analysis generates little insights into how one should think about the subjective beliefs and their representation through a probability distribution. Second, in many decision problems there may not be relevant intrinsic events that have objectively given probabilities and rich enough to determine the beliefs on the events the decision maker is uncertain about.

5.3 Savage Model

Savage develops a theory with purely subjective uncertainty. Without using any objectively given probabilities, under certain assumptions of "tightness", he derives a unique probability distribution on S that represent the decision maker's beliefs embedded in his preferences, and then using the theory of Von Neumann and Morgenstern he obtain a representation theorem — in which both utility function and the beliefs are derived from the preferences.

[1] That is, $f_A(s) = x$ whenever $s \in A$ where the lottery x assigns probability 1 to the consequence x.

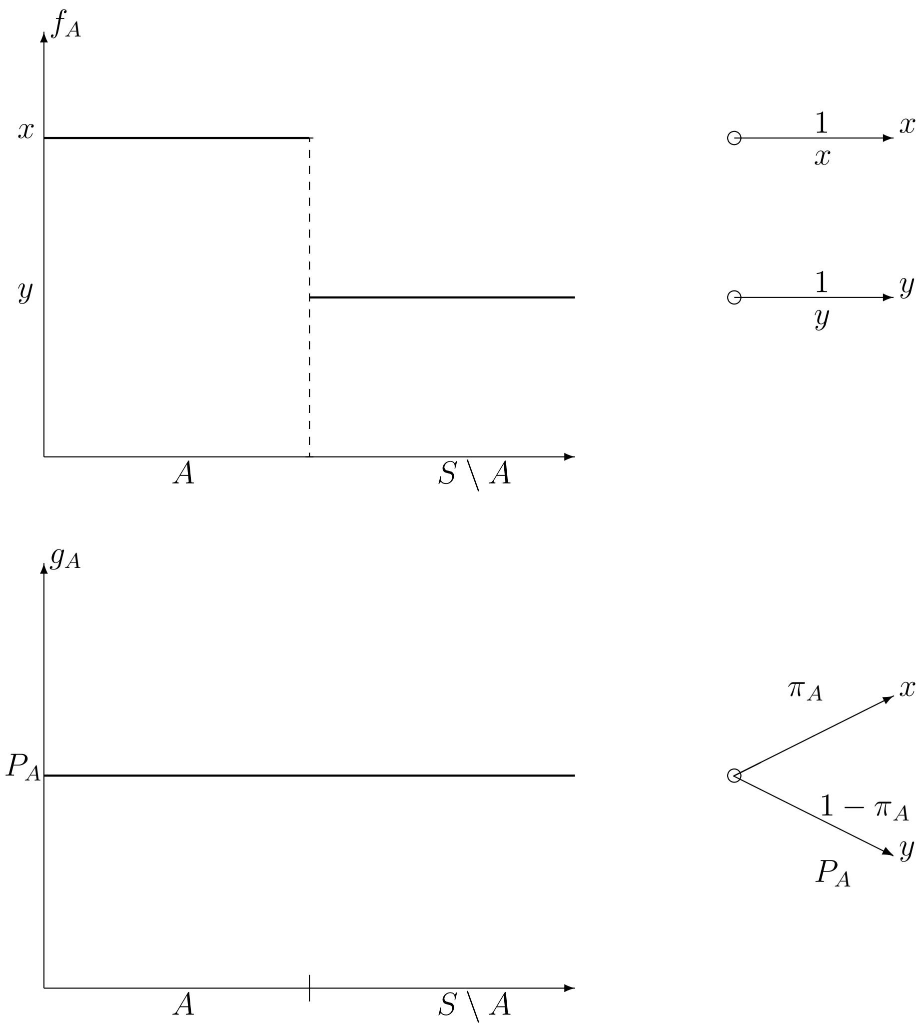

Figure 5.1: Figure for Anscombe and Aumann

Take a set S of *states* s of the world, a finite set C of *consequences* (x, y, z), and take the set $F = C^S$ of *acts* $f : S \to C$ as the set of alternatives. Fix a relation $\succeq$ on F. We would like to find necessary and sufficient conditions on $\succeq$ so that $\succeq$ can be represented by some U in the sense of (EUR); i.e., $U(f) = E[u \circ f]$. In this representation, both the utility function $u : C \to R$ and the probability distribution p on S (which determines E) are derived from $\succeq$. Theorems 1.2 and 1.3 give us the first necessary condition:

P 1 *$\succeq$ is a preference relation.*

The second condition is the central piece of Savage's theory:

5.3.1 The Sure-thing Principle

The Sure-thing Principle *If a decision maker prefers some act f to some act g when he knows that some event $A \subset S$ occurs, and if he prefers f to g when he knows that A does not occur, then he must prefer f to g when he does not know whether A occurs or not.*

This is the informal statement of the sure-thing principle. Once we determine the decision maker's probability assessments, the sure-thing principle will give us the Independence Axiom, Axiom 2.3, of Von Neumann and Morgenstern. The following formulation of Savage, P2, not only implies this informal statement, but also allows us to state it formally, by allowing us to define conditional preferences. (The conditional preferences are also used to define the beliefs.)

P 2 *Let $f, f', g, g' \in F$ and $B \subset S$ be such that*

$$f(s) = f'(s) \text{ and } g(s) = g'(s) \text{ at each } s \in B$$

and

$$f(s) = g(s) \text{ and } f'(s) = g'(s) \text{ at each } s \notin B.$$

If $f \succeq g$, then $f' \succeq g'$.

5.3.2 Conditional preferences

Using P2, we can define the conditional preferences as follows. Given any $f, g, h \in F$ and $B \subset S$, define acts $f^h_{|B}$ and $g^h_{|B}$ by

$$f^h_{|B}(s) = \begin{cases} f(s) & \text{if } s \in B \\ h(s) & \text{otherwise} \end{cases}$$

and

$$g^h_{|B}(s) = \begin{cases} g(s) & \text{if } s \in B \\ h(s) & \text{otherwise} \end{cases}.$$

That is, $f^h_{|B}$ and $g^h_{|B}$ agree with f and g, respectively, on B, but when B does not occur, they yield the same default act h.

Definition 5.1 ***(Conditional Preferences)*** $f \succeq g$ given B *iff* $f^h_{|B} \succeq g^h_{|B}$.

P2 guarantees that $f \succeq g$ given B is well-defined, i.e., it does not depend on the default act h. To see this, take any $h' \in F$, and define $f^{h'}_{|B}$ and $g^{h'}_{|B}$ accordingly. Check that

$$f^h_{|B}(s) \equiv f(s) \equiv f^{h'}_{|B}(s) \text{ and } g^h_{|B}(s) \equiv g(s) \equiv g^{h'}_{|B}(s) \text{ at each } s \in B$$

and

$$f^h_{|B}(s) \equiv h(s) \equiv g^h_{|B}(s) \text{ and } f^{h'}_{|B}(s) \equiv h'(s) \equiv g^{h'}_{|B}(s) \text{ at each } s \notin B.$$

Therefore, by P2, $f^h_{|B} \succeq g^h_{|B}$ iff $f^{h'}_{|B} \succeq g^{h'}_{|B}$.

Note that P2 precisely states that $f \succeq g$ given B is well-defined. To see this, take f and g' arbitrarily. Set $h = f$ and $h' = g'$. Clearly, $f = f^h_{|B}$ and $g' = g^{h'}_{|B}$. Moreover, the conditions in P2 define f' and g as $f' = f^{h'}_{|B}$ and $g = g^h_{|B}$. Thus, the conclusion of P2, "if $f \succeq g$, then $f' \succeq g'$", is the same as "if $f^h_{|B} \succeq g^h_{|B}$, then $f^{h'}_{|B} \succeq g^{h'}_{|B}$.

Exercise 5.1 *Show that the informal statement of the sure-thing principle is formally true: given any $f_1, f_2 \in F$, and any $B \subseteq S$,*

$$[(f_1 \succeq f_2 \textit{ given } B) \textit{ and } (f_1 \succeq f_2 \textit{ given } S \backslash B)] \Rightarrow [f_1 \succeq f_2].$$

[Hint: define $f := f_1 = f^{f_1}_{1|B} = f^{f_1}_{1|S\backslash B}$, $g' := f_2 = f^{f_2}_{2|B} = f^{f_2}_{2|S\backslash B}$, $f' := f^{f_2}_{1|B} = f^{f_1}_{2|S\backslash B}$, and $g := f^{f_1}_{2|B} = f^{f_2}_{1|S\backslash B}$. Notice that you do not need to invoke P2 (explicitly).]

Null Events Imagine that the decision maker remains indifference towards any changes made to an action within an event B. Namely, for any acts f and g, the decision maker remains indifferent between f and g, so long as f and g are identical on $S \backslash B$, no matter how widely differ on B. In that case, it is plausible to deduce that the decision maker does not think that event B obtains. Such events are called *null*.

Definition 5.2 *An event B is said to be* null *if and only if $f \sim g$ given B for all $f, g \in F$.*

Recall that our aim is to develop a theory that relates the preferences on the acts with uncertain consequences to the preferences on the consequences. (The preference relation $\succeq$ on F is extended to C by embedding C into F as constant acts. That is, we say $x \succeq x'$ iff $f \succeq f'$ where f and f' are constant acts that take values x and x', respectively.) The next postulate does this for conditional preferences:

P 3 *Given any $f, f' \in F$, $x, x' \in C$, and $B \subset S$, if $f \equiv x$, $f' \equiv x'$, and B is not null, then*

$$f \succeq f' \text{ given } B \iff x \succeq x'.$$

For $B = S$, P3 is rather trivial, a matter of definition of a consequence as a constant act. When $B \neq S$, P3 is needed as an independent postulate. Because the conditional preferences are defined by setting the outcomes of the acts to the same default act when the event does not occur, and two distinct constant acts cannot take the same value.

5.3.3 Representing beliefs with qualitative probabilities

We want to determine the decision maker's beliefs reflected in $\succeq$. Towards this end, given any two events A and B, we want to determine which event the decision maker thinks is more likely. To do this, take any two consequences $x, x' \in C$ with $x \succ x'$. The decision maker is asked to choose between the two gambles (acts) f_A and f_B with

$$f_A(s) = \begin{cases} x & \text{if } s \in A \\ x' & \text{otherwise} \end{cases}, \tag{5.1}$$

$$f_B(s) = \begin{cases} x & \text{if } s \in B \\ x' & \text{otherwise} \end{cases}.$$

If the decision maker prefers f_A to f_B, we can infer that he finds event A more likely than event B, for he prefers to get the "prize" when A occurs, rather than when B occurs.

Definition 5.3 *Take any $x, x' \in C$ with $x \succ x'$. Given any $A, B \subseteq S$, A is said to be* at least as likely as B *(denoted by $A \dot{\succeq} B$) if and only if $f_A \succeq f_B$, where f_A and f_B defined by (5.1).*

We want to make sure that this yields well-defined beliefs. That is, it should not be the case that, when we use some x and x', we infer that decision maker finds A strictly more likely than B, but when we use some other y and y', we infer that he finds B strictly more likely than A. Then next assumption guaranties that $\dot{\succeq}$ is indeed well-defined.

P 4 *Given any $x, x', y, y' \in C$ with $x \succ x'$ and $y \succ y'$, define f_A, f_B, g_A, g_B by*

$$\begin{aligned} f_A(s) &= \begin{cases} x & \text{if } s \in A \\ x' & \text{otherwise} \end{cases}, \quad g_A(s) = \begin{cases} y & \text{if } s \in A \\ y' & \text{otherwise} \end{cases} \\ f_B(s) &= \begin{cases} x & \text{if } s \in B \\ x' & \text{otherwise} \end{cases}, \quad g_B(s) = \begin{cases} y & \text{if } s \in B \\ y' & \text{otherwise} \end{cases}. \end{aligned}$$

Then,

$$f_A \succeq f_B \iff g_A \succeq g_B.$$

Finally, make sure that we can find x and x' with $x \succ x'$:

P 5 *There exist some $x, x' \in C$ such that $x \succ x'$.*

We have now a well-defined relation $\dot{\succeq}$ that determines which of two events is more likely. It turns out that, $\dot{\succeq}$ is a *qualitative probability*, defined as follows:

Definition 5.4 *A relation $\dot{\succeq}$ between the events is said to be a* qualitative probability *iff*

1. *$\dot{\succeq}$ is complete and transitive;*
2. *for any $B, C, D \subset S$ with $B \cap D = C \cap D = \emptyset$,*

$$B \dot{\succeq} C \iff B \cup D \dot{\succeq} C \cup D;$$

3. $B \dot{\succeq} \emptyset$ *for each* $B \subset S$, *and* $S \dot{\succ} \emptyset$.

Exercise 5.2 *Show that, under the postulates P1-P5, the relation* $\dot{\succeq}$ *defined in Definition 5.3 is a qualitative probability.*

5.3.4 Quantifying the qualitative probability assessments

Savage uses *finitely-additive probability measures on the discrete sigma-algebra*:

Definition 5.5 *A* probability measure *is any function* $p : 2^S \to [0, 1]$ *with*

1. *if* $B \cap C = \emptyset$, *then* $p(B \cup C) = p(B) + p(C)$, *and*

2. $p(S) = 1$.

We would like to represent our qualitative probability $\dot{\succeq}$ with a (quantitative) probability measure p in the sense that

$$B \dot{\succeq} C \iff p(B) \geq p(C) \qquad \forall B, C \subseteq S. \tag{QPR}$$

Exercise 5.3 *Show that, if a relation* $\dot{\succeq}$ *can be represented by a probability measure, then* $\dot{\succeq}$ *must be a qualitative probability.*

When S is finite, since $\dot{\succeq}$ is complete and transitive, by Theorem 1.2, it can be represented by some function p, but there might be no such function satisfying the condition 1 in the definition of probability measure. Moreover, S is typically infinite. (Incidentally, the theory that follows requires S to be infinite.)

We are interested in the preferences that can be considered coming from a decision maker who evaluates the acts with respect to their expected utility, using a utility function on C and a probability measure on S that he has in his mind. Our task at this point is to find what probability $p(B)$ he assigns to some arbitrary event B. Imagine that we ask this person whether $p(B) \geq 1/2$. Depending on his sincere answer, we determine whether $p(B) \in [1/2)$ or $p(B) \in [0, 1/2, 1]$. Given the interval, we ask whether $p(B)$ is in the upper half or the lower half of this interval, and depending on his answer, we obtain a smaller interval that contains $p(B)$. We do this ad infinitum. Since the length of the interval at the nth iteration is $1/2^n$, we learn $p(B)$ at the end.

For example, let's say that $p(B) = 0.77$. We first ask if $p(B) \geq 1/2$. He says Yes. We ask now if $p(B) \geq 3/4$. He says Yes. We then ask if $p(B) \geq 7/8$. He says No. Now, we ask if $p(B) \geq 13/16 = (3/4 + 7/8)/2$. He says No again. We now ask if $p(B) \geq 25/32 = (3/4 + 7/8)/2$. He says No. Now we ask if $p(B) \geq 49/64$. He says Yes now. At this point we know that $49/64 \cong 0.765 \leq p(B) < 25/32 \cong 0.781$. As we ask further we get a better answer.

This is what we will do, albeit in a very abstract setup. Assume that S is *infinitely divisible under* $\dot{\succeq}$. That is, S has

- a partition $\{D_1^1, D_1^2\}$ with $D_1^1 \cup D_1^2 = S$ and $D_1^1 \dot{\sim} D_1^2$,
- a partition $\{D_2^1, D_2^2, D_2^3, D_2^4\}$ with $D_2^1 \cup D_2^2 = D_1^1$, $D_2^3 \cup D_2^4 = D_1^2$, and $D_2^1 \dot{\sim} D_2^2 \dot{\sim} D_2^3 \dot{\sim} D_2^4$,
- $\vdots$
- a partition $\left\{D_n^1, \cdots, D_n^{2^n}\right\}$ with $D_n^1 \cup D_n^2 = D_{n-1}^1$, ..., $D_n^{2k-1} \cup D_n^{2k} = D_{n-1}^k$, ..., and $D_n^1 \dot{\sim} \cdots \dot{\sim} D_n^{2^n}$,
- $\vdots$

ad infinitum.

<table>
<tr><td colspan="4" align="center">S</td></tr>
<tr><td colspan="2" align="center">D_1^1</td><td colspan="2" align="center">D_1^2</td></tr>
<tr><td>D_2^1</td><td>D_2^2</td><td>D_2^3</td><td>D_2^4</td></tr>
<tr><td>$\vdots$</td><td>$\vdots$</td><td>$\vdots$</td><td>$\vdots$</td></tr>
</table>

Exercise 5.4 *Check that, if* $\dot{\succeq}$ *is represented by some* p*, then we must have* $p(D_n^r) = 1/2^n$.

Given any event B, for each n, define

$$k(n, B) = \max\left\{r | B \dot{\succeq} \bigcup_{i=1}^{r} D_n^i\right\},$$

where we use the convention that $\cup_{i=1}^r D_n^i = \emptyset$ whenever $r < 1$. Define

$$p(B) := \lim_{n\to\infty} \frac{k(n, B)}{2^n}. \tag{5.2}$$

Check that $k(n,B)/2^n \in [0,1]$ is non-decreasing in n. Therefore, $\lim_{n\to\infty} k(n,B)/2^n$ is well-defined.

Since $\dot{\succeq}$ is transitive, if $B \dot{\succeq} C$, then $k(n,B) \geq k(n,C)$ for each n, yielding $p(B) \geq p(C)$. This proves the $\Longrightarrow$ part of (QPR) under the assumption that S is infinitely divisible. The other part ($\Leftarrow$) is implied by the following assumption:

P 6' *If $B \dot{\succ} C$, then there exists a finite partition $\{D^1, \dots, D^n\}$ of S such that $B \dot{\succ} C \cup D^r$ for each r.*

Under P1-P5, P6' also implies that S is infinitely divisible. (See the definition of "tight" and Theorems 3 and 4 in Savage.) Therefore, P1-P6' imply (QPR), where p is defined by (5.2).

Exercise 5.5 *Check that, if $\dot{\succeq}$ is represented by some p', then*

$$\frac{k(n,B)}{2^n} \leq p'(B) < \frac{k(n,B)+1}{2^n}$$

at each B. Hence, if both p and p' represent $\dot{\succeq}$, then $p = p'$.

Postulate 6 will be somewhat stronger than P6'. (It is also used to obtain the continuity axiom of Von Neumann and Morgenstern.)

P 6 *Given any $x \in C$, and any $g, h \in F$ with $g \succ h$, there exists a partition $\{D^1, \dots, D^n\}$ of S such that*

$$g \succ h_i^x \text{ and } g_i^x \succ h$$

for each $i \leq n$ where

$$h_i^x(s) = \begin{cases} x & \text{if } s \in D^i \\ h(s) & \text{otherwise} \end{cases} \text{ and } g_i^x(s) = \begin{cases} x & \text{if } s \in D^i \\ g(s) & \text{otherwise} \end{cases}.$$

Take $g = f_B$ and $h = f_C$ (defined in (5.1)) to obtain P6'.

Theorem 5.1 *Under P1-P6, there exists a unique probability measure p such that*

$$B \dot{\succeq} C \iff p(B) \geq p(C) \qquad \forall B, C \subseteq S. \tag{QPR}$$

5.3.5 Expected Utility Representation

In Chapter 5, Savage shows that, when C is finite, Postulates P1-P6 imply Axioms 2.1-2.3 of Von Neumann and Morgenstern —as well as their modeling assumptions such as only the probability distributions on the set of prizes matter. In this way, he obtains the following Theorem:[2]

Theorem 5.2 *Assume that C is finite. Under P1-P6, there exist a utility function $u : C \to R$ and a probability measure $p : 2^S \to [0, 1]$ such that*

$$f \succeq g \iff \sum_{c \in C} p\left(\{s|f(s) = c\}\right) u(c) \geq \sum_{c \in C} p\left(\{s|g(s) = c\}\right) u(c)$$

for each $f, g \in F$.

[2] For the inifinte C, we need the infinite version of the sure-thing principle:

P 7 *If we have $f \succeq g(s)$ given B for each $s \in B$, then $f \succeq g$ given B. Likewise, if $f(s) \succeq g$ given B for each $s \in B$, then $f \succeq g$ given B.*

Under P1-P7, we get the expected utility representation for general case.