

LINEAR PROGRAMMING PROBLEMS FOR GENERALIZED
UNCERTAINTY

by

Phantipa Thipwiwatpotjana

M.S., Clemson University, South Carolina, USA, 2004

B.S., Chiang Mai University, Chiang Mai, Thailand, 1999

A thesis submitted to the
University of Colorado Denver
in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Mathematical and Statistical Sciences

2010

This thesis for the Doctor of Philosophy

degree by

Phantipa Thipwiwatpotjana

has been approved

by

Weldon Lodwick

Stephen Billups

George Corliss

Alexander Engau

Burt Simon

Date

Thipwiwatpotjana, Phantipa (Ph.D., Applied Mathematics)

Linear Programming Problems for Generalized Uncertainty

Thesis directed by Professor Weldon Lodwick

ABSTRACT

Uncertainty occurs when there is more than one realization that can represent an information. This dissertation concerns merely discrete realizations of an uncertainty. Different interpretations of an uncertainty and their relationships are addressed when the uncertainty is not a probability of each realization. A well known model that can handle a linear programming problem with probability uncertainty is an expected recourse model. If uncertainties in the problem have possibility interpretations, an expected average model, which is comparable to an expected recourse model, will be used. These two models can be mixed when we have both probability and possibility uncertainties in the problem, provided these uncertainties do not occur in the same constraint.

This dissertation develops three new solution methods for a linear optimization problem with generalized uncertainty. These solutions are a pessimistic, an optimistic, and a minimax regret solution. The interpretations of uncertainty in a linear programming problem with generalized uncertainty are not limited to probability and possibility. They can be also necessity, plausibility, belief, random set, probability interval, probability on sets, cloud, and interval-valued

probability measure. These new approaches can handle more than one interpretation of uncertainty in the same constraint, which has not been done before. Lastly, some examples and realistic application problems are presented to illustrate these new approaches. Some comparisons between these solutions and a solution from the mixed expected average and expected recourse model when uncertainties are probability and possibility are mentioned.

This abstract accurately represents the content of the candidate's thesis. I recommend its publication.

Signed _____
Weldon Lodwick

I dedicate this work to my parents,
Amporn and Wut.

ACKNOWLEDGMENT

I am most grateful to my adviser, Dr. Weldon A. Lodwick, who introduced me to the uncertainty world, for his guidance, support, and commitment. His invaluable advice, patience, and encouragement make this work possible.

I am thankful to Dr. Stephen Billups, Dr. George Corliss, Dr. Alexander Engau, and Dr. Burt Simon for serving on my Ph.D. committee. I appreciate their time, comments and suggestions that were very helpful in developing the research. I express my special thanks for their corrections and suggestions that improved this dissertation.

I would like to thank the Thai government officers who took care of me while I am in the US. I extend my appreciation to the Bateman family and the Department of Mathematical and Statistical Sciences, University of Colorado at Denver, who provide the financial support throughout my study after the support from the Thai government was terminated.

Finally, I am most thankful to my mother, Amporn, for her care, love, and support. Warm thanks for my sister, Yui, and my brother, Yook, for sharing their opinions. Special thanks to Iris for her kindness for a Thai girl. To, Tur, for his love, care, support, and help. Because of their encouragement, support, help, and love, my success is also theirs.

CONTENTS

Figures	x
Tables	xi
<u>Chapter</u>	
1. Introduction to the dissertation	1
1.1 Problem statement	4
1.2 Research direction and organization of the dissertation	9
2. Some selected uncertainty interpretations	12
2.1 Possibility theory	13
2.2 Random sets	24
2.2.1 Upper and lower expected values generated by random sets	39
2.2.2 The random set generated by an incomplete random set information	50
2.2.3 The random set generated by a possibility distribution or possibility of some subset of U	50
2.2.4 The random set generated by a necessity distribution or necessity of some subset of U	52
2.2.5 The random set generated by a belief distribution	53
2.2.6 The random set generated by a plausibility distribution	53
2.3 Interval-valued probability measures (IVPMs)	54
2.3.1 Upper and lower expected values generated by IVPMs	66
2.4 Clouds	70
2.5 Relationships of some selected uncertainty interpretations	75

3. Linear optimization under uncertainty: literature review	80
3.1 Deterministic model	81
3.2 Stochastic models	83
3.2.1 Stochastic program with expected recourse	85
3.3 Minimax regret model	90
3.4 Possibility model	93
3.5 Mixed model of probability and possibility	99
3.6 Mixed uncertainties model	101
4. Linear optimization under uncertainty: pessimistic, optimistic, and minimax regret approaches	107
4.1 Pessimistic and optimistic expected recourse problems	109
4.1.1 Pessimistic and optimistic expected recourse problems generated by random sets with finite realizations	115
4.1.2 Pessimistic and optimistic expected recourse problems generated by IVPMs with finite realizations	121
4.2 Minimax regret of the expected recourse problems	123
4.2.1 A relaxation procedure for minimax regret of the expected recourse problems	125
4.3 Storage or selling costs in expected recourse problems	132
4.3.1 Selling costs in expected recourse problems	133
4.3.2 Storage costs in expected recourse problems	134
4.4 Comparison between pessimistic, optimistic, minimax regret ap- proaches and the reviewed approaches	135
5. Examples and applications	139

5.1	Farming example	139
5.1.1	A random set version of the farming example	143
5.2	An application for radiation shielding design of radiotherapy treatment vaults	149
6.	Conclusion	158
	<u>Appendix</u>	
A.	How to obtain a basic probability assignment function m from a given belief measure Bel	161
B.	The nestedness requirement of focal elements for a possibility measure	163
	<u>References</u>	167

FIGURES

Figure		
1.1	Uncertainty interpretations: $A \longrightarrow B$ represents that an uncertainty interpretation B contains information given by an uncertainty interpretation A, and $A \dashrightarrow B$ represents that A is a special case of B.	10
2.1	Venn diagram of an opinion poll for a Colorado governor's election.	18
2.2	Possibility measure defined on $U = \{u_1, u_2, \dots, u_7\}$	22
2.3	A finite representation of an infinite random set.	33
2.4	A random set generated by possibility distribution.	51
2.5	A cloud on \mathbb{R} with an α -cut at $\alpha = 0.6$	71
2.6	A cloud represents an IVPM.	72
2.7	Uncertainty interpretations: $A \longrightarrow B$: there is an uncertainty interpretation B contains information given by an uncertainty interpretation A, $A \dashrightarrow B$: A is a special case of B, $A \longleftrightarrow B$: A and B can be derived from each other, and $A \cdots \rightarrow B$: B generalized A.	77
3.1	Feasible region and the unique optimal solution of the production scheduling problem (3.1).	83
5.1	An example of a radiation room.	150

TABLES

Table

2.1	Probability theory versus possibility theory: comparison of mathematical properties for finite sets	23
2.2	The number of belief and plausibility terms required for finding two density functions for the lowest and the highest expected values of θ by using Theorem 2.17 versus by solving two LP problems.	46
2.3	The differences between IVPMs and clouds	72
3.1	Productivity information shows the output of products per unit of raw materials, the unit costs of raw materials, the demand for the products, and the limitation of the total amount of raw materials.	82
3.2	Computational results for different models of the production planning problem.	105
4.1	Computational Results for the new approaches and the reviewed approach of the production planning problem.	137
5.1	Data for farming problem.	140
5.2	An optimal solution based on expected yields.	140
5.3	The scenarios of yields for all crops.	141
5.4	An optimal solution based on the expected recourse problem (5.2).	142
5.5	Optimal solutions based on the pessimistic and optimistic expected recourse of the farming problem. P = pessimistic, O = optimistic, R = minimax regret.	147

5.6	A pessimistic, an optimistic, a minimax regret, and an extreme scenario solution for the radiation shielding design problem.	157
-----	--	-----

1. Introduction to the dissertation

This dissertation develops a technique for solving linear optimizations under uncertainty problems that have discrete realizations for broad classes of uncertainty not studied and analyzed before. We study optimization under uncertainty in which the probability density mass value of each realization is not known with certainty. Information attached to an uncertainty can be categorized into different interpretations. The interpretations of uncertainty information associated with this thesis are probability, belief, plausibility, necessity, possibility, random set, probability interval, probability on sets, cloud, and interval-valued probability measure (IVPM). For convenience, we call these interpretations of uncertainty ‘*PC-BRIN*’. We develop an approach to compute a pessimistic, an optimistic, and a robust (minimum of maximum regret) solution for a linear programming (LP) problem with these uncertainty interpretations. These problems are solved based on the transformation of a linear optimization problem with uncertainty to a set of expected recourse models.

An expected recourse model is a paradigm to solve a stochastic programming problem. Stochastic programming is the study of practical procedures for decision making under uncertainty over time. Stochastic programs are mathematical programs (linear, integer, mixed-integer, nonlinear) where some of the data incorporated into the objective or constraints are uncertain with a probability interpretation. An expected recourse model requires that one makes one decision now and minimizes the expected costs (or evaluations) of the conse-

quences of that decision. We consider a two stage expected recourse model in this thesis. The first stage is as a decision that one needs to make now, and the second stage is as a decision based on what has happened. The objective is to minimize the expected costs of all decisions taken.

We refer to a *probability* interpretation of uncertainty when we can create or assume a probability for each realization from an experiment without using any prior knowledge. For example, we will not assume that a coin is fair when we know nothing about this coin. The other interpretations of an uncertainty information mentioned above (except probability) share the same behavior, i.e., the information that leads to one of those interpretations is not enough to obtain a probability for each realization. Instead, an appropriate function is created based on information provided.

Let U be a finite set of all realizations of an uncertainty \hat{u} . A *belief* interpretation of uncertainty is given in the form of a belief function, Bel , which maps from an event (a subset of U) to a number between 0 and 1. For an event A , $Bel(A)$ can be interpreted as one's degree of belief that the truth of \hat{u} lies in A . The probability may or may not coincide with the degree of belief about an event of \hat{u} . If we know the probabilities of events, then we will surely adopt them as the degrees of belief. But, if we do not know the probabilities, then it will be an extraordinary coincidence for the degrees of belief to be equal to their probability. In general, sum of the probability of two mutually disjoint events is equal to the probability of the union of those two events. This statement is relaxed when the function is a belief function. Intuitively, one's degree of belief that the truth lies in A_1 plus the degree of belief that the truth lies in A_2 is

always less than or equal to the degree of belief that the truth lies in $A_1 \cup A_2$.

G. Shafer, [60], mentioned that one's beliefs that the truth of \hat{u} lies in an event A are not fully described by one's degree of belief $Bel(A)$. One may also have some doubts about A . The degree of doubt can be expressed in the form $Dou(A) = Bel(A^c)$. A *plausibility* interpretation of uncertainty is closely related to a belief because a plausibility function, Pl , can be derived from a belief function, and vice versa, by using $Pl(A) = 1 - Bel(A^c)$. One's degree of plausibility $Pl(A)$ expresses that one fails to doubt A or one finds A plausible. Hence, Bel and Pl convey the same information, as we shall see in many examples throughout the thesis. A *necessity* and a *possibility* interpretations of uncertainty are special versions of belief and plausibility, respectively. We call belief and plausibility functions necessity and possibility functions, Nec and Pos , when for events A_1 and A_2 , $Bel(A_1 \cap A_2) = \min [Bel(A_1), Bel(A_2)]$ and $Pl(A_1 \cup A_2) = \max [Pl(A_1), Pl(A_2)]$, respectively. The mathematical definitions and some properties of belief, plausibility, necessity and possibility functions are provided in Section 2.1.

An uncertainty provided in a form of *random set* interpretation has information as a set of probabilities that are bounded above and below by plausibilities and beliefs. A *probability interval* is an interval mapping from each element of U to its corresponding interval $[\underline{a}, \bar{a}]$, where $[\underline{a}, \bar{a}] \subseteq [0, 1]$. An *IVPM* interpretation of uncertainty has information as intervals on probability of A , for every subset A of U . A *cloud* is defined differently from IVPM. However, it turns out that a cloud is an example of IVPM. More details on random set, cloud, and IVPM interpretations are in Chapter 2. *Probability on sets* is a partial information of

probability, i.e., we know a value of $P(A)$ for some $A \subseteq U$, but it is not enough to obtain the probability of each realization in U . Probability intervals and probability on sets can be viewed as examples of IVPM. Thus, the uncertainty interpretations over finite realizations we include in our analysis are: probability, belief, plausibility, necessity, possibility, random set, probability interval, probability on sets, cloud, and IVPM, or PC-BRIN, in short.

1.1 Problem statement

The problem which is the focus of this thesis is

$$\min_x \widehat{c}^T x \text{ s.t. } \widehat{A}x \geq \widehat{b}, \quad Bx \geq d, \quad x \geq 0. \quad (1.1)$$

We sometimes call (1.1) an LP problem with (generalized or mixed) uncertainty. This dissertation tries to answer the question of how to solve (1.1). To date, the theory and solution methods for (1.1) have not been developed when \widehat{A} , \widehat{b} , and \widehat{c} have only one of the PC-BRIN uncertainty interpretations, except probability and possibility. Moreover, when \widehat{A} , \widehat{b} , and \widehat{c} are mixtures of all the PC-BRIN uncertainty interpretations within one constraint, there is no theory or solution method yet to deal with this case. The significance of what is presented is that problems possessing these uncertainty interpretations can be modeled and solved directly from their true, basic, uncertainties. That is, the model is more faithful to its underlying properties. Secondly, the model is faithful to the data available.

We provide a simple LP problem with uncertainty without solving it for the purpose of showing that uncertainty information, which does not have probability interpretation, may be an integral part of in an LP model. Suppose that

a tour guide wants to minimize the transportation cost. There are usually 100 to 130 tourists the guide has to care for each day. However, the guide has to rent cars in advance without knowing the total number of the tourists. A car rental company provides the guide some information about the car types H and T based on a questionnaire of its 3,000 customers about their opinions on how many passengers the car types H and T could carry. The response from the questionnaire is indicated in the table below.

Passenger capacity	Number of responders	
	Car type H	Car type T
Up to 4 people	250	-
Up to 5 people	250	500
Up to 6 people	2500	2500

This information is similar to our Example 3 on page 22 and can be presented as a random set. Suppose that the rental prices for the car types H and T are \$34/day and \$45/day, respectively. The guide assumes that the number of tourists are equally likely to be any number between 100 and 130 people. Therefore, without knowing the age, the size, or other information about the clients, the guide sets up a small LP problem with these uncertainties as

$$\begin{aligned} \min \quad & 34H + 45T \\ \text{s.t.} \quad & \hat{a}_1H + \hat{a}_2T \geq \hat{b}, \text{ and } H, T \geq 0, \end{aligned}$$

where \hat{a}_1 can be 4, 5, or 6, and \hat{a}_2 can be 5 or 6 with the random set information above. The number of tourists \hat{b} is between 100 and 130 persons with equal

chance. There might be some other restrictions to make the problem more complicated. For example, the deal from the car rental is that a customer needs to rent at least a certain number of cars of type T to get some reduced price. The guide also may need to please his/her clients by assigning family clients in separate cars, and so on. We should be able to see now that there is an LP with uncertainty, where the uncertainty may not be interpreted as probability.

Linear program (1.1) with mixed uncertainty is a mathematical linear program, where some of parameters in the objective or constraints are uncertain with any of the interpretations mentioned earlier. Let us consider a linear programming problem with mixed uncertainty through a production planning problem, which minimizes the production cost and satisfies the demands at the same time, as a prototypical problem (1.1). Let \hat{c} be an uncertain cost vector per unit of raw material vector x , \hat{A} be a matrix of uncertain machine capacity, and \hat{b} be an uncertain demand vector. Here, we assume that components of \hat{A} , \hat{b} , and \hat{c} may possess one of the PC-BRIN uncertainty interpretations. An LP problem with uncertainty stated as the system (1.1) is not well-defined until the realizations of \hat{A} , \hat{b} , and \hat{c} are known.

Suppose that there is no uncertainty in the cost ‘ c ’ of raw materials. The model (1.1) becomes

$$\min_x c^T x \text{ s.t. } \hat{A}x \geq \hat{b}, \quad Bx \geq d, \quad x \geq 0. \quad (1.2)$$

We apply a two stage expected recourse model to an LP problem with uncertainty, when all uncertainties have probability interpretation. The first stage is to decide the amount of raw materials needed. Based on this decision, the consequent action is to make sure that these raw materials provide enough to

satisfy the demands. If not, the second action is needed, i.e, the amount of the shortages needs to be bought from a market at a (fixed) penalty prices. In the case that there is an excess amount of products left after satisfying the demands, this excess amount can be sold in the market or stored with some storage price. Therefore, the expected recourse objective function is to minimize the cost of raw materials and the expected cost of the shortages together with one of the following: (1) the expect cost of storage, or (2) negative of the expected cost of profit from selling the excesses. These two cases do not happen at the same time when we have no further planning for the excesses, because we will rather sell to make profit than spend out of the budget for the storage.

We introduce a variable z , when some component of the cost \hat{c} is uncertain with a probability interpretation and transform the stochastic program (1.1) to

$$\min_{x,z} z \text{ s.t. } \hat{A}x \geq \hat{b}, \quad z = \hat{c}^T x, \quad x, z \geq 0. \quad (1.3)$$

In this case, the first stage variables are x and z . The surplus y and slack v variables that control all realizations of constraint $z = \hat{c}^T x$ are the second stage variables, in addition to the shortage variable w . The expected recourse objective function is to minimize the expected cost of raw materials and the shortage together with either the expected cost of storage or negative of the expected cost of profit.

An expected average model [24, 76], see also in Section 3.4, is comparable to an expected recourse model. It is designed to handle the model (1.1) only when all uncertainties have a possibility interpretation. The possibility interpretation in [24] is actually a possibility distribution, which is the possibility measure for only all singleton events. If we have a possibility distribution of \hat{u} , we also have

the possibility measure of \hat{u} , which is explained in Chapter 2.

An interval expected value approach, [36, 69], is the only approach so far that takes advantage of the knowledge, ‘*possibility and necessity measures of an uncertainty \hat{u} convey the same information*’. Uncertainties in an LP problem with possibility uncertainty automatically have necessity interpretations. Possibility and necessity measures are recognized as the bounds on the cumulative distribution of probability measures. An interval expected value for a possibility uncertainty is an interval that has the left and right bounds as the smallest and the largest expected values, respectively. An interval expected value approach transforms (1.1) to an interval linear program, where the coefficients are now these interval expected values. After we study relationships among the uncertainty interpretations in Chapter 2, we can use interval expected value to handle LP problems with mixed uncertainty, since each of PC-BRIN uncertainty interpretations with finite realizations can be characterized as a set of probability measures, and hence, we can find an interval expected value of that uncertainty. We provide the details of an interval expected value approach in Section 3.6.

If all uncertainties in problem (1.1) are probability, then expected values of these uncertainties can be found. However, we do not use these expected values to represent (1.1). Instead, we represent (1.1) as a stochastic programming problem, because first of all, the expected values of each uncertainty may not even be one of the realizations of that uncertainty. Secondly, a solution obtained from the expected value representation is not the best decision in the long run. The interval expected value approach is not a good representation of an LP problem with uncertainty for similar reasons. Therefore, instead of the interval

expected value approach, we suggest three treatments for an LP problem with uncertainty (1.1); (1) optimistic approach, (2) pessimistic approach, and (3) minimax regret approach, after we are able to characterize each uncertainty interpretation as a set of probability measures.

1.2 Research direction and organization of the dissertation

We study some mathematical details of these different uncertainty interpretations, PC-BRIN, and the relationships among them, to be able to characterize each uncertainty interpretation as a closed polyhedral set of probability measures. Figure 1.1 illustrates that there is a random set that contains information given by probability, belief, plausibility, necessity, possibility, or probability interval interpretations of an uncertainty. Moreover, as we shall see, random set, probability interval, probability on sets, and cloud are special cases of IVPM interpretation. A similar figure will be seen in Chapter 2 with more details.

A literature review of linear programming with uncertainty is given in Chapter 3. We conclude this section with a word about the limitations of the approaches found in the literature.

1. The approaches in the literature are limited to probability and possibility uncertainty interpretations in linear optimization problems. Moreover, these two interpretations cannot be in the same constraint.
2. Although *Pos* and *Nec* convey the same information, there is no method (except an interval expected value approach) in the literature to handle necessity uncertainty interpretation.

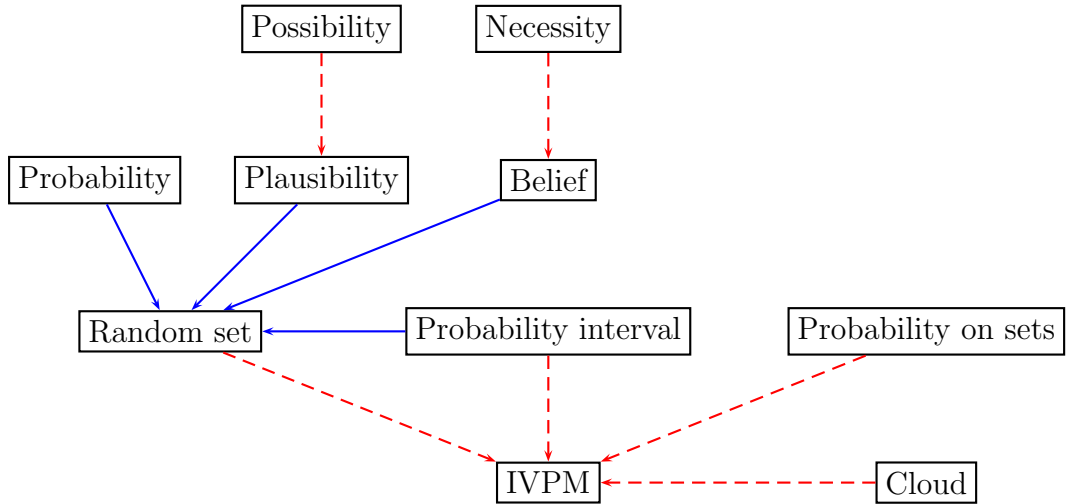


Figure 1.1: Uncertainty interpretations: $A \longrightarrow B$ represents that an uncertainty interpretation B contains information given by an uncertainty interpretation A, and $A \dashrightarrow B$ represents that A is a special case of B.

3. There is no approach (except an interval expected value approach) for solving a linear program with more than one uncertainty interpretation in one constraint.
4. An interval expected value approach is not a good representation of problem (1.1). The reason is stated at the last paragraph of the previous section.

The method presented in this dissertation overcomes these limitations, both theoretically and practically. The new approaches can handle probability, belief, plausibility, necessity, possibility, random set, probability interval, probability on sets, cloud, and interval-valued probability uncertainty interpretations in one problem.

These uncertainty interpretations tell us that although we do not know the actual probability for each realization of an uncertainty \hat{u} , we know the area where it could be. In Chapter 4, we will find two probabilities $\underline{f}_{\hat{u}}$ and $\bar{f}_{\hat{u}}$ that provide the smallest and the largest expected value of \hat{u} , respectively. Therefore, the method presented in this dissertation is based on the stochastic expected recourse programs to find a pessimistic, and an optimistic solution. We may have infinitely many expected recourse programs related to all possible probabilities. However, these two probabilities $\underline{f}_{\hat{u}}$ and $\bar{f}_{\hat{u}}$ for each uncertainty \hat{u} in a linear programming problem with uncertainty lead to the smallest and the largest expected recourse values. Moreover, we find a minimax regret solution as the best solution in the sense that without knowing the actual probability, this solution provides the minimum of the maximum regret.

Next, the comparisons of our new treatments and the other models in literature are provided through numerical examples. Some useful example and application illustrating the power and efficacy are contained in Chapter 5. Chapter 6 summarizes the results of this thesis and present plans for further research.

2. Some selected uncertainty interpretations

This dissertation focuses on *uncertainty* as it applies to optimization. When information has more than one realization, we call it an uncertainty, \hat{u} . Mathematically, uncertainty not only contains the standard probability theory but also other theories depending upon the information we have. Interpretations of uncertainty are based on the theories and information behind them. For example, a fair coin has probability 0.5 that a head (or a tail) will occur. However, if we do not have information that this coin is fair, we should not assume that it is fair. The information we really have here is only $Pr(\{\text{a head, a tail}\}) = 1$, and nothing else. In describing the outcome of a coin flip, where its fairness is in question, we could say that it is possible that a head (or a tail) will occur with degree of possibility 1. This tells us that the actual probability of a head (or a tail) of this coin can be anything between 0 and 1, which will be known only when we test the coin. Moreover, even though the degree of possibility for a fair coin that a head occurs is equal to 1, the knowledge we have that the coin is ‘fair’ is much stronger than the possibility information. This knowledge provides the exact probability, which is more useful than the range of possible probabilities.

This chapter has the aim of describing PC-BRIN interpretations of uncertainty. We will spend time on some details of these theories so that we can find an appropriate treatment to deal with linear optimization problems with mixed uncertainty interpretations as we will see in Chapter 4. Some results in

this chapter are known results. However, there also are some new contributions, which are strongly pointed out inside and at the end of the chapter. The chapter starts with a possibility measure, which is derived from a belief measure. Then, the definition and some examples of a random set are provided. We can construct a set of probability measures, whose bounds are belief and plausibility measures, given a random set. The smallest and the largest expected values of a random variable, whose probability is uncertain in the form of a random set, can be found using the density (mass) functions given in Subsection 2.2.1. The mathematical definitions of interval-valued probability measures and clouds are provided in Sections 2.3 and 2.4. We also point out that, for the case of finite set of realizations, a probability interval can be used to create a random set (not necessarily unique). Finally, we conclude with the relationships among these uncertainty interpretations. Basic probability theory is assumed.

2.1 Possibility theory

Possibility theory is a special branch of Dempster [9] and Shafer [60] evidence theory, so we will provide some details of evidence theory first. Most of the materials in this section can be found in [9, 28, 60], and [74].

Evidence theory is based on belief and plausibility measures. For a finite set U of realizations, where $\mathcal{P}(U)$ is the power set of U , a belief measure is a function

$$Bel : \mathcal{P}(U) \rightarrow [0, 1] \tag{2.1}$$

such that $Bel(\emptyset) = 0$, $Bel(U) = 1$, and having a super-additive property for all possible families of subsets of U . The super-additive property for

a belief function generated by a finite set U , where $A_1, \dots, A_n \subseteq U$, is:
 $Bel(A_1 \cup \dots \cup A_n) \geq \sum_j Bel(A_j) - \sum_{j < k} Bel(A_j \cap A_k) + \dots +$

$$(-1)^{n+1} Bel(A_1 \cap \dots \cap A_n).$$

When U is infinite, $\mathcal{P}(U)$ becomes a σ -algebra, (σ_U) . The function Bel also is required to be continuous from above in the sense that for any decreasing sequence $A_1 \supseteq A_2 \supseteq \dots$ in σ_U , if $A = \bigcap_{i=1}^{\infty} A_i \in \sigma_U$, then

$$\lim_{i \rightarrow \infty} Bel(A_i) = Bel(A). \quad (2.2)$$

The basic property of belief measures is thus a weaker version of the additive property of probability measures. Therefore, for any $A, A^c \subseteq U$, where A^c is the complement set of set A , we have

$$Bel(A) + Bel(A^c) \leq 1. \quad (2.3)$$

A plausibility measure, Pl , is defined by

$$Pl(A) = 1 - Bel(A^c), \quad \forall A \in \mathcal{P}(U). \quad (2.4)$$

Similarly,

$$Bel(A) = 1 - Pl(A^c), \quad \forall A \in \mathcal{P}(U). \quad (2.5)$$

The inequality (2.3) says that one's degree of belief that the truth lies in A together with his/her degree of doubt that the truth is not in A may not be able to capture the knowledge that s/he knows for sure the truth lies in $U = A \cup A^c$. S/he will say it is plausible that the truth lies in A , when s/he cuts off the doubt of A . The explanation will be clearer with an example.

Example 1. Consider an opinion poll for a Colorado governor's election. Let $U = \{a, b, c, d, e\}$ be the set of candidates. There are 10,000 individuals providing their preferences. They may not have made their final choice, since the poll takes place well before the election. Suppose that 3,500 individuals support candidates a and b from the Republican party, and 4,500 people support candidates c , d , and e from the Democratic party. The remaining 2,000 persons have no opinion yet. Therefore, we believe that one among the candidates from the Democratic party will become the new governor with the degree of belief 0.45, and for those who prefer that a Republican candidate will win, they doubt that the Democrat will win to the degree 0.35. That is, $Dou(\text{Democratic}) = Bel(\text{Democratic}^c) = Bel(\text{Republican}) = 0.35$. Combine the degree of belief and the degree of doubt that the Democrat will win the Colorado governor election, we obtain $0.45 + 0.35 = 0.70 < 1$. It is also plausible that the Democrat will win with $0.45 + 0.20 = 0.65$ degree of plausibility when we assume that all 2,000 voters with no opinion finally choose one of the candidates from the Democratic party. This 0.65 degree of plausibility is obtained when we subtract the 0.35 degree of doubt from the total belief of 1 that the new governor is a person in the set U . \diamond

Belief and plausibility measures also can be characterized by a *basic probability assignment function* m , which is defined on $\mathcal{P}(U)$ to $[0, 1]$, such that

$$m(\emptyset) = 0 \quad \text{and} \quad \sum_{A \in \mathcal{P}(U)} m(A) = 1. \quad (2.6)$$

It is important to understand the meaning of the basic probability assignment function, and it is essential not to confuse $m(A)$ with the probability of occur-

rence of an event A . The value $m(A)$ expresses the proportion to which all available and relevant evidence supports the claim that a particular element u of U , whose characterization in terms of relevant attributes is deficient, belongs to the set A . For Example 1, an element u can be referred to as a candidate who will win the Colorado governor's election, that is $u \in \{a, b, c, d, e\}$, where $m(\{a, b\}) = 0.35$, $m(\{c, d, e\}) = 0.45$, and $m(\{a, b, c, d, e\}) = 0.20$. It is clear that $\{a, b\}$ and $\{c, d, e\}$ are subsets of $U = \{a, b, c, d, e\}$, but $m(\{a, b\})$ and $m(\{c, d, e\})$ are greater than $m(U)$. Hence, it is allowed to have $m(A) > m(B)$ even if $A \subseteq B$.

The value $m(A)$ pertains solely to the set A . It does not imply any additional claims regarding subsets of A . One also may see $m(A)$ as the amount of probability pending over elements of A without being assigned yet, by lack of knowledge. If we had perfect probabilistic knowledge, then for every element u in a finite set U , we would have $m(\{u\}) = Pr(\{u\})$, and $\sum_{u \in U} m(\{u\}) = 1$. Thus, $m(A) = 0$, when A is not a singleton subset of U . Here are some differences between probability distribution functions and basic probability assignment functions.

- When $A \subseteq B$, it is not required that $m(A) \leq m(B)$, while $Pr(A) \leq Pr(B)$.
- It is not required that $m(U) = 1$, while $Pr(U) = 1$.
- No relationship between $m(A)$ and $m(A^c)$ is required, while $Pr(A) + Pr(A^c) = 1$.

A basic probability assignment function m is an abstract concept that helps us create belief and plausibility measures. The reason to have such an abstract

concept is for the cases when the exact probability of all sets in the universe is not known. When we do not know the probability on all elements of the universe, but we have information on some collection of subsets, it is possible to define a belief and a plausibility based on this information using the assignment function. As long as we have an assignment function (2.6), we can construct belief and plausibility functions (see (2.7) and (2.8) below).

We call a set $A \in \mathcal{P}(U)$, where $m(A) > 0$, a *focal element* of m , and denote \mathcal{F} as the set of focal elements. For the associated basic assignment function m , the pair (\mathcal{F}, m) is called a *body of evidence* or *random set* in [16]. More details about random sets are in the next section. We can formulate belief and plausibility measures uniquely from a given basic assignment m . $\forall A \in \mathcal{P}(U)$

$$Bel(A) = \sum_{B|B \subseteq A} m(B), \quad (2.7)$$

$$Pl(A) = \sum_{B|A \cap B \neq \emptyset} m(B). \quad (2.8)$$

The basic assignment function $m(A)$ characterizes the *degree of evidence* or belief that a particular element u of U belongs to the set A , while $Bel(A)$ represents the *total evidence* or belief that the element u belongs to A as well as to the various subsets of A . The plausibility measure represents not only the total evidence or belief that the element in question belongs to set A or to any of its subsets, but also the additional evidence or belief associated with sets that overlap (have at least one element in common) with A . Hence,

$$Pl(A) \geq Bel(A), \quad \forall A \in \mathcal{P}(U). \quad (2.9)$$

Example 2. Consider the group of 2,000 individuals who did not have any opinion at first, from Example 1. Suppose 500 of them admit that they will vote for the candidate a , and the other 500 will vote for either b or d , but they want to have a closer look before making final choice. Let $A_1 = \{a, b\}$, $A_2 = \{c, d, e\}$, $A_3 = \{a\}$, and $A_4 = \{b, d\}$. Figure 2.1 shows the Venn diagram of these sets. From this latest information, we obtain $m(A_1) = 0.35$, $m(A_2) = 0.45$, $m(A_3) = 0.05$, $m(A_4) = 0.05$, and $m(U) = 0.10$. Then, for instance, $Bel(A_3) = m(A_3) = 0.05$, $Bel(A_1) = m(A_1) + m(A_3) = 0.40$, and $Pl(A_1) = m(A_1) + m(A_3) + m(A_4) + m(U) = 0.55$. \diamond

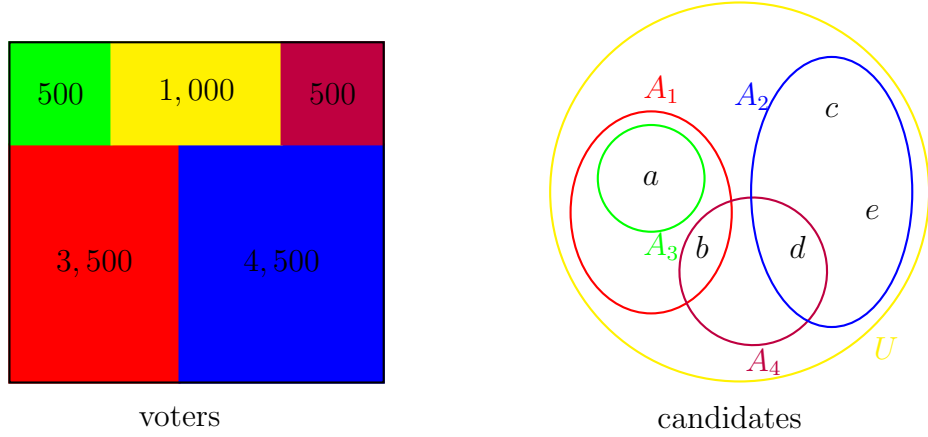


Figure 2.1: Venn diagram of an opinion poll for a Colorado governor's election.

The inverse procedure to get ' m ' from ' Bel ' is also possible. Given a belief measure Bel , the corresponding basic probability assignment m is determined for all $A \in \mathcal{P}(U)$ by the formula

$$m(A) = \sum_{B|B \subseteq A} (-1)^{|A-B|} Bel(B). \quad (2.10)$$

Hence, the basic probability assignment of the set A_1 in Example 2 is $m(A_1) = Bel(A_1) - Bel(A_3) = 0.40 - 0.05 = 0.35$, as expected. The proof of (2.10) can

be found in Appendix A.

Now, we have enough material to develop some details of possibility theory. A possibility measure is a plausibility in which all focal elements are nested. The sets defined in Example 2 are not nested. Therefore, information for Example 2 is not considered to be possibility information. Let a possibility measure and a necessity measure be denoted by the symbols Pos and Nec , respectively. The basic properties of possibility theory are the result of the nestedness of focal elements, and the proof can be found in Appendix B. These properties are as follows: $\forall A, B \in \mathcal{P}(U)$

$$Nec(A \cap B) = \min \{Nec(A), Nec(B)\}, \quad (2.11)$$

$$Pos(A \cup B) = \max \{Pos(A), Pos(B)\}. \quad (2.12)$$

However, no one has mentioned that if we have properties (2.11) and (2.12), then the nestedness of focal elements is preserved. Therefore, we prove the stronger statement (see Appendix B) that $Bel(A \cap B) = \min \{Bel(A), Bel(B)\}$, and $Pl(A \cup B) = \max \{Pl(A), Pl(B)\}$ if and only if the focal elements are nested.

The nestedness of focal elements guarantees that the meanings of necessity and possibility are preserved as in (2.11) and (2.12). A possibility measure also has a monotonicity property, which is in contrast with general basic probability assignment functions. That is

$$\text{if } A \subseteq B, \text{ then } Pos(A) \leq Pos(B).$$

This monotonicity property is a result of (2.12). It can be interpreted as when more evidence is provided, the possibility can never be lower than the possibility when less evidence is given.

Necessity measures and possibility measures satisfy these following properties since they are a special case of belief and plausibility measures:

$$Nec(A) + Nec(A^c) \leq 1, \quad (2.13)$$

$$Pos(A) + Pos(A^c) \geq 1, \quad (2.14)$$

$$Nec(A) + Pos(A^c) = 1, \quad (2.15)$$

$$\min \{Nec(A), Nec(A^c)\} = 0, \quad (2.16)$$

$$\max \{Pos(A), Pos(A^c)\} = 1, \quad (2.17)$$

$$Nec(A) > 0 \Rightarrow Pos(A) = 1, \quad (2.18)$$

$$Pos(A) < 1 \Rightarrow Nec(A) = 0. \quad (2.19)$$

The function $pos : U \rightarrow [0, 1]$ is defined on elements of the set U as

$$pos(u) = Pos(\{u\}), \quad \forall u \in U, \quad (2.20)$$

when a set-valued possibility measure Pos is given. It is called a *possibility distribution function* associated with Pos . We can construct a set-valued possibility measure from a possibility distribution in the following way. For each $A \in \mathcal{P}(U)$,

$$Pos(A) = \sup_{u \in A} \{pos(u)\}. \quad (2.21)$$

Let us start the process of constructing a possibility distribution from a basic probability assignment function m . This will help us to see the relationship between possibility/necessity measures and random sets. For simplicity, consider a set $U = \{u_1, u_2, \dots, u_n\}$. We will construct a possibility measure Pos defined on $\mathcal{P}(U)$ in terms of a given basic probability assignment m . Without loss

of generality, assume that the focal elements are some or all of the subsets in the sequence of nested subsets $A_1 \subseteq A_2 \subseteq \dots \subseteq A_n = U$, where $A_i = \{u_1, u_2, \dots, u_i\}$, $i = 1, 2, \dots, n$. We have

$$m(\emptyset) = 0, \quad \text{and} \quad \sum_{i=1}^n m(A_i) = 1.$$

However, it is not required that $m(A_i) \neq 0$ for all $i = 1, 2, \dots, n$. We define a possibility distribution function for each $u_i \in U$ using Equation (2.8) as

$$pos(u_i) = Pos(\{u_i\}) = \sum_{k=i}^n m(A_k), \quad i = 1, 2, \dots, n. \quad (2.22)$$

Written more explicitly, we have

$$\begin{aligned} pos(u_1) &= m(A_1) + m(A_2) + m(A_3) + \dots + m(A_i) + m(A_{i+1}) + \dots + m(A_n) \\ pos(u_2) &= m(A_2) + m(A_3) + \dots + m(A_i) + m(A_{i+1}) + \dots + m(A_n) \\ &\vdots \\ pos(u_i) &= m(A_i) + m(A_{i+1}) + \dots + m(A_n) \\ &\vdots \\ pos(u_n) &= m(A_n). \end{aligned}$$

This implies

$$m(A_i) = pos(u_i) - pos(u_{i+1}). \quad (2.23)$$

For example, suppose that $n = 7$. Figure 2.2 (see also in [74]) illustrates a possibility distribution constructed from the associated basic assignment defined as $m(A_1) = 0$, $m(A_2) = 0.3$, $m(A_3) = 0.4$, $m(A_4) = 0$, $m(A_5) = 0$, $m(A_6) = 0.1$, and $m(A_7) = 0.2$. The mathematical differences (for finite sets) between

probability and possibility theories are summarized in Table 2.1, which can be found in [74].

Example 3. A car manufacturer asks 100 customers on how many people they think should be able to ride in a new model car. The answers are that 4 (by 40 customers), 5 (by 30 customers), and 6 (by 30 customers) persons can ride in this car. We can interpret this information as a possibility information. This car can carry up to 4 persons is corresponding to $m(\{1, 2, 3, 4\}) = 0.4$. Similarly, we have $m(\{1, 2, 3, 4, 5\}) = 0.3$, and $m(\{1, 2, 3, 4, 5, 6\}) = 0.3$. Therefore, $pos(1) = pos(2) = pos(3) = pos(4) = 1$, $pos(5) = 0.6$, and $pos(6) = 0.3$. We can also check that the properties (2.11) - (2.19) are satisfied. \diamond

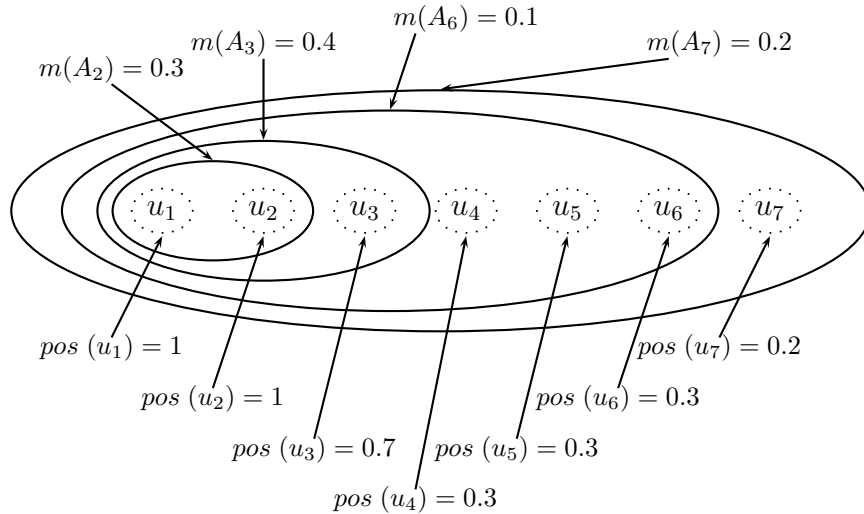


Figure 2.2: Possibility measure defined on $U = \{u_1, u_2, \dots, u_7\}$.

Klir [27] argued that probability is not adequate to capture the full scope of uncertainty, since probability theory is a subset of evidence theory when the focal elements are singletons, while possibility theory is a subset of evidence theory when the focal elements are nested. Hence, probability and possibility theories

are the same only when the body of evidence consists of one focal element that is a singleton.

Table 2.1: Probability theory versus possibility theory: comparison of mathematical properties for finite sets

Probability Theory	Possibility Theory
Based on measures of one type: Probability measures, Pr	Based on measures of two types: Possibility measures, Pos and Necessity measures, Nec
Body of evidence consists of singletons	Body of evidence consists of a family of nested subsets
Unique representation of Pr by a probability density function $pr : U \rightarrow [0, 1]$ via the formula $Pr(A) = \sum_{u \in A} pr(u)$	Unique representation of Pos by a possibility distribution function $pos : U \rightarrow [0, 1]$ via the formula $Pos(A) = \max_{u \in A} pos(u)$
Normalization: $\sum_{u \in U} pr(u) = 1$	Normalization: $\max_{u \in U} pos(u) = 1$
Additivity: $Pr(A \cup B) = Pr(A) + Pr(B) - Pr(A \cap B)$	Max rule: $Pos(A \cup B) = \max \{Pos(A), Pos(B)\}$ Min rule: $Nec(A \cap B) = \min \{Nec(A), Nec(B)\}$
Not applicable	$Nec(A) = 1 - Pos(A^c)$ $Pos(A) < 1 \Rightarrow Nec(A) = 0$ $Nec(A) > 0 \Rightarrow Pos(A) = 1$
$Pr(A) + Pr(A^c) = 1$	$Pos(A) + Pos(A^c) \geq 1$; $Nec(A) + Nec(A^c) \leq 1$ $\max \{Pos(A), Pos(A^c)\} = 1$ $\min \{Nec(A), Nec(A^c)\} = 0$
Total ignorance: $pr(u) = \frac{1}{ U }$ for all $u \in U$	Total ignorance: $pos(u) = 1$ for all $u \in U$

The nestedness property in possibility theory guarantees the maximum axiom (rather than additivity axiom in probability theory). Therefore, these two theories are not contained in one another; they are distinct. Moreover, probability and possibility theories involve different types of uncertainty and should be applied differently. For example, probability is suitable for characterizing the number of persons that are expected to ride in a particular car each day. Pos-

sibility theory, on the other hand, is suitable for characterizing the number of persons that can ride in that car at any one time. Since the physical characteristics of persons such as size or weight are intrinsically vague, it is not realistic to describe the situation by sharp probabilistic distinctions between possible and impossible instances. The readers who interest in more discussion can find full details in [27].

We explained how a basic probability assignment function m relates to belief and plausibility functions, given a finite set U . In the next section, the mathematical definition of random sets and some examples that help understanding the definition of random sets are provided. Theorem 2.14 shows that the belief (plausibility) and the necessity (possibility) functions are the same, when the focal elements are nested, (regardless the cardinality of U). We prove that belief and plausibility measures are the bounds on a set of probability measures generated by a random set. Finally, we show how to find the smallest and the largest expected values of a random variable with unknown probability from a given random set.

2.2 Random sets

We introduced in the previous section a random set as a body of evidence when U is a finite set of realizations. In this section, we give a more solid mathematical definition of a random set. Although, we limit the scope of this dissertation to a finite set U , it is worth explaining random sets in this section as generally as possible, so that we have some useful material for future research. A random set is a random variable such that an element in its range is a subset of a set X (finite or infinite). The differences between random variables and random

sets from their definitions can be seen in Definitions 2.4 and 2.5. First of all, we provide the definitions of a σ -algebra, a measurable space, a measurable mapping, and a probability space, which will be used in the definitions of a random variable and a random set. Definitions 2.1 - 2.3 can be found in any measure theory book.

Definition 2.1 Let Ω be a non-empty set. A σ -**algebra** on Ω , denoted by σ_Ω , is a family of subsets of Ω that satisfies the following properties:

- $\emptyset \in \sigma_\Omega$,
- $B \in \sigma_\Omega \Rightarrow B^c \in \sigma_\Omega$, and
- $B_i \in \sigma_\Omega$, for any countable (or finite) subset B_i of $\sigma_\Omega \Rightarrow \cup_i B_i \in \sigma_\Omega$.

A pair (Ω, σ_Ω) is called a **measurable space**.

Definition 2.2 Let (Ω, σ_Ω) be a measurable space. By a **measure** on this space, we mean a function $\mu : \sigma_\Omega \rightarrow [0, \infty]$ with the properties:

- $\mu(\emptyset) = 0$, and
- if $B_i \in \sigma_\Omega$, $\forall i = 1, 2, \dots$, are disjoint, then $\mu\left(\bigcup_{i=1}^{\infty} B_i\right) = \sum_{i=1}^{\infty} \mu(B_i)$.

We refer to the triple $(\Omega, \sigma_\Omega, \mu)$ as a **measure space**. If $\mu(\Omega) = 1$, we refer to it as a **probability space** and write it as $(\Omega, \sigma_\Omega, Pr_\Omega)$, where Pr_Ω is a **probability measure**.

Definition 2.3 Let (Ω, σ_Ω) and (U, σ_U) be measurable spaces. A function $f : \Omega \rightarrow U$ is said to be a $(\sigma_\Omega, \sigma_U)$ -**measurable mapping** if $f^{-1}(A) = \{\omega \in \Omega : f(\omega) \in A\} \in \sigma_\Omega$, for each $A \in \sigma_U$.

Definition 2.4 (see Marín [43]) *Let $(\Omega, \sigma_\Omega, Pr_\Omega)$ be a probability space and (U, σ_U) be a measurable space. A **random variable** X is a $(\sigma_\Omega, \sigma_U)$ -measurable mapping*

$$X : \Omega \rightarrow U. \quad (2.24)$$

This mapping can be used to generate a probability measure on (U, σ_U) . The probability $Pr_U(A)$ for an event $A \in \sigma_U$ is given by

$$Pr_U(A) = Pr_\Omega(X^{-1}(A)) = Pr_\Omega(\{\omega \in \Omega : X(\omega) \in A\}), \quad (2.25)$$

where $\{\omega \in \Omega : X(\omega) \in A\} \in \sigma_\Omega$.

It follows that, for $A \in \sigma_U$,

$$Pr_U(A) := \sum_{\omega \in X^{-1}(A)} Pr_\Omega(\{\omega\}) \quad (\text{discrete case}) \quad (2.26)$$

$$:= \int_{X^{-1}(A)} \underbrace{dPr_\Omega(\omega)}_{\text{c.d.f.}} \quad (\text{continuous case}), \quad (2.27)$$

where c.d.f. stands for cumulative distribution function. $Pr_\Omega(\omega)$ in (2.27) is the cumulative distribution with respect to Ω , when $\omega \in \Omega$, while $Pr_\Omega(\{\omega\})$ is the probability measure on $\{\omega\}$, which is an element in σ_Ω .

Example 4. Let Ω be the set of all outcomes of tossing two fair dice, i.e., $\Omega = \{(1, 1), (1, 2), \dots, (6, 6)\}$, with the probability $Pr_\Omega(\{(i, j)\}) = \frac{1}{36}$, where $i, j = 1, 2, \dots, 6$. A mapping from Ω to the sum of each outcome is a random variable $X : (i, j) \mapsto i + j$, where the set $U = \{2, 3, \dots, 12\}$. The sets σ_Ω and σ_U are the power set of Ω and U , respectively. Let $A \in \sigma_U$. Then $Pr_U(A) = Pr_\Omega(\{(i, j) \in \Omega : X((i, j)) \in A\})$. For instance, suppose $A = \{2, 12\}$, then $Pr_U(\{2, 12\}) = Pr_\Omega(\{(1, 1), (6, 6)\}) = \frac{2}{36}$. \diamond

Definition 2.5 (see Marín [43]) *Let $(\Omega, \sigma_\Omega, Pr_\Omega)$ be a probability space and $(\mathcal{F}, \sigma_\mathcal{F})$ be a measurable space, where $\mathcal{F} \subseteq \sigma_U$, $U \neq \emptyset$, and (U, σ_U) is a measurable space. A **random set** Γ is a $(\sigma_\Omega, \sigma_\mathcal{F})$ -measurable mapping*

$$\begin{aligned} \Gamma : \Omega &\rightarrow \mathcal{F} \\ \omega &\mapsto \Gamma(\omega). \end{aligned} \tag{2.28}$$

We can generate a probability measure on $(\mathcal{F}, \sigma_\mathcal{F})$ by using the probability measure $m(\mathcal{R}) = Pr_\Omega \{\omega : \Gamma(\omega) \in \mathcal{R}\}$, where $\mathcal{R} \in \sigma_\mathcal{F}$. Therefore, when \mathcal{F} is finite and $\sigma_\mathcal{F}$ is the power set of \mathcal{F} , we have $\sum_{\gamma \in \mathcal{F}} m(\{\gamma\}) = 1$. Finally, we use the notation (\mathcal{F}, m) to represent a **finite or infinite random set** depending on the cardinality of \mathcal{F} .

Remark 2.6 *The probability m in Definition 2.5 is called a basic probability assignment function, which is distinct from a ‘regular’ probability on U . A basic probability assignment function is actually a probability on σ_U .*

Remark 2.7 *In practice, the set \mathcal{F} can be given so that $m(\{\gamma\}) > 0$, for all $\gamma \in \mathcal{F}$. For $\omega \in \Omega$, a set $\gamma := \Gamma(\omega) \in \mathcal{F}$ is called a focal element if $m(\{\gamma\}) > 0$. A set $\{\gamma \mid m(\{\gamma\}) > 0\}$ is called a focal set.*

Remark 2.8 Γ becomes a random variable: $\Gamma(\omega) = U(\omega)$, when U is a finite non-empty set and all elements of \mathcal{F} are singletons, i.e. $\mathcal{F} = \{\{u\} : u \in U\}$. However, we cannot know the exact value of $Pr_U(A)$ for any $A \in \sigma_U$, in general, when we have information in the terms of random sets, because elements of \mathcal{F} can be any subsets of U .

There are two cases that generate a finite random set: 1) when U is finite, we have that \mathcal{F} is finite, and 2) when U is infinite (discrete or continuous), but \mathcal{F} is finite. Hence, there are a finite number of focal elements. Each of them can be a finite or an infinite set. An infinite random set is generated from an infinite set U , which also can be digested into two situations: 1) when U is infinite discrete, and 2) when U is a continuous set. Each of an infinite number of focal elements of an infinite random set can be either finite or infinite. The following examples help us understand the definition of random sets.

Example 5. Finite random set when U is finite.

Let us revisit Example 1. Consider an opinion poll for a Colorado governor's election. The set of candidates is $U = \{a, b, c, d, e\}$. There is a population Ω of 10,000 individuals that provide their preferences. They may not have made a final choice, since the poll takes place well before the election. The opinion of individual $\omega \in \Omega$ is modeled by $\Gamma(\omega) \subseteq U$. Therefore, each person ' ω ' of 3,500 Republican supporters is modeled by $\Gamma(\omega) = \{a, b\}$, each person ' ω ' of 4,500 Democrat supporters is modeled by $\Gamma(\omega) = \{c, d, e\}$, and each person ' ω ' of 2,000 undecided individuals is modeled by $\Gamma(\omega) = \{a, b, c, d, e\} = U$. Hence, the finite random set (\mathcal{F}, m) is when $\mathcal{F} = \{\{a, b\}, \{c, d, e\}, U\}$, such that $m(\{\{a, b\}\}) = 0.35$, $m(\{\{c, d, e\}\}) = 0.45$, and $m(\{U\}) = 0.20$, where $m(\{\gamma\})$ is the proportion of opinions of the form $\Gamma(\omega) = \gamma$. \diamond

Remark 2.9 *If we use different group of 15,000 people to provide the survey for Example 5 (different information), we may obtain a totally different random set from what we got in Example 5. Therefore, we need to be clear about the*

sample space Ω we are using. However, if we add these 15,000 people to the other 10,000 that we asked for the survey (more information), we will increase the size of the sample space Ω , but the updated belief (or plausibility) still could be less than, greater than or equal to the belief (or plausibility) got from random set in Example 5, depends upon the survey of these 15,000 people.

On the other hand, consider the random set information for Example 5, this time we receive more details on those of 4,500 who support Republican candidates that 2,000 will vote for candidate c , and other 2,500 still do not have final decision, so each of their votes could go to one of either c, d , or e . In this case, the new information is based on the old information, but more specific on who will vote for c . It should not difficult to see that the updated belief will be greater than or equal to the belief from the old information. Moreover, the updated plausibility will be less than or equal to the old plausibility.

Remark 2.10 We provided an example of a coin where its fairness is in question, at the beginning of this chapter. This coin has random set information as follows. Let $H := \text{head}$, and $T := \text{tail}$. Then, $\Omega = \{H, T\}$ with $Pr_{\Omega} \{H, T\} = 1$. The set U also is $\{H, T\}$, which is the only focal element for the random set, i.e., $m(\{U\}) = m(\{\{H, T\}\}) = 1$. Therefore, $Bel(\{H\}) = Bel(\{T\}) = 0$, and $Pl(\{H\}) = Pl(\{T\}) = 1$. However, if we flip this coin 100,000 times and it turns head for 43,000 times (turns tail for 57,000 times), this information is different from the information $Pr_{\Omega} \{H, T\} = 1$. Therefore, the clarification of the set Ω is very important in the conclusion of Bel and Pl values. For this experiment, Ω refers to the set $\{\text{flip}_1, \text{flip}_2, \dots, \text{flip}_{100,000}\}$, with $Pr_{\Omega}(\{\text{flip}_i\}) = 10^{-5}$. The set $U = \{H, T\}$, with $m(\{H\}) = 0.43$ and $m(\{T\}) = 0.57$. Since, the

focal elements now are singleton, we can conclude that $Bel(\{H\}) = Pl(\{H\}) = Pr_U(\{H\}) = 0.43$, and $Bel(\{T\}) = Pl(\{T\}) = Pr_U(\{T\}) = 0.57$. Based on these 100,000 trials, the probability (experiment based) of a head is 0.43, and the probability (experiment based) of a tail is 0.57. This probability is based on the experiment, and we will use it to represent this coin when the only information we have is this experiment. It may over and/or under estimate the real probability of this coin.

We conclude the thought from Remarks 2.9 and 2.10 as follows: 1) the clarification of the set Ω is important for the conclusion of Bel and Pl values, and 2) if we receive more information by digesting the old information to become more specific, then $Bel_{old}(A) \leq Bel_{updated}(A) \leq Pl_{updated}(A) \leq Pl_{old}(A)$.

Example 6. Finite random set when U is infinite.

A round target is spinning with an unknown speed and placed in a dark room. It also has numbers $0, \frac{\pi}{2}, \pi$, and $\frac{3\pi}{2}$ to indicate all angles in $[0, 2\pi)$. John throws a dart to this target for a total of 1,000 trials. The result he delivers is that the darts on the pie areas between $[0, \frac{3\pi}{4})$, $[\frac{3\pi}{4}, \frac{3\pi}{2})$, and $(\frac{3\pi}{2}, 2\pi)$ of total 350, 467, and 233, respectively. To be consistent with Definition 2.5, we consider $\Omega = \{\omega \mid \omega = 1, 2, \dots, 1,000\}$, representing trial number $1, 2, \dots, 1,000$, where $Pr_{\Omega}(\omega) = \frac{1}{1,000}$. We also have $U = [0, 2\pi)$ and $\mathcal{F} = \{[0, \frac{3\pi}{4}), [\frac{3\pi}{4}, \frac{3\pi}{2}), (\frac{3\pi}{2}, 2\pi)\}$. The information John provided can be represented as $m(\{[0, \frac{3\pi}{4})\}) = 0.350$, $m(\{[\frac{3\pi}{4}, \frac{3\pi}{2})\}) = 0.467$, and $m(\{(\frac{3\pi}{2}, 2\pi)\}) = 0.233$. This information is not enough to be able to know that the proportion that the dart will lie on a selected area, e.g., the area between $[\frac{\pi}{4}, \frac{\pi}{3}]$, for the next trial. However, we can say that the dart will point on the pie area between $[\frac{\pi}{4}, \frac{\pi}{3}]$ with plausibility of 0.350. \diamond

Example 7. Infinite random set.

An urn contains N white and M black balls. Balls are drawn randomly, one at a time, until a black one is drawn. If we assume that each selected ball is replaced before the next one is drawn, each of these independent draws has probability $p = \frac{M}{N+M}$ for being a success (a black ball is drawn). Let Ω be the set of the number of draws needed to obtain a black ball, i.e., $\Omega = \{\omega : \omega = 1, 2, 3, \dots\}$ with probability $Pr_{\Omega}(\{n\}) = p(1-p)^{n-1}$, where $n = 1, 2, 3, \dots$. In addition, we have a round target as in Example 6. This time, John provides the information that he can relate the number of draws until getting a black ball and a pie area that the dart will hit as follows:

$$\begin{aligned} \omega = 1 &\Rightarrow \text{the dart can lie somewhere in the pie area } [0, \pi] \Rightarrow \Gamma(1) = [0, \pi], \\ \omega = 2 &\Rightarrow \text{the dart can lie somewhere in the pie area } [0, \frac{3\pi}{2}] \Rightarrow \Gamma(2) = [0, \frac{3\pi}{2}], \\ \omega = 3 &\Rightarrow \text{the dart can lie somewhere in the pie area } [0, \frac{7\pi}{4}] \Rightarrow \Gamma(3) = [0, \frac{7\pi}{4}], \end{aligned}$$

and so on. Thus, $U = [0, 2\pi]$, and $\mathcal{F} = \left\{ \left[0, \frac{(2^n-1)\pi}{2^{n-1}} \right] ; n = 1, 2, 3, \dots \right\}$. Hence, $m \left(\left\{ \left[0, \frac{(2^n-1)\pi}{2^{n-1}} \right] \right\} \right) = p(1-p)^{n-1}$, and $\sum_{n=1}^{\infty} m \left(\left\{ \left[0, \frac{(2^n-1)\pi}{2^{n-1}} \right] \right\} \right) = 1$. What is the probability that the dart lies somewhere in the pie area $[\frac{9\pi}{5}, \frac{11\pi}{6}]$? We cannot get the answer for this question with the information we have. The only thing we can say is that the dart may lie in the pie area $[\frac{9\pi}{5}, \frac{11\pi}{6}]$ with the degree of possibility $1 - \sum_{n=1}^3 p(1-p)^n - 1$. \diamond

Let U be a finite set, \mathcal{F} be the power set of U , $\sigma_{\mathcal{F}}$ be the power set of \mathcal{F} , and m be a probability basic assignment function on $\sigma_{\mathcal{F}}$. The other versions of the belief and plausibility measures for a finite random set, given a finite set U ,

can be derived as: $\forall A \subseteq U$,

$$Bel_{(\mathcal{F},m)}(A) = \sum_{\substack{\gamma_i \subseteq A \\ \gamma_i \in \mathcal{F}}} m(\{\gamma_i\}) \quad (2.29)$$

$$= m(\{\gamma_i : \gamma_i \subseteq A, \gamma_i \neq \emptyset\}) \quad (2.30)$$

$$= Pr_{\Omega} \{\omega : \Gamma(\omega) \subseteq A, \Gamma(\omega) \neq \emptyset\}, \quad (2.31)$$

and

$$Pl_{(\mathcal{F},m)}(A) = \sum_{\substack{\gamma_i \cap A \neq \emptyset \\ \gamma_i \in \mathcal{F}}} m(\{\gamma_i\}) \quad (2.32)$$

$$= m(\{\gamma_i : \gamma_i \cap A \neq \emptyset\}) \quad (2.33)$$

$$= Pr_{\Omega} \{\omega : \Gamma(\omega) \cap A \neq \emptyset\}, \quad (2.34)$$

We add the subscript (\mathcal{F}, m) of *Bel* and *Pl* functions to make clear that these functions are based on the random set (\mathcal{F}, m) . A random set (\mathcal{F}, m) can be a finite random set even when U is infinite. The belief and plausibility formulas for a finite random set, with respect to a random set whose U is infinite, are the same as (2.29 - 2.34). We will sometimes say (\mathcal{F}, m) is a finite random set without mentioning the cardinality of U . An infinite random set can happen when U is an infinite set. The belief and plausibility measures of any set $A \subseteq U$ for an infinite random set also are provided by (2.30 - 2.31) and (2.33 - 2.34), respectively.

Next we will use a continuous possibility distribution $pos(u)$ on $U \subseteq \mathbb{R}^n$, when U is a continuous set, to represent an infinite random set (\mathcal{F}, m) , where \mathcal{F} is the family of all α -cuts $A_{\alpha} = \{u \in U : pos(u) \geq \alpha, \alpha \in (0, 1]\}$.

Theorem 2.12 (The strong law of large numbers, [2]) *Let X_1, X_2, \dots be a sequence of independent and identically distributed random variables, each having a finite mean $\mu = E[X_i]$. Then*

$$\lim_{n \rightarrow \infty} \frac{X_1 + X_2 + \dots + X_n}{n} \rightarrow \mu, \text{ almost surely.} \quad (2.36)$$

The strong law of large numbers is used in the proof of Theorem 2.13.

Theorem 2.13 (see Marín [43]) *Let (\mathcal{F}_n, m_n) be a finite representation of an infinite random set (\mathcal{F}, m) as defined in the Definition 2.11. The belief (and plausibility) of the random set (\mathcal{F}_n, m_n) converges almost surely to the belief (and plausibility) of the random set (\mathcal{F}, m) as $n \rightarrow \infty$, i.e.,*

$$Bel_{(\mathcal{F}, m)}(A) = \lim_{n \rightarrow \infty} Bel_{(\mathcal{F}_n, m_n)}(A), \text{ and} \quad (2.37)$$

$$Pl_{(\mathcal{F}, m)}(A) = \lim_{n \rightarrow \infty} Pl_{(\mathcal{F}_n, m_n)}(A), \quad (2.38)$$

almost surely, $\forall A \in \sigma_U$.

The next theorem tells us that in general, if the focal elements are nested (regardless of the cardinality of the focal set \mathcal{F}), then the belief measure is the necessity measure, and the plausibility measure is the possibility measure.

Theorem 2.14 (see Marín [43]) *Let (\mathcal{F}_n, m_n) be a finite representation of an infinite random set (\mathcal{F}, m) as defined in the Definition 2.11. Then*

$$Nec_U(A) = Bel_{(\mathcal{F}, m)}(A) = \lim_{n \rightarrow \infty} Bel_{(\mathcal{F}_n, m_n)}(A), \text{ and} \quad (2.39)$$

$$Pos_U(A) = Pl_{(\mathcal{F}, m)}(A) = \lim_{n \rightarrow \infty} Pl_{(\mathcal{F}_n, m_n)}(A), \quad (2.40)$$

almost surely, $\forall A \in \sigma_U$.

We skip the proofs of Theorems 2.13 and 2.14. The readers can view the proofs from [43]. Theorems 2.13 and 2.14 will be useful for future research, when we consider the continuous case of U .

Let us define the set \mathcal{M}^i of probability measures associated with A^i as

$$\mathcal{M}^i = \{Pr^i : \sigma_U \rightarrow [0, 1] \mid Pr^i(A) = 1, \text{ whenever } A^i \subseteq A, A \in \sigma_U\}.$$

A finite random set (\mathcal{F}, m) can be interpreted as the set of probability measures, $\mathcal{M}_{\mathcal{F}}$, of the form $\mathcal{M}_{\mathcal{F}} = \left\{ Pr : Pr(A) = \sum_{i=1}^L m(A^i)Pr^i(A), A \in \sigma_U \right\}$, where each Pr^i belongs to the set \mathcal{M}^i of probability measures on A^i , as claimed by many authors, see [16] for example. Many papers, e.g. [13, 16, 55, 56], claim that for all $A \in \sigma_U$, $Bel(A) \leq Pr(A) \leq Pl(A)$, $\forall Pr \in \mathcal{M}_{\mathcal{F}}$. We state the theorem and proof here. The details are as follows. The term $m(\{\gamma\})$ will be written as $m(\gamma)$, $\forall \gamma \in \mathcal{F}$ in the rest of the dissertation, for convenience.

Theorem 2.15 *Let \hat{u} be an uncertainty with a set of realizations U , where U can be finite or infinite, and σ_U is given. Suppose that (\mathcal{F}, m) is a finite random set of \hat{u} , where the focal set is $\mathcal{F} = \{A^i \in \sigma_U\}$, $i = 1, 2, \dots, L$, for some $L \in \mathbb{N}$. Then, the random set can be interpreted as the unique set of probability measure $\mathcal{M}_{\mathcal{F}}$ of the form*

$$\mathcal{M}_{\mathcal{F}} = \left\{ Pr \mid Pr(A) = \sum_{i=1}^L m(A^i)Pr^i(A), A \in \sigma_U \right\}, \quad (2.41)$$

where each Pr^i belongs to the set \mathcal{M}^i of probability measures associated with A^i , such that for all $A \in \sigma_U$, $Bel(A) \leq Pr(A) \leq Pl(A)$, $\forall Pr \in \mathcal{M}_{\mathcal{F}}$. Additionally, we have $\inf_{Pr \in \mathcal{M}_{\mathcal{F}}} Pr(A) = Bel(A)$, and $\sup_{Pr \in \mathcal{M}_{\mathcal{F}}} Pr(A) = Pl(A)$.

Proof: Please note that Pr^i is an arbitrary element of \mathcal{M}^i . First, we verify that Pr in (2.41) satisfies the probability axioms as follows:

- Since $m(A^i)$ and $Pr^i(A)$ are in $[0, 1]$, for each $A^i \in \mathcal{F}$, $A \in \sigma_U$, and $Pr^i \in \mathcal{M}^i$, $0 \leq Pr(A) = \sum_{i=1}^L m(A^i)Pr^i(A) \leq 1$.
- $Pr(U) = \sum_{i=1}^L m(A^i) = 1$, because $Pr^i(U) = 1$ for all $Pr^i \in \mathcal{M}^i$, $i = 1, 2, \dots, L$.
- For any $A_1, A_2 \in \sigma_U$ such that $A_1 \cap A_2 = \emptyset$,

$$\begin{aligned}
Pr(A_1 \cup A_2) &= \sum_{i=1}^L m(A^i)Pr^i(A_1 \cup A_2) \\
&= \sum_{i=1}^L m(A^i) (Pr^i(A_1) + Pr^i(A_2)) \\
&= \sum_{i=1}^L m(A^i)Pr^i(A_1) + \sum_{i=1}^L m(A^i)Pr^i(A_2) \\
&= Pr(A_1) + Pr(A_2).
\end{aligned}$$

Define $\underline{Pr}(A) = \inf_{Pr \in \mathcal{M}_{\mathcal{F}}} Pr(A)$ and $\overline{Pr}(A) = \sup_{Pr \in \mathcal{M}_{\mathcal{F}}} Pr(A)$, therefore

$$\begin{aligned}
\underline{Pr}(A) &= \inf \left\{ Pr(A) = \sum_{i=1}^L m(A^i)Pr^i(A), Pr^i \in \mathcal{M}^i \right\}, A \in \sigma_U, \quad \text{and} \\
\overline{Pr}(A) &= \sup \left\{ Pr(A) = \sum_{i=1}^L m(A^i)Pr^i(A), Pr^i \in \mathcal{M}^i \right\}, A \in \sigma_U.
\end{aligned}$$

We note that these two functions may not follow the axioms of probability. We can see that for a given $A \in \sigma_U$, the value of $\sum_{i=1}^L m(A^i)Pr^i(A)$ depends on the value of $Pr^i(A)$. Therefore, $\sum_{i=1}^L m(A^i)Pr^i(A)$ is the smallest value when $Pr^i(A) = 0$, as long as it does not violate the restriction of the set \mathcal{M}^i . It is not hard to see that a specific probability measure Pr_1^i , for each $i = 1, 2, \dots, L$,

$$Pr_1^i(A) = \begin{cases} 1; & \text{if } A^i \subseteq A, \\ 0; & \text{otherwise,} \end{cases}$$

provide

$$\underline{Pr}(A) = \sum_{A^i \subseteq A} m(A^i) = Bel(A).$$

Similarly, $\sum_{i=1}^L m(A^i)Pr^i(A)$ is the largest value when $Pr^i(A) = 1$, as long as it does not violate the restriction of the set \mathcal{M}^i . If this set A is such that $A^i \cap A = \emptyset$ and $Pr^i(A) > 0$, it implies that $Pr^i(A^i) < 1$, since $Pr^i(A^i) = 1 - Pr^i(A)$. Hence this $Pr^i \notin \mathcal{M}^i$. Therefore, $Pr^i(A) = 0$, when $A^i \cap A = \emptyset$, and $Pr^i \in \mathcal{M}^i$. Moreover, if $A^i \cap A \neq \emptyset$, there is $Pr^i \in \mathcal{M}^i$ such that $Pr^i(A) = 1 = Pr^i(A^i \cap A)$, because this Pr^i satisfies $Pr^i(B) = 1$, whenever $A^i \subseteq B$ (i.e., $A^i \cap B \neq \emptyset$). Hence,

$$\overline{Pr}(A) = \sum_{A^i \cap A \neq \emptyset} m(A^i) = Pl(A),$$

by using a specific probability measure Pr_2^i , for each $i = 1, 2, \dots, L$,

$$Pr_2^i(A) = \begin{cases} 1; & \text{if } A^i \cap A \neq \emptyset, \\ 0; & \text{otherwise.} \end{cases}$$

Next, we show that $\mathcal{M}_{\mathcal{F}}$ is the unique set of probability measures that has property $Bel(A) \leq Pr(A) \leq Pl(A)$, $\forall A \in \sigma_U$. Suppose that $P : \sigma_U \rightarrow [0, 1]$ is a probability measure such that $Bel(A) \leq P(A) \leq Pl(A)$, $\forall A \in \sigma_U$. Let $A \in \sigma_U$. Then,

$$Bel(A) = \sum_{i=1}^L m(A^i)Pr_1^i(A) \leq P(A) \leq \sum_{i=1}^L m(A^i)Pr_2^i(A) = Pl(A).$$

Without loss of generality, we suppose that there exists $l \leq L$ such that $A^i \subseteq A$, for each $i = 1, 2, \dots, l$. Therefore, $Bel(A) = \sum_{i=1}^l m(A^i) = \sum_{i=1}^l m(A^i)Pr_1^i(A)$. Since $A^i \not\subseteq A$, for all $i = l+1, l+2, \dots, L$, we suppose further that there exists $l < k \leq L$ such that $A^i \cap A \neq \emptyset$, for all $i = l+1, l+2, \dots, k$. Hence,

$$Bel(A) = \sum_{i=1}^l m(A^i)Pr_1^i(A) \leq P(A) \leq \sum_{i=1}^k m(A^i)Pr_2^i(A) = Pl(A).$$

If $l = k$, then $l = k = L$ and $P(A) = Bel(A) = Pl(A)$. We are done. Otherwise, set $a = P(A) - Bel(A) > 0$. $Pr^i(A) = \frac{a}{(k-l)m(A^i)} = Pr^i(A \cap A^i)$, $\forall i = l+1, l+2, \dots, k$, does not violate the restriction of \mathcal{M}^i . Thus, this $Pr^i \in \mathcal{M}^i$, for each $i = l+1, l+2, \dots, k$. Moreover, $Pr^i(A) = 0$, $\forall i = k+1, k+2, \dots, L$ also does not violate the restriction of \mathcal{M}^i , $i = k+1, k+2, \dots, L$, since $A^i \cap A = \emptyset$. Hence,

$$\begin{aligned} P(A) &= Bel(A) + a + 0 \\ &= \sum_{i=1}^l m(A^i)Pr_1^i(A) + m(A^{l+1})Pr^{l+1}(A) + \dots + m(A^k)Pr^k(A) + \\ &\quad m(A^{k+1})Pr^{k+1}(A) + \dots m(A^L)Pr^L(A) \end{aligned}$$

satisfies $Bel(A) \leq P(A) \leq Pl(A)$, $\forall A \in \sigma_U$ and is in the form of (2.41). \square

It follows from Theorems 2.13 and 2.14 that $\inf_{Pr \in \mathcal{M}_{\mathcal{F}}} Pr(A)$ and $\sup_{Pr \in \mathcal{M}_{\mathcal{F}}} Pr(A)$, where $\mathcal{M}_{\mathcal{F}}$ is interpreted from a (finite or infinite) random set induced by a possibility distribution are belief and plausibility of A , $\forall A \in \sigma_U$. We state this result in the following Corollary 2.16.

Corollary 2.16 *Assume*

H1: $(\Omega, \sigma_{\Omega}, Pr_{\Omega})$ is a probability space,

H2: U is an arbitrary set of consideration, and

H3: (\mathcal{F}, m) is a random set generated by a possibility distribution.

Then, we have the same conclusion as in Theorem 2.15.

Theorem 2.15 and Corollary 2.16 mean that a random set (\mathcal{F}, m) can be interpreted as a set of probability measures $\mathcal{M}_{\mathcal{F}}$ whose bounds on probability are

belief and plausibility measures. By these bounds, we can find a probability density mass function that generates the smallest (largest) expected value of our random variable.

2.2.1 Upper and lower expected values generated by random sets

Given a random set (\mathcal{F}, m) of $U = \{u_1, u_2, \dots, u_n\}$, and an evaluation function θ of U , where $\theta(u_1) \leq \theta(u_2) \leq \dots \leq \theta(u_n)$, the lowest and the largest expected values of θ can be evaluated by using the following density functions \underline{f} and \bar{f} of a random variable $U : \Omega \rightarrow U$,

$$\text{where } \left. \begin{aligned} \underline{f}(u_1) &= Bel(\{u_1, u_2, \dots, u_n\}) - Bel(\{u_2, u_3, \dots, u_n\}) \\ &\vdots \\ \underline{f}(u_i) &= Bel(\{u_i, u_{i+1}, \dots, u_n\}) - Bel(\{u_{i+1}, u_{i+2}, \dots, u_n\}) \\ &\vdots \\ \underline{f}(u_n) &= Bel(\{u_n\}), \end{aligned} \right\} \quad (2.42)$$

and

$$\left. \begin{aligned} \bar{f}(u_1) &= Bel(\{u_1\}) \\ &\vdots \\ \bar{f}(u_i) &= Bel(\{u_1, u_2, \dots, u_i\}) - Bel(\{u_1, u_2, \dots, u_{i-1}\}) \\ &\vdots \\ \bar{f}(u_n) &= Bel(\{u_1, u_2, \dots, u_n\}) - Bel(\{u_1, u_2, \dots, u_{n-1}\}). \end{aligned} \right\} \quad (2.43)$$

These two functions \underline{f} and \bar{f} are indeed density functions because each of them sums to 1, and are nonnegative. They are important because they are density functions that create the smallest and largest expected values $E(\theta)$, respectively.

We define Pr_g as a probability measure generated by a density function g . It was proved in [54] by Nguyen that \underline{f} in (2.42) provides the smallest expected value of θ . We combine his proof together with our contribution to show that \bar{f}

in (2.43) provides the largest expected value in Theorem 2.17 below. Example 8 illustrates the use of Theorem 2.17.

Theorem 2.17 *For a given random set (\mathcal{F}, m) of $U = \{u_1, u_2, \dots, u_n\}$, let $\theta : U \rightarrow \mathbb{R}$ be such that $\theta(u_1) \leq \theta(u_2) \leq \dots \leq \theta(u_n)$, and $E_f(\theta)$ be an expected value of the evaluation θ with respect to the density mass function f . Then the density functions \underline{f} and \bar{f} defined in Equations (2.42) and (2.43) have the property that*

$$E_{\underline{f}}(\theta) = \inf \{E_f(\theta) : Pr_f \in \mathcal{M}_{\mathcal{F}}\} \quad (2.44)$$

$$E_{\bar{f}}(\theta) = \sup \{E_f(\theta) : Pr_f \in \mathcal{M}_{\mathcal{F}}\}, \quad (2.45)$$

where \underline{f} and $\bar{f} \in \mathcal{M}_{\mathcal{F}}$, and $\mathcal{M}_{\mathcal{F}}$ is defined by Theorem 2.15.

Proof : We first show that $Pr_{\underline{f}}$ and $Pr_{\bar{f}}$ are in $\mathcal{M}_{\mathcal{F}}$. From (2.42) and (2.43), we have

$$\begin{aligned} \underline{f}(u_i) &= Bel(u_i, u_{i+1}, \dots, u_n) - Bel(u_{i+1}, u_{i+2}, \dots, u_n) \\ &= \sum_{B \subseteq \{u_i, u_{i+1}, \dots, u_n\}} m(B) - \sum_{B \subseteq \{u_{i+1}, u_{i+2}, \dots, u_n\}} m(B) \\ &= \sum_{\substack{B \subseteq \{u_i, u_{i+1}, \dots, u_n\}, \\ u_i \in B}} m(B), \end{aligned}$$

and

$$\begin{aligned} \bar{f}(u_i) &= Bel(u_1, u_2, \dots, u_i) - Bel(u_1, u_2, \dots, u_{i-1}) \\ &= \sum_{B \subseteq \{u_1, u_2, \dots, u_i\}} m(B) - \sum_{B \subseteq \{u_1, u_2, \dots, u_{i-1}\}} m(B) \\ &= \sum_{\substack{B \subseteq \{u_1, u_2, \dots, u_i\} \\ u_i \in B}} m(B). \end{aligned}$$

Let $A = \{u_i, u_{i+k_1}, \dots, u_{i+k_j}\}$, then

$$Pr_{\underline{f}}(A) = \sum_{u_i \in A} \underline{f}(u_i) = \sum_{\substack{B \subseteq \{u_i, \dots, u_n\} \\ u_i \in B}} m(B) + \sum_{\substack{B \subseteq \{u_{i+k_1}, \dots, u_n\} \\ u_{i+k_1} \in B}} m(B) + \dots + \sum_{\substack{B \subseteq \{u_{i+k_j}, \dots, u_n\} \\ u_{i+k_j} \in B}} m(B) . \quad (2.46)$$

We will show that $Pr_{\underline{f}}(A) \geq Bel(A) = \sum_{B \subseteq A} m(B)$. Let $B \subseteq A$.

- If $u_i \in B$, then $m(B)$ is included in the first term of (2.46).
- If $u_i \notin B$, and $u_{i+k_1} \in B$, then $m(B)$ is included in the second term of (2.46).
- \vdots
- If $u_i, u_{i+k_1}, \dots, u_{i+k_{j-1}} \notin B$, and $u_{i+k_j} \in B$, then $m(B)$ is included in the last term of (2.46).

Hence, $Pr_{\underline{f}}(A) \geq Bel(A)$. Moreover, since $B \cap A \neq \emptyset$ for each set B in any terms of (2.46), $Pr_{\underline{f}}(A) \leq Pl(A)$. Therefore, $Pr_{\underline{f}} \in \mathcal{M}_{\mathcal{F}}$. Likewise,

$$Pr_{\bar{f}}(A) = \sum_{u_i \in A} \bar{f}(u_i) = \sum_{\substack{B \subseteq \{u_1, \dots, u_i\} \\ u_i \in B}} m(B) + \sum_{\substack{B \subseteq \{u_1, \dots, u_{i+k_1}\} \\ u_{i+k_1} \in B}} m(B) + \dots + \sum_{\substack{B \subseteq \{u_1, \dots, u_{i+k_j}\} \\ u_{i+k_j} \in B}} m(B) . \quad (2.47)$$

We will show that $Pr_{\bar{f}}(A) \geq Bel(A) = \sum_{B \subseteq A} m(B)$. Let $B \subseteq A$.

- If $u_{i+k_j} \in B$, then $m(B)$ is included in the last term of (2.47).
- If $u_{i+k_j} \notin B$, and $u_{i+k_{j-1}} \in B$, then $m(B)$ is included in the second last term of (2.47).
- \vdots
- If $u_{i+k_j}, u_{i+k_{j-1}}, \dots, u_{i+k_1} \notin B$, and $u_i \in B$, then $m(B)$ is included in the first term of (2.47).

Hence, $Pr_{\bar{f}}(A) \geq Bel(A)$. Moreover, since $B \cap A \neq \emptyset$ for each set B in any terms of (2.47), $Pr_{\bar{f}}(A) \leq Pl(A)$. Therefore, $Pr_{\bar{f}} \in \mathcal{M}_{\mathcal{F}}$.

To prove (2.44) and (2.45), we apply the well-known expected value formula (see [54] and [59]),

$$E_f(\theta) = \int_0^{\infty} Pr_f(\{u \mid \theta(u) > t\}) dt + \int_{-\infty}^0 (Pr_f(\{u \mid \theta(u) > t\}) - 1) dt. \quad (2.48)$$

Let $\{u \mid \theta(u) > t\} = \{u_i, u_{i+1}, \dots, u_n\}$, for some $i = 1, 2, \dots, n$. We know that

$$Pr_{\underline{f}}(\{u \mid \theta(u) > t\}) = Pr_{\underline{f}}(\{u_i, \dots, u_n\}) = \sum_{j=i}^n \underline{f}(u_j) = Bel(\{u_i, \dots, u_n\}) \quad (2.49)$$

$$Pr_{\overline{f}}(\{u \mid \theta(u) > t\}) = Pr_{\overline{f}}(\{u_i, \dots, u_n\}) = \sum_{j=i}^n \overline{f}(u_j) = Pl(\{u_i, \dots, u_n\}). \quad (2.50)$$

The last equality of (2.50) holds, since

$$\begin{aligned} \sum_{j=i}^n \overline{f}(u_j) &= (Bel(\{u_1, u_2, \dots, u_i\}) - Bel(\{u_1, u_2, \dots, u_{i-1}\})) + \\ &\quad (Bel(\{u_1, u_2, \dots, u_{i+1}\}) - Bel(\{u_1, u_2, \dots, u_i\})) + \dots + \\ &\quad (Bel(\{u_1, u_2, \dots, u_n\}) - Bel(\{u_1, u_2, \dots, u_{n-1}\})) \\ &= Bel(\{u_1, u_2, \dots, u_n\}) - Bel(\{u_1, u_2, \dots, u_{i-1}\}) \\ &= 1 - Bel(\{u_1, u_2, \dots, u_{i-1}\}) \\ &= Pl(\{u_i, u_{i+1}, \dots, u_n\}). \end{aligned}$$

By Theorem 2.15, for every $Pr_f \in \mathcal{M}_{\mathcal{F}}$,

$$Pr_{\underline{f}}(\{u \mid \theta(u) > t\}) \leq Pr_f(\{u \mid \theta(u) > t\}) \leq Pr_{\overline{f}}(\{u \mid \theta(u) > t\}).$$

Without loss of generality, suppose that there exists $k < n$ such that $\theta(u_k) < 0$ and $\theta(u_{k+1}) \geq 0$. We divide t into the following subintervals

- $t \in (-\infty, \theta(u_1)) \Rightarrow \{u \mid \theta(u) > t\} = U,$
- $t \in [\theta(u_1), \theta(u_2)) \Rightarrow \{u \mid \theta(u) > t\} = \{u_2, u_3, \dots, u_n\},$
- \vdots
- $t \in [\theta(u_{k-1}), \theta(u_k)) \Rightarrow \{u \mid \theta(u) > t\} = \{u_k, u_{k+1}, \dots, u_n\},$
- $t \in [\theta(u_k), 0) \Rightarrow \{u \mid \theta(u) > t\} = \{u_{k+1}, u_{k+2}, \dots, u_n\},$

- $t \in [0, \theta(u_{k+1})) \Rightarrow \{u \mid \theta(u) > t\} = \{u_{k+1}, u_{k+2}, \dots, u_n\},$
- \vdots
- $t \in [\theta(u_{n-1}), \theta(u_n)) \Rightarrow \{u \mid \theta(u) > t\} = \{u_n\},$
- $t \in [\theta(u_n), \infty) \Rightarrow \{u \mid \theta(u) > t\} = \emptyset.$

Applying \underline{f} to (2.48), we have

$$\begin{aligned}
E_{\underline{f}}(\theta) &= \int_0^\infty Pr_{\underline{f}}(\{u \mid \theta(u) > t\}) dt + \int_{-\infty}^0 (Pr_{\underline{f}}(\{u \mid \theta(u) > t\}) - 1) dt \\
&= M_1 + M_2, \text{ where} \\
M_1 &= \int_0^{\theta(u_{k+1})} Bel(\{u_{k+1}, u_{k+2}, \dots, u_n\}) dt + \dots + \\
&\quad \int_{\theta(u_{n-1})}^{\theta(u_n)} Bel(\{u_n\}) dt + \int_{\theta(u_n)}^\infty Bel(\emptyset) dt. \\
M_2 &= \int_{-\infty}^{\theta(u_1)} (Bel(U) - 1) dt + \int_{\theta(u_1)}^{\theta(u_2)} (Bel(\{u_2, u_3, \dots, u_n\}) - 1) dt \\
&\quad + \dots + \int_{\theta(u_{k-1})}^{\theta(u_k)} (Bel(\{u_k, u_{k+1}, \dots, u_n\}) - 1) dt \\
&\quad + \int_{\theta(u_k)}^0 (Bel(\{u_{k+1}, u_{k+2}, \dots, u_n\}) - 1) dt.
\end{aligned}$$

We can see that M_1 is the smallest positive value and M_2 is the largest negative value associated with random set information. Hence Equation (2.44) hold.

Similarly, we also can apply \bar{f} to (2.48), and obtain that Equation(2.45) holds.

□

Example 8. Let Ω be the set of outcomes from tossing a die where we know only $Pr_\Omega(\{1, 6\}) = \frac{1}{3}$. Each face of this die is painted by one of the colors Black (B), Red (R), or White (W). However, we cannot see the die because it is in a dark box. Suppose that only information we have is

$$\{1, 6\} \rightarrow \{B, R, W\}, \quad \text{and} \quad \{2, 3, 4, 5\} \rightarrow \{R\},$$

i.e., we know only that faces 2, 3, 4, and 5 are all painted by Red. However, we do not know which color (B, R, or W) is used for painting face 1, and which color (B, R, or W) is used for painting face 6. We will pay \$1, \$2, or \$3 if B, R, or W appears, respectively, i.e.,

$$\theta(B) = \$1, \quad \theta(R) = \$2, \quad \text{and} \quad \theta(W) = \$3.$$

The random set (\mathcal{F}, m) for this situation is $\mathcal{F} = \{\{R\}, \{B, R, W\}\}$, where $m(\{R\}) = 2/3$ and $m(\{B, R, W\}) = 1/3$. The focal elements are nested, so we can create possibility and necessity measures using this random set.

$$\begin{array}{llll} Pos(\{B\}) & = \frac{1}{3} & Nec(\{B\}) & = 0 \\ Pos(\{R\}) & = 1 & Nec(\{R\}) & = \frac{2}{3} \\ Pos(\{W\}) & = \frac{1}{3} & Nec(\{W\}) & = 0 \\ Pos(\{B, R\}) & = 1 & Nec(\{B, R\}) & = \frac{2}{3} \\ Pos(\{B, W\}) & = \frac{1}{3} & Nec(\{B, W\}) & = 0 \\ Pos(\{R, W\}) & = 1 & Nec(\{R, W\}) & = \frac{2}{3} \\ Pos(\{B, R, W\}) & = 1 & Nec(\{B, R, W\}) & = 1. \end{array}$$

The density mass functions \underline{f} and \bar{f} require the calculations of four *Bel*'s: $Bel(\{B\})$, $Bel(\{W\})$, $Bel(\{R, W\})$, and $Bel(\{B, R\})$, i.e.,

$$\underline{f}(\$1) = \underline{f}(\{B\}) = Bel(\{B, R, W\}) - Bel(\{R, W\}) = 1 - \frac{2}{3} = \frac{1}{3},$$

$$\underline{f}(\$2) = \underline{f}(\{R\}) = Bel(\{R, W\}) - Bel(\{W\}) = \frac{2}{3} - 0 = \frac{2}{3},$$

$$\underline{f}(\$3) = \underline{f}(\{W\}) = Bel(\{W\}) = 0,$$

and

$$\bar{f}(\$1) = \bar{f}(\{B\}) = Bel(\{B\}) = 0,$$

$$\bar{f}(\$2) = \bar{f}(\{R\}) = Bel(\{B, R\}) - Bel(\{B\}) = \frac{2}{3} - 0 = \frac{2}{3},$$

$$\bar{f}(\$3) = \bar{f}(\{W\}) = Bel(\{B, R, W\}) - Bel(\{B, R\}) = 1 - \frac{2}{3} = \frac{1}{3}.$$

With respect to the information we have, \underline{f} provides the smallest $E(\theta) = 1 \cdot \frac{1}{3} + 2 \cdot \frac{2}{3} + 3 \cdot 0 = \1.67 , and \bar{f} gives the largest $E(\theta) = 1 \cdot 0 + 2 \cdot \frac{2}{3} + 3 \cdot \frac{1}{3} = \2.33 .

However, if we were to use an LP problem (2.51) to find the lowest (and largest) expected return value, we have $2(2^3 - 2) = 12$ terms of *Pl/Bel*'s to calculate. Moreover, we will need to solve 2 linear programs.

$$\begin{array}{ll}
 \min / \max & \$1 \times f(\text{B}) + \$2 \times f(\text{G}) + \$3 \times f(\text{Y}) \\
 \text{s.t.} & f(\text{B}) \in [\text{Bel}(\{\text{B}\}), \text{Pl}(\{\text{B}\})] \\
 & f(\text{G}) \in [\text{Bel}(\{\text{G}\}), \text{Pl}(\{\text{G}\})] \\
 & f(\text{Y}) \in [\text{Bel}(\{\text{Y}\}), \text{Pl}(\{\text{Y}\})] \\
 & f(\text{B}) + f(\text{G}) \in [\text{Bel}(\{\text{B}, \text{G}\}), \text{Pl}(\{\text{B}, \text{G}\})] \\
 & f(\text{B}) + f(\text{Y}) \in [\text{Bel}(\{\text{B}, \text{Y}\}), \text{Pl}(\{\text{B}, \text{Y}\})] \\
 & f(\text{G}) + f(\text{Y}) \in [\text{Bel}(\{\text{G}, \text{Y}\}), \text{Pl}(\{\text{G}, \text{Y}\})] \\
 & f(\text{B}) + f(\text{G}) + f(\text{Y}) = 1. \quad \diamond
 \end{array} \quad \left. \vphantom{\begin{array}{l} \\ \\ \\ \\ \\ \\ \\ \end{array}} \right\} \quad (2.51)$$

Let $U = \{u_1, u_2, \dots, u_n\}$ be the set of all realization of a random set uncertainty \hat{u} , and θ is an evaluation of U such that $\theta(u_1) \leq \theta(u_2) \leq \dots \leq \theta(u_n)$. Theorem 2.17 has advantage for finding the lowest and highest expected value of θ . Table 2.2 illustrates the number of belief and plausibility terms required for finding \underline{f} and \bar{f} by using Theorem 2.17 versus by solving two LP problems.

It is clear that \underline{f} and \bar{f} obtained by the construction (2.42) and (2.43), when random set information is given, require much less calculation than solving two linear programs $\min_{f \in \mathcal{M}_{\mathcal{F}}} / \max_{f \in \mathcal{M}_{\mathcal{F}}} E_f(\hat{u})$, where $\mathcal{M}_{\mathcal{F}} = \{\text{density function } f \text{ on } U : \text{Bel}(A) \leq \text{Pr}_f(A) \leq \text{Pl}(A), A \subseteq U\}$, because we need to find beliefs and plausibilities of all subsets A of U to be able to set up these LP problems. More-

Table 2.2: The number of belief and plausibility terms required for finding two density functions for the lowest and the highest expected values of θ by using Theorem 2.17 versus by solving two LP problems.

	Theorem 2.17	2 LP problems
<i>Bel</i> terms	$2(n - 1)$	$2^n - 2$
<i>Pl</i> terms	—	$2^n - 2$
Total	$2(n - 1)$	$2(2^n - 2)$

over, $\mathcal{M}_{\mathcal{F}}$ cannot be reduced to $\overline{\mathcal{M}} = \{\text{density function } f \text{ on } U : Bel(\{u_i\}) \leq Pr_f(\{u_i\}) \leq Pl(\{u_i\}), \forall u_i \in U\}$. Example 9 shows that optimal solutions of $\min_{f \in \overline{\mathcal{M}}} / \max_{f \in \overline{\mathcal{M}}} E_f(\hat{u})$ may not satisfy the random set information, since these optimal solutions may not be elements in $\mathcal{M}_{\mathcal{F}}$.

Example 9. Let $U = \{1, 2, 3, 4, 5, 6\}$ be a set of all realizations of an uncertainty \hat{u} , and suppose (\mathcal{F}, m) is a random set such that $m(\{1, 2, 3\}) = \frac{1}{2}$, $m(\{1, 4, 5\}) = \frac{1}{4}$, and $m(\{4, 6\}) = \frac{1}{4}$. Then, the construction (2.42) and (2.43) provides the smallest and the largest expected values of \hat{u} by using a probability density mass vectors $\underline{f} = [\frac{3}{4}, \frac{1}{4}, 0, 0, 0, 0]^T$, and $\overline{f} = [0, 0, \frac{1}{2}, 0, \frac{1}{4}, \frac{1}{4}]^T$. Furthermore, we have

$Bel(\{1\}) = Bel(\{2\}) = Bel(\{3\}) = Bel(\{4\}) = Bel(\{5\}) = Bel(\{6\}) = 0$, and

$$Pl(\{1\}) = \frac{3}{4}, Pl(\{2\}) = Pl(\{3\}) = Pl(\{4\}) = \frac{1}{2}, Pl(\{5\}) = Pl(\{6\}) = \frac{1}{4}.$$

An optimal solution of $\max_{f \in \overline{\mathcal{M}}} E_f(\hat{u})$ is $f_* = [0, 0, 0, \frac{1}{2}, \frac{1}{4}, \frac{1}{4}]^T$. However, Theorem 2.15 said that $Pr_{f_*}(\{1, 2, 3\}) \in [Bel(\{1, 2, 3\}), Pl(\{1, 2, 3\})] = [\frac{1}{2}, \frac{3}{4}]$, $\forall f \in \mathcal{M}_{\mathcal{F}}$, but $Pr_{f_*}(\{1, 2, 3\}) = 0 \notin [\frac{1}{2}, \frac{3}{4}]$. That is, optimal solutions for $\max_{f \in \overline{\mathcal{M}}} E_f(\hat{u})$ may not satisfy the random set information. \diamond

The next concern is about infinite random sets and finite random sets with an infinite set U . We try to build a similar theorem as Theorem 2.17 for the infinite case of U . Similar to Equations (2.49) and (2.50) in the proof of Theorem 2.17, if we can find density functions \underline{f} and \bar{f} such that

$$Pr_{\underline{f}}(\{u \mid \theta(u) > t\}) = Bel(\{u \mid \theta(u) > t\}), \text{ and}$$

$$Pr_{\bar{f}}(\{u \mid \theta(u) > t\}) = Pl(\{u \mid \theta(u) > t\}),$$

and show that these $Pr_{\underline{f}}$ and $Pr_{\bar{f}} \in \mathcal{M}_{\mathcal{F}}$, then we are done, because Bel and Pl are lower and upper bounds to probabilities as mentioned in Theorem 2.15. A plausibility measure has properties $Pl(\{u \mid \theta(u) \leq t_1\}) \leq Pl(\{u \mid \theta(u) \leq t_2\}) \forall t_1 \leq t_2$, and $Pl(U) = 1$. We can use a plausibility measure to represent a cumulative probability distribution function of θ . Likewise, a belief measure is used to construct a cumulative probability distribution function of θ . Therefore, we can get the density functions \underline{f} and \bar{f} from these cumulative probability distribution functions

$$Pr_{\underline{f}}(\{u \mid \theta(u) \leq t\}) = Pl(\{u \mid \theta(u) \leq t\}), \quad (2.52)$$

$$Pr_{\bar{f}}(\{u \mid \theta(u) \leq t\}) = Bel(\{u \mid \theta(u) \leq t\}), \quad (2.53)$$

respectively. It turns out that

$$\left. \begin{aligned} Pr_{\underline{f}}(\{u \mid \theta(u) > t\}) &= 1 - Pr_{\underline{f}}(\{u \mid \theta(u) \leq t\}) \\ &= 1 - Pl(\{u \mid \theta(u) \leq t\}) \\ &= Bel(\{u \mid \theta(u) > t\}) \end{aligned} \right\} \quad (2.54)$$

$$\left. \begin{aligned} Pr_{\bar{f}}(\{u \mid \theta(u) > t\}) &= 1 - Pr_{\bar{f}}(\{u \mid \theta(u) \leq t\}) \\ &= 1 - Bel(\{u \mid \theta(u) \leq t\}) \\ &= Pl(\{u \mid \theta(u) > t\}). \end{aligned} \right\} \quad (2.55)$$

Therefore, we have the following theorem.

Theorem 2.18 *Let a measurable space (U, σ_U) and a continuous function $\theta : U \rightarrow \mathbb{R}$ be given. Let (\mathcal{F}, m) be a random set with $\mathcal{F} \subseteq \sigma_U$. Then there exist density functions \underline{f} and $\bar{f} \in \mathcal{M}_{\mathcal{F}}$ that provide the smallest and the largest expected values of θ , and their probability distribution functions are $Pr_{\underline{f}}(\{u \mid \theta(u) \leq t\}) = Pl(\{u \mid \theta(u) \leq t\})$ and $Pr_{\bar{f}}(\{u \mid \theta(u) \leq t\}) = Bel(\{u \mid \theta(u) \leq t\})$, respectively.*

Proof: The proof of the smallest and largest expected values was presented before stating the theorem. We still need to show that \underline{f} and $\bar{f} \in \mathcal{M}_{\mathcal{F}}$. Given $A \in \sigma_U$, then $A \subseteq U$. We also can write $A := \{u \mid t_1 \leq \theta(u) \leq t_2\}$ for some $t_1 \leq t_2 \in \mathbb{R}$. We use Equations (2.53 - 2.55) to show that for any $t_1 \leq t_2 \in \mathbb{R}$,

$$Bel(A) \leq Pr_{\underline{f}}(A) \leq Pl(A), \quad \text{and} \quad (2.56)$$

$$Bel(A) \leq Pr_{\bar{f}}(A) \leq Pl(A). \quad (2.57)$$

Since

$$\begin{aligned} Pr_{\underline{f}}(A) &= Pr_{\underline{f}}(\{u \mid t_1 \leq \theta(u) \leq t_2\}) \\ &= Pr_{\underline{f}}(\{u \mid \theta(u) > t_1\}) - Pr_{\underline{f}}(\{u \mid \theta(u) > t_2\}) \\ &= Bel(A_1) - Bel(A_2); \text{ where } A_1 = \{u \mid \theta(u) > t_1\}, A_2 = \{u \mid \theta(u) > t_2\} \\ &= m(\{\gamma \in \mathcal{F} : \gamma \subseteq A_1\}) - m(\{\gamma \in \mathcal{F} : \gamma \subseteq A_2\}) \\ &= m(\{\gamma \in \mathcal{F} : \gamma \subseteq A_1, \gamma \not\subseteq A_2\}) \\ &\geq m(\{\gamma \in \mathcal{F} : \gamma \subseteq A_1 \setminus A_2\}) \\ &= Bel(\{u \mid t_1 \leq \theta(u) \leq t_2\}) \\ &= Bel(A), \end{aligned}$$

and

$$\begin{aligned}
Pr_{\underline{f}}(A) &= Pr_{\underline{f}}(\{u \mid t_1 \leq \theta(u) \leq t_2\}) \\
&= Pr_{\underline{f}}(\{u \mid \theta(u) \leq t_2\}) - Pr_{\underline{f}}(\{u \mid \theta(u) \leq t_1\}) \\
&= Pl(B_2) - Pl(B_1); \text{ where } B_1 = \{u \mid \theta(u) \leq t_1\}, B_2 = \{u \mid \theta(u) \leq t_2\} \\
&= m(\{\gamma \in \mathcal{F} : \gamma \cap B_2 \neq \emptyset\}) - m(\{\gamma \in \mathcal{F} : \gamma \cap B_1 \neq \emptyset\}) \\
&= m(\{\gamma \in \mathcal{F} : \gamma \cap B_2 \neq \emptyset, \gamma \cap B_1 = \emptyset\}) \\
&\leq m(\{\gamma \in \mathcal{F} : \gamma \cap (B_2 \setminus B_1) \neq \emptyset\}) \\
&= Pl(\{u \mid t_1 \leq \theta(u) \leq t_2\}) \\
&= Pl(A),
\end{aligned}$$

we obtain (2.56). Similarly, we also can show (2.57). \square

Uncertainty interpretations mentioned so far in this chapter are belief, plausibility, necessity, and possibility measures, and random set. A belief (plausibility, necessity, and possibility) measure assumes that there is a belief (plausibility, necessity, and possibility) about every subset of a non-empty finite set U . A random set needs the sum of the basic probability assignment function of all focal elements to be 1. Unfortunately, information we received may not have all subset details to define a belief (or other) measure. Likewise, we may have focal elements and their associated basic probability assignment values that do not add up to 1. Subsections 2.2.2 - 2.2.6, which also are our contribution in this dissertation, address these issues, so that we can continue using the constructions (2.42) and (2.43) to find the smallest and the largest expected value of \hat{u} , based on the partial information we have.

2.2.2 The random set generated by an incomplete random set information

Let $U = \{u_1, u_2, \dots, u_n\}$ be a set of all realizations of \hat{u} . An incomplete random set information is information containing focal elements and their associated basic probability assignment values that are not adding up to 1. Suppose this information is $m(A_i) = a_i > 0$, $\sum_{i=1}^L a_i < 1$, for some L . This partial information can belong to any random set that has A_i 's as its focal elements with $m(A_i) = a_i$. We can generate a random set from this partial information which guarantees the bounds of an unknown probability of each $\{u_k\}$, $k = 1, 2, \dots, n$. This random set has A_1, A_2, \dots, A_L, U as the focal elements with $m(A_i) = a_i$ and $m(U) = 1 - \sum_{i=1}^L a_i$. It is not difficult to see that this random set provides the largest set $\mathcal{M}_{\mathcal{F}}$, because this random set generates the widest interval $[Bel(A), Pl(A)]$, $\forall A \subseteq U$, and hence, it covers the unknown probability based on this partial information.

2.2.3 The random set generated by a possibility distribution or possibility of some subset of U

In some situations, we may receive only a possibility distribution instead of a possibility measure. We still are able to generate a unique random set corresponding to this possibility distribution, because possibility distribution implies possibility measure by (2.21). Hence, necessity measure is obtained. Then, we can use the formula $m(A) = \sum_{B|B \subseteq A} (-1)^{|A \setminus B|} Nec(B)$ to create the random set. More constructively, let $U = \{u_1, u_2, \dots, u_n\}$. Suppose $0 \leq a_1 < a_2 < \dots < 1$, and

$$\left. \begin{array}{l} \text{pos}(u_{k_i}) = a_1, \quad i = 1, 2, \dots, t_1, \\ \text{pos}(u_{k_i}) = a_2, \quad i = t_1 + 1, t_1 + 2, \dots, t_2, \\ \vdots \\ \text{pos}(u_{k_i}) = 1, \quad i = t_l + 1, t_l + 2, \dots, n. \end{array} \right\} \quad (2.58)$$

Let $A_j = \{u_{k_i}, i = t_{j-1} + 1, t_{j-1} + 2, \dots, t_j\}$, where $t_0 = 0$. Then, the basic probability assignment values of the associated focal elements are

$$\left. \begin{array}{l} m(U) = a_1, \\ m(U \setminus A_1) = a_2 - a_1, \\ m(U \setminus A_1 \setminus A_2) = a_3 - a_2, \\ \vdots \\ m(U \setminus A_1 \setminus A_2 \setminus \dots \setminus A_l) = 1 - a_l. \end{array} \right\} \quad (2.59)$$

Figure 2.4 illustrates a method to obtain a random set from a given possibility distribution when $U = \{u_1, u_2, u_3, u_4, u_5, u_6\}$. The dashed lines indicate the corresponding focal elements.

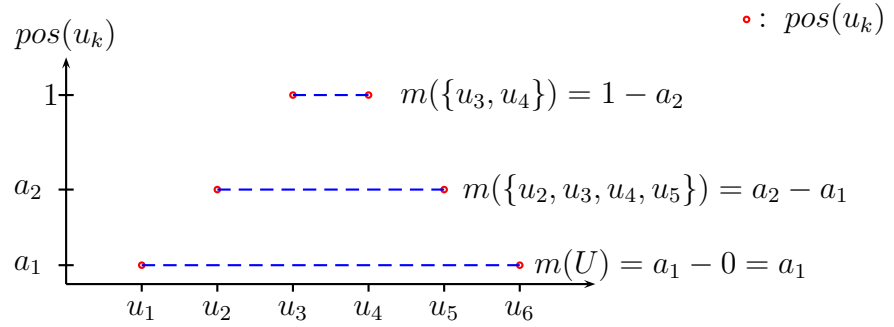


Figure 2.4: A random set generated by possibility distribution.

Uncertainty information may also be possibilities of some subsets of U . For example, let $U = \{u_1, u_2, \dots, u_n\}$, we may have only $Pos(B_1) = b_1, Pos(B_2) = b_2, \dots, Pos(B_K) = b_K$, for some K , as our information. Assume that the

information we have satisfies the possibility properties. (We need to make sure that the information we have satisfies the possibility properties, because if $U = \{1, 2, \dots, 6\}$, and we have $Pos(\{2, 4, 5\}) = \frac{1}{2}$, and $Pos(\{1, 3, 6\}) = \frac{2}{3}$, then $1 = Pos(U) = Pos(\{2, 4, 5\} \cup \{1, 3, 6\}) = \max[Pos(\{2, 4, 5\}), Pos(\{1, 3, 6\})] = \frac{2}{3}$, is wrong.) To create the largest $\mathcal{M}_{\mathcal{F}}$ based on this partial information, we assume that $pos(u_i) = b_k, \forall u_i \in B_k, k = 1, 2, \dots, K$, and $pos(u_i) = 1, \forall u_i \notin \bigcup_{k=1}^K B_k$. Then, we continue the processes (2.58) and (2.59) to obtain the random set that generates the set $\mathcal{M}_{\mathcal{F}}$.

2.2.4 The random set generated by a necessity distribution or necessity of some subset of U

Suppose the information of an uncertainty \hat{u} is given by a necessity distribution. Let $nec : U \rightarrow [0, 1]$ be a necessity distribution of an uncertainty \hat{u} with all realizations in U . Since necessity measure is a special case of belief measure, $m(\{u_k\}) = nec(u_k), \forall k = 1, 2, \dots, n$. However, the focal elements need to be nested. Hence, we cannot have more than one singleton focal element $\{u_i\}$ for some $i \in \{1, 2, \dots, n\}$, otherwise the focal elements are not nested. We can create more than one random set that satisfies this necessity distribution. However, we use the random set (\mathcal{F}, m) , where $\mathcal{F} = \{\{u_i\}, U\}$, and $m(U) = 1 - nec(u_i)$. This random set provides the largest set $\mathcal{M}_{\mathcal{F}}$. Hence, we will cover all possible expected values of \hat{u} . If the uncertainty information is just necessities of some subsets of U , we can turn them into possibilities of the complement of those sets, using $Nec(A) = 1 - Pos(A^c)$. Assume that this information satisfies the possibility properties, we can find the random set that provides the largest set $\mathcal{M}_{\mathcal{F}}$ by following the method in the previous subsection.

2.2.5 The random set generated by a belief distribution

Suppose the information of an uncertainty \check{u} is given by a belief distribution. Let $bel : U \rightarrow [0, 1]$ be a belief distribution of an uncertainty \check{u} with all realizations in U . Then, $\{u_k\}$ such that $bel(u_k) > 0$ is a focal element of a random set. We also can create more than one random set that satisfies a given belief distribution. However, we use the random set (\mathcal{F}, m) , where $\mathcal{F} = \{U, \{u_k\} \text{ such that } bel(u_k) > 0, k = 1, 2, \dots, n\}$, and $m(U) = 1 - \sum_{k=1}^n Bel(\{u_k\})$, by similar reasoning to that provided in the last subsection.

2.2.6 The random set generated by a plausibility distribution

Suppose the information of an uncertainty \hat{u} is given by a plausibility distribution. Let $pl : U \rightarrow [0, 1]$ be a plausibility distribution of an uncertainty \hat{u} with all realizations in U . We want to find the random set generated by a given plausibility distribution, which provides the widest bound on the unknown expected value. This can be done by setting $Bel(\{u_k\}) = 0, \forall k$. There are more than one random set that provides $Pr(\{u_k\}) \in [0, pl(u_k)], \forall k = 1, 2, \dots, n$, by Lemma 2.23 (presented later in the next section). Among these random sets, we can choose one that provides the smallest and the largest expected values, based on the information $Pr(\{u_k\}) \in [0, pl(u_k)], \forall k = 1, 2, \dots, n$. However, it is easier to find this biggest bound on expected values by solving two LP problems.

Subsections 2.2.2 - 2.2.6 suggest that the smallest and the largest expected values of \hat{u} are obtained by finding the two appropriate density mass functions through the random set that provides the biggest set $\mathcal{M}_{\mathcal{F}}$, based on the partial information we have. However, if the partial information is beliefs or plausibili-

ties of some subset of U , then finding the random set that provides the biggest set $\mathcal{M}_{\mathcal{F}}$, based on this partial information may be more troublesome. Instead, we may set up two LP problems to find the bounds on the expected value. The partial information in the form of beliefs or plausibilities of some subsets of U falls under the scope of an IVPM, so we will address these two in detail after the section of IVPM.

Belief and plausibility measures are considered to be the tightest bounds on a set of probability measures, $\mathcal{M}_{\mathcal{F}}$, generated by a finite random set, as in the conclusion of Theorem 2.15. However, an imprecise probability may not be presented as random set information. It could be in the form of probability intervals, e.g., see Example 10.

Example 10. Let $U = \{a, b, c\}$. Suppose someone provides the information that $Pr(\{a\}) \in [\frac{1}{6}, \frac{5}{6}]$, $Pr(\{b\}) \in [\frac{1}{3}, \frac{3}{4}]$, and $Pr(\{c\}) \in [\frac{1}{4}, \frac{2}{3}]$. This information is not considered to be a random set, and we can see that $\frac{5}{6}$ is not the tightest upper bound of the probability of $(\{a\})$, since $\frac{5}{6} + \frac{1}{3} + \frac{1}{4} > 1$. \diamond

There are some mathematical concepts that try to capture an uncertainty form of intervals of all possible probabilities, in general. Two concepts presented in this thesis are interval-valued probability measures (IVPMs) and clouds. We will first provide some basic details of these two concepts. Then we will explain the reason for using a random set and an IVPM as the unified approach for optimization with uncertainty related to this thesis.

2.3 Interval-valued probability measures (IVPMs)

Let $\text{Int}_{[0,1]}$ denote the set of all intervals contained in $[0, 1]$, i.e., $\text{Int}_{[0,1]} \equiv \{[a, b] \mid 0 \leq a \leq b \leq 1\}$. Definition 2.19 of an interval-valued probability measure

(IVPM) was first given by Weichselberger [75], who used the term R-probability as an abbreviation of ‘reasonable’ probability, since it interprets a reasonable range that an unknown probability could be.

Definition 2.19 (see Weichselberger [75]) *Given a measurable space (U, σ_U) , an interval-valued function $i : \sigma_U \rightarrow \text{Int}_{[0,1]}$ is called an **interval-valued probability measure (IVPM)** (or an R-probability) if*

- $i(A) = [i^-(A), i^+(A)] \subseteq [0, 1]$, where $A \in \sigma_U$ and $i^-(A) \leq i^+(A)$, and
- There exists a probability measure Pr such that

$$\forall A \in \sigma_U, Pr(A) \in i(A).$$

Definition 2.20 (see Weichselberger [75]) *Given a measurable space (U, σ_U) , and an IVPM i on σ_U , the set $\mathcal{M} = \{Pr : Pr(A) \in i(A), \forall A \in \sigma_U\}$ is called the **structure** of i .*

Definition 2.21 (see Weichselberger [75]) *An R-probability is called an **F-probability**, if $\forall A \in \sigma_U$*

- $i^+(A) = \sup \{Pr(A) : Pr \in \mathcal{M}\}$, and
- $i^-(A) = \inf \{Pr(A) : Pr \in \mathcal{M}\}$.

F-probability is an abbreviation of ‘feasible’ probability. In any F-probability, none of $i^+(A)$ and $i^-(A)$ are too wide, while this may be the case for an IVPM (being too wide).

Theorem 2.15 shows that a finite random set is a special example of an IVPM by using $i(A) = [Bel(A), Pl(A)]$, $\forall A \subseteq U$. Moreover, since $\mathcal{M}_{\mathcal{F}}$ is the structure \mathcal{M} , $Bel(A) = \inf \{Pr(A) : Pr \in \mathcal{M}\}$, and $Pl(A) = \sup \{Pr(A) : Pr \in \mathcal{M}\}$, a

random set generates an F-probability. We may be able to find a (not unique) random set, when the only information we have is in the form of an R or F-probability of all singleton subsets of U (instead of any subsets of U). For example, if we have information that the left and right bounds of intervals $[\frac{1}{6}, \frac{5}{12}]$, $[\frac{1}{3}, \frac{7}{12}]$, and $[\frac{1}{4}, \frac{1}{2}]$ create an F-probability of $\{a\}$, $\{b\}$, and $\{c\}$, respectively, when $U = \{a, b, c\}$, then these bounds represent beliefs and plausibilities of a, b , and c . We can construct a random set by using these bounds, e.g., $\frac{1}{6} = Bel(\{a\}) = \sum_{B \cap \{a\} \neq \emptyset} m(B) = m(\{a\})$. Similarly, $m(\{b\}) = \frac{1}{3}$, and $m(\{c\}) = \frac{1}{4}$. Together with $m(U) = 1 - \frac{1}{6} - \frac{1}{3} - \frac{1}{4} = \frac{1}{4}$, the random set (\mathcal{F}, m) , where $\mathcal{F} = \{\{a\}, \{b\}, \{c\}, U\}$, preserve the plausibilities of a, b , and c as the right bound of the intervals of this F-probability, as it should be. An R-probability of all singleton subsets of U also is called a probability interval.

Definition 2.22 (see De Campos, et al. [8]) *Let $U = \{u_1, u_2, \dots, u_n\}$ be a finite set of realizations of an uncertainty \hat{u} , and $\mathcal{L} = \{[\underline{a}_k, \bar{a}_k], \text{ where } 0 \leq \underline{a}_k \leq \bar{a}_k \leq 1, \forall k = 1, 2, \dots, n\}$. We can interpret these intervals in \mathcal{L} as a set of bounds of probability by defining the set \mathcal{P} as $\mathcal{P} = \{Pr \mid \underline{a}_k \leq Pr(\{u_k\}) \leq \bar{a}_k, \forall k = 1, 2, \dots, n\}$. We will say that \mathcal{L} is a set of probability intervals (**probability interval**, in short), and \mathcal{P} is the set of possible probabilities associated with \mathcal{L} .*

Assuming that $\mathcal{P} \neq \emptyset$, a probability interval interpretation of an uncertainty \hat{u} also can be an example of an IVPM by setting

$$i(A) = \begin{cases} [\underline{a}_k, \bar{a}_k]; & A = \{u_k\}, k = 1, 2, \dots, n \\ [0, 1]; & \text{otherwise.} \end{cases} \quad (2.60)$$

Therefore, there is a random generating the F-probability of this IVP. M.

De Campos et al., [8], showed how to generate an F-probability of all singleton subsets of U from a given IVP of all singleton subsets of U , (also called a probability interval) , when U is finite and $|U| = n$. The set of probabilities associated to this F-probability is equal to the structure

$$\mathcal{M} = \left\{ Pr \mid Pr(u_k) \in [i^-(u_k), i^+(u_k)], \sum_{k=1}^n Pr(u_k) = 1 \right\}. \quad (2.61)$$

We will present in Lemma 2.23 how to obtain an F-probability of all $\{u_k\} \subseteq U$ from a given IVP of $\{u_k\}$; $k = 1, 2, \dots, n$, in a similar way as in [8], and use it to generate a random set. Lemmer, et al. [31] provides the conditions when the bounds on a given probability interval of $\{u_k\}$ represent beliefs and plausibilities of $\{u_k\}$, $\forall k = 1, 2, \dots, n$. However, our contribution results from Lemma 2.23, and it is that we can construct a given probability interval of $\{u_k\}$, $\forall k$ as $[Bel(\{u_k\}), Pl(\{u_k\})]$, which is the tightest interval in the sense that $\mathcal{M} = \overline{\mathcal{M}}$, where $\overline{\mathcal{M}} = \left\{ Pr \mid Pr(\{u_k\}) \in [Bel(\{u_k\}), pl(\{u_k\})], \sum_{k=1}^n Pr(u_k) = 1 \right\}$.

Lemma 2.23 *Let a finite set $U = \{u_1, u_2, \dots, u_n\}$, and a corresponding R-probability of $\{u_k\}$, $k = 1, 2, \dots, n$ be given. We can construct a random set (\mathcal{F}, m) (may not be unique) from a probability interval, (the R-probability of $\{u_1\}, \{u_2\}, \dots, \{u_n\}$), such that the F-probability of $\{u_k\}$ is*

$$[Bel(\{u_k\}), Pl(\{u_k\})], \forall k = 1, 2, \dots, n.$$

Proof: Consider $[i^-(\{u_k\}), i^+(\{u_k\})]$ of an given probability interval $\forall k = 1, \dots, n$. We try to get the tightest possible subintervals of $[i^-(\{u_k\}), i^+(\{u_k\})]$ that represent the corresponding structure \mathcal{M} as in (2.61). This tightest subin-

terval is the F-probability of $\{u_k\}$, $\forall k$. Since there is a random set generating the F-probability of this probability interval, $i^-(\{u_k\}) \leq Bel(\{u_k\})$ and $i^+(\{u_k\}) \geq Pl(\{u_k\})$, $\forall k = 1, 2, \dots, n$. We note that $\mathcal{M} \neq \emptyset$, by the definition of an IVP. [8] provides conditions to check the emptiness of a probability interval.

Some of the endpoints of these intervals $[i^-(\{u_k\}), i^+(\{u_k\})]$, $k = 1, \dots, n$, can be probabilities $Pr(\{u_k\})$. We categorize them into six cases.

Case 1. $i^-(\{u_k\}) + \sum_{\substack{j=1 \\ j \neq k}}^n i^+(\{u_j\}) < 1$. We cannot increase the terms $i^+(\{u_j\})$ anymore to change ' $<$ ' to ' $=$ ', so we have to increase $i^-(\{u_k\})$. Hence, $i^-(\{u_k\}) < Bel(\{u_k\})$, and $i^+(\{u_j\}) = Pl(\{u_j\})$. We get $Bel(\{u_k\}) = 1 - \sum_{\substack{j=1 \\ j \neq k}}^n i^+(\{u_j\})$.

Case 2. $i^-(\{u_k\}) + \sum_{\substack{j=1 \\ j \neq k}}^n i^+(\{u_j\}) > 1$. We can reduce some terms of $i^+(\{u_j\})$, so that ' $>$ ' becomes ' $=$ '. This implies $i^-(\{u_k\}) = Bel(\{u_k\})$.

Case 3. $i^-(\{u_k\}) + \sum_{\substack{j=1 \\ j \neq k}}^n i^+(\{u_j\}) = 1$. Then, $i^-(\{u_k\}) = Bel(\{u_k\})$.

Therefore, we obtain $Bel(\{u_k\})$, $\forall k = 1, 2, \dots, n$. Similar to these three cases above, the later three cases provide $Pl(\{u_k\})$, $\forall k = 1, 2, \dots, n$.

Case 4. $i^+(\{u_k\}) + \sum_{\substack{j=1 \\ j \neq k}}^n i^-(\{u_j\}) < 1$. We can change some of the terms $i^-(\{u_j\})$, so that ' $<$ ' becomes ' $=$ '. Therefore, $i^+(\{u_k\}) = Pl(\{u_k\})$.

Case 5. $i^+(\{u_k\}) + \sum_{\substack{j=1 \\ j \neq k}}^n i^-(\{u_j\}) > 1$. Since we cannot reduce any of the terms $i^-(\{u_j\})$ further, the term $i^+(\{u_k\})$ has to be lower. Hence, $i^+(\{u_k\}) >$

$Pl(\{u_k\})$. We can get $Pl(\{u_k\})$ by $Pl(\{u_k\}) = 1 - \sum_{\substack{j=1 \\ j \neq k}}^n i^-(\{u_j\})$.

Case 6. $i^+(\{u_k\}) + \sum_{\substack{j=1 \\ j \neq k}}^n i^-(\{u_j\}) = 1$. Then, $i^+(\{u_k\}) = Pl(\{u_k\})$.

Only one of the first three cases can happen for each u_k , while finding $Bel(\{u_k\})$. Similarly, only one of Case 4, 5, or 6 occurs for each u_k , in the process of finding $Pl(\{u_k\})$. These cases of finding $Bel(\{u_k\})$ and $Pl(\{u_k\})$ may seem to interact to each other; for example, the updated $Bel(\{u_k\})$ from Case 1 may results the updated $Pl(\{u_j\})$ from Case 5. However, the following analysis shows that the updated $Bel(\{u_k\})$ does not effect the updated $Pl(\{u_j\})$.

- If Case 1 happens for some $k = 1, 2, \dots, n$.

Then $i^+(\{u_j\}) = Pl(\{u_j\})$, $\forall j \neq k$, and $Bel(\{u_k\}) = 1 - \sum_{\substack{j=1 \\ j \neq k}}^n i^+(\{u_j\})$. Thus, for each $j \neq k$,

$$i^+(\{u_j\}) = 1 - \sum_{\substack{l=1 \\ l \neq j, k}}^n i^+(\{u_l\}) - Bel(\{u_k\}) \quad (2.62)$$

$$\leq 1 - \sum_{\substack{l=1 \\ l \neq j, k}}^n i^-(\{u_l\}) - Bel(\{u_k\}). \quad (2.63)$$

$Bel(\{u_k\})$ is the new left bound of the interval $[i^-(\{u_k\}), i^+(\{u_k\})]$. Thus (2.63) implies

$$i^+(\{u_j\}) + \sum_{\substack{l=1 \\ l \neq j}}^n i^-(\{u_l\}) \leq 1.$$

Therefore, Case 5 will never happen, when $j \neq k$. Hence, $Pl(\{u_j\})$ is not changed and equal to $i^+(\{u_j\})$, $\forall j \neq k$. If Case 1 happens for more than

one k , says k_1 and k_2 , we can set $j := k_1$ and $k := k_2$ (and vice versa) in (2.62 - 2.63). Hence, Case 5 will never happen for any $k = 1, 2, \dots, n$, and $Pl(\{u_k\}) = i^+(\{u_k\})$. However, if Case 1 happens for only one k , then Case 5 might happen for u_k .

1. If Case 5 happens for u_k , then $Pl(\{u_k\}) = 1 - \sum_{\substack{j=1 \\ j \neq k}}^n i^-(\{u_j\})$. Moreover, for

$$\text{each } j \neq k, \quad i^-(\{u_j\}) = 1 - \sum_{\substack{l=1 \\ l \neq j, k}}^n i^-(\{u_l\}) - Pl(\{u_k\}) \quad (2.64)$$

$$\geq 1 - \sum_{\substack{l=1 \\ l \neq j, k}}^n i^+(\{u_l\}) - Pl(\{u_k\}), \quad (2.65)$$

which falls into Case 2 ,or 3. These two cases imply that $Bel(\{u_j\}) = i^-(\{u_j\})$, $\forall j \neq k$. Therefore, we obtain the tightest subinterval $[Bel(\{u_k\}), Pl(\{u_k\})]$, $\forall k = 1, 2, \dots, n$.

2. If Case 5 does not happen for u_k , then $Pl(\{u_k\}) = i^+(\{u_k\})$. We still do not know $Bel(\{u_j\})$, $j \neq k$. The case $i^-(\{u_j\}) + \sum_{\substack{l=1 \\ l \neq j}}^n i^+(\{u_l\}) < 1$, $\exists j \neq k$ will never happen, since we are considering when Case 1 happens just for k . Hence, $Bel(\{u_j\}) = i^-(\{u_j\})$, $\forall j \neq k$. We also obtain the tightest subinterval $[Bel(\{u_k\}), Pl(\{u_k\})]$, $\forall k = 1, 2, \dots, n$.

- If Case 1 has never happened for u_k , $\forall k = 1, 2, \dots, n$.

Then $Bel(\{u_k\}) = i^-(\{u_k\})$, $\forall k$. If Case 5 also has never happened, then $Pl(\{u_k\}) = i^+(\{u_k\})$, $\forall k$. If Case 5 happens for some $k = 1, 2, \dots, n$, then $Pl(\{u_k\}) = 1 - \sum_{\substack{j=1 \\ j \neq k}}^n i^-(\{u_j\})$ and $Pl(\{u_j\}) = i^+(\{u_j\})$, $\forall j \neq k$. Moreover (2.64) and (2.65) are true for these k . Suppose that there is more than one

k satisfies (2.64) and (2.65), says k_1 and k_2 . Setting $j := k_1$ and $k := k_2$ (and vice versa), we can see that Case 1 will never re-happen, and it results the unchanges $Bel(\{u_k\}) = i^-(\{u_k\}), \forall k$. If Case 5 happens for only one k , then Case 1 will never re-happen for $u_j, \forall j \neq k$ because of (2.64) and (2.65). We also get the unchanged $Pl(\{u_j\}) = i^+(\{u_j\}), \forall j \neq k$, which means that $i^-(\{u_k\}) + \sum_{\substack{j=1 \\ j \neq k}}^n i^+(\{u_j\})$ remains unchanged and hence, it does not fall into Case 1. Therefore, Case 1 will never re-occur. We obtain the tightest subinterval $[Bel(\{u_k\}), Pl(\{u_k\})], \forall k = 1, 2, \dots, n$. We conclude from this analysis that the updated $Bel(\{u_k\})$ does not effect the updated $Pl(\{u_j\})$.

Let $\overline{\mathcal{M}} = \{Pr \mid Pr(\{u_k\}) \in [Bel(\{u_k\}), Pl(\{u_k\})], \forall k = 1, 2, \dots, n\}$. It is easy to see that $\overline{\mathcal{M}} \subseteq \mathcal{M}$. On the other hand, if $Pr \in \mathcal{M}$, then because of the restriction $\sum_{k=1}^n Pr(\{u_k\}) = 1$, Cases 1 - 6 imply that $Pr \in \overline{\mathcal{M}}$. Thus, $\overline{\mathcal{M}} = \mathcal{M}$. Next, we construct a random set (\mathcal{F}, m) from $[Bel(\{u_k\}), Pl(\{u_k\})], k = 1, 2, \dots, n$, by first setting $\mathcal{F} = \emptyset$, then adding $\{u_k\}$ to \mathcal{F} , when $m(\{u_k\}) = Bel(\{u_k\}) > 0, \forall k = 1, 2, \dots, n$. So,

$$1 - \sum_{k=1}^n m(\{u_k\}) = \sum_{\substack{B \subseteq U \\ B \neq \{u_k\}, k=1, \dots, n}} m(B). \quad (2.66)$$

Since, for each $k = 1, 2, \dots, n, Pl(\{u_k\}) = \sum_{B \cap \{u_k\} \neq \emptyset} m(B)$,

$$Pl(\{u_k\}) - Bel(\{u_k\}) = \sum_{\substack{B \cap \{u_k\} \neq \emptyset \\ B \neq \{u_k\}}} m(B), \quad \forall k = 1, 2, \dots, n. \quad (2.67)$$

The system of Equations (2.66) and (2.67) has the total of $n + 1$ rows and $2^n - n - 1$ columns. We check that the right-hand-side matrix of this system has the full row-rank $n + 1$ in Lemma 2.24. Therefore, there are infinitely many

solutions to $2^n - n - 1$ terms of $m(B)$. We can construct one of the solutions of the system as follows.

1. Set $M := m(X) = \min \{Pl(\{u_k\}) - Bel(\{u_k\}), k = 1, 2, \dots, n\}$. If $M \geq 1 - \sum_{k=1}^n m(\{u_k\})$, set $m(U) = 1 - \sum_{k=1}^n m(\{u_k\})$, and $m(B) = 0, \forall B \subseteq U, B \neq U, B \neq \{u_k\}, k = 1, 2, \dots, n$, then we are done. Otherwise, let $\Lambda = \{1, 2, \dots, n\}$ and $\Lambda_1 = \{k \mid M = Pl(\{u_k\}) - Bel(\{u_k\})\}$. We have $m(B) = 0; \forall B \subseteq U, B \neq U, B \cap \{u_k\} \neq \emptyset$, and $B \neq \{u_k\}, \forall k \in \Lambda_1$.
2. Now, we consider the system (2.67) when $k \in \Lambda - \Lambda_1$. Choose an unknown value $m(B_1)$ for some $B_1 \subseteq U$, such that $m(B_1)$ is a term in $\sum_{\substack{B \cap \{u_k\} \neq \emptyset \\ B \neq \{u_k\}}} m(B)$, for each $k \in \Lambda - \Lambda_1$. Such a set B_1 always exists. Set $m(B_1) = \min \{Pl(\{u_k\}) - Bel(\{u_k\}) - M, \forall k \in \Lambda - \Lambda_1\}$. If $m(B_1) \geq 1 - \sum_{k=1}^n m(\{u_k\}) - M$, set $m(B_1) = 1 - \sum_{k=1}^n m(\{u_k\}) - M$, and $m(B) = 0$ for the rest of unassigned value $m(B)$'s, then we are done. Otherwise, set $M := m(B_1)$, and $\Lambda_2 := \{k \mid M = Pl(\{u_k\}) - Bel(\{u_k\})\}$. Again, we have $m(B) = 0$ for every unassigned value $m(B)$ in $\sum_{\substack{B \cap \{u_k\} \neq \emptyset \\ B \neq \{u_k\}}} m(B), k \in \Lambda_2$.
3. Set $\Lambda_1 := \Lambda_1 \cup \Lambda_2$, return to step 2.

We add set $B \subseteq U$ into \mathcal{F} , whenever $m(B) > 0$. Hence, we obtain a random set (\mathcal{F}, m) . The procedure of constructing this random set will terminate in a finite number of iterations, since U is finite. It is indeed a random set, since $\sum_{A \in \mathcal{F}} m(A) = 1$. \square

Lemma 2.24 *The system of Equations (2.66 - 2.67) has the full row-rank.*

Proof: It is trivial, when $|U| = 1$ or 2 . Consider when $n = 3$, so $X = \{u_1, u_2, u_3\}$.

The left-hand-side of Equations (2.66 - 2.67) are nonnegative numbers between $[0, 1]$. The right-hand-side can be expanded as:

$$\begin{array}{l} m(\{u_1, u_2\}) + m(\{u_1, u_3\}) + m(\{u_2, u_3\}) + m(U) \\ m(\{u_1, u_2\}) + m(\{u_1, u_3\}) \qquad \qquad \qquad + m(U) \\ m(\{u_1, u_2\}) \qquad \qquad \qquad + m(\{u_2, u_3\}) + m(U) \\ \qquad \qquad \qquad m(\{u_1, u_3\}) + m(\{u_2, u_3\}) + m(U) \end{array} \Rightarrow \begin{array}{c} \left[\begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{array} \right] := A^{(3)}. \end{array} \quad (2.68)$$

The terms $m(B)$ in (2.68) are considered to be unknown variables. The matrix $A^{(3)}$ in (2.68) represents the coefficient matrix of these unknown variables. It is not difficult to transform the matrix $A^{(3)}$ to the 4×4 identity matrix, I_4 . Thus, the right-hand-side coefficient matrix of Equations (2.66 - 2.67), when $n = 3$, has the full row-rank.

When $n = 4$, $U = \{u_1, u_2, u_3, u_4\}$. A part of the right-hand-side of Equations (2.66 - 2.67) is

$$\begin{array}{l} m(\{u_1, u_2, u_3\}) + m(\{u_1, u_2, u_4\}) + m(\{u_1, u_3, u_4\}) + m(\{u_2, u_3, u_4\}) + m(U) \\ m(\{u_1, u_2, u_3\}) + m(\{u_1, u_2, u_4\}) + m(\{u_1, u_3, u_4\}) \qquad \qquad \qquad + m(U) \\ m(\{u_1, u_2, u_3\}) + m(\{u_1, u_2, u_4\}) \qquad \qquad \qquad + m(\{u_2, u_3, u_4\}) + m(U) \\ m(\{u_1, u_2, u_3\}) \qquad \qquad \qquad + m(\{u_1, u_3, u_4\}) + m(\{u_2, u_3, u_4\}) + m(U) \\ \qquad \qquad \qquad m(\{u_1, u_2, u_4\}) + m(\{u_1, u_3, u_4\}) + m(\{u_2, u_3, u_4\}) + m(U), \end{array}$$

whose coefficient matrix can be written as

$$\left[\begin{array}{c|cccc} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ \hline 0 & 1 & 1 & 1 & 1 \end{array} \right] = \left[\begin{array}{c|cccc} 1 & & & & \\ 1 & & & & \\ 1 & & & & \\ 1 & & & & \\ \hline 0 & 1 & 1 & 1 & 1 \end{array} \right] := A^{(4)}.$$

Again, we can transform $A^{(4)}$ to I_5 by using the Gauss - Jordan elimination method. Hence, the right-hand-side coefficient matrix of Equations (2.66 - 2.67), when $n = 4$, has the full row-rank.

In general, for $|X| = k = 4, 5, \dots, n - 1$, we have $U = \{u_1, u_2, \dots, u_k\}$. A part of the right-hand-side of Equations (2.66 - 2.67) can be written as the following square matrix of dimension $(k + 1) \times (k + 1)$, which associates with k of unknown terms $m(B)$, where $|B| = k - 1$, and $m(U)$.

$$\left[\begin{array}{c|c} 1 & \\ 1 & A^{(k-1)} \\ \vdots & \\ 1 & \\ \hline 0 & 1 \ 1 \ \dots \ 1 \end{array} \right] := A^{(k)}, \quad \forall k = 1, 2, \dots, n - 1.$$

By mathematical induction, suppose that $A^{(k)}$ can be transformed to the identity matrix, $I_{(k+1)}$, for each $k = 1, 2, \dots, n - 1$. Thus, for $k = n$, the matrix

$$\left[\begin{array}{c|c} 1 & \\ 1 & A^{(n-1)} \\ \vdots & \\ 1 & \\ \hline 0 & 1 \ 1 \ \dots \ 1 \end{array} \right] := A^{(n)}$$

can be transformed to $I_{(n+1)}$ by Gauss - Jordan elimination. Hence, the right-hand-side coefficient matrix of Equations (2.66 - 2.67) has the full row-rank, when $|U| = 3, 4, \dots, n$. \square

Lemma 2.23 helps transform probability interval information to (not unique) random set information. However, we may not get two probability density

mass functions from this random set to obtain the largest bound on the expected value of \hat{u} . To guarantee the smallest and the largest expected values of \hat{u} given probability intervals information, we solve the LP problems: $\min_f / \max_f E_f(\hat{u})$ s.t. $f(u_i) \in [\underline{a}_i, \bar{a}_i]$, $\sum_{i=1}^n f(u_i) = 1$, and $f(u_i) \geq 0, \forall i$.

Let $\mathcal{A} = \{A_1, A_2, \dots, A_i\} \subseteq \mathcal{P}(U)$, for some $i \in \mathbb{N}$, $i \leq 2^{|U|} - 1$. Suppose we know the values of probabilities on these sets in \mathcal{A} , $Pr(A_k) = p_k$, $k = 1, 2, \dots, i$, but the information is not enough to provide the values of $Pr(\{u_k\}), \forall k$. Then, we may not be able to find any random set to represent probability on sets information, because $\mathcal{M}_{\mathcal{A}} \neq \mathcal{M}$, in general, where $\mathcal{M}_{\mathcal{A}} = \{Pr : Pr(A_k) = p_k, \forall A_k \in \mathcal{A}\}$, and \mathcal{M} is the set of probability generated by a random set. Therefore, a probability on sets interpretation of an uncertainty \hat{u} is not a special case of a random set interpretation. It can be an example of an IVPM by setting

$$i(A) = \begin{cases} [Pr(A), Pr(A)]; & A \in \mathcal{A}, \\ [0, 1]; & A \in \mathcal{P}(U) \setminus \mathcal{A}. \end{cases} \quad (2.69)$$

To guarantee the smallest and the largest expected values of \hat{u} given a probability on sets information, we solve the LP problems:

$$\min_f / \max_f E_f(\hat{u}) \text{ s.t. } \sum_{u_i \in A_k} f(u_i) = p_k, \forall A_k \in \mathcal{A}, \text{ and } \sum_{i=1}^n f(u_i) = 1.$$

Let $Bel(A_k) = \alpha_k$, and $Pl(A_k) = \beta_k, \forall A_k \in \mathcal{A}$. Suppose we have partial information in the form of belief (or plausibility) of some subsets of U , then for an unknown probability Pr that obeys this partial information, $\alpha_k = Bel(A_k) \leq \sum_{u_i \in A_k} Pr(u_i), \forall A_k \in \mathcal{A}$, (or $\beta_k = Pl(A_k) \geq \sum_{u_i \in A_k} Pr(u_i), \forall A_k \in \mathcal{A}$). We

can write belief of some subsets of U as the following IVPM:

$$i(A) = \begin{cases} [Bel(A), 1]; & A \in \mathcal{A}, \\ [0, 1]; & A \in \mathcal{P}(U) \setminus \mathcal{A}. \end{cases} \quad (2.70)$$

To guarantee the smallest and the largest expected values of \hat{u} given a probability on sets information, we solve the LP problems:

$$\min_f / \max_f E_f(\hat{u}) \text{ s.t. } \sum_{u_i \in A_k} f(u_i) \geq \alpha_k, \forall A_k \in \mathcal{A}, \text{ and } \sum_{i=1}^n f(u_i) = 1.$$

A similar approach applies to a partial information of plausibility on subsets of the set U .

2.3.1 Upper and lower expected values generated by IVPMs

We consider the case when the incomplete information is interpreted as an IVPM, i.e., suppose that a finite set U of n realizations of \hat{u} is defined by $U := \{u_1, u_2, \dots, u_n\}$, where $u_1 \leq u_2 \leq \dots \leq u_n$, and have the probability information in a nonempty set \mathcal{P} as follows:

$$\mathcal{P} = \left\{ (Pr_1, Pr_2, \dots, Pr_n) \left| \begin{array}{l} \underline{a}_i \leq \sum_{k \in \Lambda_i} Pr_k \leq \bar{a}_i, \quad \forall i = 1, 2, \dots, t, \\ \sum_{k=1}^n Pr_k = 1, \quad Pr_k \geq 0, \quad \forall k = 1, 2, \dots, n. \end{array} \right. \right\} \quad (2.71)$$

Λ_i in (2.71) is an index set which is a subset of $\{1, 2, \dots, n\}$, where t is the total number of index sets, i.e., $t \in \mathbb{N}$ and $t \leq 2^n - 2$. The unknown probability of $\{u_k\}$ is Pr_k . The sum of these probabilities in Λ_i is $\lambda_i \in [0, 1]$. Let Pr be any probability satisfying (2.71), and $E_{Pr}(\hat{u})$ denote the expected values of \hat{u} , with respect to Pr . We show in Theorem 2.25 and Corollary 2.26 below that there are density mass functions Pr_* and Pr^* such that the expected value of \hat{u} with respect to Pr_* and Pr^* are the lower and the upper bounds on all expected values of \hat{u} with respect to any probability satisfies (2.71), i.e., $E_{Pr_*}(\hat{u}) \leq E_{Pr}(\hat{u}) \leq E_{Pr^*}(\hat{u})$.

Theorem 2.25 *Suppose that the system (2.71) is nonempty. Then, there is a fixed density mass $Pr_* = (Pr_{1*}, Pr_{2*}, \dots, Pr_{n*})$, which is an optimal solution for the problem $P_U := \min_{Pr \in \mathcal{P}} u_1 Pr_1 + u_2 Pr_2 + \dots + u_n Pr_n$, for any uncertainty \hat{u} that has n realizations u_1, u_2, \dots, u_n such that $-\infty < u_1 \leq u_2 \leq \dots \leq u_n < \infty$.*

Proof: An optimal solution for the problem P_U exists because the set \mathcal{P} is a closed and bounded nonempty polyhedral set. We construct a probability Pr_* using a greedy algorithm as follows.

step 1. Choose all Pr 's in \mathcal{P} that have the highest value possible of Pr_1 . Define the set of all these Pr 's as \mathcal{P}_1

step 2. From \mathcal{P}_1 , choose all Pr 's that have the highest value possible of Pr_2 . Define the set of all these Pr 's as \mathcal{P}_2 .

step 3. From \mathcal{P}_2 , choose all Pr 's that have the highest value possible of Pr_3 . Define the set of all these Pr 's as \mathcal{P}_3 .

⋮

step i . From \mathcal{P}_{i-1} , choose all Pr 's that have the highest value possible of Pr_i . Define the set of all these Pr 's as \mathcal{P}_i .

⋮

step n . We will get one value of Pr_n , by following these steps.

Hence, we obtain one probability measure from this construction. Let us define this probability as Pr_* . We suppose that this Pr_* is not an optimal solution for P_U , (and try to find a contradiction). Therefore, each component i has changed from Pr_{i*} to $Pr_{i*} - \delta_i + \Delta_i$, where $\delta_i, \Delta_i \geq 0, \forall i$. Suppose that this new changed

probability is optimal for P_U . Let Z be the minimal objective value of P_U . This requires the analysis of n cases.

Case 1. Suppose that Pr_{1^*} changes to $Pr_{1^*} - \delta_1 + \Delta_1$.

We know that some other components i , $i \neq 1$, also need to be changed, so that this changed probability still satisfies the conditions in \mathcal{P} . The 1st component can only be reduced, i.e., $Pr_{1^*} - \delta_1$, for some $\delta_1 > 0$, since Pr_{1^*} is the highest value by the construction above. Thus,

$$Z = u_1 (Pr_{1^*} - \delta_1) + u_2 (Pr_{2^*} - \delta_2 + \Delta_2) + u_3 (Pr_{3^*} - \delta_3 + \Delta_3) + \dots + u_n (Pr_{n^*} - \delta_n + \Delta_n).$$

As we know that $Pr_{1^*} - \delta_1$ can be increased, let us increase it by λ_1 , where $0 < \lambda_1 \leq \delta_1$, and to be able to satisfy \mathcal{P} , $Pr_{i^*} - \delta_i + \Delta_i$ may be reduced by $\lambda_i \geq 0$, such that $\sum_{i=2}^n \lambda_i = \lambda_1$. Hence,

$$\begin{aligned} Z_1 &= u_1 (Pr_{1^*} - \delta_1 + \lambda_1) + u_2 (Pr_{2^*} - \delta_2 + \Delta_2 - \lambda_2) + u_3 (Pr_{3^*} - \delta_3 + \Delta_3 - \lambda_3) \\ &\quad + \dots + u_n (Pr_{n^*} - \delta_n + \Delta_n - \lambda_n) \\ &= u_1 (Pr_{1^*} - \delta_1) + u_2 (Pr_{2^*} - \delta_2 + \Delta_2) + u_3 (Pr_{3^*} - \delta_3 + \Delta_3) + \dots + \\ &\quad u_n (Pr_{n^*} - \delta_n + \Delta_n) + (u_1 - u_2)\lambda_2 + (u_1 - u_3)\lambda_3 + \dots + (u_1 - u_n)\lambda_n \\ &= Z + (u_1 - u_2)\lambda_2 + (u_1 - u_3)\lambda_3 + \dots + (u_1 - u_n)\lambda_n. \end{aligned}$$

Now, consider $\Theta = (u_1 - u_2)\lambda_2 + (u_1 - u_3)\lambda_3 + \dots + (u_1 - u_n)\lambda_n$. If $u_1 < u_2 < \dots < u_n$, then $\Theta < 0$. Hence, $Z_1 < Z$. If $u_i = u_{i+1} = u_{i+2} = \dots = u_{i+j} = u$, for some $i \geq 1$ and $i + j \leq n$, then $\Theta = (u_1 - u_2)\lambda_2 + (u_1 - u_3)\lambda_3 + \dots + (u_1 - u_{i-1})\lambda_{i-1} + (u_1 - (j - i + 1)u)(\lambda_i + \lambda_{i+1} + \dots + \lambda_j) + (u_i - u_{j+1})\lambda_{j+1} + \dots + (u_1 - u_n)\lambda_n < 0$. Hence, $Z_1 < Z$.

Case 2. Suppose that Pr_{2^*} changes to $Pr_{2^*} - \delta_2 + \Delta_2$, where $\Delta_2 - \delta_2 \neq 0$.

We know that some other components i , $i \neq 2$, also need to be changed, so that this changed probability still satisfies the conditions in \mathcal{P} . The 1st component can only be reduced, i.e., $Pr_{1^*} - \delta_1$, for some $\delta_1 > 0$, since Pr_{1^*} is the highest value by the construction above. If Pr_{1^*} is reduced, then it becomes the previous case. If there are no changes in Pr_{1^*} , then this changed probability is in \mathcal{P}_1 . Pr_{2^*} can only be decreased, otherwise this changed probability is not in \mathcal{P}_1 . So, Pr_{2^*} changes to $Pr_{2^*} - \delta_2$, for some $\delta_2 > 0$. Thus,

$$Z = u_1(Pr_{1^*}) + u_2(Pr_{2^*} - \delta_2) + u_3(Pr_{3^*} - \delta_3 + \Delta_3) + \dots + u_n(Pr_{n^*} - \delta_n + \Delta_n).$$

By a similar pattern as in **Case 1**, we know that $Pr_{2^*} - \delta_2$ can be increased. Let us increase it by λ_2 , where $0 < \lambda_2 \leq \delta_2$, and to be able to satisfy \mathcal{P} , $Pr_{i^*} - \delta_i + \Delta_i$ may be reduced by $\lambda_i \geq 0$, such that $\sum_{i=3}^n \lambda_i = \lambda_2$. Hence,

$$\begin{aligned} Z_2 &= u_1(Pr_{1^*}) + u_2(Pr_{2^*} - \delta_2 + \lambda_2) + u_3(Pr_{3^*} - \delta_3 + \Delta_3 - \lambda_3) \\ &\quad + \dots + u_n(Pr_{n^*} - \delta_n + \Delta_n - \lambda_n) \\ &= u_1(Pr_{1^*}) + u_2(Pr_{2^*} - \delta_2) + u_3(Pr_{3^*} - \delta_3 + \Delta_3) + \dots + \\ &\quad u_n(Pr_{n^*} - \delta_n + \Delta_n) + (u_2 - u_3)\lambda_3 + (u_2 - u_4)\lambda_4 + \dots + (u_2 - u_n)\lambda_n \\ &= Z + (u_2 - u_3)\lambda_3 + (u_2 - u_4)\lambda_4 + \dots + (u_2 - u_n)\lambda_n. \end{aligned}$$

Hence, $Z_2 < Z$. We apply a similar argument to the rest of the cases. Then, the changed probability is not optimal, a contradiction. Thus, Pr_* is an optimal solution for P_U . \square

Corollary 2.26 *Suppose that the system (2.71) is nonempty. Then, there is a fixed density mass $Pr^* = (Pr_1^*, Pr_2^*, \dots, Pr_n^*)$ as an optimal solution for the problem $P_V := \max_{Pr \in \mathcal{P}} v_1 Pr_1 + v_2 Pr_2 + \dots + v_n Pr_n$, for any uncertainty \hat{v} that has n realizations v_1, v_2, \dots, v_n such that $-\infty < v_1 \leq v_2 \leq \dots \leq v_n < \infty$.*

Proof: Apply Theorem 2.26 by using $v_1 = -u_n$, $v_2 = -u_{n-1}$, \dots , and $v_n = -u_1$. \square

Next, the concept of clouds is presented. We point out some differences between clouds and IVPMs and give an example to shows that a cloud could be represented as an IVPM.

2.4 Clouds

The concept of a cloud, which tries to combine probability and possibility into one conceptual basis, was introduced by Neumaier in [52]. The general definition of a cloud is given below.

Definition 2.27 (see Neumaier, [52]) *Suppose all realizations of an uncertainty parameter are in a set U . A **cloud** over set U is an interval mapping $c : U \rightarrow \text{Int}_{[0,1]}$ such that*

1. $\forall u \in U; c(u) = [\underline{c}(u), \bar{c}(u)]$ with $\underline{c}(u), \bar{c}(u) \in \mathbb{R}$, $0 \leq \underline{c}(u) \leq \bar{c}(u) \leq 1$, and
2. $(0, 1) \subseteq \bigcup_{u \in U} c(u) \subseteq [0, 1]$.

In addition, a random variable Y taking values in U is said to belong to cloud c , (written as $Y \in c$), if and only if

$$\forall \alpha \in [0, 1]; Pr(\underline{c}(Y) \geq \alpha) \leq 1 - \alpha \leq Pr(\bar{c}(Y) > \alpha). \quad (2.72)$$

There is at least one random variable belonging to any given cloud. The proof of this statement is given in [51]. This random variable is proved to exist, not assumed to exist as in the definition of an IVP. We can restate the property (2.72) by saying that

- the probability that a random variable $Y \in c$ belongs to the *upper α -cut* $\bar{A}_\alpha := \{u \in U \mid \bar{c}(u) > \alpha\}$ is at least $1 - \alpha$, and
- the probability that a random variable $X \in c$ belongs to the *lower α -cut* $\underline{A}_\alpha := \{u \in U \mid \underline{c}(u) \geq \alpha\}$ is at most $1 - \alpha$.

Figure 2.5 shows a cloud over \mathbb{R} with an α -cut.

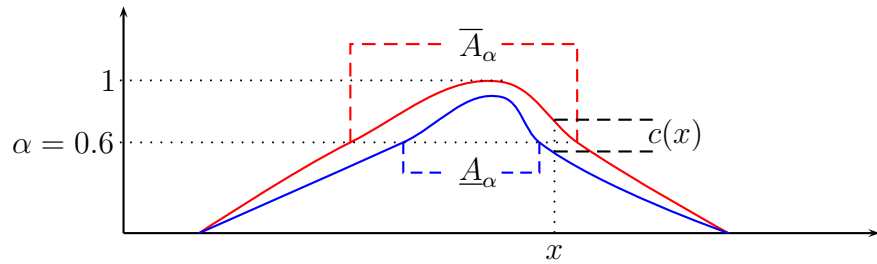


Figure 2.5: A cloud on \mathbb{R} with an α -cut at $\alpha = 0.6$.

Definition 2.28 (see Jamison [24]) *Let $pos : U \rightarrow [0, 1]$ be a regular possibility distribution function, i.e., $\sup_u \{pos(u) \in U\} = 1$, with associated possibility measure Pos and necessity measure Nec , (no need to be the dual of Pos). Then pos is said to be **consistent** with a random variable Y if for every measurable set A , $Nec(A) \leq Pr(Y \in A) \leq Pos(A)$.*

A pair of necessity and possibility measures in Definition 2.28 is an example of an IVP. The reference [36] shows that every cloud c can be defined as $c(u) = [\underline{n}(u), \bar{p}(u)]$, where \underline{p} and \bar{p} are regular possibility distribution functions

Table 2.3: The differences between IVPMs and clouds

	IVPMs	Clouds
Domain	σ_X	X
Range	$i(A) = [i^-(A), i^+(A)] \subseteq [0, 1],$ $\forall A \in \sigma_X$	$c(x) = [\underline{c}(x), \bar{c}(x)], \forall x \in X,$ and $(0, 1) \subseteq \bigcup_{x \in X} c(x) \subseteq [0, 1]$
Random variable	assumed to a part of IVPM's definition	proved to be in any clouds
Example	$[Nec(A), Pos(A)]$	$[nec(x), pos(x)],$ when pos is regular possibility distribution

on X such that $\forall u \in U; \underline{p}(u) + \bar{p}(u) \geq 1$ and $\underline{n}(u) = 1 - \underline{p}(u)$, and their associated possibility measures are consistent with every random variable belonging to c . More specifically, Destercke et al., [11], proves that a cloud c is representable by the pair of possibility distributions $1 - \underline{c}$ and \bar{c} . We distinguish some differences between the definitions of an IVPM and a cloud in Table 2.3.

Example 11. Suppose $U = \{1, 2, 3, 4, 5, 6\}$. A cloud over set U is given in Figure 2.6. We can see that

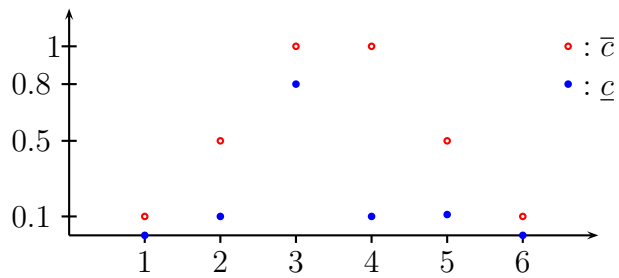


Figure 2.6: A cloud represents an IVPM.

$$\begin{aligned}
\alpha = 0 &\Rightarrow Pr(U) \leq 1 - 0 \leq Pr(U), \\
\alpha \in (0, 0.1) &\Rightarrow Pr(\{2, 3, 4, 5\}) \leq 1 - \alpha \leq Pr(U), \\
\alpha = 0.1 &\Rightarrow Pr(\{2, 3, 4, 5\}) \leq 1 - 0.1 \leq Pr(\{2, 3, 4, 5\}), \\
\alpha \in (0.1, 0.5) &\Rightarrow Pr(\{3\}) \leq 1 - \alpha \leq Pr(\{2, 3, 4, 5\}), \\
\alpha = 0.5 &\Rightarrow Pr(\{3\}) \leq 1 - 0.5 \leq Pr(\{3, 4\}), \\
\alpha \in (0.5, 0.8] &\Rightarrow Pr(\{3\}) \leq 1 - \alpha \leq Pr(\{3, 4\}), \\
\alpha \in (0.8, 1) &\Rightarrow 1 - \alpha \leq Pr(\{3, 4\}).
\end{aligned}$$

Hence, $Pr(U) = 1$, $Pr(\{2, 3, 4, 5\}) = 0.9$, $Pr(\{1, 6\}) = 0.1$, $Pr(\{3\}) \leq 0.2$, and $Pr(\{3, 4\}) \geq 0.2$. This information can be considered as an IVP, because we can set $i(\{3\}) = [0, 0.2]$, $i(\{3, 4\}) = [0.2, 1]$, and $i(A) = [0, 1]$, whenever any nonempty subset A of U is not one of the sets $U, \{2, 3, 4, 5\}, \{1, 6\}, \{3, 4\}$, or $\{3\}$. \diamond

Any cloud can be represented as an IVP, in general. Hence, we conclude that clouds are special examples of IVPs. Walley [73] introduced his imprecise probability theories to characterize a set of gambles of interest, where a gamble is a bounded real-valued function defined on a set U . However, his approach is too general for uncertainties concerned in this dissertation. We will not present the details of Walley's imprecise probability in this thesis. Interested readers can find out more in [73]. A discussion of relationships of these concepts in the sense of theory 'A' is more general than theory 'B' is well explained in [10] and [11]. We are interesting in arguing the relationships of different interpretations of an uncertainty \hat{u} . These interpretations are

1. possibility measure

- (a) possibility distribution
- (b) possibility on some subsets of U
- 2. necessity measure
 - (a) necessity distribution
 - (b) necessity on some subsets of U
- 3. plausibility measure
 - (a) plausibility distribution
 - (b) plausibility on some subsets of U
- 4. belief measure
 - (a) belief distribution
 - (b) belief on some subsets of U
- 5. random set
- 6. probability interval
- 7. probability on sets
- 8. cloud
- 9. IVP.

We wish to use an appropriate approach to obtain two special probability density mass functions that provide the smallest and the largest expected values of \hat{u} .

2.5 Relationships of some selected uncertainty interpretations

Some results in this chapter are generalized to be able to handle the continuous case of U . However for simplicity, we reduce the scope of this dissertation to the finite case of U . Since an interval interpretation of uncertainty falls into the continuous case of U , we will not consider interval uncertainty in this research. The continuous case is for further research.

We summarize the review of the interpretations of an uncertainty from Sections 2.1 - 2.4 as follows.

- We can derive a unique random set from possibility, necessity, plausibility, or belief measures by using the formula (2.10).
- We can derive a unique possibility, necessity, plausibility, or belief measures by using the formula (2.7) or (2.8).
- We can generate the random set that provides the largest set $\mathcal{M}_{\mathcal{F}}$ from a given partial information of a random set, as explained in Subsection 2.2.2.
- We can generate the random set that provides the largest set $\mathcal{M}_{\mathcal{F}}$ from a given a possibility, a necessity, a plausibility, or a belief distribution, as explained in Subsections 2.2.3 - 2.2.6.
- There is the random set generated by partial information in the form of possibility or necessity on some subsets of U . This random set also provides the largest set $\mathcal{M}_{\mathcal{F}}$.
- We can construct a random set from a given probability interval.

- Random set, probability interval, probability on sets, and cloud are examples of IVPMs.
- We consider partial information in the form of belief or plausibility on some subsets of U as an IVPM information.
- Probability is just a random set when all focal elements are singletons.

We now provide a full relationship diagram of all different uncertainty interpretations related to this thesis in Figure 2.7, which enhances the basic diagram of Figure 1.1 in Chapter 1.

We will never know with certainty the probability of a parameter when information we received is one of the uncertainty interpretations mentioned in this chapter. However, we can at least find out two probability density mass functions from \mathcal{M} that provide the smallest and the largest expected values. There are two approaches to obtain these two probability density mass functions. The first approach is by the constructions (2.42) and (2.43) when an uncertainty interpretation can be viewed as a random set. The second approach is by setting up two associated LP problems to solve for the minimum and maximum expected values of \hat{u} , when an uncertainty interpretation is considered to be an IVPM.

Our contribution results in this chapter are summarized below.

1. Remarks 2.9 and 2.10 provide the insight that $Bel(A)$ and $Pl(A)$, for each $A \subseteq U$, depend on the set Ω . Moreover, if we receive more information where the old information becomes more specific, then $Bel_{old}(A) \leq Bel_{updated}(A) \leq Pl_{updated}(A) \leq Pl_{old}(A)$.

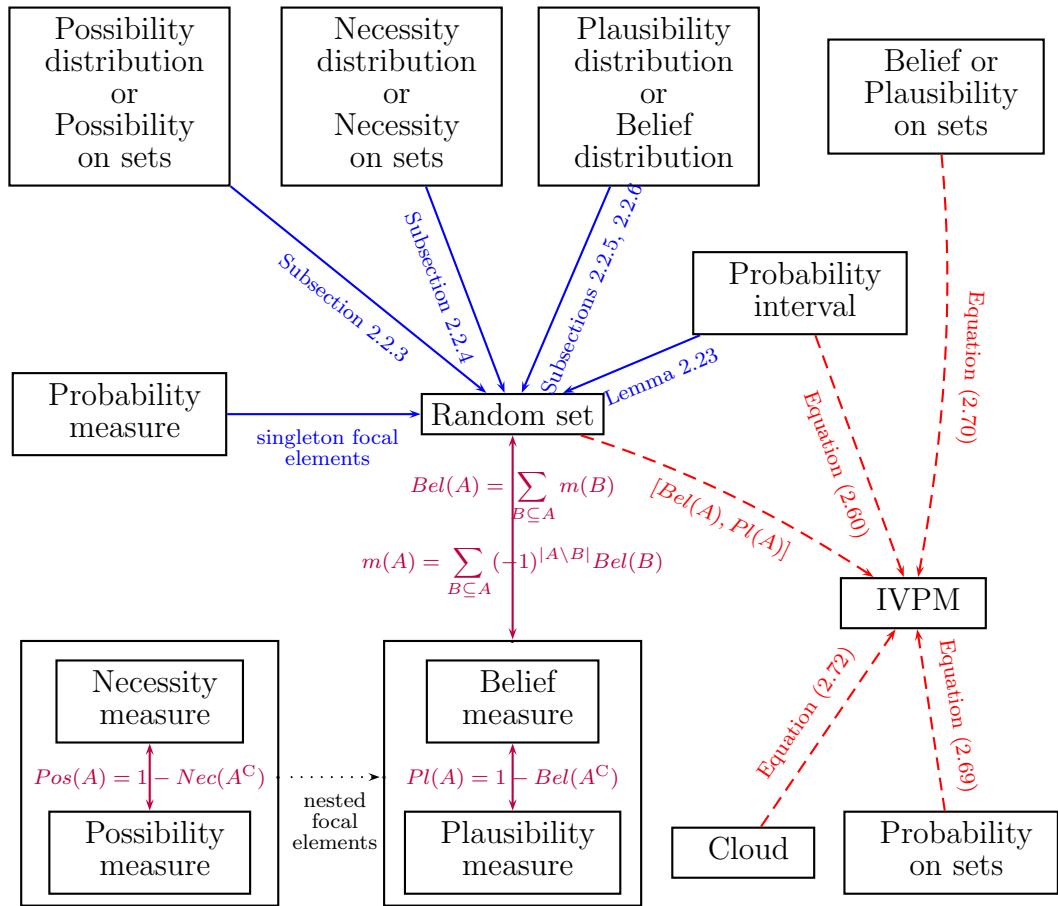


Figure 2.7: Uncertainty interpretations: $A \longrightarrow B$: there is an uncertainty interpretation B contains information given by an uncertainty interpretation A, $A \dashrightarrow B$: A is a special case of B, $A \longleftrightarrow B$: A and B can be derived from each other, and $A \cdots \rightarrow B$: B generalized A.

2. We provide a stronger statement to ensure the meanings of possibility and necessity, i.e., $Nec(A \cap B) = \min \{Nec(A), Nec(B)\}$ and $Pos(A \cup B) = \max \{Pos(A), Pos(B)\}$ if and only if the focal elements are nested. The proof of this statement is in Appendix B.
3. Based on a given uncertainty \hat{u} with random set interpretation, we prove in Theorem 2.15 that the lower and upper functions of \mathcal{M} are belief and plausibility functions.
4. Theorem 2.17 and 2.18 guarantee that probability density functions \underline{f} and \bar{f} provide the lowest and the highest expected values for a given uncertainty \hat{u} with random set interpretation. For a finite set of realizations of \hat{u} , Table 2.2 illustrates the number of *Bel* and *Pl* terms needed to derive \underline{f} and \bar{f} , which is much less than the number of *Bel* and *Pl* terms needed to state LP problems to find the lowest and the highest expected values.
5. Subsections 2.2.2 - 2.2.6 emphasize that if we have only partial information about belief (or other) measure or random set, then we still can find the random set generated by this partial information that provides the largest set $\mathcal{M}_{\mathcal{F}}$.
6. We can find a (not unique) random set that contains information given by a probability interval, see Lemma 2.23.
7. Theorem 2.25 and Corollary 2.26 are used for finding \underline{f} and \bar{f} , when \hat{u} has IVPM interpretation.

8. We provide the relationships of PC-BRIN interpretations of uncertainty in Figure 2.7.

Chapter 4 explains how to solve an LP problem with uncertainty to get a pessimistic and an optimistic result. It involves using these two probability density mass functions and two expected recourse models. Moreover, a minimax regret approach for an LP problem with uncertainty is presented to provide a minimax regret solution when the true probability of uncertainty is unknown, but its bound is known. The next chapter is devoted to a literature review of linear programming problems with uncertainty. Uncertainty presented in the review is limited to probability measure, possibility distribution, and interval (which is not developed in this thesis).

3. Linear optimization under uncertainty: literature review

We provide a review of literature dealing with modeling LP problems with uncertainty. If the objective is to find a solution for an LP problem with uncertainty that will minimize the maximum regret of using this solution due to the uncertainty, we may apply a minimax regret model to the problem and ignore any interpretations of uncertainty. However, if we are interesting in predicting the average of objective values in the long run, then we must consider uncertainty interpretations. Toward this end, we categorize uncertainty in LP problems depending on the information available about the uncertainties. They fall into the following cases.

- Uncertainty interpretations in the problem are probability. The modeling approach is an expected recourse model.
- Uncertainty interpretations in the problem are possibility. The modeling approach is an expected average model.
- Uncertainty interpretations in the problem can be both probability and possibility, but they are not in the same constraint. The modeling approach is an expected recourse-average model.
- Uncertainty interpretations in the problem can be both probability and possibility in the same constraint. The modeling approach is an interval expected value model.

To date, there is no modeling concept that capture LP problems with other uncertainty interpretations, except the ones listed above. We try to overcome this disadvantage by using a new approach (based on an expected recourse model), which uses the knowledge we had from Chapter 2 that each of PC-BRIN uncertainty can be represented as a set of all probability measures associated with their information. The details of the new approach are provided in Chapter 4.

In this chapter, the modeling concepts for the LP problems with uncertainty categorized above are presented. We begin with an illustration of a deterministic model of a production planning example. Then, after we explain the modeling concept for each of the LP problems with uncertainty categorized above, we provide an example for each concept by modifying this production planning example, which leads us to different types of models depending on the interpretations of uncertainties. The solutions of these models are done in GAMS, which is a program language for solving general optimization problems. For simplicity, we assume that we have information that already has been classified into probability and possibility interpretations of uncertainty. The differences should be seen in modeling and semantics of each example.

3.1 Deterministic model

A small production scheduling example is presented for ease and clarity of presentation. From two raw materials x_1 and x_2 , (for example, supplies of two grades of mineral oil), a refinery produces two different goods, cream #1 and cream #2. Table 3.1 shows the output of products per unit of the raw materials, the unit costs of raw materials (yielding the production cost z), the demands

for the products, and the maximal total amount of raw materials (production capacity) that can be processed. We assume that the mineral oil is mixed with other fluids (e.g., water) at no production cost to manufacture. Hence, these ingredients are not in the model.

Table 3.1: Productivity information shows the output of products per unit of raw materials, the unit costs of raw materials, the demand for the products, and the limitation of the total amount of raw materials.

Mineral oil (fl.oz.)	Products		Costs (\$/fl.oz.)	The limit amount of mineral oil that can be processed (fl.oz.)
	cream #1 (oz./fl.oz.)	cream #2 (oz./fl.oz.)		
x_1	2	3	2	1
x_2	6.1667	3	3	1
relation	\geq	\geq	=	\leq
	Demands			
	174.83	161.75	z	100

A manager wants to know how many units of raw materials to order to satisfy the demands and minimize the total production cost. Therefore, we have the linear programming problem (3.1) with the unique optimal solution $(x_1^*, x_2^*) = (37.84, 16.08)$, and $z^* = \$123.92$. Figure 3.1 illustrates the feasible region of the system (3.1) and its unique optimal solution.

$$\left. \begin{aligned}
 \min \quad & z := 2x_1 + 3x_2 \\
 \text{s.t.} \quad & 2x_1 + 6.1667x_2 \geq 174.83, \\
 & 3x_1 + 3x_2 \geq 161.75, \\
 & x_1 + x_2 \leq 100, \\
 & x_1, x_2 \geq 0.
 \end{aligned} \right\} \quad (3.1)$$

The values shown in Table 3.1 are fixed for the deterministic problem (3.1). However, this is not always a realistic assumption. For instance, if the produc-

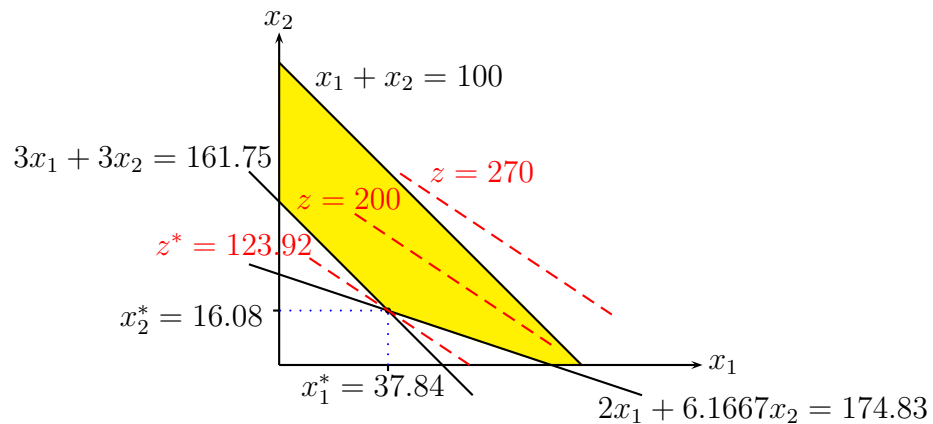


Figure 3.1: Feasible region and the unique optimal solution of the production scheduling problem (3.1).

tivities and demands can vary within certain limits or even have different types of uncertainty, and/or we need to formulate a production plan before knowing the exact values of these data, then this deterministic model is not adequate. We have to reformulate the system (3.1) so that it can produce a reasonable result. In Sections 3.2-3.6, we present different types of models depending on the interpretations of uncertainty to put the main thrust of this thesis, pessimistic, optimistic and minimax regret of expected recourse problems, in the context of other optimization under uncertainty.

3.2 Stochastic models

The aim of stochastic programming is to find an optimal decision in problems involving uncertain data of probability interpretation. The terminology ‘*stochastic*’ is opposed to ‘*deterministic*’ and means that some data are random. There are two well-known stochastic programming models. One is a *stochastic program with expected recourse*, which transforms the randomness contained in a stochastic program into an expectation of the distribution of some random vector. The

other is a *stochastic program with chance constraints*, which the constraints can satisfy the requirements with some probability or reliability level. One of these models is chosen for a particular problem under appropriate assumptions. However, we will not present a stochastic program with chance constraints in this dissertation. The details on this topic can be found in many books, e.g., [5, 26].

A general formulation for a *stochastic linear program* is the same as the model (1.2), when \hat{A} and \hat{b} have probability interpretations. We restate this problem here:

$$\min_x c^T x \quad \text{s.t.} \quad \hat{A}x \geq \hat{b}, \quad Bx \geq d, \quad x \geq 0, \quad (3.2)$$

where $\hat{A}x \geq \hat{b}$ contains m constraints of uncertain inequalities. This problem is not well-defined, because of the uncertainties. To be more specific, let us assume the following in addition to what we had in the problem (3.1).

- A1** The raw materials for the weekly production process rely on the supply of two grades of mineral oil, denoted by x_1 and x_2 , respectively.
- A2** The refinery uses two grades of mineral oil (and other ingredients at no cost) to produce cream #1 and cream #2.
- A3** The weekly demands of cream #1 and cream #2 vary randomly and mutually independent on each other (for simplicity).
- A4** The production (the output of the product per unit of the raw materials) of cream #1 varies randomly.
- A5** The production per unit of cream #2 is fixed.

Given the above assumptions, the problem becomes a stochastic linear program (3.3), where \hat{a}_{11} and \hat{a}_{12} are the production of cream #1 per unit of mineral oil x_1 and x_2 , and \hat{b}_1 and \hat{b}_2 are random demands. They possess known probability of demands for cream #1 and cream #2, respectively. It is not clear how to solve the minimization problem (3.3), since it is not a well-defined problem before knowing the realizations of $(\hat{a}_{11}, \hat{a}_{12}, \hat{b}_1, \hat{b}_2)$.

$$\left. \begin{array}{ll} \min 2x_1 + 3x_2 \\ \text{s.t. } \hat{a}_{11}x_1 + \hat{a}_{12}x_2 \geq \hat{b}_1, & 3x_1 + 3x_2 \geq \hat{b}_2, \\ x_1 + x_2 \leq 100, & x_1, x_2 \geq 0. \end{array} \right\} \quad (3.3)$$

We have to formulate a production plan *under uncertainty*, since we only have the probability distributions of the random demands.

3.2.1 Stochastic program with expected recourse

The terminology ‘*recourse*’ refers to a second (or further) action that helps improve a situation whenever the first action (taken before knowing the realization of uncertainty) does not satisfy the requirements. An example of the situation and requirements in the model is when the manufacturer is required to satisfy the customer’s demands, which are uncertain, while trying to minimize the production cost. The manager has to make a decision on the amount of raw materials (the first action) without knowing the actual demands. Later on, if these amounts of raw materials do not satisfy the actual demands, the manager needs to buy any shortage amount (recourse variables, or second action) from an outside market with some price attached.

An expected recourse model minimizes the cost of the first action and the expected cost of the second action. Define the random vector $\xi =$

$(\widehat{a}_{11}, \widehat{a}_{12}, \dots, \widehat{a}_{mn}, \widehat{b}_1, \widehat{b}_2, \dots, \widehat{b}_m)^T$, and the penalty price vector for m constraints as $s = (s_1, s_2, \dots, s_m)^T \geq 0$. Let $\Psi = \{x \mid Bx \geq d, x \geq 0\}$. An *expected recourse model* has the general formula

$$\min_{x \in \Psi} c^T x + E_\xi Q(x, \xi), \quad (3.4)$$

where $E_\xi Q(x, \xi)$ is the expected value of $Q(x, \xi)$ with respect to random variable ξ , and $Q(x, \xi) = s^T \left(\max \left[(\widehat{b} - \widehat{A}x), 0 \right] \right)$. More precisely, suppose there are α_{ij} and β_i finite realizations of each uncertainty \widehat{a}_{ij} in the matrix \widehat{A} and \widehat{b}_i in the vector \widehat{b} , respectively. Thus, there are $\prod_{j=1}^n \prod_{i=1}^m \alpha_{ij} \beta_i = N$ scenarios of ξ , which are denoted as $\xi^k, \forall k = 1, 2, \dots, N$. Each scenario has probability of occurrence pr_k . Hence, $Q(x, \xi) = (Q(x, \xi^1), Q(x, \xi^2), \dots, Q(x, \xi^N))$, where A^k and b^k are the matrix and vector at the k^{th} scenario, and $Q(x, \xi^k) = s^T \left(\max \left[(b^k - A^k x), 0 \right] \right)$. Let $y(\xi) = \max \left[(\widehat{b} - \widehat{A}x), 0 \right]$. The variable $y(\xi)$ is a *recourse variable* vector of shortages. Model (3.4) can be rewritten as

$$\min_{x \in \Psi} c^T x + E_\xi (s^T y(\xi)) \quad \text{s.t.} \quad y(\xi) \geq \widehat{b} - \widehat{A}x, \quad y(\xi) \geq 0. \quad (3.5)$$

We illustrate an example of an expected recourse problem by supposing that the assumptions A1-A5 apply to this subsection and assume further as follows.

A6 Fixed amounts x_1 and x_2 of two grades of mineral oil must be chosen in advance of knowing actual demand and cannot change during the week.

A7 The customers expect their actual demands of cream #1 and cream #2 to be satisfied.

A8 Fixed ‘*penalty*’ prices will be incurred if the customer’s demands cannot be covered by the production, since the amount of shortages has to be bought from a market.

We assume here that the penalty prices per unit of shortages are $s_1 = \$7$ and $s_2 = \$12$ for cream #1 and cream #2, respectively. For simplicity, we suppose further that there is no storage cost or selling cost for the over production.

Assume that the realizations of the random productivity of cream #1 and the random demands of both creams are finite with their probabilities as follows:

$$\left. \begin{array}{l} \widehat{a}_{11} = \left\{ \begin{array}{l} 1 \quad Pr(\{1\}) = 1/4 \\ 2 \quad Pr(\{2\}) = 1/2 \\ 3 \quad Pr(\{3\}) = 1/4 \end{array} \right. \\ \widehat{b}_1 = \left\{ \begin{array}{l} 149 \quad Pr(\{149\}) = 5/12 \\ 180 \quad Pr(\{180\}) = 1/3 \\ 211 \quad Pr(\{211\}) = 1/4 \end{array} \right. \\ \widehat{a}_{12} = \left\{ \begin{array}{l} 5 \quad Pr(\{5\}) = 1/6 \\ 6 \quad Pr(\{6\}) = 1/2 \\ 7 \quad Pr(\{7\}) = 1/3 \end{array} \right. \\ \widehat{b}_2 = \left\{ \begin{array}{l} 138 \quad Pr(\{138\}) = 1/4 \\ 162 \quad Pr(\{162\}) = 1/2 \\ 185 \quad Pr(\{185\}) = 1/4. \end{array} \right. \end{array} \right\} \quad (3.6)$$

Define the random vector $\xi = (\widehat{a}_{11}, \widehat{a}_{12}, \widehat{b}_1, \widehat{b}_2)^\top = \{\xi^1 = (1, 5, 149, 138)^\top, \xi^2 = (1, 5, 149, 162)^\top, \dots, \xi^{81} = (3, 7, 211, 185)^\top\}$, and introduce recourse variables $y_1(\xi)$ and $y_2(\xi)$ measuring the shortages of cream #1 and cream #2, respectively. The $y_i(\xi), i = 1, 2$, are themselves random variables since shortage depends on the realizations of random vector ξ . We also refer to the expected costs due to any shortage of production (or due to the amount of violation in the constraints, in general) as the expected *recourse costs*. We can replace the stochastic problem (3.3) by the well-defined stochastic program with expected recourse (3.7) using E_ξ as an expected value with respect to the distribution of ξ .

$$\left. \begin{array}{l} \min \quad 2x_1 + 3x_2 + E_\xi Q(x_1, x_2, \xi) \\ \text{s.t.} \quad x_1 + x_2 \leq 100, \\ \quad \quad x_1, x_2 \geq 0, \end{array} \right\} \quad (3.7)$$

where for each $j = 1, 2, \dots, 81$,

$$\begin{aligned}
Q(x_1, x_2, \xi^j) &:= 7 \max \{b_1^j - a_{11}^j x_1 - a_{12}^j x_2, 0\} + 12 \max \{b_2^j - 3x_1 - 3x_2, 0\} \\
&= \min \quad 7y_1 + 12y_2 \\
&\quad s.t. \quad a_{11}^j x_1 + a_{12}^j x_2 + y_1(\xi^j) \geq b_1^j, \\
&\quad \quad \quad 3x_1 + 3x_2 + y_2(\xi^j) \geq b_2^j, \\
&\quad \quad \quad y_1, y_2 \geq 0.
\end{aligned}$$

The mathematical program (3.7) is a linear programming problem which yields an optimal solution of $x_1^* = 31.80$ and $x_2^* = 29.87$, with the corresponding production (*first-stage*) costs of \$153.21 and the sum of the first-stage costs and the expected recourse costs of \$155.38. This optimal solution violates $x_1 + 5x_2 \geq 211$, which results in the recourse variable $y_1(\xi^3) = 29.85$. However, this constraint happens with a very small probability of $\frac{1}{4} \frac{1}{6} \frac{1}{4} = \frac{1}{96}$. Hence, we can interpret this optimal solution as that in the long run, the manufacturer will spend on average of \$153.21 for the production planning cost by using 31.80 units of mineral oil grade 1 and 29.87 units of mineral oil grade 2.

Suppose instead of using the expected recourse model, we apply the mean values of these demands and the productivity of cream #1. The problem becomes the deterministic model (3.1), and the supply solutions are $x_1 = 37.84$ and $x_2 \cong 16.08$. Then suppose further that it turns out later that the customer needs 200 units of cream #1 and demands 155 units of cream #2. Moreover, the realizations of the productivity for cream #1 are $a_{11} = 2$ and $a_{12} = 5$. We, as the adviser for a manager of this production plan, would need to deal with the excess of cream #2 and the shortage of cream #1 in some way. Un-

der the assumption of no cost for excess production, if the priority is to meet the customer's demand, then $2 \times 37.84 + 5 \times 16.08 = 156.08$, but we are short $200 - 156.08 = 43.92$ units of cream #1. We need to spend $\$(7 \times 43.92) = \307.44 extra, incurring a total cost $\$123.92 + \$307.44 = \$431.36$.

The recourse problem handles all possible cases that might occur and provides the long run average of what the supply should be. If the same situation happens, the supplies $x_1 \cong 31.80$ and $x_2 = 29.87$ will provide enough to satisfy the client's demand because $2x_1 + 5x_2 = 212.95 > 200$ and $3x_1 + 3x_2 = 185 > 155$. Therefore, it is better to choose the recourse model (3.7) over the deterministic model (3.1) when the actual demands are unknown, since when $x_1 = 31.80$ and $x_2 = 29.87$ are chosen, we are able to satisfy most variation in the client's demands while minimizing the production cost. On the other hand, if we decide a priori, we will incur $\$307.44$ extra costs to be able to satisfy the demand.

Stochastic optimization has the assumption that probability distributions of uncertainties, (random variables), are known exactly using statistical data, (in our example, they are discrete). However, in many applications, the decision-maker does not have this information. For instance, in assessing customer demand for a product, the lack of historical data for new items is an obvious challenge to estimating probabilities. Moreover, historical data is not certain to guarantee the same distribution applies to a future decision. Even well-established product lines can face sudden changes in demand, due to the market entry by a competitor or negative/positive publicity. In these (and other) cases, the manufacturer may use other types of uncertainties based on the information it has.

3.3 Minimax regret model

A minimax regret model for an LP problem with uncertainty is the model that tries to find a solution that minimizes the maximum regret of using this solution due to uncertainty. The model seeks to mitigate the effects of uncertainty rather than anticipating it. Therefore, the probability (or other interpretations) of each uncertainty is not required for this model. More information on minimax regret approaches can be found in [49], for example.

The uncertainty model (3.2) can be expanded into N deterministic LP problems, for some positive interger N . Define

$$z(\xi^k) := \min c^T x, \text{ s.t. } A^k x \geq b^k, Bx \geq d \text{ and } x \geq 0,$$

where $\xi^k = (A^k, b^k)$ is the data for the k^{th} deterministic LP problem. The regret function $r(x)$ shows the amount a candidate solution x deviates form the true objective value $z(\xi^k)$, i.e.,

$$r(x) = c^T x - z(\xi^k).$$

Therefore, the maximum (worst) regret of a candidate solution x is

$$R(x) = \max \{c^T x - z(\xi^k), k = 1, 2, \dots, N\}.$$

Hence, a general *minimax regret model* is

$$\begin{aligned} \min \quad & R \\ \text{s.t.} \quad & R \geq c^T x - z(\xi^k), \quad k = 1, 2, \dots, N, \\ & A^k x \geq b^k, \quad k = 1, 2, \dots, N, \\ & Bx \geq d, \quad x \geq 0. \end{aligned}$$

In the case of a production plan example, let us define $z(\xi^k)$ by dropping the superscript k for an easier reading as follows:

$$\begin{aligned} z(\xi^k) = z_{(a_{11}, a_{12}, b_1, b_2)} &:= \min 2x_1 + 3x_2 \\ \text{s.t. } a_{11}x_1 + a_{12}x_2 &\geq b_1, \quad 3x_1 + 3x_2 \geq b_2, \\ x_1 + x_2 &\leq 100, \quad x_1, x_2 \geq 0, \end{aligned}$$

where the scenarios of demands for cream #1 and cream #2 are $b_1 \in \{149, 180, 211\}$ and $b_2 \in \{138, 162, 185\}$, respectively. The scenarios of the productivity of cream #1 are $a_{11} = \{1, 2, 3\}$ and $a_{12} = \{5, 6, 7\}$. The regret function shows the amount a candidate solution (x_1, x_2) deviates from the true objective value $z_{(a_{11}, a_{12}, b_1, b_2)}$, i.e.,

$$r(x_1, x_2) = 2x_1 + 3x_2 - z_{(a_{11}, a_{12}, b_1, b_2)}.$$

Therefore, the maximum (worst) regret of a candidate solution (x_1, x_2) is

$$R(x_1, x_2) = \max_{\substack{a_{11} \in \{1, 2, 3\} \\ a_{12} \in \{5, 6, 7\} \\ b_1 \in \{149, 180, 211\} \\ b_2 \in \{138, 162, 185\}}} 2x_1 + 3x_2 - z_{(a_{11}, a_{12}, b_1, b_2)}.$$

The minimax regret model is a model that finds the best of the worst regret over all candidate solutions. We formulate the minimax regret for the production scheduling problem as:

$$\left. \begin{aligned} \min_{(x_1, x_2)} R(x_1, x_2) \\ \text{s.t. } a_{11}x_1 + a_{12}x_2 &\geq b_1, \quad \forall a_{11} \in \{1, 2, 3\}, \forall a_{12} \in \{5, 6, 7\}, \\ &\quad \forall b_1 \in \{149, 180, 211\}, \\ 3x_1 + 3x_2 &\geq b_2, \quad \forall b_2 \in \{138, 162, 185\}, \\ x_1 + x_2 &\leq 100, \quad x_1, x_2 \geq 0. \end{aligned} \right\} (3.8)$$

The minimum value of $z_{(a_{11}, a_{12}, b_1, b_2)}$ is 94.74, which happens when $a_{11} = 3$, $a_{12} = 6$, $b_1 = 149$, and $b_2 = 138$. So $R(x_1, x_2) = 2x_1 + 3x_2 - 94.74$, and we can rewrite the model (3.8) as

$$\left. \begin{aligned} \min_{(x_1, x_2)} R(x_1, x_2) &:= 2x_1 + 3x_2 - 94.74 \\ \text{s.t. } a_{11}x_1 + a_{12}x_2 &\geq b_1, & \forall a_{11} \in \{1, 2, 3\}, \forall a_{12} \in \{5, 6, 7\}, \\ & & \forall b_1 \in \{149, 180, 211\}, \\ 3x_1 + 3x_2 &\geq b_2, & \forall b_2 \in \{138, 162, 185\}, \\ x_1 + x_2 &\leq 100, & x_1, x_2, \geq 0. \end{aligned} \right\} \quad (3.9)$$

The mathematical model (3.9) can be reduced to the system (3.10), since every feasible solution of (3.10) satisfies all other constraints in (3.9).

$$\left. \begin{aligned} \min_{(x_1, x_2)} R(x_1, x_2) &:= 2x_1 + 3x_2 - 94.74 \\ \text{s.t. } 1x_1 + 5x_2 &\geq 211, \\ 3x_1 + 3x_2 &\geq 185, \\ x_1 + x_2 &\leq 100, \\ x_1, x_2, &\geq 0. \end{aligned} \right\} \quad (3.10)$$

The minimax regret value provided by (3.10) is $R^* = 65.91$ with $x_1^* = 24.33$ and $x_2^* = 37.33$. We interpret this information as when we choose $x_1^* = 24.33$ and $x_2^* = 37.33$ without knowing the realization of a_{11} , a_{12} , b_1 , and b_2 , the worst that we may regret based upon this decision costs \$65.91.

A manager can choose a suitable model depending on operating priorities. The manager may choose the expected recourse model (3.7) over the minimax regret (3.10) to spend as little as possible in the long run. Even if the manager may need to pay the price of $\$7 \times 29.85 = \208.95 for the shortage of cream

#1 when the realization of (a_{11}, a_{12}, b_1) is $(1, 5, 211)$, this realization has a small chance of happening. Therefore, in the long run, the manager will save more money by the identified optimal solution for the production planning problem. However, the manager may choose to use the model (3.10) over (3.7), to avoid taking any chances of spending the amount of \$208.95. Moreover, concerns about the model complexity may lead to a preference of (3.10) over (3.7).

3.4 Possibility model

An LP problem with possibility uncertainty (in the form of possibility distribution) treated by many researchers provides a solution as a function of an acceptance level of uncertainty, see for example in [12, 57, 64, 65] and [66]. However, this solution is not applicable for a user with an as yet undecided satisfaction level. There is another approach using the expected average (to be explained shortly) which coincides with the expected recourse approach when uncertainties have probability interpretations.

Let \hat{u} be an uncertainty with realizations in the set $U = \{u_1, u_2, \dots, u_n\} \subseteq \mathbb{R}$, and a possibility distribution $pos : U \rightarrow [0, 1]$. Define an α -level set of \hat{u} as $U_\alpha = \{u_i \in U \mid pos(u_i) \geq \alpha\}$. It is noted that U_α is an ordinary subset of U . Yager [76] defines the mean value of U_α as $M(U_\alpha) = \frac{1}{|U_\alpha|} \sum_{u_i \in U_\alpha} u_i$. $M(U_\alpha)$ expresses the average value of \hat{u} at the satisfaction level α . An *expected average EA* of \hat{u} is the mean of $M(U_\alpha)$ for all α -levels, i.e.,

$$EA(\hat{u}) = \int_0^1 M(U_\alpha) d\alpha. \quad (3.11)$$

Suppose \hat{v} is an uncertainty with the set of realizations $V = \{v_1, v_2, \dots, v_m\} \subseteq \mathbb{R}$, and a possibility distribution $pos_v : V \rightarrow [0, 1]$. Yager [76] proves that an α -level

set of $\hat{u} + \hat{v}$ is $U_\alpha \oplus V_\alpha$, where $U_\alpha \oplus V_\alpha$ is the set consisting of the sum of each element of U_α with all the elements of v_α . Jamison [24] states a general *expected average model*, which coincides with an expected recourse model for probability interpretations of uncertainties, as follows:

$$\min_{x \in \Psi} c^T x + EA \left(s^T \left(\max \left\{ \hat{b} - \hat{A}x, 0 \right\} \right) \right). \quad (3.12)$$

When U is continuous, Carlsson and Fullér [6] propose another concept for an expected average. However, they do not explain the discrete case of U . There also is no possibility model based on this approach. Therefore, we do not discuss further their expected average approach.

Suppose that the productivity of cream #1 and the demands in the production planning example are known to have the following possibility distributions:

$$\left. \begin{aligned} \hat{a}_{11} &= \begin{cases} 1 & Pos(\{1\}) = \frac{1}{3}, \\ 2 & Pos(\{2\}) = 1, \\ 3 & Pos(\{3\}) = \frac{1}{2}, \end{cases} & \hat{a}_{12} &= \begin{cases} 5 & Pos(\{5\}) = \frac{1}{3}, \\ 6 & Pos(\{6\}) = \frac{3}{4}, \\ 7 & Pos(\{7\}) = 1, \end{cases} \\ \hat{b}_1 &= \begin{cases} 149 & Pos(\{149\}) = \frac{1}{2}, \\ 180 & Pos(\{180\}) = 1, \\ 211 & Pos(\{211\}) = \frac{2}{3}, \end{cases} & \hat{b}_2 &= \begin{cases} 138 & Pos(\{138\}) = \frac{1}{3}, \\ 162 & Pos(\{162\}) = \frac{3}{4}, \\ 185 & \text{with } Pos(\{185\}) = 1. \end{cases} \end{aligned} \right\} \quad (3.13)$$

Let $s = [7, 12]^T$ and $\hat{g} = [\hat{g}_1, \hat{g}_2]^T$ such that $\hat{g}_1 = \max \left\{ \hat{b}_1 - \hat{a}_{11}x_1 - \hat{a}_{12}x_2, 0 \right\}$, and $\hat{g}_2 = \max \left\{ \hat{b}_2 - 3x_1 - 3x_2, 0 \right\}$. Therefore, the expected average model for production planning problem becomes

$$\left. \begin{aligned} \min & 2x_1 + 3x_2 + EA(s^T \hat{g}) \\ \text{s.t.} & x_1 + x_2 \leq 100, \quad x_1, x_2 \geq 0. \end{aligned} \right\} \quad (3.14)$$

Given $(x_1, x_2) \in \{(x_1, x_2) \mid x_1 + x_2 \leq 100, x_1, x_2 \geq 0\}$, the α -level sets U_α^1 for $\widehat{b}_1 - \widehat{a}_{11}x_1 - \widehat{a}_{12}x_2$ and the α -level sets U_α^2 for $\widehat{b}_2 - 3x_1 - 3x_2$ are as follows.

For $\widehat{b}_1 - \widehat{a}_{11}x_1 - \widehat{a}_{12}x_2$,

$$\begin{aligned} \alpha \in [0, \frac{1}{3}] &\Rightarrow U_\alpha^1 = \{149 - 1x_1 - 5x_2, 149 - 1x_1 - 6x_2, \dots, 211 - 3x_1 - 7x_2\} \\ &\Rightarrow M(U_\alpha^1) = 180 - 2x_1 - 6x_2, \end{aligned}$$

$$\begin{aligned} \alpha \in (\frac{1}{3}, \frac{1}{2}] &\Rightarrow U_\alpha^1 = \{149 - 2x_2 - 6x_2, 149 - 3x_1 - 7x_2, \dots, 211 - 3x_1 - 7x_2\} \\ &\Rightarrow M(U_\alpha^1) = 180 - 2.5x_1 - 6.5x_2, \end{aligned}$$

$$\begin{aligned} \alpha \in (\frac{1}{2}, \frac{2}{3}] &\Rightarrow U_\alpha^1 = \{180 - 2x_1 - 6x_2, 180 - 2x_1 - 7x_2, \dots, 211 - 2x_1 - 6x_2, \\ &\quad 211 - 2x_1 - 7x_2\} \\ &\Rightarrow M(U_\alpha^1) = 195.50 - 2x_1 - 6.5x_2, \end{aligned}$$

$$\begin{aligned} \alpha \in (\frac{2}{3}, \frac{3}{4}] &\Rightarrow U_\alpha^1 = \{180 - 2x_1 - 6x_2, 180 - 2x_1 - 7x_2\} \\ &\Rightarrow M(U_\alpha^1) = 180 - 2x_1 - 6.5x_2, \end{aligned}$$

$$\begin{aligned} \alpha \in (\frac{3}{4}, 1] &\Rightarrow U_\alpha^1 = \{180 - 2x_1 - 7x_2\} \\ &\Rightarrow M(U_\alpha^1) = 180 - 2x_1 - 7x_2. \end{aligned}$$

For $\widehat{b}_2 - 3x_1 - 3x_2$,

$$\begin{aligned} \alpha \in [0, \frac{1}{3}] &\Rightarrow U_\alpha^2 = \{138 - 3x_1 - 3x_2, 162 - 3x_1 - 3x_2, 185 - 3x_1 - 3x_2\} \\ &\Rightarrow M(U_\alpha^2) = 161.67 - 3x_1 - 3x_2, \end{aligned}$$

$$\begin{aligned} \alpha \in (\frac{1}{3}, \frac{3}{4}] &\Rightarrow U_\alpha^2 = \{162 - 3x_1 - 3x_2, 185 - 3x_1 - 3x_2\} \\ &\Rightarrow M(U_\alpha^2) = 173.50 - 3x_1 - 3x_2, \end{aligned}$$

$$\begin{aligned} \alpha \in (\frac{3}{4}, 1] &\Rightarrow U_\alpha^2 = \{185 - 3x_1 - 3x_2\} \\ &\Rightarrow M(U_\alpha^2) = 185 - 3x_1 - 3x_2. \end{aligned}$$

The expansion of the objective function of (3.14) is

$$2x_1 + 3x_2 + 7 \int_0^{\frac{1}{3}} \max\{180 - 2x_1 - 6x_2, 0\} d\alpha + 7 \int_{\frac{1}{3}}^{\frac{1}{2}} \max\{180 - 2.5x_1 - 6.5x_2, 0\} d\alpha$$

$$\begin{aligned}
&+7 \int_{\frac{1}{2}}^{\frac{2}{3}} \max \{195.50 - 2x_1 - 6.5x_2, 0\} d\alpha + 7 \int_{\frac{2}{3}}^{\frac{3}{4}} \max \{180 - 2x_1 - 6.5x_2, 0\} d\alpha \\
&+7 \int_{\frac{3}{4}}^1 \max \{180 - 2x_1 - 7x_2, 0\} d\alpha + 12 \int_0^{\frac{1}{3}} \max \{161.67 - 3x_1 - 3x_2, 0\} d\alpha \\
&+12 \int_{\frac{1}{3}}^{\frac{3}{4}} \max \{173.50 - 3x_1 - 3x_2, 0\} d\alpha + 12 \int_{\frac{3}{4}}^1 \max \{185 - 3x_1 - 3x_2, 0\} d\alpha,
\end{aligned}$$

which is the same as

$$\begin{aligned}
&2x_1 + 3x_2 + \frac{7}{3} \max \{180 - 2x_1 - 6x_2, 0\} + \frac{7}{6} \max \{180 - 2x_1 - 6.5x_2, 0\} \\
&+ \frac{7}{6} \max \{195.50 - 2x_1 - 6.5x_2, 0\} + \frac{7}{12} \max \{180 - 2x_1 - 6.5x_2, 0\} \\
&+ \frac{7}{4} \max \{180 - 2x_1 - 7x_2, 0\} + 4 \max \{161.67 - 3x_1 - 3x_2, 0\} \\
&+ 5 \max \{173.50 - 3x_1 - 3x_2, 0\} + 3 \max \{185 - 3x_1 - 3x_2, 0\}.
\end{aligned}$$

Therefore we can rewrite the system (3.14) as

$$\left. \begin{aligned}
&\min \quad 2x_1 + 3x_2 + \frac{7}{3}y_1 + \frac{7}{6}y_2 + \frac{7}{6}y_3 + \frac{7}{12}y_4 + \frac{7}{4}y_5 + 4y_6 + 5y_7 + 3y_8 \\
&\quad y_1 \geq 180 - 2x_1 - 6x_2, \quad y_2 \geq 180 - 2x_1 - 6.5x_2, \\
&\quad y_3 \geq 195.50 - 2x_1 - 6.5x_2, \quad y_4 \geq 180 - 2x_1 - 6.5x_2, \\
\text{s.t.} \quad &y_5 \geq 180 - 2x_1 - 7x_2, \quad y_6 \geq 161.67 - 3x_1 - 3x_2, \\
&y_7 \geq 173.50 - 3x_1 - 3x_2, \quad y_8 \geq 185 - 3x_1 - 3x_2, \\
&x_1 + x_2 \leq 100, \quad x, y \geq 0.
\end{aligned} \right\} (3.15)$$

Thus, for the same penalty prices as the stochastic expected recourse model, the system (3.15) has a unique optimal solution $(x_1^*, x_2^*) = (45.63, 16.04)$ and the objective value of \$139.37. The interpretation of this objective value is that the cost of two grades of mineral oil together with the expected average cost of shortages over all the satisfaction levels, with the given possibility uncertainty information, is \$139.37.

We are now in a position to compare the solutions obtained by expected recourse and expected average models. That is, suppose we had used a pos-

sibility distribution given by (3.13) instead of the probability density given by (3.6) page 87 and see how much the objective value from the possibility model differs from the objective value from the expected recourse model. For the production planning example, the objective value of (3.14) is lower than (3.7) by $155.38 - 139.37 = \$16.01$. However, we also can see that the expected average model (3.15) does not use all realizations in its constraints. For example, -1 as a realization of $-\widehat{a}_{11}$, which is the coefficient of x_1 , does not appear in any of the constraints of (3.15). Moreover, some parameters in the constraints are not the realizations of their associated uncertainties, e.g., 173.50 and 195.50 are not realizations of \widehat{b}_1 . They are in fact the mean values of some α -level sets. Therefore, the optimal objective values for an expected recourse model and an expected average model are not related to each other, in general, even if the probabilities in the expected recourse model are trapped in the information of possibility distributions of the expected average model. The optimal objective value of one model could be less than or greater than the other, depending upon the information they have.

Moreover, Theorem 2.15 confirms that probabilities from (3.6) are trapped between necessity and possibility, i.e.

$$\widehat{a}_{11} = \begin{cases} 1, & Pos(\{1\}) = \frac{1}{3} & Pr(\{1\}) \in [0, \frac{1}{3}], Pr(\{1, 2\}) \in [0, 1], \\ 2, & Pos(\{2\}) = 1 \Rightarrow Pr(\{2\}) \in [\frac{1}{2}, 1], Pr(\{1, 3\}) \in [0, \frac{1}{2}], \\ 3, & Pos(\{3\}) = \frac{1}{2} & Pr(\{3\}) \in [0, \frac{1}{2}], Pr(\{2, 3\}) \in [\frac{1}{2}, 1], \end{cases}$$

$$\widehat{a}_{12} = \begin{cases} 5, & Pos(\{5\}) = \frac{1}{3} & Pr(\{5\}) \in [0, \frac{1}{3}], Pr(\{5, 6\}) \in [0, \frac{3}{4}], \\ 6, & Pos(\{6\}) = \frac{3}{4} \Rightarrow Pr(\{6\}) \in [0, \frac{3}{4}], Pr(\{5, 7\}) \in [\frac{1}{4}, 1], \\ 7, & Pos(\{7\}) = 1 & Pr(\{7\}) \in [\frac{1}{4}, 1], Pr(\{6, 7\}) \in [\frac{2}{3}, 1], \end{cases}$$

$$\widehat{b}_1 = \begin{cases} 149, Pos(\{149\}) = \frac{1}{2} & Pr(\{149\}) \in [0, \frac{1}{2}], Pr(\{149, 180\}) \in [\frac{1}{3}, 1], \\ 180, Pos(\{180\}) = 1 \Rightarrow Pr(\{180\}) \in [\frac{1}{3}, 1], Pr(\{149, 211\}) \in [0, \frac{2}{3}], \\ 211, Pos(\{211\}) = \frac{2}{3} & Pr(\{211\}) \in [0, \frac{2}{3}], Pr(\{180, 211\}) \in [\frac{1}{2}, 1], \\ \text{and} \end{cases}$$

$$\widehat{b}_2 = \begin{cases} 138, Pos(\{138\}) = \frac{1}{3} & Pr(\{138\}) \in [0, \frac{1}{3}], Pr(\{138, 162\}) \in [0, \frac{3}{4}], \\ 162, Pos(\{162\}) = \frac{3}{4} \Rightarrow Pr(\{162\}) \in [0, \frac{3}{4}], Pr(\{138, 185\}) \in [\frac{1}{4}, 1], \\ 185, Pos(\{185\}) = 1 & Pr(\{185\}) \in [\frac{1}{4}, 1], Pr(\{162, 185\}) \in [\frac{2}{3}, 1]. \end{cases}$$

Probability interpretations of uncertainties $\widehat{a}_{11}, \widehat{a}_{12}, \widehat{b}_1$, and \widehat{b}_2 of the expected recourse problem (3.7) are in the bounds of the above information. The expected recourse model (3.7) provides an optimal solution that helps predicting the production planning in the long run with respect to the given probabilities in the bounds of these necessity and possibility measures. The set of all probabilities trapped between these necessity and possibility measures are the same information as the possibility interpretations of the expected average model (3.14). However, we loose information by changing it to expected averages. Hence, the expected average model (3.14) does not carry information along its calculation that well. Based on the information as the bounds on the probabilities, a question is raised about the meaning of an optimal solution of an expected average model, rather than the semantic that has the origin from the mathematical definition of an expected average. We are unable to explain the meaning of an optimal solution of an expected average model with respect to the possibility interpretation of uncertainty as the bounds of probabilities. More discussions on this topic are provided in Section 4.4 after the presentation of new approaches that can handle possibility information as bounds on its associated probabilities.

3.5 Mixed model of probability and possibility

For this section, some uncertainties may be viewed as possibility information, and some may have enough statistical data to impute probability distributions. However, there is no approach in the literature that handles these two uncertainty interpretations in one constraint, except an interval expected value approach [36, 69] presented in the next section. When possibility and probability interpretations of uncertainty are in different constraints of an LP problem with uncertainty, the transformation of this LP problem to a deterministic model is by combining an expected recourse and expected average models together.

Let $\widehat{A}x \geq \widehat{b}$ be m uncertain inequalities with probability interpretations, and $\widehat{A}'x \geq \widehat{b}'$ be m' uncertain inequalities with possibility interpretations. A general form of a mixed LP problem with possibility and probability interpretations in different constraints is

$$\min_x c^T x \quad \text{s.t.} \quad \widehat{A}x \geq \widehat{b}, \quad \widehat{A}'x \geq \widehat{b}', \quad Bx \geq d, \quad x \geq 0. \quad (3.16)$$

Then, an *expected recourse-average model* of (3.16) is

$$\min_{x \in \Psi} c^T x + E_{\xi} Q(x, \xi) + EA \left(s'^T \left(\max \left\{ \widehat{b}' - \widehat{A}'x, 0 \right\} \right) \right), \quad (3.17)$$

where, $s' = (s'_1, s'_2, \dots, s'_{m'})$ is the penalty price vector of $\widehat{A}'x \geq \widehat{b}'$.

Suppose that the uncertainty information is the same as presented on page 94, except that the demand of cream #2 varies randomly with the same probability of \widehat{b}_2 as presented on page 87. Hence, the mixed probabilistic and possibility model is stated as in the system (3.18).

$$\left. \begin{array}{ll} \min 2x_1 + 3x_2 & \\ \text{s.t. } \widehat{a}_{11}x_1 + \widehat{a}_{12}x_2 \geq \widehat{b}_1, & 3x_1 + 3x_2 \geq \widehat{b}_2, \\ x_1 + x_2 \leq 100, & x_1, x_2 \geq 0, \end{array} \right\} \quad (3.18)$$

where

$$\widehat{a}_{11} = \begin{cases} 1 & \text{with } Pos(\{1\}) = \frac{1}{3}, \\ 2 & \text{with } Pos(\{2\}) = 1, \text{ and} \\ 3 & \text{with } Pos(\{3\}) = \frac{1}{2}, \end{cases} \quad \widehat{a}_{12} = \begin{cases} 5 & \text{with } Pos(\{5\}) = \frac{1}{3}, \\ 6 & \text{with } Pos(\{6\}) = \frac{3}{4}, \\ 7 & \text{with } Pos(\{7\}) = 1. \end{cases}$$

$$\widehat{b}_1 = \begin{cases} 149 & \text{with } Pos(\{149\}) = \frac{1}{2}, \\ 180 & \text{with } Pos(\{180\}) = 1, \text{ and} \\ 211 & \text{with } Pos(\{211\}) = \frac{2}{3}, \end{cases} \quad \widehat{b}_2 = \begin{cases} 138 & \text{with } Pr(\{138\}) = \frac{1}{4}, \\ 162 & \text{with } Pr(\{162\}) = \frac{1}{2}, \\ 185 & \text{with } Pr(\{185\}) = \frac{1}{4}. \end{cases}$$

Problem (3.18) is not well-defined until we know a realization of the uncertainties. It is also not the same as problem (3.14), since \widehat{b}_2 now has probability interpretation. Applying recourse variables to the probabilistic constraints and using the expected average for possibilistic constraints, the expected recourse-average of (3.18) is

$$\left. \begin{array}{l} \min \ 2x_1 + 3x_2 + E_{\widehat{b}_2} Q(x, \widehat{b}_2) + EA \left(7 \max \left\{ \widehat{b}_1 - \widehat{a}_{11}x_1 - \widehat{a}_{12}x_2, 0 \right\} \right) \\ \text{s.t. } \ x_1 + x_2 \leq 100, \\ \quad \quad x_1, x_2 \geq 0, \end{array} \right\} (3.19)$$

where $Q(x, \widehat{b}_2) = \max 12y$ s.t. $3x_1 + 3x_2 + y(b_2^i) \geq b_2^i$, $y \geq 0$, and $b_2^1 = 138$, $b_2^2 = 162$, and $b_2^3 = 185$.

An optimal solution for (3.19) is $(x_1^*, x_2^*) = (45.63, 16.04)$, and the optimal objective value is \$139.37. This result happens to be the same as one for (3.14). However, the semantic of the objective value for (3.14), through the mathematical definitions of expected value and expected average, is different from the semantic of the objective value for (3.19), as you may see. A similar question as in the last section has raised here as well. Moreover, this approach (of treating a

probability as a recourse model and a possibility as a penalized expected average model) is suitable only when each constraint has only one type of uncertainty, (either probability or possibility).

3.6 Mixed uncertainties model

We can apply an expected recourse function to the objective function of an LP problem with uncertainty when all uncertainties in a constraint have probability interpretations. Similarly, we use expected average when all uncertainties in a constraint are possibilities. However, it might happen that there is more than one interpretation of uncertainty in one constraint. In this case, an expected recourse-average approach is no longer suitable. Lodwick and Jamison [36] introduced the concept of interval expected value to handle this type of problem.

Let $U = \{u_1 < u_2 < \dots < u_l\}$ be the set of all realizations of an uncertainty \hat{u} , and $i : \mathcal{P}(U) \rightarrow \text{Int}_{[0,1]}$ be an IVP generated from consistent necessity, *Nec*, and possibility, *Pos*, measures (see Definition 2.28). Then, the *interval expected value* of this IVP is constructed from possibility and necessity measures as the specific upper and lower cumulative probability distributions, respectively. The upper cumulative probability distribution denoted by \bar{F} is defined as $\bar{F}(u_k) = \text{Pos}(\{u_1, u_2, \dots, u_k\})$. The lower cumulative probability distribution \underline{F} is defined as $\underline{F}(u_k) = \text{Nec}(\{u_1, u_2, \dots, u_k\})$. The interval expected value is defined in [36] as $E(i) = \left[\sum_{k=1}^l u_k \underline{f}(u_k), \sum_{k=1}^l u_k \bar{f}(u_k) \right]$, where \underline{f} and \bar{f} refer to the probability density mass functions associated with \bar{F} and \underline{F} , respectively. It turns out that these \underline{f} and \bar{f} are the same density functions from

the construction (2.42) and (2.43) on page 39, because

$$\begin{aligned}
\underline{f}(u_k) &= \overline{F}(u_k) - \overline{F}(u_{k-1}) \\
&= Pos(\{u_1, u_2, \dots, u_k\}) - Pos(\{u_1, u_2, \dots, u_{k-1}\}) \\
&= (1 - Nec(\{u_{k+1}, u_{k+2}, \dots, u_n\})) - (1 - Nec(\{u_k, u_{k+1}, \dots, u_n\})) \\
&= Nec(\{u_k, u_{k+1}, \dots, u_n\}) - Nec(\{u_{k+1}, u_{k+2}, \dots, u_n\}), \text{ and} \\
\overline{f}(u_k) &= \underline{F}(u_k) - \underline{F}(u_{k-1}) \\
&= Nec(\{u_1, u_2, \dots, u_k\}) - Nec(\{u_1, u_2, \dots, u_{k-1}\}).
\end{aligned}$$

Let $E(A)$ be an interval expected value matrix generated by an uncertainty matrix \widehat{A} , and $E(b)$ be an interval expected value vector generated by an uncertainty vector \widehat{b} . [36, 69] transform an LP with mixed possibility and probability interpretations of uncertainty to an interval linear programming by changing all uncertainties to their interval expected values, as shown in (3.20). We call this transformed model as an *interval expected value model*.

$$\left. \begin{array}{ll}
\min c^T x & \min c^T x \\
\text{s.t. } \widehat{A}x \geq \widehat{b} & \text{s.t. } E(A)x \geq E(b) \\
Bx \geq d & Bx \geq d \\
x \geq 0 & x \geq 0.
\end{array} \right\} \quad (3.20)$$

Although the uncertainty interpretations presented in [36, 69] are only probability and possibility, the relationships among different uncertainty interpretations in Chapter 2 imply that an interval expected value approach is suitable for other PC-BRIN uncertainty interpretations as well. However, for the same reason why we are not using the expected values to represent an LP problem with probability uncertainty, the interval expected value approach is a limited

representation for an LP problem with generalized uncertainty and not as robust as the new approaches developed in the next chapter.

Returning to the production planning problem, suppose that we allow three interpretations of uncertainty: possibility, probability, and probability interval, in the first constraint and assume that the demand of cream #2 in the second constraint varies randomly with the same probability of \widehat{b}_2 . The mixed uncertainties model, involving three interpretations of uncertainties; probability possibility, and probability interval, is the system (3.21)

$$\left. \begin{array}{l} \min 2x_1 + 3x_2 \\ \text{s.t. } \widehat{a}_{11}x_1 + \widehat{a}_{12}x_2 \geq \widehat{b}_1, \\ 3x_1 + 3x_2 \geq \widehat{b}_2, \\ x_1 + x_2 \leq 100, \\ x_1, x_2 \geq 0, \end{array} \right\} \quad (3.21)$$

where

$$\widehat{a}_{11} = \begin{cases} 1 & \text{with } Pos(\{1\}) = \frac{1}{3}, \\ 2 & \text{with } Pos(\{2\}) = 1, \\ 3 & \text{with } Pos(\{3\}) = \frac{1}{2}, \end{cases} \quad \text{and} \quad \widehat{a}_{12} = \begin{cases} 5 & \text{with } Pr(\{5\}) = \frac{1}{6}, \\ 6 & \text{with } Pr(\{6\}) = \frac{1}{2}, \\ 7 & \text{with } Pr(\{7\}) = \frac{1}{3}, \end{cases}$$

and

$$\widehat{b}_1 = \begin{cases} 149 & \text{with } Pr(\{149\}) = [\frac{1}{3}, \frac{1}{2}] \\ 180 & \text{with } Pr(\{180\}) = \frac{1}{3}, \\ 211 & \text{with } Pr(\{211\}) = [\frac{1}{6}, \frac{1}{3}], \end{cases} \quad \text{and} \quad \widehat{b}_2 = \begin{cases} 138 & \text{with } Pr(\{138\}) = \frac{1}{4}, \\ 162 & \text{with } Pr(\{162\}) = \frac{1}{2}, \\ 185 & \text{with } Pr(\{185\}) = \frac{1}{4}. \end{cases}$$

The uncertainty \widehat{a}_{11} has the density function \underline{f} as $\underline{f}(1) = \frac{1}{3}$, $\underline{f}(2) = \frac{2}{3}$, $\underline{f}(3) = 0$, and the density function \bar{f} as $\bar{f}(1) = 0$, $\bar{f}(2) = \frac{1}{2}$, $\bar{f}(3) = \frac{1}{2}$, which were

obtained by applying the construction (2.42) and (2.43). Therefore, the interval expected value of \widehat{a}_{11} is $[1.6667, 2.5]$. Similarly, we can get the interval expected value of \widehat{b}_1 as $[169.6667, 180]$. The LP problem with mixed uncertainty (3.21) becomes an interval expected value model, which is the following interval linear programming problem after applying the interval expected value for each of the uncertainties.

$$\left. \begin{aligned} \min & 2x_1 + 3x_2 \\ \text{s.t.} & [1.6667, 2.5]x_1 + 6.1667x_2 \geq [169.6667, 180], \\ & 3x_1 + 3x_2 \geq 161.75, \\ & x_1 + x_2 \leq 100, \\ & x_1, x_2 \geq 0. \end{aligned} \right\} \quad (3.22)$$

The model (3.22) is also not well-defined. However, the bound on the objective value is $[\$117.35, \$127.87]$ with an optimal solution $x_1 \in [33.89, 44.41]$, and $x_2 \in [9.51, 20.03]$, which are obtained by solving four LP problems generated by the endpoints of the first constraint of (3.22). We can see for example that ‘1’ and ‘3’ as the realizations of \widehat{a}_{11} are not in $[1.6667, 2.5]$. It is also clear that

$$\widehat{a}_{11} = \begin{cases} 1 & \text{with } Pr(\{1\}) = \frac{1}{4} \in [Nec(\{1\}), Pos(\{1\})] = [\underline{f}(1), \overline{f}(1)] = [0, \frac{1}{3}], \\ 2 & \text{with } Pr(\{2\}) = \frac{1}{2} \in [Nec(\{2\}), Pos(\{2\})] = [\underline{f}(2), \overline{f}(2)] = [\frac{1}{2}, \frac{2}{3}], \\ 3 & \text{with } Pr(\{3\}) = \frac{1}{4} \in [Nec(\{3\}), Pos(\{3\})] = [\underline{f}(3), \overline{f}(3)] = [0, \frac{1}{2}], \end{cases}$$

$$\widehat{b}_1 = \begin{cases} 149 & \text{with } Pr(\{149\}) = \frac{5}{12} \in [\frac{1}{3}, \frac{1}{2}], \\ 180 & \text{with } Pr(\{180\}) = \frac{1}{3}, \\ 211 & \text{with } Pr(\{211\}) = \frac{1}{4} \in [\frac{1}{6}, \frac{1}{3}], \end{cases}$$

but in the long run, an optimal solution based on this information (see Subsection 3.2.1), is $x_1 = 31.80 \notin [33.89, 44.41]$, and $x_2 = 29.87 \notin [9.51, 20.03]$.

Therefore, the interval solution for (3.22) is not a reasonable representation of (3.21) in the long run. The approaches developed in the next chapter, not only correct this problem, they are more consistent with the bounds generated by uncertainties.

Table 3.2: Computational results for different models of the production planning problem.

Types of models, depending upon uncertainty interpretations	Raw materials		Production costs
	x_1	x_2	
1. Deterministic model	37.84	16.08	\$123.92
2. Expected recourse model	31.80	29.87	\$155.38
3. Minimax regret model	24.33	37.33	\$160.65 with $R = \$65.91$
4. Expected average model	45.63	16.04	\$139.37
5. Expected recourse-average model	45.63	16.04	\$139.37
6. Interval expected value model	[33.89, 44.41]	[9.51, 20.03]	[\$117.35, \$127.87]

These model examples in Section 3.1 - 3.6 illustrate the variety of approaches that are available. What is important, of course, are the semantics of the problems and data availability. The solution results of all the example uncertainty models are summarized in Table 3.2. At this point, some observations about the solutions and their meaning (semantics) are in order.

If we do not count an interval expected value model (which is not a reasonable model to represent an LP problem with mixed uncertainty), there is no

research on an approach dealing with different types of uncertainty in one constraint. Furthermore, as the argument made at the end of Section 3.4, we do not see that an expected average model and an expected recourse-average model are reasonable representations for their associated uncertainty information. Therefore, we introduce three expected recourse based models: (1) optimistic expected recourse model, (2) pessimistic expected recourse model, and (3) minimax regret problem of all expected recourse models, with the help of Theorems 2.17, and 2.25. We develop the theories and illustrate some applications of linear optimizations under generalized uncertainty in Chapters 4 and 5.

4. Linear optimization under uncertainty: pessimistic, optimistic, and minimax regret approaches

Stochastic optimization is a well-known approach for mathematical programs with uncertainty. It is suitable when uncertainties are random parameters with known probability distributions. The uncertainty with a known probability distribution is called a *probability* interpretation of uncertainty. However, we may not have complete information on these random parameters sufficient to model them as probability distributions. This dissertation considers linear programs with PC-BRIN uncertainty interpretations. Consider a general form of a linear problem with generalized uncertainty (1.2). Let us expand the uncertainty constraints of (1.2) out term-by-term as follows:

$$\left. \begin{aligned}
 & \min_x c_1x_1 + c_2x_2 + \dots + c_nx_n \\
 & \text{s.t. } \widehat{a}_{11}x_1 + \widehat{a}_{12}x_2 + \dots + \widehat{a}_{1n}x_n \geq \widehat{b}_1, \\
 & \quad \widehat{a}_{21}x_1 + \widehat{a}_{22}x_2 + \dots + \widehat{a}_{2n}x_n \geq \widehat{b}_2, \\
 & \quad \vdots \\
 & \quad \widehat{a}_{m1}x_1 + \widehat{a}_{m2}x_2 + \dots + \widehat{a}_{mn}x_n \geq \widehat{b}_m, \\
 & \quad Bx \geq 0, \\
 & \quad x \geq 0,
 \end{aligned} \right\} \quad (4.1)$$

where \widehat{a}_{ij} (and \widehat{b}_i), $i = 1, 2, \dots, m$ and $j = 1, 2, \dots, n$, can be given as one of PC-BRIN uncertainty interpretations. It could be that more than one interpretation of uncertainty appears in one constraint. We introduce a new variable z when there are uncertainties in an objective coefficient. Then, we add a constraint

$\widehat{c}_1x_1+\widehat{c}_2x_2+\dots+\widehat{c}_nx_n = z$ to the problem (4.1) and change the objective function to $\min_{x,z} z$. It is also assumed that all uncertain parameters in the constraints of the problem are independent. We define independence of any two uncertain parameters in Definition 4.1

Definition 4.1 *Let \widehat{u} and \widehat{v} be two different uncertain parameters with any interpretation in PC-BRIN interpretation (not necessarily the same). The sets of all realizations of \widehat{u} and \widehat{v} are $U = \{u_1, u_2, \dots, u_{l_1}; \exists l_1 \in \mathbb{N}\}$, and $V = \{v_1, v_2, \dots, v_{l_2}; \exists l_2 \in \mathbb{N}\}$, respectively. Furthermore, let \mathcal{M}_U and \mathcal{M}_V be the set of all probability density mass functions generated from the uncertainty interpretation of \widehat{u} and \widehat{v} , respectively. We call \widehat{u} (or \widehat{v}) a random variable associated with $f_{\widehat{u}} \in \mathcal{M}_U$ (or $f_{\widehat{v}} \in \mathcal{M}_V$). Let $A \subseteq U$, $B \subseteq V$, $f_{\widehat{u}} \in \mathcal{M}_U$ and $f_{\widehat{v}} \in \mathcal{M}_V$. We define $Pr_{\widehat{u},\widehat{v}}(\widehat{u} \in A, \widehat{v} \in B)$ as a **joint probability** of \widehat{u} and \widehat{v} with respect to $f_{\widehat{u}} \in \mathcal{M}_U$ and $f_{\widehat{v}} \in \mathcal{M}_V$. Then \widehat{u} and \widehat{v} are **independent** if and only if for every $A \subseteq U$ and $B \subseteq V$,*

$$Pr_{\widehat{u},\widehat{v}}(\widehat{u} \in A, \widehat{v} \in B) = Pr_{\widehat{u}}(A) \cdot Pr_{\widehat{v}}(B) = \sum_{u \in A} f_{\widehat{u}}(u) \cdot \sum_{v \in B} f_{\widehat{v}}(v); \forall f_{\widehat{u}} \in \mathcal{M}_U, f_{\widehat{v}} \in \mathcal{M}_V.$$

The problem (4.1) is not a well defined problem due to the unknown uncertainties. Three solution approaches are developed. The three deterministic problems associated with (4.1) are:

1. the *optimistic expected recourse problem*, which is the problem that provides the minimum of the expected recourse values for every possible distribution instantiation of the given parameters,
2. the *pessimistic expected recourse problem*, which is the problem that provides the maximum of the expected recourse values for every possible distribution instantiation of the given parameters, and

3. the *minimax regret problem of all expected recourse models*, which is the problem that provides the minimum of the maximum regret due to not knowing the actual probability to establish an expected recourse model.

We discuss each of these problems in Sections 4.1 and 4.2.

4.1 Pessimistic and optimistic expected recourse problems

We explain the problem (4.1) as a refinery production plan problem that has finite realizations of uncertainties. The first priority of the refinery is to satisfy the demands of its clients while minimizing its productivity cost. So, the amount of any shortage has to be bought from a market, incurring penalty (recourse) costs to the refinery. We suppose that penalty costs are fixed throughout the dissertation. We will not consider the storage cost or the selling cost of the excess products at this moment.

We first assume that each of \widehat{a}_{ij} has α_{ij} realizations, $a_{ij}^1, a_{ij}^2, \dots, a_{ij}^{\alpha_{ij}}$, with probability density mass values $g_{ij}(a_{ij}^1), g_{ij}(a_{ij}^2), \dots, g_{ij}(a_{ij}^{\alpha_{ij}})$, and each of \widehat{b}_i has β_i realizations, $b_i^1, b_i^2, \dots, b_i^{\beta_i}$, with probability density mass values $h_i(b_i^1), h_i(b_i^2), \dots, h_i(b_i^{\beta_i})$, where $\sum_{k=1}^{\alpha_{ij}} g_{ij}(a_{ij}^k) = 1$, and $\sum_{k=1}^{\beta_i} h_i(b_i^k) = 1$. Instead of random variables \widehat{a}_{ij} and \widehat{b}_i , $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, we define the random vector $\widehat{\xi} = (\widehat{a}_{11}, \widehat{a}_{12}, \dots, \widehat{a}_{mn}, \widehat{b}_1, \widehat{b}_2, \dots, \widehat{b}_m)$. This random vector $\widehat{\xi}$ creates $\prod_{j=1}^n \prod_{i=1}^m \alpha_{ij} \cdot \beta_i = N$ scenarios for constraints of (4.1). We introduce a recourse variable $y_i^k = \max \{0, b_i(\xi^k) - a_{i1}(\xi^k)x_1 - a_{i2}(\xi^k)x_2 - \dots - a_{in}(\xi^k)x_n\}$ for each of the i^{th} constraint in (4.1) and each k^{th} scenario, $i = 1, 2, \dots, m$, $k = 1, 2, \dots, N$. The recourse variable measures the corresponding shortage in production if there is any. Assume further that the positive penalty price per unit

of shortages is the fixed value s_i for each i^{th} constraint of (4.1), i.e., the penalty price s_1 corresponds with recourse variables $y_1^1, y_1^2, \dots, y_1^N$, the penalty price s_2 corresponds with recourse variables $y_2^1, y_2^2, \dots, y_2^N$, and so on. Hence, the expected recourse problem of (4.1) when all uncertainties have finite discrete probabilities is to minimize the sum of the production (first-stage) costs and the expected recourse costs and can be formulated as the following problem where (ξ^k, pr_k) (corresponding to y^k) contains a realization of the k^{th} scenario and its joint probability density mass value pr_k such that $\sum_{k=1}^N pr_k = 1$:

$$\left. \begin{aligned}
 & \min c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \sum_{k=1}^N pr_k [s_1 y_1^k + s_2 y_2^k + \dots + s_m y_m^k] \\
 & \text{s.t. } a_{11}(\xi^k) x_1 + a_{12}(\xi^k) x_2 + \dots + a_{1n}(\xi^k) x_n + y_1^k \geq b_1(\xi^k), \quad \forall k \\
 & \quad a_{21}(\xi^k) x_1 + a_{22}(\xi^k) x_2 + \dots + a_{2n}(\xi^k) x_n + y_2^k \geq b_2(\xi^k), \quad \forall k \\
 & \quad \vdots \\
 & \quad a_{m1}(\xi^k) x_1 + a_{m2}(\xi^k) x_2 + \dots + a_{mn}(\xi^k) x_n + y_m^k \geq b_m(\xi^k), \quad \forall k \\
 & \quad Bx \geq d, \quad x, y \geq 0.
 \end{aligned} \right\} (4.2)$$

We rewrite the problem (4.2) by keeping recourse variables on one side and moving the rest to the other side of the inequalities, i.e.,

$$\left. \begin{aligned}
 & \min c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \sum_{k=1}^N pr_k [s_1 y_1^k + s_2 y_2^k + \dots + s_m y_m^k] \\
 & \text{s.t. } y_1^k \geq b_1(\xi^k) - a_{11}(\xi^k) x_1 - a_{12}(\xi^k) x_2 - \dots - a_{1n}(\xi^k) x_n, \quad \forall k \\
 & \quad y_2^k \geq b_2(\xi^k) - a_{21}(\xi^k) x_1 - a_{22}(\xi^k) x_2 - \dots - a_{2n}(\xi^k) x_n, \quad \forall k \\
 & \quad \vdots \\
 & \quad y_m^k \geq b_m(\xi^k) - a_{m1}(\xi^k) x_1 - a_{m2}(\xi^k) x_2 - \dots - a_{mn}(\xi^k) x_n, \quad \forall k \\
 & \quad Bx \geq d, \quad x, y \geq 0.
 \end{aligned} \right\} (4.3)$$

Define $K_i = \prod_{j=1}^n \alpha_{ij} \cdot \beta_i$ to be the number of scenarios with respect to the i^{th} constraint of (4.1), for each $i = 1, 2, \dots, m$. According to constraints in (4.3), there are $M_i = \prod_{t=1, t \neq i}^m \prod_{j=1}^n \alpha_{tj} \cdot \beta_t$ constraints from

$$y_i^k \geq b_i(\xi^k) - a_{i1}(\xi^k)x_1 - a_{i2}(\xi^k)x_2 - \dots - a_{in}(\xi^k)x_n, \quad k = 1, 2, \dots, N, \quad (4.4)$$

that have the same scenario with respect to the i^{th} constraint of (4.1). Thus there are K_i groups of (4.4), each containing M_i constraints with the same scenario of the i^{th} constraint, i.e., $K_i M_i = N$, for each $i = 1, 2, \dots, m$. More precisely, suppose that $y_i^{k_1}, y_i^{k_2}, \dots, y_i^{k_{M_i-1}}$ and $y_i^{k_{M_i}}$ have the same scenario with respect to the i^{th} constraint, when $b_i^l, a_{i1}^{l_1}, a_{i2}^{l_2}, \dots, a_{in}^{l_n}$ are realizations of $\widehat{b}_i, \widehat{a}_{i1}, \widehat{a}_{i2}, \dots, \widehat{a}_{in}$, respectively. Therefore,

$$\begin{aligned} y_i^{k_1} &\geq b_i^l - a_{i1}^{l_1}x_1 - a_{i2}^{l_2}x_2 - \dots - a_{in}^{l_n}x_n, \\ y_i^{k_2} &\geq b_i^l - a_{i1}^{l_1}x_1 - a_{i2}^{l_2}x_2 - \dots - a_{in}^{l_n}x_n, \\ &\vdots \\ y_i^{k_{M_i}} &\geq b_i^l - a_{i1}^{l_1}x_1 - a_{i2}^{l_2}x_2 - \dots - a_{in}^{l_n}x_n. \end{aligned}$$

However, since $y_i^k = \max\{0, b_i^l - a_{i1}^{l_1}x_1 - a_{i2}^{l_2}x_2 - \dots - a_{in}^{l_n}x_n\}$ for each $k = k_1, k_2, \dots, k_{M_i}$, it turns out that $y_i^{k_1} = y_i^{k_2} = \dots = y_i^{k_{M_i}}$. We define $w_i^L = y_i^{k_1}$, where L is the scenario $(l, l_1, l_2, \dots, l_n)$. The joint probability for w_i^L is f_i^L , where

$$f_i^L = pr_{k_1} + pr_{k_2} + \dots + pr_{k_{M_i}} = h_i(b_i^l)g_{i1}(a_{i1}^{l_1})g_{i2}(a_{i2}^{l_2}) \dots g_{in}(a_{in}^{l_n}).$$

We provide a small numerical example here for a better understanding of this paragraph.

There are 4 scenarios with respect to constraint (4.5), i.e., $(a_{11}, b_1) = (1, 170)$, $(2, 170)$, $(1, 180)$ and $(2, 180)$. We define $w_1^1 = y_1^1 = y_1^2 = y_1^3 = y_1^4$, $w_1^2 = y_1^5 = y_1^6 = y_1^7 = y_1^8$, $w_1^3 = y_1^9 = y_1^{10} = y_1^{11} = y_1^{12}$, and $w_1^4 = y_1^{13} = y_1^{14} = y_1^{15} = y_1^{16}$. The joint probabilities of w_1^1, w_1^2, w_1^3 and w_1^4 are

$$f_1^1 = \binom{1111}{3624} + \binom{1113}{3624} + \binom{1111}{3624} + \binom{1113}{3624} = \frac{1}{3} \cdot \frac{1}{6} = \frac{1}{18},$$

$$f_1^2 = \binom{1511}{3624} + \binom{1513}{3624} + \binom{1511}{3624} + \binom{1513}{3624} = \frac{1}{3} \cdot \frac{5}{6} = \frac{5}{18},$$

$$f_1^3 = \binom{2111}{3624} + \binom{2113}{3624} + \binom{2111}{3624} + \binom{2113}{3624} = \frac{2}{3} \cdot \frac{1}{6} = \frac{2}{18},$$

$$f_1^4 = \binom{2511}{3624} + \binom{2513}{3624} + \binom{2511}{3624} + \binom{2513}{3624} = \frac{2}{3} \cdot \frac{5}{6} = \frac{10}{18}.$$

Similarly, we define $w_2^1 = y_2^1 = y_2^5 = y_2^9 = y_2^{13}$, $w_2^2 = y_2^2 = y_2^6 = y_2^{10} = y_2^{14}$, $w_2^3 = y_2^3 = y_2^7 = y_2^{11} = y_2^{15}$, and $w_2^4 = y_2^4 = y_2^8 = y_2^{12} = y_2^{16}$. The joint probabilities of w_2^1, w_2^2, w_2^3 , and w_2^4 are $f_2^1 = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$, $f_2^2 = \frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}$, $f_2^3 = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$, and $f_2^4 = \frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}$, respectively. We can transform the problem (4.7) to

$$\left. \begin{array}{l} \min \ 2x_1 + 3x_2 + 5 \sum_{L=1}^4 f_1^L w_1^L + 7 \sum_{L=1}^4 f_2^L w_2^L \\ \text{s.t.} \quad w_1^1 \geq 170 - 1x_1 - 6x_2, \quad w_2^1 \geq 160 - 3x_1 - 2x_2, \\ \quad \quad w_1^2 \geq 170 - 2x_1 - 6x_2, \quad w_2^2 \geq 160 - 3x_1 - 3x_2, \\ \quad \quad w_1^3 \geq 180 - 1x_1 - 6x_2, \quad w_2^3 \geq 162 - 3x_1 - 2x_2, \\ \quad \quad w_1^4 \geq 180 - 2x_1 - 6x_2, \quad w_2^4 \geq 162 - 3x_1 - 3x_2, \quad x, w \geq 0. \quad \diamond \end{array} \right\} (4.8)$$

The transformation of (4.5) and (4.6) into (4.8) indicates how, in general, the problem (4.3) can be remodeled. The reduced expected recourse model is:

$$\left. \begin{aligned}
& \min_{x \in \Psi} c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \\
& \quad s_1 \sum_{L=1}^{K_1} f_1^L w_1^L + s_2 \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + s_m \sum_{L=1}^{K_m} f_m^L w_m^L \\
& \text{s.t. } w_1^1 \geq b_1^1 - a_{11}^1 x_1 - a_{12}^1 x_2 - \dots - a_{1n}^1 x_n, \\
& \quad \vdots \\
& \quad w_1^{K_1} \geq b_1^{\beta_1} - a_{11}^{\alpha_{11}} x_1 - a_{12}^{\alpha_{12}} x_2 - \dots - a_{1n}^{\alpha_{1n}} x_n, \\
& \quad \vdots \\
& \quad w_m^1 \geq b_m^1 - a_{m1}^1 x_1 - a_{m2}^1 x_2 - \dots - a_{mn}^1 x_n, \\
& \quad \vdots \\
& \quad w_m^{K_m} \geq b_m^{\beta_m} - a_{m1}^{\alpha_{m1}} x_1 - a_{m2}^{\alpha_{m2}} x_2 - \dots - a_{mn}^{\alpha_{mn}} x_n, \quad w \geq 0,
\end{aligned} \right\} (4.9)$$

where $\sum_{L=1}^{K_i} f_i^L = 1$ for each $i = 1, 2, \dots, m$. The original expected recourse model (4.3) has n components of variable x and mN terms of the recourse variable y . Thus, it has the total of $mN + n$ unknown variables and mN constraints (with respected to the uncertainties). On the other hand, the reduced expected recourse model (4.9) has $n + \sum_{i=1}^m K_i$ unknown variables with $\sum_{i=1}^m K_i$ constraints (with respected to the uncertainties). Since $M_i \geq 2, \forall i = 1, 2, \dots, m, mN = K_1 M_1 + K_2 M_2 + \dots + K_m M_m \geq 2 \sum_{i=1}^m K_i$. Hence, (4.9) has at least $\sum_{i=1}^m K_i$ less variables and constraints than (4.3). The remodeled problem (4.9) reduces the number of operations needed in each iteration of the simplex method compared to the original expected recourse problem (4.3), in general, because the size of the coefficient matrix of (4.9) is smaller than that of (4.3).

Uncertainties in the problem (4.1) also can have other interpretations in PC-BRIN besides probability. Thus, instead of a single joint probability distribution $f_i^L, L = 1, 2, \dots, K_i, i = 1, 2, \dots, m$, as in the expected recourse problem (4.9),

we have a set of joint probabilities.

4.1.1 Pessimistic and optimistic expected recourse problems generated by random sets with finite realizations

We consider any uncertain information in Figure 2.7 that can be interpreted as a random set. Suppose that uncertainties \widehat{a}_{ij} and \widehat{b}_i ; $\forall i = 1, 2, \dots, m, \forall j = 1, 2, \dots, n$, can be interpreted as random sets. Let $\Theta_{a_{ij}} = \{a_{ij}^1, a_{ij}^2, \dots, a_{ij}^{\alpha_{ij}}\}$ be the set of all realizations of \widehat{a}_{ij} , and $\Theta_{b_i} = \{b_i^1, b_i^2, \dots, b_i^{\beta_i}\}$ be the set of all realizations of \widehat{b}_i . Then, the sets of all probability density mass functions satisfying the random set information of \widehat{a}_{ij} and \widehat{b}_i are $\mathcal{M}_{a_{ij}}$ and \mathcal{M}_{b_i} , respectively, where

$$\mathcal{M}_{a_{ij}} = \left\{ g_{ij} : \Theta_{a_{ij}} \rightarrow [0, 1] \mid \sum_{a_{ij}^k \in A} g_{ij}(a_{ij}^k) \in [Bel_{ij}(A), Pl_{ij}(A)], \forall A \subseteq \Theta_{a_{ij}} \right\},$$

$$\mathcal{M}_{b_i} = \left\{ h_i : \Theta_{b_i} \rightarrow [0, 1] \mid \sum_{b_i^k \in A} h_i(b_i^k) \in [Bel_i(A), Pl_i(A)], \forall A \subseteq \Theta_{b_i} \right\}.$$

Therefore, the term f_i , $\forall i = 1, 2, \dots, m$, in the objective function of (4.9) is an element of the set \mathcal{M}_i of all joint probabilities satisfying random set information on the i^{th} constraint of (4.1), where

$$\mathcal{M}_i = \left\{ f_i : \Theta_{b_i} \times \prod_{j=1}^n \Theta_{a_{ij}} \rightarrow [0, 1] \mid \sum_{L=1}^{K_i} f_i^L = 1, f_i^L = h_i^L g_{i1}^L g_{i2}^L \dots g_{in}^L \right\},$$

$h_i \in \mathcal{M}_{b_i}$, $g_{ij} \in \mathcal{M}_{a_{ij}}$, $K_i = \prod_{j=1}^n \alpha_{ij} \beta_i$ is the number of scenarios with respect to the i^{th} constraint of (4.1), $h_i^L = h_i(b_i^{k_1})$, $\exists b_i^{k_1} \in \Theta_{b_i}$, and for each j , $g_{ij}^L = g_{ij}(a_{ij}^{k_2})$, $\exists a_{ij}^{k_2} \in \Theta_{a_{ij}}$.

Example 13 illustrates $\mathcal{M}_{a_{ij}}$, \mathcal{M}_{b_i} , and \mathcal{M}_i using a small set of realizations.

Example 13. Suppose that uncertainties \widehat{a}_{11} and \widehat{b}_1 in constraint (4.5) of Example 12 do not have probability interpretations. Instead, two realizations, 1 and 2, of \widehat{a}_{11} and other two realizations, 170 and 180, of \widehat{b}_1 have random set information as follows: $m(\{1\}) = \frac{1}{2}, m(\{1, 2\}) = \frac{1}{2}, m(\{170\}) = \frac{1}{3}$, and $m(\{180\}) = \frac{2}{3}$. So,

$$\begin{aligned}\mathcal{M}_{a_{11}} &= \left\{ g_{11} \mid g_{11}(1) + g_{11}(2) = 1, g_{11}(1) \in \left[\frac{1}{2}, 1 \right], g_{11}(2) \in \left[0, \frac{1}{2} \right] \right\}, \\ \mathcal{M}_{b_1} &= \left\{ h_1 \mid h_1(170) + h_1(180) = 1, h_1(170) \in \left[\frac{1}{3}, 1 \right], h_1(180) \in \left[0, \frac{2}{3} \right] \right\}.\end{aligned}$$

There are $K_1 = 4$ scenarios for constraint (4.5): 1) $1x_1 + 6x_2 \geq 170$, 2) $1x_1 + 6x_2 \geq 180$, 3) $2x_1 + 6x_2 \geq 170$, and 4) $2x_1 + 6x_2 \geq 180$. Thus, f_1 in the objective function of (4.8) is

$$\mathcal{M}_1 = \left\{ f_1 \mid \sum_{L=1}^4 f_1^L = 1, f_1^L = h_1^L g_{11}^L, \text{ s.t. } h_1 \in \mathcal{M}_{b_1}, g_{11} \in \mathcal{M}_{a_{11}} \right\}.$$

- When $L = 1$, $g_{11}^L = g_{11}(1)$, and $h_1^L = h_1(170)$.
- When $L = 2$, $g_{11}^L = g_{11}(1)$, and $h_1^L = h_1(180)$.
- When $L = 3$, $g_{11}^L = g_{11}(2)$, and $h_1^L = h_1(170)$.
- When $L = 4$, $g_{11}^L = g_{11}(2)$, and $h_1^L = h_1(180)$.

Suppose $g_{11}^1 = g_{11}^2 = \frac{2}{3}$, and $h_1^1 = h_1^3 = \frac{2}{3}$, then $g_{11}^3 = g_{11}^4 = \frac{1}{3}$, and $h_1^2 = h_1^4 = \frac{1}{3}$. Hence, $f_1 \in \mathcal{M}_1$, where $f_1^1 = \frac{4}{9}, f_1^2 = \frac{2}{9}, f_1^3 = \frac{2}{9}$, and $f_1^4 = \frac{1}{9}$. \diamond

Random sets generate the set \mathcal{M}_i of joint probabilities for each i^{th} constraint of (4.1). The uncertainty interpretations that can be presented as random sets are listed in Figure 2.7. Without any further information about

these uncertainties, we define Ξ as the feasible set of the constraints in (4.9), and $\mathcal{M} := \{f = (f_1, f_2, \dots, f_m) \mid f_i \in \mathcal{M}_i, i = 1, 2, \dots, m\}$. Two important approaches are the pessimistic and optimistic ones for the expected recourse problems (4.9), $\forall f \in \mathcal{M}$. The recourse problem for an optimistic solution is

$$\min_{f \in \mathcal{M}} \min_{(x,w) \in \Xi} \left(\begin{array}{l} c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \\ s_1 \sum_{L=1}^{K_1} f_1^L w_1^L + s_2 \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + s_m \sum_{L=1}^{K_m} f_m^L w_m^L \end{array} \right), \quad (4.10)$$

and the recourse problem for a pessimistic solution is

$$\max_{f \in \mathcal{M}} \min_{(x,w) \in \Xi} \left(\begin{array}{l} c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \\ s_1 \sum_{L=1}^{K_1} f_1^L w_1^L + s_2 \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + s_m \sum_{L=1}^{K_m} f_m^L w_m^L \end{array} \right). \quad (4.11)$$

We refer to the problems (4.10) and (4.11) as the optimistic and pessimistic expected recourse problems, respectively.

Theorem 4.2 *Assume that all uncertainties are random sets. Let $s = [s_1, s_2, \dots, s_m]^T$, $f = (f_1, f_2, \dots, f_m) \in \mathcal{M}$, $f_i = [f_i^1, f_i^2, \dots, f_i^{K_i}]$ such that $\sum_{L=1}^{K_i} f_i^L = 1$, and $w_i = [w_i^1, w_i^2, \dots, w_i^{K_i}]^T$, $i = 1, 2, \dots, m$. Suppose that*

$$\min_{(x,w) \in \Xi} \left(\begin{array}{l} c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \\ s_1 \sum_{L=1}^{K_1} f_1^L w_1^L + s_2 \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + s_m \sum_{L=1}^{K_m} f_m^L w_m^L \end{array} \right) \quad (4.12)$$

is bounded for each $f \in \mathcal{M}$. Then, there exist \underline{f} and $\bar{f} \in \mathcal{M}$, which can be constructed, such that

$$\min_{f \in \mathcal{M}} \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} f_1 w_1 \\ \vdots \\ f_m w_m \end{pmatrix} = \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} \underline{f}_1 w_1 \\ \vdots \\ \underline{f}_m w_m \end{pmatrix}, \quad (4.13)$$

and

$$\max_{f \in \mathcal{M}} \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} f_1 w_1 \\ \vdots \\ f_m w_m \end{pmatrix} = \min_{x,w \in \Xi} c^T x + s^T \begin{pmatrix} \bar{f}_1 w_1 \\ \vdots \\ \bar{f}_m w_m \end{pmatrix}. \quad (4.14)$$

Proof: First, we rearrange the realizations $1, 2, \dots, \beta_i$ of the uncertainty \widehat{b}_i so that $b_i^1 \leq b_i^2 \leq \dots \leq b_i^{\beta_i}; \forall i = 1, 2, \dots, m$. Similarly, the realizations $1, 2, \dots, \alpha_{ij}$ of $-\widehat{a}_{ij}$ are rearranged to $-a_{ij}^1 \leq -a_{ij}^2 \leq \dots \leq -a_{ij}^{\alpha_{ij}}; \forall i = 1, 2, \dots, m, \forall j = 1, 2, \dots, n$. Then we use the construction (2.42) to define $\underline{h}_i(b_i^1), \underline{h}_i(b_i^2), \dots, \underline{h}_i(b_i^{\beta_i})$, and $\underline{g}_{ij}(a_{ij}^1), \underline{g}_{ij}(a_{ij}^2), \dots, \underline{g}_{ij}(a_{ij}^{\alpha_{ij}})$. We do the same for \bar{g}_{ij} and \bar{h}_i by using the construction (2.43). Therefore, we construct \underline{f} and \bar{f} as follows.

- $\underline{f} = (\underline{f}_1, \underline{f}_2, \dots, \underline{f}_m)$, where for each $i = 1, 2, \dots, m$, $\underline{f}_i = [\underline{f}_i^1, \underline{f}_i^2, \dots, \underline{f}_i^{K_i}]$ is the joint density function vector generated by the redefined \underline{h}_i , and \underline{g}_{ij} .
- $\bar{f} = (\bar{f}_1, \bar{f}_2, \dots, \bar{f}_m)$, where for each $i = 1, 2, \dots, m$, $\bar{f}_i = [\bar{f}_i^1, \bar{f}_i^2, \dots, \bar{f}_i^{K_i}]$ is the joint density function vector generated by the redefined \bar{h}_i , and \bar{g}_{ij} .

Since $\underline{h}_i, \bar{h}_i \in \mathcal{M}_{b_i}$, and $\underline{g}_{ij}, \bar{g}_{ij} \in \mathcal{M}_{a_{ij}}$, we have that $\underline{f}, \bar{f} \in \mathcal{M}$.

Let x_1, x_2, \dots, x_n be given. Let $\widehat{\theta}_i = \widehat{b}_i - \widehat{a}_{i1}x_1 - \widehat{a}_{i2}x_2 - \dots - \widehat{a}_{in}x_n$ be an uncertainty with the set of realizations $\{\theta_i^1, \theta_i^2, \dots, \theta_i^{K_i}\}$, where $\theta_i^L = b_i^L - a_{i1}^Lx_1 - a_{i2}^Lx_2 - \dots - a_{in}^Lx_n$. Since the lowest expected value of \widehat{b}_i is $E_{\underline{h}_i}(\widehat{b}_i)$, the lowest expected value of $-\widehat{a}_{ij}$ is $E_{\underline{g}_{ij}}(\widehat{a}_{ij})$, for each $j = 1, 2, \dots, n$, and $\underline{f} = \underline{h}_i \underline{g}_{i1} \underline{g}_{i2} \dots \underline{g}_{in}$, it follows directly from Theorem 2.17, and the linear property of an expected value function, that the lowest expected value of $\widehat{\theta}_i$ is

$$\inf_{f \in \mathcal{M}} \left\{ E[\widehat{\theta}_i] \right\} = \inf_{f \in \mathcal{M}} \left\{ E[\widehat{b}_i - \widehat{a}_{i1}x_1 - \widehat{a}_{i2}x_2 - \dots - \widehat{a}_{in}x_n] \right\}$$

$$\begin{aligned}
\inf_{f \in \mathcal{M}} \left\{ E[\widehat{\theta}_i] \right\} &= \inf_{f \in \mathcal{M}} \left\{ E_{h_i}[\widehat{b}_i] + E_{g_{i1}}[-\widehat{a}_{i1}]x_1 + E_{g_{i2}}[-\widehat{a}_{i2}]x_2 + \dots + E_{g_{in}}[-\widehat{a}_{in}]x_n \right\} \\
&\quad \text{where } h_i \in \mathcal{M}_{b_i} \text{ and } g_{ij} \in \mathcal{M}_{a_{ij}}, \quad j = 1, 2, \dots, n \\
&= E_{h_i}[\widehat{b}_i] + E_{g_{i1}}[-\widehat{a}_{i1}]x_1 + E_{g_{i2}}[-\widehat{a}_{i2}]x_2 + \dots + E_{g_{in}}[-\widehat{a}_{in}]x_n \\
&= E_f[\widehat{\theta}_i] \\
&= \sum_{L=1}^{K_i} \underline{f}_i^L \theta_i^L.
\end{aligned}$$

The largest expected value of $\widehat{\theta}_i$ can be obtained in similar pattern. Therefore, the expected value of θ_i with respect to any density function vector $f_i = [f_i^1, f_i^2, \dots, f_i^{K_i}] \in \mathcal{M}_i$ is

$$\sum_{L=1}^{K_i} \underline{f}_i^L \theta_i^L \leq \sum_{L=1}^{K_i} f_i^L \theta_i^L \leq \sum_{L=1}^{K_i} \overline{f}_i^L \theta_i^L. \quad (4.15)$$

We can also see that by using joint density function, we can rewrite this lowest expected value as

$$\begin{aligned}
\underline{E}(\widehat{\theta}_i) &= \sum_{\substack{\alpha_{ij} \\ k_j=1 \\ \forall j=1,2,\dots,n}} \left[\left(\prod_{j=1}^n g_{ij}^{k_j} \right) \left(\overbrace{h_i^1 (b_i^1 - a_{i1}^{k_1} x_1 - a_{i2}^{k_2} x_2 - \dots - a_{in}^{k_n} x_n)}^{\theta_i^{1,k_1,k_2,\dots,k_n}} \right) \right. \\
&\quad \left. + \dots + \overbrace{h_i^{\beta_i} (b_i^{\beta_i} - a_{i1}^{k_1} x_1 - a_{i2}^{k_2} x_2 - \dots - a_{in}^{k_n} x_n)}^{\theta_i^{\beta_i,k_1,k_2,\dots,k_n}} \right], \quad (4.16)
\end{aligned}$$

which means that the term $\underline{h}_i^1 \theta_i^{1,k_1,k_2,\dots,k_n} + \dots + \underline{h}_i^{\beta_i} \theta_i^{\beta_i,k_1,k_2,\dots,k_n}$ is the smallest, for each $k_j = 1, \dots, \alpha_{ij}$. Hence, the statement (4.15) also holds if we change θ_i^L to 0 whenever $\theta_i^L < 0$. Here is the reason. For the fixed scenario indexes k_1, k_2, \dots, k_n of uncertainties $\widehat{a}_{i1}, \widehat{a}_{i2}, \dots, \widehat{a}_{in}$, suppose that $b_i^k - a_{i1}^{k_1} x_1 - a_{i2}^{k_2} x_2 - \dots - a_{in}^{k_n} x_n < 0$, for some $k = 1, 2, \dots, \beta_i$. Then, it means $b_i^j - a_{i1}^{k_1} x_1 - a_{i2}^{k_2} x_2 - \dots - a_{in}^{k_n} x_n < 0$, for

all $j = 1, 2, \dots, k$, because $b_i^j \leq b_i^k$. Set these terms that less than zero to be zero.

We will have $\underbrace{0, 0, \dots, 0}_{k \text{ terms}}, \theta_i^{k+1, k_1, k_2, \dots, k_n}, \dots, \theta_i^{\beta_i, k_1, k_2, \dots, k_n}$, which is a nondecreasing order. Hence \underline{h}_i still provides the smallest value for

$$\underline{h}_i^1 \cdot 0 + \dots + \underline{h}_i^k \cdot 0 + \underline{h}_i^{k+1} \theta_i^{1, k_1, k_2, \dots, k_n} + \dots + \underline{h}_i^{\beta_i} \theta_i^{\beta_i, k_1, k_2, \dots, k_n}.$$

Therefore, by (4.16), the statement (4.15) holds if we change θ_i^L to 0 whenever it is less than 0.

Define $w_i^L = \max \{b_i^L - a_{i1}^L x_1 - a_{i2}^L x_2 - \dots - a_{in}^L x_n, 0\}$, $L = 1, 2, \dots, K_i$.

Thus,

$$\sum_{L=1}^{K_i} \underline{f}_i^L w_i^L \leq \sum_{L=1}^{K_i} f_i^L w_i^L \leq \sum_{L=1}^{K_i} \bar{f}_i^L w_i^L. \quad (4.17)$$

Let $(\underline{x}, \underline{w})$ and (\bar{x}, \bar{w}) be optimal solutions for the recourse problems on the right-hand-side of (4.13) and (4.14), respectively. Since both problems have the same set of constraints, these optimal solutions are feasible in both problems. Given $f \in \mathcal{M}$ and its corresponding optimal solution (x_f^*, w_f^*) , we have

$$\begin{aligned} & c^T \underline{x} + s_1 \sum_{k=1}^{K_1} \underline{f}_1^k \underline{w}_1^k + s_2 \sum_{k=1}^{K_2} \underline{f}_2^k \underline{w}_2^k + \dots + s_m \sum_{k=1}^{K_m} \underline{f}_m^k \underline{w}_m^k \\ & \leq c^T x^* + s_1 \sum_{k=1}^{K_1} f_1^k w_{f_1}^{*k} + s_2 \sum_{k=1}^{K_2} f_2^k w_{f_2}^{*k} + \dots + s_m \sum_{k=1}^{K_m} f_m^k w_{f_m}^{*k} \quad (\text{property of 'min'}) \\ & \leq c^T x^* + s_1 \sum_{k=1}^{K_1} f_1^k w_{f_1}^{*k} + s_2 \sum_{k=1}^{K_2} f_2^k w_{f_2}^{*k} + \dots + s_m \sum_{k=1}^{K_m} f_m^k w_{f_m}^{*k} \quad (\text{property of } \underline{f} \\ & \quad \text{in Theorem 2.17}) \\ & \leq c^T \bar{x} + s_1 \sum_{k=1}^{K_1} f_1^k \bar{w}_1^k + s_2 \sum_{k=1}^{K_2} f_2^k \bar{w}_2^k + \dots + s_m \sum_{k=1}^{K_m} f_m^k \bar{w}_m^k \quad (\text{property of 'max'}) \\ & \leq c^T \bar{x} + s_1 \sum_{k=1}^{K_1} \bar{f}_1^k \bar{w}_1^k + s_2 \sum_{k=1}^{K_2} \bar{f}_2^k \bar{w}_2^k + \dots + s_m \sum_{k=1}^{K_m} \bar{f}_m^k \bar{w}_m^k \quad (\text{property of } \bar{f} \\ & \quad \text{in Theorem 2.17}). \quad \square \end{aligned}$$

4.1.2 Pessimistic and optimistic expected recourse problems generated by IVPMs with finite realizations

Suppose now that uncertainties \widehat{a}_{ij} and \widehat{b}_i ; $\forall i = 1, 2, \dots, m$, and $\forall j = 1, 2, \dots, n$, can be interpreted as IVPMs. We use the same notation $\mathcal{M}_{a_{ij}}$ and \mathcal{M}_{b_i} for the sets of all probability measures satisfying the IVPM information of \widehat{a}_{ij} and \widehat{b}_i , respectively, where

$$\mathcal{M}_{a_{ij}} = \left\{ g_{ij} : \Theta_{a_{ij}} \rightarrow [0, 1] \mid \sum_{a_{ij}^k \in A} g_{ij}(a_{ij}^k) \in [L_{ij}(A), U_{ij}(A)], \forall A \subseteq \Theta_{a_{ij}} \right\},$$

and

$$\mathcal{M}_{b_i} = \left\{ h_i : \Theta_{b_i} \rightarrow [0, 1] \mid \sum_{b_i^k \in A} h_i(b_i^k) = 1 \in [L_i(A), U_i(A)], \forall A \subseteq \Theta_{b_i} \right\}.$$

$L_{ij}(A)$ and $U_{ij}(A)$ are the lower and upper bounds on probability g_{ij} of an event $A \subseteq \Theta_{a_{ij}}$. $0 \leq L_{ij}(A) \leq U_{ij}(A) \leq 1$. Similarly, $L_i(A)$ and $U_i(A)$ are the lower and upper bounds on probability h_i of an event $A \subseteq \Theta_{b_i}$. $0 \leq L_i(A) \leq U_i(A) \leq 1$. Then, using the same notation \mathcal{M}_i for the set of all joint probabilities satisfying IVPM information on the i^{th} constraint of (4.1), we also have

$$\mathcal{M}_i = \left\{ f_i : \Theta_{b_i} \times \prod_{j=1}^n \Theta_{a_{ij}} \rightarrow [0, 1] \mid \sum_{L=1}^{K_i} f_i^L = 1, f_i^L = h_i^L g_{i1}^L g_{i2}^L \dots g_{in}^L \right\},$$

Define $\mathcal{M} := \{f = (f_1, f_2, \dots, f_m) \mid f_i \in \mathcal{M}_i, i = 1, 2, \dots, m\}$ as in the last subsection, but keep in mind that information in this subsection is IVPM. We state a similar theorem to Theorem 4.2 as follows.

Theorem 4.3 *Assume that all uncertainties are IVPMs. Let $s = [s_1, s_2, \dots, s_m]^T$, $f = (f_1, f_2, \dots, f_m) \in \mathcal{M}$, $f_i = [f_i^1, f_i^2, \dots, f_i^{K_i}]$ such that $\sum_{L=1}^{K_i} f_i^L = 1$, and $w_i = [w_i^1, w_i^2, \dots, w_i^{K_i}]^T$, $i = 1, 2, \dots, m$. Suppose that*

$$\min_{(x,w) \in \Xi} \left(\begin{array}{l} c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \\ s_1 \sum_{L=1}^{K_1} f_1^L w_1^L + s_2 \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + s_m \sum_{L=1}^{K_m} f_m^L w_m^L \end{array} \right)$$

is bounded for each $f \in \mathcal{M}$. Then, there exist \underline{f} and $\bar{f} \in \mathcal{M}$, which can be constructed, such that

$$\min_{f \in \mathcal{M}} \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} f_1 w_1 \\ \vdots \\ f_m w_m \end{pmatrix} = \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} \underline{f}_1 w_1 \\ \vdots \\ \underline{f}_m w_m \end{pmatrix},$$

and

$$\max_{f \in \mathcal{M}} \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} f_1 w_1 \\ \vdots \\ f_m w_m \end{pmatrix} = \min_{x,w \in \Xi} c^T x + s^T \begin{pmatrix} \bar{f}_1 w_1 \\ \vdots \\ \bar{f}_m w_m \end{pmatrix}.$$

Proof: Similar to the proof of Theorem 4.2, but use Theorem 2.25 instead of Theorem 2.17 inside the proof. \square

Therefore, we can combine both Theorems 4.2 and 4.3 as Theorem 4.4, so that the new theorem is applicable for any uncertainty interpretation in Figure 2.7. Use the same notation \mathcal{M} to define the set of all joint probabilities. Hence, $\mathcal{M} := \{f = (f_1, f_2, \dots, f_m) \mid f_i \in \mathcal{M}_i, i = 1, 2, \dots, m\}$, where \mathcal{M}_i in this set \mathcal{M} is $\mathcal{M}_i = \left\{ f_i \mid \sum_{L=1}^{K_i} f_i^L = 1, f_i^L = h_i^L g_{i1}^L g_{i2}^L \dots g_{in}^L, \text{ s.t. } h_i \in \mathcal{M}_{b_i}, g_{ij} \in \mathcal{M}_{a_{ij}} \right\}$ such that \mathcal{M}_{b_i} and $\mathcal{M}_{a_{ij}}$ are generated by either random set or IVPm information. The proof of Theorem 4.4 is the combination of the proofs of Theorems 4.2 and 4.3.

Theorem 4.4 Assume that \hat{a}_{ij} and \hat{b}_i can have PC-BRIN interpretation. Let $s = [s_1, s_2, \dots, s_m]^T$, $f = (f_1, f_2, \dots, f_m) \in \mathcal{M}$, $f_i = [f_i^1, f_i^2, \dots, f_i^{K_i}]$ such that $\sum_{L=1}^{K_i} f_i^L = 1$, and $w_i = [w_i^1, w_i^2, \dots, w_i^{K_i}]^T$, $i = 1, 2, \dots, m$. Suppose that

$$\min_{(x,w) \in \Xi} \left(\begin{array}{l} c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \\ s_1 \sum_{L=1}^{K_1} f_1^L w_1^L + s_2 \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + s_m \sum_{L=1}^{K_m} f_m^L w_m^L \end{array} \right)$$

is bounded for each $f \in \mathcal{M}$. Then, there exist \underline{f} and $\bar{f} \in \mathcal{M}$, which can be constructed, such that

$$\min_{f \in \mathcal{M}} \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} f_1 w_1 \\ \vdots \\ f_m w_m \end{pmatrix} = \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} \underline{f}_1 w_1 \\ \vdots \\ \underline{f}_m w_m \end{pmatrix},$$

and

$$\max_{f \in \mathcal{M}} \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} f_1 w_1 \\ \vdots \\ f_m w_m \end{pmatrix} = \min_{(x,w) \in \Xi} c^T x + s^T \begin{pmatrix} \bar{f}_1 w_1 \\ \vdots \\ \bar{f}_m w_m \end{pmatrix}.$$

4.2 Minimax regret of the expected recourse problems

The uncertainty interpretations we have lead us to random sets and IVPMs, which provide only the bound on our unknown probability. Therefore, it is natural to find the lower and upper expected objective values of the uncertainty problem (4.1) when we reformulate the problem to two recourse problems based on two density functions provided in the last section. The extreme values of the pessimistic and optimistic solutions are useful to obtain an idea of the spread of the uncertainty. However, we are not often using the extremes. With a collection of probabilities, an optimal solution that minimizes the maximum regret is another reasonable result that we can provide to a user. This is a fruitful approach because its solution tells a user that when s/he does not know the exact probability, his/her expected objective value using this solution is the

best of the worst regret. We have the aim to solve the minimum of the maximum regret (the minimax regret, for short) of the expected recourse problems (4.9), with unknown probabilities of each uncertainty \widehat{a}_{ij} and \widehat{b}_i in the set $\mathcal{M}_{a_{ij}}$ and \mathcal{M}_{b_i} . The uncertainty interpretations of \widehat{a}_{ij} and \widehat{b}_i can be a random set or an IVP information.

Define the objective function of (4.9) as

$$z(f, x, w) := c^T x + s_1 \sum_{k=1}^{K_1} f_1^k w_1^k + s_2 \sum_{k=1}^{K_2} f_2^k w_2^k + \dots + s_m \sum_{k=1}^{K_m} f_m^k w_m^k.$$

We now consider the problem

$$\min_{(x,w) \in \Xi} z(f, x, w), \quad (4.18)$$

which has uncertain probability $f \in \mathcal{M}$. The regret with respect to a known probability f , which shows the amount that a candidate solution (x, w) deviates from the true objective value, can be expressed by

$$r(f, x, w) := \left(z(f, x, w) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right), \quad (4.19)$$

where, $\min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega)$ is the true solution with this known probability f . When the true probability f is unknown, the worst (maximum) regret of a candidate solution (x, w) is

$$R(x, w) := \max_{f \in \mathcal{M}} r(f, x, w). \quad (4.20)$$

The best of the worst regret over all candidate solutions is

$$\min_{(x,w) \in \Xi} R(x, w). \quad (4.21)$$

We rewrite the problem (4.21) as

$$\min_{(x,w) \in \Xi} \max_{f \in \mathcal{M}} \left(z(f, x, w) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right). \quad (4.22)$$

Since $-\min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) = \max_{(\chi, \omega) \in \Xi} (-z(f, \chi, \omega))$, the problem (4.22) becomes

$$\begin{aligned} & \min_{(x,w) \in \Xi} \max_{f \in \mathcal{M}} \left(z(f, x, w) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right) \\ &= \min_{(x,w) \in \Xi} \max_{f \in \mathcal{M}} \left(z(f, x, w) + \max_{(\chi, \omega) \in \Xi} (-z(f, \chi, \omega)) \right) \\ &= \min_{(x,w) \in \Xi} \max_{f \in \mathcal{M}} \left(\max_{(\chi, \omega) \in \Xi} (-z(f, \chi, \omega)) + z(f, x, w) \right) \\ &= \min_{(x,w) \in \Xi} \max_{f \in \mathcal{M}} \left(\max_{(\chi, \omega) \in \Xi} (-z(f, \chi, \omega) + z(f, x, w)) \right) \\ &= \min_{(x,w) \in \Xi} \max_{f \in \mathcal{M}} \left(\max_{(\chi, \omega) \in \Xi} (z(f, x, w) - z(f, \chi, \omega)) \right) \\ &= \min_{(x,w) \in \Xi} \max_{f \in \mathcal{M}} \max_{(\chi, \omega) \in \Xi} (z(f, x, w) - z(f, \chi, \omega)) \\ &= \min_{(x,w) \in \Xi} \max_{\substack{f \in \mathcal{M} \\ (\chi, \omega) \in \Xi}} (z(f, x, w) - z(f, \chi, \omega)). \end{aligned} \quad (4.23)$$

4.2.1 A relaxation procedure for minimax regret of the expected recourse problems

Literature on finding a minimax regret solution to a linear programming problem, whose objective coefficients are not certain in a closed convex set, appears when the linear program has interval objective coefficients. The paper [1] in 2009 surveys the recent work on this problem, and the latest found are in [21, 22, 45, 46], and [62] which are based on a relaxation procedure to solve a minimax problem of an LP with interval objective coefficients. The relaxation procedure is effective for a class of problems with a very large number of constraints. We can transform the minimax regret problem (4.23) into the following

optimization problem with an infinite number of constraints:

$$\left. \begin{aligned} & \min_{R, (x, w)} R \\ \text{s.t.} \quad & R \geq (z(f, x, w) - z(f, \chi, \omega)), \quad \forall f \in \mathcal{M} \text{ and } (\chi, \omega) \in \Xi \\ & (x, w) \in \Xi \\ & R \geq 0. \end{aligned} \right\} (4.24)$$

Instead of solving this problem (4.24), we solve a series of the following relaxed versions of the problem:

$$\left. \begin{aligned} & \min_{R, (x, w)} R \\ \text{s.t.} \quad & (x, w) \in \Xi \\ & R \geq 0 \\ & R \geq (z(f^{(i)}, x, w) - z(f^{(i)}, \chi^{(i)}, \omega^{(i)})); f^{(i)} \in \mathcal{M} \text{ and } (\chi^{(i)}, \omega^{(i)}) \in \Xi, \\ & \quad \quad \quad i = 1, 2, \dots, k. \end{aligned} \right\} (4.25)$$

Let $(x^{(k)}, w^{(k)}, R^{(k)})$ be optimal for the relaxed problem (4.25). If $(x^{(k)}, w^{(k)}, R^{(k)})$ is feasible for the problem (4.24), then it must be optimal for the problem (4.24). Hence, $(x^{(k)}, w^{(k)})$ is a minimax regret solution to the problem (4.23). On the other hand, if there exists at least one constraint from $R \geq (z(f, x, w) - z(f, \chi, \omega)), \forall f \in \mathcal{M} \text{ and } (\chi, \omega) \in \Xi$ violates $(x^{(k)}, w^{(k)}, R^{(k)})$, then the most violated constraint will be added to the constraints of the relaxed problem. Next, solve the updated problem setting $k := k + 1$ and repeating the process. We can check which constraint is the most violated one by

1. if $\max_{\substack{f \in \mathcal{M} \\ (\chi, \omega) \in \Xi}} (z(f, x^{(k)}, w^{(k)}) - z(f, \chi, \omega)) \leq R^{(k)}$, then obviously $(z(f, x^{(k)}, w^{(k)}) - z(f, \chi, \omega)) \leq R^{(k)}$, for all $f \in \mathcal{M}$ and $(\chi, \omega) \in \Xi$,

2. if $\max_{\substack{f \in \mathcal{M} \\ (\chi, \omega) \in \Xi}} (z(f, x^{(k)}, w^{(k)}) - z(f, \chi, \omega)) > R^{(k)}$,
then the most violated constraint to $(x^{(k)}, w^{(k)}, R^{(k)})$ is
 $(z(f^{(k+1)}, x^{(k)}, w^{(k)}) - z(f^{(k+1)}, \chi^{(k+1)}, \omega^{(k+1)})) > R^{(k)}$, where
 $(f^{(k+1)}, \chi^{(k+1)}, \omega^{(k+1)})$ is a solution to $\max_{\substack{f \in \mathcal{M} \\ (\chi, \omega) \in \Xi}} (z(f, x^{(k)}, w^{(k)}) - z(f, \chi, \omega))$.

We conclude the relaxation procedure for minimax regret of expected recourse problem in the following steps:

Algorithm 4.5 Relaxation procedure.

step 1: Initialization. Choose $f^{(1)} = \underline{f}$, and $(\chi^{(1)}, \omega^{(1)})$ is an optimal solution to the problem $\min_{(\chi, \omega) \in \Xi} z(\underline{f}, \chi, \omega)$. Set $k = 1$.

step 2: Solve the current relaxed problem

$$\begin{aligned} & \min_{R, (x, w)} R \\ & \text{s.t.} \quad (x, w) \in \Xi \\ & R \geq 0 \\ & R \geq (z(f^{(i)}, x, w) - z(f^{(i)}, \chi^{(i)}, \omega^{(i)})), \quad i = 1, 2, \dots, k, \end{aligned}$$

and obtain an optimal solution $(x^{(k)}, w^{(k)}, R^{(k)})$ for the relaxed problem.

step 3: Solve $Z^{(k)} = \max_{\substack{f \in \mathcal{M} \\ (\chi, \omega) \in \Xi}} (z(f, x^{(k)}, w^{(k)}) - z(f, \chi, \omega))$
 $= \max_{f \in \mathcal{M}} \left(z(f, x^{(k)}, w^{(k)}) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right)$,

and obtain an optimal solution $(f^{(k+1)}, \chi^{(k+1)}, \omega^{(k+1)})$ and its optimal objective value $Z^{(k)}$.

step 4: If $Z^{(k)} \leq R^{(k)}$, terminate. A minimax regret optimal solution is $(x^{(k)}, w^{(k)})$. If $Z^{(k)} > R^{(k)}$, set $k = k + 1$ and go back to step 2. \diamond

Step 3 is the most computationally demanding part of the algorithm since $f_i = h_i g_{i1} g_{i2} \dots g_{in}$ is nonlinear.

When f is an interval objective coefficient vector, step 3 requires solving a quadratic programming problem. One approach that can handle this interval case is to consider the problem as a mutiparametric linear programming problem. The method provided in [18] for solving a mutiparametric linear programming problem can be used to find optimal solutions (χ_*, ω_*) of $\min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega)$ and the regions of f that satisfy these optimal solutions. Another approach to solve the problem in step 3 when f is interval vector, without going through the mutiparametric linear programming, is based on the idea of Inuiguchi and Sakawa [22] in 1996. Again, their work is limited to interval objective coefficients. They introduced the dual of the subproblem $\min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega)$ and used complementary slackness to yield an equivalent problem that contains a complementary slackness constraint. Then they relax this problem to get rid of the complementary slackness constraint and use branch-and-bound method to add/remove the partial information gotten from the complementary slackness constraint. Mausser and Laguna, [45] and [46], adopt this idea by creating a binary variable to handle the complementary slackness constraint. Hence, they deal with a mixed integer programming version of the problem in step 3.

We may apply these approaches to our minimax regret problem when the objective coefficients are bounds on probabilities generated by random sets and IVPs. However, we will leave them for future research. In this dissertation, we solve step 3 of Algorithm 4.5 by the help of Theorems 2.17 and 2.25. The

details follow. Consider

$$\begin{aligned}
Z^{(k)} &= \max_{\substack{f \in \mathcal{M} \\ (\chi, \omega) \in \Xi}} \left(z(f, x^{(k)}, w^{(k)}) - z(f, \chi, \omega) \right) \\
&= \max_{\substack{f \in \mathcal{M} \\ (\chi, \omega) \in \Xi}} \left(c^T(x^{(k)} - \chi) + s_1 \sum_{j=1}^{K_1} f_1^k(w_1^{(k)^j} - \omega_1^j) + s_2 \sum_{j=1}^{K_2} f_2^k(w_2^{(k)^j} - \omega_2^j) \right. \\
&\quad \left. + \dots + s_m \sum_{j=1}^{K_m} f_m^k(w_m^{(k)^j} - \omega_m^j) \right).
\end{aligned}$$

We do not know the order of the components in the vector $(w_i^{(k)} - \omega_i)$, for each i . Therefore, we can not apply Theorem 4.4 directly to obtain $Z^{(k)}$. However, f_i is a joint density of \hat{a}_{ij} , $\forall j = 1, 2, \dots, n$ and \hat{b}_i . Suppose an uncertainty \hat{v} has the set of all realizations $V = \{v_1, v_2, \dots, v_{\alpha_{ij}}\}$, but we do not know the ordering of these realizations. Suppose further that \hat{v} has the same random set or IVP information as \hat{a}_{ij} (or \hat{b}_i , when $V = \{v_1, v_2, \dots, v_{\beta_i}\}$). That is, \mathcal{M}_v has the same restrictions as $\mathcal{M}_{a_{ij}}$, where \mathcal{M}_v is the set of all probability densities satisfying the random set or IVP information of \hat{v} . Hence, v_k is equivalent to a_{ij}^k , $\forall k = 1, 2, \dots, \alpha_{ij}$, and $A_v = \{v_{k_1}, v_{k_2}, \dots, v_{k_t}\} \subseteq V$ is equivalent to $A = \{a_{ij}^{k_1}, a_{ij}^{k_2}, \dots, a_{ij}^{k_t}\} \subseteq \Theta_{a_{ij}}$. For

$$\mathcal{M}_{a_{ij}} = \left\{ g_{ij} : \Theta_{a_{ij}} \rightarrow [0, 1] \mid \sum_{a_{ij}^k \in A} g_{ij}(a_{ij}^k) \in [L_{ij}(A), U_{ij}(A)], \forall A \subseteq \Theta_{a_{ij}} \right\},$$

we can write \mathcal{M}_v as

$$\mathcal{M}_v = \left\{ g : V \rightarrow [0, 1] \mid \sum_{v_k \in A_v} g(v_k) \in [L(A_v), U(A_v)], \forall A_v \subseteq V \right\},$$

such that $g(v_k) = g_{ij}(a_{ij}^k)$, $\forall k = 1, 2, \dots, \alpha_{ij}$, $L(A_v) = L_{ij}(A)$, and $U(A_v) = U_{ij}(A)$. Using Theorems 2.17 or 2.25 (depending upon the interpretation of uncertainty \hat{v}) to solve the problems $\min_{g \in \mathcal{M}_v} g^T v$, with all $\alpha_{ij}!$ orderings of v , we

obtain $\alpha_{ij}!$ optimal solutions $g^{(1)}, g^{(2)}, \dots, g^{(\alpha_{ij}!)} \in \mathcal{M}_v$. Therefore, we know that $g_{ij}^{(1)}, g_{ij}^{(2)}, \dots, g_{ij}^{(\alpha_{ij}!)} \in \mathcal{M}_{a_{ij}}$. Similarly, we have $h_i^{(1)}, h_i^{(2)}, \dots, h_i^{(\beta_i!)} \in \mathcal{M}_{b_i}$. For example, suppose we know that \widehat{b}_i has three realizations b_i^1, b_i^2, b_i^3 with the density function $h_i(b_i^1) \in [\frac{1}{3}, \frac{1}{2}]$, $h_i(b_i^2) \in [\frac{1}{6}, \frac{2}{3}]$, and $h_i(b_i^3) \in [\frac{1}{6}, \frac{1}{2}]$, then $V = \{v_1, v_2, v_3\}$ can have $3! = 6$ possible orderings with the associated probabilities that provide the lowest expected values with respect to those orderings as follows:

$$\begin{aligned}
v_1 \leq v_2 \leq v_3 &\Rightarrow \underline{h}^{(1)}(v_1) = \frac{1}{2}, \underline{h}^{(1)}(v_2) = \frac{1}{3}, \underline{h}^{(1)}(v_3) = \frac{1}{6}, \\
v_1 \leq v_3 \leq v_2 &\Rightarrow \underline{h}^{(2)}(v_1) = \frac{1}{2}, \underline{h}^{(2)}(v_2) = \frac{1}{6}, \underline{h}^{(2)}(v_3) = \frac{1}{3}, \\
v_2 \leq v_1 \leq v_3 &\Rightarrow \underline{h}^{(3)}(v_1) = \frac{1}{3}, \underline{h}^{(3)}(v_2) = \frac{1}{2}, \underline{h}^{(3)}(v_3) = \frac{1}{6}, \\
v_2 \leq v_3 \leq v_1 &\Rightarrow \underline{h}^{(4)}(v_1) = \frac{1}{3}, \underline{h}^{(4)}(v_2) = \frac{1}{2}, \underline{h}^{(4)}(v_3) = \frac{1}{6}, \\
v_3 \leq v_1 \leq v_2 &\Rightarrow \underline{h}^{(5)}(v_1) = \frac{1}{3}, \underline{h}^{(5)}(v_2) = \frac{1}{6}, \underline{h}^{(5)}(v_3) = \frac{1}{2}, \\
v_3 \leq v_2 \leq v_1 &\Rightarrow \underline{h}^{(6)}(v_1) = \frac{1}{3}, \underline{h}^{(6)}(v_2) = \frac{1}{6}, \underline{h}^{(6)}(v_3) = \frac{1}{2}.
\end{aligned}$$

These probabilities of \widehat{a}_{ij} and \widehat{b}_i generate $t_i = \beta_i! \times \alpha_{i1}! \times \alpha_{i2}! \times \dots \times \alpha_{im}!$ joint probabilities $f_i^{(1)}, f_i^{(2)}, \dots, f_i^{(t_i)} \in \mathcal{M}_i$, for each $i = 1, 2, \dots, m$. Therefore, we consider $t = t_1 \times t_2 \times \dots \times t_m$ of joint probability vectors $f = (f_1, f_2, \dots, f_m)$. For each of these joint probability vectors $f^{(l)}$, $l = 1, 2, \dots, t$, we solve for an optimal solution $(\chi_{f^{(l)}}, \omega_{f^{(l)}})$ of the problem

$$\min_{(\chi, \omega) \in \Xi} z(f^{(l)}, \chi, \omega).$$

Algorithm 4.6 A method for solving step 3 of Algorithm 4.5.

step 3.1: Let $\mathcal{M}^t = \{f^{(1)}, f^{(2)}, \dots, f^{(t)}\}$. Then

$$Z^{(k)} = \max_{f \in \mathcal{M}} \left(z(f, x^{(k)}, w^{(k)}) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right) \quad (4.26)$$

$$= \max_{f \in \mathcal{M}^t} \left(z(f, x^{(k)}, w^{(k)}) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right) \quad (4.27)$$

$$= \max \begin{pmatrix} z(f^{(1)}, x^{(k)}, w^{(k)}) - z(f^{(1)}, \chi_{f^{(1)}}, \omega_{f^{(1)}}), \\ z(f^{(2)}, x^{(k)}, w^{(k)}) - z(f^{(2)}, \chi_{f^{(2)}}, \omega_{f^{(2)}}), \\ \vdots \\ z(f^{(t)}, x^{(k)}, w^{(k)}) - z(f^{(t)}, \chi_{f^{(t)}}, \omega_{f^{(t)}}) \end{pmatrix} \quad (4.28)$$

$$= z(f^{(k+1)}, x^{(k)}, w^{(k)}) - z(f^{(k+1)}, \chi^{(k+1)}, \omega^{(k+1)}).$$

Thus, $f^{(k+1)} = f^{(l)}$ and $(\chi^{(k+1)}, \omega^{(k+1)}) = (\chi_{f^{(l)}}, \omega_{f^{(l)}}) \exists l \in \{1, 2, \dots, t\}$ is an optimal solution for (4.26). \diamond

Since the possible terms of $(f^{(k)}, \chi^{(k)}, \omega^{(k)})$ that provide optimal $Z^{(k)}$ are finite as shown in (4.28), Algorithm 4.5 terminates in the finite number of iterations. Theorem 4.7 shows why (4.26) and (4.28) are equal. Moreover, if $(\chi_{f^{(r)}}, \omega_{f^{(r)}})$ is an optimal solution for both problems $\min_{(\chi, \omega) \in \Xi} z(f^{(r_1)}, \chi, \omega)$ and $\min_{(\chi, \omega) \in \Xi} z(f^{(r_2)}, \chi, \omega)$, but $\min_{(\chi, \omega) \in \Xi} z(f^{(r_1)}, \chi, \omega) \leq \min_{(\chi, \omega) \in \Xi} z(f^{(r_2)}, \chi, \omega)$, then the size of the set \mathcal{M}^t can be reduced by removing $f^{(r_2)}$ from the set.

Theorem 4.7 *Let $(x^{(k)}, w^{(k)})$ be given. Suppose that*

$$Z^{(k)} = \max_{f \in \mathcal{M}} \left(z(f, x^{(k)}, w^{(k)}) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right), \text{ and}$$

$$Z_{\mathcal{M}^t}^{(k)} = \max_{f \in \mathcal{M}^t} \left(z(f, x^{(k)}, w^{(k)}) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right).$$

Then $Z^{(k)} = Z_{\mathcal{M}^t}^{(k)}$.

Proof: It is clear that $Z^{(k)} \geq Z_{\mathcal{M}^t}^{(k)}$, since $\mathcal{M} \supseteq \mathcal{M}^t$. For each $f \in \mathcal{M}$, suppose that (χ_f, ω_f) is an optimal solution for $\min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega)$. Then, for each $f \in \mathcal{M}$,

$$\begin{aligned}
Z_f^{(k)} &= \left(z(f, x^{(k)}, w^{(k)}) - \min_{(\chi, \omega) \in \Xi} z(f, \chi, \omega) \right) \\
&= (z(f, x^{(k)}, w^{(k)}) - z(f, \chi_f, \omega_f)) \\
&= c^T(x^{(k)} - \chi_f) + s_1 \sum_{L=1}^{K_1} f_1^L(w^{(k)}_1^L - \omega_{f_1}^L) + s_2 \sum_{L=1}^{K_2} f_2^L(w^{(k)}_2^L - \omega_{f_2}^L) \\
&\quad + \dots + s_m \sum_{L=1}^{K_m} f_m^L(w^{(k)}_m^L - \omega_{f_m}^L) \\
&\leq c^T(x^{(k)} - \chi_f) + s_1 \sum_{L=1}^{K_1} f^{(r)}_1^L(w^{(k)}_1^L - \omega_{f_1}^L) + s_2 \sum_{L=1}^{K_2} f^{(r)}_2^L(w^{(k)}_2^L - \omega_{f_2}^L) \\
&\quad + \dots + s_m \sum_{L=1}^{K_m} f^{(r)}_m^L(w^{(k)}_m^L - \omega_{f_m}^L), \quad (\text{where } f^{(r)} \text{ is obtained by Theorem 2.17 or 2.25} \\
&\hspace{15em} \text{for some } r = 1, 2, \dots, t) \\
&= (z(f^{(r)}, x^{(k)}, w^{(k)}) - z(f^{(r)}, \chi_f, \omega_f)) \\
&\leq (z(f^{(r)}, x^{(k)}, w^{(k)}) - z(f^{(r)}, \chi_{f^{(r)}}, \omega_{f^{(r)}})) \quad ((\chi_{f^{(r)}}, \omega_{f^{(r)}}) \text{ is an optimal for} \\
&\hspace{15em} \min_{(\chi, \omega) \in \Xi} z(f^{(r)}, \chi, \omega)).
\end{aligned}$$

Hence, $Z^{(k)} \leq Z_{\mathcal{M}^t}^{(k)}$. \square

We provided, in Sections 4.1 and 4.2, a pessimistic, an optimistic, and a minimax regret solution for an LP problem with uncertainty by transforming the problem to expected recourse problems (with all possible probabilities in the set \mathcal{M}). The recourse variables for these expected recourse problems are the amount of shortages needed to buy from a market with a particular price. In the next section, we will consider also the case of storage price or selling price when we have an over production (some excess occurs).

4.3 Storage or selling costs in expected recourse problems

We will not combine the recourse variables of selling the excess products and storage the excess products in one problem. If a refinery's manager has both

choices of selling and storing the excess products, s/he will just sell the products to make more profits, instead of spending money for storage. Therefore, we will consider the following two situations on top of the priority that the refinery needs to satisfy demands of its client while minimizing its production cost.

1. The manager sells the excess products after he satisfying the demands with the selling price α .
2. The manager stores the excess products after satisfying the demands with the storage price β .

4.3.1 Selling costs in expected recourse problems

We assume that the selling price of the excess products α per unit of product is less than or equal to the buying price s from outside market when shortages occur. Otherwise, we can just buy with the lower price and sell with higher price to make a profit. Let w be a recourse variable measuring the shortages in production, and v be a recourse variable measuring the excesses in production. We obtain an analogous expected recourse model of the model (4.9) as follows, where $s_i \geq \alpha_i$, $i = 1, 2, \dots, m$.

$$\left. \begin{array}{l}
 \min_{x \in \Psi} \quad c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \\
 \quad s_1 \sum_{L=1}^{K_1} f_1^L w_1^L + s_2 \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + s_m \sum_{L=1}^{K_m} f_m^L w_m^L \\
 \quad - \alpha_1 \sum_{L=1}^{K_1} f_1^L v_1^L - \alpha_2 \sum_{L=1}^{K_2} f_2^L v_2^L - \dots - \alpha_m \sum_{L=1}^{K_m} f_m^L v_m^L \\
 \text{s.t.} \quad w_1^1 - v_1^1 \geq b_1^1 - a_{11}^1 x_1 - a_{12}^1 x_2 - \dots - a_{1n}^1 x_n \\
 \quad \vdots \\
 \quad w_1^{K_1} - v_1^{K_1} \geq b_1^{\beta_1} - a_{11}^{\alpha_{11}} x_1 - a_{12}^{\alpha_{12}} x_2 - \dots - a_{1n}^{\alpha_{1n}} x_n \\
 \quad \vdots \\
 \quad w_m^1 - v_m^1 \geq b_m^1 - a_{m1}^1 x_1 - a_{m2}^1 x_2 - \dots - a_{mn}^1 x_n \\
 \quad \vdots \\
 \quad w_m^{K_m} - v_m^{K_m} \geq m_1^{\beta_m} - a_{m1}^{\alpha_{m1}} x_1 - a_{m2}^{\alpha_{m2}} x_2 - \dots - a_{mn}^{\alpha_{mn}} x_n, \\
 \quad v, w \geq 0.
 \end{array} \right\} (4.29)$$

The objective function of (4.29) can be rewritten as

$$\begin{aligned}
& c_1x_1 + c_2x_2 + \dots + c_nx_n + \\
& (s_1 - \alpha_1) \sum_{L=1}^{K_1} f_1^L w_1^L + (s_2 - \alpha_2) \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + (s_m - \alpha_m) \sum_{L=1}^{K_m} f_m^L w_m^L \\
& + \alpha_1 \sum_{L=1}^{K_1} f_1^L (w_1^L - v_1^L) + \alpha_2 \sum_{L=1}^{K_2} f_2^L (w_2^L - v_2^L) + \dots + \alpha_m \sum_{L=1}^{K_m} f_m^L (w_m^L - v_m^L).
\end{aligned}$$

It is not hard to see that for each $i = 1, 2, \dots, m$, $w_i = (w_i^1, w_i^2, \dots, w_i^{K_i})$ and $w_i - v_i = (w_i^1 - v_i^1, w_i^2 - v_i^2, \dots, w_i^{K_i} - v_i^{K_i})$ have the same ordering. Hence, we can apply any of the approaches presented in Sections 4.1 and 4.2 to (4.29).

4.3.2 Storage costs in expected recourse problems

An expected recourse model with recourse variables w and v that represent the shortages and the storage excesses is

$$\left. \begin{aligned}
& \min_{x \in \Psi} c_1x_1 + c_2x_2 + \dots + c_nx_n + \\
& s_1 \sum_{L=1}^{K_1} f_1^L w_1^L + s_2 \sum_{L=1}^{K_2} f_2^L w_2^L + \dots + s_m \sum_{L=1}^{K_m} f_m^L w_m^L \\
& + \beta_1 \sum_{L=1}^{K_1} f_1^L v_1^L + \beta_2 \sum_{L=1}^{K_2} f_2^L v_2^L + \dots + \beta_m \sum_{L=1}^{K_m} f_m^L v_m^L \\
& \text{s.t. } w_1^1 - v_1^1 = b_1^1 - a_{11}^1 x_1 - a_{12}^1 x_2 - \dots - a_{1n}^1 x_n \\
& \quad \vdots \\
& w_1^{K_1} - v_1^{K_1} = b_1^{\beta_1} - a_{11}^{\alpha_{11}} x_1 - a_{12}^{\alpha_{12}} x_2 - \dots - a_{1n}^{\alpha_{1n}} x_n \\
& \quad \vdots \\
& w_m^1 - v_m^1 = b_m^1 - a_{m1}^1 x_1 - a_{m2}^1 x_2 - \dots - a_{mn}^1 x_n \\
& \quad \vdots \\
& w_m^{K_m} - v_m^{K_m} = m_1^{\beta_m} - a_{m1}^{\alpha_{m1}} x_1 - a_{m2}^{\alpha_{m2}} x_2 - \dots - a_{mn}^{\alpha_{mn}} x_n, \\
& v, w \geq 0.
\end{aligned} \right\} (4.30)$$

The constraints of (4.30) need to be equality constraints. Because at optimality, if it turns out that $w_i^L - v_i^L \geq b_i^L - a_{i1}^L x_1 - a_{i2}^L x_2 - \dots - a_{in}^L x_n$, $\exists i =$

$1, 2, \dots, m$, $\exists L = 1, 2, \dots, K_i$, while $a_{i1}^L x_1 - a_{i2}^L x_2 - \dots - a_{in}^L x_n > b_i^L$, then $v_i^L = 0$ would provides the minimum objective value instead of the correct excess amount $v_i^L = a_{i1}^L x_1 - a_{i2}^L x_2 - \dots - a_{in}^L x_n - b_i^L$. An similar argument applies for the case of ' \leq '.

In order to find a pessimistic, optimistic, and minimax regret solutions for (4.30), we first rewrite its objective function to

$$c_1 x_1 + c_2 x_2 + \dots + c_n x_n + \sum_{L=1}^{K_1} f_1^L (s_1 w_1^L + \beta_1 v_1^L) + \sum_{L=1}^{K_2} f_2^L (s_2 w_2^L + \beta_2 v_2^L) + \dots + \sum_{L=1}^{K_m} f_m^L (s_m w_m^L + \beta_m v_m^L).$$

Since we do not know the ordering of $(s_i w_i^1 + \beta_i v_i^1, s_i w_i^2 + \beta_i v_i^2, \dots, s_i w_i^{K_i} + \beta_i v_i^{K_i})$, what we can do is to apply Theorems 2.17 and 2.25 to find all possible probabilities, as explained in the second paragraph of page 129.

4.4 Comparison between pessimistic, optimistic, minimax regret approaches and the reviewed approaches

The reviewed approaches in Chapter 3 that incorporate uncertainty information into the models are expected recourse, expected average, expected recourse-average, and interval expected value models. An expected recourse model is for an LP problem with probability uncertainty. An expected average model is used when uncertainties in an LP problem with uncertainty are possibility. A mixture of probability and possibility in an LP with uncertainty can be treated by an expected recourse-average model when these two uncertainties are not in the same constraint, otherwise we apply an interval expected value model.

However, we had already argued in Chapter 3 that an interval expected value model is limited in its ability to find the expected objective value in the

long run. Moreover, optimal solutions and objective values of an expected average model and an expected recourse-average model are not reasonable with respect to the set of probabilities generated by possibility interpretations of uncertainty. On the other hand, the pessimistic, optimistic and minimax regret approaches explained in Sections 4.1 and 4.2 are more reasonable and faithful to the information available. They provide three solutions as information for a user that in the long run, the user can expect his/her objective value to be between pessimistic and optimistic values. In addition, the user can use the minimax solution to minimize the regret of not knowing the actual probability.

Table 4.1 supports the argument made in the previous paragraph that the reviewed approaches are not reasonable for LP problems with mixed uncertainty. From the production planning problem in Chapter 3, although an optimal solution $(x_1, x_2) = (45.63, 16.04)$ for the expected average model is in the ranges $[24.33, 50.33]$ and $[11.33, 37.33]$ (obtained from the associated pessimistic and optimistic models) where the long run solution x_1 and x_2 is supposed to be, the objective value $z = \$139.37$ for the expected average model in $[\$134.67, \$160.65]$ does not have a reasonable semantic. We may say that there is a probability in the set \mathcal{M} generated by uncertainties presented in (3.13) such that an expected recourse model with respect to this probability provides the objective value of $\$139.37$. However, we do not have any theories to support that an optimal first stage variable solution (x_1, x_2) obtained by this expected recourse model also is optimal for the expected average model. Moreover, if we were to use an optimal solution from the expected average model to represent the LP problem with possibility uncertainty, we will not have the second stage solution for the con-

Table 4.1: Computational Results for the new approaches and the reviewed approach of the production planning problem.

Types of models, depending upon uncertainty interpretations	Raw materials		Production costs
	x_1	x_2	
1. Expected recourse model	31.80	29.87	\$155.38
2. Expected average model	45.63	16.04	\$139.37
2.1 Pessimistic expected recourse	24.33	37.33	\$160.65
2.2 Optimistic expected recourse	50.33	11.33	\$134.67
2.3 Minimax regret of expected recourses	33.15	28.51	$R = \$17.18$
3. Expected recourse-average model	45.63	16.04	\$139.37
3.1 Pessimistic expected recourse	24.33	37.33	\$160.65
3.2 Optimistic expected recourse	50.33	11.33	\$134.67
3.3 Minimax regret of expected recourses	33.15	28.51	$R = \$17.18$
4. Interval expected value model	[33.89, 44.41]	[9.51, 20.03]	[\$117.35, \$127.87]
4.1 Pessimistic expected recourse	31.80	29.87	\$157.07
4.2 Optimistic expected recourse	39.75	21.92	\$147.38
4.3 Minimax regret of expected recourses	33.61	28.05	$R = \$4.35$

sequence of the first stage solution. Similar discussion applies for the expected recourse-average model.

The objective value in the range [\\$117.35, \\$127.87] provided by the interval expected value model is much lower than the optimistic objective value \$147.38, because the interval expected value model does not have any penalty on its objective function. Moreover, since the probabilities for the expected recourse model are trapped in the sets of probabilities for the interval expected value model, an optimal solution and its objective value of the expected recourse model are in the range of the results from pessimistic and optimistic approaches associated with the interval expected value model.

We now can handle an LP problem with generalized uncertainty by using the knowledge from Chapter 2 that each of PC-BRIN interpretations of uncertainty is in fact a set of probabilities associated with that uncertainty. The uncertainty in the problem does not have to be limited to probability and possibility. It can be any PC-BRIN interpretations. The problem can have more than one interpretation of uncertainty in one constraint. The solutions provide to a user when we do not know the actual probability of each uncertainty are pessimistic, optimistic, and minimax regret solutions.

In Chapter 5, we apply the pessimistic, optimistic, and minimax regret approaches to a farming problem and a radiation shielding design problem.

5. Examples and applications

In this chapter, two application problems are presented using the results of the dissertation: the pessimistic, the optimistic, and the minimax regret approaches. The first application is a farming example, which is adopted from a farming example in [5]. The second application is a radiation shielding design problem from Newman, et al. [53]. The data in these two references are fixed or have probability interpretations. We impute other uncertainties to the problems, where they are reasonable as an easy model of the real applications. The solutions are obtained using GAMS.

5.1 Farming example

The following probability version example is adopted from a farming example in [5]. A farmer raises wheat, corn, and sugar beets on 500 acres of land. He knows that at least 200 tons of wheat and 240 tons of corn are needed for cattle feed. These amounts can be raised on the farm or bought from a market. Any production in excess of the feeding requirement will be sold. Selling prices are \$170 and \$150 per ton of wheat and corn, respectively. The purchase price is higher, \$238 per ton of wheat and \$210 per ton of corn. Another profitable crop is sugar beet, which sells at \$36 per ton. However, the state commission imposes a quota on sugar beet production of 6000 tons. Any amount in excess of the quota can be sold only at \$10 per ton. The mean yield on the farmer's land is 2.5, 3, and 20 tons per acre for wheat, corn, and sugar beets, respectively. Table 5.1 summarizes the data and the planting costs for these crops.

Table 5.1: Data for farming problem.

Available land: 500 acres	Wheat	Corn	Sugar Beets
Yield (tons/acre)	2.5	3	20
Planting cost (\$/acre)	150	230	260
Selling price (\$/ton)	170	150	36 (under 6000 tons) 10 (above 6000 tons)
Purchase price (\$/ton)	238	210	-
Minimum requirement (tons)	200	240	-

To formulate this problem as a linear programming model, let

x_1 = acres of land devoted to wheat,

x_2 = acres of land devoted to corn,

x_3 = acres of land devoted to sugar beets,

w_1 = tons of wheat purchased,

w_2 = tons of corn purchased,

u_1 = tons of wheat sold,

u_2 = tons of corn sold,

u_3 = tons of sugar beets sold at \$36, and

u_4 = tons of sugar beets sold at \$10.

Table 5.2: An optimal solution based on expected yields.

Overall profit: \$118,600	Wheat	Corn	Sugar Beets
Devoted land (acres)	120	80	300
Yield (tons)	300	240	6000
Sales (tons)	100	0	6000 (at \$36/ton)
Purchase (tons)	0	0	-

The problem can be formulated as the model (5.1), and its optimal solution is displayed in Table 5.2.

$$\begin{array}{ll}
 \min & 150x_1 + 230x_2 + 260x_3 + 238w_1 - 170u_1 + 210w_2 - 150u_2 \\
 & -36u_3 - 10u_4 \\
 \text{s.t.} & x_1 + x_2 + x_3 \leq 500 \\
 & 2.5x_1 + w_1 - u_1 \geq 200 \\
 & 3x_2 + w_2 - u_2 \geq 240 \\
 & 20x_3 - u_3 - u_4 \geq 0 \\
 & w_3 \leq 6000 \\
 & x, u, w \geq 0.
 \end{array} \quad \left. \vphantom{\begin{array}{l} \min \\ \text{s.t.} \end{array}} \right\} (5.1)$$

Table 5.3: The scenarios of yields for all crops.

	Below Average	Average	Above Average
Wheat (tons/acre)	2	2.5	3
Corn (tons/acre)	2.4	3	3.6
Sugar Beets (tons/acre)	16	20	24
Probability	1/3	1/3	1/3

Now, suppose that the scenarios of yields for all crops vary over years; below average, average, and above average yields, which are shown in Table 5.3. Then, the corresponding expected recourse problem is

$$\begin{array}{ll}
 \min & 150x_1 + 230x_2 + 260x_3 + \\
 & \frac{1}{3}(238w_{11} - 170u_{11} + 210w_{21} - 150u_{21} - 36u_{31} - 10u_{41}) + \\
 & \frac{1}{3}(238w_{12} - 170u_{12} + 210w_{22} - 150u_{22} - 36u_{32} - 10u_{42}) + \\
 & \frac{1}{3}(238w_{13} - 170u_{13} + 210w_{23} - 150u_{23} - 36u_{33} - 10u_{43})
 \end{array} \quad \left. \vphantom{\begin{array}{l} \min \\ \frac{1}{3} \end{array}} \right\} (5.2)$$

$$\left. \begin{array}{ll}
\text{s.t. } 2x_1 + w_{11} - u_{11} \geq 200 & 2.4x_2 + w_{21} - u_{21} \geq 240 \\
2.5x_1 + w_{12} - u_{12} \geq 200 & 3x_2 + w_{22} - u_{22} \geq 240 \\
3x_1 + w_{13} - u_{13} \geq 200 & 3.6x_2 + w_{23} - u_{23} \geq 240 \\
16x_3 - u_{31} - u_{41} \geq 0 & u_{31} \leq 6000 \\
20x_3 - u_{32} - u_{42} \geq 0 & u_{32} \leq 6000 \\
24x_3 - u_{33} - u_{43} \geq 0 & u_{33} \leq 6000 \\
x_1 + x_2 + x_3 \leq 500 & x, u, w \geq 0.
\end{array} \right\} \quad (5.3)$$

Table 5.4: An optimal solution based on the expected recourse problem (5.2).

Overall profit: \$108,390	Wheat	Corn	Sugar Beets
Devoted land (acres)	170	80	250
Below Average			
Yield (tons)	340	192	4000
Sales (tons)	140	0	4000 (at \$36/ton)
Purchase (tons)	0	48	-
Average			
Yield (tons)	425	240	5000
Sales (tons)	225	0	5000 (at \$36/ton)
Purchase (tons)	0	0	-
Above Average			
Yield (tons)	510	288	6000
Sales (tons)	310	48	6000 (at \$36/ton)
Purchase (tons)	0	0	-

We assign $k = 1, 2, 3$ corresponding to the below average, average, and above average scenarios, respectively. The variables w_{ik} , $i = 1, 2$ refer to the

purchase (tons) of crop i at scenario k , when $i = 1$ is wheat and $i = 2$ is corn. Likewise, the variables u_{jk} , $j = 1, 2, 3, 4$ are the sales (tons) for each scenario. As an example, u_{42} represents the amount of sugar beets sold at \$10/ton when yields are average. An optimal solution of the expected recourse problem (5.2) with the constraints (5.3) is given in Table 5.4. With overall expected profit of \$108,390, the farmer must determined the first stage planting areas of 170, 80, and 250 acres of wheat, corn, and sugar beets, respectively, before realizing the crop yields. The yields, sales, and purchases are called the second stage because they are recovered later after knowing the realization.

5.1.1 A random set version of the farming example

Suppose now that the farmer collects the weather data for the past 300 years of his town from the local government sources. The farmer finds out that there are 100 years that have bad weather, but he could not see the pattern when bad weather may happen. There are 150 years that have no bad things happening at all; no monster storms, no devastating hail, no outbreak of insects. Besides, the weather seems to incorporate with the planting schedule during these years. However, since there are no data of the yields, the farmer does not know for sure that these 150 years have above average or on average yields. Moreover, the weather data show that there are 50 years of chaos; mixed good, bad, and quite days for farming. The farmer cannot decide that these 50 years provide below average, average, or above average crops' yields. For this information, the farmer has a random set with respect to the uncertainty \hat{v} :

$$\hat{v} = \left\{ v_1 = [2, 2.4, 16]^T, v_2 = [2.5, 3, 20]^T, v_3 = [3, 3.6, 24]^T \right\}$$

such that

$$m(\{v_1\}) = \frac{1}{3}, \quad m(\{v_2, v_3\}) = \frac{1}{2}, \quad m(\{v_1, v_2, v_3\}) = \frac{1}{6}.$$

We can see that $v_1 \leq v_2 \leq v_3$ and

$$\begin{aligned} Bel(\{v_1, v_2, v_3\}) &= 1, & Pl(\{v_1, v_2, v_3\}) &= 1, \\ Bel(\{v_1, v_2\}) &= \frac{1}{3}, & Pl(\{v_1, v_2\}) &= 1, \\ Bel(\{v_1, v_3\}) &= \frac{1}{3}, & Pl(\{v_1, v_3\}) &= 1, \\ Bel(\{v_2, v_3\}) &= \frac{1}{2}, & Pl(\{v_2, v_3\}) &= \frac{2}{3}, \\ Bel(\{v_1\}) &= \frac{1}{3}, & Pl(\{v_1\}) &= \frac{1}{2}, \\ Bel(\{v_2\}) &= 0, & Pl(\{v_2\}) &= \frac{2}{3}, \\ Bel(\{v_3\}) &= 0, & Pl(\{v_3\}) &= \frac{2}{3}. \end{aligned}$$

The probabilities (proportions) of below average, average, and above average yields over 300 years can be in the intervals as follows:

$$f(v_1) \in \left[\frac{1}{3}, \frac{1}{2} \right], \quad f(v_2) \in \left[0, \frac{2}{3} \right], \quad f(v_3) \in \left[0, \frac{2}{3} \right], \quad (5.4)$$

and therefore,

$$\left. \begin{aligned} \underline{f}(\{-v_1\}) &= \frac{1}{3}, & \bar{f}(\{-v_1\}) &= \frac{1}{2}, \\ \underline{f}(\{-v_2\}) &= 0, & \bar{f}(\{-v_2\}) &= \frac{1}{2}, \\ \underline{f}(\{-v_3\}) &= \frac{2}{3}, & \bar{f}(\{-v_3\}) &= 0. \end{aligned} \right\} \quad (5.5)$$

Please note that the proportion $f(v_1)$ cannot be continuous on $\left[\frac{1}{3}, \frac{1}{2} \right]$, it could be one from $\left\{ \frac{100}{300}, \frac{101}{300}, \dots, \frac{150}{300} \right\}$. Similar arguments apply for $f(v_2)$ and $f(v_3)$.

Therefore (5.4) can be written more precisely as

$$\left. \begin{aligned} f(v_1) &\in \left\{ \frac{100}{300}, \frac{101}{300}, \dots, \frac{150}{300} \right\}, \\ f(v_2) &\in \left\{ 0, \frac{1}{300}, \dots, \frac{200}{300} \right\}, \\ f(v_3) &\in \left\{ 0, \frac{1}{300}, \dots, \frac{200}{300} \right\}. \end{aligned} \right\} \quad (5.6)$$

In this farming problem, we have the excess variables u_{jk} that make the problem more challenging than the production planning problem in Chapter 4 that motivated our theories, which has only the shortage variables w_{ik} . Next, we explain that \underline{f} and \overline{f} constructed in Theorem 4.2 are still valid to be the probabilities that provide the lowest and highest expected values of the set of expected recourse problems, respectively.

Given any probability density function f that satisfies (5.6), at an optimality, we know that

$$\begin{aligned} 2x_1^* + w_{11}^* - u_{11}^* &= 200, & 2.4x_2^* + w_{21}^* - u_{21}^* &= 240, & 16x_3^* - u_{31}^* - u_{41}^* &= 0, \\ 2.5x_1^* + w_{12}^* - u_{12}^* &= 200, & 3x_2^* + w_{22}^* - u_{22}^* &= 240, & 20x_3^* - u_{32}^* - u_{42}^* &= 0, \\ 3x_1^* + w_{13}^* - u_{13}^* &= 200, & 3.6x_2^* + w_{23}^* - u_{23}^* &= 240, & 24x_3^* - u_{33}^* - u_{43}^* &= 0. \end{aligned}$$

Hence, $w_{i3}^* - u_{i3}^* \leq w_{i2}^* - u_{i2}^* \leq w_{i1}^* - u_{i1}^*$, $i = 1, 2$. The solution can be in these two forms: $w_{ik}^* \geq 0, u_{ik}^* = 0$ or $w_{ik}^* = 0, u_{ik}^* \geq 0$, $i = 1, 2, k = 1, 2, 3$. Therefore, for $i = 1, 2$,

- $w_{i3}^* \leq w_{i2}^* \leq w_{i1}^*$, and
- $-u_{i3}^* \leq -u_{i2}^* \leq -u_{i1}^*$,

Similarly, $-u_{33}^* - u_{43}^* \leq -u_{32}^* - u_{42}^* \leq -u_{31}^* - u_{41}^*$.

- If $u_{31}^* = u_{32}^* = u_{33}^* = 6000$, then $-u_{43}^* \leq -u_{42}^* \leq -u_{41}^*$.
- If u_{31}^*, u_{32}^* , and u_{33}^* are less than 6000, then $-u_{43}^* = -u_{42}^* = -u_{41}^* = 0$ and $-u_{33}^* \leq -u_{32}^* \leq -u_{31}^*$.

Hence, all recourse variables are in an increasing order, and we can now apply Theorem 4.2 to the problem. The expected recourse problems that provide the lowest and the largest expected objective values are

$$\left. \begin{aligned} \min_{(x,u,w) \in \Xi} & 150x_1 + 230x_2 + 260x_3 + \\ & \frac{1}{3} (238w_{11} - 170u_{11} + 210w_{21} - 150u_{21} - 36u_{31} - 10u_{41}) + \\ & 0 (238w_{12} - 170u_{12} + 210w_{22} - 150u_{22} - 36u_{32} - 10u_{42}) + \\ & \frac{2}{3} (238w_{13} - 170u_{13} + 210w_{23} - 150u_{23} - 36u_{33} - 10u_{43}) \end{aligned} \right\} (5.7)$$

and

$$\left. \begin{aligned} \min_{(x,u,w) \in \Xi} & 150x_1 + 230x_2 + 260x_3 + \\ & \frac{1}{2} (238w_{11} - 170u_{11} + 210w_{21} - 150u_{21} - 36u_{31} - 10u_{41}) + \\ & \frac{1}{2} (238w_{12} - 170u_{12} + 210w_{22} - 150u_{22} - 36u_{32} - 10u_{42}) + \\ & 0 (238w_{13} - 170u_{13} + 210w_{23} - 150u_{23} - 36u_{33} - 10u_{43}), \end{aligned} \right\} (5.8)$$

where $\Xi = \{(x, u, w) : (x, u, w) \text{ satisfies the constraints (5.3)}\}$. Optimal solutions for these pessimistic and optimistic cases can be seen in Table 5.5.

We also provide an optimal solution based on minimax regret with respect to not knowing probability of each scenario. For a given probability density function f , we define

$$\begin{aligned} z(f, x, u, w) = & 150x_1 + 230x_2 + 260x_3 + \\ & f(v_1) (238w_{11} - 170u_{11} + 210w_{21} - 150u_{21} - 36u_{31} - 10u_{41}) + \\ & f(v_2) (238w_{12} - 170u_{12} + 210w_{22} - 150u_{22} - 36u_{32} - 10u_{42}) + \\ & f(v_3) (238w_{13} - 170u_{13} + 210w_{23} - 150u_{23} - 36u_{33} - 10u_{43}). \end{aligned}$$

This minimax regret of farming problem is stated as

$$\min_{(x,u,w) \in \Xi} \max_{\substack{f \in \mathcal{M} \\ (\chi,v,\omega) \in \Xi}} (z(f, x, u, w) - z(f, \chi, v, \omega)), \quad (5.9)$$

Table 5.5: Optimal solutions based on the pessimistic and optimistic expected recourse of the farming problem. P = pessimistic, O = optimistic, R = minimax regret.

	Wheat			Corn			Sugar Beets		
	P	R	O	P	R	O	P	R	O
Devoted land (acres)	100	145.98	183.33	100	82.32	66.67	300	271.70	250
Below Average									
Yield (tons)	200	291.95	366.67	240	197.57	160	4800	4347.25	4000
Sales (tons)	0	91.95	166.67	0	0	0	4800	4347.25	4000
Purchase (tons)	0	0	0	0	42.43	80	-	-	-
Average									
Yield (tons)	250	364.94	458.33	300	246.97	200	6000	5434.06	5000
Sales (tons)	50	164.94	258.33	60	6.97	0	6000	5434.06	5000
Purchase (tons)	0	0	0	0	0	40	-	-	-
Above Average									
Yield (tons)	300	437.93	550	360	296.36	240	7200	6520.87	6000
Sales (tons)	100	237.93	350	120	56.36	0	6000	6000	6000
Purchase (tons)	0	0	0	0	0	0	1200*	520.87*	-

* Sugar beets sold at \$10/ton.

Pessimistic profit	\$87,150
Optimistic profit	\$127,677.78
Minimax regret value	\$4656.51

where $\mathcal{M} = \{f : \sum_{i=1}^3 f(v_i) = 1 \text{ and } f \text{ satisfies the system (5.6)}\}$, and $\Xi = \{(x, u, w) : (x, u, w) \text{ satisfies the constraints in (5.2)}\}$. An optimal solution of the regret problem also is given in Table 5.5.

Therefore, the farmer knows that to maintain the profit of his yields to be in $[\$87,150, \$127,677.78]$, he should use his land between 100 and 183.33 acres for wheat, 66.67 and 100 acres for corn, and 250 and 300 acres for sugar beets. Moreover, since the farmer does not know the weather in advance, he can use information that he has about these past 300 years to plan how many acres that he should use for each crop so that he minimizes the maximum regret of his decision in the long run. For example, if for some reason, the farmer knows that the below average, average, and above average years have equal probability of $\frac{1}{3}$, but he uses the minimax regret strategy instead of Table 5.4. He would regret no more than \$4656.51 in the long run.

5.2 An application for radiation shielding design of radiotherapy treatment vaults

This application problem is based on the real shielding design of radiotherapy treatment vaults problem that the head of Denver Health's oncologists, Francis Newman and his team faced to minimize the cost of their shielding design. Newman et al. [53] provide three LP problems for minimizing the cost, the dose of radiation beam, and the thickness of the shielding materials. We will apply our pessimistic, optimistic, and minimax regret approaches to minimize the dose of radiation beam beyond the shielding when the costs of shielding materials are not certain.

A shielding expert wants to optimize the design of shielding, so that the x-ray dose at the reference point beyond the protective barrier has the lowest amount. Figure 5.1 illustrates an example of a radiation room with a shielded wall. The x-ray dose beyond the protective barrier also needs to be lower than the allowable x-ray dose of 1.0×10^{-4} Sv/week, which is restricted by law. Newman et al. [53] explain that there are peculiarities of shielding design that present challenges to his optimization model. Neutrons, in particular, introduce modeling problems.

To formulate the model, the formula (5.10) outlined in the references [14, 47] is needed. The barrier transmission for the primary barrier is

$$B_{\text{primary}} = \frac{(P)(d_{\text{primary}})^2}{(WU)_{\text{primary}}(T)}, \quad (5.10)$$

where P (Sv/week) = the allowable dose beyond the primary barrier, d (m) = distance from the x-ray target to the protected point for the primary barrier, W (Gy/week) = workload, T = occupancy factor, and U = use factor.

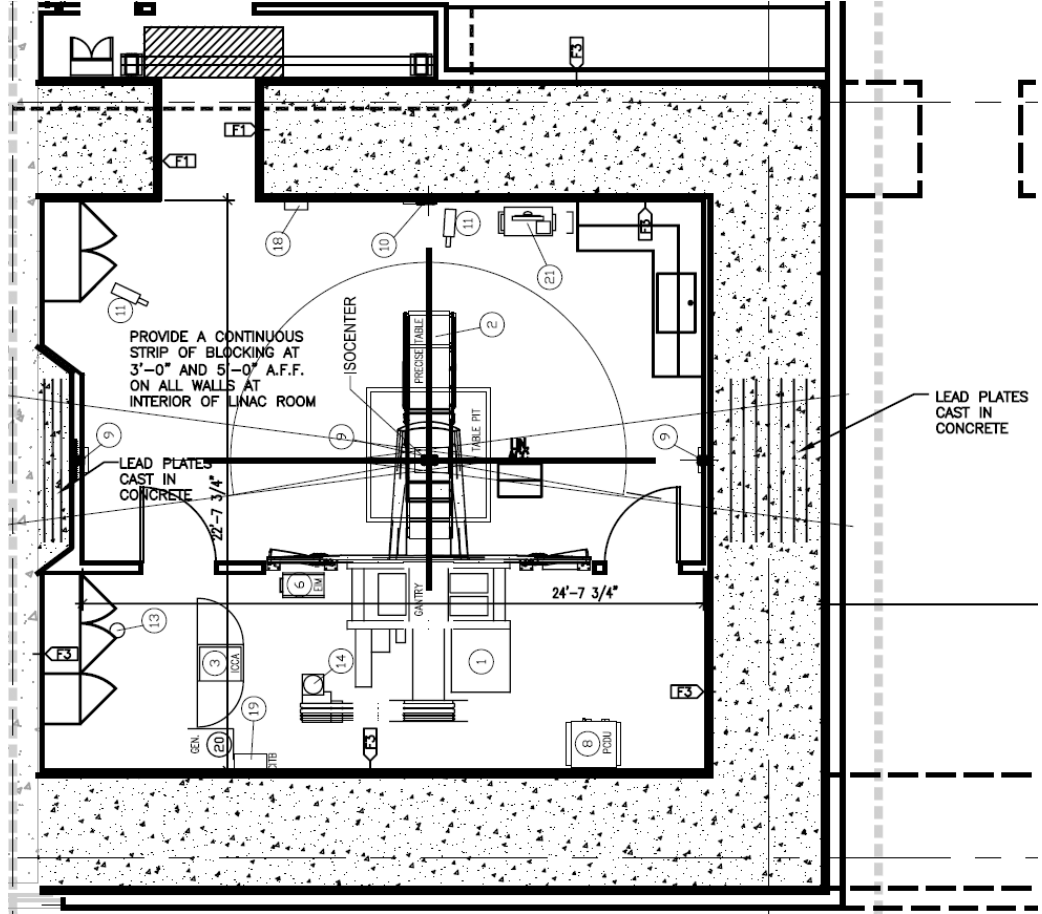


Figure 5.1: An example of a radiation room.

H_{tr} is the transmitted x-ray dose beyond a laminated barrier. With a given B_{new} , H_{tr} is computed by replacing P with H_{tr} in Eq.(5.10), and solving for H_{tr} . To earn B_{new} , we need to know the total number n_{total} of *TVLs*, which can be obtained from solving the LP problem (5.11). $B_{new} = 2.7 \times 10^{-n_{total}}$.

Let TVL_c , TVL_s , and TVL_l be the tenth value layers (the thickness per TVL) of concrete, steel, and lead, respectively, and let n_c , n_s , and n_l be the number of $TVLs$. $n_{\text{total}} = n_c + n_s + n_l$. The number of $TVLs$ needed for a shielding design is $n_B = -\log_{10} B$, where B is calculated from Eq. (5.10). Furthermore, let t_{design} be the total barrier thickness allowed for a particular shielding design, and C_t be the total cost the project can accommodate. Finally, we let c_c , c_s and c_l be the unit costs (\$/square foot/inch of thickness) of concrete, steel, and lead, respectively. Then, the basic linear programming model minimizing dose subject to the dose limit (the first row of the constraints), thickness, and cost constraints is

$$\left. \begin{aligned}
 \min & -n_c - n_s - n_l \\
 \text{s.t. } & n_c + n_s + n_l \geq n_B \\
 & c_c TVL_c n_c + c_s TVL_s n_s + c_l TVL_l n_l \leq C_t \\
 & n_c TVL_c + n_s TVL_s + n_l TVL_l \leq t_{\text{design}} \\
 & n_c, n_s, n_l \geq 0.
 \end{aligned} \right\} \quad (5.11)$$

Suppose that the costs of concrete, steel, and lead per square foot per 1 inch of thickness are \$1, \$80, and \$200, respectively. We consider a controlled area and a 15MV beam. The following input data with respect to a 15 MV beam were used: $P = 1.0 \times 10^{-4}$ Sv/week, $W_{\text{primary}} = 500$ Gy/week, $T = 1$, $U_{\text{primary}} = 1/4$, $d_{\text{primary}} = 5\text{m}$, $TVL_c = 0.4246$ m = 17 inches, $TVL_s = 0.11$ m = 4.33 inches, $TVL_l = 0.057$ m = 2.24 inches, $t_{\text{design}} = 1.5$ m = 60 inches, and $C_t = \$1,000/\text{square foot}$. Therefore,

$$B = \frac{(1 \times 10^{-4})(5^2)}{(500)(1/4)(1)} = 2 \times 10^{-5} \Rightarrow n_B = -\log_{10}(2 \times 10^{-5}) = 4.7.$$

Using this n_B in (5.11), the optimization returns the values $n_c = 2.847$ (1.197 m of concrete), $n_s = 2.750$ (0.303 m of steel), and $n_l = 0$. So, $n_{\text{total}} = 2.847 + 2.750 = 5.597$. Hence,

$$B_{\text{new}} = 2.7 \times 10^{-5.597} = 6.829 \times 10^{-6}, \text{ and}$$

$$H_{\text{tr}} = \frac{(500)(1/4)(1)(6.829 \times 10^{-6})}{5^2} = 3.414 \times 10^{-5}.$$

The optimization returns a dose of 3.414×10^{-5} Sv/week, which is 34.14% of the allowable 1.0×10^{-4} Sv/week.

Suppose that the costs of concrete, steel, and lead per inch of thickness are not uncertain due to the market fluctuations (steel dumping cost, cost of scrap metal, metal lead mining strikes, etc). The goal is to manage all materials within the allowable accommodate cost and satisfy the number needed for a shielding design n_B , while minimizing the radiation dose beyond the barrier. Therefore, we define the second stage variables y_c, y_s , and y_l as the number of TVLs that need to be reduced so that the total cost for each scenario is still in the given budget C_t . We hope that with these reductions, we still satisfy the number needed for a shielding design n_B , and the total allowed barrier thickness t_{design} . Otherwise, the recourse model is infeasible.

Suppose that the uncertain cost of concrete per inch of thickness has the realizations \$0.80, \$1, and \$ 1.20, with the possibility distribution $pos(0.80) = 1$, $pos(1) = \frac{2}{3}$, and $pos(1.20) = 1$. The cost of steel can be \$75, \$80, or \$85 with probabilities $Pr(\{75\}) \in [0, \frac{1}{3}]$, $Pr(\{80\}) \in [\frac{2}{3}, 1]$, and $Pr(\{85\}) \in [0, \frac{1}{3}]$, while the cost of lead are \$190 and \$200, which have the same probability $Pr(\{190\}) = Pr(\{200\}) = \frac{1}{2}$. The possibility distribution provides the random set (\mathcal{F}, m) such that $m(\{1\}) = \frac{2}{3}$, and $m(\{0.80, 1, 1.20\}) = 1$. Therefore,

we obtain the following probability density mass functions:

$$\begin{aligned}
\underline{f}_c^{(0.80)} &= \underline{f}_c(\{0.80\}) = \frac{1}{3}, & \overline{f}_c^{(0.80)} &= \overline{f}_c(\{0.80\}) = 0, \\
\underline{f}_c^{(1)} &= \underline{f}_c(\{1\}) = \frac{2}{3}, & \overline{f}_c^{(1)} &= \overline{f}_c(\{1\}) = \frac{2}{3}, \\
\underline{f}_c^{(1.20)} &= \underline{f}_c(\{1.20\}) = 0, & \overline{f}_c^{(1.20)} &= \overline{f}_c(\{1.20\}) = \frac{1}{3}, \\
\underline{f}_s^{(75)} &= \underline{f}_s(\{75\}) = \frac{1}{3}, & \overline{f}_s^{(75)} &= \overline{f}_s(\{75\}) = 0, \\
\underline{f}_s^{(80)} &= \underline{f}_s(\{80\}) = \frac{2}{3}, & \overline{f}_s^{(80)} &= \overline{f}_s(\{80\}) = \frac{2}{3}, \\
\underline{f}_s^{(85)} &= \underline{f}_s(\{85\}) = 0, & \overline{f}_s^{(85)} &= \overline{f}_s(\{85\}) = \frac{1}{3}, \\
\underline{f}_1^{(190)} &= \underline{f}_1(\{190\}) = \frac{1}{2} = \overline{f}_1(\{190\}) = \overline{f}_1^{(190)}, \text{ and} \\
\underline{f}_1^{(200)} &= \underline{f}_1(\{200\}) = \frac{1}{2} = \overline{f}_1(\{200\}) = \overline{f}_1^{(200)}.
\end{aligned}$$

These uncertainties could happen because of various situations. For example, Company A sets the concrete's price according to the following three cases: case 1, case 2, and case 3 as \$0.8 , \$1, and \$ 1.20 per inch of thickness of concrete, respectively. It provides information about the availability of these three cases as that at least one of case 1 or case 3 is available $\frac{1}{3}$ of the time, and $\frac{2}{3}$ of the time, either case 1, case 2, or case 3 is available. Company B, who delivers steel, also sets the steel's price according to the other three cases: case 1' (cost \$75) is available less than $\frac{1}{3}$ of the time, case 2' (cost \$85) is available less than $\frac{1}{3}$ of the time, and case 3' (cost \$80) is available more than $\frac{2}{3}$ of the time. Finally, each of two cases that company C uses for pricing lead is available $\frac{1}{2}$ of the time. Therefore, we have the uncertainty information as stated at the beginning of the paragraph.

We transform the problem (5.11) to the following expected recourse problem, where f can be replaced by \underline{f} and \overline{f} .

$$\begin{aligned}
\min \quad & -n_c - n_s - n_l + f_c^{(0.80)} f_s^{(75)} f_1^{(190)} \left(y_c^{(0.80)} + y_s^{(75)} + y_1^{(190)} \right) + \dots + \\
& f_c^{(1.20)} f_s^{(85)} f_1^{(200)} \left(y_c^{(1.20)} + y_s^{(85)} + y_1^{(200)} \right) \\
\text{s.t.} \quad & n_c, n_s, n_l, y_c, y_s, y_1 \geq 0, \\
& n_c - y_c^{(0.80)} + n_s - y_s^{(75)} + n_l - y_1^{(190)} \geq n_B, \\
& n_c - y_c^{(0.80)} + n_s - y_s^{(75)} + n_l - y_1^{(200)} \geq n_B, \\
& \vdots \\
& n_c - y_c^{(1.20)} + n_s - y_s^{(85)} + n_l - y_1^{(200)} \geq n_B, \\
& 0.80TVL_c(n_c - y_c^{(0.80)}) + 75TVL_s(n_s - y_s^{(75)}) + 190TVL_1(n_l - y_1^{(190)}) \leq C_t, \\
& 1TVL_c(n_c - y_c^{(1)}) + 80TVL_s(n_s - y_s^{(80)}) + 200TVL_1(n_l - y_1^{(200)}) \leq C_t, \\
& \vdots \\
& 1.20TVL_c(n_c - y_c^{(1.20)}) + 85TVL_s(n_s - y_s^{(85)}) + 200TVL_1(n_l - y_1^{(200)}) \leq C_t, \\
& (n_c - y_c^{(0.80)})TVL_c + (n_s - y_s^{(75)})TVL_s + (n_l - y_1^{(190)})TVL_1 \leq t_{\text{design}}, \\
& (n_c - y_c^{(0.80)})TVL_c + (n_s - y_s^{(75)})TVL_s + (n_l - y_1^{(200)})TVL_1 \leq t_{\text{design}}, \\
& \vdots \\
& (n_c - y_c^{(1.20)})TVL_c + (n_s - y_s^{(85)})TVL_s + (n_l - y_1^{(200)})TVL_1 \leq t_{\text{design}}.
\end{aligned}$$

The objective function can be rewritten as

$$\left. \begin{aligned}
\min \quad & -n_c - n_s - n_l + f_c^{(0.80)} y_c^{(0.80)} + f_c^{(1)} y_c^{(1)} + f_c^{(1.20)} y_c^{(1.20)} + \\
& f_s^{(75)} y_s^{(75)} + f_s^{(80)} y_s^{(80)} + f_s^{(85)} y_s^{(85)} + f_1^{(190)} y_1^{(190)} + f_1^{(200)} y_1^{(200)}.
\end{aligned} \right\} (5.12)$$

An optimal solution of the expected recourse problem with \underline{f} provides $n_c = 2.8294$, $n_s = 2.9332$, $y_c^{(1.20)} = 0.7185$, $y_s^{(80)} = 0.1833$, $y_s^{(85)} = 0.3451$, and the other variables are zero. The objective value with respect to \underline{f} is -5.6404. Hence, the optimization returns the average values $\bar{n}_c = 2.8294 - \frac{1}{3} \cdot 0 - \frac{2}{3} \cdot 0 - 0 \cdot 0.7185 = 2.8294$ (1.20 m of concrete), and $\bar{n}_s = 2.9332 - \frac{1}{3} \cdot 0 - \frac{2}{3} \cdot 0.1833 - 0 \cdot 0.3451 = 2.811$

(0.31 m of steel). So,

$$B_{\text{new}} = 2.7 \times 10^{-(2.8294+2.811)} = 6.1796 \times 10^{-6}, \text{ and}$$

$$H_{\text{tr}} = \frac{(500)(1/4)(1)(6.1796 \times 10^{-6})}{5^2} = 3.090 \times 10^{-5}.$$

An optimal solution of the expected recourse problem with \bar{f} is $n_c = 2.8848$, $n_s = 2.7193$, $y_s^{(75)} = 0.9051$, $y_s^{(85)} = 0.1600$, and the other variables are zero. The objective value with respect to \bar{f} is -5.5508. Hence, the optimization returns the average values $\bar{n}_c = 2.8848 - \frac{1}{3} \cdot 0 - \frac{2}{3} \cdot 0 - 0 \cdot 0 = 2.8848$ (1.225 m of concrete), and $\bar{n}_s = 2.7193 - 0 \cdot 0.9051 - \frac{2}{3} \cdot 0 - \frac{1}{3} \cdot 0.1600 = 2.666$ (0.29 m of steel). So,

$$B_{\text{new}} = 2.7 \times 10^{-(2.8848+2.666)} = 7.5956 \times 10^{-6}, \text{ and}$$

$$H_{\text{tr}} = \frac{(500)(1/4)(1)(7.5956 \times 10^{-6})}{5^2} = 3.798 \times 10^{-5}.$$

Therefore, the bound on the x-ray dose beyond the barrier is $[3.090, 3.798] \times 10^{-5}$ Sv/week, or between 30.90% and 37.98% of the allowable 1×10^{-4} Sv/week.

We do not have the actual probability when we have uncertain costs of concrete, and steel. Cost of concrete has random set information. Therefore, its actual probability is in the set

$$\mathcal{M}_c = \{Pr \mid Nec(A) \leq Pr(A) \leq Pos(A), A \subseteq \{0.8, 1, 1.2\}\}.$$

Cost of steel has probability interval information, so its actual probability is in the set

$$\mathcal{M}_s = \left\{ Pr \mid Pr(\{75\}) \in \left[0, \frac{1}{3}\right], Pr(\{80\}) \in \left[\frac{2}{3}, 1\right], \text{ and } Pr(\{85\}) \in \left[0, \frac{1}{3}\right] \right\}.$$

The minimum of maximum regret due to the unknow probability is

$$\min_{(n,y) \in \Xi} \max_{\substack{f \in S \\ (\chi,v) \in \Xi}} (z(f, n, y) - z(f, \chi, v)), \quad (5.13)$$

where, $\Xi = \{(n, y) \mid (n, y) \text{ satisfies the constraints in (5.12)}\}$,

$$S = \left\{ f = (f_c, f_s, f_1) \mid f_c \in \mathcal{M}_c, f_s \in \mathcal{M}_s, \text{ and } f_1^{(190)} = f_1^{(200)} = \frac{1}{2} \right\}, \text{ and}$$

$$z(f, n, y) = -n_c - n_s - n_1 + f_c^{(0.80)} y_c^{(0.80)} + f_c^{(1)} y_c^{(1)} + f_c^{(1.20)} y_c^{(1.20)} + \\ f_s^{(75)} y_s^{(75)} + f_s^{(80)} y_s^{(80)} + f_s^{(85)} y_s^{(85)} + f_1^{(190)} y_1^{(190)} + f_1^{(200)} y_1^{(200)}.$$

Applying the minimax regret procedure to the problem (5.13), the minimum of the maximum regret is 0.0296, when $n_{c_*} = 2.8717$, $n_{s_*} = 2.7699$, $y_{c_*}^{(1.20)} = 0.0624$, $y_{s_*}^{(80)} = 0.0462$, $y_{s_*}^{(85)} = 0.2064$, and the other variables are zero. No matter what probability in S that we pick to represent the uncertainty, we will not regret more than 0.0296 of total number of *TVLs* with this solution (n_*, y_*) .

However, for the safety, any person who does not want to get exposed to the radiation would want to know the highest number of the x-ray dose beyond the barrier. Therefore, we may provide the pessimistic solution, which will tell that it is not more than 3.798×10^{-5} Sv/week on average of the x-ray dose beyond the barrier. If we were to use the minimax regret solution to represent the problem, then the pessimistic and the optimistic amount of the x-ray dose beyond the barrier with respect to this minimax regret solution will 4.0657×10^{-5} and 3.3078×10^{-5} Sv/week, which are 33.08% and 40.66% of the allowable dose. These percentage numbers are shifted from the actual numbers 30.90% and 37.98% by at most 3%.

Some people who are concerned about the extreme scenario that provides the highest dose beyond the barrier may not use the expected recourse approach, but instead use $c_c = \$1.20$, $c_s = \$85$, and $c_1 = \$200$ in (5.11), which will result 5.6148×10^{-5} Sv/week of the x-ray dose beyond the barrier, or 56.15% of the allowable by law, by using $n_c = 2.8240$ and $n_s = 2.5570$.

We include all the solutions obtained in this section in Table 5.6, for an easy reading.

Table 5.6: A pessimistic, an optimistic, a minimax regret, and an extreme scenario solution for the radiation shielding design problem.

Average number of $TVLs$	Pessimistic solution	Optimistic solution
Concrete (\bar{n}_c)	2.8848 (122.50 cm of concrete)	2.8294 (120.14 cm of concrete)
Steel (\bar{n}_s)	2.6660 (29 cm of concrete)	2.8110 (31 cm of steel)
X-ray dose beyond the barrier (Sv/week)	3.798×10^{-5} (37.98%)	3.090×10^{-5} (30.90%)

Number of $TVLs$	Minimax regret solution	Extreme scenario solution
Concrete (n_c)	2.8717 (122 cm of concrete)	2.8240 (119.90 cm of concrete)
Steel (n_s)	2.7699 (30.47 cm of steel)	2.5570 (28.13 cm of steel)
X-ray dose beyond the barrier (Sv/week)	$[3.3078, 4.0657] \times 10^{-5}$ ([33.08%, 40.66%])	5.6148×10^{-5} (56.15%)

6. Conclusion

The dissertation has developed linear programming problems with coefficients whose uncertainties are restricted to a finite set of realizations. The uncertainties in the scope of this dissertation have had PC-BRIN interpretations. We explained in Chapter 2 that each of PC-BRIN interpretations of uncertainty can be written as a set of probability measures \mathcal{M} associated with that uncertainty. When an uncertainty \hat{u} has random set interpretation (or other interpretations that can be written as a random set, see Figure 2.7), we have proved that any probability in the set of probability measures associated with this random set is trapped between belief and plausibility functions.

Two density functions for the lowest and highest expected values of \hat{u} (or an evaluation function of U , where U is the set of all finite realizations of \hat{u}) can be constructed by Theorem 2.17, when \hat{u} has random set interpretation. These constructed density functions use fewer calculations than the process of establishing an associated LP problems for finding the lowest and highest expected values of this random set uncertainty. Theorem 2.25 and Corollary 2.26 handle the IVP case.

Theorems 2.17 and 2.25 are used in the proof of remodeling pessimistic and optimistic recourse problems as linear programs. The theorems also are used to tackle the relaxation procedure for a minimax regret of expected recourse models. We apply these theorems with reduced expected recourse models which have much smaller size of variables and constraints than ordinary expected recourse

models.

A pessimistic, an optimistic, and a minimax regret of expected recourse models are introduced to be able to handle an LP problem with generalized uncertainty, where more than one uncertainty interpretation can be in the same constraint. Therefore, we overcome the limitation of all approaches in literature. Since we do not know the actual probability of uncertainty, the three solution approaches provided by this thesis are useful information for a user. The user will have extended knowledge about the consequences of various decisions and associated costs in the long run that will be between the pessimistic and the optimistic solution. The minimax regret approach of a set of expected recourse models also is a reasonable solution when we want to see a solution that minimizes the maximum regret of not knowing the actual probability.

This dissertation focused on uncertainty over a finite case. However, we have some results that could apply to continuous realizations of uncertainty, and the full development of these ideas to the continuous case is a logical next step. Moreover, the penalty price also could be uncertain, which means that an expected recourse model becomes a nonlinear program, and the developed theorems may no longer apply.

In addition to this research, we should find the computational complexity of all three approaches and compare them with an expected recourse model, to see how much harder than the computation of a single expected recourse model we need to work on. Hence, we can compare our approaches with a simulation approach. We may randomly choose a probability in the set of probabilities \mathcal{M} , use it in an expected recourse model, re-do the process of choosing a proba-

bility and finding its associated expected recourse solution for a large number of iterations. We would see an approximated pessimistic and an approximated optimistic objective value.

We expect that an expected recourse solution whose associated probability belongs to set \mathcal{M} is trapped between a pessimistic and an optimistic solution. We also should check that a minimax regret solution is trapped between a pessimistic and an optimistic solution, in general. It may have a nice structure that helps reduce the size of the density functions in Algorithm 4.6, so that the complexity of the relaxation procedure is improved. We may apply the idea of using duality and complementarity slackness to the relaxation procedure when an LP's objective coefficients belong to \mathcal{M} . Moreover, we may write a general code in MATLAB (or other programming languages) for LP problems with generalized uncertainty.

Lastly, another direction is to find an extension to the expected average approach so that its solution not only provides a richer semantic but also guarantees to lie within the bounds of associated pessimistic and optimistic solutions.

APPENDIX A. How to obtain a basic probability assignment function m from a given belief measure Bel

The materials in this Appendix can be found in Shafer [60].

Theorem A.1 *Suppose U is a finite set and Bel and m are a belief measure, and a basic probability assignment function on $\mathcal{P}(U)$, respectively. Then*

$$m(A) = \sum_{B|B \subseteq A} (-1)^{|A-B|} Bel(B),$$

for all $A \subseteq X$.

We need the following lemmas in order to solve Theorem A.1.

Lemma A.2 *If A is a finite set, then*

$$\sum_{B \subseteq A} (-1)^{|B|} = \begin{cases} 1 & \text{if } A = \emptyset, \\ 0 & \text{otherwise.} \end{cases}$$

Proof: We know that for a positive number n ,

$$\binom{n}{0} - \binom{n}{1} + \binom{n}{2} + \dots + (-1)^n \binom{n}{n} = (1 - 1)^n = 0.$$

So, whenever $A = \{\theta_1, \theta_2, \dots, \theta_n\}$ is a finite non-empty set,

$$\begin{aligned} \sum_{B \subseteq A} (-1)^{|B|} &= (-1)^{|\emptyset|} + \sum_i (-1)^{|\{\theta_i\}|} + \sum_{i < j} (-1)^{|\{\theta_i, \theta_j\}|} + \dots + (-1)^{|A|} \\ &= \binom{n}{0} - \binom{n}{1} + \binom{n}{2} + \dots + (-1)^n \binom{n}{n} = (1 - 1)^n = 0. \end{aligned}$$

And when $A = \emptyset$,

$$\sum_{B \subseteq A} (-1)^{|B|} = (-1)^{|A|} = (-1)^0 = 1. \quad \square$$

Lemma A.3 *If A is a finite set and $B \subseteq A$, then*

$$\sum_{\substack{C \\ B \subseteq C \subseteq A}} (-1)^{|C|} = \begin{cases} (-1)^{|A|} & \text{if } A = B, \\ 0 & \text{otherwise.} \end{cases}$$

Proof: This lemma is seen to follow from Lemma A.2 once it is noticed that

$$\sum_{\substack{C \\ B \subseteq C \subseteq A}} (-1)^{|C|} = \sum_{D \subseteq A-B} (-1)^{|B \cup D|} = (-1)^{|B|} \sum_{D \subseteq A-B} (-1)^{|D|}. \quad \square$$

Proof of Theorem A.1

We know that $Bel(A) = \sum_{B \subseteq A} m(B)$. So,

$$\begin{aligned} \sum_{B \subseteq A} (-1)^{|A-B|} Bel(B) &= (-1)^{|A|} \sum_{B \subseteq A} (-1)^{|B|} Bel(B) \\ &= (-1)^{|A|} \sum_{B \subseteq A} (-1)^{|B|} \sum_{C \subseteq B} m(C) \\ &= (-1)^{|A|} \sum_{C \subseteq A} m(C) \sum_{\substack{B \\ C \subseteq B \subseteq A}} (-1)^{|B|} \\ &= (-1)^{|A|} m(A) (-1)^{|A|} \\ &= m(A). \quad \square \end{aligned}$$

APPENDIX B. The nestedness requirement of focal elements for a possibility measure

The requirement that focal elements be nested restricts belief and plausibility measures to necessity and possibility measures. Theorem B.1 can be found in Chapter 7 of [28].

Theorem B.1 *Let a given finite body of evidence (\mathcal{F}, m) be nested. Then, the associated belief and plausibility measures have the following properties for all $A, B \in \mathcal{P}(U)$:*

(a) $Bel(A \cap B) = \min [Bel(A), Bel(B)];$

(b) $Pl(A \cup B) = \max [Pl(A), Pl(B)].$

Proof:

(a) Since the focal elements in (F) are nested, they may be linearly ordered by the subset relationship. Let $\mathcal{F} = \{A_1, A_2, \dots, A_n\}$, and assume that $A_i \subset A_j$ whenever $i < j$. Consider now arbitrary subsets A and B of U . Let i_1 be the largest integer i such that $A_i \subseteq A$, and let i_2 be the largest integer i such that $A_i \subseteq B$. Then, $A_i \subseteq A$ and $A_i \subseteq B$ if and only if $i \leq i_1$ and $i \leq i_2$, respectively. Moreover, $A_i \subseteq A \cap B$ if and only if $i \leq \min(i_1, i_2)$. Hence,

$$\begin{aligned} Bel(A \cap B) &= \sum_{i=1}^{\min(i_1, i_2)} m(A_i) = \min \left[\sum_{i=1}^{i_1} m(A_i), \sum_{i=1}^{i_2} m(A_i) \right] \\ &= \min [Bel(A), Bel(B)]. \end{aligned}$$

(b) Assume that (a) holds. Then, using $Pl(A) = 1 - Bel(A^c)$ for all $A, B \in \mathcal{P}(U)$, we have

$$\begin{aligned}
Pl(A \cup B) &= 1 - Bel((A \cup B)^c) \\
&= 1 - Bel(A^c \cap B^c) \\
&= 1 - \min[Bel(A^c), Bel(B^c)] \\
&= \max[1 - Bel(A^c), 1 - Bel(B^c)] \\
&= \max[Pl(A), Pl(B)]. \quad \square
\end{aligned}$$

Therefore, we have the properties for possibility and necessity measures:

(a') $Nec(A \cap B) = \min[Nec(A), Nec(B)];$

(b') $Pos(A \cup B) = \max[Pos(A), Pos(B)].$

We can prove the converse statement of Theorem B.1.

Theorem B.2 *If for all $A, B \in \mathcal{P}(U)$, we have*

(a) $Bel(A \cap B) = \min[Bel(A), Bel(B)],$ or

(b) $Pl(A \cup B) = \max[Pl(A), Pl(B)],$

then the focal elements are nested.

Proof: The proof is based on having information (b). (We can derive (a) from (b)). Let $U = \{u_1, u_2, \dots, u_n\}$. Suppose that U can be partitioned into k sets for some $k \leq n$ as follows:

$$\begin{aligned}
U^0 &= \{u^0 \mid Pl(\{u^0\}) = 1\}, \\
U^1 &= \{u^1 \mid Pl(\{u^1\}) = a_1 < 1\},
\end{aligned}$$

$$\begin{aligned}
U^2 &= \{u^2 \mid Pl(\{u^2\}) = a_2 < a_1\}, \\
&\vdots \\
U^k &= \{u^k \mid Pl(\{u^k\}) = a_k < a_{k-1}\}.
\end{aligned}$$

Let $A^0 \subsetneq U^0$, and $A^0 \neq \emptyset$. For any $u \in \bigcup_{i=1}^k U^i$, we have $1 = Pl(U \setminus A^0) = \max[Pl(U \setminus (A^0 \cup \{u\})), Pl(\{u\})]$. Since $Pl(\{u\}) < 1$,

$$1 = Pl(U \setminus (A^0 \cup \{u\})) = 1 - Bel(A^0 \cup \{u\}) = 1 - \sum_{B \subseteq (A^0 \cup \{u\})} m(B) \Rightarrow \sum_{B \subseteq (A^0 \cup \{u\})} m(B) = 0.$$

Thus, for $k_i = |U^i|$, we achieve

- $m(\{u_i\}) = 0, \forall i = 1, 2, \dots, n,$
- $m(\{u_i^0, u_j^0\}) = 0, \forall i, j = 1, 2, \dots, k_0,$ and
 $m(\{u_i^0, u_j^1\}) = 0, \forall i = 1, 2, \dots, k_0, j = 1, 2, \dots, k_1,$
- $m(\{u_i^0, u_j^0, u_k^0\}) = 0, \forall i, j, k = 1, 2, \dots, k_0,$ and
 $m(\{u_i^0, u_j^0, u_k^1\}) = 0, \forall i, j = 1, 2, \dots, k_0, k = 1, 2, \dots, k_1,$
 \vdots
- $m(A_i^0) = 0,$ where $A_i^0 = U^0 \setminus \{u_i^0\}, \forall i = 1, 2, \dots, k_0,$ and
 $m((A_i^0 \setminus \{u_j^0\}) \cup \{u_k^1\}) = 0, \forall i, j = 1, 2, \dots, k_0, k = 1, 2, \dots, k_1,$
- $m(A_i^0 \cup \{u_j^1\}) = 0, \forall i = 1, 2, \dots, k_0, j = 1, 2, \dots, k_1.$

We also know that $1 > Pl(U \setminus U^0) = 1 - Bel(U^0)$. Hence, $0 < Bel(U^0) = \sum_{B \subseteq U^0} m(B) < 1$. Since $m(B) = 0, \forall B \subsetneq U^0$ (as listed above), $Bel(U^0) = m(U^0)$, and $0 < m(U^0) < 1$. We achieve the first focal element as the set U^0 .

We expect that other focal elements will contain U^0 .

$Pl(U \setminus U^0) = Pl(U^1 \cup U^2 \cup \dots \cup U^k) = 1 - m(U^0)$ means that $a_1 = Pl(U^1) = 1 - m(U^0)$. For every $A^1 \subsetneq U^1$, and $A^1 \neq \emptyset$, we have $1 - m(U^0) = a_1 = Pl(A^1) = 1 - Bel(U \setminus A^1)$, which implies that $m(U^0) = Bel(U \setminus A^1) = \sum_{B \subseteq (U \setminus A^1)} m(B)$. Hence, we have $m(B) = 0, \forall B \neq U^0$, such that $B \subseteq (U \setminus A^1), \forall A^1 \subsetneq U^1, A^1 \neq \emptyset$. We also know that $1 - m(U^0) > Pl(U \setminus (U^0 \cup U^1)) = 1 - Bel(U^0 \cup U^1) \Rightarrow m(U^0) < Bel(U^0 \cup U^1) = \sum_{B \subseteq (U^0 \cup U^1)} m(B)$. But, $m(B) = 0, \forall B \subseteq U^0 \cup U^1, B \neq U^0, B \neq U^0 \cup U^1$. Therefore, $m(U^0 \cup U^1) > 0$. The second focal element is the set $U^0 \cup U^1$. $Bel(U^0 \cup U^1) = m(U^0) + m(U^0 \cup U^1)$.

The next focal element is $\sum_{i=0}^2 U^i$. Here is the reason. $a_2 = Pl(U \setminus (U^0 \cup U^1)) = 1 - Bel(U^0 \cup U^1) = 1 - m(U^0) - m(U^0 \cup U^1)$ For every $A^2 \subsetneq U^2$, and $A^2 \neq \emptyset$, we have $1 - m(U^0) - m(U^0 \cup U^1) = a_2 = Pl(A^2) = 1 - Bel(U \setminus A^2)$, which implies that $m(U^0) + m(U^0 \cup U^1) = Bel(U \setminus A^2) = \sum_{B \subseteq (U \setminus A^2)} m(B)$. Hence, we have $m(B) = 0, \forall B \neq U^0, B \neq (U^0 \cup U^1)$ such that $B \subseteq (U \setminus A^2), \forall A^2 \subsetneq U^2, A^2 \neq \emptyset$. We also know that $1 - m(U^0) - m(U^0 \cup U^1) > Pl(U \setminus (U^0 \cup U^1 \cup U^2)) = 1 - Bel(U^0 \cup U^1 \cup U^2) \Rightarrow m(U^0) + m(U^0 \cup U^1) < Bel(U^0 \cup U^1 \cup U^2) = \sum_{B \subseteq (U^0 \cup U^1 \cup U^2)} m(B)$. But, $m(B) = 0, \forall B \subseteq U^0 \cup U^1 \cup U^2, B \neq U^0, B \neq U^0 \cup U^1, B \neq \sum_{i=0}^2 U^i$. Therefore, $m(U^0 \cup U^1 \cup U^2) > 0$. The third focal element is the set $U^0 \cup U^1 \cup U^2$. $Bel(U^0 \cup U^1 \cup U^2) = m(U^0) + m(U^0 \cup U^1) + m(U^1 \cup U^1 \cup U^2)$.

Follow the similar pattern as in the previous two paragraphs, we can conclude that

$$\mathcal{F} = \left\{ U^0, \sum_{i=0}^1 U^i, \sum_{i=0}^2 U^i, \dots, \sum_{i=0}^k U^i = U \right\}. \quad \square$$

REFERENCES

- [1] H. Aissi, C. Bazgan, and D. Vanderpooten. Min-max and min-max regret versions of combinatorial optimization problems: A survey. *European Journal of Operational Research*, 197(2):427–238, 2009.
- [2] R. Banuelos. *Measure Theory and Probability*. Department of Mathematics, Purdue University, 2003. Lecture notes.
- [3] M.S. Bazaraa, J. J. Jarvis, and H. D. Sherali. *Linear Programming and Network Flows*. John Wiley & Sons, Inc, Canada, 2 edition, 1990.
- [4] R. E. Bellman and L. A. Zadeh. Decision-making in a fuzzy environment. *Management Sci.*, Serial B 17:141–164, 1970.
- [5] J. R. Brige and F. Louveaux. *Introduction to Stochastic Programming*. Springer, 1997.
- [6] C. Carlsson and R. Fullér. On possibilistic mean value and variance of fuzzy numbers. *Fuzzy Sets and Systems*, (122):315–326, 2001.
- [7] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, 1963.
- [8] L. M. de Campos, J. F. Huete, and S. Moral. Probability intervals: A tool for uncertain reasoning. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 2(2):167–196, 1994.
- [9] A. P. Dempster. Upper and lower probability induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339, 1967.
- [10] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations: I. generalized p-boxes. *International Journal of Approximate Reasoning*, 49:649–663, 2008.
- [11] S. Destercke, D. Dubois, and E. Chojnacki. Unifying practical uncertainty representations: II. clouds. *International Journal of Approximate Reasoning*, 49:664–677, 2008.

- [12] D. Dubois and H. Prade. Systems of linear fuzzy constraints. *Fuzzy Sets and Systems*, (3):37–48, 1980.
- [13] Didier Dubois and Henri Prade. Random sets and fuzzy interval analysis. *Fuzzy Sets and Systems*, (42):87–101, 1991.
- [14] J. A. Dye, J. E. Rodgers, R. K. Wu, P. J. Biggs, P. H. McGinley, and R. M. McCall. *Recommendations of the National Council on Radiation Protection and Measurements, Structural Shielding Design and Evaluation for Megavoltage X- and Gamma-Ray Radio-therapy Facilities*. NCRP Report No. 151, Bethesda, MD, 2005.
- [15] L. Dymowa and M. Dolata. The transportation problem under probabilistic and fuzzy uncertainties. http://zsiie.icis.pcz.pl/artykuly/md/md_wrodaw.pdf.
- [16] Th. Fetz and M. Oberguggenberger. Propagation of uncertainty through multivariate functions in the framework of sets of probability measures. *Reliability Engineering & System Safety*, (85):73–87, 2004.
- [17] M. Fiedler, J. Nedoma, J. Ramik, J. Rohn, and K. Zimmermann. *Linear Optimization Problems with Inexact Data*. Springer, 2006.
- [18] T. Gal and J. Nedoma. Multiparametric linear programming. *Management Science*, (18):406–422, 1972.
- [19] D. M. Gay. Solving interval linear equations. *SIAM J. Number Anal.*, 19(4):858–870, 1982.
- [20] M. Inuiguchi, H. Ichihashi, and H. Tanaka. Stochastic versus fuzzy approaches to multi-objective mathematical programming under uncertainty. In R. Slowinski and J. Teghem, editors, *Fuzzy Programming: A Survey of Recent Developments*, pages 45–68. Kluwer Academic Publishers, 1990.
- [21] M. Inuiguchi and M. Sawaka. Minimax regret solution to linear programming problems with an interval objective function. *European Journal of Operational Research*, (86):526–536, 1995.
- [22] M. Inuiguchi and M. Sawaka. Minimax regret analysis in linear programs with an interval objective function. In *Proceedings of IWSCI'96*, pages 308–317. IEEE, 1996.

- [23] M. Inuiguchi, T. Tanino, and M. Sakawa. Membership function elicitation in possibilistic programming problems. *Fuzzy Sets and Systems*, (111):29–45, 2000.
- [24] K. D. Jamison. *Modeling uncertainty using probabilistic based possibility theory with applications to optimization*. PhD thesis, Department of Mathematics, University of Colorado at Denver, Denver, CO, 1998.
- [25] K.D. Jamison and W. A. Lodwick. The construction of consistent possibility and necessity measures. *Fuzzy Set and Systems*, 132:1–10, 2002.
- [26] P. Kall and S. W. Wallace. *Stochastic Programming*. John Wiley & Sons, Inc, 1994.
- [27] G. J. Klir. Is there more to uncertainty than some probability theorists might have us believe? *Int. J. General Systems*, 15:347–378, 1989.
- [28] G.J. Klir and Bo Yuan. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentic Hall P T R, New Jersey, 1995.
- [29] R. Krawczyk. Interval mathematics. In *Lecture Notes in Computer Science 29*, pages 215–222. Springer Verlag, New York, 1975.
- [30] K. K. Lai et al. A class of linear interval programming problems and its application to portfolio selection. *IEEE Transactions on Fuzzy Systems*, 10(6):698–703, 2002.
- [31] J. F. Lemmer and H. E. Kyburg Jr. Conditions for the existence of belief functions corresponding to intervals of belief. In *AAAI-91 Proceedings*, pages 488–493. AAAI, 1991.
- [32] Y. P. Li, G. H. Huang, H. N. Xiao, and X.S. Qin. Municipal solid waste management under uncertainty: an interval-fuzzy two-stage stochastic programming approach. *Environmental Informatics Archives*, 5:131–145, 2007.
- [33] Baoding Liu. *Uncertain Programming*. John Wiley & Sons, Inc, New York, 2000.
- [34] W. A. Lodwick. Analysis of structure in fuzzy linear programs. *Fuzzy Sets and Systems*, (38):15–26, 1990.
- [35] W. A. Lodwick and K. A. Bachman. Solving large-scale fuzzy and possibilistic optimization problems. *Fuzzy Optimization and Decision Making*, 4:258–275, 2005.

- [36] W. A. Lodwick and K. D. Jamison. Interval valued probability in the analysis of problems containing a mixture of possibilistic, probabilistic and interval uncertainty. In *Fuzzy Information Processing Society, 2006. NAFIPS 2006. Annual Meeting of the North American*, pages 617–620. IEEE, 2006.
- [37] W. A. Lodwick and K. D. Jamison. Theoretical and semantic distinctions of fuzzy, possibilistic, and mixed fuzzy/possibilistic optimization. *Fuzzy Sets and Systems*, (158):1861–1872, 2007.
- [38] W. A. Lodwick and K. D. Jamison. The use of interval-valued probability measures in optimization under uncertainty for problems containing a mixture of fuzzy, possibility and interval uncertainty. In *Proceedings of the 12th International Fuzzy Systems Association World Congress on Foundation of Fuzzy Logic and Soft Computing*, Lecture Notes In Artificial Intelligence, pages 361–370, 2007.
- [39] W. A. Lodwick and Elizabeth Untiedt. *Introduction to Fuzzy and Possibilistic Optimization*. Department of Mathematics, University of Colorado at Denver, September 9, 2009. Monograph Draft.
- [40] D. G. Luenberger. *Linear and Nonlinear Programming*. Springer, 2 edition, 2003.
- [41] Chr. Mandl. Number of operations for updating the elimination form of the basis-inverse of the revised simplex algorithm. *Computing*, 18:365–366, 1977.
- [42] I. Maqsood, G. H. Huang, and J. S. Yeomans. An interval-parameter fuzzy two-stage stochastic program for water resources management under uncertainty. *European Journal of Operational Research*, 167:208–225, 2005.
- [43] Diego Marín Andrés Alvarez. *Infinite random sets and applications in uncertainty analysis*. PhD thesis, The Fakultät für Bauingenieurwissenschaften of the Leopold-Franzens Universität Innsbruck, 2007.
- [44] H. E. Mausser. *Minimizing maximum regret for linear programs with interval objective function coefficients*. PhD thesis, Graduate School of Business, University of Colorado, Boulder, CO, 1997.
- [45] H. E. Mausser and M. Laguna. A new mixed integer formulation for the maximum regret problem. *International Transactions in Operational Research*, 5(5):389–403, 1998.

- [46] H. E. Mausser and M. Laguna. A heuristic to minimax absolute regret for linear programs with interval objective function coefficients. *European Journal of Operational Research*, (117):157–174, 1999.
- [47] P. H. McGinley. *Shielding techniques for radiation oncology facilities*. Medical Physics Pub, Madison, WI, second edition, 2002.
- [48] R. E. Moore. *Methods and Applications of Interval Analysis*. SIAM, Philadelphia, 1979.
- [49] T. Morrison and H. J. Greenberg. Robust optimization. In A. R. Ravindran, editor, *Operations Research and Management Science Handbook*, The Operations Research Series, chapter 14. CRC Press, Boca Raton, FL, 2008.
- [50] S. G. Nash and A. Sofer. *Linear and Nonlinear Programming*. The McGraw-Hill Companies, Inc., New York, 1996.
- [51] A. Neumaier. On the structure of clouds. <http://www.mat.univie.ac.at/~neum/papers.html>. 05/06/2010.
- [52] A. Neumaier. Clouds, fuzzy sets and probability intervals. *Reliable Computing*, 4(10):249–272, 2004.
- [53] F. Newman and M. Asadi-Zeydabadi. An optimization model and solution for radiation shielding design of radiotherapy treatment vaults. *American Association of Physicists in Medicine*, 35(1):171–180, 2008.
- [54] Hung T. Nguyen. *An Introduction to Random Sets*. Chapman & Hall/CRC, Boca Raton, FL, 2006.
- [55] M. Oberguggenberger. The mathematics of uncertainty: models, methods and interpretations. In *Analyzing Uncertainty in Civil Engineering*, pages 51–72. Springer Berlin Heidelberg, 2005.
- [56] M. Oberguggenberger and W. Fellin. Reliability bounds through random sets: non-parametric methods and geotechnical applications. *Computers & Structures*, 86:1093–1101, 2008.
- [57] J. Ramik and J. Rimanek. Inequality relation between fuzzy numbers and its use in fuzzy optimization. *Fuzzy Sets and Systems*, 16(2):123–138, 1985.
- [58] J. Rohn. Solvability of systems of linear interval equations. *SIAM J. Matrix Anal. Appl.*, 25(1):237–245, 2003.

- [59] Sheldon Ross. *A First Course in Probability*. Prentice Hall, Upper Saddle River, New Jersey, 1997.
- [60] G. Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, New Jersey, 1976.
- [61] G.J. Shih and R. A. S. Wangsawidjaja. Mixed fuzzy-probabilistic programming approach for multiobjective engineering optimization with random variables. *Computer & Structures*, 59(2):283–290, 1996.
- [62] K. Shimizu and E. Aiyoshi. Necessary conditions for min-max problems and algorithms by a relaxation procedure. *IEEE Transactions on Automatic Control*, AC-25(1):62–66, February 1980.
- [63] R. E. Steuer. Algorithms for linear programming problems with interval objective function coefficients. *Mathematics of Operations Research*, 6(3):333–348, August 1981.
- [64] H. Tanaka and K. Asai. Fuzzy linear programming with fuzzy numbers. *Fuzzy Sets and Systems*, (13):1–10, 1984.
- [65] H. Tanaka, H. Ichahashi, and K. Asai. A formulation of fuzzy linear programming problem based on comparison of fuzzy numbers. *Control and Cybernetics*, 3(13), 1984.
- [66] H. Tanaka, H. Ichihashi, and K. Asai. Fuzzy decision in linear programming with trapezoid fuzzy parameters. In J. Kacprzyk and R. R. Yager, editors, *Management Decision Support Systems Using Fuzzy Sets and Possibility Theory*. Verlag TUV, Koln, 1985.
- [67] H. Tanaka, T. Okuda, and K. Asia. On fuzzy mathematical programming. *Trans. Soc. Instrum. Control Engineers*, 5(9):607–613, 1973. (In Japanese).
- [68] H. Tanaka, T. Okuda, and K. Asia. On fuzzy mathematical programming. *J. Cybernetics*, (3):34–46, 1974.
- [69] P. Thipwiwatpotjana and W. A. Lodwick. Algorithm for solving optimization problems using interval valued probability measure. In *Fuzzy Information Processing Society, 2008. NAFIPS 2008. Annual Meeting of the North American*. IEEE, 2008.
- [70] E. A. Untiedt. Fuzzy and possibilistic programming techniques in the radiation therapy problem: an implementation-based analysis. Master’s thesis,

Department of Mathematics, University of Colorado at Denver, Denver, CO, 2006.

- [71] S. Vajda. *Probabilistic Programming*. Academic Press, New York, 1972.
- [72] C. van de Panne. A node method for multiparametric linear programming. *Management Science*, (21):1014–1020, 1975.
- [73] P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall, London, 1991.
- [74] Z. Wang and G. J. Klir. *Fuzzy Measure Theory*. Plenum Press, New York, 1992.
- [75] K. Weichselberger. The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning*, (24):149–170, 2000.
- [76] R. R. Yager. A procedure of ordering fuzzy subsets of the unit interval. *Information Sciences*, (24):143–161, 1981.
- [77] P. L. Yu and M. Zeleny. Linear multiparametric programming by multicriteria simplex method. *Management Science*, (23):150–170, 1976.
- [78] H. Zimmermann. Description and optimization of fuzzy systems. *Internat. J. Gen. Systems*, (2):209–215, 1976.