

Lecture Notes on Info-Gap Uncertainty

Yakov Ben-Haim
Faculty of Mechanical Engineering
Technion — Israel Institute of Technology
Haifa 32000 Israel
yakov@technion.ac.il

Source material: Yakov Ben-Haim, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, 2nd edition, Academic Press. Chapter 2.

A Note to the Student: These lecture notes are not a substitute for the thorough study of books. These notes are no more than an aid in following the lectures.

Contents

1	Central Ideas	2
2	Risk and Uncertainty	3
2.1	Three-Box Riddle: Probabilistic Version	3
2.2	Three-Box Riddle: Non-Probabilistic Version	7
3	Principle of Indifference	8
3.1	2-Envelope Riddle	9
3.2	Keynes' Example	10
3.3	Shackle-Popper Indeterminism	12
3.4	Influenza Pandemic	13
4	Historical Perspectives on Reasoning and Uncertainty	15
5	Info-gap Uncertainty, Probability and Fuzziness	17
6	Some Info-gap Models of Uncertainty	21
6.1	Energy-bound Models	22
6.2	Envelope-bound Models	23
6.3	Slope-bound info-gap models	25
6.4	Fourier-bound info-gap models	25
7	Why Convex Info-gap Models?	26
7.1	Most Info-Gap Models are Convex: Examples	26
7.2	Why Are Most Info-Gap Models Convex?	28

1 Central Ideas

¶ Decisions.

- **Animals:**
 - Evolve subject to real world constraints.
 - Make decisions.
- **Homo sapiens:**
 - Thinker.
 - Tool maker.
 - Decision maker.
 - Psychology: relation to evolution.
 - Knowledge: relation to real world constraints.
- **Decisions:**
 - Have consequences: survival, reward, punishment, etc.
 - Are made under uncertainty, competition, resource deficiency.
- **Decision strategies:** what strategies are available?

¶ Decision strategies:

- *Satisficing:* Doing good enough. Beating the competition.
- *Optimizing:*
 - Use best model to find most rewarding decision.
 - Psychology: Optimizers have more than satisficers, but satisficers are happier (Schwarz).
- *Windfalling:* Exploiting opportunity; facilitating better-than-expected outcomes.
- Which strategy to use? When?

¶ Robust-Satisficing:

- Satisficing is sub-optimal, therefore multiple solutions.
- Use extra DoF's to robustify against surprise.
- Robustness and probability of success:
 - Animal foraging seems to be robust-satisficing, not intake-maximization.
 - Why? Survival value: greater probability of attaining critical need.

¶ Robustness and Opportuneness: Antagonism or sympathy?

¶ The big challenge: uncertainty.

- What is uncertainty?
 - Epistemic.
 - Ontological.
- How to model and manage uncertainty?
 - Probability and statistics.
 - Info-gap theory.
 - Fuzzy logic.
 - Dempster-Shafer theory.
 - More.
- We will focus on probability and info-gap theory.

2 Risk and Uncertainty

2.1 Three-Box Riddle: Probabilistic Version

¶ Background to the riddle:

- Prize in 1 of 3 boxes.
- We must choose a box:

?
?
?

- We know prior probabilities are equal: $P_1 = P_2 = P_3 = 1/3$.
 P_i = probability that prize is in i th box.
- For choice of box i , the probability of failure is:

$$P_f = 1 - P_i = \frac{2}{3} \tag{1}$$

- We choose a box, call it ‘C’.

C
?
?

- Then the M.C., who knows where the prize is, opens a box, ‘E’, and shows us that it is empty:

C
E
T

The third box, ‘T’, remains closed.

¶ The Riddle: is there a rational basis for changing our choice? $C \implies T$?

¶ Probabilistic analysis:

- Our prior knowledge is: $\text{Prob}(C) = 1/3$.
- By normalization of the probability: $\text{Prob}(\text{not } C) = 1 - \text{Prob}(C) = 2/3$.
- The prize being in ‘E’ or in ‘T’ are disjoint events. Thus:
 $\text{Prob}(\text{not } C) = \text{Prob}(E \text{ or } T) = \text{Prob}(E) + \text{Prob}(T)$
- After the M.C. opens ‘E’, we know that $\text{Prob}(E) = 0$.
- Thus $\text{Prob}(T) = 2/3$.
- Thus changing from ‘C’ to ‘T’ increases the probability of success by a factor of two!

Hence: change the choice.

¶ **Skeptical? Consider the 1000-box version:**

- Prize in 1 of 1000 boxes.
- We know prior probabilities are equal: $P_1 = \dots = P_{1000} = 1/1000$.
- For choice of box i , the probability of failure is:

$$P_f = 1 - P_i = 0.999 \quad (2)$$

- We choose a box, call it 'C'.
- Then the M.C., who knows where the prize is, opens and shows us 998 empty boxes, leaving 'C' and 'T' closed.
- **Question:** is there a rational basis for changing our choice? $C \implies T$?

¶ **Probabilistic analysis:**

- Our prior knowledge is: $\text{Prob}(C) = 0.001$.
 - By normalization of the probability: $\text{Prob}(\text{not } C) = 1 - \text{Prob}(C) = 0.999$.
 - Let 'E' denote the 998 empty boxes.
 - The prize being in 'E' or in 'T' are disjoint events. Thus:

$$\text{Prob}(\text{not } C) = \text{Prob}(E \text{ or } T) = \text{Prob}(E) + \text{Prob}(T)$$
 - After the M.C. opens 998 empty boxes, we know that their cumulative probability is zero: $\text{Prob}(E) = 0$.
 - Thus $\text{Prob}(T) = 0.999$.
 - Thus changing from 'C' to 'T' increases the probability of success by a factor of 999.
- Hence: change the choice.

¶ **Still skeptical?** Do an experiment with 5 boxes.

- Two players. One is the M.C., the other chooses a box.
- The “Chooser” picks a strategy before play, either ‘change after opening’, or ‘don’t change after opening’.
- The M.C. uses a random number generator to pick the winning box, but does not reveal the winning box to the Chooser. For 5 boxes, the winning box is i if the random number r is in the interval $(i - 1)/5 \leq r < i/5$.
- The Chooser chooses a box.
- The M.C. opens 3 boxes different from the chosen box and from the winning box.
- After the Chooser implements his strategy, record whether the Chooser won or not.
- Repeat this a number of times to test the Chooser’s strategy. With 5 boxes, 10 or 15 repetitions should convince you which is the best strategy. Use table 1 on p.5.

Trial #	Chosen box (C)	Remaining Box (T)	True Box	Success (Y/N)
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

Table 1: Experimental confirmation of optimal switching strategy.

¶ **Return to the 3-box problem**, in more general form:

- Prize in 1 of 3 boxes, and we must choose a box.
- We know prior probabilities, P_1, P_2, P_3 , but they are not necessarily equal.
- For choice of box i , the probability of failure is:

$$P_f = 1 - P_i \quad (3)$$

- We choose a box, call it 'C'.
- Then the M.C., who knows where the prize is, opens an empty box, 'E'.
- **Question:** is there a rational basis for changing our choice? $C \implies T$?

¶ **General Probabilistic analysis:**

- Known prior probability distribution: P_1, P_2, P_3 .
- Our prior knowledge is: $\text{Prob}(C) = P_C$ which is known.
- By normalization of the probability: $\text{Prob}(\text{not } C) = 1 - P_C$.
- The prize being in 'E' or in 'T' are disjoint events. Thus:

$$\text{Prob}(\text{not } C) = \text{Prob}(E \text{ or } T) = \text{Prob}(E) + \text{Prob}(T)$$
- After the M.C. opens 'E', we know that $\text{Prob}(E) = 0$.
- Thus $\text{Prob}(T) = 1 - P_C$.
- Change from 'C' to 'T' if and only if $1 - P_C > P_C$:

$$\text{Change if:} \quad P_C < 1/2 \quad (4)$$

$$\text{Don't change if:} \quad P_C > 1/2 \quad (5)$$

$$\text{Indifferent if:} \quad P_C = 1/2 \quad (6)$$

• **Conclusions:**

- There is an infinity of prior distributions P_1, P_2, P_3 for which we **would not** change the chosen box.
- There is an infinity of prior distributions P_1, P_2, P_3 for which we **would** change the chosen box.
- The choice of a strategy depends upon knowledge of the prior probability distribution.

2.2 Three-Box Riddle: Non-Probabilistic Version

¶ Background to the problem:

- Prize in 1 of 3 boxes.
- We must choose a box: $\boxed{?}$ $\boxed{?}$ $\boxed{?}$
- We **do not know** the prior probabilities, P_1, P_2, P_3 .
- If $P_i > 1/2$ then we would choose box i and stick with it.

¶ Can we guess the prior probabilities?

• Suppose we have a rough guess of the prior distribution. How should we use this information?

- Can we use the principle of indifference to choose the prior distribution?

• **Principle of indifference:** ascribe equal probabilities to elementary events about which no other information is available. Hence uniform distribution:

$$(P_1, P_2, P_3) = (1/3, 1/3, 1/3) \quad (7)$$

- Uniform distribution has maximum entropy:

$$H(P_1, P_2, P_3) = - \sum_{N=1}^3 P_N \ln P_N$$

From among all possible probability distributions.

- ‘Max entropy’ does not imply ‘no information’.
- Knowing a specific probability distribution is knowledge.

¶ The uncertainty is an **information gap**

between what we **do know** and **what we need to know**.

- We have no information, not even probabilistic.

3 PRINCIPLE of INDIFFERENCE

§ Question:

Is ignorance probabilistic?

§ Principle of indifference (Bayes, LaPlace, Jaynes, ...):

- Elementary events, about which nothing is known, are assigned **equal probabilities**.
- Uniform distribution represents complete ignorance.

§ The info-gap contention:

The probabilistic domain of discourse does not encompass all epistemic uncertainty.

3.1 *2-Envelope Riddle*

§ The riddle:

- You are presented with two envelopes.
 - Each contains a positive sum of money.
 - One contains twice the contents of the other.
- You choose an envelope, open it, and find \$50.
- Would you like to switch envelopes?

§ You reason as follows:

- Other envelope contains either \$25 or \$100.
- Principle of indifference:
- Assume equal probabilities.
The expected utility upon switching is:

$$\text{E.U.} = \frac{1}{2} \$25 + \frac{1}{2} \$100 = \$62.50.$$

$$\$62.50 > \$50.$$
- Yes! Let's switch, you say.

§ The riddle, re-visited:

- You are presented with two envelopes.
 - Each contains a positive sum of money.
 - One contains twice the contents of the other.
- You choose an envelope, but do not open it.
- Would you like to switch envelopes?

§ You reason as follows:

- This envelope contains \$ X > \$0.
- Other envelope contains either \$ $2X$ or \$ $\frac{1}{2}X$.
- Principle of indifference:
- Assume equal probabilities.

The expected utility upon switching is:

$$\text{E.U.} = \frac{1}{2} \$2X + \frac{1}{2} \$\frac{1}{2}X = \$\left(1 + \frac{1}{4}\right)X > X.$$

- Yes! Let's switch, you say.

§ You wanna switch again? And again? And again?

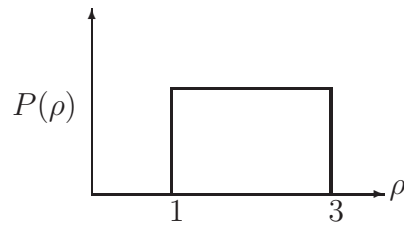
3.2 *Keynes' Example*

§ ρ = specific gravity [g/cm³] is unknown:

$$1 \leq \rho \leq 3$$

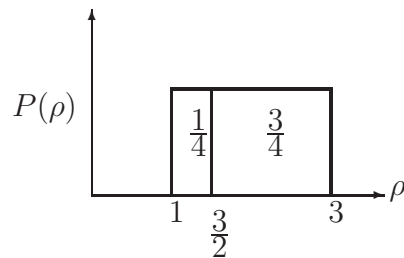
§ Principle of indifference:

Uniform distribution in [1, 3], so:



§ Uniform distribution in [1, 3], so:

$$\mathbf{PROB} \left(\frac{3}{2} \leq \rho \leq 3 \right) = \frac{3}{4}$$

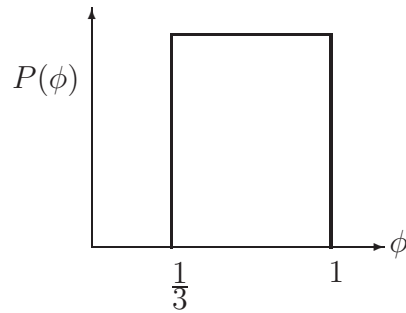


§ ϕ = specific volume [cm^3/g] is unknown:

$$\frac{1}{3} \leq \phi \leq 1$$

§ Principle of indifference:

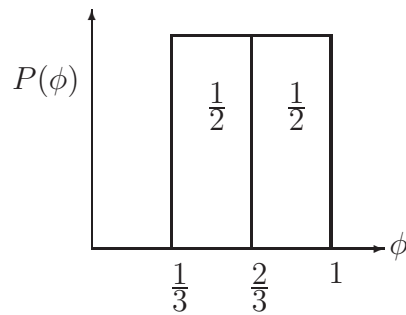
Uniform distribution in $\left[\frac{1}{3}, 1\right]$, so:



§ Principle of indifference:

Uniform distribution in $\left[\frac{1}{3}, 1\right]$, so:

$$\text{PROB} \left(\frac{1}{3} \leq \phi \leq \frac{2}{3} \right) = \frac{1}{2}$$



§ Contradiction:

$$\frac{1}{2} = \underbrace{\text{PROB} \left(\frac{1}{3} \leq \phi \leq \frac{2}{3} \right)}_{\text{Specific volume}} = \underbrace{\text{PROB} \left(\frac{3}{2} \leq \rho \leq 3 \right)}_{\text{Specific gravity}} = \frac{3}{4}$$

§ The Culprit:

- Principle of indifference.

§ The resolution:

- Ignorance is not probabilistic.
- Ignorance is an info-gap.

3.3 *Shackle-Popper Indeterminism*

§ Intelligence:

What people know,
influences how they behave.

§ Discovery:

What will be discovered tomorrow
cannot be known today.

§ Indeterminism:

Tomorrow's behavior cannot be
modelled completely today.

§ Information-gaps, indeterminisms,
sometimes

cannot be modelled probabilistically.

§ Ignorance is not probabilistic.

3.4 INFLUENZA PANDEMIC

This section is based on the following World Health Organization (WHO) website:

http://www.who.int/csr/disease/avian_influenza/en

The following is a quote from that site:

Estimating the impact of the next influenza pandemic: enhancing preparedness

Geneva, 8 December 2004

Influenza pandemics are recurring and unpredictable calamities. WHO and influenza experts worldwide are concerned that the recent appearance and widespread distribution of an avian influenza virus, Influenza A/H5N1, has the potential to ignite the next pandemic.

Given the current threat, WHO has urged all countries to develop or update their influenza pandemic preparedness plans (see information on web pages below) for responding to the widespread socioeconomic disruptions that would result from having large numbers of people unwell or dying.

— Pandemic preparedness

Central to preparedness planning is an estimate of how deadly the next pandemic is likely to be. Experts' answers to this fundamental question have ranged from 2 million to over 50 million. All these answers are scientifically grounded. The reasons for the wide range of estimates are manifold.

Some estimates are based on extrapolations from past pandemics but significant details of these events are disputed, including the true numbers of deaths that resulted. The most precise predictions are based on the pandemic in 1968 but even in this case estimates vary from one million to four million deaths. Similarly, the number of deaths from the Spanish flu pandemic of 1918 is posited by different investigators to range from 20 million to well over 50 million. Extrapolations are problematic because the world in 2004 is a different place from 1918. The impact of greatly improved nutrition and health care needs to be weighed against the contribution the increase in international travel would have in terms of global spread. The specific characteristics of a future pandemic virus cannot be predicted. It may affect between 20-50% of the total population. It is also unknown how pathogenic a novel virus would be, and which age groups will be affected. The level of preparedness will also influence the final death toll. Even moderate pandemics can inflict a considerable burden on the unprepared and disadvantaged. Planning to maintain health care systems will be especially crucial. Good health care will play a central role in reducing the impact, yet the pandemic itself may disrupt the supply of essential medicines and health care workers may fall ill. Because of these factors, confidently narrowing the range of estimates cannot be done until the pandemic emerges. Therefore, response plans need to be both strong and flexible.

Even in the best case scenarios of the next pandemic, 2 to 7 million people would die and tens of millions would require medical attention. If the next pandemic virus is a very virulent strain, deaths could be dramatically higher.

The global spread of a pandemic cannot be stopped but preparedness will reduce its impact. WHO will continue to urge preparedness and assist Member States in these activities. In the next few weeks, WHO will be publishing a national assessment tool to evaluate and focus national preparedness efforts. WHO will also be providing guidance on stockpiling antivirals and vaccines. Next week, WHO will be convening an expert meeting on preparedness planning. WHO is also working to advance development of pandemic virus vaccines, and to expedite research efforts to understand the mechanisms of emergence and spread of influenza pandemics.

It is of central importance that Member States take the necessary steps to develop their own preparedness plans. Some have already developed structures and processes to counter this threat but some plans are far from complete and many Member States have yet to begin. WHO believes the appearance of H5N1, which is now widely entrenched in Asia, signals that the world has moved closer to the next pandemic. While it is impossible to accurately forecast the magnitude of the next pandemic, we do know that much of the world is unprepared for a pandemic of any size.

4 Historical Perspectives on Reasoning and Uncertainty

(Ch. 2, sec. 1, pp.9–12)

¶ Four stages:

- I. Deduction: ancient Greece ...
- II. Forward probability: 1600, Pascal, Fermat, ...
- III. Inverse probability: 1760, Bayes, ...
- IV. Modern inference.

I. Symbolic Logic, mathematical deduction:

- Axiomatic geometry (Euclid, c.200 BCE)
- Symbolic logic, syllogism (Aristotle, c.200 BCE):
 1. All men are mortal.
 2. Socrates is a man.
 3. Therefore Socrates is mortal.

II. Games of chance (Pascal, Fermat, etc., 1600 ...)

- Forward probability:
Given **known coin**, what is the probability of any specified sequence?
- Probabilistic syllogism for sequence of coin flips with **known coin**: p_H, p_T known.
 1. Probability of specific sequence of Heads and Tails is calculated as $p_H^{N_H} p_T^{N_T}$.
 2. Given the specific sequence $HHTH$.
 3. Hence the probability of $HHTH$ is known and equals ...

III. Inference: Inverse probability.

- Given **unknown coin**, which has produced a given sequence, what are the probabilities of the individual basic events? (Bayes, 1760, studying dice).
- Least-squares estimation:
Given noisy measurements of celestial motions, what is the best estimate of an elliptical orbit? (Gauss, Legendre, 1800)
- Central limit theorem (Laplace 1815):
Given measured mean and variance, we know the pdf.
- Many other distributions and associated statistical tests. (Karl Pearson and many others, 19th, 20th c.)
- Probabilistic syllogism for sequence of coin flips:
 1. **If** probabilities of Heads and Tails are P_H, P_T ,
Then probability of observation is $P(HHTH|P_H, P_T)$.
 2. Observed sequence $HHTH$.
 3. Hence estimates of P_H, P_T are ...

IV. Modern design, decision and inference.

- Verbal uncertainty (Lukaceiwicz, 1920; Zadeh, 1960).
“If it rattles, tighten it.”

Uncertainties:

What is a rattle?

How tight is tight?

- Innovation under **time-pressure**.
- **New manufacturing technologies:**
Exchangeable parts, assembly line, automation, robotics, etc.
- **New managerial techniques:**
planning, distributed control, data bases, networks, etc.
- **New materials:**
plastics, superconductors, etc.
- **New scientific discoveries:**
solid state physics, quantum effects, etc.
- **Uncertainties:**
 - Lack of data, information, understanding.
 - Mutually inconsistent data.
 - Disparity between
what **is known**
and what **must be known**
for a good decision.

This is an **information-gap uncertainty**.

¶ Each stage evolved

new modes of reasoning
in order to deal with
new problems

¶ Summary of stages:

I. Deduction:

Mathematical question with unique inescapable answer.

II. Forward probability:

Math applied to uncertainty, with unique inescapable answer.

III. Inverse probability:

Math applied to uncertainty, with non-unique or incomplete answer.

IV. Modern high-speed decision-making:

Severe uncertainty; info-gaps.

¶ Example: Toxic waste.

- Toxic waste is released from industrial plant into river.
- The waste causes damage to plants, animals, humans.
- The waste is removed by various processes:
 - Chemical reaction, adsorption, absorption.
- These are very complex processes. Depend on many factors:
 - Time, temperature, concentration, surface conditions, etc.
- Accurate model of removal rate: very expensive.
- $\tilde{r}(x, t, w)$ = current best model of removal rate.
- x = position, t = time, w = other parameters.
- $\tilde{r}(x, t, w)$ may be simple or complex.
- $r(x, t, w)$ = unknown true model of removal rate.
- $\tilde{r}(x, t, w)$ and $r(x, t, w)$ differ in unknown ways.
- Too little info and understanding to evaluate:
 - probability of deviation of \tilde{r} from r .
 - plausibility of deviation of \tilde{r} from r .
- Neither probabilistic nor fuzzy uncertainty model is accessible.
- Severe uncertainty. Use info-gap model for uncertainty in $r(x, t, w)$.
- Family of nested sets of r -functions.

¶ Fractional-error info-gap model for uncertainty in $r(x, t, w)$:

$$\mathcal{U}(h, \tilde{r}) = \left\{ r(x, t, w) : \left| \frac{\tilde{r}(x, t, w) - r(x, t, w)}{\tilde{r}(x, t, w)} \right| \leq h \right\}, \quad h \geq 0 \quad (8)$$

h is the unknown **horizon of uncertainty**.

$\tilde{r}(x, t, w)$ is the **nominal model**.

¶ Two levels of uncertainty:

- Unknown fractional error of \tilde{r} in $\mathcal{U}(h, \tilde{r})$ at given h .
- Unknown horizon of uncertainty, h .

¶ Nesting of info-gap models:

$$h < h' \implies \mathcal{U}(h, \tilde{r}) \subseteq \mathcal{U}(h', \tilde{r}) \quad (9)$$

¶ Contraction of info-gap models:

$$\mathcal{U}(0, \tilde{r}) = \{\tilde{r}\} \quad (10)$$

¶ Info-gap models organize events into sets.

¶ Probability also organizes events into sets.

E.g.: What is prob. of r : $r \leq \rho$?

¶ Likewise in possibility theory with linguistic uncertainty:

Is the set of r -functions with 'low' values 'highly' possible?

¶ The **difference between info-gap** uncertainty and **probability or possibility** theory is that in info-gap theory we do not use measure functions

to assess uncertainty.

Clustering: yes.

Measures: no.

Insufficient information for measure functions.

An info-gap model is a weaker characterization of uncertainty than probability or possibility models.

The distinction between info-gap models and probability or possibility models is particularly important in dealing with **rare events**.

¶ **Form** of the pdf:

- Once the **form** of a pdf is known, its parameters (e.g. μ , σ for normal distribution) are determined from the bulk of “usual” events.
- Then the frequencies of rare events are predicted.
- A probability model entails very informative predictive power.
- However, the tails may err greatly if the **form** is wrong.
- This can lead to bad decisions.
- Usually rare events — catastrophes — concern the decision maker.
- Probability models are useful when we know the **form** of the model.
E.g.: sampling problems.

6 Some Info-gap Models of Uncertainty

(Section 2.3)

¶ An info-gap model is a family of nested sets all having a common property.

¶ In general terms, the structure of the sets is chosen as:

the smallest or strictest family of sets consistent with given prior information.

6.1 Energy-bound Models

(pp.16–17)

¶ Transient uncertainty:

- Very uncertain.
- Temporary.
- Possibly large disturbance.
- Possibly rapid variation.

¶ Examples:

- Seismic loads.
- Currency fluctuations.
- Severe meteorological anomalies (e.g. extreme wind loads.)

¶ Cumulative energy-bound info-gap model for scalar functions.

- What is known:
 - $\tilde{u}(t)$ = “typical” or “nominal” function.
 - $u(t)$ = “actual” function.

$$\mathcal{U}(h, \tilde{u}) = \left\{ u(t) : \int_0^\infty [u(t) - \tilde{u}(t)]^2 dt \leq h^2 \right\}, \quad h \geq 0 \quad (11)$$

The bounded integral is an “energy of deviation”.

E.g.: unknown variation in seismic energy.

Family of nested sets:

$$h < h' \quad \implies \quad \mathcal{U}(h, \tilde{u}) \subseteq \mathcal{U}(h', \tilde{u}) \quad (12)$$

Contraction:

$$\mathcal{U}(0, \tilde{u}) = \{\tilde{u}\} \quad (13)$$

¶ Two levels of uncertainty:

- Unknown u in $\mathcal{U}(h, \tilde{u})$ at given h .
- Unknown horizon of uncertainty, h .

¶ Cumulative energy-bound info-gap model for vector functions.

$$\mathcal{U}(h, \tilde{u}) = \left\{ u(t) : \int_0^\infty [u(t) - \tilde{u}(t)]^T V [u(t) - \tilde{u}(t)] dt \leq h^2 \right\}, \quad h \geq 0 \quad (14)$$

¶ Instantaneous energy-bound info-gap model for vector functions.

$$\mathcal{U}(h, \tilde{u}) = \left\{ u(t) : [u(t) - \tilde{u}(t)]^T V [u(t) - \tilde{u}(t)] \leq h^2 \right\}, \quad h \geq 0 \quad (15)$$

¶ **Differences** between functions in cumulative energy-bound

and

instantaneous energy-bound models:

Cumulative:

may be unbounded.

asymptotically vanishing.

Instantaneous:

bounded only.

may be persistent.

6.2 Envelope-bound Models

(pp.18–20)

¶ Sometimes prior information about the uncertain functions is of the type:

- “Nominal” or “typical” function $\tilde{u}(t)$.
- Known **shape** of bounding envelope.
- Unknown **size** of bounding envelope.

E.g. $u(t)$ may tend to vary within an interval $\propto \tilde{u}(t)$:

$$\mathcal{U}(h, \tilde{u}) = \{u(t) : |u(t) - \tilde{u}(t)| \leq h\tilde{u}(t)\}, \quad h \geq 0 \quad (16)$$

More generally:

$$\mathcal{U}(h, \tilde{u}) = \{u(t) : |u(t) - \tilde{u}(t)| \leq h\psi(t)\}, \quad h \geq 0 \quad (17)$$

where $\psi(t)$ is the known envelope function.

¶ Example: uncertain variation constrained to the domain $a \leq t \leq b$.

$$\psi(t) = \begin{cases} 1, & a \leq t \leq b \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

¶ Example: unknown fractional variation.

$$\psi(t) = |\tilde{u}(t)| \quad (19)$$

So:

$$\mathcal{U}(h, \tilde{u}) = \left\{ u(t) : \left| \frac{u(t) - \tilde{u}(t)}{\tilde{u}(t)} \right| \leq h \right\}, \quad h \geq 0 \quad (20)$$

¶ Example: an envelope-bound info-gap model for vectors:

$$\mathcal{U}(h, \tilde{u}) = \{u(t) : |u_n(t) - \tilde{u}_n(t)| \leq h\psi_n(t), \quad n = 1, \dots, N\}, \quad h \geq 0 \quad (21)$$

¶ Envelope-bound info-gap models can be formulated in another way.

We illustrate this as follows.

Return to eq.(17) on page 23:

$$\mathcal{U}(h, \tilde{u}) = \{u(t) : |u(t) - \tilde{u}(t)| \leq h\psi(t)\}, \quad h \geq 0 \quad (22)$$

Note that:

$$|u(t) - \tilde{u}(t)| \leq h\psi(t) \quad (23)$$

is the same as:

$$\tilde{u}(t) - h\psi(t) \leq u(t) \leq \tilde{u}(t) + h\psi(t) \quad (24)$$

We will construct eq.(22) differently,

and then generalize for vectors.

¶ Define the set:

$$S = \{u(t) : -\psi(t) \leq u(t) \leq \psi(t)\} \quad (25)$$

Define:

$$hS = \{u(t) : -h\psi(t) \leq u(t) \leq h\psi(t)\} \quad (26)$$

(Multiply each element of S by h)

Define:

$$hS + \tilde{u}(t) = \{u(t) : \tilde{u}(t) - h\psi(t) \leq u(t) \leq \tilde{u}(t) + h\psi(t)\} \quad (27)$$

(Add $\tilde{u}(t)$ to each element of S .)

Comparing eqs.(22) and (27) we see:

$$\mathcal{U}(h, \tilde{u}) = \{u(t) : |u(t) - \tilde{u}(t)| \leq h\psi(t)\}, \quad h \geq 0 \quad (28)$$

$$= hS + \tilde{u}(t) \quad (29)$$

¶ The set S has the following 2 basic properties:

1. S contains the zero element:

$$0 \in S \quad (30)$$

2. S is contained in its own expansions:

$$0 < h < h' \implies hS \subseteq h'S \quad (31)$$

¶ Now, for **any** set S obeying (30) and (31), define:

$$\mathcal{U}(h, \tilde{u}) = hS + \tilde{u}(t) \quad (32)$$

This is a **generalized envelope-bound info-gap model**.

Note: S need not be convex.

Note: S can be a set of vectors,

in which case $\mathcal{U}(h, \tilde{u})$ is a vector info-gap model.

6.3 Slope-bound info-gap models

(p.21)

¶ The elements of all info-gap models so far contain:
functions of **unbounded variation**.

Sometimes the prior information is:

- Nominal or typical $\tilde{u}(t)$.
- Unknown but **bounded slope**: constrained rate of variation.

$$\mathcal{U}(h, \tilde{u}) = \left\{ u(t) : \left| \frac{d}{dt}[u(t) - \tilde{u}(t)] \right| \leq h\psi(t) \right\}, \quad h \geq 0 \quad (33)$$

On the one hand, this is an envelope-bound info-gap model.
On the other hand, an important difference from other envelope-bound info-gap models is:
rate of variation of the functions is **constrained**.

6.4 Fourier-bound info-gap models

(pp.21–22)

¶ Another important type of partial information
on rate of variation : **bounded Fourier coefficients**.

¶ Truncated Fourier series:

$$u(t) = \sum_{n=n_1}^{n_2} [a_n \cos n\pi t + b_n \sin n\pi t] \quad (34)$$

Define:

$$\phi^T(t) = (\cos n_1\pi t, \dots, \cos n_2\pi t, \sin n_1\pi t, \dots, \sin n_2\pi t) \quad (35)$$

$$c^T = (a_{n_1}, \dots, a_{n_2}, b_{n_1}, \dots, b_{n_2}) \quad (36)$$

Thus:

$$u(t) = c^T \phi(t) \quad (37)$$

¶ Fourier ellipsoid-bound info-gap model:

Surface of an ellipsoid in \mathfrak{R}^N :

$$c^T W c = h \quad (38)$$

where W is a known, real, symmetric, positive definite matrix.

Solid ellipsoid in \mathfrak{R}^N :

$$c^T W c \leq h \quad (39)$$

Ellipsoid-bound info-gap model:

$$\mathcal{U}(h, \tilde{u}) = \left\{ u(t) = \tilde{u}(t) + c^T \phi(t) : c^T W c \leq h^2 \right\}, \quad h \geq 0 \quad (40)$$

7 Why Convex Info-gap Models?

7.1 Most Info-Gap Models are Convex: Examples

§ Convexity: definition.

A set S is **convex** if:

- For any elements $u, v \in S$, the line segment joining u and v is also entirely in S .

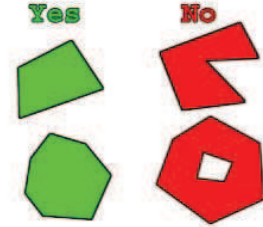
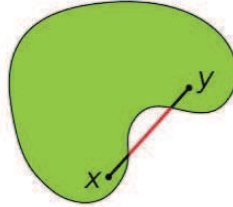
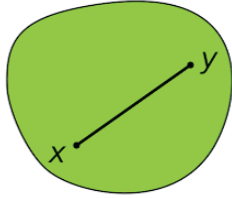


Figure 1: Convex set.

Figure 2: Non-convex set.

Figure 3: Convex and non-convex sets.

- Algebraically, S is convex if:

$$\gamma u + (1 - \gamma)v \in S \quad \text{for all } u, v \in S, \quad 0 \leq \gamma \leq 1 \quad (41)$$

§ It is no coincidence that the sets in all the info-gap models we have encountered, and many we will see, are convex. Explanation of relation between uncertainty and convexity: section 7.2, p.28.

§ Example of convexity: Fractional error info-gap model for scalars.

$$\mathcal{U}(h) = \left\{ u \in \mathfrak{R} : \left| \frac{u - \tilde{u}}{s} \right| \leq h \right\}, \quad h \geq 0 \quad (42)$$

- For any $u, v \in \mathcal{U}(h)$, define $z = \alpha u + (1 - \alpha)v$ for some $\alpha \in [0, 1]$.
- From membership in $\mathcal{U}(h)$:

$$\tilde{u} - sh \leq u \leq \tilde{u} + sh \quad (43)$$

$$\tilde{u} - sh \leq v \leq \tilde{u} + sh \quad (44)$$

$$\implies \tilde{u} - sh \leq \alpha u + (1 - \alpha)v \leq \tilde{u} + sh \quad (45)$$

$$\implies z \in \mathcal{U}(h) \quad (46)$$

§ Example of convexity: Fractional error info-gap model for vectors.

$$\mathcal{U}(h) = \left\{ u \in \mathfrak{R}^N : \left| \frac{u_i - \tilde{u}_i}{s_i} \right| \leq h, \quad i = 1, \dots, N \right\}, \quad h \geq 0 \quad (47)$$

- Demonstrate convexity by same argument as eqs.(43)–(46).

§ Example of convexity: Energy-bound info-gap model.

$$\mathcal{U}(h) = \left\{ u(t) : \int_{-\infty}^{\infty} [u(t) - \tilde{u}(t)]^2 dt \leq h^2 \right\}, \quad h \geq 0 \quad (48)$$

• For any $u(t), v(t) \in \mathcal{U}(h)$, define $z(t) = \alpha u(t) + (1 - \alpha)v(t)$ for any $\alpha \in [0, 1]$. We must show that $z(t) \in \mathcal{U}(h)$.

• We see that:

$$z(t) - \tilde{u}(t) = \alpha \underbrace{[u(t) - \tilde{u}(t)]}_{x(t)} + (1 - \alpha) \underbrace{[v(t) - \tilde{u}(t)]}_{y(t)} \quad (49)$$

which defines $x(t)$ and $y(t)$. Thus:

$$\int_{-\infty}^{\infty} [z(t) - \tilde{u}(t)]^2 dt = \alpha^2 \int x^2 dt + 2\alpha(1 - \alpha) \int xy dt + (1 - \alpha)^2 \int y^2 dt \quad (50)$$

• By the Cauchy-Schwarz inequality:

$$\int xy dt \leq \sqrt{\int x^2 dt} \sqrt{\int y^2 dt} \quad (51)$$

• Because $u(t), v(t) \in \mathcal{U}(h)$ we see that:

$$\int x^2 dt \leq h^2 \quad \text{and} \quad \int y^2 dt \leq h^2 \quad (52)$$

• Thus eq.(50) becomes:

$$\int_{-\infty}^{\infty} [z(t) - \tilde{u}(t)]^2 dt \leq \alpha^2 h^2 + 2\alpha(1 - \alpha)h^2 + (1 - \alpha)^2 h^2 = [\alpha + (1 - \alpha)]^2 h^2 = h^2 \quad (53)$$

• Thus $z(t) \in \mathcal{U}(h)$ which proves that the set is convex.

§ **Example of convexity: info-gap model of a pdf.**

$$\mathcal{U}(h) = \left\{ p(x) : p(x) \geq 0, \int_{-\infty}^{\infty} p(x) dx = 1, |p(x) - \tilde{p}(x)| \leq \tilde{p}(x)h \right\}, \quad h \geq 0 \quad (54)$$

• For any $p(x), q(x) \in \mathcal{U}(h)$, define $z(x) = \alpha p(x) + (1 - \alpha)q(x)$ for some $\alpha \in [0, 1]$.

• Clearly $z(x) \geq 0$. Also:

$$\int_{-\infty}^{\infty} z(x) dx = \alpha \int_{-\infty}^{\infty} p(x) dx + (1 - \alpha) \int_{-\infty}^{\infty} q(x) dx = 1 \quad (55)$$

• Finally:

$$(1 - h)\tilde{p}(x) \leq p(x) \leq (1 + h)\tilde{p}(x) \quad (56)$$

$$(1 - h)\tilde{p}(x) \leq q(x) \leq (1 + h)\tilde{p}(x) \quad (57)$$

$$\implies (1 - h)\tilde{p}(x) \leq \alpha p(x) + (1 - \alpha)q(x) \leq (1 + h)\tilde{p}(x) \quad (58)$$

• Hence $z(x) \in \mathcal{U}(h)$.

7.2 Why Are Most Info-Gap Models Convex?

¶ One explanation/motivation for convex info-gap models:
Superpositions of many microscopic events tend to cluster convexly.

¶ We can understand this by analogy to the **central limit theorem**:

Given i.i.d. random variables x_1, x_2, \dots with mean μ and finite variance σ^2 .

The distribution of x_1, x_2, \dots is arbitrary.

The x_i may be either discrete or continuous.

Define the **n -fold linear superposition**:

$$S_n = \frac{1}{\sigma\sqrt{n}} \sum_{i=1}^n (x_i - \mu) \quad (59)$$

The distribution of S_n converges asymptotically to a **standard normal distribution**, regardless of the distribution of x_1, x_2, \dots :

$$\lim_{n \rightarrow \infty} S_n \stackrel{\text{prob}}{=} \mathcal{N}(0, 1) \quad (60)$$

Interpretation:

Linear superposition of many arbitrarily distributed random variables converges to a specific macroscopic distribution: normal.

¶ Now consider the following **info-gap-analog** of the central limit theorem:

Let \mathcal{X} be an arbitrary set (not necessarily convex).

Define the **n -fold linear superposition**:

$$\xi_n = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{for } x_i \in \mathcal{X} \quad (61)$$

Consider the **set** of all such n -fold superpositions:

$$S_n = \left\{ \xi : \xi = \frac{1}{n} \sum_{i=1}^n x_i \quad \text{for all } x_i \in \mathcal{X} \right\} \quad (62)$$

S_n is not convex unless \mathcal{X} is convex.

However, the sequence of sets S_1, S_2, \dots

converges to a convex set:

$$\lim_{n \rightarrow \infty} S_n = \text{ch}(\mathcal{X}) \quad (63)$$

Interpretation:

Superposition of many arbitrary microscopic events of arbitrary dispersion tends to a convex cluster.