Working Paper No. MS-1-06

# What's Wrong with Info-Gap? An Operations Research Perspective*

Moshe Sniedovich

Department of Mathematics and Statistics

The University of Melbourne

Melbourne, Vic 3010, Australia

moshe@ms.unimelb.edu.au

www.ms.unimelb.edu.au/∼moshe

December 31, 2006

Last update: September 21, 2007

**Abstract**

*Info-Gap* is purported to be a new theory for decision-making under severe uncertainty. Its claim to fame is that it is non-probabilistic in nature and thus offers an alternative to all current theories for decision-making under uncertainty. In this essay I examine this theory from an Operations Research point of view. I show that:

- *Info-Gap* is neither new nor radically different from current decision theories. Specifically, I formally prove that *Info-Gap*'s decision theoretic model is a simple application of *Wald's Maximin Principle*, the most celebrated Principle in decision-making under strict uncertainty.
- *Info-Gap*'s uncertainty model is fundamentally flawed and consequently there is no reason to believe that the solutions it generates are likely to be robust.

Apparently *Info-Gap* followers are unaware of the huge discrepancy (gap?) between what *Info-Gap* claims to be and do and what it actually is and does.

The lesson for Operations Research is that it cannot take it for granted that its established methods and techniques stand ready to be adapted into other disciplines.

**Keywords:** voodoo decision-making, severe uncertainty, maximin, worst-case analysis, robust optimization, Pareto tradeoffs, non-probabilistic methods, info-gap.

*This paper was presented at the 2006 ASOR Recent Advances mini-conference, December 1, 2006, Melbourne, Australia.

# Contents

# 1 Prologue

On several occasions over the past three years I have been asked to express my views on *Info-Gap*. The reactions to my position have surprised me, and I find myself in a situation where I need to devote a considerable effort to explain this position.

The publication this year of the second edition of the *Info-Gap* book (Ben-Haim [2006]) convinced me that I should make these views more accessible to *Info-Gap* enthusiasts and to the public at large.

So here is what I have to say.

I presented this essay under the titled *What exactly is Info-Gap? An Operational Research Perspective* at the *2006 ASOR Recent Advances* mini-conference (Dec 1, 2006, Melbourne Australia).

My plan is to ultimately integrate it into a short book entitled *"Worst-Case Analysis for decision-making Under Severe Uncertainty"*. When ready, I'll put it on my website so that I can update it from time to time. This will be an on-going project. The URL of this book and related resources is

`www.ms.unimelb.edu.au/~moshe/frame_maximin.html`

The site is now open to the public. If everything goes according to plan, the book should be ready before the end of 2007.

The target audience of this mini-campaign is mainly the group of dedicated *Info-Gap* devotees out there who actually believe that *Info-Gap* is cool. My objective is not to convert them – God forbid – to another *decision-making Religion*.

I realize that this is a mission impossible.

My objective is much less ambitious: I believe that the *Info-Gap* community badly needs a short and focused introspection session, the aim of which is two-fold:

- To look behind the *Info-Gap* jargon and highfalutin declarations with a view to identify what *Info-Gap* in fact does.

- To examine how *Info-Gap*'s way of doing things is related to other methods and techniques used for the modeling, analysis and solution of problems involving decision-making under severe uncertainty.

I argue that this is long overdue.

Now, before I express my thoughts on *Info-Gap*, it is important that I tell you something about ... myself, namely my involvement and interest in *Info-Gap*.

The thing is that I am not interested in *Info-Gap* as such. I have my hands full with other projects and I do not plan at this stage to add *Info-Gap* to my list of research/teaching interests. But one of those other projects that I am working on is the promotion of *Operations Research* (OR) as an academic discipline and a profession, both locally in Australia and internationally. To clarify: OR is not my religion – it is merely my profession.

Why am I telling you these very personal things?

I am telling you this because in my view the *Info-Gap* literature exhibits serious misconceptions about the state of the art in *Operations Research* and related areas of applied mathematics. What is more disturbing though is that this literature also exhibits serious misconceptions about the state of the art in CLASSICAL DECISION THEORY proper.

It is important, however, that you do not misconstrue this fact. As I indicated above, I am not trying to convert you here and now to a new *decision-making Religion* or to give the old theory a second chance. To the contrary, my message is that if you carefully examine what *Info-Gap* actually does – in contrast to what it claims it does – you very quickly discover that . . . there is nothing new under the sun. So, no conversion is necessary.

What is necessary though is to appreciate the body of knowledge that classical decision theory already offers us. There is no need to re-invent the wheel.

To repeat, in my view, the most disturbing aspect of *Info-Gap* is its failure to appreciate the fact that classical decision theory already offers a probabilistic-free approach to decision-making under severe uncertainty.

Another difficulty with *Info-Gap* is that it is self contradictory: it advises us to refrain from doing precisely that which its generic recipe instructs us to do.

**Time Out!**

I have been asked on a number of occasions the following intriguing questions:

**FAQ 1** *How do you explain the fact that all those things have not been discovered so far? How do you explain the fact that this theory has already been embraced by so many people?"*

These are very good questions and I shall be delighted to discuss them with you over a cup of coffee. However, I shall not address them in this essay. The only thing I can say here is that the facts that I bring to light speak for themselves.

The objective of this longish essay is to give you, dear reader, something to think about. As I indicated above, I plan to put a copy of a book on this topic on my website so you should regard this essay as a pre-release trailer of the full monty.

For the benefit of readers who are too busy to read this essay in full, here is a summary of what is on the agenda in this discussion:

| | |
|---|---|
| Section 1. | Prologue<br>This section. |
| Section 2. | Worst-Case Analysis<br>Informal discussion on the "play it safe" approach to decision-making under uncertainty. |
| Section 3. | decision-making Under Strict Uncertainty<br>Description of the generic problem associated with decision-making under strict uncertainty. |
| Section 4. | What is *Info-Gap*?<br>Explanation of *Info-Gap*'s prescription for tackling decision-making situations subject to severe uncertainty. A description of its generic model and a discussion of a number of apparent difficulties with this theory. |

Section 5. Things that are definitely wrong in *Info-Gap*
Two things are mentioned that are definitely not-right, namely wrong, in *Info-Gap*. The first is that rather than being a new theory, *Info-Gap* is an instance of the more than 60 year old *Wald's Maximin Principle*. The second is that *Info-Gap*'s uncertainty model is fundamentally flawed so that there is no ground to believe that the solutions generated by *Info-Gap* are likely to be robust.

Section 6. Wald's Maximin Principle
A formal look at this stalwart of decision theory.

Section 7. Maximin vs Info-Gap
A formal proof that the generic *Info-Gap* model is an instance of the generic *Maximin* model.

Section 8. Practice What You Preach!
Explanation outlining the grounds for doubting the robustness of the solutions generated by *Info-Gap* and a reminder that by employing a single point estimate of the unknown true value of a parameter *Info-Gap* practices the opposite of what it instructs us against.

Section 9. Satisficing vs Optimizing
A formal proof is provided that any satisficing problem can be easily transformed into an equivalent optimization problem, hence *Info-Gap*'s "satisficing is better than optimizing" campaign is counter productive.

Section 10. Probabilistic vs Non-probabilistic Models
A reminder that any uncertainty model based on *Wald's Maximin Principle* is, at bottom, probabilistic in nature. Therefore it is not surprising at all that *Info-Gap*'s model of uncertainty has an intuitively simple probabilistic interpretation and formal representation.

Section 11. Discussion
A summary of the implications of the conclusions from the preceding sections.

Section 12. Bibliographical Notes
A short discussion on the references cited in this essay.

Section 13. Epilogue
Wrap-up of the discussion with some conclusions regarding *Info-Gap* and the lessons the OR community should draw from it.

Appendix A. The Art and Science of Worst-case Analysis
A formal analysis of the claims that *Info-Gap* is not *Maximin*, showing that these claims are based on misconceptions about worst-case analysis and its application in *Info-Gap*.

Enjoy the tour.

# 2 Worst-Case Analysis

Whether we like it or not, *worst-case analysis* is an intuitive and important conceptual decision-making framework. It is definitely one of the most basic and well-studied tools of classical decision theory. And if you are in the business of decision-making under *severe uncertainty*, you cannot leave home without it.

But *worst-case analysis* is applicable outside decision theory as well. For example, this kind of analysis dominates areas such as *computational complexity theory* and *analysis of algorithms* (see Cormen at al [2001] for details).

Of course the origins of this basic idea predate modern decision theory. As noted by Rustem and Howe (2002):

> The gods to-day stand friendly, that we may,
> Lovers of peace, lead on our days to age!
> But, since the affairs of men rest still incertain,
> Let's reason with the worst that may befall.
>
> William Shakespeare (1564 - 1616)
> Julius Caesar, Act 5, Scene 1

In less poetic terms, this basic concept is encapsulated in the popular adage *When in doubt, assume the worst!*

Here is the first section of the page generated by WIKIPEDIA for the search of the phrase *worst-case analysis*:

> In computer science, best, worst and average cases of a given algorithm express what the resource usage is at least, at most and on average, respectively. Usually the resource being considered is running time, but it could also be memory or other resource.
>
> In real-time computing, the worst case execution time is often of particular concern since it is important to know how much time might be needed in the worst case to guarantee that the algorithm would always finish on time.
>
> Average performance and worst-case performance are the most used in algorithm analysis. Less widely found is best-case performance, but it does have uses, for example knowing the best cases of individual tasks can be used to improve accuracy of an overall worst-case analysis. Computer scientists use probabilistic analysis techniques, especially expected value, to determine expected average running times.
>
> `http://en.wikipedia.org/wiki/Worst_case_analysis`

The following is the complete abstract of a paper entitled *Specifying design conservatism: Worst case versus probabilistic analysis*:

> Design conservatism is the difference between specified and required performance, and is introduced when uncertainty is present. The classical approach of worst-case analysis for specifying design conservatism is presented, along with the modern approach of probabilistic analysis. The appropriate degree of design conservatism is a tradeoff between the required resources and the probability and consequences of a failure. A probabilistic analysis properly models this tradeoff, while a worst-case

analysis reveals nothing about the probability of failure, and can significantly overstate the consequences of failure. Two aerospace examples will be presented that illustrate problems that can arise with a worst-case analysis.

Miles [1993]
`http://adsabs.harvard.edu/abs/1993gdss.proc..703M`

And here is a short quote from the NASA Office of Logic Design website:

> Goals
>
> * Detailed design review and worst-case analysis are the best tools for ensuring mission success.
> * The goal here is not to make more work for the designer, but to:
>   - Enhance efficiency of reviews
>   - Make proof of design more clear
>   - Make design more transferable
>   - Improve design quality
>
> `http://klabs.org/DEI/References/design_guidelines/design_analysis_test_guides.htm`

In short, *worst-case analysis* figures prominently wherever truly tough decisions are to be made in the face of severe uncertainty.

It is therefore puzzling that the *Info-Gap* books keep mum on this stalwart of classical decision theory.

Perhaps the explanation for this omission is that *Info-Gap* actually holds that in the framework of its uncertainty model there is no worst case at all:

> It is important to emphasize that the robustness $\tilde{h}(R, c)$ is *not* a minimax algorithm. In minimax robustness analysis, one *minimizes* the *maximum* adversity. This is not what info-gap robustness does. There is no maximal adversity in an info-gap model of uncertainty: the worst case at any horizon of uncertainty $h$ is less damaging than some realization at a greater horizon of uncertainty. Since the horizon of uncertainty is unbounded, there is no worst case and the info-gap analysis cannot and does not purport to ameliorate a worst case.
>
> Ben-Haim [2005, p. 392]

What would you say then, dear reader, if I tell you (see Appendix A) that the argument

"...Since the horizon of uncertainty is unbounded, there is no worst case ..."

is not only misleading but downright absurd?

And what would you say, dear reader, if I tell you that when you clear away the fog so as to enable you to make sense of what is actually going on there, you'll find that *Info-Gap* is no more and no less than ... *worst-case analysis* a lá *Maximin*?

This, no doubt, may greatly disappoint you, as it did in my case, when I discovered this fact for the first time. I was expecting better from a theory that presents itself as new and radically different from all current decision theories.

The sad thing is that *Info-Gap* itself is not aware of this fact.

The end result is that you do not find any reference, hint, suggestion or indication in the official *Info-Gap* literature (Beb-Haim [2001, 2006] ) that this supposedly new theory is basically … *worst-case analysis*. But this, dear reader, is all that *Info-Gap* is.

The trouble is that in view of its numerous declarations in the abstract, revealing the true nature of *Info-Gap* is not an easy job. The job is further complicated by the fact that although the notion *worst-case analysis* is simple and intuitive, contrary to its appearance, it is highly flexible from a mathematical modeling point of view. So if you do not have first hand experience with it you may underestimate what it can do for you as a modeling paradigm.

In case you are completely unfamiliar with *worst-case analysis*, here is how it is described by the philosopher John Rawls in the discussion on his theory of justice:

> The maximin rule tells us to rank alternatives by their worst possible outcomes: we are to adopt the alternative the worst outcome of which is superior to the worst outcome of the others.
>
> Rawls [1971, p. 152]

Of course *Maximin* is a fancy technical term used to describe the most classical type of *worst-case analysis*. It reflects the fact that the duo *max* and *min* appear next to each other in the problem statement.

Such problems were studied extensively in the late 1920s by von Neumann [1944] in connection with his work on *Game Theory*, more specifically zero-sum 2-person games.

With a stroke of imagination, Wald [1945a] introduced this notion into decision problems under uncertainty by regarding *Nature* as the second, antagonistic player, representing *uncertainty*.

We shall discuss the relevant aspects of Maximin theory in due course.

In preparation for our formal analysis of *Info-Gap*, we quickly examine in this section three examples illustrating *worst-case analysis* in action. The first example is straightforward, in fact naive. The second is a bit more subtle, and the third represents the type of worst-case analysis associated with classical *Maximin* problems.

The objective of this preliminary discussion is to pave the way for an analysis of *decision-making under severe uncertainty*.

However, as we shall see, the scope of operation of *worst-case analysis* and *Maximin* formulations goes beyond decision-making under uncertainty. The key term here is *variability* rather than *uncertainty*, observing that variability is not necessarily triggered by uncertainty.

And now to the first, very simple, textbook example of a *worst-case analysis*.

## 2.1 Example

Suppose that one fine morning you find a note and four envelopes on your doorstep. The full text of the note is displayed below. Table 1 depicts the information Joe provided on the four envelopes.

> Good morning Sir/Madam:
>
> I left on your doorstep four envelopes. Each contains some money. You are welcome to open any one of these envelopes and keep the money you find there.
>
> Please note that as soon as you open an envelope the other three will automatically self-destruct, so think carefully about which of these envelopes you should open.
>
> To assist you with your decision, I printed on each envelope the possible values of the amount of money (in Australian dollars) you may find in it. The amount that is actually there is equal to one of these figures.
>
> Unfortunately the entire project is under severe uncertainty so I cannot tell you more than this.
>
> Good luck!
>
> Joe.

Full text of Joe's Note

| Envelope | *Possible Amounts (Australian dollars)* |
|---|---|
| $E1$ | $20, 10, 300, 786$ |
| $E2$ | $2, 4000000, 102349, 500000000, 99999999, 56435432$ |
| $E3$ | $201, 202$ |
| $E4$ | $200$ |

Table 1: Easy Problem

So what do you do Dear Sir/Madam? Which envelope should you open?

Let us see what will happen if we decide to resolve this dilemma via *worst-case analysis*. Observe then that conceptually this analysis involves two steps:

- Determine the worst outcome for each of the decisions available to us.
- Select the decision whose worst outcome is best.

Here is how we organize in a tabular form the data and the results for our little problem:

| *Envelope* | *Possible Amounts* | *Worst outcome* | *Best worst outcome* |
|---|---|---|---|
| $E1$ | $20, \boxed{10}, 300, 786$ | 10 | |
| $E2$ | $\boxed{2}, 4000000, 10234$ | 2 | |
| $E3$ | $\boxed{201}, 202$ | $\boxed{201}$ | $\boxed{201}$ |
| $E4$ | $\boxed{200}$ | 200 | |

According to this back-of-the-envelope analysis, we should open $E3$. Our analysis indicates that the worst that can happen in this case is that we shall find 201 Australian dollars in the envelope. This is better than the worst case outcomes associated with the other three envelopes.

## 2.2 Example

You plan to buy a present for you dog Rex, a beautiful 7 year old German shepherd. This year you decided to buy him an educational game. There are two brands in your local pet shop, **Charisma** and **Agility**. The manuals for these games provide the operating charts shown in Figure 1. The games are suitable only for dogs whose $BI$ and $IQ$ scores are within the shaded areas on the charts[1].

The question is: which brand should you buy for Rex – **Charisma** or **Agility**?

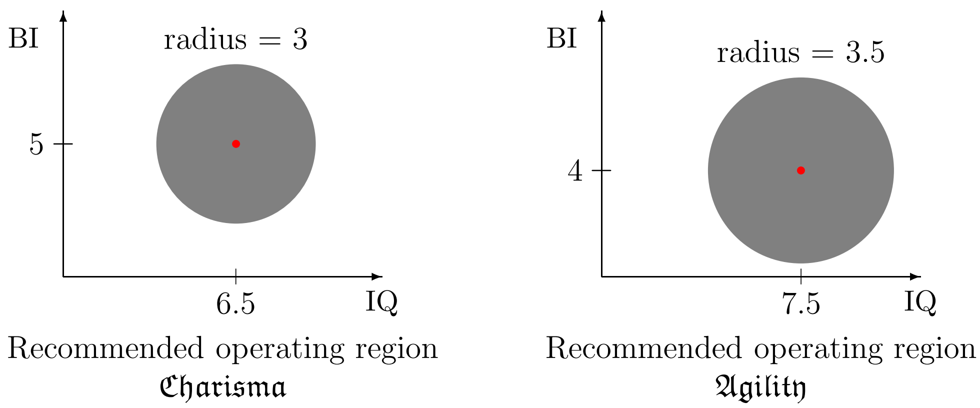


Figure 1: Operating Charts for the two brands

We shall now discuss briefly four versions of this problem. Three are associated with different levels of *uncertainty* pertaining to Rex's $BI$ and $IQ$ scores, and one is totally unrelated to uncertainty.

**Version 1:** Certainty.

[1] $BI$ is short for *Barking Index*.

Here we assume that we have the exact scores for Rex. The choice seems to be obvious: we can choose any brand as long as Rex's scores are within the specified operating region of the brand.

**Version 2:** Strict Uncertainty.

Suppose that we have no information at all about Rex's scores. Given the extreme level of uncertainty, it seems that the best thing to do is to go for the brand whose operating region is the largest. In our case this is the operating region of $\mathfrak{Agility}$, so it looks like this would be the best choice.

**Version 3:** Pretty good estimates.

Suppose that we do not have the exact values of Rex's scores, but we do have pretty good estimates of these scores. Let $\underline{a}$ denote the estimate of the $IQ$ score and let $\underline{b}$ denote the estimate of the $BI$ score. In this case we may wish to *play it safe* and select the brand providing the *largest* SAFE *deviation* from the estimates.

I illustrate the *worst-case analysis* graphically. You are encouraged to do it analytically on your own.

Suppose that the two estimates are $(\underline{a}, \underline{b}) = (6, 6)$. How far can we go from these estimates and still be in the operating region of a brand? To answer this question, we can draw circles centered at the point $(6, 6)$ on the charts. We increase the radius of these circles until they are not FULLY contained in the operating regions, as shown in Figure 2.

Clearly, for these estimates $\mathfrak{Charisma}$ seems to be far safer as the radius of the largest safe circle on its chart is much larger than the radius of the largest safe circle on the $\mathfrak{Agility}$ chart.

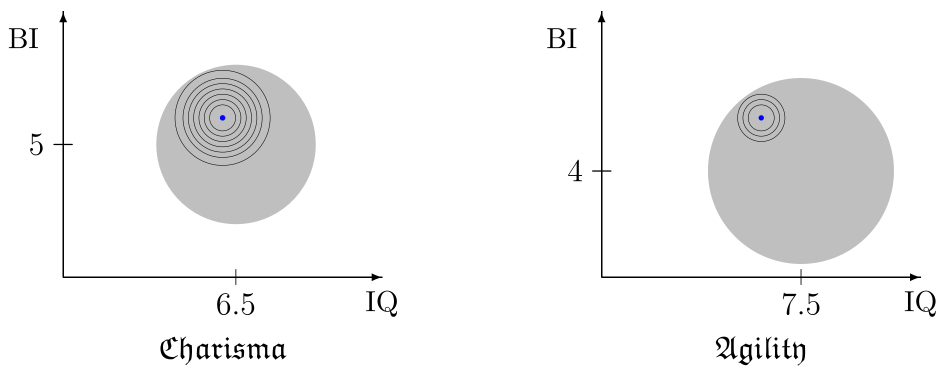


Figure 2: worst-case analysis

It is important to note that if the estimates available are not good, the analysis could be much more complicated. For example, suppose that the estimates are poor and all we know is that the true values are somewhere on the line segment connecting the two end points $(6, 6)$ and $(9, 5)$ on the $IQ/BI$ plane.

Figure 3 displays the *worst-case analysis* for these two end points. Note that $\mathfrak{Charisma}$ seems to be the better brand for the point $(6, 6)$ whereas $\mathfrak{Agility}$ seems to be the better brand for the point $(9, 5)$.

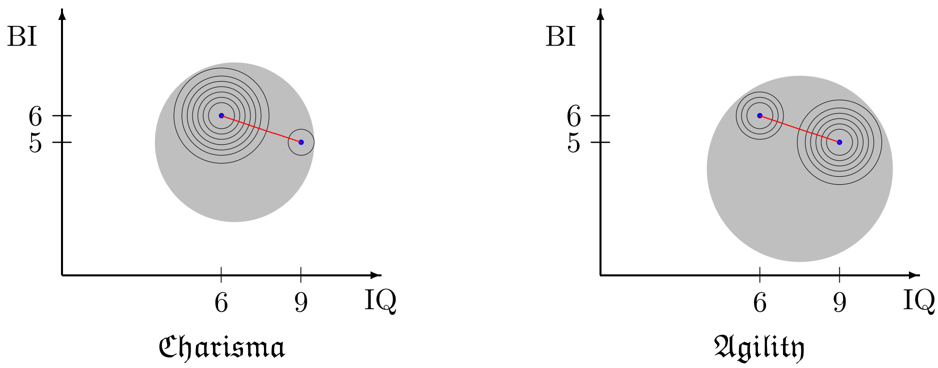


Figure 3: worst-case analysis

So, which brand should you buy for Rex?

The next version we examine is wholly unrelated to uncertainty. I discuss it so as to emphasize that the scope of *worst-case analysis* goes beyond problems dealing with uncertainty.

**Version 4:** Variability

Consider the case where before you went shopping with Rex, you called Jack the Vet and were told that Rex's scores are precisely 7 for the $IQ$ and 7 for $BI$.

What do you do in this case?

Since both brands are suitable for Rex in this case, it really makes no difference which brand you decide to buy. But, on second thought, how about ... Rex's friends? They always play with Rex's toys and games.

In view of this additional consideration, you now want a brand that will be suitable not only for Rex, but also for his friends. In short, you want a brand that will be suitable for Rex but capable of handling the largest possible variability from Rex's scores. Since Rex's friends have similar $IQ$ and $BI$ scores, you decided to conduct the *worst-case analysis* in the *immediate neighborhood* of Rex's scores on the $IQ/BI$ plane.

So formally, you are interested in the brand whose operation chart can cope (safely) with the largest deviation from the point $(7,7)$ on the $IQ/BI$ plane.

The solution generated by the *worst-case analysis* for this version of the problem is shown in Figure 4. The clear winner is no doubt **Charisma**, and so it looks like Rex and his friends will play **Charisma** for the rest of the year.

As promised by the manufactures, this should increase their $IQ$ scores and decrease their $BI$ scores.[2]

## 2.3 Example

Suppose that you want to maximize a function $f$ over a domain $Z = X \times Y$ where $X$ and $Y$ are some given sets. The difficulty is that you have control only over $x \in X$,

[2]In subsequent papers on this subject I'll report on Rex's progress on the $IQ/BI$ front.

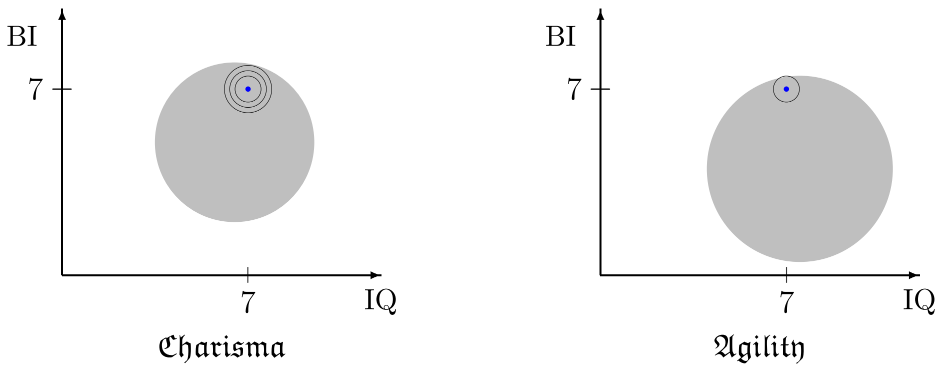


Figure 4: worst-case analysis

so if you select some $x \in X$ your payoff is $f(x, y)$ where $y$ is selected from $Y$ by a process, call it $\mathcal{P}$, over which you have no control. Furthermore, when you select $x \in X$ you do not know what value of $y$ will be selected by $\mathcal{P}$. The only thing that you do know is that $\mathcal{P}$ will select $y$ from $Y$. We therefore regard $Y$ as the *region of uncertainty*.

What should you do? What is the best $x$ in $X$ from your perspective?

Should you decide to resolve this dilemma by a *worst-case analysis*, you will assume that if you select $x \in X$, then $\mathcal{P}$ will select the worst value of $y$ in $Y$ pertaining to $x$. Since you attempt to maximize your payoff, this means that $\mathcal{P}$ will select a $y$ in $Y$ that minimizes your payoff, given that you selected $x$.

In short, under this scheme, if you select $x \in X$, your payoff will be

$$p(x) := \min_{y \in Y} f(x, y) \tag{1}$$

So your best policy is to select an $x \in X$ that maximizes $p(x)$. The recipe for such a decision, call it $x^*$, is then as follows:

$$\begin{aligned} x^* &: = \arg \max_{x \in X} \ p(x) && (2) \\ &= \arg \max_{x \in X} \ \min_{y \in Y} f(x, y) && (3) \end{aligned}$$

For instance, consider the case where $X = Y = \mathbb{R}$ and

$$f(x, y) := -(x - 4)^2 + 2xy + y^2 \tag{4}$$

where $\mathbb{R}$ denotes the real line.

Then,

$$p(x) : = \min_{y \in Y} -(x - 4)^2 + 2xy + y^2 \ , \ x \in \mathbb{R} \tag{5}$$

It is easy to verify that $y = -x$ is the minimizing $y$, hence

$$\begin{aligned} p(x) &: = \left. -(x - 4)^2 + 2xy + y^2 \right|_{y=-x} && (6) \\ &= -(x - 4)^2 + 2x(-x) + (-x)^2 && (7) \\ &= -2x^2 + 8x - 16 && (8) \end{aligned}$$

Therefore,

$$x^* = \arg \max_{x \in \mathbb{R}} -2x^2 + 8x - 16 \tag{9}$$
$$= 2 \tag{10}$$

This means that under the worst-case scenario it is best for you to set $x^* = 2$. The worst outcome (payoff) pertaining to this decision is

$$p(x^*) = -2x^2 + 8x - 16\big|_{x=2} \tag{11}$$
$$= -8 \tag{12}$$

Regarding the assertion (Ben-Haim [2005]) that no worst case exists if the region of uncertainty is unbounded, note that the region of uncertainty here, namely $Y = \mathbb{R}$, is unbounded, yet there is definitely a worst case for each decision. For more on this issue see Appendix A.

## 2.4 Mathematical formalism

There are various ways to formalize, in a rigorous mathematical manner, the basic idea behind *worst-case analysis*. We shall do it shortly.

For the time being let us recall that in the context of decision-making problems, *worst-case analysis* consists of two interrelated but distinct tasks:

· Evaluation of each alternative (decision) under the worst-case scenario pertaining to it.

· Selection of the alternative (decision) whose worst-case performance is best.

Typically, each of these two tasks involves *optimization* and, as implied by the term *Maximin*, the nature of the optimization – whether it is *maximization* or *minimization* – is reversed in these two logical steps. If one involves *maximization* then the other involves minimization and vice versa.

To explain this feature, let $\mathbb{D}$ denote the set of feasible *decisions* available to the decision maker and let $S(d)$ denotes the set of all possible *scenarios* pertaining to decision $d$. Also, suppose that the decision maker is interested in *maximizing* (with respect to $d$) a real valued function $f$ whose arguments are $d$ and $s \in S(d)$. In this case, the worst-case scenario for decision $d$ is a scenario $s^* \in S(d)$ that *minimizes* $f(d,s)$ over $S(d)$.

This means that the *worst-case analysis* will assign to decision $d \in \mathbb{D}$ the following payoff:

$$v(d) := \min_{s \in S(d)} f(d,s) \ , \ d \in \mathbb{D} \tag{13}$$

It is the smallest payoff for decision $d$ with respect to the set of scenarios associated with it, $S(d)$. In this framework the best decision is one whose $v(d)$ value is the largest. This setup leads to the *Maximin Rule* mentioned in the quote from Rawls [1971].

If the decision maker is interested in *minimizing* the objective function $f$, then the worst-case scenario for decision $d$ would be a scenario $s^*$ that *maximizes* $f(d,s)$ over $S(d)$. Hence, the *worst-case analysis* will assign to decision $d \in \mathbb{D}$ the following payoff:

$$w(d) := \max_{s \in S(d)} f(d,s) \ , \ d \in \mathbb{D} \tag{14}$$

In this case the best decision is one whose $w(d)$ value is the smallest. This leads to the *Minimax Rule*.

We shall focus on the *Maximin* formulation, observing that *Minimax* formulation can be obtained in a similar manner simply by reversing the mode of the optimization: max is transformed to min and vice versa.

Now, according to this preliminary analysis, it looks like the process of setting up a *worst-case analysis* for a given decision-making problem is a straightforward matter. All we have to do is formulate three objects:

- A *decision space*, namely a set $\mathbb{D}$ consisting of all the decisions available to us.
- A collection of sets $(S(d), d \in \mathbb{D})$, where $S(d)$ denotes the set containing all the *scenarios* associated with decision $d$.
- An *objective function*, $f$, assigning a reward, $f(d,s)$, for each (decision,scenario) pair.

In short, for the purposes of our discussion it is convenient to formulate the generic *worst-case analysis* model as follows:

$$v^* := \max_{d \in \mathbb{D}} \ \min_{s \in S(d)} f(d,s) \tag{15}$$

In this framework $S(d)$ is interpreted as the *region of uncertainty pertaining to decision* $d$.

Experience has shown that the construction of such models for real-world decision problems under severe uncertainty is not a trivial matter. Like other *mathematical modeling* enterprises, it is often ... *easier said than done*.

As an exercise, dear reader, you may wish to set up a formal model for the *worst-case analysis* we conducted in the preceding section in connection with the Rex problem. If you have not done such things before, you may find this exercise a rather challenging mathematical modeling project.

## 2.5 Discussion

The body of knowledge available to us on *worst-case analysis* is enormous. Therefore it is important to be clear on whether a decision-making methodology we plan to use is based on this type of analysis. Furthermore, since *Wald's Maximin Principle* is the most classical paradigm for *worst-case analysis*, it is also important to identify the *Maximin* content, if any, of such a methodology.

Indeed, this is imperative in view of the gentlemen's agreement that decision-making based on *worst-case analysis* or *Maximin* is often much too conservative:

> It should be mentioned that Wald advocated the minimax principle in a tentative way and because of certain formal advantages. I am informed that he was still interested in finding a less conservative and more satisfactory principle for statistical inference.
>
> To my mind, it is somewhat doubtful if principles of this kind are really applicable in the social sciences [26]. They are without any doubt applicable in industrial applications (quality control, etc.)
>
> Tintner [1952, p. 24]

and

> It should also be remarked that the minimax principle even if it is applicable leads to an extremely conservative policy.
>
> Tintner [1952, p. 25]

The reference to "social sciences" is interesting not the least because of the role *Wald's Maximin Prilciple* plays in Rawls' [1971] *Theory of Justice*.

The following quotes eloquently summarize the on-going debate on "expected value optimization vs worst-case analysis via maximin" dilemma:

> The conventional approach to decision under uncertainty is based on expected value optimization. The main problem with this concept is that it neglects the worst-case effect of the uncertainty in favor of expected values. While acceptable in numerous instances, decisions based on expected value optimization may often need to be justified in view of the worst-case scenario. This is especially important if the decision to be made can be influenced by such uncertainty that, in the worst case, might have drastic consequences on the system being optimized. On the other hand, given an uncertain effect, some worst-case realizations might be so improbable that dwelling on them might result in unnecessarily pessimistic decisions. Nevertheless, even when decisions based on expected value optimization are to be implemented, the worst-case scenario does provide an appropriate benchmark indicating the risks.
>
> Rustem and Howe [2002, p. xiii]
>
> Through its inherent pessimism, the minimax strategy may lead to a serious deterioration of performance. Alternatively, the realization of the worst-case scenario may result in an unacceptable performance deterioration for the strategy based on expected value optimization. *Neither minima nor expected value optimization provide a substitute for wisdom.* At best, they can be regarded as risk management tools for analyzing the effects of uncertain events.
>
> Rustem and Howe [2002, p. xiv]

As we shall see, *Info-Gap* seems to be unaware of the fact that it is a *Maximin Principle* in disguise. It is not surprising therefore that the *Info-Gap* literature is totally oblivious to the wealth of knowledge available on the limitations and problematic aspects of *worst-case analysis*.

This is a pity because this very rich knowledge-base is definitely relevant to the way things are done in *Info-Gap*.

In the next section we take a quick look at decision-making under strict uncertainty. This will pave the way for a formal examination of *Info-Gap*.

# 3 Decision-making Under Strict Uncertainty

*Who is a wise man? He who sees that which will happen!*

Talmud, Tamid 32(a)

Classical decision theory distinguishes between three types of decision-making situations:

· decision-making under *certainty.*

· decision-making under *risk.*

· decision-making under *strict uncertainty.*

The first case represents situations where we pretend that no uncertainty exist at all in the decision-making situation.

The second case represents situations where the uncertainty in the decision-making situation can be described and quantified by conventional statistical and probabilistic models and/or methods.

The third case is the most difficult. Here our knowledge of the consequences of our decisions is poor to such an extent that there is precious little to work with in order to develop a solid, comprehensive, and useful decision-making methodology.

The classical model for situations of this kind is therefore very austere. It consists of three very simple ingredients:

· a DECISION SPACE, $\mathbb{D}$;

· a STATE SPACE, $\mathbb{S}$;

· a REWARD FUNCTION, $f$;

where $\mathbb{D}$ and $\mathbb{S}$ are some sets and $f$ is a real valued function on $\mathbb{D} \times \mathbb{S}$.

The conceptual model is this: you, the decision maker (DM), must select a *decision* $d \in \mathbb{D}$. In return you obtain a *reward* $f(d, s)$ whose value depends on your decision as well as on the *state of nature* $s \in \mathbb{S}$. The difficulty is that the true value of the state $s$ is under strict uncertainty, meaning that we are ignorant of the true value of $s$, hence of the true value of the reward $f(d, s)$.

So what do we do?

Needless to say, classical decision theory proposes no magic wand to tackle such situations. What it does offer is a pair of fundamental approaches, to wit PRINCIPLES, that can be considered in situations like this. Over the years these two principles have become famous, some would say infamous, because both are very problematic. In any case, here is the celebrated duo:

· **Laplace's Principle of insufficient Reason** (1825)

· **Wald's Maximin Principle** (1945)

In brief, *Laplace's Principle* suggests that if you really have no inkling as to the true state of nature, then it is reasonable to assume that all the states are *equally likely*. This means that you can regard the state variable $s$ as a *random variable* associated with a *uniform* probability distribution function over the state space $\mathbb{S}$. There are of course cases where this is impossible, eg. $\mathbb{S} = \mathbb{R}$, where $\mathbb{R}$ denotes the real line, observing that it is impossible to create a uniform probability density function on $\mathbb{R}$.

The appeal of this principle is that it transforms a difficult problem (decision-making under strict uncertainty) into a relatively "easy" problem (decision-making under risk).

The *Maximin Principle* goes much further than that: it transforms a decision-making situation under *strict uncertainty* into a decision-making situation under *certainty*.

It performs this trick by following my dear wife's dictum: *in situations under strict uncertainty, it is reasonable to assume that the* WORST *possible thing will happen.* That is, this principle assumes that MOTHER NATURE is playing *against us* in that it always selects the *least favorable* state $s \in \mathbb{S}$ pertaining to our choice of $d \in \mathbb{D}$. This, of course, is a very pessimistic view of how *Mother Nature* works.

The appeal of this principle is that it transforms a difficult problem (decision-making under strict uncertainty) into a "very easy" problem (decision-making under certainty). That is, we exploit the fact that *Mother Nature* is antagonistic to such an extent that it becomes completely predictable. This removes the uncertainty altogether and we are left with a simple deterministic problem.

In summary, these are the two basic principles offered by classical decision theory for decision-making under strict uncertainty. Of course, there are many variations on these two basic themes.

Now, since *Info-Gap* claims to be a new theory, one that is *radically different* from *all* current decision theories, we expect it to be radically different from these two principles. And since it claims to be a *probabilistic-free theory*, it is only natural to expect *info-Gap* to explain in what way it radically differs from *Wald's Maximin Principle*.

But as we shall soon see, contrary to its claim, *Info-Gap* offers no new, radical and exciting ideas. And what is more alarming is its failure to recognize its close affinity to worst-case analysis and *Wald's Maximin Principle*.

Having examined, albeit very briefly, the two basic principles provided by decision theory for decision-making under severe uncertainty, let us now examine the basic ideas behind *Info-Gap*.

Fasten your seat belts!

# 4 What is *Info-Gap*?

To reiterate, *Info-Gap* is supposedly a brand new methodology for generating robust solutions to problems depicting situations under severe uncertainty. The main references are the two editions of the book by Ben-Haim [2001, 2006].

To give you a flavor of how *Info-Gap* views itself vis-a-vis classical theories for decision-making under uncertainty, here are some quotes from the 2nd edition of the *Info-Gap* book.

Regarding its radical departure from all the classical theories:

> Info-gap decision theory is radically different from all current theories of decision under uncertainty. The difference originates in the modelling of uncertainty as an information gap rather than as a probability.
>
> Ben-Haim [2006, p.xii]

The alternative it offers to the more classical decision theories and its big leap forward:

> In this book we concentrate on the fairly new concept of information-gap uncertainty, whose differences from more classical approaches to uncertainty are real and deep. Despite the power of classical decision theories, in many areas such as engineering, economics, management, medicine and public policy, a need has arisen for a different format for decisions based on severely uncertain evidence.
>
> Ben-Haim [2006, p. 11]

Its new perspective on classical decision theories:

> The emergence of info-gap decision theory as a viable alternative to probabilistic methods helps to reconcile Knight's dichotomy between risk and uncertainty. But more than that, while info-gap models of severe lack of information serve to quantify Knight's 'unmeasurable uncertainty', they also provide new insight into risk, gambling and the entire pantheon of classical probabilistic explanada. We realize the full potential of the new theory when we see that it provides new ways of thinking about old problems.
>
> Ben-Haim [2006, p. 342]

And the follow-up:

> Conversely, the greatest difficulty in assimilating the new methodology of info-gap decision theory is presented by classical problem-formulation which are incompatible with the new approach. This incompatibility results primarily from the fact that classical methods capture only part of the phenomena of uncertainty: probabilistic theories generate probabilistic questions, while uncertainty is not exclusively probabilistic.
>
> Ben-Haim [2006, p. 342]

Got the drift?

Now, back to earth.

*Info-Gap* deals with the following abstract decision-making situation (I deliberately employ *Info-Gap*'s standard notation): You have to select a decision $q \in \mathbb{Q}$. In return you obtain a reward $R(q,u)$ that depends on $q$ but also on a parameter $u$ whose true value is unknown and subject to severe uncertainty. It is assumed that the larger the reward the better it is.

So the question is: what do you do? What is the best decision?

To answer this question *Info-Gap* introduces two additional parameters, and a family of nested sets, namely:

· A critical reward value, $r_c$.

· An estimate $\tilde{u}$ of the true value of $u$.

· A parametric family of nested sets $\mathcal{U}(\alpha,\tilde{u}) \subseteq \mathfrak{U}, \alpha \geq 0$.

The generic *Info-Gap* model based on these constructs can be formulated mathematically as follows: First we have a recipe for determining the *robustness* of any given decision $q \in \mathbb{Q}$:

$$\hat{\alpha}(q,r_c) := \max\left\{\alpha : r_c \leq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right\} \tag{16}$$

And then we are instructed that the best thing to do is to select the decision whose robustness is the largest:

$$\hat{\alpha}(r_c) := \max_{q\in\mathbb{Q}} \ \hat{\alpha}(q,\tilde{u}). \tag{17}$$

Interpret this model as follows: we must select a *decision* $q$ from a given set of feasible decisions $\mathbb{Q}$. The *reward* generated by a decision, $R(q,u)$, depends on the value of some parameter $u \in \mathfrak{U}$. In selecting $q$ we have to make sure that the reward is not below a given critical value $r_c$.

So far so good.

The difficulty is that the true value of $u$ is unknown and is subject to SEVERE uncertainty. All we have is a POOR estimate, $\tilde{u}$, of the true value of $u$, and this estimate is likely to be SUBSTANTIALLY WRONG.

So what should we do?

We create nested regions of uncertainty, $\mathcal{U}(\alpha,\tilde{u}), \alpha \geq 0$, around the estimate $\tilde{u}$ with the property that $\mathcal{U}(0,\tilde{u}) = \{\tilde{u}\}$ and $\mathcal{U}(\alpha,\tilde{u})$ is increasing in size with $\alpha$, namely $\alpha > \alpha'$ implies $\mathcal{U}(\alpha',\tilde{u}) \subseteq \mathcal{U}(\alpha,\tilde{u})$. Intuitively regard $\alpha$ as a measure of the "size" of the region $\mathcal{U}(\alpha,\tilde{u})$.

Using these regions of uncertainty we define the robustness of a decision $q$ as the largest value of $\alpha$ such that the reward requirement $r_c \leq R(q,u)$ is satisfied for all $u \in \mathcal{U}(\alpha,\tilde{u})$.

We use the robustness as the preference criterion for the selection of the best decision, namely the best decision is one whose robustness is the largest.

The notation $\hat{\alpha}(q, r_c)$ and $\hat{\alpha}(r_c)$ indicates that the critical reward $r_c$ should be viewed as a "soft" requirement. In order to determine which value should be used to decide on the best decision, it might therefore be necessary to solve the problem for a range of $r_c$ values, and then conduct a tradeoff analysis.

In short, the issue here is to determine what decision $q \in \mathbb{Q}$ gives us the best $(\tilde{\alpha}(q, r_c), r_c)$ pair. The point is that we would like both $\tilde{\alpha}(q, r_c)$ and $r_c$ to be as large as possible, but this is usually impossible because $\tilde{\alpha}(q, r_c)$ is non-increasing with $r_c$.

# 5 Things that are definitely wrong in *Info-Gap*

There are many things that are not right in *Info-Gap*, but some cry out to be rectified more than others. The list below consists of items that I classify as *definitely* wrong, to be distinguished from those that are just not right, or semi-definitely wrong.

Needless to say, this classification is inherently subjective, so feel free to declassify the items on this list and/or use your own classification.

**Things that are definitely wrong in *Info-Gap***

- W-1 *Info-Gap* has serious misconceptions about the state of the art in decision theory and optimization theory.
- W-2 For all intents and purposes *Info-Gap* is a re-invention of a simple instance of *Wald's Maximin Principle*.
- W-3 The *Info-Gap* uncertainty model is fundamentally flawed. *Info-Gap* suffers from a severe case of split personality in relation to its treatment of severe uncertainty.
- W-4 *Info-Gap* does not deal with severe uncertainty: it simply and unceremoniously ignores it.
- W-5 There is no reason to believe that the solutions generated by *Info-Gap* are likely to be robust.
- W-6 *Info-Gap*'s "satisficing is better than optimizing" crusade is outdated and pointless.
- W-7 The *Info-Gap* subject index and bibliography are severely deficient.

My advice to *Info-Gap* aficionados is that the items on this list are important and should be heeded. I discuss them in detail in my book on decision-making under strict uncertainty (Sniedovich [2007])

In any case, as we shall soon see, it is quite easy to show, in fact formally "prove", the following two things:

- The generic *Info-Gap* model is an instance of the famous *Wald's Maximin Principle*.
- The *Info-Gap* uncertainty model is fundamentally flawed.

Let us then examine the *Principle* and then turn our attention to the generic *Info-Gap* model and the way it is presented in the official *Info-Gap* literature.

# 6 Wald's Maximin Principle

In my humble opinion, every proponent of *Info-Gap* should have a *Wald's Maximin Principle* sticker on their car. If you do not have a car put it on your computer. If you do not have a computer put it on your fridge. If you do not have a fridge put it on the cover of your *Info-Gap* book.

Feel free to copy the sticker and distribute it widely.

I Love

$$v^* := \overset{You}{\max_{d\in\mathbb{D}}}\ \overset{Mama}{\min_{s\in S(d)}}\ f(d,s)$$

Wald's Maximin Principle

To refresh your memory, *Wald's Maximin Principle* [1945, 1950] is one of the most important principles in decision theory. It was inspired by von Neumann's (1928) formulation of *Maximin* problems in connection with his work on *Game Theory*.

As we already noted, this is how it is described by the philosopher John Rawls, in his famous *A Theory of Justice*:

> The maximin rule tells us to rank alternatives by their worst possible outcomes: we are to adopt the alternative the worst outcome of which is superior to the worst outcome of the others.
>
> Rawls [1971, p. 152]

And here is how the general *Maximin Principle* is described in the Free On-Line Dictionary of Philosophy (www.swif.it/foldop):

> maximin principle
> <logic, mathematics, game-theory> supposition that the preferable alternative is one whose worst outcome is least harmful, originally in mathematical and game-theoretical contexts. Thus, when success in any venture is uncertain, it is better to choose courses of action that risk the least, even if they don't offer a chance at the most. Rawls argued that this maximization of the minimum gain to be achieved is a rational guide for social decision-making. Recommended Reading: V. F. Dem'Yanov and V. N. Malozemov, Introduction to Minimax (Dover, 1990); Stephen Simons, Minimax and Monotonicity (Springer Verlag, 1999); Ronald Christensen, General Description of Entropy Minimax (Entropy, 1981); and John Rawls, A Theory of Justice (Belknap, 1999).
>
> http://www.swif.uniba.it/lei/foldop/foldoc.cgi?maximin+principle

And here is how the related *Minimax Principle* is described in the online encyclopedia WIKIPEDIA (www.wikipedia.com):

> Minimax in the face of uncertainty
> Minimax theory has been extended to decisions where there is no other player, but where the consequences of decisions depend on unknown facts. For example, deciding to prospect for minerals entails a cost which will be wasted if the minerals are not present, but will bring major rewards if they are. One approach is to treat this as a game against nature, and using a similar mindset as Murphy's law, take an approach which minimizes the maximum expected loss, using the same techniques as in the two-person zero-sum games.
>
> en.wikipedia.org/wiki/Minimax#Minimax_in_the_face_of_uncertainty

In short, the *Maximin Principle* is a major celebrity in decision theory. It is therefore not surprising that it appears in standard introductory OR/MS textbooks.

It certainly appears in my lecture notes for the subject *620-262: decision-making*, is used in my lecture notes for the subject *620-261: Introduction to Operations Research*, and is deployed here and there in my journal articles ...to get rid of uncertainty.

I therefore cannot imagine how anyone doing research in the area of decision-making under uncertainty can survive without having some knowledge of this famous mega star of decision theory.

Back to base.

The conceptual framework for the interpretation of *Wald's Maximin Principle* is as follows.

- You are required to select a *decision* $d$ from a set of feasible decisions $\mathbb{D}$.
- The *reward* generated by a decision, $f(d,s)$, depends also on the *state of nature*, $s$, whose true value is unknown, except that it lies in the state space $\mathbb{S}$, or more specifically in a subset $S(d)$ of $\mathbb{S}$.

· The question is then: what is the best decision, given that the larger the reward the better it is?

If we assume that *Mother Nature* selects the worst state associated with our decision $d$, then the reward generated by $d \in \mathbb{D}$ would be

$$v(d) := \min_{s \in S(d)} f(d,s) \tag{18}$$

We refer to $v(s)$ as the *security level* of decision $d$. This is the worst possible reward if we select decision $d$.

*Wald's Maximin Principle* argues that the best decision is one whose security level is the largest. As a result, we can find this best decision by solving the following optimization problem:

$$v^* := \max_{d \in \mathbb{D}} v(d) \tag{19}$$

$$= \max_{d \in \mathbb{D}} \min_{s \in S(d)} f(d,s) \tag{20}$$

In short, by assuming that "Mother Nature is playing against us", Wald removed completely the uncertainty regarding the true value of $s$: the true value of $s$ is (assumed to be) the worst state associated with our decision. In case of a tie between different states for this honorary title, we break it arbitrarily.

Note that the formal formulation given in (20) to *Wald's Maximin Principle* is identical to the formulation given in (15) for the *worst-case analysis*. This is not an accident: *Wald's Maximin Principle* is the "most classical" framework for *worst-case analysis*.

Now, it should be pointed out that this form of the classical *Maximin* model can be modified to reflect specific properties of the instance under consideration. In particular, if the objective function has a *composite* structure, say

$$f(d,s) = g(d, \rho(d,s)) \ , \ d \in \mathbb{D}, \ s \in S(d) \tag{21}$$

such that $g(d, \rho(d,s))$ is *non-decreasing* with $\rho(d,s)$, then we have

$$v^* := \max_{d \in \mathbb{D}} \min_{s \in S(d)} g(d, \rho(d,s)) \tag{22}$$

$$= \max_{d \in \mathbb{D}} g\left(d, \min_{s \in S(d)} \rho(d,s)\right) \tag{23}$$

In this case *Mother Nature*'s worst-case policy would be

$$\sigma(d) := \arg \min_{s \in S(d)} \rho(d,s) \tag{24}$$

In other words, in this framework *Mother Nature* is deploying $\rho$ as a proxy for the official objective function $f$. An example of this form of the *Maximin* model can be found in Sniedovich [2003].

# 7 Maximin vs *Info-Gap*

To examine the relationship between *Wald's Maximin Principle* and *Info-Gap*, it is instructive to express the generic *Info-Gap* model in a more compact form so that we can see more clearly what happens here.

So let $\preceq$ denote the binary operation defined by:

$$a \preceq b := \begin{cases} 1 & , \quad a \leq b \\ 0 & , \quad a > b \end{cases} \quad , \ a, b \in \mathbb{R} \tag{25}$$

and consider the real valued function $\varphi$ defined as follows:

$$\varphi(q, \alpha, u) := \alpha \cdot (r_c \preceq R(q, u)) \ , \ q \in \mathbb{Q}, \alpha \geq 0, u \in \mathcal{U}(\alpha, \tilde{u}) \tag{26}$$

where $\cdot$ denotes scalar multiplication.

Then, by construction $\varphi(q, \alpha, u)$ is non-decreasing with $R(q, u)$ and therefore

$$\begin{aligned}
\beta(r_c) :&= \max_{q \in \mathbb{Q}, \alpha \geq 0} \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \varphi(q, \alpha, u) && (27)\\
&= \max_{q \in \mathbb{Q}, \alpha \geq 0} \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \alpha \cdot (r_c \preceq R(q, u)) && (28)\\
&= \max_{q \in \mathbb{Q}, \alpha \geq 0} \alpha \cdot \left( r_c \preceq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right) && (29)\\
&= \max_{q \in \mathbb{Q}} \max_{\alpha \geq 0} \alpha \cdot \left( r_c \preceq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right) && (30)\\
&= \max_{q \in \mathbb{Q}} \max \left\{ \alpha : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right\} && (31)\\
&= \hat{\alpha}(r_c) && (32)
\end{aligned}$$

In other words, utilizing $\varphi$ as the objective function of the *Maximin* model we can represent the generic *Info-Gap* model compactly as follows:

$$\hat{\alpha}(r_c) := \max_{q \in \mathbb{Q}, \alpha \geq 0} \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \alpha \cdot (r_c \preceq R(q, u)) \tag{33}$$

Suffice it to say that representations of this type are common in *Operations Research* modeling and our undergraduate students are taught how to use them for such purposes.

If you have nothing else to do this weekend you might consider spending a couple of minutes examining Exhibit 1.

| *Wald's Maximin Principle* | *Generic Info-Gap Model* |
|---|---|
| $v^* := \max_{d \in \mathbb{D}} \min_{s \in S(d)} f(d, s)$ | $\hat{\alpha}(r_c) = \max_{q \in \mathbb{Q}, \alpha \geq 0} \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \alpha \cdot (r_c \preceq R(q, u))$ |

Exhibit 1

What you see here is a pair of mathematical models. One is the famous *Wald's Maximin Principle,* the other is the generic *Info-Gap* model expressed compactly in a civilized form.

So what are we to make of this?

**Theorem 1** *The generic Info-Gap model is an instance of Wald's Maximin Principle.*

**Proof.** You do not have to be an experienced mathematician to conclude that these two models are very similar. Just in case, the correspondence is spelled out for you in Table 2.

More formally, let $\mathbb{R}_+$ denote the non-negative part of the real line and consider the following instances of the basic ingredients of a *Maximin* model:

$$\mathbb{D} = \mathbb{Q} \times R_+ \tag{34}$$

$$S(q,\alpha) = \mathcal{U}(\alpha,\tilde{u}) \ , \ q \in \mathbb{Q}, \alpha \geq 0 \tag{35}$$

$$f(q,\alpha,u) = \alpha \cdot (r_c \preceq R(q,u)) \tag{36}$$

Now, consider the instance of the *Maximin* model specified by these objects:

$$\begin{aligned}
v^* := & \max_{d\in\mathbb{D}} \min_{s\in S(d)} f(d,s) && (37)\\
= & \max_{q\in\mathbb{Q},\alpha\geq 0} \min_{u\in\mathcal{U}(\alpha,\tilde{u})} f(q,\alpha,u) && (38)\\
= & \max_{q\in\mathbb{Q},\alpha\geq 0} \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \alpha \cdot (r_c \preceq R(q,u)) && (39)\\
= & \max_{q\in\mathbb{Q},\alpha\geq 0} \alpha \cdot \left( r_c \preceq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u) \right) && (40)\\
= & \max_{q\in\mathbb{Q}} \max_{\alpha\geq 0} \alpha \cdot \left( r_c \preceq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u) \right) && (41)\\
= & \max_{q\in\mathbb{Q}} \max \left\{ \alpha : r_c \leq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u) \right\} && (42)
\end{aligned}$$

And this is the end of the story. There is no danger of mistaken identity here! This is none other but the generic *Info-Gap* model. ........................ 𝔔𝔈𝔇

In words, *Info-Gap*'s generic model is an instance of *Wald's Maximin Principle* characterized by a number of specific features the most important of which is the structure of the objective function, namely

$$f(q,\alpha,u) = \alpha \cdot \big(r_c \preceq R(q,u)\big) \tag{43}$$

This representation highlights the conflict between the decision maker, who attempts to maximize the value of $\alpha$, and *Mother Nature* who attempts to minimize the value $\alpha$ by minimizing $R(q,u)$ within the region of uncertainty stipulated by $\alpha$.

It is a typical *worst-case analysis*: each decision is evaluated by the worst outcome associated with it: *Mother Nature* selects a $u$ in $\mathcal{U}(\alpha,\tilde{u})$ that minimizes $f(q,\alpha,u)$ over

| *Wald's Maximin Principle* | *Generic Info-Gap Model* |
|---|---|
| $d$ | $(q,\alpha)$ |
| $s$ | $u$ |
| $\mathbb{D}$ | $\mathbb{Q}\times\mathbb{R}_+$ |
| $S(d)$ | $\mathcal{U}(\alpha,\tilde{u})$ |
| $f(d,s)$ | $\alpha\cdot(r_c\preceq R(q,u))$ |
| Here $\preceq$ denotes a binary relation yielding 1 if the relation is satisfied, 0 otherwise. $\mathbb{R}_0$ denotes the non-negative part of the real line. | |

Table 2: Correspondence between the generic *Info-Gap* and Maximin models

$\mathcal{U}(\alpha,\tilde{u})$. In this framework the values of $q$ and $\alpha$ are fixed, so minimizing $f(q,\alpha,u) = \alpha\cdot(r_c\preceq R(q,u))$ over $u\in\mathcal{U}(\alpha,\tilde{u})$ amounts to minimizing $R(q,u)$ over $u\in\mathcal{U}(\alpha,\tilde{u})$.

To see more clearly what is going on here, consider a given $q\in\mathbb{Q}$ and its robustness:

$$\hat{\alpha}(q,r_c) := \max_{\alpha\geq 0}\ \min_{u\in\mathcal{U}(\alpha,\tilde{u})}\ \alpha\cdot(r_c\preceq R(q,u)) \tag{44}$$

$$= \max_{\alpha\geq 0}\ \alpha\cdot \min_{u\in\mathcal{U}(\alpha,\tilde{u})} (r_c\preceq R(q,u)) \tag{45}$$

$$= \max_{\alpha\geq 0}\ G(\alpha)\cdot H(q,\alpha) \tag{46}$$

where

$$G(\alpha) := \alpha\ ,\ \ \alpha\geq 0 \tag{47}$$

$$H(q,\alpha) := \min_{u\in\mathcal{U}(\alpha,\tilde{u})} (r_c\preceq R(q,u))\ \ ,\ \ \mathbb{Q},\alpha\geq 0 \tag{48}$$

observing that the nesting property of the regions of uncertainty implies that for a given $q$, $H(q,\alpha)$ is a step function of $\alpha$, as shown in Figure 5.

This implies that $G(\alpha)\cdot H(q,\alpha)$ consists of two linear parts: on the interval $[0,\hat{\alpha}(q,r_c)]$ this function is equal to $G$. And then on the interval $(\hat{\alpha}(q,r_c),\infty)$ the function is equal to 0, as shown in Figure 6.

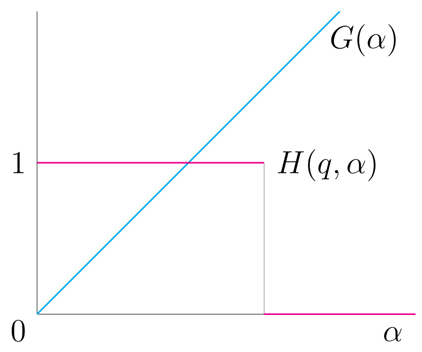


Figure 5: $G = G(\alpha)$ and $H = H(q,\alpha)$, $q$ is fixed.

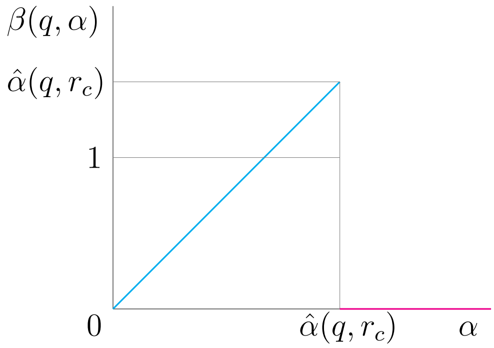


Figure 6: $\beta(q,\alpha) := G(\alpha)\cdot H(q,\alpha)$, $q$ is fixed.

In summary, the fact that the generic *Info-Gap* model is a special case of *Wald's Maximin Principle* is crystal clear and requires no further discussion.

What is unclear – in fact very odd – is the absence of any reference whatsoever to this decision theoretic celebrity in the official *Info-Gap* literature (Ben-Haim[2001, 2006])! It is as if this literature decided to banish this lovely and friendly tool of thought.

What a shame!

**Remark**:

In fact, the situation is far more serious than that. Although the official *Info-Gap* literature (Ben-Haim [2001, 2006]) does not mention the *Maximin* connection, elsewhere in the literature there are claims that *Info-Gap* is not *Maximin*.

As we show in Appendix A, these ill-founded claims are based on serious misconceptions *Info-Gap* has about *worst-case analysis* and by implication *Wald's Maximin Principle*.

# 8 Practice What You Preach!

*Info-Gap* goes out of its way – and rightly so – to stress that under SEVERE uncertainty it is "wrong" to base our decision on a *single point estimate* of the uncertainty parameter under consideration. The argument is simple: under severe uncertainty estimates are of poor quality and are likely to be substantially wrong.

To illustrate the issues involved, suppose that you have to choose an option, or alternative, from a given collection of $n$ options (alternatives). Let $v_i$ denote the value of option $i$, $i=1,\dots,n$, and assume that "larger is better," so ideally – under conditions of strict certainty – you would select the option whose $v_i$ value is largest.

But what should we do in cases where the true values of $v_i$, $i=1,2\dots,n$ are unknown and are subject to SEVERE uncertainty? For example, consider the following concrete case:

| option, $i$ | $v_i$ | $\hat{v}_i$ |
|---|---|---|
| 1 | ? | 17 |
| 2 | ? | 21 |
| 3 | ? | 18 |

assuming that the estimates $\hat{v}_i, i = 1, 2, 3$, are subject to severe uncertainty.

Here is *Info-Gap*'s advice on the fundamental issue posed by such simple decision-making problems:

> The value of $v_i$ is highly uncertain and possibly varying in time, so that historical evidence is of limited utility. The best estimate of the value of option $i$ is $\tilde{v}_i$. For instance, this might be an historical mean, perhaps over a limited time window, and perhaps with temporal lag. Since things change, or since the long-range mean deviates greatly from the mean on short time intervals, the estimate is a poor indication of the true value that will accrue from option $i$ the next time a choice is made.
>
> Ben-Haim [2006, p. 280]

In other words, *Info-Gap* argues the obvious: estimates obtained under severe uncertainty should be regarded as POOR approximations of the true values they represent.

Now, let $\tilde{v}_i$ denote the best estimate we have for the true value of $v_i$ and let $i^*$ denote the option whose $\tilde{v}_i$ value is largest, namely let $i^* = \arg max\{\tilde{v}_i : i = 1, \dots, n\}$.

Here is what *Info-Gap* says about the choice of option $i^*$ as the best (optimal) option:

> The large value of $\tilde{v}_{i^*}$ is desirable. But $\tilde{v}_{i^*}$ is only an estimate of the value of option $i$, and this estimate is likely to be substantially wrong. An additional reason that large $\tilde{v}_{i^*}$ is attractive is the implicit assumption that, since $\tilde{v}_{i^*}$ is large, then the actual value of option $i^*$ is also large even if $\tilde{v}_{i^*}$ errs. This of course is not necessarily true.
>
> Ben-Haim [2006, p. 281]

In other words, *Info-Gap* warns us against the simplistic policy of ranking alternatives on the basis of poor estimates resulting from severe uncertainty. The reason: these estimates are likely to be SUBSTANTIALLY WRONG.

Who can argue against this sound advice?

The question is then: what do we do given that in severe uncertainty the estimates we use are invariably "wild guesses"?

Clearly, *Info-Gap* comes out against the use of "point estimates" of the rewards or payoffs for ranking alternatives. So how exactly does *Info-Gap* resolve the dilemma? How does *Info-Gap* evaluate how good/bad an alternative is?

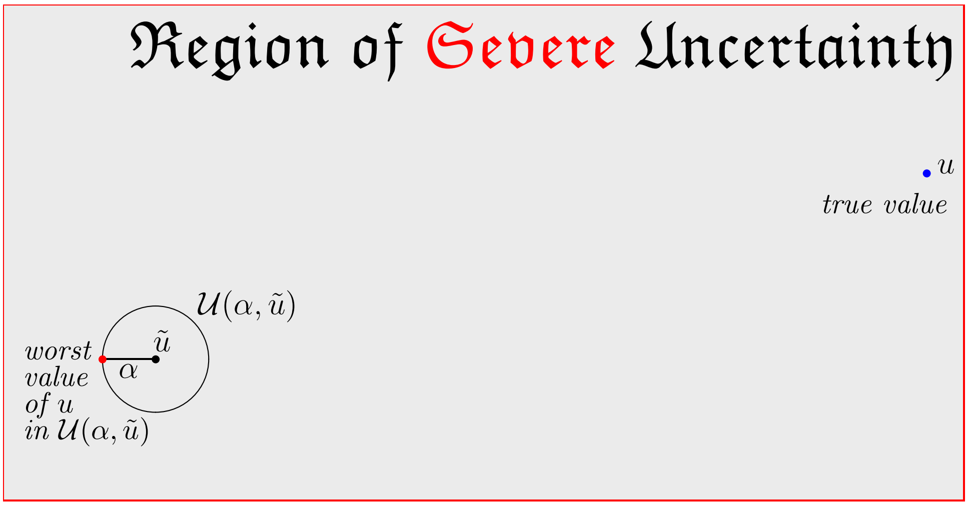


Figure 7: The region of uncertainty

In line with its generic model, *Info-Gap* ranks alternatives according to their *robustness*:

$$\hat{\alpha}(q, r_c) := \max \left\{ \alpha : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right\} \tag{49}$$

where $U(\alpha, \tilde{u})$ denotes the region of uncertainty around the estimate $\tilde{u}$ specified by the parameter $\alpha$.

Since by construction $\mathcal{U}(0, \tilde{u}) = \{\tilde{u}\}$ and $U(\alpha, \tilde{u})$ is increasing in size with $\alpha$, you can regard $\alpha$ as the "radius" of the region of uncertainty.

In short, the *Info-Gap* recipe instructs us to rank alternatives on the basis of how they behave in the neighborhood of the estimate $\tilde{u}$.

But doesn't this fly in the face of *Info-Gap* arguing that under severe uncertainty these estimates are POOR and can be SUBSTANTIALLY WRONG?!

It should be realized that the dilemma cannot be resolved even if, in addition to the point estimate $\tilde{u}$, the immediate neighborhood of $\tilde{u}$, namely $\mathcal{U}(\alpha, \tilde{u})$, is included in the analysis. This is still a *local* analysis, whereas under severe uncertainty we have to take a *global* view at the region of uncertainty.

In fact, if $\tilde{u}$ is far from the true value of $u$, this would only make the situation worse, because the distance from the worst point in $\mathcal{U}(\alpha, \tilde{u})$ to the true value of $u$ is typically greater than the distance from $\tilde{u}$ to the true value of $u$.

The fundamental flaw in *Info-Gap*'s uncertainty model is depicted in Figure 7. Note how small the region of uncertainty $\mathcal{U}(\alpha, \tilde{u})$ is compared to the entire region of uncertainty, and how far it can be from the true value of $u$.

I conducted numerical experiments to confirm this fact and I can report that... this is indeed the case. This and other related issues are discussed in (Sniedovich [2007]).

This being the case, the conclusion is therefore that there is no reason to believe that the solutions generated by *Info-Gap* are robust. At best they are robust in the neighborhood of the estimate $\tilde{u}$.

Note that it is true that for very large values of $\alpha$, the region $\mathcal{U}(\alpha, \tilde{u})$ can cover a significant part of the entire region of uncertainty. However, the largest region of uncertainty investigated by *Info-Gap* is the one determined by

$$\hat{\alpha}(r_c) := \max_{q \in \mathbb{Q}} \hat{\alpha}(q, r_c) \tag{50}$$

Thus, the fact that the region of uncertainty $\mathcal{U}(\alpha, \tilde{u})$ can be large for large $\alpha$ does not imply that $\hat{\alpha}(r_c)$ is large, and therefore there is no a priori guarantee that the *Info-Gap* model will investigate a large portion of the entire region of uncertainty.

Hence,

**Theorem 2** *The Info-Gap uncertainty model is fundamentally flawed.*

**Proof.** This follows immediately from the – justified — *Info-Gap* campaign against the use of best point estimates to rank alternatives (see Ben-Haim [2006, pp. 280-281]) . $\mathfrak{QED}$

My quick back of the envelope analysis indicates in no uncertain terms that in practice this is precisely the case in situations where robust analysis is important. After all, our main concern is the case where the value of $\hat{\alpha}(r_c)$ is small.

To sum-up, *Info-Gap* cannot have it both ways: if it claims to be a tool for decision-making under severe uncertainty then it cannot use a model based on a single point estimate. If the best point estimate is so good that the recipe indeed yields robust solutions, then it cannot be claimed that we are in a situation under severe uncertainty.

The focus on a point estimate and its immediate neighborhood exposes *Info-Gap* to the whims of *local analysis* whereas the severe uncertainty feature of the problem requires a *global analysis*. It is akin to using *local* search to identify a *global* optimum. It does not work.

I should add that the notation in (49) conceals the fact that the robustness deployed by *Info-Gap* is *local* in nature. This is an important issue and the notation used should reflect it. In short, to be reader-friendly $\hat{\alpha}(q, r_c)$ should be re-written as $\hat{\alpha}(q, r_c|\tilde{u})$.

One of the consequences of the local nature of *Info-Gap*'s uncertainty model is that the generic *Info-Gap* model is *completely oblivious* to the "size" of the total region of uncertainty $\mathfrak{U}$ in relation to the "size", $\hat{\alpha}(r_c)$, of the optimal region of uncertainty $\mathcal{U}(\hat{\alpha}(r_c), \tilde{u})$. More precisely,

**Theorem 3** *Info-Gap does not deal with severe uncertainty, it simply ignores it. More precisely, the generic Info-Gap model is invariant with the "size" of total region of uncertainty $\mathfrak{U}$: the value of $\hat{\alpha}(r_c)$ does not vary with $\mathfrak{U}$ for all $\mathfrak{U}$ such that $\mathcal{U}(\hat{\alpha}(r_c) + \varepsilon, \tilde{u}) \subseteq \mathfrak{U}$ for some $\varepsilon > 0$.*

Proof. Let $\alpha^* := \hat{\alpha}(r_c)$ and $\mathfrak{U}^* := \mathcal{U}(\alpha^* + \varepsilon, \tilde{u}), \varepsilon > 0$. We have to show that $\alpha^*$ does not vary with $\mathfrak{U}$ for all $\mathfrak{U}$ such that $\mathfrak{U}^* \subseteq \mathfrak{U}$. This follows immediately from

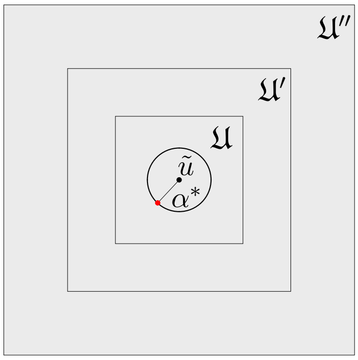


Figure 8: Illustration of Theorem 3

the nesting property of the regions of uncertainty $\mathcal{U}(\alpha, \tilde{u}), \alpha \geq 0$ and the worst-case characteristic of robustness stipulated in the definition of $\hat{\alpha}(r_c)$. $\mathfrak{QED}$

This point is illustrated in Figure 8 where three regions of uncertainty are displayed, $\mathfrak{U} \subset \mathfrak{U}' \subset \mathfrak{U}''$. The same solution, $\alpha^*$, is obtained for any region of uncertainty containing the set $\mathcal{U}(\alpha^*, \tilde{u})$ represented by the circle.

To appreciate the implication of this fact consider the following case: you have just solved a decision making problem under severe uncertainty using *Info-Gap* and obtain an optimal decision $q^*$ whose robustness is $\alpha^* = \hat{\alpha}(r_c) = \hat{\alpha}(q^*, r_c)$. Then you discover the bad news that you actually severely underestimated the severity of the uncertainty associated with your problem: the level of uncertainty is 1000-fold larger! That is, the true total region of uncertainty $\mathfrak{U}$ is 1000-fold larger.

Since this news means that the updated total region of uncertainty contains the old one, there is no change in the *Info-Gap*'s analysis and the same results will be generated: there is no change is the optimal decision and there is no change in its robustness.

Isn't this ridiculous?

**Time Out!**

You may be tempted, dear reader, to defend the *Info-Gap* recipe by arguing as follows:

**FAQ 2** *$\tilde{u}$ is the best estimate we have, so what else can we do?*

I have heard this argument before and I cannot hide my smile when I hear it.

My answer is as follows:

> I trust that you indeed use the best estimate you have. The trouble is not with the choice or quality of your estimate. This is the best estimate you have.
>
> Ask yourself:

> Am I using the best METHODOLOGY currently available to me for the purpose of dealing with severe uncertainty?
>
> This is the question you should ask yourself, and the answer is a resounding ... NO! You are not!
>
> As I suggested already, you should consult the *Robust Optimization* literature for inspiration and guidance on how to tackle this issue.

The point is that the literature on robust optimization advises us to select scenarios with care: scenarios should be selected so as to provide an adequate "coverage" of the total region of uncertainty.

> If the forecaster tries to specify too many discrete forecasts, in an attempt to cover most possibilities, discrete minimax may yield too pessimistic strategies or even run into numerical, or computational, problems due to the resulting numerous scenarios. Similarly, as the upper and lower bounds on a range of forecasts get wider, to provide coverage to a wider set of possibilities, the minimax strategy may become pessimistic. Thus, scenarios have to be chosen with care, among genuinely likely values. The minimax strategy will then answer the legitimate question of what the best strategy should be, in view of the worst case.
>
> Rustem and Howe [2002, p. xiii]

And the message to *Info-Gap* users is this: in decision-making under SEVERE uncertainty you cannot obtain a robust solution by investigating a small neighborhood around the best point estimate you have at hand.

For your convenience, I provide in Exhibit 2 a very naive first-aid package for this purpose. Use it with imagination.

The nice thing about this simple recipe is that you can now use it as a seed for 10 or so PhD dissertations, each proposing and experimenting with slightly different coverage and averaging schemes.

But do not forget to mention to your PhD students that they should read the *Robust Optimization* literature for inspiration and guidance regarding more sophisticated schemes.

**Remark:**

It should be stressed that the fault in the *Info-Gap* uncertainty model does not lie in the (unwitting) deployment of *Wald's Maximin Principle.* Rather the fault is in the use of a single point estimate and its immediate neighborhood as an approximation of the entire region of uncertainty. Indeed, the *Principle* is used extensively in *Robust Optimization* to generate robust solutions for decision-making situations under severe uncertainty. In other words, the culprit here is not the Principle, it is the *local* nature of the worst-case analysis conducted by *Info-Gap.*

*Emergency First Aid Package*

1. Instead of using a single estimate, cover the total region of uncertainty with say $m$ point estimates, call them $\tilde{u}^{(j)}, j = 1, 2, \ldots, m$, making sure that these estimates provide "adequate" coverage of the entire region.
2. Find the optimal solution for each of these estimates. Let $q^{(j)}$ denote the optimal decision associated with estimate $\tilde{u}^{(j)}$.
3. Evaluate the performance of each of these optimal decisions, $q^{(j)}$, in relation to each of the $(m-1)$ other estimates $q^{(i)}, i \neq j$.
4. Use some measure of "averaging" to select the "overall" best solution.

Exhibit 2

# 9 Satisficing vs Optimizing

*Info-Gap*'s *Satisficing is better than Optimizing* campaign and its connection with the promotion of *Robustness* as a cure for severe uncertainty is counter productive and some 30 years late. It reminds me of the heated discussions in the early 1970's about the future of Operations Research, the role of optimization in decision-making, and life in general.

The following quote comprises the complete abstract of a paper published in the journal *Operational Research Quarterly*, now called *Journal of the Operational Research Society* (JORS). The title of the paper is *"Robustness and Optimality as Criteria for Strategic Decisions"*:

> The use of "optimality" as an operational research criterion is insufficiently discriminating. Ample evidence exists that for many problems simple optimization (particularly profit maximization) does not represent the aims of management. In this paper we discuss the nature of the problem situations for which alternative decision criteria are more appropriate. In particular the structure of strategic planning problems is analyzed. The provisional commitment involved in a plan (in contrast to the irrevocable commitment of a decision) leads to the development of a particular criterion, *robustness* – a measure of the flexibility which an initial decision of a plan maintains for achieving near - optimal states in conditions of uncertainty. The robustness concept is developed through the case study of a sequential factory location problem.
>
> Rosenhead et al [1972, p. 413]

And the last paragraph in this paper reads as follows:

> Robustness and stability are two criteria which are appropriate in particular circumstances. Optimality is a criterion which will continue to have wide and useful application. Our argument is that criteria must be matched to circumstances; that more criteria are available than are often considered; and that new criteria can be developed when the need exists.

> If the criteria are related to the real requirements of the problem situation, their novelty need not be a bar to their understanding and acceptance by management.
>
> Rosenhead et al [1972, p. 430]

Needless to say, our trusted aids *Wald's Maximin Principle*, *tradeoffs*, multiple objective and so on play their usual roles in this paper. For example,

> One possible criteria for uncertainty situation is the *minimax* criterion, under which the decision-alternative to choose is that for which the lowest level of benefit (taken across all possible competitive decisions or external events) is as high as possible. Use of the minimax criterion necessarily results in conservative decisions, based as it is on an anticipation that the worst might well happen. In a competitive situation this may be appropriate – if your competitor's interests conflict with yours and he pursues them rationally, he will choose polices which will reduce your gains to a minimum.
>
> Rosenhead et al [1972, p. 416]

In any case, some 35 years later it is best to summarize the "Satisficing vs Optimizing" issue as follows:

**Theorem 4** *The "Satisficing vs Optimizing" issue is a non-issue. Any "satisficing" problem can be transformed into an equivalent "optimizing" problem.*

There are many possible proofs for this important theorem. For example, one can run along these lines: take your "satisficing" problem and add to it an *arbitrary* objective function that attains a positive constant value over the feasible region of the "satisficing" problem. For decisions that are not feasible let the objective function assign a negative value. Clearly, any maximal solution to this optimization problem is a feasible solution to the given "satisficing" problem.

But in the context of our discussion it is more instructive to use a special objective function of this type. The formalities are as follows.

Let $X'$ denote the "unconstrained" domain of the decision variables of the "satisficing problem" and let $X \subseteq X'$ denote the feasible domain of the decision variables, so that $x \in X'$ is feasible iff $x \in X$. For example, let $X' = \mathbb{R}^n$ and let $X$ be a subset of $\mathbb{R}^n$ determined by a system of linear constraints, like those deployed in linear programming.

Associated with this framework define

$$\mathcal{I}(x) := \begin{cases} 1 & , \ x \in X \\ -\infty & , \ x \notin X \end{cases} \quad , \ x \in X' \tag{51}$$

where, as usual in maximization problems $\mathcal{I}(x) = -\infty$ is interpreted as *infeasibility*.

So what we have is this:

| *Generic Satisficing Problem* | *Generic Optimization Problem* |
|---|---|
| *Find an* $x^* \in X'$ *such that* $x^* \in X$ | $x^* \in \text{Arg} \max_{x \in X'} \mathcal{I}(x)$ |

where the capital A in Arg indicates that we collect *all* the optimal solutions into a set.

In other words, by construction, $x^* \in X'$ is a solution to the satisficing problem iff it is a solution to the optimization problem.

In short, the "Satisficing vs Optimizing" issue is an *issue without a cause*. The real issue is $\boxed{\text{WHAT}}$ are we aiming to satisfy and $\boxed{\text{WHAT}}$ are we aiming to optimize.

Therefore, the focus should be on what we deploy as hard and soft $\boxed{\text{CONSTRAINTS}}$ and what we deploy as soft and hard $\boxed{\text{OBJECTIVES}}$ or goals.

Now, *Info-Gap* insists that it is better to optimize the robustness and satisfy the reward requirements. Yet, at the same time *Info-Gap* calls for a Pareto-tradeoff between robustness and rewards.

But the same Pareto frontier is obtained if we do it *the other way around*, namely if we regard robustness as a requirement and maximize the reward given this requirement:

$$\hat{r}(\alpha) := \max_{q\in\mathbb{Q},r\in\mathbb{R}} \min_{u\in\mathcal{U}(\alpha,\tilde{u})} r \cdot (r \preceq R(q,u)) \ , \ \alpha \geq 0 \tag{52}$$

By definition, $\hat{r}(\alpha)$ is the maximum reward that can be generated by a feasible decision under the worst case scenario associated with the region of uncertainty of the given value of $\alpha$.

So what is the big idea of insisting that you maximize the robustness and satisfy the reward? What is wrong with doing it the other way around?!?

In any case the entire enterprise can be stated as follows:

$$z^*(\tilde{u}) := \underset{q\in\mathbb{Q},\alpha\geq 0,r\in[\underline{r},\overline{r}]}{\text{p-max}} \min_{u\in\mathcal{U}(\alpha,\tilde{u})} (r, \alpha \cdot (r \preceq R(q,u))) \tag{53}$$

where p-max denotes the *Pareto Max* operation and $[\underline{r},\overline{r}]$ is the interval of interest for the reward requirement $r$.

Observe that because when $r$ is fixed the minimization with respect to $u$ involves only $\alpha$, the min operation is a single-objective one, hence there is no need for the *Pareto* generalization here.

The inclusion of $\tilde{u}$ in $z^*(\tilde{u})$ is a reminder that this set consists of *Pareto* optimal $(r,\alpha)$ pairs if the true value of $u$ is equal to $\tilde{u}$.

# 10 Probabilistic vs Non-probabilistic Models

This is a another non-issue that should have been kept a non-issue in our textbooks, journal articles, presentations and public pronouncements. But we need to discuss it here at some length because *Info-Gap* makes a big deal out of it.

Readers who are well versed in the "Probabilistic vs Non-probabilistic" debate may wish to skip this section and proceed directly to the next section.

There are two aspects to the well established distinction between probabilistic and non-probabilistic models and it is important not to confuse them:

· The fact that the underlying model deals with *uncertainty*.

· The fact that no probabilistic constructs are used formally in the formulation and quantification of the uncertainty.

The bottom line is that in the framework of a model dealing with uncertainty, a deterministic process can be regarded, and treated formally, as a (degenerate) probabilistic process. For example, in the case of *Wald's Maximin Principle* this hidden probability process selects the worst state of nature with probability one. The same, of course, is the case in the *Info-Gap* model.

So it does not matter what model you use to describe the uncertainty with regard to the state of nature. Under the rules and regulations of *Wald's Maximin Principle* what counts is not how you describe the uncertainty itself, and how it is distributed over the region of uncertainty. Rather, what counts is the set of "feasible" states: *Mother Nature* will select the worst state with probability 1.

Whether such a model is probabilistic or non-probabilistic is a matter of *interpretation*. It is a pity that our textbooks do not make this point crystal clear.

In any case, formally speaking this distinction should be approached with caution. In the context of decision-making under SEVERE uncertainty it is instructive to view *Wald's Maximin Principle* as an uncertainty model whereby the worst state of nature is observed with probability 1. This way, the distinction between this model of uncertainty and the famous *Laplace's Principle of Insufficient Reason* can be explained along the same lines:

· *Laplace's Principle of Insufficient Reason* assumes that all the states are *equally likely*.

· *Wald's Maximin Principle* assumes that the *worst* state is observed with *probability 1*.

In other words, both principles are probabilistic: *Laplace* distributes the random variable under consideration (the state) uniformly over the region of uncertainty (state space) whereas Wald puts it in a single spot.

This reminds me of the two methods we used in the army to clean our barracks for the weekly inspections: one was to spread the dust thinly all over the place in the hope that it will be thin enough to be missed. The other was to collect the dust in one, well hidden, spot, in the hope that it will not be discovered.

We used to have long arguments about the merit of each method and we adopted a mixed strategy. Generally speaking, both worked quite well.

Now back to *Info-Gap*.

In the context of *Info-Gap* the "probabilisitic vs non-probabilistic" issue is important because, as indicated above, *Info-Gap* regards this aspect of the theory as crucial, in fact its hallmark:

> Info-gap decision theory is radically different from all current theories of decision under uncertainty. The difference originates in the modelling of uncertainty as an information gap rather than as a probability.
>
> Ben-Haim [2006, p.xii]

The objective of the discussion in this section is to show that the probabilistic vs non-probabilistic issue lacks real substance.

The legal argument is that not only does *Info-Gap* deploy a probabilistic approach to handle severe uncertainty, the probabilistic construct it uses is a very simple probability distribution function, to wit:

> The *Info-Gap* model ALWAYS generates the WORST VALUE of $u$ (with respect to given $q$, $r_c$,$\tilde{u}$ and $\alpha$) with PROBABILITY 1.

That is, let

$$u^*(q, r_c, \tilde{u}, \alpha) := \arg \min_{u \in \mathcal{U}(\alpha,\tilde{u})} R(q, u) \tag{54}$$

so that by definition, $u^*(q, r_c, \tilde{u}, \alpha)$ denotes the worst value of $u$ in $\mathcal{U}(\alpha, \tilde{u})$ for decision $q$.

Then formally, the probability distribution function used by *Info-Gap* to simulate $u$ is as follows:

$$Prob(u|q, r_c, \tilde{u}, \alpha) := \begin{cases} 1 & , \quad u = u^*(q, r_c, \tilde{u}, \alpha) \\ 0 & , \quad otherwise \end{cases} \quad , \ q \in \mathbb{Q}, \alpha \geq 0 \tag{55}$$

This is the same probability function deployed by *Wald's Maximin Principle* and the ground for our claim that *Info-Gap* is *Wald's Maximin Principle* in disguise.

Given the nesting property of the regions of uncertainty $\{\mathcal{U}(\alpha, \tilde{u})\}, \alpha \geq 0$ deployed by *Info-Gap*, it is possible to formulate infinitely many other probabilistic models – indeed fancier ones– to capture the essential spirit of the *Info-Gap* model.

To see how this can be done, recall that the regions $\{\mathcal{U}(\alpha, \tilde{u})\}, \alpha \geq 0$ have two basic properties:

$$\mathcal{U}(0, \tilde{u}) = \{\tilde{u}\} \tag{56}$$

$$\alpha > \alpha' \Longrightarrow \mathcal{U}(\alpha', \tilde{u}) \subseteq \mathcal{U}(\alpha', \tilde{u}) \tag{57}$$

As we shall see soon, this meta-structure is readily amenable to a classical probabilistic interpretation.

The central question is:

what is the *meaning* of the parameter $\alpha$ in this framework?
What exactly does $\alpha$ *represent*?

To fix ideas, consider the following instance of the case considered in Ben-Haim [2006, p.71]:

$$\mathcal{U}(\alpha, \tilde{u}) = \left\{u \in \mathbb{R}^2 : (u_1 - \tilde{u})^2 + (u_2 - \tilde{u}_2)^2 \leq \alpha^2\right\} \ , \ \alpha \geq 0 \tag{58}$$

so $\mathcal{U}(\alpha, \tilde{u})$ is a circle of radius $\alpha$ centered at $\tilde{u} \in \mathbb{R}^2$ and consequently the regions of uncertainty constitute a collection of infinitely many concentric circles.

Given that $\tilde{u}$ is interpreted as the estimate of the true value of $u$, intuitively it is natural to think about these concentric circles as *contours* of a real valued function defined on the total region of uncertainty $\mathcal{U}$. This function describes how "likely" it is that the true value of $u$ lies in a circle of radius $\alpha$ centered at $\tilde{u}$.

This is in line with the way *Info-Gap* describes the meaning of $\alpha$ in relation to similar regions of uncertainty:

> The larger the value of $\alpha$, the greater the range of unknown variation, so $\alpha$ is called the *uncertainty parameter* or *horizon of uncertainty*. However, quite often the value of $\alpha$ itself is not known so in fact (2.3) is not a single set but a rather an unbounded family of nested sets of functions. The degree of nesting, as well as well as the level of uncertainty, is expressed by the uncertainty parameter $\alpha$.
>
> Ben-Haim [2006, p. 17]

So, if indeed $\alpha$ represents the *level of uncertainty* associated with the true value of $u$, why don't we call a spade a spade, and regard $\alpha$ as a realization of a *random variable* that stipulates how far the true value of $u$ is from the estimate $\tilde{u}$? And while we are at it, why don't we assign to $\alpha$ a proper *probability distribution function*?

Of course, the reader may object to this idea, arguing that under *severe* uncertainty it would be impossible to justify the use of any such function. But then, how can we justify – under the same conditions of *severe* uncertainty – an uncertainty model based on a good estimate, $\tilde{u}$, of the true value of $u$ and a relatively small neighborhood around it?

In any case, the purpose of the following exercise is not to convince the reader that this approach makes sense, but rather that it ends up with a model that is equivalent to the model deployed by *Info-Gap*. That is, the purpose of the exercise is to show that it is very easy indeed to give the uncertainty model deployed by *Info-Gap* a probabilistic interpretation par excellence.

As explained above, this is so because by always selecting the worst case, *Wald's Maximin Principle* does its work with complete disregard for the stipulated probabilistic structure. So basically we can use whatever model we fancy as long as we keep the same state space.

To imitate the *Info-Gap* uncertainty model using a formal probabilistic construct, we regard the parameter $u \in \mathcal{U}$ as a realization of a random variable $\dot{u}$. We shall now show how the probability density function of $\dot{u}$ can be constructed so that it imitates the *Info-Gap* uncertainty model.

The basic constructs are as follows:

- A random variable $\dot{\alpha}$ induced by a probability density function $\rho$ on $\mathbb{R}_0$.

- A conditional probability distribution function $\varphi(u|q, r_c, \tilde{u}, \alpha)$ on $\mathcal{U}(\alpha, \tilde{u})$. Let $\dot{u}(q, r_c, \tilde{u}, \alpha)$ denote the random variable induced by this distribution.

The conceptual model for generating a realization of the unknown parameter $u$ – now viewed as a realization of random variable $\dot{u}$ – is then as follows:

**Step 1**: Generate a realization $\alpha$ of $\dot{\alpha}$ on $\mathbb{R}_0$.
**Step 2**: Generate a realization $u$ of $\dot{u}(q, r_c, \tilde{u}, \alpha)$ in $\mathcal{U}(\alpha, \tilde{u})$.

That is, conceptually we think about $\dot{u}$ as a random variable whose realizations are determined by a two-stage process. In the first stage a realization $\alpha$ of $\dot{\alpha}$ is generated in accordance with the density $\rho$ of $\dot{\alpha}$. Then, the conditional density $\varphi(\cdot|q, r_c, \tilde{u}, \alpha)$ generates a realization of $\dot{u}(q, r_c, \tilde{u}, \alpha)$ in $\mathcal{U}(\alpha, \tilde{u})$.

Thus, by definition,

$$Prob(\dot{\alpha} \leq \alpha) = \int_0^{\alpha} \rho(\beta) d\beta \ , \ \alpha \geq 0 \tag{59}$$

and

$$\varphi(u|q, r_c, \tilde{u}) = \int_0^{\infty} \varphi(u|q, r_c, \tilde{u}, \alpha)\rho(\alpha) d\alpha \ , \ u \in \mathcal{U} \tag{60}$$

As far as robustness is concerned, the event of interest to us is $r_c \leq R(q, \dot{u})$, so let us examine the probability of this event given the above probabilistic model:

$$\begin{aligned} Prob(r_c \leq R(q, \dot{u})) &= \int_0^{\infty} Prob(r_c \leq R(q, \dot{u})|\alpha)\rho(\alpha) d\alpha && (61) \\ &= \int_0^{\infty} Prob(r_c \leq R(q, \dot{u}(q, r_c, \tilde{u}, \alpha)))\rho(\alpha) d\alpha && (62) \\ &= \int_0^{\hat{\alpha}(q, r_c, \tilde{u})} Prob(r_c \leq R(q, \dot{u}(q, r_c, \tilde{u}, \alpha)))\rho(\alpha) d\alpha \\ &\quad + \int_{\hat{\alpha}(q, r_c, \tilde{u})}^{\infty} Prob(r_c \leq R(q, \dot{u}(q, r_c, \tilde{u}, \alpha)))\rho(\alpha) d\alpha && (63) \end{aligned}$$

Next, to imitate *Info-Gap*'s worst-case philosophy, consider the special case where

$$\varphi(u|q, r_c, \tilde{u}, \alpha) = \begin{cases} 1 & , \quad u = u^*(q, r_c, \tilde{u}, \alpha) \\ 0 & , \quad u \neq u^*(q, r_c, \tilde{u}, \alpha) \end{cases} \quad , \quad u \in \mathcal{U}(\alpha, \tilde{u}) \tag{64}$$

where $u^*(q, r_c, \tilde{u}, \alpha)$ is defined by (54). In this case we have

$$Prob(r_c \leq R(q, \dot{u}(q, r_c, \tilde{u}, \alpha))) = \begin{cases} 1 & , \quad \alpha \leq \hat{\alpha}(q, r_c, \tilde{u}) \\ 0 & , \quad otherwise \end{cases} \tag{65}$$

Thus, it follows from (63) that under *Info-Gap*'s worst-case scenario

$$\begin{aligned} Prob(r_c \leq R(q, \dot{u})) &= \int_0^{\hat{\alpha}(q, r_c, \tilde{u})} \rho(\alpha) d\alpha && (66) \\ &= Prob(\dot{\alpha} \leq \hat{\alpha}(q, r_c, \tilde{u})) && (67) \end{aligned}$$

And because *Info-Gap*'s regions of uncertainty are concentric, this entails that

$$Prob(r_c \leq R(q, \dot{u})) = \max\{\alpha \geq 0 : Prob(r_c \leq R(q, u^*(q, r_c, \tilde{u}, \alpha))) = 1\} \tag{68}$$

In short,

**Theorem 5**

$$\hat{\alpha}(q,r_c) = \max_{\alpha\geq 0} \ \alpha \cdot Prob(r_c \leq R(q,u^*(q,r_c,\tilde{u},\alpha))) \tag{69}$$

$$\hat{\alpha}(r_c) = \max_{q\in\mathbb{Q},\alpha\geq 0} \ \alpha \cdot Prob(r_c \leq R(q,u^*(q,r_c,\tilde{u},\alpha))) \tag{70}$$

In other words, we can express the *Info-Gap* robustness indices in simple probabilistic terms and the generic *Info-Gap* model as a whole as a classical probabilistic model.

Observe that the missing min of the standard *Maximin* format is hidden in the definition of $u^*(q,r_c,\tilde{u},\alpha)$, namely in (54).

# 11 Discussion

From an Operations Research point of view *Info-Gap* is a simple instance of *Wald's Maximin Principle*. Indeed, this idea can be immediately employed to "generalize" the generic model of *Info-Gap*.

To appreciate the gain to *Info-Gap*, note that *Info-Gap* seems to have difficulties dealing with general types of requirements. Its multiple reward model deals only with $\leq$ constraints. This is due to the fact that in order to apply the *Maximin* formulation it is necessary to minimize the reward function, namely to assume that "more is better".

But how do you deal with $=$ constraints? And how do you deal with $\in$ constraints? And how do you deal with more general constraints?

From an OR perspective this is not an issue because all you have to do is represent the requirement constraints by the *indicator function* of these constraints, call it $\mathcal{I}$. That is, let $\mathcal{I}$ denote the function defined as follows:

$$\mathcal{I}(q,u) := \begin{cases} 1 & , \ \textit{all the constraints are satisfied, given } (q,u) \\ -\infty & , \ \textit{Otherwise} \end{cases} \tag{71}$$

In this case the explicit *Maximin* representation of the *Info-Gap* model will be as follows:

$$\hat{\alpha}(r_c) = \max_{q\in\mathbb{Q},\alpha\geq 0} \ \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \alpha\,\mathcal{I}(q,u) \tag{72}$$

Also, it is puzzling that *Pareto Optimization* is not mentioned in the first edition of the *Info-Gap* book in spite of the extensive discussions therein on tradeoff analysis à la Pareto.

There is some progress in the second edition of the book. The term "Pareto efficient" is mentioned, albeit toward the end of the book (Ben-Haim [2006, p. 281, 312]). Still . . . no references are cited and the term "Pareto" does not appear in the subject index.

Equally important, given these extensive discussions on tradeoffs between rewards and robustness, it is very odd that nowhere is any reference given or indication made that the robustness is *with respect to the assumed value of the estimate* $\tilde{u}$. Given the severe uncertainty, it is only natural to explore the uncertainty region rather than focus the analysis on a single point estimate and its immediate neighborhood.

In any case, this "local" interpretation of robustness should be reflected in the notation used. That is, $\hat{\alpha}(q, r_c|\tilde{u})$ and $\hat{\alpha}(r_c|\tilde{u})$ should be used to denote robustness, rather than $\hat{\alpha}(q, r_c)$ and $\hat{\alpha}(r_c)$, respectively. This is not a case of being pedantic about notation. It is about using proper notation to help readers understand the implications of the assumptions used in the formulation of the model.

Another big gap in the *Info-Gap* literature is the lack of indication as to how one should go about *solving* the optimization problems induced by the *Info-Gap* model, namely optimization problems of the form

$$\hat{\alpha}(r_c) := \max_{q \in \mathbb{Q}, \alpha \geq 0} \alpha \tag{73}$$

$$r_c \leq R(q, u) \ , \forall u \in \mathcal{U}(\alpha, \tilde{u}) \tag{74}$$

Observe that in general these problems are nonlinear.

Strange though it may be, *Info-Gap* is clearly unaware of the area of *Robust Optimization.*

This is a pity.

Not only does *Robust Optimization* provide useful ideas for developing solution methods for robust optimization problems, it also provides alternative approaches for defining "robustness". In fact *Info-Gap* can profit greatly by consulting the *Operations Research* and *Mathematical Programming* literature on *Robust Optimization.* Specifically, this literature can enlighten *Info-Gap* on how to fix its uncertainty model.

You may wonder, dear reader, how *Info-Gap* can be so lax about its links to relevant decision theoretic oriented disciplines such as *Operations Research*, *Mathematical Programming* and *Robust Optimization.* The clue can be found in the *Info-Gap* book:

> Info-Gap models of uncertainty originated in the technological sciences, and the early work on decision-making with info-gap uncertainty concentrated on engineering analysis and design. This is in rather marked contrast to the development of most current decision theories, which have been intensively pursued by economists and other social scientists, psychologists, management and operations researchers and related scholars in the supporting disciplines of mathematics and statistics. Info-Gap decision theory has been heavily influenced by the classical theories, primarily in the identification of the roles and goals of a decision theory, and much less in the formulation of questions or methods of solutions. Many concepts from classical theories such as risk aversion, value of information, and learning, have identifiable but different manifestations in info-gap.
>
> Ben-Haim [2006, p. 3]

This is a pity! And it shows!

As we have seen, such declarations are based on serious misconceptions not only about the state of the art in *Operations Research*, *Mathematical Programming* and *Robust Optimization*, but also about *classical decision theory proper*.

Indeed, the idea that classical decision theory does not offer a non-probabilistic approach to decision-making under uncertainty is preposterous. Furthermore, the failure to refer to this approach, let alone discuss it, in a book on decision-making under severe uncertainty is inexplicable.

I find it instructive at this stage to illustrate how this is done in our textbooks. Take for example the very soft textbook entitled *Spreadsheet Modeling Decision Analysis* by Ragsdale [2004].

Chapter 15 in this book, entitled *Decision Analysis*, deals, among other related things, with *probabilistic methods* and *nonprobabilistic methods*. Here is the first paragraph of the sixth section in this chapter:

> **15.6 NONPROBABILISTIC METHODS**
>
> The decision rules we will discuss can be divided into two categories: those that assume that probabilities of occurrences can assigned to the states of nature in a decision problem (**probabililistic methods**), and those that do not (**nonprobabililistic methods**). We will discuss the nonprobabilistic methods first.
>
> Ragsdale [2004, p. 760]

*Wald's Maximin Principle*, and its many variations, appear in this section.

So what exactly are we to make of *Info-Gap*'s claim:

> Info-gap decision theory is radically different from all current theories of decision under uncertainty. The difference originates in the modelling of uncertainty as an information gap rather than as a probability.
>
> Ben-Haim [2006, p.xii]

How exactly does *info-Gap* handle the information-gap if not by assuming the *worst case scenario*?

Here is how Ragsdale describes the *Maximin Rule*:

> **15.6.2 The Maximin Decision Rule**
>
> A more conservative approach to decision-making is given by the **maximin decision rule**, which pessimistically assumes that nature will always be "against us" regardless of the decision we make. The decision rule can be used to hedge against the worst possible outcome of a decision.
>
> Ragsdale [2004, p. 761]

I tried on several occasions to find standard decision theoretic terms such as "worst case", "Mother nature", "state of nature", "against us" and so on, in the *Info-Gap* books, but I failed. Perhaps it is yet another sign that I am getting old, but I doubt it.

I must confess that while reading the *Info-Gap* books I do not feel that I am in a decision theoretic environment.

# 12 Bibliographical Notes

I provide references to a number of general introductory Operation Research textbooks (Hillier and Lieberman [2005], Markland and Sweigart [1987], Ragsdale [2004], Winston [1994]). This should give you a general overview of OR and how *Wald's Maximin Principle* fits in this framework.

For the same reason I provide references to more specialized decision theory books (French [1998], Grünig and Kühn [ 2005]) and a reference to a multicriteria optimization book (Steuer [1985]).

The game theory book (von Neumann and Morgenstern [1944]) is on my list because although it is cited in the *Info-Gap* literature ... this literature makes no mention of the *Maximin* paradigm developed by von Neumann in the late 1920s.

I encourage *Info-Gap* devotees to take a quick look at the references I provide for robust optimization (Ben-Tal et al [2006], Rustem and Howe (2002), Kouvelis and Yu [1997], Rosenhead et al [1972], Vladimirou ad Zenios [1997]) for reassurance that optimization theory is fully aware of the robustness issue, a fact that the *Info-Gap* should acknowledge and appreciate.

My paper on the famous Egg Dropping puzzle (Sniedovich [2003]) is on the list because ... by coincidence it refers explicitly to the *robustness* issue. It shows that in the case of this puzzle there is a robust policy: the same policy is optimal both with respect to the *Maximin* criterion and with respect to *Laplace's Insufficient Reason* criterion.

The reference to Gilboa and Schmeidler [1989] is important and requires a special explanation.

This paper is cited in the 2nd edition of the *Info-Gap* book, but not in the first edition. Its title is *Maxmin expected utility with non-unique prior.* As indicated by its title, it refers explicitly to *Wald's maximin criterion.*

And here is how this reference is mentioned in the *Info-Gap* book:

> Info-gap models are not the only possible way to quantify Knightian uncertainty. On the contrary, Gilboa and Schmeidler [78], Epstein and Wang [72], Epstein and Miao [71] and others, achieve uninsurable uncertainty of a clearly Knaightian type by replacing a single prior probability distribution with a set of distributions. These approaches are Knightian "true uncertainty" since the absence of a probability measure on the set of probability distributions make the uncertainty uninsurable. Nonetheless, an info-gap model of uncertainty is a more extreme departure of the probabilistic tradition. In our formulation, preferences are generated by the robustness function without any distribution functions at all.
>
> Ben-Haim [2006, p. 294]

Given that *Wald's Maximin Principle* equally does not deploy "... any distribution functions at all ..." and that Gilboa and Schmeidler's formulation refers explicitly to *Wald's Maximin criterion*, it is puzzling that the *Info-Gap* book does not refer to *Wald's Maximin criterion.*

There is hope, though, that the strong connection between *Info-Gap* and *Wald's Maximin Principle* will be acknowledged and will receive proper treatment in the next edition of the book.

# 13 Epilogue

My more than 30 years of experience using decision theory, teaching various aspects of decision theory and developing a large body of related courseware, have taught me an important lesson.

One must appreciate to the full the difficulties associated with decision-making under uncertainty in general and severe uncertainty in particular.

In my humble opinion, given the state of the art in *decision theory, operations research and robust optimization*, proponents of *Info-Gap* should reassess this theory and its role and place in decision theory.

Indeed, it is a pity that *Info-Gap* does not recognize the strong links between its goal of obtaining robust solutions under severe uncertainty and what has been happening in Operations Research in this area over the past 50 years.

*Info-Gap* should acknowledge, appreciate and exploit the fact that its generic model is an instance of *Wald's Maximin Principle*, that its tradeoff analysis is standard *Pareto analysis* and that there is a well developed area called *Robust Optimization.*

It should also practice what it preaches and refrain from basing its severe uncertainty analysis solely on a single point estimate and its neighborhood. And while at it, it should bring to a happy end the unproductive *Satisficing is Better than Optimizing* campaign.

It is important to take note, dear reader, of the severity of my critique. I question not only *Info-Gap*'s contribution to the state of the art in decision theory, but also its very familiarity with this state of the art.

But this is unavoidable. The two go hand in hand.

How are we to judge a theory for decision-making under uncertainty that does not recognize *Wald's Maximin Principle* and is completely unaware of the extensive research work in such areas as *Operations Research* and *Robust Optimization* in relation to decision-making under severe uncertainty?

But there is a lesson here for *Operations Research* enthusiasts as well, in fact the OR community as a whole.

We must be doing something very wrong in the way we market our products and technologies. For how is it that these are not recognized and appreciated by professionals from other disciplines for what they are and what they can do?

Why is it that other disciplines have to re-invent standard OR methods and techniques, rather than use our existing products – products that are discussed in our textbooks?

And how is it that OR analysts do not recognize these methods and techniques when they come disguised by jargon !?!?

And to my dear colleagues:

I am fully aware of the fact that you'll soon tell me that the flaws in *Info-Gap* are so OBVIOUS that you do not understand what the fuss is all about. After all, this longish essay is just about two ways of expressing the same thing:

$$\max\left\{\alpha : r_c \leq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right\} \qquad\qquad \max_{\alpha\geq 0}\ \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \alpha\cdot(r_c \preceq R(q,u))$$

There is no fuss here at all. I regard this analysis as part of my ongoing effort to promote the idea that *mathematical modeling* is immensely important in decision-making. What we have here is a vivid illustration of the *mathematical modeling* subtleties associated with *Wald's Maximin Principle* and *worst-case analysis.*

I shall therefore be extremely pleased to hear that you find the conclusions obvious. The more obvious they are to you the more pleased I'll be.

So make my day and send me a note!

# 14 Conclusions

Given the claims made in the official *Info-Gap* literature about its role and place in decision theory, our investigation was guided by the following three fundamental questions:

Q1 Is the generic *Info-Gap* model *new*?

Q2 Is it radically *different* from the classical models of decision theory?

Q3 How well does it represent *severe uncertainty*?

Having carefully examined the generic *info-Gap* model from an *Operations Research* point of view, we are now in a position to provide answers to these questions:

A1 Not only is it the case that the generic *Info-Gap* model is *not new*, it is a simple instance of none other than the most famous model in decision-making under severe uncertainty, namely *Wald's Maximin model.*

A2 For the very same reason, the generic *Info-Gap* model is *not radically different* from classical models for decision-making under severe uncertainty.

A3 The generic *Info-Gap* model does not deal with the severe uncertainty aspect of the decision problem. It simply and unceremoniously ignores it.

So it turns out that when you clear all the fog surrounding the essence of what *Info-Gap* actually does, you discover that conceptually its generic decision model consists of two very simple ingredients:

1. Replacing severe uncertainty by a poor estimate of the parameter under consideration that can be substantially wrong.

2. Conducting a simple vanilla *Maximin* analysis in the neighborhood of this estimate.

The first step amounts to practicing *voodoo decision-making*: instead of dealing with severe uncertainty by properly exploring the region of uncertainty under consideration, you simply …… ignore the severe uncertainty altogether and base the entire analysis on a single point estimate of the parameter in question and its immediate neighborhood. The picture is this:

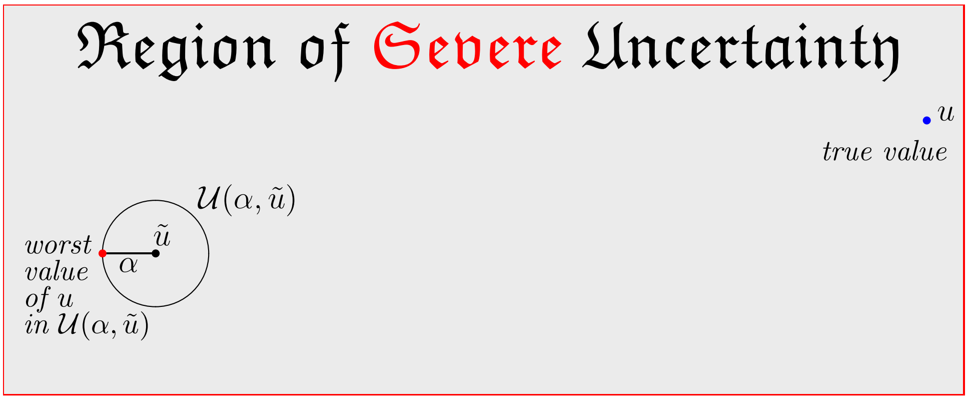


Figure 9: Voodoo Decision-Making

Indeed, *Info-Gap*'s generic model addresses the following question: how far can we move from a given point $\tilde{u} \in \mathcal{U}$ without violating the performance requirement $r_c \leq R(q,u), \forall u \in \mathcal{U}(\alpha,\tilde{u})$?

You can ask this question in the context of a *deterministic* problem where there is no uncertainty at all in the parameter under consideration.

In short, *Info-Gap* does not deal with severe uncertainty – it simply ignores it.

The implication is then this: all that *Info-Gap*'s generic model does is to conduct a simple, vanilla, worst-case analysis a la *Maximin* in the neighborhood of a given point estimate of the parameter under consideration.

So technically *Info-Gap*'s generic model is just an instance of Wald's [1945] *Maximin* model.

However, the idea to use such a model under severe uncertainty, where the parameter is poor and is likely to be substantially wrong amounts to *voodoo decision-making*.

So what are we to make of the claim that *Info-Gap* is a new theory for decision-making under severe uncertainty that is radically different from all existing theories?

I don't know about you, dear reader. I think that this claim is preposterous!

More than anything else, it exhibits severe *Info-Gap* about the state of the art in decision-making under severe uncertainty.

**Acknowledgment**. I would like to thank my two summer students, Jaeger Renn-Jones and Michael Clark, for their constructive comments on various versions of this and other related papers.

**Note**

A copy of a short book on

*Worst-Case Analysis for decision-making Under Strict Uncertainty*

based on this essay should be available on my website at

`www.ms.unimelb.edu.au/~moshe/maximin/`

before the end of 2007.

If it is not there by then, send me a note and I shall make sure that it is there!

I shall be delighted to give a presentation on this topic at your place.

# References


[1] Ben-Haim Y. *Information Gap Decision Theory*, Academic Press, 2001.

[2] Ben-Haim Y. Value-at-risk with info-gap uncertainty, *The Journal of Risk Finance*, 6(5), 388-403, 2005

[3] Ben-Haim Y. *Info-Gap Decision Theory*, Elsevier, 2006.

[4] Ben-Tal A, El Ghaoui L, Nemirovski A. *Mathematical Programming*, Special issue on *Robust Optimization*, Volume 107(1-2), 2006.

[5] Carmel Y, Ben-Haim Y. Info-Gap Robust-Satisficing Model of Foraging Behavior: Do Foragers or Satisfice?, *The American Naturalist*, 166(5), 663-641, 2005.

[6] Cormen TH, Leiserson CE, Rivest RL, Stein C. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001.

[7] French SD. *Decision Theory*, Ellis Horwood, 1988.

[8] Gilboa I, Schmeidler D. Maxmin expected utility with non-unique prior, *Journal of Mathematical Economics*, 18, 141-153, 1989.

[9] Grünig R, Kühn R. *Successful decision-making*, Springer-Verlag, 2005.

[10] Hillier SF and Lieberman GJ. *Introduction to Operations Research*, 8th edition, McGraw Hill, 2005.

[11] Markland RE, Sweigart JR. *Quantitative Methods: Applications to Managerial decision-making*, John Wiley, 1987.

[12] Miles RF Jr. Specifying design conservatism: worst case versus probabilistic analysis, in *Proceedings of the Second International Symposium on Ground Data Systems for Space Mission Operations*, 703-709, 1993.

[13] Kouvelis P, Yu G. *Robust Discrete Optimization and Its Applications*, Kluwer, 1997.

[14] Ragsdale, CT. *Spreadsheet Modeling and Decision Analysis*, Thomson, 2004.

[15] Rawls J. *A Theory of Justice*, Belknap, 1971.

[16] Rosenhead MJ, Elton M, Gupta SK. Robustness and Optimality as Criteria for Strategic Decisions, *Operational Research Quarterly*, 23(4), 413-430, 1972.

[17] Rustem B, Howe M. *Algorithms for Worst-case Design and Applications to Risk Management*, Princeton University Press, 2002.

[18] Sniedovich M. OR/MS Games: 3. The Counterfeit coin problem, *INFORMS Transactions in Education*, Vol. 3(2), 32-41, 2003.

[19] Sniedovich M. *Worst-Case Analysis for decision-making Under Severe Uncertainty*, under preparation, 2007.

[20] Steuer R. *Multiple Criteria Optimization: Theory, Computation and Application*, John Wiley, 1985.

[21] Tintner G. Abraham Wald's contributions to econometrics, *The Annals of Mathematical Statistics,* 23(1), 21-28, 1952.

[22] Vladimirou H, Zenios SA. Stochastic Programming and Robust Optimization, Chapter 12 in *Advances in Sensitivity Analysis and Parametric Programming*, edited by Gal T, Greenberg HJ, Kluwer, 1997.

[23] von Neumann J. Zur theories der gesellschaftsspiele, *Math. Annalen*, Volume 100, 295-320, 1928.

[24] von Neumann J, Morgenstern O. *Theory of Games and Economic Behavior*, Princeton University Press, 1944.

[25] Wald A. Statistical decision functions which minimize the maximum risk, *The Annals of Mathematics,* 46(2), 265-280, 1945a.

[26] Wald A. Generalization of a Theorem by v. Neuman Concerning zero-sum two person games, *The Annals of Mathematics,* 46(2), 281-286, 1945b.

[27] Wald A. *Statistical Decision Functions*, John Wiley, 1950.

[28] Winston WL. *Operations Research: Applications and Algorithms*, Duxbury Press, 1994.

# Appendix

## A The 𝔄ℜ𝔗 and 𝕊ℂ𝕀𝔼ℕℂ𝔼 of worst-case analysis

I have already shown how easy it is to formulate the entire *Info-Gap* model as an instance of *Wald's Maximin Principle* and, by implication, how natural it is to view *Info-Gap* as a simple *worst-case analysis*.

Recall that in this framework the decision maker is playing against an ANTAGONISTIC *Mother Nature*. That is, in line with the worst-case dogma, *Mother Nature* always selects the worst state (for the decision maker) associated with the decision made by the decision maker.

In this appendix I need to enlarge on my discussion of this issue. The reason for this is that although the *Info-Gap* books (Ben-Haim [ 2001, 2006]) are completely oblivious to the obvious *Maximin/worst-case analysis* connection, there are claims in a recent article that in the *Info-Gap* framework there is no *worst case* and therefore *Info-Gap* is not *Maximin*.

So apparently what I consider "simple" and "easy" is not so simple and not so easy after all.

My plan for this appendix is first to examine quickly the reasoning behind these claims and then to explain what is amiss with them. Since these claims are invalid on many fronts, my discussion is a bit lengthy. So bear with me.

### A.1 The claims

As I have already indicated on several occasions, one of the oddest things about the official *Info-Gap* literature (Ben-Haim [ 2001, 2006]), is that it takes no cognizance of the existence of *Maximin* and *worst-case analysis* and the important role they play in decision-making under severe uncertainty and robust optimization.

It should be noted, though, that this connection is touched on in other *Info-Gap* publications. For example, consider this:

> The info-gap model is unbounded in the sense that there is no largest set and there is no worst case.
>
> Carmel and Ben-Haim [2005, p. 635]

And this:

> It is important to emphasize that the robustness $\tilde{h}(R, c)$ is *not* a minimax algorithm. In minimax robustness analysis, one *minimizes* the *maximum* adversity. This is not what info-gap robustness does. There is no maximal adversity in an info-gap model of uncertainty: the worst case at any horizon of uncertainty $h$ is less damaging than some realization at a greater horizon of uncertainty. Since the horizon of uncertainty is unbounded, there is no worst case and the info-gap analysis cannot and does not purport to ameliorate a worst case.
>
> Ben-Haim [2005, p. 392]

And this:

> Info-Gap robustness analysis is a a stress-testing tool with similarities to the maximum-loss and worst-case methods. However, it is important to point out two fundamental differences between info-gap analysis and these methods, differences that make the info-gap approach a useful supplement. First the robustness is *not* a worst-case or minimax assessment. There is no worst case in an info-gap model of uncertainty: as the horizon of uncertainty $h$ grows, the uncertainty sets $\mathcal{F}(h, \tilde{f})$ become more inclusive. The robustness function $\tilde{h}(R_*, c)$ does not identify a worst case. What is evaluated is the greatest horizon of uncertainty up to which the performance is acceptable. This in no way asserts that the real variation is acceptable. The utility of the robustness function is in comparing alternative investment in order to determine which is more immune and which is less, and in assessing capital requirements in terms of Knightian uncertainty in the estimated PDF.
>
> The second basic difference between info-gap and extreme-value methods is that info-gap analysis deals non-probabilistically with severe uncertainty. An info-gap model quantifies the Knightian uncertainty – the lack of information and understanding of unmeasured future changes or surprises – that accompanies an estimate of the PDF. The info-gap robustness assess the impact of Knightian uncertainty without introducing measure functions or probabilistic assumptions and requirements such as normality or large samples.
>
> Ben-Haim [2005, p. 401]

**Remark**: in the context of the model used in this paper, read $h$ as $\alpha$; $c$ as $q$; $R_*$ as $r_c$; $\tilde{h}(R_*, c)$ as $\hat{\alpha}(q, r_c)$; and $\mathcal{F}(h, \tilde{f})$ as $\mathcal{U}(\alpha, \tilde{u})$.

In view of the discussion, these claims are clearly invalid: I have formally proved that *Info-Gap* is a simple instance of *Wald's Maximin Principle*, hence that it is a classical *worst-case analysis*.

So the question is not whether the above claims are valid or not. The question is rather: what exactly is wrong in these claims? Where exactly is the fault in the specific arguments these claims are based on?

Before I address these questions in detail it is instructive to provide a more vivid demonstration that these claims are wrong. For this purpose I now supplement the formal proof that *Info-Gap* is *maximin* in disguise with a semi-formal description of *Info-Gap* as a typical *worst-case analysis*.

## A.2 Info-Gap as a typical worst-case analysis

Recall that in the context of the *Info-Gap* model, the decisions are ranked according to their robustness, and the recipe for this is as follows:

$$\hat{\alpha}(q, r_c) := \max\left\{\alpha : r_c \le \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q, u)\right\} \ , \ q \in \mathbb{Q}, \alpha \ge 0 \tag{75}$$

This is the robustness of decision $q \in \mathbb{Q}$. With no loss of generality assume that $r_c \leq R(q, \tilde{u}), \forall q \in \mathbb{Q}$.

To interpret this recipe as a *worst-case analysis* we view the optimization problem on the right-hand side of (75) as a game between *Decision Maker* and *Mother Nature*. In this framework, *Decision Maker* controls the value of $\alpha$ and *Mother Nature* controls the value of $u$. Note that in this framework $q$ is given and is therefore treated as a fixed parameter.

The payoff function for this game is stipulated by

$$f(q, \alpha, u) := \alpha \cdot (r_c \preceq R(q, u)) \ , \ q \in \mathbb{Q}, \alpha \geq 0, u \in \mathcal{U}(\alpha, \tilde{u}) \tag{76}$$

Note that the payoff $f(q, \alpha, u)$ is equal to either $\alpha$ or 0 depending on whether the performance requirement $r_c \leq R(q, u)$ is satisfied or not, respectively.

The *Decision Maker* is aiming to maximize $f(q, \alpha, u)$ with respect to $\alpha$ whereas *Mother Nature* is attempting to minimize $f(q, \alpha, u)$ with respect to $u$.

Since *Mother Nature* is minimizing $f(q, \alpha, u)$ with respect to $u$ given that $q$ and $\alpha$ have already been fixed by the *Decision Maker*, it follows that the worst outcome given $(q, \alpha)$ is generated by

$$\tilde{u}(q, \alpha) : = \arg \min_{u \in \mathcal{U}(\alpha, \tilde{u})} f(q, \alpha, u) \tag{77}$$

$$= \arg \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \alpha \cdot (r_c \preceq R(q, u)) \tag{78}$$

$$= \arg \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \tag{79}$$

This is a reflection of the fact that $f(q, \alpha, u)$ is non-decreasing with $R(q, u)$.

So if we wish to quickly write a script for this recipe, it will be along these lines:

*The Script for the Info-Gap Game*

A play in three Acts

Act 1: *Decision Maker* selects a $q \in \mathbb{Q}$ and a $\alpha \geq 0$.

Act 2: In response *Mother Nature* selects the worst $u$ in $\mathcal{U}(\alpha, \tilde{u})$, namely $\tilde{u}(q, \alpha)$, according to (79).

Act 3: The *Referee* awards *Decision Maker* the payoff determined by $(q, \alpha, u)$, namely $f(q, \alpha, u)$, as specified in (76).

In short, for each decision $\alpha$, there is a *worst-case analysis* with respect to function $f$ on the region of uncertainty $\mathcal{U}(\alpha, \tilde{u})$.

Whether *Info-Gap* likes it or not, this conceptual framework is *worst-case analysis* par excellence.

Indeed, given the intuitive appeal and expressive power of the classical game theoretic metaphor, it is a a great pity that the *Info-Gap* literature does not use this metaphor to describe the *worst-case analysis* features of the generic *Info-Gap* model.

What a shame!

## A.3 What exactly is wrong with the claims?

Ben-Haim [2005, p. 392] claims that (a) there is no worst case in the *Info-Gap* analysis and (b) that *Info-Gap* is not *Maximin*. The task is to find where exactly does the flaw lie. I shall therefore address some of the arguments on which these claims are based in detail.

### A.3.1 Unbounded horizon of uncertainty

For some strange reason *Info-Gap* mistakenly holds that, as a rule, the region of uncertainty $\mathcal{U}(\alpha, \tilde{u})$ is STRICTLY expanding with $\alpha$, namely that

$$\alpha'' > \alpha' \Longrightarrow \mathcal{U}(\alpha', \tilde{u}) \subset \mathcal{U}(\alpha'', \tilde{u}) \tag{80}$$

where $\subset$ denotes *strict inclusion*.

This is definitely not so.

For example, consider the case where $u$ represents the *probability* of a certain event, $0 < \tilde{u} < 1$, and the regions of uncertainty are defined as follows (see Ben-Haim [2006, p. 256-257]):

$$\mathcal{U}(\alpha, \tilde{u}) = \{u \in [0,1] : (1-\alpha)\tilde{u} \leq u \leq (1+\alpha)\tilde{u}\} \ , \ \alpha \geq 0 \tag{81}$$

Clearly, there is a finite $\alpha^*$ such that $\mathcal{U}(\alpha, \tilde{u}) = [0,1], \forall \alpha \geq \alpha^*$. In fact it is easy to conclude that

$$\alpha^* = \max\left\{1, \frac{1-\tilde{u}}{\tilde{u}}\right\} \tag{82}$$

can be used for this purpose.

Consider for instance the concrete case where $\tilde{u} = 0.5$. Here we have $\alpha^* = 1$ and consequently

$$\mathcal{U}(\alpha, \tilde{u}) = \begin{cases} [0,1] & , \quad \alpha \geq 1 \\ [0.5 - 0.5\alpha, 0.5 + 0.5\alpha] & , \quad 0 \leq \alpha \leq 1 \end{cases} \tag{83}$$

So clearly, the region of uncertainty $\mathcal{U}(\alpha, \tilde{u})$ reaches its largest size when $\alpha = \alpha^* = 1$ and does not expand any further as $\alpha$ is increasing above $\alpha^* = 1$.

In short, in this example we have

$$\mathfrak{U} = \bigcup_{\alpha \geq 0} \mathcal{U}(\alpha, \tilde{u}) = [0,1] \tag{84}$$

So much then for the idea (Ben-Haim [2005, p. 392] ) that the region of uncertainty is increasing in size with $\alpha$ and therefore the total region of uncertainty is unbounded.

### A.3.2 Existence of worst case

For some other strange reason *Info-Gap* mistakenly holds that, as a rule, if the region of uncertainty is unbounded then there is no worst case.

Now this contention is astounding.

As we regularly teach our first year students, *saddle points* exist on (the unbounded) $\mathbb{R}^2$. My favorite saddle point (for teaching purposes) is the beauty associated with the following *Maximin* problem:

$$v^* := \max_{x\in\mathbb{R}} \min_{y\in\mathbb{R}} \left\{y^2 + 2xy - x^2\right\} \tag{85}$$

To obtain the optimal (worst) value of $y$ for a given $x$, we conduct the *worst-case analysis* on the $\boxed{\text{unbounded}}$ feasible range of values of $y$, namely on $Y = \mathbb{R}$:

$$y^*(x) := \arg \min_{y\in\mathbb{R}} \left\{y^2 + 2xy - x^2\right\} \ , \ x \in \mathbb{R} \tag{86}$$

$$= -x \tag{87}$$

So the recipe for the optimal value of $x$ is as follows:

$$x^* := \arg \max_{x\in\mathbb{R}} \left\{y^2 + 2xy - x^2\right\}\Big|_{y=-x} \tag{88}$$

$$= \arg \max_{x\in\mathbb{R}} \left\{(-x)^2 + 2x(-x) - x^2\right\} \tag{89}$$

$$= \arg \max_{x\in\mathbb{R}} \left\{-2x^2\right\} \tag{90}$$

$$= 0 \tag{91}$$

The optimal solution is then the saddle point $(x^*, y^*) = (0, 0)$. Figure 10 displays this point in its full glory.

Note that in this framework the region of uncertainty is $Y = \mathbb{R}$ and this set is $\boxed{\text{unbounded}}$.

So much then for the assertion that *worst-case analysis* cannot be conducted on unbounded uncertainty regions.

More generally, . . .

### A.3.3 Objective function

When you conduct a *worst-case analysis* under severe uncertainty, the existence of a worst case is not determined only by the region of uncertainty under consideration, but also by the *objective function* under consideration.

The point is that even in cases where the region of uncertainty is unbounded there could be a worst case if the objective function is bounded on that region.

For example, suppose that the region of uncertainty is the real line $\mathbb{R}$ and the objective function is $\sin(x)$. Since sin is bounded on the real line there would be a worst case in this instance.

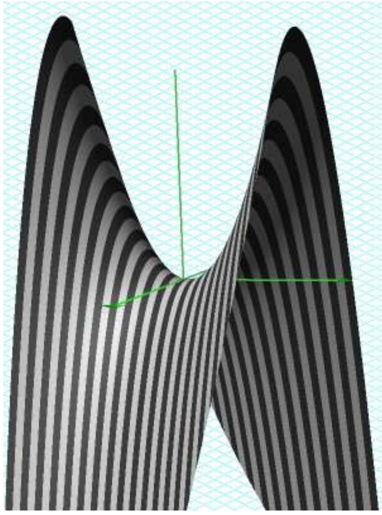

Figure 10: My favorite Saddle Point

In fact, as shown above, a worst case can exist even if the objective function is not bounded: it is sufficient that the function satisfy certain *convexity* conditions (see for example Rustem and Howe [2006]).

But in the framework of *Info-Gap* we need not worry about the general case. It is sufficient to consider the specific objective function deployed by *Info-Gap*. As we have already noted, this function is of the form

$$f(q,\alpha,u) := \alpha\,(r_c \preceq R(q,u)) \ \ , \ q \in \mathbb{Q}, \alpha \geq 0, u \in \mathcal{U}(\alpha,\tilde{u}) \tag{92}$$

For any given pair ($q \in \mathbb{Q}, \alpha \geq 0$) the worst (smallest) value this function takes on $\mathcal{U}(\alpha,\tilde{u})$ is then either 0 or 1. So clearly, for any such pair, there is a worst case.

To see more clearly, consider again the formal explanation. So focus on a given $q \in \mathbb{Q}$ and its robustness:

$$\hat{\alpha}(q,r_c) : = \max_{\alpha\geq 0} \ \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \ \alpha \cdot (r_c \preceq R(q,u)) \tag{93}$$

$$= \max_{\alpha\geq 0} \ \alpha \cdot \min_{u\in\mathcal{U}(\alpha,\tilde{u})} (r_c \preceq R(q,u)) \tag{94}$$

$$= \max_{\alpha\geq 0} \ G(\alpha) \cdot H(q,\alpha) \tag{95}$$

where

$$G(\alpha) := \alpha \ , \ \alpha \geq 0 \tag{96}$$

$$H(q,\alpha) := \min_{u\in\mathcal{U}(\alpha,\tilde{u})} (r_c \preceq R(q,u)) \ \ , \ \mathbb{Q}, \alpha \geq 0 \tag{97}$$

observing that the nesting property of the regions of uncertainty implies that for a given $q$, $H(q,\alpha)$ is a *step function* of $\alpha$, as shown in Figure 11.

This implies that $G(\alpha) \cdot H(q,\alpha)$ consists of two linear parts: on the interval $[0, \hat{\alpha}(q,r_c)]$ this function is equal to $G$. And then on the interval $(\hat{\alpha}(q,r_c), \infty)$ the function is equal to 0, as shown in Figure 12.

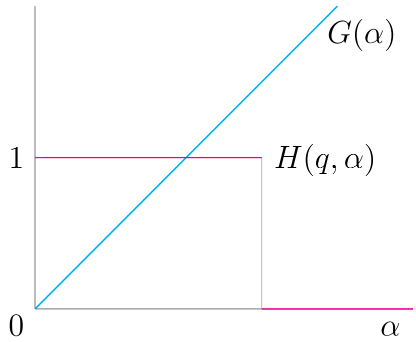


Figure 11: $G = G(\alpha)$ and $H = H(q,\alpha)$, $q$ is fixed.

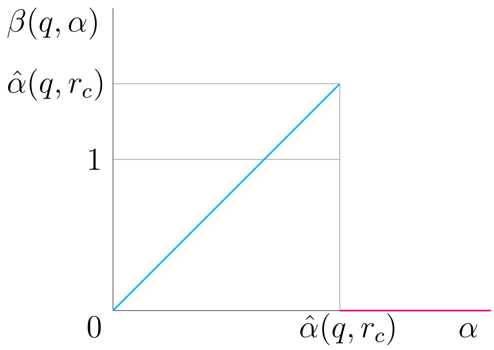


Figure 12: $\beta(q,\alpha) := G(\alpha)\cdot H(q,\alpha)$, $q$ is fixed.

It follows then that the function $\beta = \beta(q,\alpha)$ defined by

$$\beta(q,\alpha) := G(\alpha)\cdot H(q,\alpha)\ ,\ q \in \mathbb{Q}, \alpha \geq 0 \tag{98}$$

initially grows as $G(\alpha) = \alpha$ and then vanishes forever into 0 for $\alpha > \hat{\alpha}(q,r_c)$, as shown in Figure 12. Note that by construction

$$\hat{\alpha}(q,r_c) = \max_{\alpha\geq 0} \beta(q,\alpha)\ ,\ q \in \mathbb{Q} \tag{99}$$

Thus, the robustness of decision $q$, namely $\hat{\alpha}(q,r_c)$, is the *discontinuity* point on the graph of $\beta(q,\alpha)$, where the value of $\beta(q,\alpha)$ drops from $\alpha$ to 0.

In short, for any given value of $q$ and $\alpha$, the objective function of the *Info-Gap* model is bounded below by 0 and above by $\alpha$. You can increase the value of $\alpha$ from here to eternity, yet for each value of $\alpha$ a worst case exists in the region of uncertainty associated with $\alpha$.

The case where $\beta(q,\alpha) = \alpha, \forall \alpha \geq 0$ is of no interest because in this case the robustness of decision $q$ is unbounded, so robustness is not an issue: the requirement $r_c \leq R(q,u)$ is satisfied for all $u \in \mathfrak{U}$.

So much then for the claim (Ben-Haim [2005, p. 392] ) that because *Info-Gap*'s region of uncertainty is unbounded there is no worst case.

Moreover, the *worst-case analysis* conducted by *Info-Gap* is not carried out over the entire region of uncertainty $\mathfrak{U}$, but over the regions $\mathcal{U}(\alpha, \tilde{u}), \alpha > 0$, one region at a time – so to speak.

This brings us to the more fundamental issue.

### A.3.4 Limited vs unlimited worst-case analysis

It is important to point out at the outset that there are various means to control the scope of a *worst-case analysis*. This means that a decision model based on *worst-case analysis* is not necessarily a doomsday scenario predicting the final destruction of the Universe. It is not like that at all.

The availability of tools of thought aimed at limiting the scope of *worst-case analysis* are of the essence if *worst-case analysis* is to be of any use. We should be able, for example, to conduct a *worst-case analysis* of the 100-year flood in a given region even though more extreme floods are possible, such as the 425-year flood and the 2034-year flood, and perhaps Noah's flood.

We have already indicated that *Wald's Maximin Principle* provides a simple mechanism for this purpose.

Let me reiterate this point.

Consider our *Maximin* model

$$v^* := \max_{d \in \mathbb{D}} \min_{s \in S(d)} f(d, s) \tag{100}$$

Here we can control the scope of the *worst-case analysis* by controlling the feasible values that the state of nature $s$ can take given $d$. This is done by defining the sets $S(d) \subseteq \mathbb{S}, d \in \mathbb{D}$ to suit our needs.

That is, in this model not all the states in $\mathbb{S}$ are required to be associated with all the decisions in $\mathbb{D}$. This means that this model admits a *limited worst-case analysis* in the sense that the choices available to *Mother Nature* can be limited by the decision $d \in \mathbb{D}$ made by the decision maker.

Clearly *Info-Gap* is doing exactly this: the robustness of decision $q \in \mathbb{Q}$ involves a LIMITED *worst-case analysis*, to wit:

$$\hat{\alpha}(q, r_c) = \max \left\{ \alpha \geq 0 : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right\} \tag{101}$$

$$= \max_{\alpha \geq 0} \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \alpha \left( r_c \preceq R(q, u) \right) \tag{102}$$

Observe that the values of the uncertainty parameter $u$ are restricted to the region $\mathcal{U}(\alpha, \tilde{u})$ rather than to the total region of uncertainty $\mathfrak{U}$.

If we regard $u$ as the *state of nature* – as we should – then this formulation means that given $q$ and $\alpha$ *Mother Nature* is not allowed to select the worst state in $\mathfrak{U}$. It must select the worst state in $\mathcal{U}(\alpha, \tilde{u})$.

Thus, as in the case of the *Maximin* model where, given $d \in \mathbb{D}$, the state is restricted to $S(d)$, in the *Info-Gap* model, given $q$ and $\alpha$, the uncertainty parameter $u$ is restricted to $\mathcal{U}(\alpha, \tilde{u})$.

In short, as explained above, in the context of the *Info-Gap* model, the worst outcome ($u$) pertaining to a $(q, \alpha)$ pair is determined as follows:

$$\tilde{u}(q, \alpha) := \arg \min_{u \in \mathcal{U}(\alpha, \tilde{u})} \alpha \left( r_c \preceq R(q, u) \right) \tag{103}$$

$$= \arg \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \tag{104}$$

So *Info-Gap* conducts a *limited worst-case analysis* par excellence.

Clearly then, the assertion (Ben-Haim [2005, p. 491]) that "...the robustness function is *not* a worst-case or minimax assessment ..." is groundless. The robustness function specified by (102) is patently a typical *worst-case* assessment. The decision variables $q$ and $\alpha$ are assessed by the worst value of $u$ in $\mathcal{U}(\alpha, \tilde{u})$.

Here is how this notion is described in the *Info-Gap* book:

> The robustness $\hat{\alpha}(q)$ of decision vector $q$ is the largest value of the horizon of uncertainty $\alpha$ for which a specified minimal requirement is always satisfied.
>
> Ben-Haim [2006, p. 38]

The key term here is "always": this minimal requirement must be satisfied by $\boxed{\text{all}}$ $u$ in $\mathcal{U}(\alpha, \tilde{u})$. This means that the acceptability of the value of $\alpha$ under consideration is assessed by the worst performance level over the region of uncertainty under consideration, namely $\mathcal{U}(\alpha, \tilde{u})$.

Indeed, we can rephrase the quote as follows:

> The robustness $\hat{\alpha}(q)$ of decision vector $q$ is the largest value of the horizon of uncertainty $\alpha$ for which the worst performance in this region is not below the specified requirement $r_c$.

The issue here is not which of these two similar descriptions is more informative. The point is that the *worst performance* over the uncertainty region is used to assess whether a given $\alpha$ value is acceptable.

The *Maximin* representation of the *Info-Gap* model makes this point crystal clear. The generic *Info-Gap* formulation says the same thing but is not so explicit about it.

## A.4 An important modeling issue

As we have demonstrated above, *Info-Gap* conducts a "limited" *worse-case analysis*: given $q$ and $\alpha$, the worst case of $u$ is restricted to the region of uncertainty $\mathcal{U}(\alpha, \tilde{u})$ which could be much smaller than the complete region of uncertainty $\mathfrak{U}$.

So how would you respond to the following plea for help?

In any case, the good news is that such a recipe is available. In fact, the recipe is straight forward.

Dear Sir/Madam:

It was just announced here that starting 7:00AM tomorrow, any limited *worst-case analysis* will be prohibited in the Institute. Only unlimited *worst-case analysis* will be allowed.

Violation of this new regulation will be dealt with severely.

My problem is that I really like limited *worst-case analysis* and I use it extensively. I would therefore like to continue doing this kind of analysis in the future.

Do you know of any recipe that will allow me to disguise my limited *worst-case analysis* as an unlimited *worst-case analysis*?

I am looking forward to your advice on this matter.

Best wishes

Borat

A call for help!

All you have to do, Borat, is write the *Info-Gap* robustness as follows:

$$\hat{\alpha}(q, r_c) = \max\left\{\alpha \geq 0 : r_c \leq \min_{u \in \mathfrak{U}} \tilde{R}(q, \alpha, u)\right\} \tag{105}$$

where

$$\tilde{R}(q, \alpha, u) := \begin{cases} R(q, u) & , \quad u \in \mathcal{U}(\alpha, \tilde{u}) \\ -\infty & , \quad u \notin \mathcal{U}(\alpha, \tilde{u}) \end{cases} \tag{106}$$

Note that since here *Mother Nature* is allowed to select the worst state in $\mathfrak{U}$, rather than in $\mathcal{U}(\alpha, \tilde{u})$, this is indeed an *unlimited worst-case analysis*.

If you have nothing else to do this evening, you might consider spending a minute or two showing formally that

$$\max\left\{\alpha : r_c \leq \min_{u \in \mathfrak{U}} \tilde{R}(q, \alpha, u)\right\} = \max\left\{\alpha : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, \alpha, u)\right\} \tag{107}$$

for all $q \in \mathbb{Q}$ and $\alpha \geq 0$, thus confirming that *Info-Gap* is indeed also a typical *unlimited worst-case analysis.*

Wooooohhhhaaa!!!!

**A note to my students:**
This is an excellent item for the final exam. Make sure, therefore, that you complete this exercise on your own.

## A.5 Implications

The fact that *Info-Gap* is a *limited worst-case analysis* is of no major significance, and should be regarded as a modeling technicality. Note that as we have shown, we can also express *Info-Gap* as an *unlimited worst-case analysis* by slightly modifying the model.

However, the nature of the limitation on the scope of the *worst-case analysis* incorporated in the *Info-Gap* model is significant, in fact crucial, for the *interpretation of the results*. Given that the worst-case analysis is limited to the regions $\mathcal{U}(\alpha, \tilde{u}), 0 \leq \alpha \leq \hat{\alpha}(r_c)$, the ROBUSTNESS of the solution is LOCAL rather than GLOBAL in nature. Typically, only regions in the immediate neighborhood of the estimate $\tilde{u}$ are explored.

This is the reason that in Section 8 I drew attention to the fact that there is no ground to believe that under severe uncertainty the solutions generated by *Info-Gap* are likely to be robust: under severe uncertainty this estimate is expected to be a poor indication of the true value of $u$ and is likely to be substantially wrong.

As a result, the very limited scope of its *worst-case analysis* prevents *Info-Gap* from probing the complete region of uncertainty $\mathfrak{U}$. Therefore, there is no reason to believe ...

I strongly recommend that users of this theory should warn the public about this aspect of *Info-Gap* by incorporating the estimate $\tilde{u}$ in the notation for the robustness. Thus,

$$\hat{\alpha}(q, r_c) := \max \left\{ \alpha \geq 0 : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right\} \tag{108}$$

should be re-written as

$$\hat{\alpha}(q, r_c | \tilde{u}) := \max \left\{ \alpha \geq 0 : r_c \leq \min_{u \in \mathcal{U}(\alpha, \tilde{u})} R(q, u) \right\} \tag{109}$$

and

$$\hat{\alpha}(r_c) := \max_{q \in \mathbb{Q}} \hat{\alpha}(q, r_c) \tag{110}$$

should be written as

$$\hat{\alpha}(r_c | \tilde{u}) := \max_{q \in \mathbb{Q}} \hat{\alpha}(q, r_c | \tilde{u}) \tag{111}$$

And more importantly, the following sticker should be attached to all reports generated by *Info-Gap* models:

> ***Public Warning***
>
> Be careful when you interpret the results generated by *Info-Gap* models. The worst-case analysis incorporated in these models is very limited in scope so that the robustness defined by these models is inherently very *local* in nature. Therefore, there is no reason to believe that, under severe uncertainty, the solutions generated by such models are likely to be robust.

And the usual picture I show in support of this argument should be posted next to this warning sign.

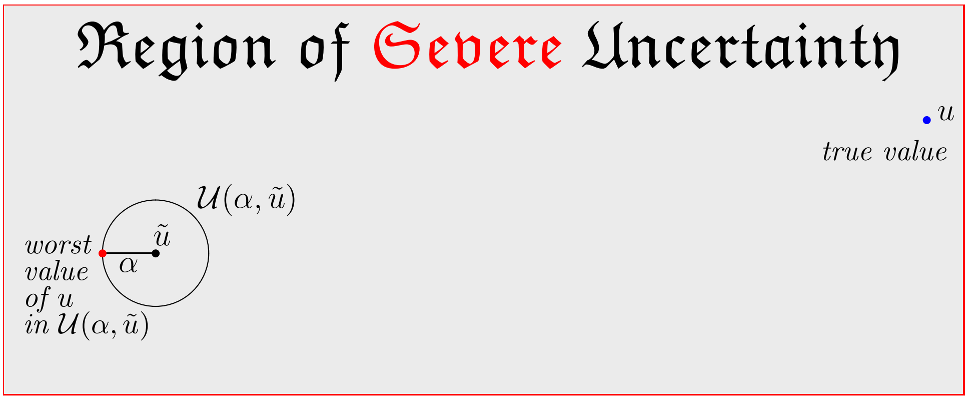


Figure 13: See Public Warning

## A.6 Summary

There are several ways to interpret the generic *Info-Gap* recipe for determining the robustness of a decision $q$:

$$\hat{\alpha}(q,r_c) := \max\left\{\alpha \geq 0 : r_c \leq \min_{u\in\mathcal{U}(\alpha,\tilde{u})} R(q,u)\right\} \tag{112}$$

I have shown here that a natural interpretation of this recipe views the *Info-Gap* model as a vanilla application of the classical *Maximin Principle*. That is, it is natural to view the generic *Info-Gap* problem as *Maximin* game a decision maker plays against *Nature*, with the implied *worst-case analysis* connection:

$$\hat{\alpha}(q,r_c) := \max_{\alpha\geq 0} \min_{u\in\mathcal{U}(\alpha,\tilde{u})} \alpha\cdot(r_c \preceq R(q,u)) \tag{113}$$

The advantage of such an interpretation is that it brings with it an enormous body of knowledge and literature that has been growing steadily over the past 60 years or so.

I have also shown that the arguments put forward by *Info-Gap* to dissociate itself from *worst-case analysis* vividly illustrate the degree to which *Info-Gap* is unaware of ... what it actually does.

So here is a provocative question to all *Info-Gap* aficionados out there who despite my effort here still believe that *Info-Gap* is not all a simple *worse-case analysis*:

What is the worst value of $u$ in $\mathcal{U}(\alpha,\tilde{u})$ for a decision maker whose objective in life is to maximize the payoff $f(q,\alpha,u)$ stipulated by

$$f(q,\alpha,u) := \begin{cases} \alpha & , \ r_c \leq R(q,u) \\ -\infty & , \ \textit{otherwise} \end{cases} \quad , \alpha \geq 0 \tag{114}$$

assuming that the values of $q$ and $\alpha$ have already been determined?

If your answer is something along the following lines, you should accept the fact that *Info-Gap* is a simple instance of Maximin:

> Clearly, for any given $q$ and $\alpha$ the value of $f(q, \alpha, u)$ is non-decreasing with $R(q, u)$. Thus, the worst value of $u$ in $\mathcal{U}(\alpha, \tilde{u})$ for the decision maker is one that makes $R(q, u)$ as small as possible. Therefore, to hurt the decision maker as much as possible I should select a $u$ in $\mathcal{U}(\alpha, \tilde{u})$ that minimizes $R(q, u)$ over $u$ in $\mathcal{U}(\alpha, \tilde{u})$.

If you accept this interpretation you are in good company!

Furthermore, in this case you will be able to utilize the well established *Robust Optimization* literature where similar recipes are designated at the outset as *Maximin*-based ideas (see for example Restum and Howe [2002] and Kouvelis and Yu [1997]). This literature will also furnish you with guidance and inspiration for fixing the fundamental flaw in the *Info-Gap* model resulting from the very local nature of its worst-case analysis.