Decisions, Decisions, Decisions . . . .

Info-Gap Theory

and

Severe Uncertainty

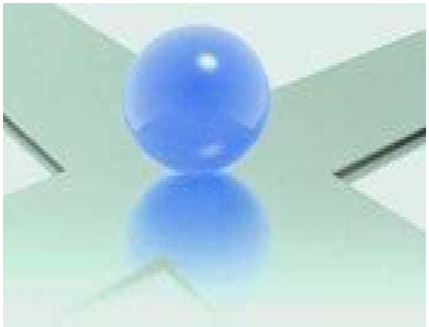
Yakov Ben-Haim

Technion

Israel Institute of Technology

yakov@technion.ac.il,   http://info-gap.com

http://www.technion.ac.il/yakov



Contents

1 *Highlights*

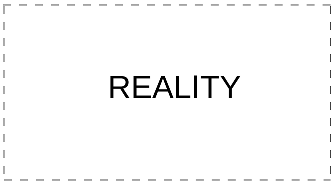
In the beginning,

God created the heavens and the earth.

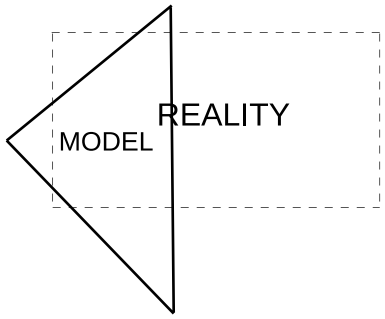
And the earth was total confusion . . .

. . . so humans started making models . . .

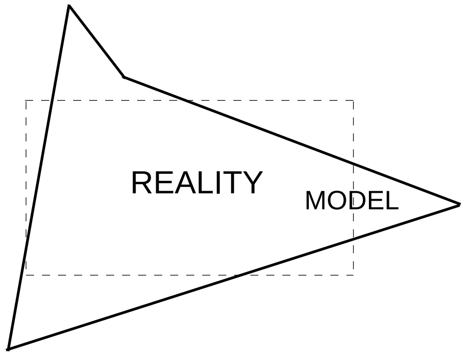
Out there is a . . .



We build models which, well, . . .



. . . but over time . . .



§ Scientific optimism, philosophical positivism:

$$\lim_{t \to \infty} \textbf{MODEL} = \textbf{REALITY}$$

§ Scientific optimism, philosophical positivism:

$$\lim_{t \to \infty} \textbf{MODEL} = \textbf{REALITY}$$

§ Motivations:

- **Truth.**
- **Utility.**

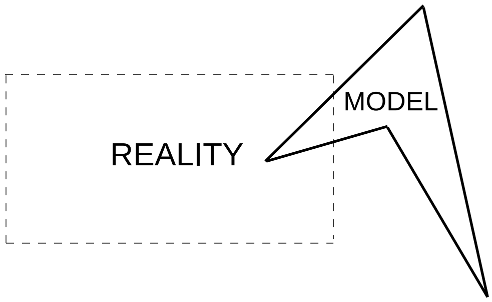
But...

§ **The art of designing, deciding, planning:**

Use the wrong model

to make the right decision

(when the right model is unknown).



§ **Evaluate and select a design under severe uncertainty.**

§ **Info-gaps:**

- **Incomplete understanding.**
- **Erroneous data.**
- **Changing conditions.**
- **Sur p rises.**

§ **Info-gap decision strategies:**

- **Robust-satisficing:**

  protect against uncertainty.

- **Opportune-windfalling:**

  exploit uncertainty.

§ **Issues:**

- **Modelling uncertainty.**
- **Robustness and probability of survival.**
- **Robust-satisficing and min-max.**
- **Preference reversal under competition.**
- **Opportuneness: Other side of uncertainty.**
- **Applications of info-gap theory.**

2 *Principle of Indifference*

§ Question:

Is ignorance probabilistic?

[0] \lectures\talks\lib\indif5.tex 26.11.2008

§ **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.

2.1 *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 value upon switching is:

  E.V. $= \frac{1}{2}$ **$** 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 value upon switching is:

  E.V. $= \frac{1}{2}\ \$\,2X + \ \frac{1}{2}\ \$\,\frac{1}{2}X = \$\,\left(1+\frac{1}{4}\right)X \ > \ X$.
- **Yes! Let's switch, you say.**

§ **You 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 value upon switching is:

  E.V. $= \frac{1}{2}\ \$\, 2X + \ \frac{1}{2}\ \$\, \frac{1}{2}X = \$\left(1 + \frac{1}{4}\right)X \ > \ X$.
- **Yes! Let's switch, you say.**

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

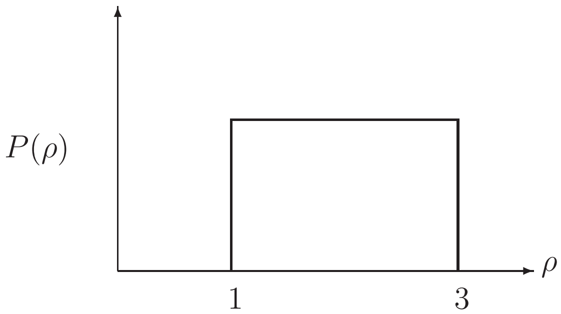
2.2 *Keynes' Example*

§ ρ = **specific gravity [g/cm^3] is unknown:**

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

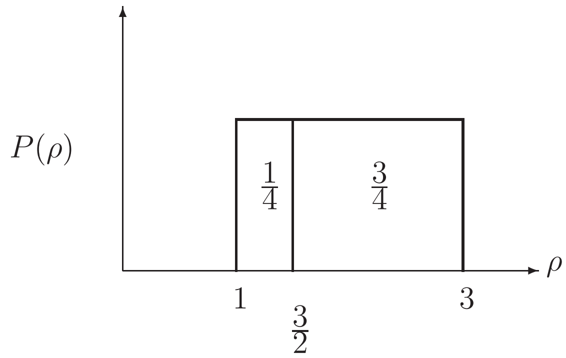
§ **Principle of indifference:**

Uniform distribution in $[1, \ 3]$**, so:**



§ Uniform distribution in $[1,\ 3]$, so:

$$\textbf{Prob}\left(\frac{3}{2} \ \leq \ \rho \ \leq \ 3\right) = \frac{3}{4}$$

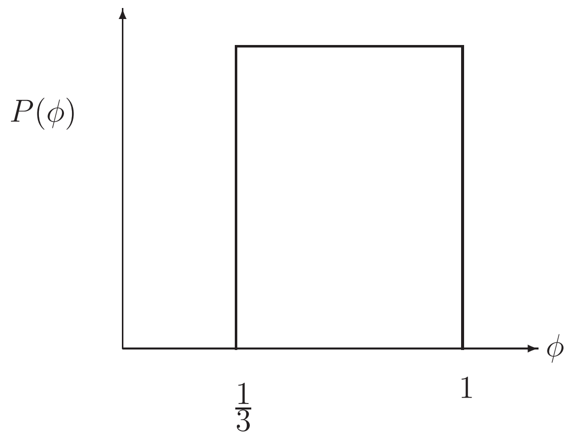


§ ϕ = specific volume [cm^3/g] is unknown:

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

§ Principle of indifference:

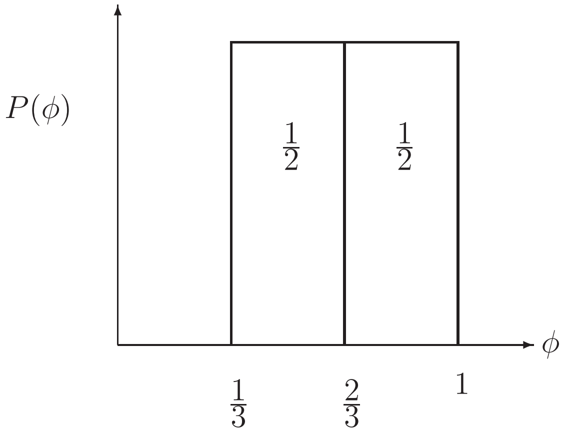
Uniform distribution in $\left[\frac{1}{3},\ 1\right]$, so:



§ Principle of indifference:

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

$$\textbf{Prob}\left(\frac{1}{3} \le \phi \le \frac{2}{3}\right) = \frac{1}{2}$$



§ Contradiction:

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

§ Contradiction:

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

§ The Culprit:

- **Principle of indifference.**

§ The resolution:

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

2.3 *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 *DESIGN ANALYSIS with INFO-GAPS*

§ The problem: Surprises!

§ The solution: Caution.

§ Our method: Info-gap robust-satisficing.

[0] lectures\talks\lib\design02.tex 25.7.2007

3.1 ROBUSTNESS to UNCERTAINTY

§ System model:

$$Y = GD + Z$$

Y= **Variable to be controlled.**

D= **Design variable to be chosen.**

G, Z= **Highly uncertain model variables.**

§ More generally:

System	Uncertainty
$Y = G^T D$	**Vector** G
$Y = D^T G D$	**Matrix** G
$Y = \int G(x) D(x)\, \mathrm{d}x$	**Function** G **e.g. pdf**
$Y = \min_{f \in G} \int f(x) D(x)\, \mathrm{d}x$	**Set** G **e.g. min exp util**

§ **System model:**

$$Y = GD + Z$$

Y= **Variable to be controlled.**

D= **Design variable to be chosen.**

G, Z= **Highly uncertain model variables.**

§ **Info-gap uncertainty** in G, Z:

- **Known** typical values, $\widetilde{G}$, $\widetilde{Z}$.
- **Unknown** range of error: surprises.
- Unknown probability distribution.

§ **Info-gap uncertainty in G, Z:**

- **Known typical values, $\widetilde{G}$, $\widetilde{Z}$.**
- **Unknown range of error: surprises.**
- **Unknown probability distribution.**
- **Unknown fractional error info-gap model:**

$$\mathcal{U}(\alpha, \widetilde{G}, \widetilde{Z}) = \left\{ G, Z : \ \left| \frac{G - \widetilde{G}}{\widetilde{G}} \right| \le \alpha, \ \left| \frac{Z - \widetilde{Z}}{\widetilde{Z}} \right| \le \alpha \right\}, \quad \alpha \ge 0$$

α = Unknown horizon of uncertainty.

§ **Another info-gap model:**

Estimated PDF with uncertain tails.

§ Performance function:

$$E^2(D,G,Z) = \left[Y(D,G,Z) - Y^{\bullet}\right]^2$$

$Y^{\bullet}$ **= Target value.**

§ Satisfice performance:

$$E(D,G,Z) \le E_{\rm c}$$

§ Robustness questions:

- **How wrong can the model $(\widetilde{G}, \widetilde{Z})$ be w/o jeopardizing performance $(E \leq E_{\rm c})$?**

- **What performance can be reliably anticipated?**

- **Design implication?**

§ Robustness:

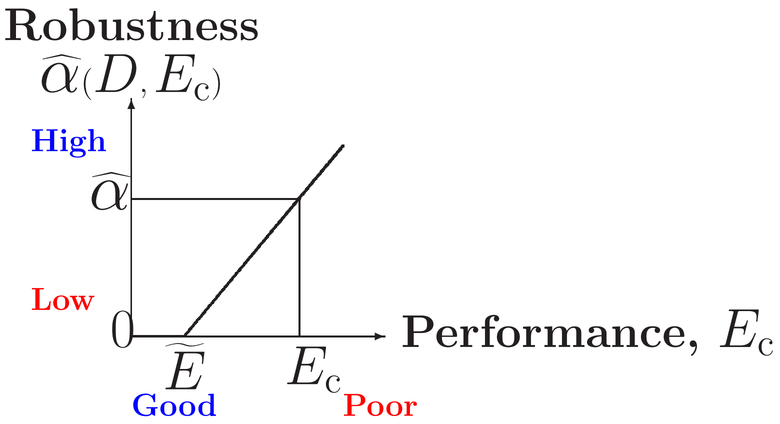
- $\mathcal{U}(\alpha, \widetilde{G}, \widetilde{Z})$ = **info-gap uncertainty.**
- α = **unknown horizon of uncertainty.**
- **Robustness = max. tolerable α:**

$$\widehat{\alpha}(D, E_{\rm c}) = \max\left\{\alpha : \ \left(\max_{G,Z\in\mathcal{U}(\alpha,\widetilde{G},\widetilde{Z})} E(D,G,Z)\right) \le E_{\rm c}\right\}$$

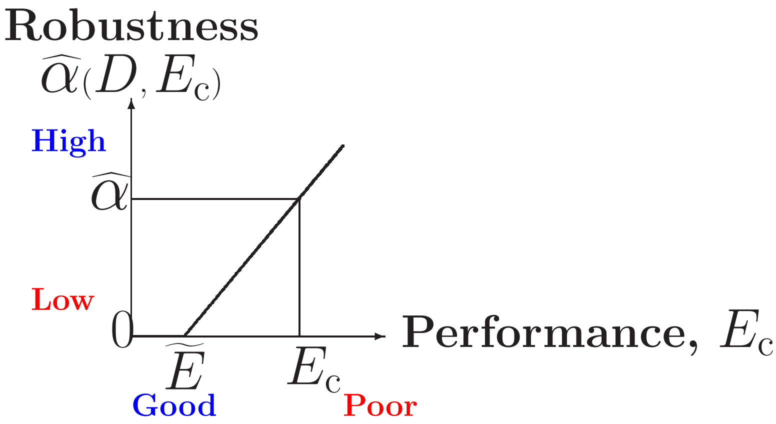
§ Preferences:

$$D \succ D^{\bullet} \quad \textbf{if} \quad \widehat{\alpha}(D, E_{\rm c}) > \widehat{\alpha}(D^{\bullet}, E_{\rm c})$$

- **Satisfice performance at $E_{\rm c}$.**
- **Maximize robustness $\widehat{\alpha}(D, E_{\rm c})$.**



§ **Trade-off: robustness vs. performance.**



§ **Trade-off: robustness vs. performance.**

§ **Anticipated outcome:** $\widetilde{E} = E(D, \widetilde{G}, \widetilde{Z})$

Zero robustness.

§ **Sub-optimal outcome:** $E_c > E(D, \widetilde{G}, \widetilde{Z})$

Positive robustness.

§ Two design choices, D, $D^{\bullet}$:

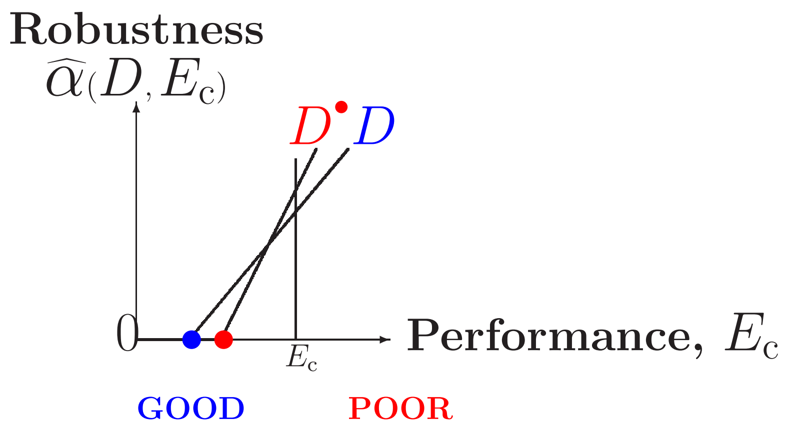
Best-model preference:

$$E(D,\widetilde{G},\widetilde{Z}) < E(D^{\bullet},\widetilde{G},\widetilde{Z}) \quad \Longrightarrow \quad D \succ D^{\bullet}$$

§ Two design choices, D, $D^\bullet$:

Best-model preference:

$$E(D, \widetilde{G}, \widetilde{Z}) < E(D^\bullet, \widetilde{G}, \widetilde{Z}) \quad \Longrightarrow \quad D \succ D^\bullet$$



- **Robust preference at E_c:**

$$\widehat{\alpha}(D^\bullet, E_c) > \widehat{\alpha}(D, E_c) \quad \Longrightarrow \quad D \prec D^\bullet$$

- **Calculate optimum (D)**

  but (possibly) do something else ($D^\bullet$).

3.2 OPPORTUNENESS from UNCERTAINTY

§ Two faces of uncertainty:

- **Pernicious: threaten failure.**
- **Propitious: enable windfall.**

§ Two faces of uncertainty:

- Pernicious: threaten failure.
- Propitious: enable windfall.

§ **Robust-satisficing:**

- **Satisfice: require** $E \le E_{\rm c}$.
- **Maximize robustness,** $\widehat{\alpha}(D, E_{\rm c})$.

§ **Opportune-windfalling:**

- **Windfall (v.): facilitate** $E \le E_{\rm w}$.
  $(E_{\rm w} < E_{\rm c})$.
- **Maximize opportuneness,** $\widehat{\beta}(D, E_{\rm w})$.

§ Pernicious uncertainty.

Robustness: maximum tolerable α.

$$\widehat{\alpha}(D, E_{\rm c}) = \max\left\{\alpha : \ \left(\max_{G,Z \in \mathcal{U}(\alpha,\widetilde{G},\widetilde{Z})} E(D,G,Z) \right) \leq E_{\rm c} \right\}$$

§ Propitious uncertainty.

Opportuneness: minimum required α.

$$\widehat{\beta}(D, E_{\rm w}) = \min\left\{\alpha : \ \left(\min_{G,Z \in \mathcal{U}(\alpha,\widetilde{G},\widetilde{Z})} E(D,G,Z) \right) \leq E_{\rm w} \right\}$$

§ Antagonistic immunities, $\widehat{\alpha}(D, E_{\rm c})$, $\widehat{\beta}(D, E_{\rm w})$:

- **Change D: degrade $\widehat{\alpha}$, improve $\widehat{\beta}$.**
- **Trade-off: sell rob., buy opportuneness.**

§ **Antagonistic immunities,** $\widehat{\alpha}(D, E_{\rm c})$**,** $\widehat{\beta}(D, E_{\rm w})$**:**

- **Change** D**: degrade** $\widehat{\alpha}$**, improve** $\widehat{\beta}$**.**
- **Trade-off: sell rob., buy opportuneness.**

§ **Sympathetic immunities,** $\widehat{\alpha}(D, E_{\rm c})$**,** $\widehat{\beta}(D, E_{\rm w})$**:**

- **Change** D**: improve** $\widehat{\alpha}$ **and** $\widehat{\beta}$**.**

3.3 ROBUSTNESS and PROBABILITY of DESIGN SUCCESS

§ Probability of design success

- **Model:** $Y = GD + Z$.
- **Use design** D **to reach target** $Y^{\bullet}$.
- **Error:** $E^2(G, Z) = (Y - Y^{\bullet})^2$.
- **Require:** $E(G, Z) \le E_{\rm c}$.
- (G, Z) **uncertain:**
  - $\mathcal{U}(\alpha, \widetilde{G}, \widetilde{Z})$ = **known info-gap model.**
  - $F(G, Z)$ = **unknown prob. distribution.**
- **Probability of design success:**

$$P_{\rm s}(D, E_{\rm c}) = F[E(G, Z) \le E_{\rm c}]$$

- **Probability of design success:**

$$P_s(D, E_c) = F[E(G, Z) \le E_c]$$

- $P_s(D, E_c)$ **unknown.** $\widehat{\alpha}(D, E_c)$ **known.**

§ **"Theorems":**

$$\left(\frac{\partial \widehat{\alpha}(D, E_c)}{\partial D}\right)\left(\frac{\partial P_s(D, E_c)}{\partial D}\right) \ge 0$$

- **Robustness proxies for prob. of success.**
- **We can't calculate** $P_s(D, E_c)$**.**
- **We can calculate** $\widehat{\alpha}(D, E_c)$**, so**
- **We can maximize** $P_s(D, E_c)$**.**

3.4 NEURAL NET FAULT DIAGNOSIS

§ Collaborators:

- **S. Gareth Pierce, Strathclyde.**
- **Keith Worden, Sheffield.**
- **Graeme Manson, Sheffield.**

§ Reference:

Evaluation of Neural Network Robust Reliability Using Information-Gap Theory,

IEEE Transactions on Neural Networks, **2006.**

§ Neural Net Diagnosis:

- Given training data.
- Choose parameters of multilayer perceptron.
- Diagnose faults in real data.

§ **Problem.** **Training data are:**

- **Noisy.**
- **Limited.**
- **System specific.**
- **Idiosyncratic.**

  (Future faults differ from past.)

Last three items are info-gaps.

§ Applications:

- **Simulated 2-D data.**
- **Real 9-D vib. of aircraft wing data.**
- **Real 2-D breast cancer incidence data.**

§ Conventional design of neural net:

- q = **parameters of MLP.**
- $\widetilde{u}$ = **training data.**
- **Performance: % success,** $R(q, \widetilde{u})$**.**
- $q^{\star}$ = **conventional optimization of MLP:**
  - **Maximum likelihood training.**
  - **Bayesian evidence training.**

§ Info-gap robust-satisficing design of NN:

- **Training data, $\widetilde{u}$.**
- **Info-gap model for uncertainty in $\widetilde{u}$:**

$$\mathcal{U}(\alpha, \widetilde{u}), \quad \alpha \geq 0$$

- **Performance requirement: $R(q, u) \geq R_{\rm c}$.**
- **Robustness of q:**

Max tolerable horizon of uncertainty, α:

$$\widehat{\alpha}(q, R_{\rm c}) = \max \left\{ \alpha : \left(\min_{u \in \mathcal{U}(\alpha, \widetilde{u})} R(q, u) \right) \geq R_{\rm c} \right\}$$

- **Robust-satisficing preference on designs:**

$$q \succ_{\rm r} q^{\bullet} \quad \textbf{if} \quad \widehat{\alpha}(q, R_{\rm c}) > \widehat{\alpha}(q^{\bullet}, R_{\rm c})$$

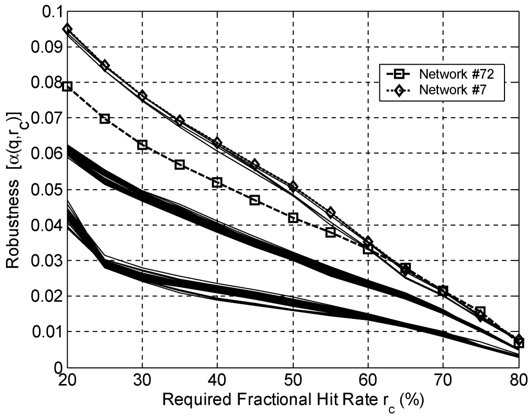


Fig. 5. Robustness functions for the 100 networks using maximum-likelihood training on the 1000-point Gaussian data.

§ Simulated data.

- **ML Optimum: network #72, □.**
- **IGRS Design: network #7, ◇.**
- **Robust-satisficing preferences:**
  - **#72 $\succ_{\rm r}$ #7 at high $R_{\rm c}$, low $\widehat{\alpha}$.**
  - **#7 $\succ_{\rm r}$ #72 at low $R_{\rm c}$, high $\widehat{\alpha}$.**

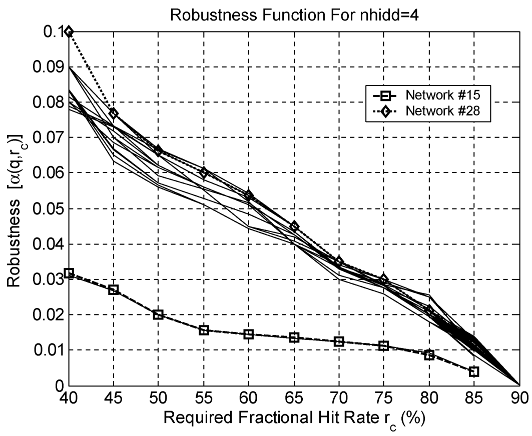


Fig. 7. Robustness functions for the 100 networks using Bayesian-evidence training on the 100-point Gaussian data.

§ Simulated data.

- **ML Optimum: network #15, □.**
- **IGRS Design: network #28, ◇.**
- **Robust-satisficing preferences:**
  - **#15 $\succ_{\rm r}$ #28 at very hi $R_{\rm c}$, very lo $\widehat{\alpha}$.**
  - **#28 $\succ_{\rm r}$ #15 else.**

§ Conclusions:

- **Robustness trades-off against performance.**
- **Opt. NN** $\succ_{\rm r}$ **IGRS at hi** $R_{\rm c}$**, low** $\widehat{\alpha}$**.**
- **IGRS** $\succ_{\rm r}$ **Opt. NN else.**
- **Similar results in real applications.**

4 *Historical Perspectives on Reasoning and Uncertainty*

Sources:

- ***IGDT*,**[1] **Ch. 2, sec. 1, pp.9–12.**

- **Hacking, Ian, *The Emergence of Probability: A Philosophical Study of Early Ideas About Probability, Induction and Statistical Inference*, 1975, Cambridge University Press.**

- **Stigler, Stephen M., 1986, *The History of Statistics: The Measurement of Uncertainty before 1900.* The Belknap Press of Harvard University Press.**

---

[0] \lectures\talks\lib\history02.tex 5.10.2010

[1] Yakov Ben-Haim, 2006, *Info-Gap Decision Theory: Decisions Under Severe Uncertainty*, 2nd edition, Academic Press, London.

§ Four stages:

I. Deduction: ancient Greece . . .

II. Forward probability: 1600, Pascal, Fermat, . . .

III. Inverse probability: 1760, Bayes, . . .

IV. Modern design and inference.

I. Deduction:

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

II. Forward probability:

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

- **Forward probability:**

  Given known die, what is the probability of any specified sequence i, j, k, . . . ?

- **Probabilistic syllogism for sequence:**
  1. **Prob(i) $= P_i$.**
  2. **Given the sequence i, j, k, ...**
  3. **Hence Prob(i, j, k, . . .) is $P_i P_j P_k \ldots$.**

III. Inverse probability: Inference.

- **Given unknown die that produced a given sequence,**

  what are the prob's of i, j, k, ...?

  (Bayes, 1760).

- **Least-squares estimation:**

  Given noisy meas's of planetary motion,

  what is best estimate of elliptical orbit?

  (Gauss, Legendre, 1800)

(cont.)

- **Central limit theorem (Laplace 1815):**

Given measured mean and variance, we know the pdf.

- **Many distributions and statistical tests.**

(Karl Pearson and others, 19th, 20th c.)

- **Probabilistic syllogism for die sequence:**

  1. **If prob of i is P_i,**

     Then prob of observation is $P(ijk|P_i, P_j, P_k)$.

  2. **Observed sequence ijk.**

  3. **Hence estimates of P_i, P_j, P_k are . . . .**

IV. Modern design, decision and inference.

- **Verbal uncertainty (Lukaceiwizc, 1920; Zadeh, 1960).**

"If it rattles, tighten it."

What is a rattle? How tight is tight?

(cont.)

- **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, quantum, nano, etc.
  - **Geo-political and economic crises.**

(cont.)

- **Uncertainties:**
  - **Lack of data and understanding.**
  - **Mutually inconsistent data.**
  - **Change over time.**
  - **Disparity between**

    what is known

    and what must be known

    for a good decision.
- **This is info-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.

5 *Info-gaps, Probability and Fuzziness*

§ Three models/conceptions of uncertainty:

- **Probability:**
  - **Frequency of recurrence.**

    $P(X)$ = **frequency with which** X **recurs.**
  - **Degree of belief.**

    $P(X)$ = **degree of belief** X **will occur.**
  - **In both cases:**

    $P(X) + P(\neg X) = 1$. **(‘**$\neg X$**’ $\equiv$ ‘not** X**’)**

(cont.)

- **Fuzzy logic:**[2]
  - **Linguistic ambiguity.**
  - **Possibility (less strict than prob).**

    $M(X) =$ **possibility that** X

    e.g., rain tomorrow may occur.

    $M(X) + M(\neg X) > 1$.

    'rain' and 'not rain' highly possible.
  - **Necessity (more strict than prob).**

    $M(X) =$ **neccesity that** X

    e.g., rain tomorrow will occur.

    $M(X) + M(\neg X) < 1$.

    Maybe dew; neither 'rain' nor 'not rain'.

(cont.)

[2]Dubois, Didier and Henri Prade, 1986, *Possibility Theory: An Approach to Computerized Processing of Uncertainty.* With the collaboration of H. Farreny, R. Martin-Clouaire and C. Testemale. Translated by E.F. Harding, Plenum Press, New York.

- **Info-gap:**
  - **Clustering of events.**
  - **Unbounded uncertainty:**
    No known worst case.
  - **No measure functions of uncertainty.**
  - **Disparity between the**
    known and the learnable.

We will briefly consider an example.

§ Example: Toxic waste.

- **Toxic waste released into river.**
- $\tilde{r}(x, t, w)$ = **model of removal rate.**
  - x = **position,** t = **time,** w = **parameters.**
- $r(x, t, w)$ = **true removal rate.**
- $\tilde{r}(x, t, w) - r(x, t, w)$**: unknown.**
- **Info-gap model for uncertain** $r(x, t, w)$**.**
- **Family of nested sets of** r**-functions.**

§ **An info-gap model for uncertain** $r(x,t,w)$**:**

$$\mathcal{U}(h,\tilde{r}) = \{r(x,t,w): \ |\tilde{r}(x,t,w) - r(x,t,w)| \le h\}, \quad h \ge 0$$

h **is the uncertainty parameter.**

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

§ **Two levels of uncertainty:**

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

§ **Axioms of info-gap models:**

- **Nesting:**

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

- **Contraction:** $\mathcal{U}(0,\tilde{r}) = \{\tilde{r}\}$

6 *Why are Info-Gap Models Convex?*

§ **One explanation for convexity:**[3]

Superpositions of many microscopic events tend to cluster convexly.

§ **Convexity:**

A set S is convex if:

For any elements $u, v \in S$

The line joining u and v is entirely in S.

Algebraically, S is convex if:

$$\gamma u + (1-\gamma)v \in S \quad \textbf{for all} \quad u, v \in S, \quad 0 \le \gamma \le 1$$

Most info-gap models are convex. Why?

[3] IGDT, pp.25–26.

§ The Central Limit Theorem:

Given i.i.d. random variables $x_1, x_2, \ldots$ **.**

The distribution of $x_1, x_2, \ldots$ **is arbitrary.**

The x_i **may be either discrete or continuous.**

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

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i$$

S_n **converges asymptotically to a**

normal variate for any $x_1, x_2, \ldots$ **.**

§ Interpretation:

Superposition of many micro iid variables converges to specific macro distribution: normal.

§ Info-gap analog of the CLT:

Let $\mathcal{X}$ be a set (maybe not convex).

Define the n-fold linear superposition:

$$\xi_n = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \textbf{for} \quad x_i \in \mathcal{X}$$

The set of all such n-fold superpositions:

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

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

However, the sequence of sets S_1, $S_2, \ldots$

converges to a convex set:

$$\lim_{n\to\infty} S_n = \mathbf{ch}(\mathcal{X})$$

§ Interpretation:

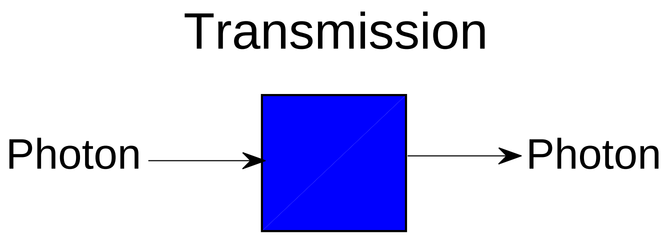
Superposition of many microscopic events of arbitrary dispersion tends to a convex cluster.

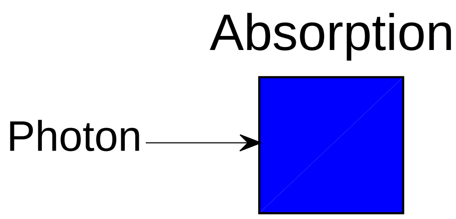
7 *Quantum Indeterminism*

[3] \lectures\talks\lib\quantum03shrt.tex 13.6.2010

§ Polarized Photons on Tourmaline

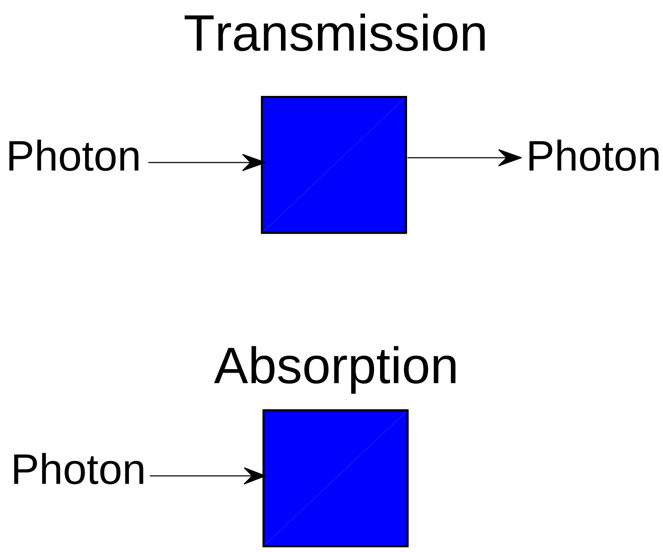
- **Identical photons; different outcomes.**





§ Polarized Photons on Tourmaline

- **Identical photons; different outcomes.**



- **What happened to Causality?**
- **Aren't there Laws of Nature?**

§ Classical Physics:

- **Natural law: Deterministic.**
- **Individual events: causal relations.**

§ Classical Physics:

- **Natural law: Deterministic.**
- **Individual events: causal relations.**

§ Standard Interp of Quantum Theory:

- **Natural law: Probabilistic.**
- **Individual events: indeterminate.**
- **Individual causality: lost.**

§ Classical Physics:

- **Natural law: Deterministic.**
- **Individual events: causal relations.**

§ Standard Interp of Quantum Theory:

- **Natural law: Probabilistic.**
- **Individual events: indeterminate.**
- **Individual causality: lost.**

§ Alternative Interp of Quantum Theory:

- **Natural law: Indeterminate.**
- **Individual events causally determined.**

§ Mother Nature's Problem.

- **Principle of Least Action:**

  Optimize action integral: $S = \int_0^t L(x, \tau)\, \mathrm{d}\tau$

- L, **Lagrangian, from Law of Nature.**
- **Suppose Law of Nature indeterminate?**

§ Mother Nature's Problem.

- **Principle of Least Action:**

  Optimize action integral: $S = \int_0^t L(x, \tau)\, \mathrm{d}\tau$
- L, **Lagrangian, from Law of Nature.**
- **Suppose Law of Nature indeterminate?**

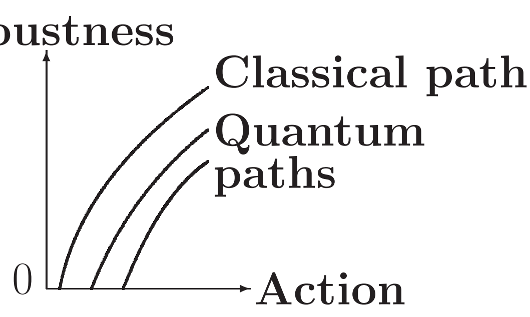
§ Mother Nature's Solution.

- **Satisfice action integral.**
- **Maximize robustness to uncertain** L**.**
- **Wave function becomes:**

$$\psi(x) = c \exp\left(\frac{i}{\hbar} S[\widehat{h}(x), \widetilde{L}]\right)$$

(Extension of Feynman's argument:

$\max \widehat{h} \Longrightarrow$ **Stationary** $S \Longrightarrow$ **Large** $|\psi|$**)**



§ **Trade-off:** **Robustness vs. action.**

§ **Zero robustness at classical action.**

§ **Quantum Paths:**

Low action = hi robustness = hi probability.

§ Nature:

- **Abhores an indeterminism.**
- **Robust-satisfices:**
  - **Satisfice the action.**
  - **Maximize robustness to indeterminacy.**

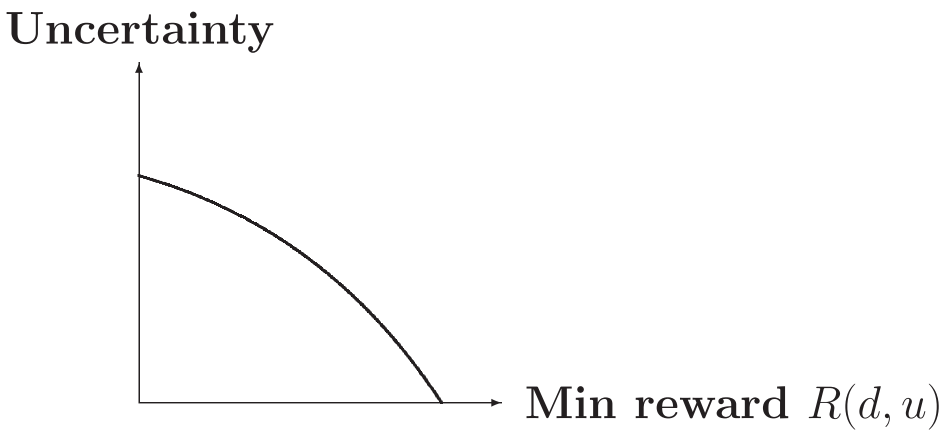
8 *Max-Min and Robust-Satisficing*

§ **Task: make a decision.**

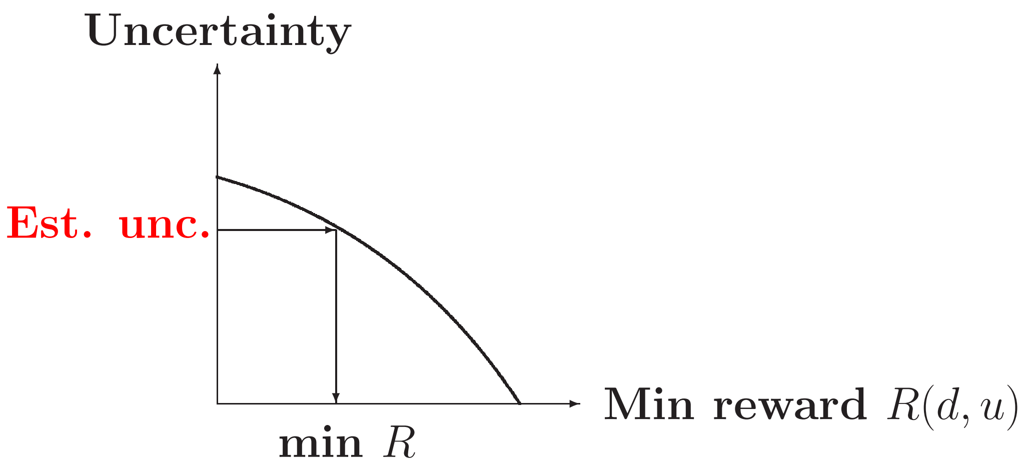
- d = **decision.**
- u = **uncertain parameters, functs., sets.**
- $R(d, u)$ = **reward.**

[3] lectures\talks\lib\maxmin-rs03shrt.tex 5.4.2012

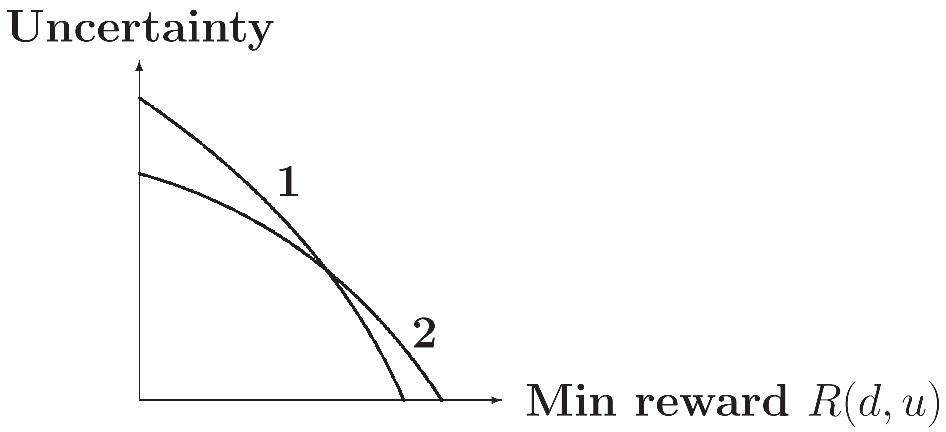
§ Trade-off: uncertainty vs. min reward.



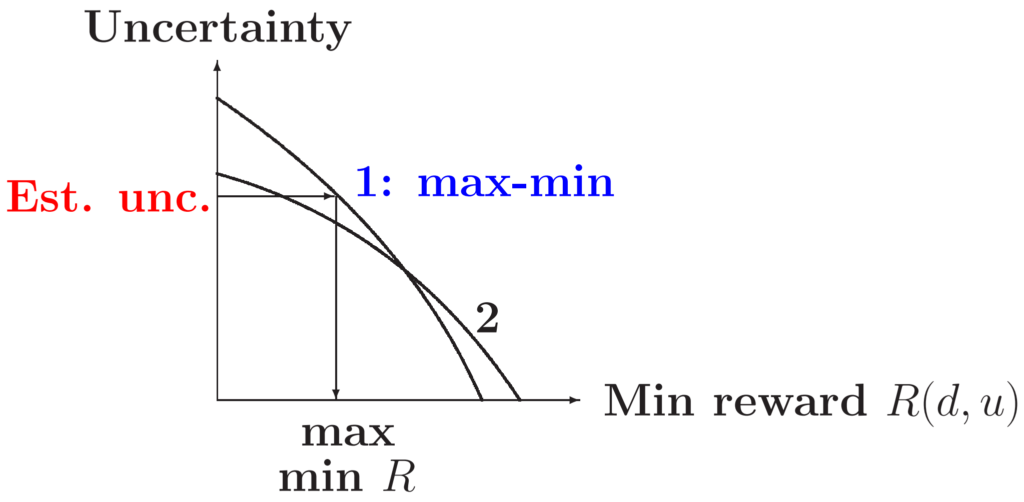
§ Trade-off: uncertainty vs. min reward.



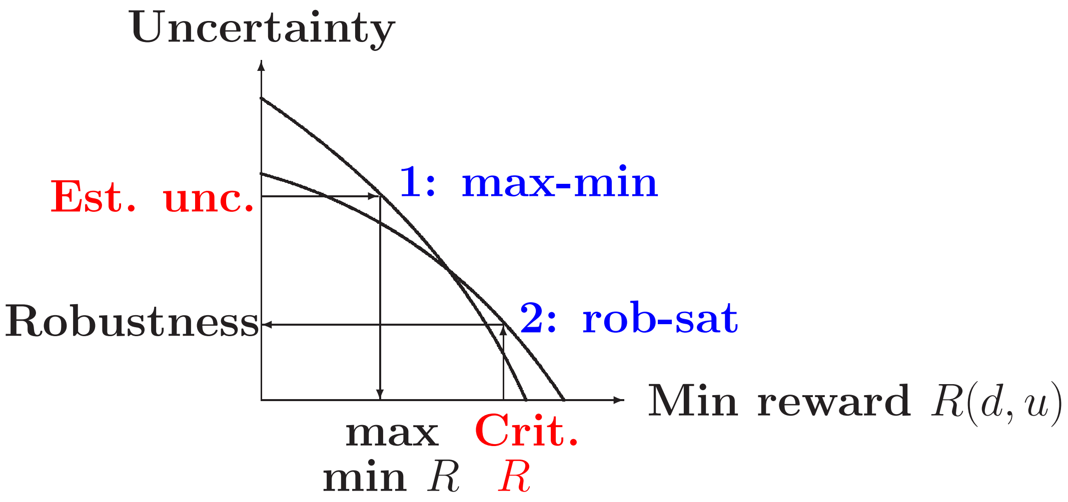
§ Choose from 2 decisions: d_1, d_2.



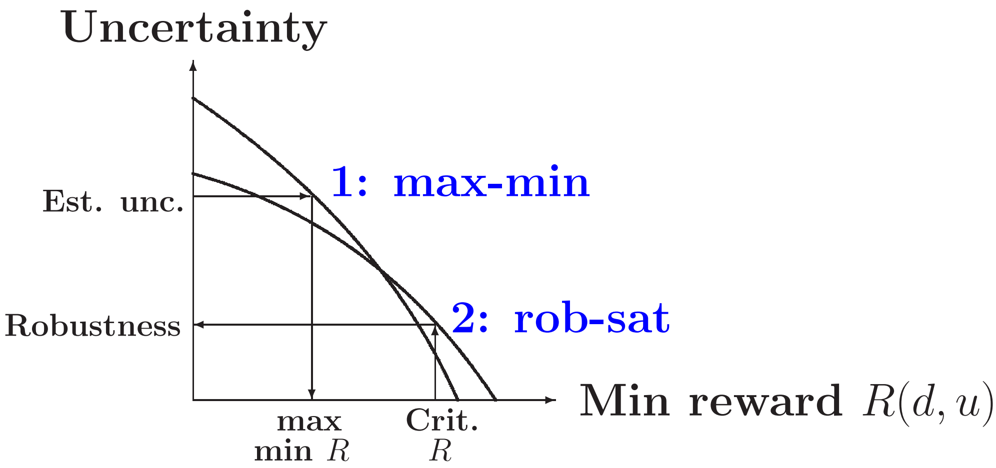
§ Choose from 2 decisions: d_1, d_2.



§ Choose from 2 decisions: d_1, d_2.



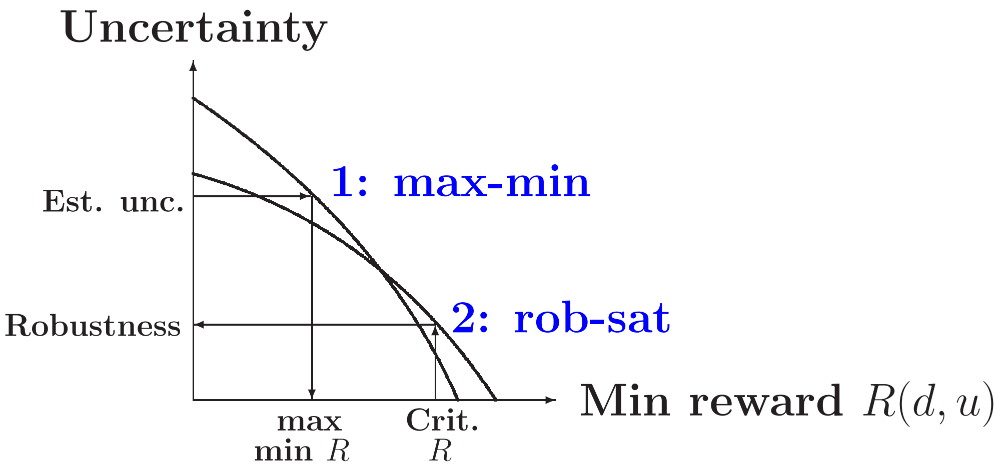
§ **Choose from 2 decisions:** d_1, d_2.



§ **Modeller's equivalence: description.**

- **Max-min can always describe rob-sat (by adjusting prior beliefs).**
-

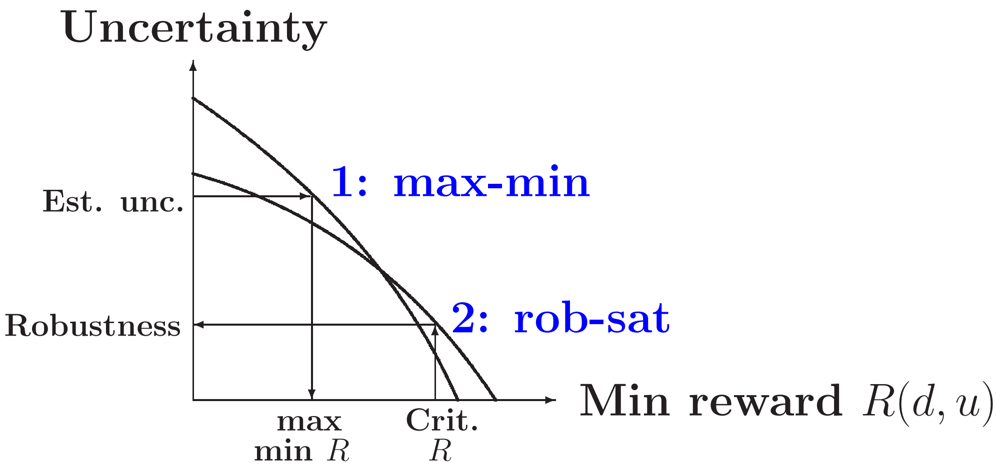
§ Choose from 2 decisions: d_1, d_2.



§ Modeller's equivalence: description.

- Max-min can always describe rob-sat (by adjusting prior beliefs).
- Rob-sat can always describe max-min (by adjusting requirements).

§ **Choose from 2 decisions:** d_1, d_2.



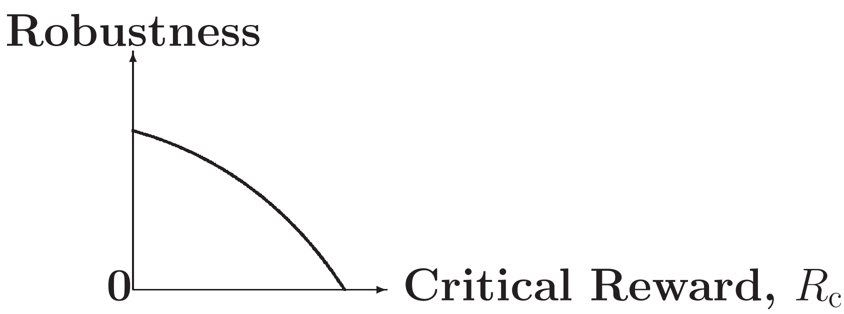
§ **Modeller's equivalence: description.**

§ **Decision-maker's duality: prescription.**

Max-min and rob-sat differ if:

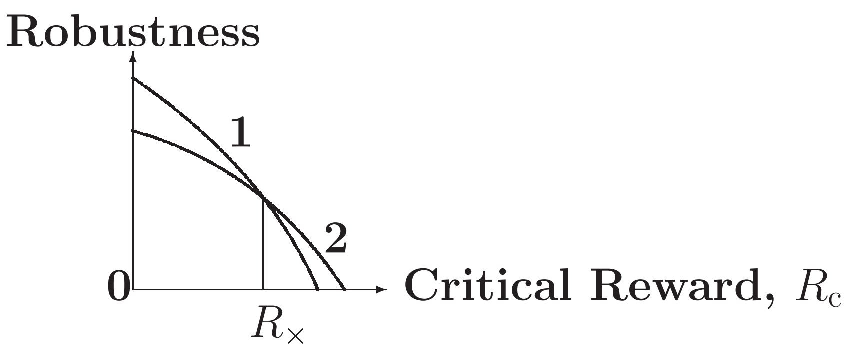
- **Max-min gain too low, or,**
- **Worst case is uncertain.**

§ Optimizing vs Robustifying



- **Trade off:** Robustness vs performance.
- **Zeroing:** No rbs of predicted reward.

§ Optimizing vs Robustifying



- **Trade off:** Robustness vs performance.
- **Zeroing:** No rbs of predicted reward.
- **Predicted optimum:** 2.
- **Robust-satisficing optimum:** 2 iff $R_c > R_\times$.

§ Info-gap robustness is non-probabilistic.

Is it a good bet?

§ **Info-gap robustness is non-probabilistic.**

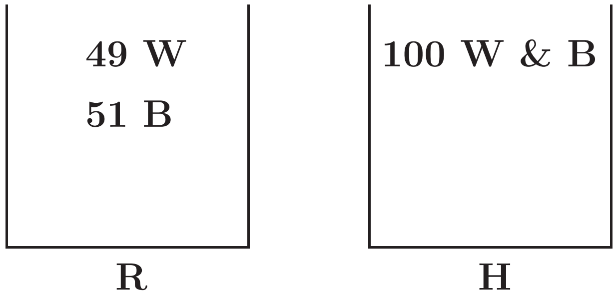
Is it a good bet?

§ **Evolutionary advantage of robustness:**

- **Robustness may proxy for Probability of survival.**
- **Proxy theorems.**

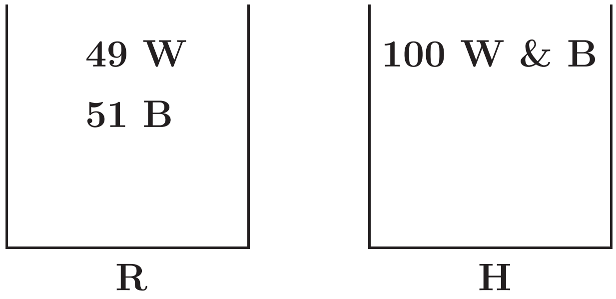
9 *ELLSBERG'S PARADOX*

§ Ellsberg's urn experiment:



Urns with 100 white and black balls.

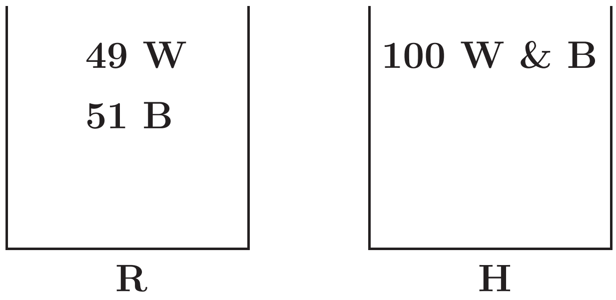
[3] lectures\talks\lib\ellsberg01.tex 18.3.2010



§ Choose urn and 1 ball. Which urn?

- **1st experiment: win $1000 if black.**

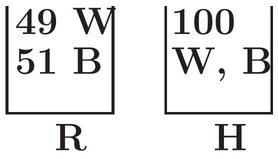
§ What would YOU do?



§ **Choose urn and 1 ball. Which urn?**

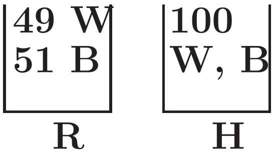
- **1st experiment: win $1000 if black.**
- **2nd experiment: win $1000 if white.**

§ **What would YOU do?**



§ Ellsberg's observations:

- **1st exp (win on B): most choose R urn.**
- **2nd exp (win on W): most choose R urn.**



§ Ellsberg's observations:

- **1st exp (win on B): most choose R urn.**
- **2nd exp (win on W): most choose R urn.**

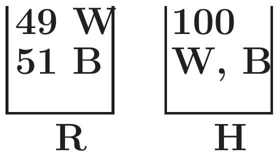
§ Ellsberg's paradox:

- **R urn: P(W)= 0.49. P(B) = 0.51.**
- **H urn: P(W)= π. P(B) = $1 - \pi$.**
- **1st experiment implies $\pi > 0.49$.**
- **Should choose H in 2nd exp. Most don't!?!**

§ Consider Ellsberg's 1st experiment.

§ Expected utility:

- U = **utility of $1000.**
- **0** = **utility of $0.**
- P = **probability of success.**
- UP = **expected utility.**



§ **Subjective uncertainty in P**

- **Low in R urn.**
- **High in H urn.**

§ **Aspiration:**

Acceptable expected utility:

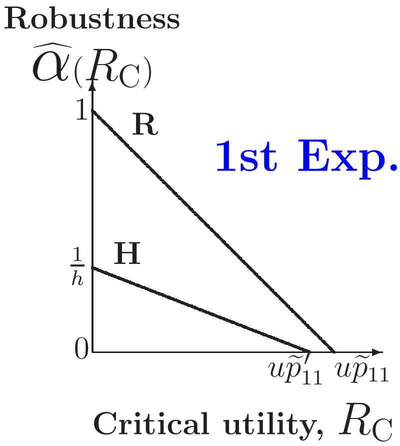
$$PU \geq R_{\rm c}$$

§ **Robustness:**

Maximum tolerable error in P.

§ **Aspiration:** **Acceptable EU:** $PU \geq R_c$.

§ **Robustness:** **Max tolerable error in** P.



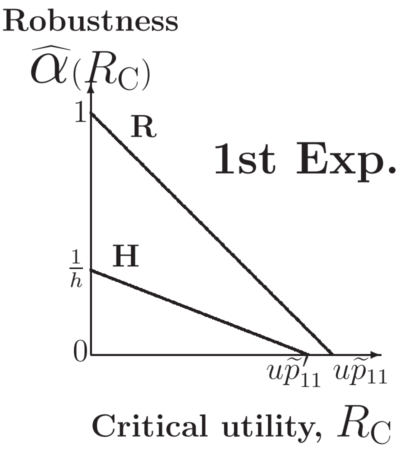
§ **Trade-off: robustness vs utility.**

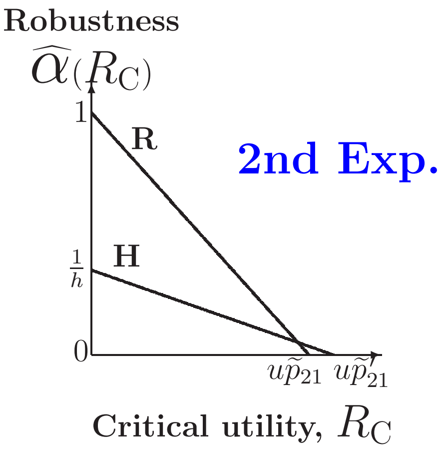
§ **No robustness of anticipated utility.**

§ **R more robust than H for all** R_C.

- **Robust-satisficer chooses R not H.**
- **Utility-optimizer chooses R not H.**
- **Agree on action, for different reasons.**

§ Now consider Ellsberg's 2nd experiment.





- **Utility-optimizer chooses H not R.**
- **Robust-satisficer chooses R not H.**
- **Disagree on action.**
- **Ellsberg's observation:**

  Most choose R both times.
- **Most are robust-satisficers in both exps.**

§ **We have resolved Ellsberg's paradox:**

Most agents robust-satisfice.

§ **Questions:**

- **Is this human mental frailty or the guile of *Homo sapiens*?**
- **Is satisficing a last resort, or strategically advantageous?**

10 *INFO-GAP FORECASTING*

Yakov Ben-Haim, 2009,
Info-gap forecasting,
European Journal of Operational Research.

Yakov Ben-Haim, 2010,
Info-Gap Economics:
An Operational Introduction,
Palgrave-Macmillan.

[3] lectures talks lib forecast02.tex 4.10.2010

10.1 *1-D Example*

§ "True" scalar system:

$$y_t = \lambda_t y_{t-1}$$

§ Historical data: $\lambda_t = \widetilde{\lambda}$ **for** $t \leq T$**.**

§ Contextual understanding:

λ **could drift upwards.**

§ Fractional-error info-gap model:

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_t, t > T : \; 0 \leq \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \leq h \right\}, \qquad h \geq 0$$

- **Unbounded family of sets.**
- **No worst case.**

§ **Slope-adjusted (erroneous) forecaster:**

$$y_t^{\mathrm{s}} = \ell y_{t-1}^{\mathrm{s}}$$

§ **Contrast with historically estimated model:**

$$y_t = \widetilde{\lambda}_t y_{t-1}$$

How to choose $\ell \ge \widetilde{\lambda}$?

§ **Robust satisficing:**

Satisfice the forecast error:

$$\left| y_{T+k}^{\mathrm{s}} - y_{T+k} \right| \le \varepsilon_{\mathrm{c}}$$

Maximize robustness to future surprise.

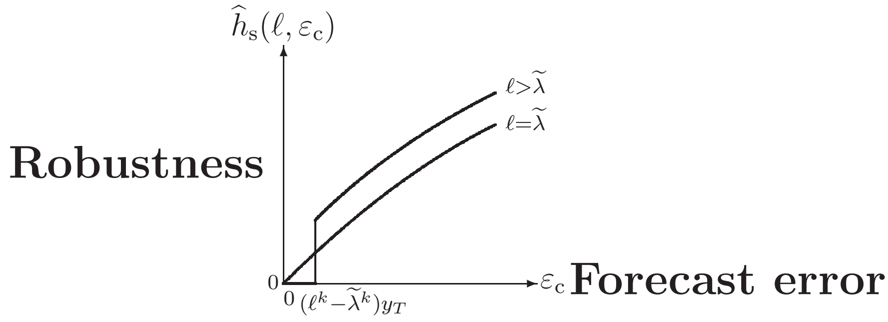
§ Robustness of forecast ℓ:

Max h up to which all λ_{T+i} in $\mathcal{U}(h, \widetilde{\lambda})$

satisfice forecast error at ε_{c}:

$$\widehat{h}_{\mathrm{s}}(\ell, \varepsilon_{\mathrm{c}}) = \max\left\{ h : \left(\max_{\substack{\lambda_{T+i} \in \mathcal{U}(h, \widetilde{\lambda}) \\ i=1, \ \ldots, \ k}} \left| y^{\mathrm{s}}_{T+k} - y_{T+k} \right| \right) \le \varepsilon_{\mathrm{c}} \right\}$$

§ Preference: $\ell \succ \ell'$ if $\widehat{h}_{\mathrm{s}}(\ell, \varepsilon_{\mathrm{c}}) > \widehat{h}_{\mathrm{s}}(\ell', \varepsilon_{\mathrm{c}})$



§ **Trade off:** robustness vs. forecast error.

§ **Zeroing:** Estim outcome has 0 robustness.

§ **Crossing robustness curves:** $\ell \succ \widetilde{\lambda}$.

- **Preference reversal.**
- Robustness-advantage of sub-optimal (erroneous) model.

§ **Robustness is proxy for success-probability.**

§ Numerical example.

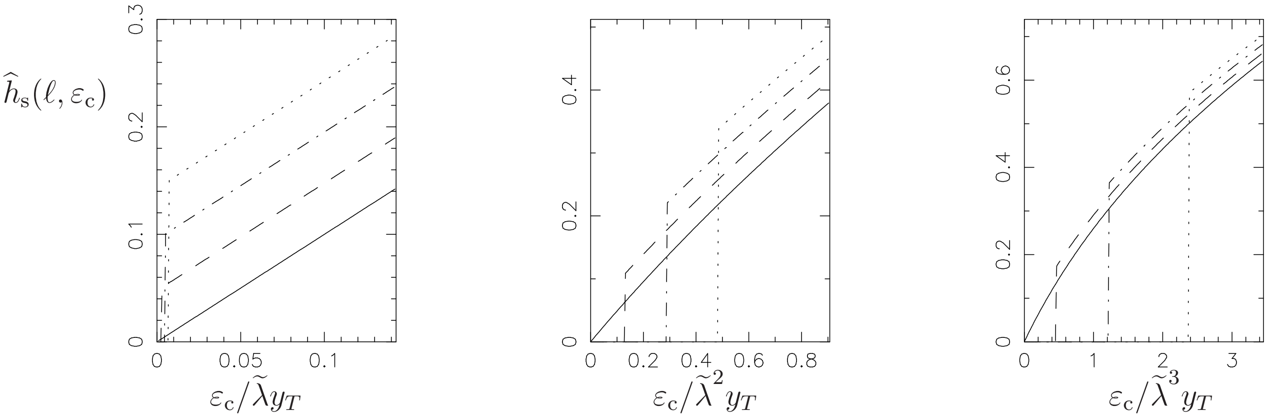


Figure 1: Robustness vs. normalized forecast error for $\ell = 1.05$, 1.1, 1.15, 1.2 from bottom to top curve. $\widetilde{\lambda} = 1.05$, $y_T = 1$. $k = 1$ (left), 2(mid), 3(right).

- **Preference reversal at all time horizons, k.**
- **Robustness premium decreases with k.**
- **Reversal-ε_c increases with k.**

Robustness & Probability of Forecast Success

§ **Future growth coefficients:** $\lambda_{T,k} = (\lambda_{T+1}, \ldots, \lambda_{T+k})$.

$\lambda_{T,k}$ **is random vector on domain** D.

$F(\lambda_{T,k})$ = **cumulative probability distrib.**

§ **Forecast success set:**

$$\mathcal{Y}(\ell) = \{\lambda_{T,k} \in D : \ |y^{\mathrm{s}}_{T+k}(\ell) - y_{T+k}| \le \varepsilon_{\mathrm{c}}\}$$

§ **Probability of success:**

$$P_{\mathrm{s}}(\ell) = F[\mathcal{Y}(\ell)]$$

§ **Goal:**

Choose ℓ to maximize success prob.

§ **Problem:**

$F(\lambda_{T,k}) =$ **is unknown.**

§ **Solution:**

- $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c})$ **is known.**
- $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c})$ **proxies for success prob.**

§ Theorem.

Probability of successful forecast, $P_{\rm s}(\ell)$, increases with increasing info-gap robustness, $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c})$.
Given: (a) The domain of $F(\cdot)$ is contained in the info-gap model of eq.(10.1). (b) $y_T > 0$, $\widetilde{\lambda} > 0$. (c) ℓ and ℓ' are two slope parameters for which:

$$\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c}) > \widehat{h}_{\rm s}(\ell', \varepsilon_{\rm c}) > 0$$

Then:

$$P_{\rm s}(\ell) \geq P_{\rm s}(\ell')$$

§ Robustness is proxy for success-probability.

Summary so far:

§ **Forecasters do better if they robust-satisfice.**

§ **Satisficing is not a last resort.**

It is strategically advantageous.

10.2 *European Central Bank: Overnight Rate*

Date	Interest rate	Implied λ
1 Jan 1999	4.50	
9 Apr 1999	3.50	0.778
5 Nov 1999	4.00	1.143
4 Feb 2000	4.25	1.063
17 Mar 2000	4.50	1.059
28 Apr 2000	4.75	1.056
9 Jun 2000	5.25	1.105
28 Jun 2000	5.25	1.000
1 Sep 2000	5.50	1.048
6 Oct 2000	5.75	1.045
11 May 2001	5.50	0.957
31 Aug 2001	5.25	0.955

§ Typical change: 25 basis points.

§ Largest change: 100 basis points.

Date	Interest rate	Implied λ
1 Jan 1999	4.50	
9 Apr 1999	3.50	0.778
5 Nov 1999	4.00	1.143
4 Feb 2000	4.25	1.063
17 Mar 2000	4.50	1.059
28 Apr 2000	4.75	1.056
9 Jun 2000	5.25	1.105
28 Jun 2000	5.25	1.000
1 Sep 2000	5.50	1.048
6 Oct 2000	5.75	1.045
11 May 2001	5.50	0.957
31 Aug 2001	5.25	0.955

§ **9Jun'00–31Aug'01: $\mu = 5.4\%$, $\sigma = 0.19\%$.**

§ **On 9/12/2001, (1 day after 9/11)**

predict next interest rate.

Rate down, but by how much?

§ Historical model:

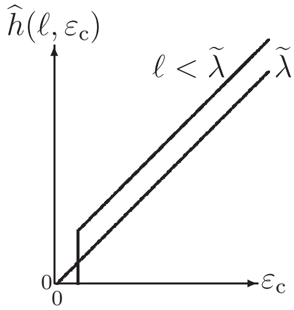
$$y_t = \lambda_t y_{t-1}, \qquad \lambda_t = \widetilde{\lambda} = 1$$

§ Info-gap model:

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_T : \ (1-h)\widetilde{\lambda} \le \lambda_T \le \widetilde{\lambda} \right\}, \qquad h \ge 0$$

§ Robustness:

$$\widehat{h}(\ell, \varepsilon_{\rm c}) = \begin{cases} 0 & \varepsilon_{\rm c} < \varepsilon_\times \\ \dfrac{\varepsilon_{\rm c}}{\widetilde{\lambda} y_T} + \dfrac{\widetilde{\lambda} - \ell}{\widetilde{\lambda}} & \varepsilon_\times \le \varepsilon_{\rm c} \end{cases}$$



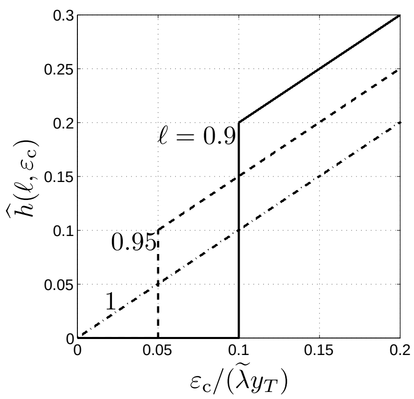


Figure 2: **Robustness vs normalized forecast error.** $\widetilde{\lambda} = 1$, $y_T = 5.25$.

§ $\ell = 1.0 \Longrightarrow$ **0% robustness at 0% error.**

§ $\ell = 0.95 \Longrightarrow$ **10% robustness at 5% error.**

§ $\ell = 0.9 \Longrightarrow$ **20% robustness at 10% error.**

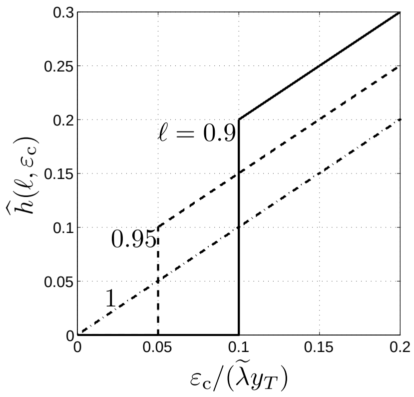


Figure 3: **Robustness vs normalized forecast error.** $\widetilde{\lambda} = 1$, $y_T = 5.25$.

§ $\ell = 0.9 \Longrightarrow$ **20% robustness at 10% error.**

§ **Forecast:** $y^{\rm s}_{T+1} = 0.9 y_T =$ **4.725.**

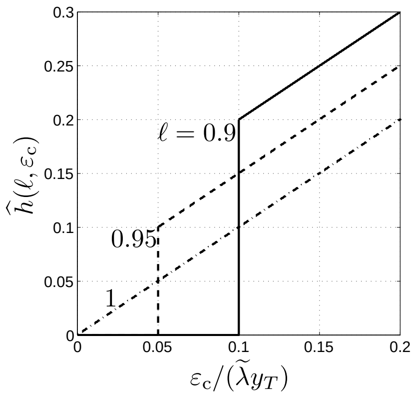


Figure 4: **Robustness vs normalized forecast error.** $\widetilde{\lambda} = 1$, $y_T = 5.25$.

§ $\ell = 0.9 \Longrightarrow$ **20% robustness at 10% error.**

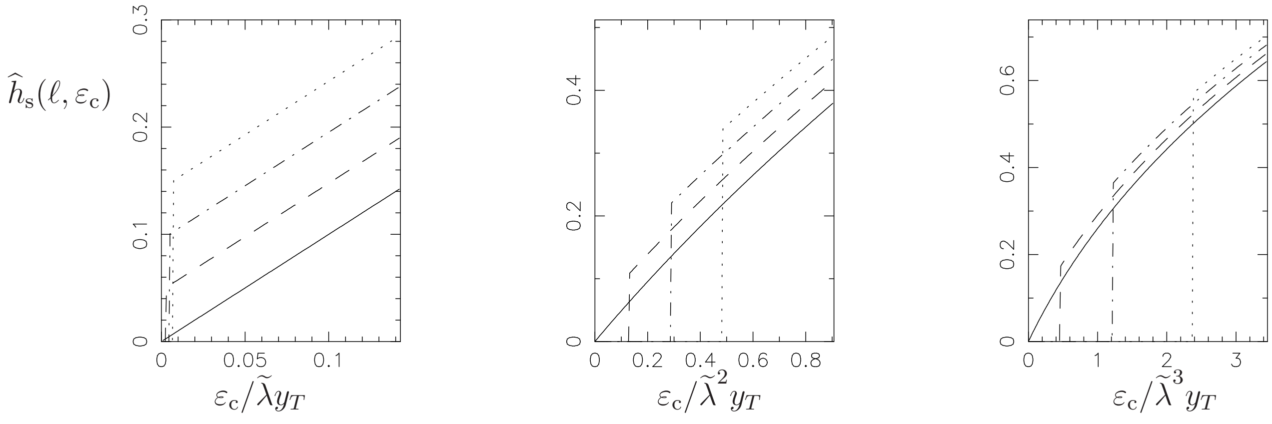
§ **Forecast:** $y^{\rm s}_{T+1} = 0.9 y_T =$ **4.725.**

§ **Outcome:**

- $y_{T+1} =$ **4.75 on 18.9.2001.**
- -0.5% **forecast error.**

10.3 *Crossing Robustness Curves: General Case*

§ Recall numerical example:



§ Crossing robustness curves:

Advantage of sub-optimal model.

§ How general?

§ System:

$$y_t = A_t y_{t-1}, \quad y_t \in \Re^N$$

§ Incorporate inputs into state vector.

§ Ignore zero-mean, additive, random disturbances.

§ Goal:

Given y_T and historical $\widetilde{A}$,

predict y_{T+k} within $\pm\varepsilon_c$.

§ Problem: Uncertain future A_t.

§ **Info-gap model,** $\mathcal{U}(h, \widetilde{A})$, $h \geq 0$.

Axioms:

Nesting: $h < h'$ **implies** $\mathcal{U}(h, \widetilde{A}) \subset \mathcal{U}(h', \widetilde{A})$

Contraction: $\mathcal{U}(0, \widetilde{A}) = \{\widetilde{A}\}$

§ **Two levels of uncertainty:**

- **Horizon of uncertainty,** h, **unknown.**
- **Realization unknown.**

§ Example: Unbounded-interval info-gap model:

$$\mathcal{U}(h, \widetilde{A}) = \left\{ A_t, \ t > T : \right.$$

$$\widetilde{A}_{ij} - h v_{ij} \le [A_t]_{ij} \le \widetilde{A}_{ij} + h w_{ij},$$

$$\left. i, j = 1, \ldots, N \right\}, \quad h \ge 0$$

§ Historically estimated model:

$$y_t = \widetilde{A} y_{t-1}$$

§ Slope-adjusted predictor:

$$y_t^{\mathrm{s}} = B y_{t-1}^{\mathrm{s}}$$

B **chosen by forecaster.**

§ Forecast error:

$$\eta_k(B, A_t) = y_{T+k}^{\mathrm{s}} - y_{T+k} = \left(B^k - \prod_{i=1}^{k} A_{T+i} \right) y_T$$

§ Satisficing forecast requirement:

$$|\eta_{k,m}(B, A_t)| \le \varepsilon_{\rm c}$$

§ Robustness: Max tolerable uncertainty.

$$\widehat{h}(B, \varepsilon_{\rm c}) = \max \left\{ h : \left(\max_{\substack{A_{T+i} \in \mathcal{U}(h, \widetilde{A}) \\ i=1, \ldots, k}} |\eta_{k,m}(B, A_t)| \right) \le \varepsilon_{\rm c} \right\}$$

§ k-Step transition matrices:

$$\mathcal{U}_k(h, \widetilde{A}^k) = \left\{ A = \prod_{i=1}^{k} A_i : \ \ A_i \in \mathcal{U}(h, \widetilde{A}) \right\}, \qquad h \geq 0$$

Lemma 1 ***If*** $\mathcal{U}(h, \widetilde{A})$ ***is an info-gap model. Then*** $\mathcal{U}_k(h, \widetilde{A}^k)$ ***is an info-gap model.***

§ 1-step thm based on

nesting and contraction axioms

is k-step thm.

§ Robustness-premium theorem. Define:

$$\theta_{\rm c}(h) = \max_{A_{T+1}\in\mathcal{U}(h,\widetilde{A})} \sum_{n=1}^{N} [A_{T+1} - \widetilde{A}]_{mn} y_{T,n}$$

$$\theta_{\rm a}(h) = -\min_{A_{T+1}\in\mathcal{U}(h,\widetilde{A})} \sum_{n=1}^{N} [A_{T+1} - \widetilde{A}]_{mn} y_{T,n}$$

- **Contraction:** $\theta_{\rm a}(0) = \theta_{\rm c}(0) = 0$.
- **Nesting:** $\theta_{\rm a}(h)$ **and** $\theta_{\rm c}(h)$ **increase with** h.
- $\theta_{\rm c}(h)$ **large: IGM coherent with** y_T.
- $\theta_{\rm a}(h)$ **large: IGM anti-coherent with** y_T.

Theorem 1 ***Sub-optimal models are more robust than optimal models.***

Given:

- $y_T \neq 0$.
- $\mathcal{U}(h, \widetilde{A})$ ***is an info-gap model.***
- $\theta_{\rm c}(h)$ ***and*** $\theta_{\rm a}(h)$ ***are continuous, at least one is unbounded, and either*** $\theta_{\rm c}(h) \geq \theta_{\rm a}(h)$ ***or*** $\theta_{\rm a}(h) \geq \theta_{\rm c}(h)$ ***for all*** $h > 0$.

Then: ***for any*** $\varepsilon_{\rm c} > 0$ ***for which*** $\delta_{\times}(\varepsilon_{\rm c}) \neq 0$, ***there is a*** B ***such that:***

$$\widehat{h}(B, \varepsilon_{\rm c}) > \widehat{h}(\widetilde{A}, \varepsilon_{\rm c})$$

§ **Evaluating the robustness.** **Define:**

$$\theta_{\rm c}(h) = \max_{A_{T+1}\in\mathcal{U}(h,\widetilde{A})} \sum_{n=1}^{N} [A_{T+1} - \widetilde{A}]_{mn} y_{T,n}$$

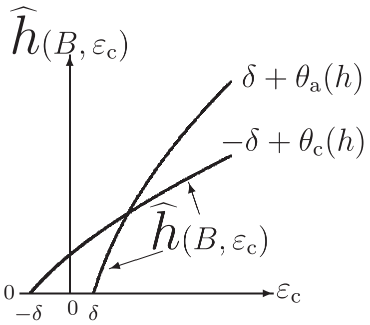
$$\theta_{\rm a}(h) = - \min_{A_{T+1}\in\mathcal{U}(h,\widetilde{A})} \sum_{n=1}^{N} [A_{T+1} - \widetilde{A}]_{mn} y_{T,n}$$

$$\delta = \sum_{n=1}^{N} [B - \widetilde{A}]_{mn} y_{T,n}$$

- **Contraction:** $\theta_{\rm a}(0) = \theta_{\rm c}(0) = 0$.
- **Nesting:** $\theta_{\rm a}(h)$ **and** $\theta_{\rm c}(h)$ **increase with** h.
- $\theta_{\rm c}(h)$ **large: IGM coherent with** y_T.
- $\theta_{\rm a}(h)$ **large: IGM anti-coherent with** y_T.
- δ **controlled by forecaster.**

§ Robustness function:

$$\widehat{h}(B, \varepsilon_{\rm c}) = \max \left\{ h : \ \delta + \theta_{\rm a}(h) \le \varepsilon_{\rm c} \ \textbf{and} \ -\delta + \theta_{\rm c}(h) \le \varepsilon_{\rm c} \right\}$$



§ When is the robustness zero?

§ Anticipated 1-step prediction error:

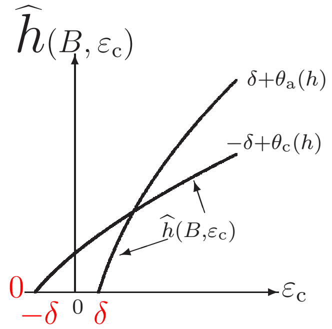
$$\eta_{1,m}(B,\widetilde{A}) = \underbrace{\sum_{n=1}^{N} \left(B - \widetilde{A}\right)_{mn} y_{T,n}}_{\delta}$$

§ **When is the robustness zero?**

§ **Anticipated 1-step prediction error:**

$$\eta_{1,m}(B,\widetilde{A}) = \underbrace{\sum_{n=1}^{N} \left(B-\widetilde{A}\right)_{mn} y_{T,n}}_{\delta}$$

$\varepsilon_{\rm c} = \delta$ **has zero robustness:**

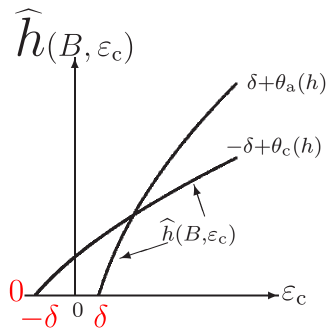


§ **Positive robustness only at greater-than-predicted forecast error.**

§ Choose $B = \widetilde{A}$ to minimize anticipated prediction error, $\eta_{1,m}(B, \widetilde{A})$???

§ **Choose $B = \widetilde{A}$ to minimize anticipated prediction error, $\eta_{1,m}(B, \widetilde{A})$???**

- **Predicted error is $\delta = 0$.**
- **$\varepsilon_c = \delta$ still has zero robustness:**



§ **Positive robustness only at greater-than-predicted forecast error.**

10.4 *Robustness and Probability of Forecast Success*

§ Probability of Forecast Success.

- **Forecast success:** $|\eta_{1,m}| \le \varepsilon_{\rm c}$**, or:**

$$\delta - \varepsilon_{\rm c} \le \underbrace{\sum_{n=1}^{N} \left[A_{T+1} - \widetilde{A} \right]_{mn} y_{T,n}}_{u} \le \delta + \varepsilon_{\rm c}$$

$u =$ **random variable.**

$F(u) =$ **cdf (unknown).**

$\delta(B) =$ **anticip. forecast error (chosen).**

- **Probability of forecast success:**

$$P_{\rm s}(B) = F(\delta + \varepsilon_{\rm c}) - F(\delta - \varepsilon_{\rm c})$$

Thus:

$$\frac{{\rm d}P_{\rm s}(B)}{{\rm d}\delta} > 0 \quad \textbf{if and only if} \quad f(\delta+\varepsilon_{\rm c}) > f(\delta-\varepsilon_{\rm c})$$

§ $\theta_{\mathrm{c}}(h) > \theta_{\mathrm{a}}(h)$ implies:

- $[A_{T+1} - \widetilde{A}]_{mn}$ **coherent with sgn**(y_{Tn}).
- u **tends to be positive.** $u = \Sigma_{n=1}^{N} \left[A_{T+1} - \widetilde{A}\right]_{mn} y_{T,n}$
- $f(u)$ **tends to increase around** $u = 0$.

Definition 1 $\mathcal{U}(h, \widetilde{A})$ ***and*** $F(u)$ ***coherent at*** $(\delta, \varepsilon_{\mathrm{c}})$ ***if:***

$$[\theta_{\mathrm{c}}(h) - \theta_{\mathrm{a}}(h)]\,[f(\delta + \varepsilon_{\mathrm{c}}) - f(\delta - \varepsilon_{\mathrm{c}})] \geq 0 \;\; \textbf{\textit{for all}}\; h > 0$$

- **Coherence:**

  The info-gap model

  weakly reveals the pdf.

Theorem 2 ***Robustness is a proxy for probability of forecast success. Given:***

- $y_T \neq 0$, $\varepsilon_{\rm c} \geq 0$, $\delta_\times(\varepsilon_{\rm c}) \neq 0$, $|\delta| < |\delta_\times(\varepsilon_{\rm c})|$.
- $\widehat{h}(B, \varepsilon_{\rm c}) > 0$.
- $\theta_{\rm a}(h)$ ***and*** $\theta_{\rm c}(h)$ ***are continuous and at least one is unbounded.***
- $\mathcal{U}(h, \widetilde{A})$ ***and*** $F(u)$ ***are coherent.***

Then:

$$\frac{{\rm d}\widehat{h}(B, \varepsilon_{\rm c})}{{\rm d}\delta} > 0 \quad \textit{if and only if} \quad \frac{{\rm d}P_{\rm s}(B)}{{\rm d}\delta} > 0$$

§ Importance of proxy thm:

- $P_s(B)$ **unknown.**
- $\widehat{h}(B, \varepsilon_c)$ **known.**
- B **chosen by forecaster.**

§ Coherence of $\mathcal{U}(h, \widetilde{A})$ and $F(u)$ implies:

P_s **is not known but can be optimized.**

§ Many proxy theorems.

- A = model, data: uncertain.
- $\mathcal{U}(h, \widetilde{A})$ = info-gap model for uncertainty.
- B = design, decision, strategy.
- $\eta(B, A)$ = outcome: $\eta(B, A) \le \varepsilon_{\rm c}$.
- $\widehat{h}(B, \varepsilon_{\rm c})$ = robustness function.
- $P_{\rm s}(B)$ = probability of success.
- Proxy "theorem":

$$\left(\frac{\partial \widehat{h}(B, \varepsilon_{\rm c})}{\partial B}\right)\left(\frac{\partial P_{\rm s}(B)}{\partial B}\right) \ge 0$$

- Fine print: e.g. $\mathcal{U}(h, \widetilde{A})$ & $F(A)$ coherent.

§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \widetilde{\lambda}$.
- **Future:** $\lambda_t = \widetilde{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \textbf{unknown}, & u_t \geq \lambda_\star \end{cases}$.

E.g. $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \lambda_\star / u^2, & u_t \geq \lambda_\star \end{cases}$.

§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \widetilde{\lambda}$.
- **Future:** $\lambda_t = \widetilde{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \textbf{unknown}, & u_t \geq \lambda_\star \end{cases}$.

  E.g. $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \lambda_\star / u^2, & u_t \geq \lambda_\star \end{cases}$.
- **Forecaster:** $x_t^{\rm s} = \ell x_{t-1}^{\rm s}$.
- δ = **anticipated forecast error** = $(\ell - \widetilde{\lambda}) y_T$.

§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \widetilde{\lambda}$.
- **Future:** $\lambda_t = \widetilde{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \textbf{unknown}, & u_t \geq \lambda_\star \end{cases}$.
- **Forecaster:** $x_t^{\mathrm{s}} = \ell x_{t-1}^{\mathrm{s}}$.
- $\delta =$ **anticipated forecast error** $= (\ell - \widetilde{\lambda}) y_T$.
- $\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda : \ \widetilde{\lambda} \leq \lambda \leq (1+h)\widetilde{\lambda} \right\}, \ h \geq 0$.
- $\mathcal{U}(h, \widetilde{\lambda})$ **and** $F(u)$ **coherent if:**

$$\widetilde{\lambda} + \lambda_\star - \frac{\varepsilon_{\mathrm{c}}}{y_T} \ < \ \ell \ < \ \widetilde{\lambda} + \lambda_\star + \frac{\varepsilon_{\mathrm{c}}}{y_T}.$$

§ Coherence: Example.

- **System:** $x_t = \lambda_t x_{t-1}$. **Historically:** $\lambda_t = \widetilde{\lambda}$.
- **Future:** $\lambda_t = \widetilde{\lambda} + u_t$, $f(u_t) = \begin{cases} 0, & u_t < \lambda_\star \\ \textbf{unknown}, & u_t \geq \lambda_\star \end{cases}$.
- **Forecaster:** $x_t^{\mathrm{s}} = \ell x_{t-1}^{\mathrm{s}}$.
- $\delta =$ **anticipated forecast error** $= (\ell - \widetilde{\lambda}) y_T$.
- $\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda : \ \widetilde{\lambda} \leq \lambda \leq (1+h)\widetilde{\lambda} \right\}, \ h \geq 0$.
- $\mathcal{U}(h, \widetilde{\lambda})$ **and** $F(u)$ **coherent if:**

  $\widetilde{\lambda} + \lambda_\star - \frac{\varepsilon_{\mathrm{c}}}{y_T} \ < \ \ell \ < \ \widetilde{\lambda} + \lambda_\star + \frac{\varepsilon_{\mathrm{c}}}{y_T}$.
- **Then:** $\dfrac{\partial P_{\mathrm{s}}}{\partial \ell} > 0$.
- $P_{\mathrm{s}}(\ell)$ **unknown but improvable.**

10.5 *Regression Prediction*

Yakov Ben-Haim,

Info-Gap Economics:

An Operational Introduction,

2010, Palgrave-Macmillan, section 6.1.

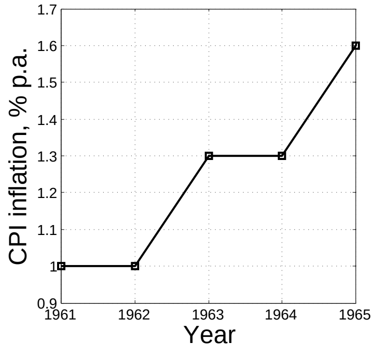


Figure 5: **US inflation vs. year, 1961–1965.**

§ US inflation '61–'65: Linear?

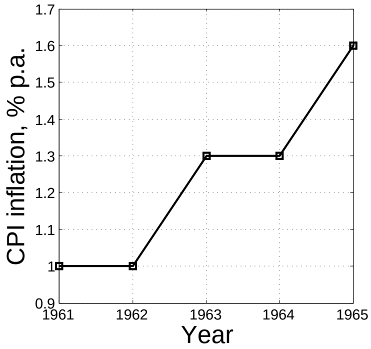


Figure 6: **US inflation vs. year, 1961–1965.**

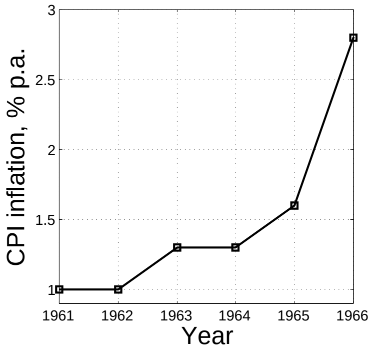


Figure 7: **US inflation vs. year, 1961–1966.**

§ US inflation '61–'65: Linear?

§ US inflation '61–'66: Quadratic?

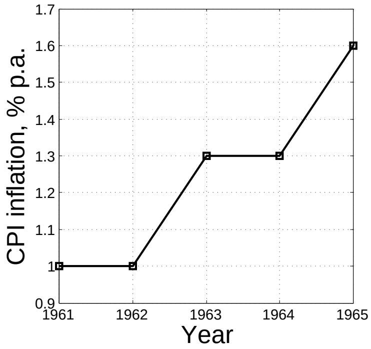


Figure 8: **US inflation vs. year, 1961–1965.**

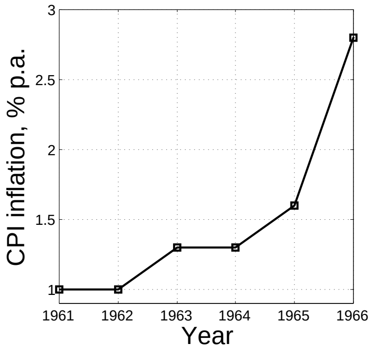


Figure 9: **US inflation vs. year, 1961–1966.**

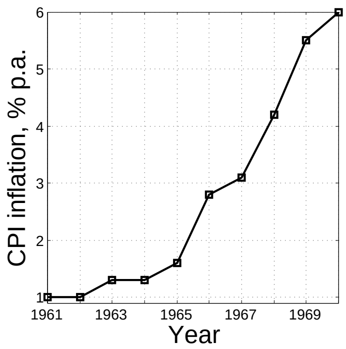


Figure 10: **US inflation vs. year, 1961–1970.**

§ US inflation '61–'65: Linear?

§ US inflation '61–'66: Quadratic?

§ US inflation '61–'70: Piece-wise linear?

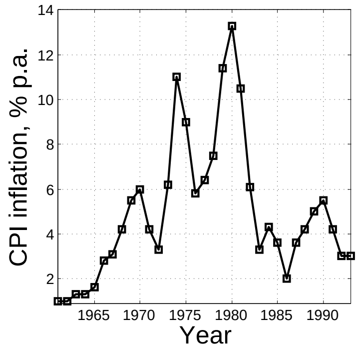


Figure 11: **US inflation vs. year, 1961–1993.**

§ US inflation '61–'93: A mess?

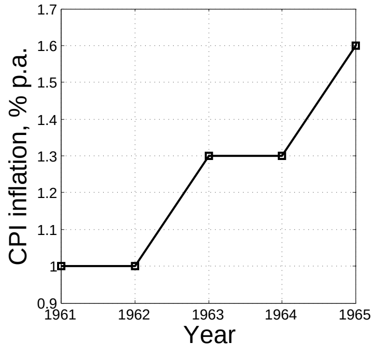


Figure 12: **US inflation vs. year, 1961–1965.**

§ US inflation '61–'65:

- **Linear? Quadratic?**
- **Model '61–'65 for predicting '66:**

$$y_i^{\mathrm{r}} = c_0 + c_1 t_i + c_2 t_i^2$$

§ System model: MSE.

$$S_N^2(c) = \frac{1}{N} \sum_{i=1}^{N} (y_i - y_i^{\mathrm{r}})^2$$

$N = 5$ **for '61–'65.**

§ If we knew y_{N+1} ('66):

$$S^2_{N+1}(c) = \frac{1}{N+1}\sum_{i=1}^{N+1}\left(y_i - y_i^{\mathrm{r}}\right)^2$$

$$= \frac{N}{N+1}S^2_N(c) + \frac{\left(y_{N+1} - y^{\mathrm{r}}_{N+1}\right)^2}{N+1}$$

§ **If we knew y_{N+1} ('66):**

$$S^2_{N+1}(c) = \frac{1}{N+1}\sum_{i=1}^{N+1}(y_i - y^{\mathrm{r}}_i)^2$$

$$= \frac{N}{N+1}S^2_N(c) + \frac{(y_{N+1} - y^{\mathrm{r}}_{N+1})^2}{N+1}$$

§ **All we know is contextual info:**

y_{N+1} **may well exceed prediction,** $\widetilde{y}^{\mathrm{r}}_{N+1}$**.**

§ **If we knew** y_{N+1} **('66):**

$$S^2_{N+1}(c) = \frac{1}{N+1}\sum_{i=1}^{N+1}(y_i - y_i^{\mathrm{r}})^2$$
$$= \frac{N}{N+1}S^2_N(c) + \frac{(y_{N+1} - y^{\mathrm{r}}_{N+1})^2}{N+1}$$

§ **All we know is soft info:**

y_{N+1} **may well exceed prediction,** $\widetilde{y}^{\mathrm{r}}_{N+1}$**.**

§ **Info-gap model of uncertain** y_{N+1}**:**

$$\mathcal{U}(h) = \left\{y_{N+1} : \; 0 \le y_{N+1} - \widetilde{y}^{\mathrm{r}}_{N+1} \le h\right\}, \quad h \ge 0$$

- **Unbounded family of nested sets.**
- **No worst case.**

§ Performance requirement:

$$S_{N+1}(c) \leq S_{\mathrm{c}}$$

§ Performance requirement:

$$S_{N+1}(c) \le S_{\rm c}$$

§ Robustness of regression c:

Greatest tolerable uncertainty.

$$\widehat{h}(c, S_{\rm c}) = \max\left\{h : \ \left(\max_{y_{N+1}\in\mathcal{U}(h)} S_{N+1}(c)\right) \le S_{\rm c}\right\}$$

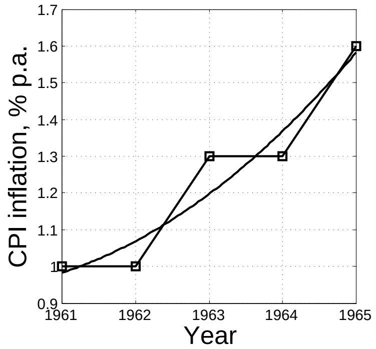


Figure 13: **US inflation vs. year, 1961–1965, and least squares fit.**

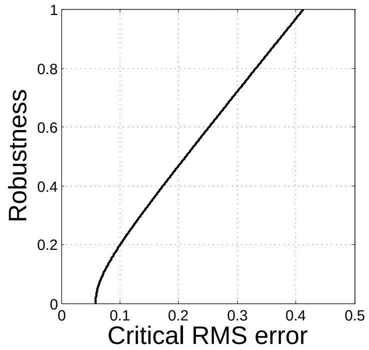


Figure 14: **Robustness vs. critical root mean squared error for inflation 1961–1965.**

§ Least squares fit: fig. 13.

§ Robust of LS fit: fig. 14.

§ Trade off: Greater rbs. ≡ greater MSE.

§ Zeroing: No robustness of est. MSE.

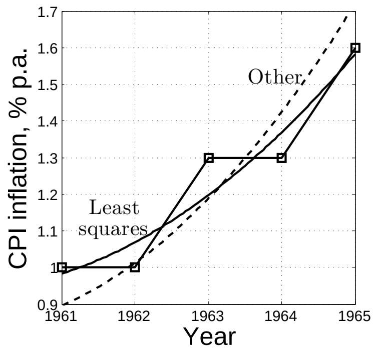


Figure 15: **US inflation vs. year, 1961–1965, and least squares fit (solid) and other fit (dash).**

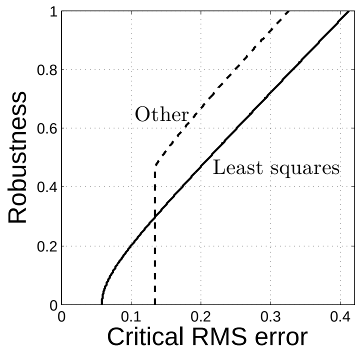


Figure 16: **Robustness vs. critical root mean squared error for inflation 1961–1965 for least squares fit (solid) and other fit (dash).**

§ Least squares and other fit: fig. 15.

§ Robust of LS and other fit: fig. 16.

Curve-crossing: preference reversal.

11 *INFO-GAP ROBUST-SATISFICING FORAGING*

¶ **Joint work with Dr. Yohay Carmel.**

***American Naturalist*, 2005, 166: 633–641.**

¶ **Foraging decision:**

- **Animal currently feeding "here".**
- **To move or not to move?**

  That is the question.

[3] lectures talks lib foraging02.tex 20.8.2010

§ The animal's problem:

Survive.

§ Optimization strategy:

Maximize energy intake.

§ Robust-satisficing strategy:

- **Satisfice energy intake.**
- **Maximize robustness to uncertainty.**

¶ **Variables:**

G_0 = **energy gain here (highly uncertain).**

G_1 = **energy gain there (highly uncertain).**

C = **cost of moving.**

T = **remaining time.**

¶ Optimal foraging:

- **Maximize best estimate of energy.**
- **Leave patch when:**

$$\widetilde{G}_0 T < \widetilde{G}_1 T - C$$

¶ Robust satisficing:

- E_{min} = **least acceptable energy.**
- **Reliably achieve** E_{min}**.**

§ Robustness question:

For any energy E_{min},

how wrong can $\widetilde{G}_0, \widetilde{G}_1$ be, and

energy gain is at least E_{min}

with contemplated action, A?

§ **Robustness question:**

For any energy $E_{\min}$,

how wrong can $\widetilde{G}_0, \widetilde{G}_1$ be, and

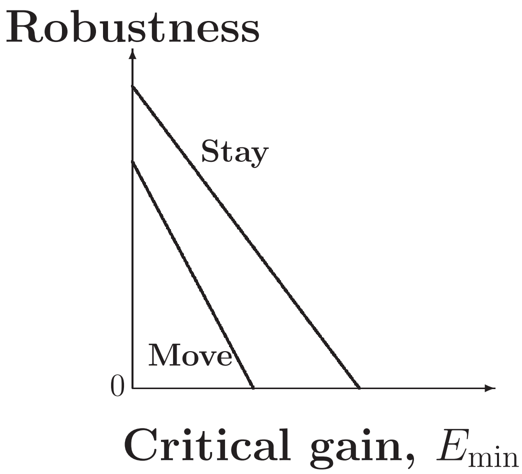
energy gain is at least $E_{\min}$

with contemplated action, A?

§ **Answer: robustness function.**

§ Two cases:

- **Optimal to stay:** $\widetilde{G}_1 T - C < \widetilde{G}_0 T$.
- **Optimal to move:** $\widetilde{G}_1 T - C > \widetilde{G}_0 T$.

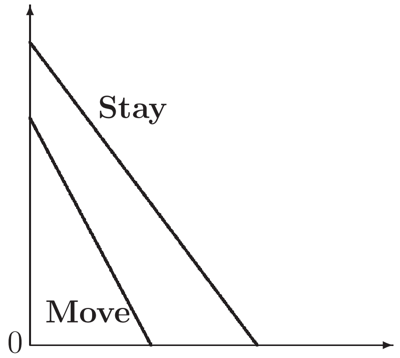


Optimal to stay.

$$\widetilde{G}_1 T - C < \widetilde{G}_0 T$$

Satisficer also stays.

Robustness



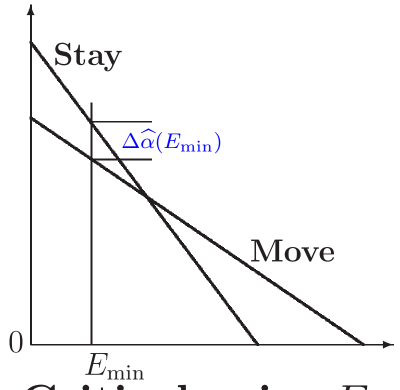
Critical gain, $E_{\rm min}$

Optimal to stay.

$$\widetilde{G}_1 T - C < \widetilde{G}_0 T$$

Satisficer also stays.

Robustness



Critical gain, $E_{\rm min}$

Optimal to move.

$$\widetilde{G}_1 T - C > \widetilde{G}_0 T$$

Satisficer may stay.

¶ **Literature survey of 34 studies suggests support for robust-satisficing in many taxa.**

¶ Discussion

¶ Why optimize?

More is better than less.

Hence most is best.

¶ **Why not optimize?**

(Even though most is best.)

¶ **Bounded rationality (Simon):**

- **Poor information.**
- **Poor information processing.**
- **Optimization unreliable, infeasible.**

¶ **Why not optimize?**

(Even though most is best.)

¶ **Bounded rationality. Optim. unreliable.**

¶ **Robust-satisficing is proxy for**

Probability of success:

$$\textbf{max Robustness} \implies \textbf{max Prob of Success}$$
$$\implies \textbf{max Prob of Survival}$$

¶ Darwin and satisficing:

- **Evolution:**
  - **Survival of more fit over less fit (*O of S*).**
  - **Don't optimize. Satisfice:**
    - — **Beat competition.**
    - — **Robustify against surprises.**

¶ Darwin and satisficing:

- **Species distribution:**
  - **"Great fact"** (*O of S*)**:**

Similar habitats in old & new worlds have "widely different . . . living productions!".

¶ Darwin and satisficing:

- **Species distribution:**
  - **"Great fact" (*O of S*):**

Similar habitats in old & new worlds have "widely different . . . living productions!".

  - **Optimization:**

Similar phenotypes under similar constraints.

  - **Robust satisficing:**

Diversity from performance-sub-optimality.

¶ Foraging animals seem to robust-satisfice.

¶ Questions:

- **Is this mental frailty of animals or evolutionary guile?**
- **It seems that satisficing is strategically advantageous.**

12 *Conservation Management, or: Save the Sumatran Rhinoceros*

§ COLLABORATORS:

- **Helen M. Regan**
- **Bill Langford**
- **Will G. Wilson**
- **Per Lundberg**
- **Sandy J. Andelman**
- **Mark A. Burgman**

***Ecological Applications*, 2005, vol.15(4): 1471–1477.**

[3] \lectures\talks\lib\rhinoceros02.tex 27.2.2012

12.1 The Problem: Endangered Species

§ Sumatran rhinoceros: Endangered species.

§ We will first develop expected utility.

- We need:
  - Probabilities.
  - Utilities.
- Problem: very limited knowledge.

§ Solution: embed expected utility in info-gap robust-satisficing decision.

12.2 Expected Utility

§ Three elements:

- **States.**
- **Actions.**
- **Utilities.**

§ States of the world:

Possible causes of decline of Sumatran rhinoceros:

- **Poaching.**
- **Loss of habitat.**
- **Demographic accidents.**
- **Disease.**

P_J = **Probability that world in state** J**.**

§ **Actions,** A_1, A_2, . . ., **to protect rhino:**

- **Translocation of population to new region.**
- **Extension of current reserve.**
- **Captive breeding.**

§ Utilities:

- V_{IJ} = **utility of action** A_I **in state** J.
  = **Prob of survival for specified duration.**
- V_{IJ} **is unnormalized probability. e.g.:**
  - **all the** V_{IJ} **may be near 0.**
  - **all the** V_{IJ} **may be near 1.**

§ **Expected utility** of action A_I:

$$\mathrm{E}(A_I) = \sum_J V_{IJ} P_J$$

§ **Expected utility of optimal action,** $A^\star$:

$$A^\star = \arg\max_{A_I} \mathrm{E}(A_I) \quad (1)$$

$$= \arg\max_{A_I} \sum_J V_{IJ} P_J \quad (2)$$

12.3 Uncertainties

§ Expected utility deals with:

- **Uncertainty in state of the world:**

  Probability P_J.

- **Uncertainty in survival:**

  Probability V_{IJ}.

§ Another very important uncertainty:

P_J and V_{IJ} are poorly known:

large info-gaps.

§ Info-gaps:

Unknown fractional error.

$$\frac{|P_J - \widetilde{P}_J|}{\widetilde{P}_J} \leq h, \qquad \frac{|V_{IJ} - \widetilde{V}_{IJ}|}{\widetilde{V}_{IJ}} \leq h$$

§ **Info-gaps:**

Unknown fractional error.

$$\frac{|P_J - \widetilde{P}_J|}{\widetilde{P}_J} \le h, \qquad \frac{|V_{IJ} - \widetilde{V}_{IJ}|}{\widetilde{V}_{IJ}} \le h$$

§ **Info-gap models** $\mathcal{P}(h, \widetilde{P})$ **and** $\mathcal{V}(h, \widetilde{V})$**:**

$$\mathcal{P}(h) = \left\{ P : \frac{|P_J - \widetilde{P}_J|}{\widetilde{P}_J} \le h, \ P_J \in [0,1] \ \forall \ J, \sum_J P_J = 1 \right\}, \quad h \ge 0$$

$$\mathcal{V}(h) = \left\{ V : \frac{|V_{IJ} - \widetilde{V}_{IJ}|}{\widetilde{V}_{IJ}} \le h, \ V_{IJ} \in [0,1] \ \forall \ I, J, \right\}, \quad h \ge 0$$

- **Contraction.**
- **Nesting: no known worst case.**

12.4 Robustness

§ The problem:

- **Best estimates of probabilities: $\widetilde{P}$ and $\widetilde{V}$.**
- **Best estimate of expected utility: $\mathrm{E}(A_I, \widetilde{P}, \widetilde{V})$.**
- **Actual expected utility of action A_I: $\mathrm{E}(A_I, P, V)$.**
- **Actual and anticipated E.U. differ:**

$$\mathrm{E}(A_I, P, V) \neq \mathrm{E}(A_I, \widetilde{P}, \widetilde{V})$$

§ Robustness question:

- **How wrong can $\widetilde{P}$ and $\widetilde{V}$ be without causing failure?**
- **This is an epistemic question, not an ontological question.**

§ E_c: **Lowest acceptable utility (survival prob).**

§ **Robustness of action A_I,**
to uncertainty in P_J and V_{IJ}:
Max horizon of uncertainty h
up to which $\mathrm{E}(A_I, P, V)$ is acceptable.

$$\widehat{h}(A_I, E_c) = \max \left\{ h : \ \left(\min_{\substack{P \in \mathcal{P}(h, \widetilde{P}) \\ V \in \mathcal{V}(h, \widetilde{V})}} \sum_j V_{IJ} P_J \right) \geq E_c \right\}$$

§ **Robust optimal action,** $\widehat{A}(E_{\rm c})$**:**

$$\widehat{A}(E_{\rm c}) = \arg\max_{A_I} \widehat{h}(A_I, E_{\rm c})$$

§ $\widehat{A}(E_{\rm c})$ **depends on aspiration** $E_{\rm c}$**.**

§ **Usually:**

$$\widehat{A}(E_{\rm c}) \;\neq\; A^{\star}$$

12.5 Robustness: Derivation at a Glance

§ Recall the robustness definition:

$$\widehat{h}(A_I, E_{\rm c}) = \max\left\{ h : \ \left(\min_{\substack{P \in \mathcal{P}(h, \widetilde{P}) \\ V \in \mathcal{V}(h, \widetilde{V})}} \sum_j V_{IJ} P_J \right) \geq E_{\rm c} \right\}$$

§ Recall the robustness definition:

$$\widehat{h}(A_I, E_c) = \max\left\{ h : \left(\min_{\substack{P \in \mathcal{P}(h,\widetilde{P}) \\ V \in \mathcal{V}(h,\widetilde{V})}} \sum_j V_{IJ} P_J \right) \geq E_c \right\}$$

§ Inverse of robustness: Inner min, $\mu(h)$.

- **IGM's nested $\Longrightarrow$**

  $\mu(h)$ decreases as h increases.

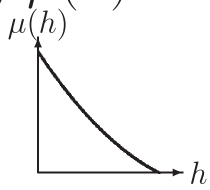


Figure 17: Robustness and its inverse.

§ Recall the robustness definition:

$$\widehat{h}(A_I, E_c) = \max \left\{ h : \left(\min_{\substack{P \in \mathcal{P}(h, \widetilde{P}) \\ V \in \mathcal{V}(h, \widetilde{V})}} \sum_j V_{IJ} P_J \right) \geq E_c \right\}$$

§ Inverse of robustness: Inner min, $\mu(h)$.

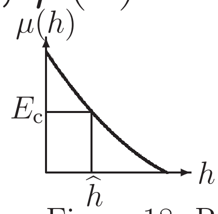


Figure 18: Robustness and its inverse.

- **IGM's nested $\Longrightarrow$**

  $\mu(h)$ decreases as h increases.

- **$\widehat{h} = \max h$ @ $\mu(h) = E_c$.**

§ Recall the robustness definition:

$$\widehat{h}(A_I, E_c) = \max\left\{ h : \left(\min_{\substack{P \in \mathcal{P}(h,\widetilde{P}) \\ V \in \mathcal{V}(h,\widetilde{V})}} \sum_j V_{IJ} P_J \right) \geq E_c \right\}$$

§ Inverse of robustness: Inner min, $\mu(h)$.

- **IGM's nested** $\Longrightarrow$

  $\mu(h)$ **decreases as** h **increases.**

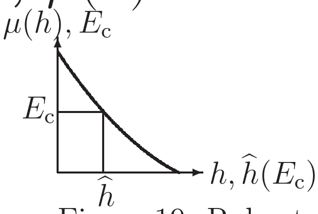


Figure 19: Robustness and its inverse.

- $\widehat{h} = \max h$ **@** $\mu(h) = E_c$.
- $\widehat{h}(E_c) = h$ **iff** $\mu(h) = E_c$:

Plot of h vs $\mu(h)$ $\equiv$ Plot of $\widehat{h}(E_c)$ vs E_c.

§ Recall the info-gap models:

$$\mathcal{P}(h) = \left\{ P : \frac{|P_J - \widetilde{P}_J|}{\widetilde{P}_J} \leq h, \; P_J \in [0,1] \; \forall \; J, \right.$$
$$\left. \sum_J P_J = 1 \right\}, \quad h \geq 0$$

$$\mathcal{V}(h) = \left\{ V : \frac{|V_{IJ} - \widetilde{V}_{IJ}|}{\widetilde{V}_{IJ}} \leq h, \right.$$
$$\left. V_{IJ} \in [0,1] \; \forall \; I, J, \right\}, \quad h \geq 0$$

§ Choose $V \in \mathcal{V}$, $P \in \mathcal{P}$ to minimize $\mu(h)$:

- $\mu(h) = \min \Sigma_j V_{ij} P_j$.

§ Choose $V \in \mathcal{V}$, $P \in \mathcal{P}$ to minimize $\mu(h)$:

- $\mu(h) = \min \sum_j V_{ij} P_j$.
- **Choose** $V_{ij} = (1-h)\widetilde{V}_{ij}$. $(h \leq 1)$.

§ **Choose $V \in \mathcal{V}$, $P \in \mathcal{P}$ to minimize $\mu(h)$:**

- $\mu(h) = \min \Sigma_j V_{ij} P_j$.
- **Choose** $V_{ij} = (1-h)\widetilde{V}_{ij}$. $(h \le 1)$.
- **Choose** P_j **by**

  Robin Hood Principle on $\widetilde{V}_{ij}\widetilde{P}_j$.

  - **If** $\widetilde{V}_{ij}\widetilde{P}_j$ **large,**

    choose P_j **small, near** $(1-h)\widetilde{P}_j$.

§ **Choose $V \in \mathcal{V}$, $P \in \mathcal{P}$ to minimize $\mu(h)$:**

- $\mu(h) = \min \Sigma_j V_{ij} P_j$.
- **Choose** $V_{ij} = (1-h)\widetilde{V}_{ij}$. $(h \le 1)$.
- **Choose** P_j **by**

  Robin Hood Principle on $\widetilde{V}_{ij}\widetilde{P}_j$.

  - **If** $\widetilde{V}_{ij}\widetilde{P}_j$ **large,**

    choose P_j **small, near** $(1-h)\widetilde{P}_j$.
  - **If** $\widetilde{V}_{ij}\widetilde{P}_j$ **small,**

    choose P_j **large, near** $(1+h)\widetilde{P}_j$.

§ **Choose $V \in \mathcal{V}$, $P \in \mathcal{P}$ to minimize $\mu(h)$:**

- $\mu(h) = \min \Sigma_j V_{ij} P_j$.
- **Choose** $V_{ij} = (1-h)\widetilde{V}_{ij}$. $(h \le 1)$.
- **Choose** P_j **by**

  Robin Hood Principle on $\widetilde{V}_{ij}\widetilde{P}_j$.

  - **If** $\widetilde{V}_{ij}\widetilde{P}_j$ **large,**

    choose P_j **small, near** $(1-h)\widetilde{P}_j$.
  - **If** $\widetilde{V}_{ij}\widetilde{P}_j$ **small,**

    choose P_j **large, near** $(1+h)\widetilde{P}_j$.
  - **Normalize** P.

12.6 Example

State (Cause of decline)	Prob. $\widetilde{P}_J$ of that state	Cond. Prob. $\widetilde{V}_{1J}$ (Translocation) (A_1)	Cond. Prob. $\widetilde{V}_{2J}$ (New reserve) (A_2)	Cond. Prob. $\widetilde{V}_{3J}$ (Captive breeding) (A_3)
Poaching	0.1	0.3	0.25	0.9
Loss of habitat	0.3	0.1	0.2	0.2
Demographic accidents	0.5	0.05	0.09	0.01
Disease	0.1	0.1	0.1	0.4
Expected utility		$\sum_J \widetilde{V}_{1J}\widetilde{P}_J =$ **0.095**	$\sum_J \widetilde{V}_{2J}\widetilde{P}_J =$ **0.14**	$\sum_J \widetilde{V}_{3J}\widetilde{P}_J =$ **0.195**

§ **Expected utility optimum:** $A^\star = A_3$.

§ Questions:

- **Is $A^\star$ a good strategy?**
- **How reliable is attainment of survival prob $\geq$ $\mathrm{E}(A^\star) = 0.195$?**
- **How costly is robustness?**

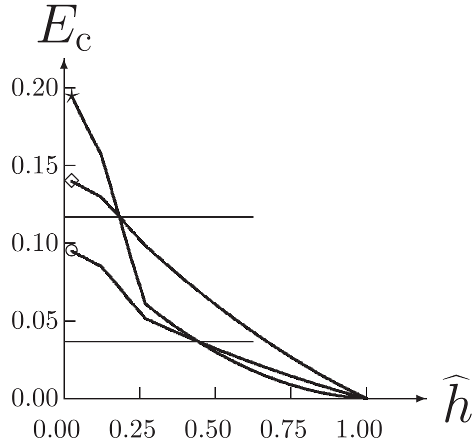


Figure 20: Robustness curves for actions 1 (○), 2 (◇) and 3 (⋆).

§ **Trade-off between performance (E_c) and robustness ($\widehat{h}$).**

§ **No robustness at nominal expected utility.**

§ **No robustness at EU optimum, $A^\star = A_3$.**

§ **Reversal of preference:**

- $E_c > 0.12$: $\widehat{A} = A_3$.
- $E_c < 0.12$: $\widehat{A} = A_2$.

§ Methodological summary:

- **Hybrid uncertainty:**
  Info-gaps in probabilistic model.
- **Info-gap model** of uncertainty:
  Unbounded family of nested sets of events.
- **2-tiered decision** analysis:
  - Core decision rule: expected utility.
  - Supervisory decision rule:
    Info-gap robust-satisficing.
- . . .

§ Methodological summary (continued):

- **Trade-off between performance and robustness:**
  - **Performance-maximization is always robustness minimization.**
  - **E.U. maximum has zero robustness.**
  - **Preferred action depends on aspiration: Preference reversal.**
- **Max-min vs Rob-Sat: may differ.**

13 *Probabilistic Reliability in Engineering with Info-Gaps*

§ Source:

Yakov Ben-Haim, *Info-Gap Economics: An Operational Introduction*, Palgrave, 2010.

[3] lectures\talks\lib\va-risk-engr01.tex 26.8.2012

13.1 The Problem: Uncertainties

§ The problem:

- **Fat tails:**
  - Extreme outcomes too frequent.
  - High percentiles under-estimated.
-

§ The problem:

- **Fat tails:**
  - Extreme outcomes too frequent.
  - High percentiles under-estimated.
- **Past vs future:**
  - Processes vary in time.
  - Data are revised.
  - Shackle-Popper indeterminism.
-

§ The problem:

- **Fat tails:**
  - Extreme outcomes too frequent.
  - High percentiles under-estimated.
- **Past vs future:**
  - Processes vary in time.
  - Data are revised.
  - Shackle-Popper indeterminism.
- **Joint probabilities:**
  - Uncertain common-mode failures.
  - Uncertain correlations.

§ Two foci of uncertainty:

- **Statistical fluctuations:**
  - **Randomness, “noise”.**
  - **Estimation uncertainty.**
-

§ **Two foci of uncertainty:**

- **Statistical fluctuations:**
  - **Randomness, "noise".**
  - **Estimation uncertainty.**
- **Knightian uncertainty:**
  - **Surprises.**
  - **Structural changes.**
  - **Historical data used to predict future.**

§ **Info-gap theory to manage Knightian uncertainty.**

§ Outline:

- **Discrete system with 2 sub-units:**
  - **Reliability.**
  - **Redundancy.**
  - **Uncertain correlation.**
- **Origin of fat tails:**

  Parameter uncertainty.
- **Quantile uncertainty.**
  - **Fat tails.**
  - **Thin tails.**

13.2 *Reliability, Redundancy, Uncertain Correlation*

§ **System with 2 sub-units:**

- **Serial: both units essential.**
- **Parallel: either unit sufficient.**

§

§ System with 2 sub-units:

- **Serial: both units essential.**
- **Parallel: either unit sufficient.**

§ Probabilities of failure:

- F_i = **marginal prob of failure of unit** i.
- F_{12} = **prob of joint failure.**

§

§ System with 2 sub-units:

- **Serial: both units essential.**
- **Parallel: either unit sufficient.**

§ Probabilities of failure:

- F_i = **marginal prob of failure of unit** i.
- F_{12} = **prob of joint failure.**

§ Probabilities of system failure:

- **Serial:** $F_{\rm s} = F_1 + F_2 - F_{12}$.
  (Explain: Venn)
- **Parallel:** $F_{\rm p} = F_{12}$

§ Example.

- **Estimates:** $\widetilde{F}_i = 0.01$. $\widetilde{F}_{12} = 0.0001$
- **Serial:** $F_{\rm s} = 0.0199$. **Not too good.**
- **Parallel:** $F_{\rm p} = 0.0001$. **Not too bad.**

§

§ Example.

- **Estimates:** $\widetilde{F}_i = 0.01$. $\widetilde{F}_{12} = 0.0001$
- **Serial:** $F_{\rm s} = 0.0199$. **Not too good.**
- **Parallel:** $F_{\rm p} = 0.0001$. **Not too bad.**

§ Problem:

Estimates of F_i and F_{12} uncertain.

$$\left|F_i - \widetilde{F}_i\right| \le h, \;\; i = 1, \, 2$$

$$\left|F_{12} - \widetilde{F}_{12}\right| \le h$$

Horizon of uncertainty, h, is unknown.

§ Info-gap model:

$$\mathcal{U}(h) = \left\{ F_1, F_2, F_{12} : \begin{array}{l} F_1 \geq 0, \quad F_2 \geq 0, \quad F_{12} \geq 0. \\ F_{12} \leq \min[F_1,\ F_2]. \\ F_1 + F_2 - F_{12} \leq 1. \\ \left|F_i - \widetilde{F}_i\right| \leq h,\ \ i = 1,\ 2. \\ \left|F_{12} - \widetilde{F}_{12}\right| \leq h \end{array} \right\}, \quad h \geq 0$$

- **Contraction:** $\mathcal{U}(0) = \{\widetilde{F}_1, \widetilde{F}_2, \widetilde{F}_{12}\}$.
- **Nesting:** $h < h' \implies \mathcal{U}(h) \subseteq \mathcal{U}(h')$.
- **Family of nested sets.**
- **No known worst case.**

§ Performance requirements:

- **Acceptable failure probabilities:**
  - **Serial:** $F_{\text{s}} \leq F_{\text{cs}}$
  - **Parallel:** $F_{\text{p}} \leq F_{\text{cp}}$

§

§ Performance requirements:

- **Acceptable failure probabilities:**
  - **Serial:** $F_{\mathrm{s}} \le F_{\mathrm{cs}}$
  - **Parallel:** $F_{\mathrm{p}} \le F_{\mathrm{cp}}$

§ Robustness:

- **Maximum tolerable uncertainty.**
- **Max horizon of uncertainty at which failure probability acceptable.**
- **Serial and parallel robustness:**

$$\widehat{h}_{\mathrm{s}}(F_{\mathrm{cs}}) = \max\left\{h : \left(\max_{F_i, F_{12} \in \mathcal{U}(h)} F_{\mathrm{s}} \right) \le F_{\mathrm{cs}} \right\}$$

$$\widehat{h}_{\mathrm{p}}(F_{\mathrm{cp}}) = \max\left\{h : \left(\max_{F_i, F_{12} \in \mathcal{U}(h)} F_{\mathrm{p}} \right) \le F_{\mathrm{cp}} \right\}$$

§ Derive parallel robustness.

§ Ramp function:

$$\overline{r}(x) = \begin{cases} 0 & x < 0 \\ x & 0 \le x \le 1 \\ 1 & 1 < x \end{cases}$$

§ Inverse of robustness:

$$\mu(h) = \max_{F_i, F_{12} \in \mathcal{U}(h)} F_{\mathrm{p}}, \qquad F_{\mathrm{p}} = F_{12}$$

§

§ **Derive parallel robustness.**

§ **Ramp function:**

$$\overline{r}(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & 1 < x \end{cases}$$

§ **Inverse of robustness:**

$$\mu(h) = \max_{F_i, F_{12} \in \mathcal{U}(h)} F_{\mathrm{p}}, \qquad F_{\mathrm{p}} = F_{12}$$

§ **Uncertainty and robustness:**

$$|F_{12} - \widetilde{F}_{12}| \leq h \quad \Longrightarrow \quad \mu(h) = \overline{r}(\widetilde{F}_{12} + h)$$

$$\mu(h) \leq F_{\mathrm{cp}} \quad \Longrightarrow \quad \widehat{h}_{\mathrm{p}}(F_{\mathrm{cp}}) = F_{\mathrm{cp}} - \widetilde{F}_{12}$$

or zero if negative.

§ **Derive serial robustness.**

§ **Inverse of robustness:**

$$\mu(h) = \max_{F_i, F_{12} \in \mathcal{U}(h)} F_{\rm s}, \qquad F_{\rm s} = F_1 + F_2 - F_{12}$$

§

§ Derive serial robustness.

§ Inverse of robustness:

$$\mu(h) = \max_{F_i, F_{12} \in \mathcal{U}(h)} F_{\rm s}, \qquad F_{\rm s} = F_1 + F_2 - F_{12}$$

§ Uncertainty and inverse robustness:

$$F_i, F_{12} \in \mathcal{U}(h) \quad \Longrightarrow \tag{3}$$
$$\mu(h) = \overline{r}\left[\overline{r}(\widetilde{F}_1 + h) + \overline{r}(\widetilde{F}_2 + h) - \overline{r}(\widetilde{F}_{12} - h)\right]$$

§ Robustness:

- **Plot h vs. $\mu(h)$.**
- **No need to invert $\mu(h)$.**

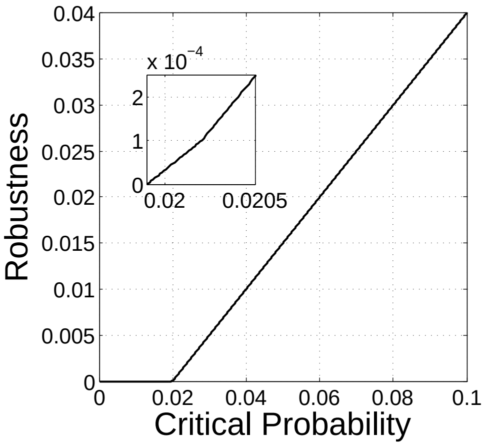


Figure 21: Serial robustness. $\widetilde{F}_i = 0.01$, $\widetilde{F}_{12} = 0.0001$.

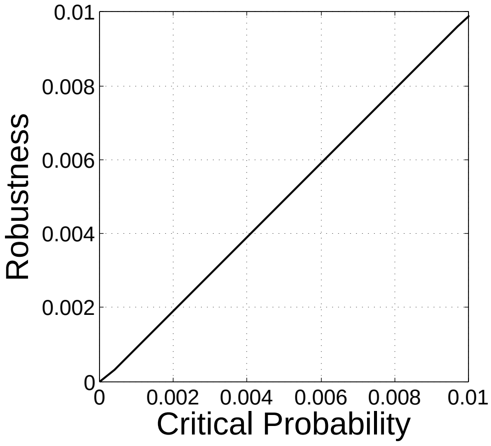


Figure 22: Parallel robustness. $\widetilde{F}_i = 0.01$, $\widetilde{F}_{12} = 0.0001$.

§ Trade-off: Robustness vs. critical prob.

§ Zeroing: No rbs. of est. failure prob.

§ High cost of rbs: large slope.

- **Serial:** $\Delta\widehat{h}/\Delta F_{\text{cs}} = \mathbf{0.5}$.
- **Parallel:** $\Delta\widehat{h}/\Delta F_{\text{cp}} = \mathbf{1}$.

§ Summary of Reliability, Redundancy, Uncer Correlation:

Analysis of uncertainty is essential.

13.3 *Parameter Uncertainty and* FAT *Tails*

§ **Basic idea:**

- **Fat tails:**
  - **Tail decays slower than exponential.**
  - **Not all moments exist.**
- **Thin-tail pdf may have uncert param's.**
- **Total pdf may be fat tailed.**

Example of Fat Tails

§ Exponential distribution of t.

- t is a random variable, e.g. lifetime:

$$f(t|\lambda) = \lambda \mathrm{e}^{-\lambda t}, \quad t \geq 0$$

- **All moments of $t|\lambda$ exist.**

§

Example of Fat Tails

§ Exponential distribution of t.

- t is a random variable, e.g. lifetime:

$$f(t|\lambda) = \lambda e^{-\lambda t}, \quad t \geq 0$$

- All moments of $t|\lambda$ exist.

§ Gamma distribution of λ

- λ is uncertain.

  E.g., mixture of populations.

$$\pi(\lambda) = \frac{\alpha}{\Gamma(k)} (\alpha\lambda)^{k-1} e^{-\alpha\lambda}, \quad \lambda \geq 0, \quad \underbrace{\alpha > 0, \quad k > 0}_{\text{parameters}}$$

- All moments of λ exist.

§ What is marginal distribution of t?

- $t|\lambda$ is **exponential** (thin).
- λ is **Gamma** (thin).
- t is **Pareto** (fat).
-

§ What is marginal distribution of t?

- $t|\lambda$ is **exponential** (thin).
- λ is **Gamma** (thin).
- t is **Pareto** (fat).
- **Define:** $Y = 1 + \frac{T}{\alpha}$
- Y **is Pareto:** $f(y) = k\left(\frac{1}{y}\right)^{k+1}, \quad y \geq 1$
- **Mean and variance:**

$$\mathrm{E}(y) = \frac{k}{k-1} \quad \textbf{if} \quad k > 1$$

$$\mathrm{var}(y) = \frac{k}{k-2} - \left(\frac{k}{k-1}\right)^2 \quad \textbf{if} \quad k > 2$$

- **Not all moments exist.**
- **Parameter uncertainty implies fat tails.**

13.4 *Quantile Uncertainty*

§ **Outcome, reward, r:**

- **Random variable.**
- **Large is better than small.**

§

§ **Outcome, reward, r:**

- **Random variable.**
- **Large is better than small.**

§ **Decision: Choose system: mean, variance.**

§ **Problem: Distribution uncertain.**

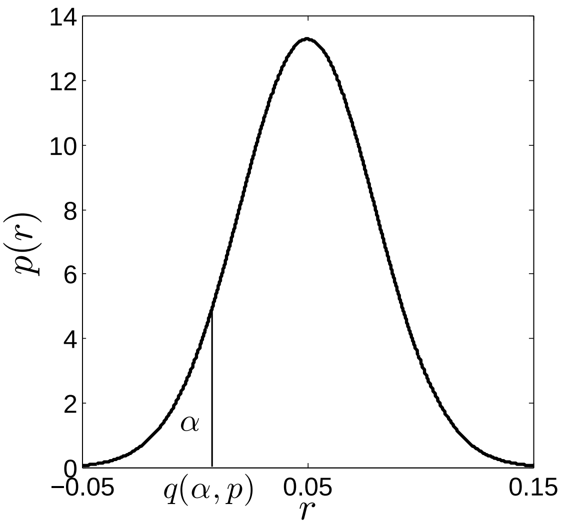


Figure 23: α quantile, $q(\alpha, p)$.

§ α Quantile, $q(\alpha, p)$:

$$\alpha = \int_{-\infty}^{q(\alpha,p)} p(r)\, \mathrm{d}r$$

- **α is probability of failure.**
- **$q(\alpha, p)$ is critical value of r.**

§ **Small quantile (far left) $\Longleftrightarrow$ high risk.**

§ **Performance requirement:**

$$q(\alpha, p) \geq R_{\mathrm{C}}$$

§

§ **Small quantile (far left) $\Longleftrightarrow$ high risk.**

§ **Performance requirement:**

$$q(\alpha, p) \geq R_{\mathrm{C}}$$

§ **Problem:**

- **pdf of r highly uncertain.**
- **May have fat tails.**
- **Hence $q(\alpha, p)$ highly uncertain.**

§ **Question:**

Is system reliable?

§ Info-gap model of uncertainty:

- $\tilde{p}(r)$ **is estimated pdf; e.g. normal.**
- **Envelope-bound uncertainty:**

$$|p(r) - \tilde{p}(r)| \leq g(r)h$$

§

§ Info-gap model of uncertainty:

- $\tilde{p}(r)$ is **estimated pdf;** e.g. normal.
- **Envelope-bound uncertainty:**

$$|p(r) - \tilde{p}(r)| \le g(r)h$$

§ How to choose envelope?

- **Engineering judgment.**
- **Dimensional analysis.**
- **Analogical reasoning.**
- **Assess "equivalent risk".**

§ Info-gap model of uncertainty:

- $\tilde{p}(r)$ is **estimated pdf;** e.g. normal.
- **Envelope-bound uncertainty:**

$$|p(r) - \tilde{p}(r)| \le g(r)h$$

§ Example: $g(r)$ is "$1/r^2$" on tails:

$$g(r) = \begin{cases} \dfrac{\tilde{p}(\mu - r_{\rm s})(\mu - r_{\rm s})^2}{r^2} & \textbf{if } r < \mu - r_{\rm s} \\ \tilde{p}(r) & \textbf{if } |r - \mu| \le r_{\rm s} \\ \dfrac{\tilde{p}(\mu + r_{\rm s})(\mu + r_{\rm s})^2}{r^2} & \textbf{if } r > \mu + r_{\rm s} \end{cases}$$

Mean and variance may be unbounded.

§ Robustness:

Max horizon of uncertainty at which loss is acceptable:

$$\widehat{h}(\alpha, R_{\rm C}) = \max\left\{h :\ \left(\min_{p\in\mathcal{U}(h)} q(\alpha, p)\right) \geq R_{\rm C}\right\}$$

§

§ **Robustness:**

Max horizon of uncertainty at which loss is acceptable:

$$\widehat{h}(\alpha, R_{\mathrm{C}}) = \max\left\{h: \ \left(\min_{p \in \mathcal{U}(h)} q(\alpha, p)\right) \geq R_{\mathrm{C}}\right\}$$

§ **Robustness depends on:**

- **Underlying design, e.g.** μ, σ.
- **Designated failure probability,** α.
- **Critical loss,** R_{C}.

§

§ Robustness:

Max horizon of uncertainty at which loss is acceptable:

$$\widehat{h}(\alpha, R_C) = \max\left\{h : \left(\min_{p\in\mathcal{U}(h)} q(\alpha, p)\right) \geq R_C\right\}$$

§ Robustness depends on:

- **Underlying design, e.g.** μ, σ.
- **Designated failure probability,** α.
- **Critical loss,** R_C.

§ Robustness is a decision function:

- **Satisfy requirements.**
- **Maximize robustness.**

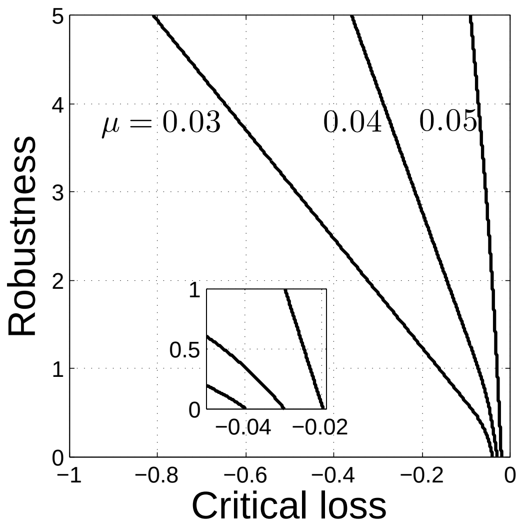


Figure 24: Robustness, $\widehat{h}$, vs. critical loss, R_{C}, for 3 values of μ. $\sigma = 0.03$.

§ **Trade-off:** robustness vs critical loss.

§ **Zeroing:** No rbs of estim'd critical loss.

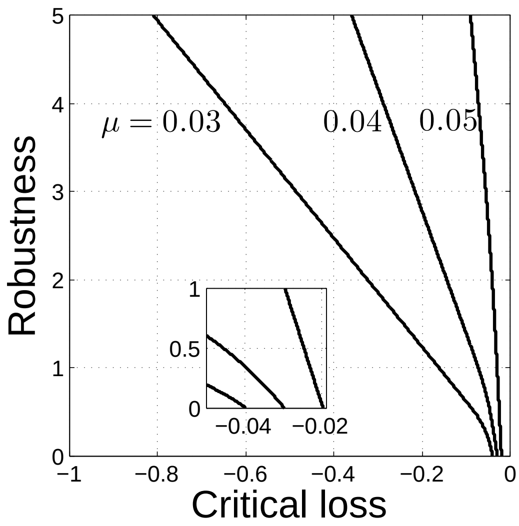


Figure 25: Robustness, $\widehat{h}$, vs. critical loss, R_C, for 3 values of μ. $\sigma = 0.03$.

§ Effects of increasing the mean:

- **Shift to higher robustness.**
- **Increase slope: reduce cost of rbs.**

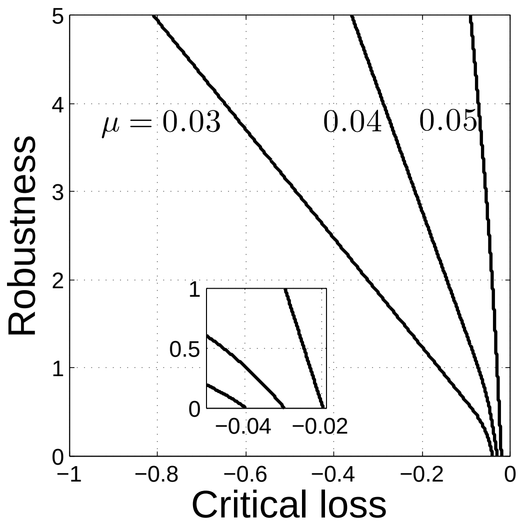


Figure 26: Robustness, $\widehat{h}$, vs. critical loss, R_{C}, for 3 values of μ. $\sigma = 0.03$.

§ Calibrating the robustness.

- **Is robustness of 3 or 5 "large"?**
- $\widehat{h} = 3$ **means 3** $\times$ **tail tolerable.**
- **Standard normal below** -2σ**: 0.0228.**
- **Fat tail below** -2σ**: 0.1080.**

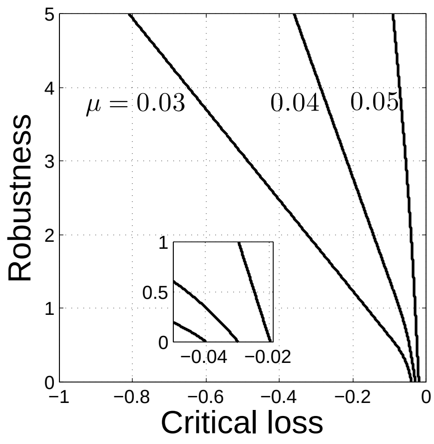


Figure 27: Robustness, $\widehat{h}$, vs. critical loss, R_C, for 3 values of μ. $\sigma = 0.03$.

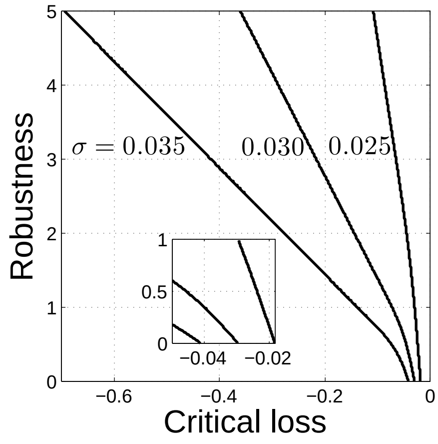


Figure 28: Robustness, $\widehat{h}$, vs. critical loss, R_C, for 3 values of σ. $\mu = 0.04$.

§ Mean-variance trade-off:

Increasing mean by 0.01

roughly equivalent to

decreasing st. dev by 0.005.

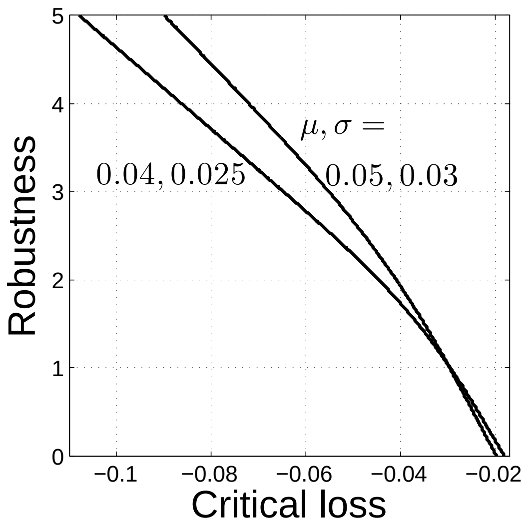


Figure 29: Robustness, $\widehat{h}$, vs. critical loss, $R_{\rm C}$, for two different combinations of μ and σ.

§ Preference reversal:

- $(\mu, \sigma) = (0.04, 0.025) \succ_{\mathbf{est.}} (0.05, 0.03)$.
-

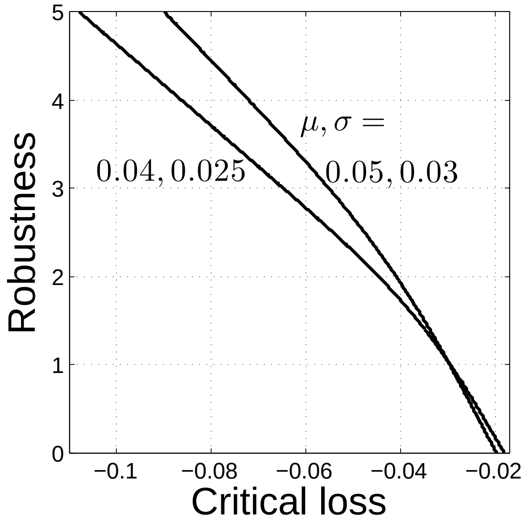


Figure 30: Robustness, $\widehat{h}$, vs. critical loss, $R_{\rm C}$, for two different combinations of μ and σ.

§ Preference reversal:

- $(\mu, \sigma) = (0.04, 0.025) \succ_{\textbf{est.}} (0.05, 0.03)$.
- $(\mu, \sigma) = (0.05, 0.03) \succ_{\textbf{rbs.}} (0.04, 0.025)$ **at** $\widehat{h} > 1$.

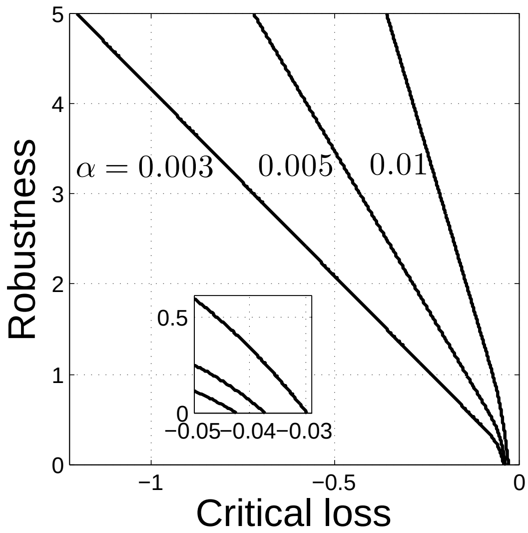


Figure 31: Robustness, $\widehat{h}$, vs. critical loss, $R_{\rm C}$, for three different probabilities of failure, α. $\mu = 0.04$ and $\sigma = 0.03$.

§ Effect of greater prob of failure, α:

- **Greater robustness** (curves shift right).
- **Lower cost of robustness** (steeper curves).
- **Trade-offs:** μ, σ and α.

§ Less severe uncertainty:

- **Thin tails.**
- **Known pdf family.**
- **Uncertain moments of normal pdf:**

$$\mathcal{U}(h) = \left\{ p(r) \sim N(\mu, \sigma^2) : \ \left| \frac{\mu - \widetilde{\mu}}{\varepsilon_\mu} \right| \le h, \ \left| \frac{\sigma - \widetilde{\sigma}}{\varepsilon_\sigma} \right| \le h, \ \sigma \ge 0 \right\}$$

$$h \ge 0$$

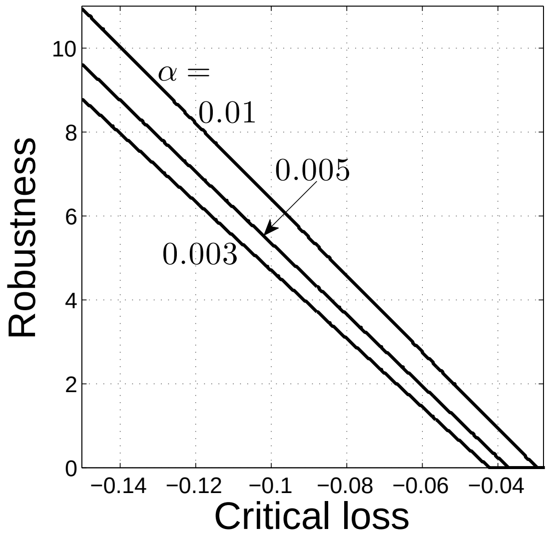


Figure 32: Robustness, $\widehat{h}$, vs. critical loss, R_{C}, for three different probabilities of failure, α. $\widetilde{\mu} = 0.04$, $\widetilde{\sigma} = 0.03$, $\varepsilon_{\mu} = 0.004$, $\varepsilon_{\sigma} = 0.003$.

§ **Trade-off: Robustness vs. critical loss.**

§ **Zeroing: No rbs. of est. critical loss.**

§ **Robustness: Greater than for fat tails.**

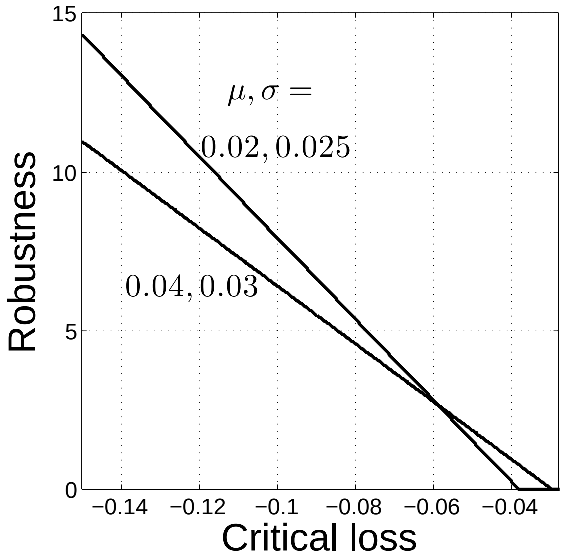


Figure 33: Robustness, $\widehat{h}$, vs. critical loss, R_{C}, for two different combinations of μ and σ. $\alpha = 0.01$.

§ Preference reversal:

- $(\mu, \sigma) = (0.04, 0.03) \ \succ_{\mathbf{est.}} \ (0.02, 0.025)$.
- $(\mu, \sigma) = (0.02, 0.025) \succ_{\mathbf{rbs.}} (0.04, 0.03)$

  at $\widehat{h} > 2$.

§ Summary:

- **Uncertain pdf: fat tails.**
- **Nominal estimates: zero robustness.**
- **Trade-offs:**
  - **Robustness vs performance.**
  - **Mean vs variance.**
  - **Mean vs failure probability.**
- **Cost of robustness.**
- **Preference reversal between options.**
- **Uncertainty judgment, info-gap model.**

14 Statistical Tests with Distributional Uncertainty

[3] lectures\talks\lib\stat-tests02.tex 27.2.2012

14.1 Distributional Uncertainty

§ **Uncertainty, two foci:**

- **Randomness: structured uncertainty.**
- **Info-gaps: Surprise, ignorance, indeterminism.**

§ Distributional Uncertainty:

Unknown sampling distribution due to:

- **Non-independence of observations.**

  E.g. unknown causal pathways.

-

§ Distributional Uncertainty:

Unknown sampling distribution due to:

- **Non-independence of observations.**

  E.g. unknown causal pathways.

- **Non-stationarity of population.**

  E.g. unknown evolution over time.

-

§ Distributional Uncertainty:

Unknown sampling distribution due to:

- **Non-independence of observations.**
  E.g. unknown causal pathways.
- **Non-stationarity of population.**
  E.g. unknown evolution over time.
- **Variability of observer.**
  E.g. professional/non-professional.
-

§ **Distributional Uncertainty:**

Unknown sampling distribution due to:

- **Non-independence of observations.**

  E.g. unknown causal pathways.

- **Non-stationarity of population.**

  E.g. unknown evolution over time.

- **Variability of observer.**

  E.g. professional/non-professional.

- **Non-asymptotic data.**

§

§ Distributional Uncertainty:

Unknown sampling distribution due to:

- **Non-independence of observations.**
  E.g. unknown causal pathways.
- **Non-stationarity of population.**
  E.g. unknown evolution over time.
- **Variability of observer.**
  E.g. professional/non-professional.
- **Non-asymptotic data.**

§ The challenge:

Design statistical test of hypothesis.

§ Example: Chronic Wasting Disease.

- **Given n nulls at t, test no-disease hypo.**
-

§ Example: Chronic Wasting Disease.

- **Given n nulls at t, test no-disease hypo.**
- **Antler extract from diseased deer induces disease in mice.**
- **Time to expression in mice:**

Diseased Animal	Time to Expression
1	442 ± 16 (6/8)
2	> 594 (0/5)
3	463 ± 23 (2/3)
4	> 601 (0/6)

Table 1: Mice expression of deer prion protein from antler velvet of diseased animals. Angers *et al.*, 2009.

  - **Uncertain pdf.**
  - **Unstable sample moments?**

§ Example: Long-term bio-monitoring.

- **Given 200 y's of data test no-change hypo.**
- **Data:**
  - **Naturalists' logs.**
  - **Amateurs' diaries.**
  - **Museum collections.**
- **Uncertainty:**
  - **Museum policy changes over time.**
  - **Observers' habits are variable.**
  - **Variable observers: pros, amateurs.**
  - **Protocol and purpose of observation.**

§ Example: Detect invasive species.

- **Decisions:**
  - Choose traps and deployment.
  - Allocate resources:
    - Professional vs non-professional.
    - Detection vs irradication.
  - Port-of-entry search strategy.
  - Interpret finds (e.g. nulls).
- **Uncertainties:**
  - Transport mechanisms.
  - Entry mechanisms.
  - Habitat suitability.

§ Sources:

• Lior Davidovitch, Richard Stoklosa, Jonathan Majer, Alex Nietrzeba, Peter Whittle, Kerrie Mengersen and Yakov Ben-Haim, 2009, Info-Gap theory and robust design of surveillance for invasive species: The case study of Barrow Island, *Journal of Environmental Management,* vol.90, #8, pp.2785–2793.

• Yemshanov, Denys, Frank H. Koch, Yakov Ben-Haim and William D. Smith, 2010, Robustness of risk maps and survey networks to knowledge gaps about a new invasive pest, *Risk Analysis: An International Journal,* vol.30, #2, pp.261–276.

14.2 Some Info-Gap Models of Distributional Uncertainty

- **Really severe distributional uncertainty:**

  ◦ **Unbounded moments, fat tails, multi-modal, atoms.**

  ◦ **Uniform-bound in the cdf:**

$$\mathcal{U}(h) = \left\{ F(y) : F(y) \in \mathcal{P}, \ |F(y) - \widetilde{F}_i(y)| \le h, \right\}, \ h \ge 0$$

- **Severe distributional uncertainty:**
  - **Unbounded moments, fat tails, multi-modal, no atoms.**
  - **Uniform-bound in pdf:**

$$\mathcal{U}(h) = \{\, f(y) :\ f(y) \in \mathcal{D},\ |f(y) - \widetilde{f}_i(y)| \leq h f_i^{\star}, \}\,, \quad h \geq 0$$

-

- **Severe distributional uncertainty:**
  - **Unbounded moments, fat tails, multi-modal, no atoms.**
  - **Uniform-bound in pdf:**

$$\mathcal{U}(h) = \left\{ f(y) : \ f(y) \in \mathcal{D}, \ |f(y) - \widetilde{f}_i(y)| \le h f_i^{\star}, \right\}, \quad h \ge 0$$

- **Moderate distributional uncertainty:**
  - **Bounded moments, ordinary tails, multi-modal, atoms.**
  - **Envelope-bound in cdf:**

$$\mathcal{U}(h) = \left\{ F(y) : \ F(y) \in \mathcal{P}, \ |F(y) - \widetilde{F}_i(y)| \le h\psi(y), \right\}, \ h \ge 0$$

- **Light distributional uncertainty:**
  - **Bounded moments, ordinary tails, multi-modal, no atoms.**
  - **Fractional-bound in pdf:**

$$\mathcal{U}(h) = \left\{ f(y) : \ f(y) \in \mathcal{D}, \ \left| f(y) - \widetilde{f}_i(y) \right| \le h \widetilde{f}_i(y), \right\}, \ h \ge 0$$

§

- **Light distributional uncertainty:**
  - **Bounded moments, ordinary tails, multi-modal, no atoms.**
  - **Fractional-bound in pdf:**

$$\mathcal{U}(h) = \left\{ f(y) : \ f(y) \in \mathcal{D}, \ |f(y) - \widetilde{f}_i(y)| \le h\widetilde{f}_i(y), \right\}, \ \ h \ge 0$$

§ Choosing an info-gap model:

- **Unbounded family of nested sets:**
  - **Contraction:** $\mathcal{U}(0) = \{\widetilde{F}_i\}$.
  - **Nesting:** $h < h'$ **implies** $\mathcal{U}_i(h) \subseteq \mathcal{U}(h')$.
  - $h =$ **unknown horizon of uncertainty.**
- **Tightest family consistent with info.**
- **Judgment.**

14.3 Test of False Nulls

§ Example: Chronic Wasting Disease.

- Antler extract from diseased deer induces disease in mice.
- Time to expression: uncertain pdf.

§ Sources:

• Ben-Haim, Yakov, 2009, Tests of the Mean with Distributional Uncertainty: An Info-Gap Approach, ISIPTA, Durham, UK.

• Yakov Ben-Haim, Interpreting Null Results from Measurements with Uncertain Correlations: An Info-Gap Approach, *Risk Analysis,* vol.31 (1), pp.78–85

• L. Joe Moffitt and Yakov Ben-Haim, Robustness Analysis of Expert Dispute About Incubation Time, working paper.

§ **Question:**

- n **innoculated mice.**
- **No PrP expression after incubation times** $t_1, \ldots, t_n$.
- **How confident that CWD is not present?**

§

§ **Question:**

- n **innoculated mice.**
- **No PrP expression after incubation times** $t_1, \ldots, t_n$.
- **How confident that CWD is not present?**

§ **System model: probability of false null:**

$$P_{\text{fn}}(t_1, \ldots, t_n) = \prod_{i=1}^{n} [1 - P(t_i)]$$

$P(t) =$ **prob. of expression by time** t.

§

§ **Question:**

- n **innoculated mice.**
- **No PrP expression after incubation times** $t_1, \ldots, t_n$.
- **How confident that CWD is not present?**

§ **System model:** **probability of false null:**

$$P_{\text{fn}}(t_1, \ldots, t_n) = \prod_{i=1}^{n} [1 - P(t_i)]$$

$P(t)$ = **prob. of expression by time** t.

§ **Problem:** $P(t)$ **highly uncertain.**

§ Uncertainty model.

Uncertain pdf of time to expression:

Diseased Animal	Time to Expression
1	442 ± 16 (6/8)
2	> 594 (0/5)
3	463 ± 23 (2/3)
4	> 601 (0/6)

Table 2: Mice expression of deer prion protein from antler velvet of diseased animals. Angers *et al.*, 2009.

- **No moments?**
- **Fat tails.**

$$\mathcal{U}(h) = \left\{ p : p \in \mathcal{P}, p(t) \le \tilde{p}(t) + \frac{t_{\text{s}} h}{t^2} \forall t \ge t_{\text{s}} \right\}, \quad h \ge 0$$

§ Robustness.

Maximum tolerable uncertainty:

$$\widehat{h}(n, P_{\text{fnc}}) = \max\left\{h : \ \left(\max_{p \in \mathcal{U}(h)} P_{\text{fn}}\right) \leq P_{\text{fnc}}\right\}$$

§ Deriving the robustness.

• $m(h)$ = **inner maximum in eq.(14.3); inverse of** $\widehat{h}(n, P_{\rm fnc})$**.**

•

§ Deriving the robustness.

- $m(h)$ = **inner maximum in eq.(14.3); inverse of** $\widehat{h}(n, P_{\rm fnc})$**.**

- **Plot of** $m(h)$ **vs.** h **equivalent to plot of** $P_{\rm fnc}$ **vs.** $\widehat{h}(n, P_{\rm fnc})$**.**

-

§ Deriving the robustness.

- $m(h)$ = **inner maximum in eq.(14.3); inverse of** $\widehat{h}(n, P_{\rm fnc})$**.**
- **Plot of** $m(h)$ **vs.** h **equivalent to plot of** $P_{\rm fnc}$ **vs.** $\widehat{h}(n, P_{\rm fnc})$**.**
- **Assume** $t_i > t_{\rm s}$ **for all** i**.**
-

§ Deriving the robustness.

- $m(h)$ = **inner maximum in eq.(14.3); inverse of** $\widehat{h}(n, P_{\rm fnc})$.

- **Plot of** $m(h)$ **vs.** h **equivalent to plot of** $P_{\rm fnc}$ **vs.** $\widehat{h}(n, P_{\rm fnc})$.

- **Assume** $t_i > t_{\rm s}$ **for all** i.

- **Thus** $m(h)$ **evaluated at upper envelope at horizon of uncertainty** h**, provided that this distribution can be normalized.**

-

§ Deriving the robustness.

- $m(h)$ = **inner maximum in eq.(14.3); inverse of** $\widehat{h}(n, P_{\rm fnc})$**.**

- **Plot of** $m(h)$ **vs.** h **equivalent to plot of** $P_{\rm fnc}$ **vs.** $\widehat{h}(n, P_{\rm fnc})$**.**

- **Assume** $t_i > t_{\rm s}$ **for all** i**.**

- **Thus** $m(h)$ **evaluated at upper envelope at horizon of uncertainty** h**, provided that this distribution can be normalized.**

- **For each individual observation:**

$$\max_{p \in \mathcal{U}(h)} [1 - P(t_i)] = \min \left[1, \ \int_{t_i}^{\infty} \left(\tilde{p}(t) + \frac{t_{\rm s} h}{t^2} \right) \, {\rm d}t \right] \quad (4)$$

$$= \min \left[1, \ 1 - \widetilde{P}(t_i) + \frac{t_{\rm s} h}{t_i} \right] \quad (5)$$

- **For each individual observation:**

$$\max_{p\in\mathcal{U}(h)} [1 - P(t_i)] \;=\; \min\left[1,\;\, 1 - \widetilde{P}(t_i) + \frac{t_{\rm s}h}{t_i}\right] \quad (6)$$

- n **observations independent so:**

$$m(h) = \prod_{i=1}^{n} \min\left[1,\;\, 1 - \widetilde{P}(t_i) + \frac{t_{\rm s}h}{t_i}\right]$$

-

- **For each individual observation:**

$$\max_{p\in\mathcal{U}(h)} [1 - P(t_i)] \;=\; \min\left[1,\;\; 1 - \widetilde{P}(t_i) + \frac{t_{\rm s} h}{t_i}\right] \quad (7)$$

- n **observations independent so:**

$$m(h) = \prod_{i=1}^{n} \min\left[1,\;\; 1 - \widetilde{P}(t_i) + \frac{t_{\rm s} h}{t_i}\right]$$

- t_i **large so** $1 - \widetilde{P}(t_i) \approx 0$**. Then, for** $h \le 1$**:**

$$m(h) \approx \frac{t_{\rm s}^n h^n}{\prod_{i=1}^{n} t_i}$$

-

- **For each individual observation:**

$$\max_{p \in \mathcal{U}(h)} [1 - P(t_i)] \;=\; \min \left[1,\ 1 - \widetilde{P}(t_i) + \frac{t_{\rm s} h}{t_i} \right] \quad (8)$$

- n **observations independent so:**

$$m(h) = \prod_{i=1}^{n} \min \left[1,\ 1 - \widetilde{P}(t_i) + \frac{t_{\rm s} h}{t_i} \right]$$

- t_i **large so** $1 - \widetilde{P}(t_i) \approx 0$**. Then, for** $h \le 1$**:**

$$m(h) \approx \frac{t_{\rm s}^n h^n}{\prod_{i=1}^{n} t_i}$$

- **Equate to** $P_{\rm fnc}$**, solve for** h**:**

$$\widehat{h}(n, P_{\rm fnc}) \approx \frac{1}{t_{\rm s}} \left(P_{\rm fnc} \prod_{i=1}^{n} t_i \right)^{1/n}$$

-

- **For each individual observation:**

$$\max_{p \in \mathcal{U}(h)} [1 - P(t_i)] \ = \ \min \left[1, \ 1 - \widetilde{P}(t_i) + \frac{t_\mathrm{s} h}{t_i} \right] \quad (9)$$

- n **observations independent so:**

$$m(h) = \prod_{i=1}^{n} \min \left[1, \ 1 - \widetilde{P}(t_i) + \frac{t_\mathrm{s} h}{t_i} \right]$$

- t_i **large so** $1 - \widetilde{P}(t_i) \approx 0$**. Then, for** $h \le 1$**:**

$$m(h) \approx \frac{t_\mathrm{s}^n h^n}{\prod_{i=1}^n t_i}$$

- **Equate to** P_fnc**, solve for** h**:**

$$\widehat{h}(n, P_\mathrm{fnc}) \approx \frac{1}{t_\mathrm{s}} \left(P_\mathrm{fnc} \prod_{i=1}^{n} t_i \right)^{1/n}$$

- $\overline{t}_\mathrm{gm}$ = **geometric mean of** $t_1, \ldots, t_n$**, so:**

$$\widehat{h}(n, P_\mathrm{fnc}) \approx \frac{\overline{t}_\mathrm{gm}}{t_\mathrm{s}} P_\mathrm{fnc}^{1/n}$$

§ Results:

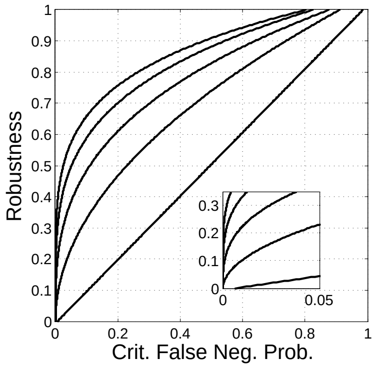


Figure 34: $\widehat{h}(n, P_{\text{fnc}})$ vs P_{fnc}, $n = 1$ to 5 (bottom to top).

- **Data:** $t_i =$ **500, 530, 510, 520, 505 days.**
- **Estimated distribution** $\mathcal{N}(450, 400)$**,** $t_{\text{s}} = 490$**.**
- **Trade-off: rbs vs false-negative prob.**
- **Zeroing: no rbs of estim false-neg prob.**

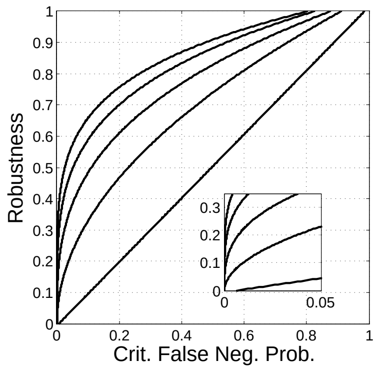


Figure 35: $\widehat{h}(n, P_{\rm fnc})$ vs $P_{\rm fnc}$, $n = 1$ to 5 (bottom to top).

$$\widehat{h}(n, P_{\rm fnc}) \approx \frac{\overline{t}_{\rm gm}}{t_{\rm s}} P_{\rm fnc}^{1/n}$$

§ Sample size effect:

- $\overline{t}_{\rm gm}$ **fluctuates with sample size.**
- $P_{\rm fnc}^{1/n}$ **is dominant:**

  grows fast w/ n **if** n **and** $P_{\rm fnc}$ **small.**

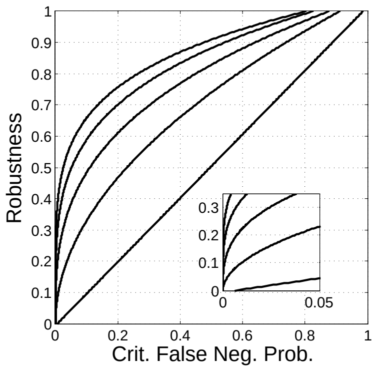


Figure 36: $\widehat{h}(n, P_{\rm fnc})$ vs $P_{\rm fnc}$, $n = 1$ to 5 (bottom to top).

$$\widehat{h}(n, P_{\rm fnc}) \approx \frac{\overline{t}_{\rm gm}}{t_{\rm s}} P_{\rm fnc}^{1/n}$$

§ Cost of robustness:

- **Slope of $\widehat{h}$ vs $P_{\rm fnc}$: cost of robustness.**
- **When $P_{\rm fnc}$ is very small,**
  slope increases as n increases.
- **Thus cost of robustness,**
  in units of increased $P_{\rm fnc}$,
  is small when n is large.

14.4 Statistical Test of the Mean with Distributional Uncertainty

§ Source: Yakov Ben-Haim, 2009, Tests of the Mean with Distributional Uncertainty: An Info-Gap Approach, ISIPTA, Durham, UK.

§ **Sample:** $X = \{x_1, \ldots, x_n\}$.

Not necessarily random.

§

§ **Sample:** $X = \{x_1, \ldots, x_n\}$.

Not necessarily random.

§ **Two simple hypotheses:**

$$H_0 : \quad \mu = T_0$$

$$H_1 : \quad \mu = T_1$$

T_i **specified,** $T_1 > T_0$.

§

§ **Sample:** $X = \{x_1, \ldots, x_n\}$.

Not necessarily random.

§ **Two simple hypotheses:**

$$H_0: \quad \mu = T_0$$

$$H_1: \quad \mu = T_1$$

T_i **specified,** $T_1 > T_0$.

§ **Errors:**

- **Type I: falsely reject** H_0.
- **Type II: falsely accept** H_0.

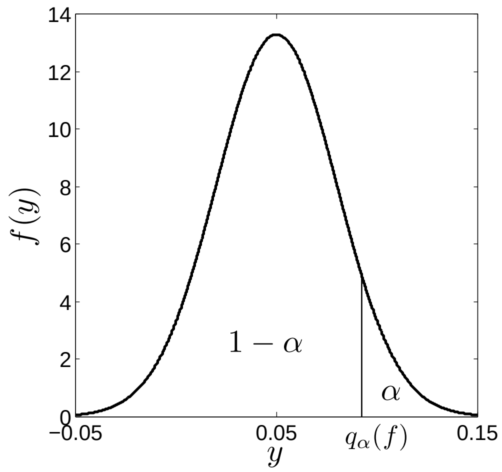


Figure 37: α quantile, $q_\alpha(f)$.

§ **Statistic:** $y(X)$. **E.g.** $y(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$.

§ **$(1-\alpha)$ quantile of $F(y)$:**

$$1-\alpha = \int_{-\infty}^{q_\alpha(F)} f(y)\,\mathrm{d}y$$

Or:

$$q_\alpha(F) = \inf\,\{y: \ F(y) \ge 1-\alpha\}$$

§ Estimated distributions:

- $\widetilde{F}_0(y)$ **if** H_0 **true.**
- $\widetilde{F}_1(y)$ **if** H_1 **true.**

§ Nominal threshold tests, using $\widetilde{F}_0$ and $\widetilde{F}_1$:

- **Test of nominal size $\alpha^\star$ rejects H_0 when:**

$$y \geq q_{\alpha^\star}(\widetilde{F}_0)$$

Falsely rejects H_0 with prob $\alpha^\star$.

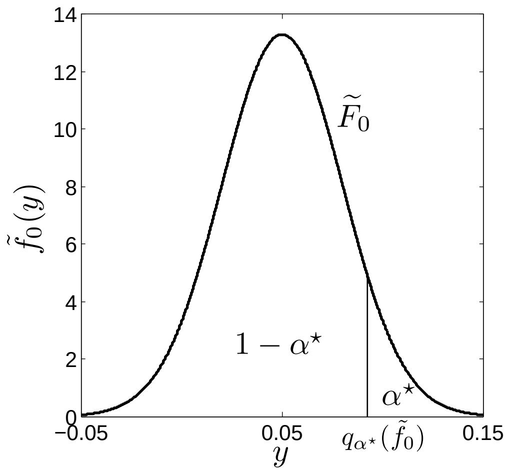


Figure 38: Nominal distribution for H_0, $\widetilde{F}_0(y)$.

§ Nominal threshold tests, using $\widetilde{F}_0$ and $\widetilde{F}_1$:

- **Test of nominal size $\alpha^\star$ rejects H_0 when:**

$$y \geq q_{\alpha^\star}(\widetilde{F}_0)$$

Falsely rejects H_0 with prob $\alpha^\star$.

- **Test of nominal power $1-\beta^\star$**

correctly rejects H_0 with prob $1-\beta^\star$:

$$1-\beta^\star = 1-\widetilde{F}_1[q_{\alpha^\star}(\widetilde{F}_0)]$$

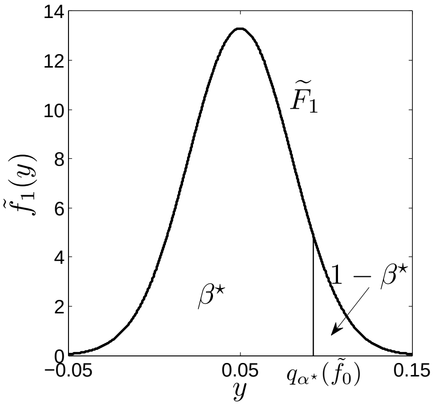


Figure 39: Nominal distribution for H_1, $\widetilde{F}_1(y)$.

§ Nominal threshold tests, using $\widetilde{F}_0$ and $\widetilde{F}_1$:

- **Test of nominal size** $\alpha^\star$ **rejects** H_0 **when:**

$$y \geq q_{\alpha^\star}(\widetilde{F}_0)$$

Falsely rejects H_0 **with prob** $\alpha^\star$**.**

- **Test of nominal power** $1 - \beta^\star$

correctly rejects H_0 **with prob** $1 - \beta^\star$**:**

$$1 - \beta^\star = 1 - \widetilde{F}_1[q_{\alpha^\star}(\widetilde{F}_0)]$$

- $\alpha^\star$ **small: low prob of type I error.**
- $1 - \beta^\star$ **large: low prob of type II error.**

§ Nominal threshold tests, using $\widetilde{F}_0$ and $\widetilde{F}_1$:

- **Test of nominal size $\alpha^\star$ rejects H_0 when:**

$$y \geq q_{\alpha^\star}(\widetilde{F}_0)$$

Falsely rejects H_0 with prob $\alpha^\star$.

- **Test of nominal power $1-\beta^\star$**

correctly rejects H_0 with prob $1-\beta^\star$:

$$1-\beta^\star = 1-\widetilde{F}_1[q_{\alpha^\star}(\widetilde{F}_0)]$$

- **$\alpha^\star$ small: low prob of type I error.**
- **$1-\beta^\star$ large: low prob of type II error.**

§ Problem with nominal tests:

$\widetilde{F}_0(y)$ and $\widetilde{F}_1(y)$ highly uncertain.

§ **Info-gap models for $F_0(y)$ and $F_1(y)$:**

$\mathcal{U}_0(h)$, $\mathcal{U}_1(h)$.

§ Robustness of type I error:

maximum h at which

test at nominal size $\alpha^\star$

falsely rejects H_0 with probability $\leq \alpha$:

$$\widehat{h}_0(\alpha^\star, \alpha) = \max\left\{h : \left(\min_{F \in \mathcal{U}_0(h)} F[q_{\alpha^\star}(\widetilde{F}_0)]\right) \geq 1 - \alpha\right\}$$

- $\alpha^\star$**: nominal size.**
- α**: effective size.**

§ Robustness of type II error:

maximum h at which

test at nominal size $\alpha^\star$

falsely accepts H_0 with probability $\le \beta$:

$$\widehat{h}_1(\alpha^\star, \beta) = \max\left\{h : \ \left(\max_{F \in \mathcal{U}_1(h)} F[q_{\alpha^\star}(\widetilde{F}_0)]\right) \le \beta\right\}$$

- $\alpha^\star$: **nominal size.**
- $1 - \beta$: **effective power.**

$\alpha^\star = 0.01$		$\alpha^\star = 0.05$	
n	$1-\beta^\star$	n	$1-\beta^\star$
5	0.10	5	0.54
7	0.32	7	0.75
31	0.998	31	0.9997

Table 3: Size and power in the absence of distributional uncertainty.

§ Size–power trade-off.

No distributional uncertainty.

$\alpha^\star = 0.01$		$\alpha^\star = 0.05$	
n	$1-\beta^\star$	n	$1-\beta^\star$
5	0.10	3	0.18
7	0.32	4	0.37
9	0.54	5	0.54
12	0.76	7	0.75
31	0.998	31	0.9997

Table 4: Size and power in the absence of distributional uncertainty.

§ Effect of sample size.

No distributional uncertainty.

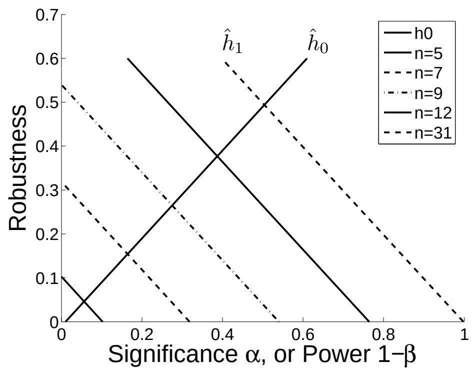


Figure 40: Robustness curves for the t test, $\widehat{h}_0(\alpha^\star, \alpha)$ for falsely rejecting H_0, and $\widehat{h}_1(\alpha^\star, \alpha)$ for falsely rejecting H_1. Nominal size is $\alpha^\star = 0.01$. $\widehat{h}_1(\alpha^\star, \alpha)$ calculated at 5 different sample sizes: $n = 5$, 7, 9, 12 and 31. $\delta = 1$.

§ Robustness curves:

- **Trade offs:**
  - **Positive slope of $\widehat{h}_0$:**

    Robustness trades off with significance.
  -

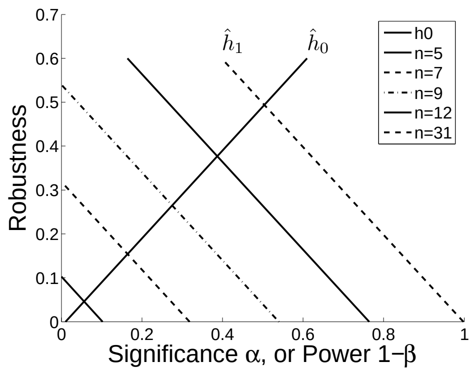


Figure 41: Robustness curves for the t test, $\widehat{h}_0(\alpha^\star, \alpha)$ for falsely rejecting H_0, and $\widehat{h}_1(\alpha^\star, \alpha)$ for falsely rejecting H_1. Nominal size is $\alpha^\star = 0.01$. $\widehat{h}_1(\alpha^\star, \alpha)$ calculated at 5 different sample sizes: $n = 5$, 7, 9, 12 and 31. $\delta = 1$.

§ Robustness curves:

- **Trade offs:**
  - **Positive slope of $\widehat{h}_0$:**

    Robustness trades off with significance.
  - **Negative slope of $\widehat{h}_1$:**

    Robustness trades off with power.

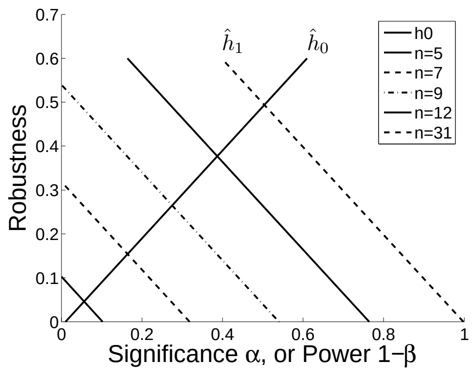


Figure 42: Robustness curves for the t test, $\widehat{h}_0(\alpha^\star, \alpha)$ for falsely rejecting H_0, and $\widehat{h}_1(\alpha^\star, \alpha)$ for falsely rejecting H_1. Nominal size is $\alpha^\star = 0.01$. $\widehat{h}_1(\alpha^\star, \alpha)$ calculated at 5 different sample sizes: $n = 5$, 7, 9, 12 and 31. $\delta = 1$.

§ Robustness curves:

- **Zeroing:**

  Estimated significance or power has no robustness to distributional uncertainty.

§ **Decisions and judgments.**

- **Two decisions** to

  determine the decision threshold $q_{\alpha^\star}(\widetilde{F}_0)$**:**

  - **Nominal test size** $\alpha^\star$**.**
  - **Sample size** n**.**

- **Two judgments:**

  - **Effective size** α**.**
  - **Effective power** $1-\beta$

  $\alpha =$ **prob of falsely rejecting** H_0**.**

  $1-\beta =$ **prob of correctly rejecting** H_0**.**

§ **Use robustness functions** $\widehat{h}_0(\alpha^\star, \alpha)$ **and** $\widehat{h}_1(\alpha^\star, \beta)$**.**

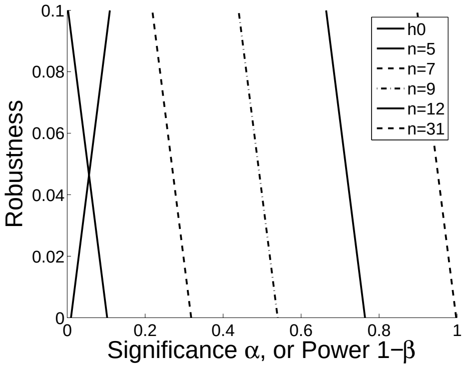


Figure 43: Expanded from fig. 40.

§ Nominal and effective size:

- **Decide: nominal size** $\alpha^\star = 0.01$.
- **Judge: effective size** $\alpha = 0.05$ **is adequate:**
  $\widehat{h}_0(0.01, 0.05) = 0.04$.

§ Effective power and sample size:

- **Apply this rbs to type II error:**
  $\widehat{h}_1(\alpha^\star, \beta) = 0.04$.
- **Judge effective power = 0.72 or 0.96.**
- **Decide: sample size = 12 or 31.**

14.5 Recap

- **Distributional uncertainty.**
- **Statistical tests.**
- **Decisions and judgments.**
- **Robust-satisficing:**
  - **Satisfice effective size and power,**
  - **Maximize robustness.**

15 *SUMMARY*

§ **Models:**

Attributes of model correspond to attributes of reality.

§ **Model-based decision:**

Adapt decision to attributes of model.

§ **Optimization:**

Use best model to
choose decision with best outcome.

§ Uncertainty:

- **Randomness:** **structured uncertainty.**
- **Info-gaps:** **surprises, ignorance.**

§ Fallacy of optimal model-based decision:

- **Severe uncertainty:**
  - **Best model errs seriously.**
  - **Some model attributes are correct.**
  - **Some model attributes err greatly.**
- **Best-model optimization**
  - **exploits all model attributes to extreme.**
  - **Vulnerable to model error.**

§ Resolution: Info-gap decision theory.

- **Satisfice** performance.
- **Optimize** robustness to uncertainty.
- Model and manage sur$^{\text{p}}$rises.

§ Robust-satisficing syllogism:

- Adequate performance is necessary.
- More reliable adequate performance is better than less reliable adeq. perf.
- Thus maximum reliability is best.

§ Proxy theorems:

max robustness ≡ max survival prob.

§ **Sources: http://info-gap.com**

§ **Applications of info-gap theory:**

- **Engineering design.**
- **Fault detection and diagnosis.**
- **Project management.**
- **Homeland security.**
- **Sampling, assay design.**
- **Statistical hypothesis testing.**
- **Monetary economics.**
- **Financial stability.**
- **Biological conservation.**
- **Medical decision making.**