Lecture 5

Linear Regression and Forecasting with Info-Gaps

Yakov Ben-Haim

Technion

Israel Institute of Technology

[0]lectures\talks\lib\paris2023Lec05-002.tex 21.5.2023

Contents

1 *Linear Regression*

1.1 *Preliminary Discussion*

[0] lectures\talks\lib\regression001.tex 21.5.2023 Derived from \lectures\info-gap-methods\lectures\estim03.tex, section 1.

§ Modeling is a decision problem. 2 examples:

- **Modeling WLAN[1] client tracking and prediction.**
- **Modeling population behavior (Phillips curve[2]).**

[1]Wireless Local Area Network.
[2]Source: Yakov Ben-Haim, 2010, *Info-Gap Economics: An Operational Introduction,* Palgrave-Macmillan.

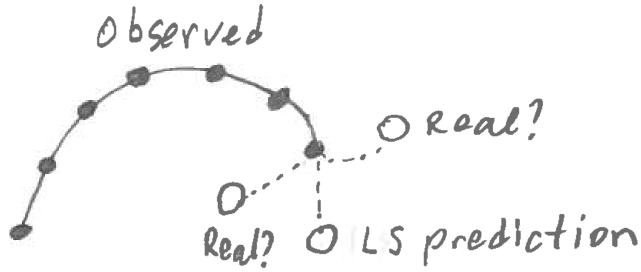


Figure 1: WLAN client motion.

§ WLAN client tracking and prediction:

§ Challenge: Two foci of uncertainty:

-

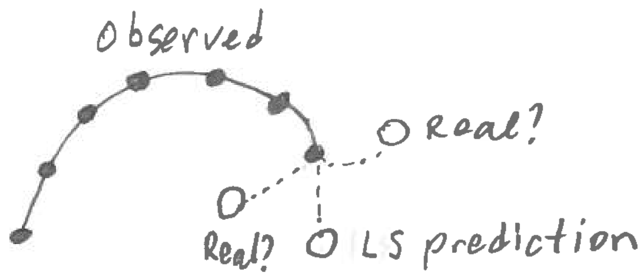


Figure 2: WLAN client motion.

§ **WLAN client tracking and prediction:**

§ **Challenge: Two foci of uncertainty:**

- **Randomness:**
  - **Noisy data (statistics).**
-

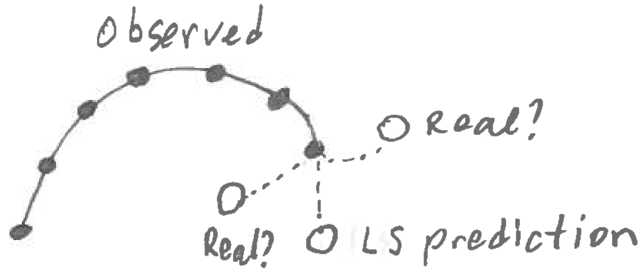


Figure 3: WLAN client motion.

§ **WLAN client tracking and prediction:**

§ **Challenge: Two foci of uncertainty:**

- **Randomness:**
  - **Noisy data (statistics).**
- **Info-gaps:**
  - **Changing plans and intentions of client.**
  - **Interaction with other people.**
  - **Environmental variability.**

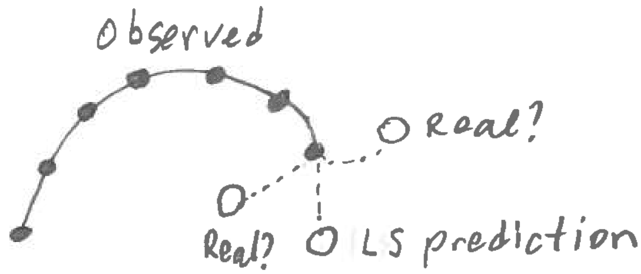


Figure 4: WLAN client motion.

§ Questions:

- **How to use empirical data to model uncertain past motion?**
-

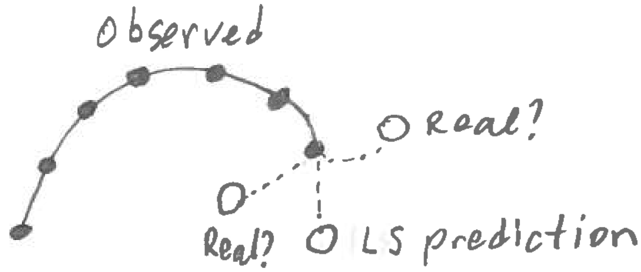


Figure 5: WLAN client motion.

§ Questions:

- **How to use empirical data to model uncertain past motion?**
- **Is optimal estimation (e.g. least-squares) a good strategy for predicting future position?**
-

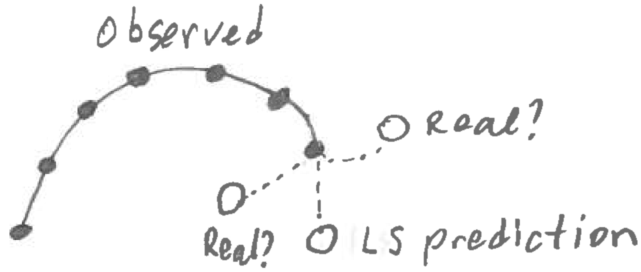


Figure 6: WLAN client motion.

§ Questions:

- **How to use empirical data to model uncertain past motion?**
- **Is optimal estimation (e.g. least-squares) a good strategy for predicting future position?**
- **Can we do better?**
-

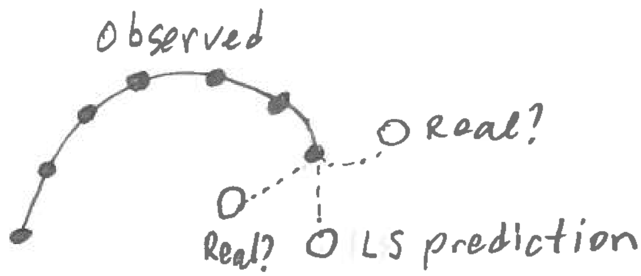


Figure 7: WLAN client motion.

§ Questions:

- **How to use empirical data to model uncertain past motion?**
- **Is optimal estimation (e.g. least-squares) a good strategy for predicting future position?**
- **Can we do better?**
- **How to manage both statistical and info-gap uncertainty?**
-

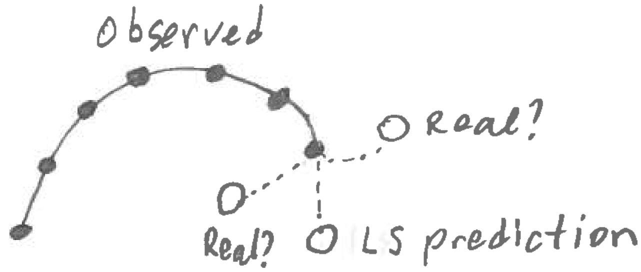


Figure 8: WLAN client motion.

§ **Questions:**

- **How to use empirical data to model uncertain past motion?**
- **Is optimal estimation (e.g. least-squares) a good strategy for predicting future position?**
- **Can we do better?**
- **How to manage both statistical and info-gap uncertainty?**
- **How to evaluate estimates vis à vis info-gaps?**

§ Modeling population behavior (Phillips curve):

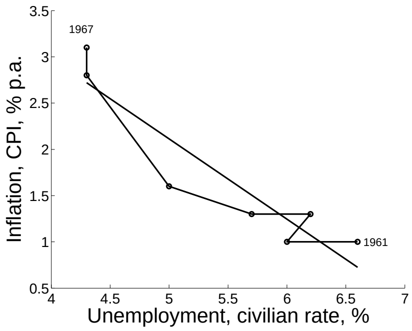


Figure 9: Inflation vs. unemployment in the US, 1961–1967.

§ Inflation vs. unemployment, US, '61–'67:

- **Approximately linear.**
- **Slope ≈ -0.87 %CPI/%unemployment.**

§

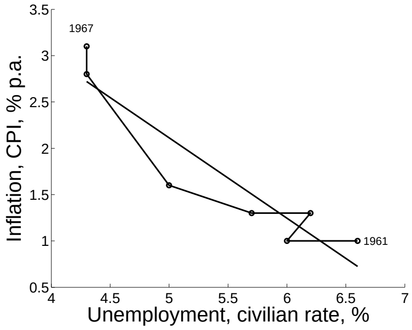


Figure 10: Inflation vs. unemployment in the US, 1961–1967.

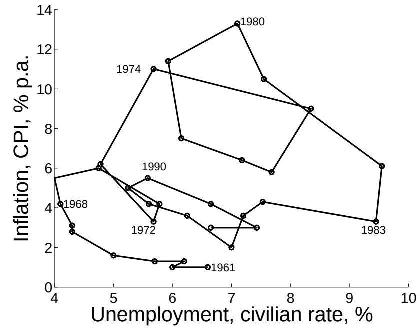


Figure 11: Inflation vs. unemployment in the US, 1961–1993.

§ **Modeling population behavior (Phillips curve):**

§ **Inflation vs. unemployment, US, '61–'67:**

- **Approximately linear.**
- **Slope ≈ −0.87 %CPI/%unemployment.**

§ **Slope changes greatly over time:**

- **'61–'67: −0.87**
- **'80–'83: −3.34**
- **'85–'93: −1.08**
- **'70–'78: ???**

§

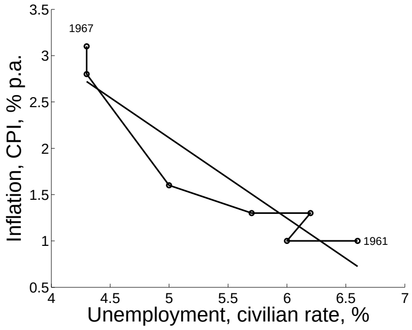


Figure 12: Inflation vs. unemployment in the US, 1961–1967.

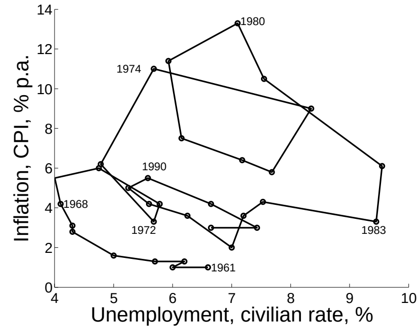


Figure 13: Inflation vs. unemployment in the US, 1961–1993.

§ Modeling population behavior (Phillips curve):

§ Inflation vs. unemployment, US, '61–'67:

- **Approximately linear.**
- **Slope ≈ -0.87 %CPI/%unemployment.**

§ Slope changes greatly over time:

- **'61–'67: -0.87**
- **'80–'83: -3.34**
- **'85–'93: -1.08**
- **'70–'78: ???**

§ Same challenges: Randomness and info-gaps.

§ Same questions.

1.2 *Robustness with Fractional-Error Parameter Uncertainty*

1.2.1 *Formulation of the Problem*

§ Paired data, fig. 14. E.g.:

- **Recovery time vs dose.**
- **Mental illness vs income.**
- **CPI vs unemployment.**

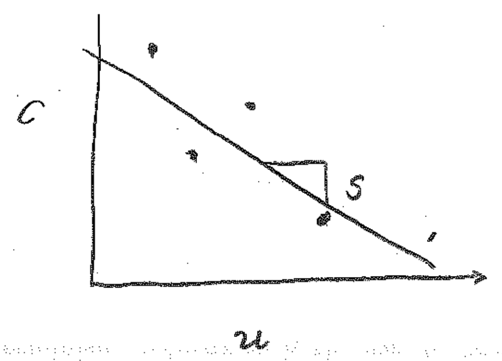


Figure 14: Paired data.

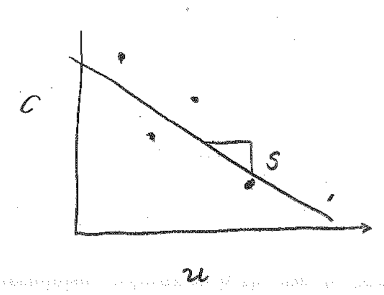


Figure 15: Paired data.

§ **Least-squares estimate of slope. E.g.**

Decrement of illness for increment of income.

•

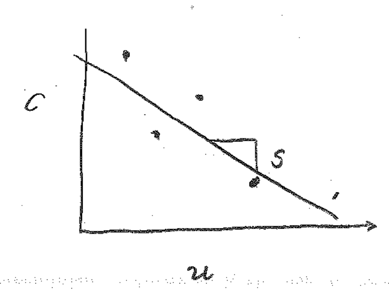


Figure 16: Paired data.

§ **Least-squares estimate of slope.** **E.g.**

Decrement of illness for increment of income.

- **Linear regression:** $c = su + b$.
-

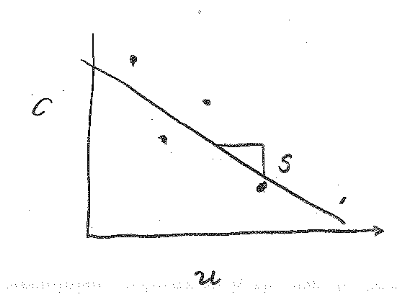


Figure 17: Paired data.

§ **Least-squares estimate of slope. E.g.**

Decrement of illness for increment of income.

- **Linear regression:** $c = su + b$.
- **Mean squared error: MSE** $= \frac{1}{N}\sum_{i=1}^{N}[c_i - (su_i + b)]^2$
-

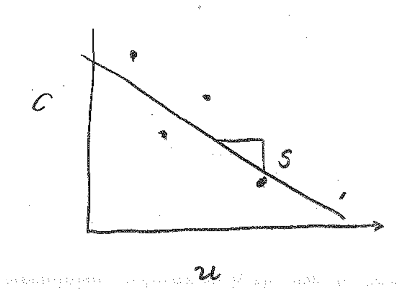


Figure 18: Paired data.

§ **Least-squares estimate of slope.** **E.g.**

Decrement of illness for increment of income.

- **Linear regression:** $c = su + b$.
- **Mean squared error:** $\mathbf{MSE} = \frac{1}{N}\sum_{i=1}^{N}[c_i - (su_i + b)]^2$
- **MSE of estimate of the slope:**

$$\tilde{s} = \arg\min_s \mathbf{MSE} = \frac{\mathrm{cov}(u, c)}{\mathrm{var}(u)} \qquad (1)$$

-

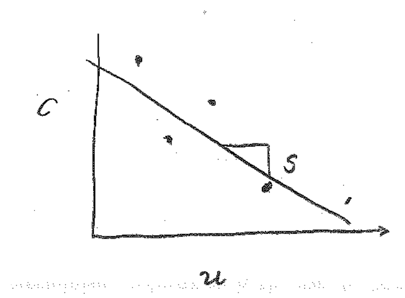


Figure 19: Paired data.

§ **Least-squares estimate of slope.** **E.g.**

Decrement of illness for increment of income.

- **Linear regression:** $c = su + b$.
- **Mean squared error:** $\mathbf{MSE} = \frac{1}{N}\sum_{i=1}^{N}[c_i - (su_i + b)]^2$
- **MSE of estimate of the slope:**

$$\widetilde{s} = \arg\min_s \mathbf{MSE} = \frac{\mathrm{cov}(u, c)}{\mathrm{var}(u)} \tag{2}$$

- **Uncertainties:** **info-gaps and statistical variability.**
-

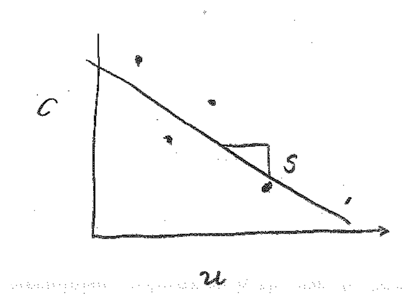


Figure 20: Paired data.

§ **Least-squares estimate of slope.** **E.g.**

Decrement of illness for increment of income.

- **Linear regression:** $c = su + b$.
- **Mean squared error:** $\mathbf{MSE} = \frac{1}{N} \sum_{i=1}^{N} [c_i - (su_i + b)]^2$
- **MSE of estimate of the slope:**

$$\widetilde{s} = \arg\min_{s} \mathbf{MSE} = \frac{\mathrm{cov}(u,c)}{\mathrm{var}(u)} \tag{3}$$

- **Uncertainties:** **info-gaps and statistical variability.**
- **Response:** **info-gap robust-satisficing.**

1.2.2 *Formulation of the Robustness Function*

§ Robustness question:

- **How much can the data err due to info-gaps, and the slope's error will be acceptable?**
-

§ Robustness question:

- **How much can the data err due to info-gaps, and the slope's error will be acceptable?**
- **Two uncertainties: statistical and info-gap. We focus on non-statistical info-gaps.**
-

§ Robustness question:

- **How much can the data err due to info-gaps, and the slope's error will be acceptable?**
- **Two uncertainties: statistical and info-gap. We focus on non-statistical info-gaps.**
- **We are not trying to minimize the error.**
-

§ Robustness question:

- **How much can the data err due to info-gaps, and the slope's error will be acceptable?**
- **Two uncertainties: statistical and info-gap. We focus on non-statistical info-gaps.**
- **We are not trying to minimize the error.**
- **We satisfice the error and optimize robustness to uncertainty.**

§ Notation for moments:

$\gamma =$ **covariance,** $\mathrm{cov}(u, c)$**.** $\widetilde{\gamma} =$ **estimate.**

$\sigma^2 =$ **variance,** $\mathrm{var}(u)$**.** $\widetilde{\sigma}^2 =$ **estimate.**

§

§ Notation for moments:

$\gamma =$ **covariance,** $\mathrm{cov}(u, c)$**.** $\widetilde{\gamma} =$ **estimate.**

$\sigma^2 =$ **variance,** $\mathrm{var}(u)$**.** $\widetilde{\sigma}^2 =$ **estimate.**

§ Consider info-gaps in data:

Unknown fractional errors of moments:

$$\left| \frac{\gamma - \widetilde{\gamma}}{\widetilde{\gamma}} \right| = \mathbf{???}, \quad \left| \frac{\sigma^2 - \widetilde{\sigma}^2}{\widetilde{\sigma}^2} \right| = \mathbf{???} \tag{4}$$

§

§ Notation for moments:

$\gamma =$ **covariance,** $\mathrm{cov}(u, c)$**.** $\widetilde{\gamma} =$ **estimate.**

$\sigma^2 =$ **variance,** $\mathrm{var}(u)$**.** $\widetilde{\sigma}^2 =$ **estimate.**

§ Consider info-gaps in data:

Unknown fractional errors of moments:

$$\left| \frac{\gamma - \widetilde{\gamma}}{\widetilde{\gamma}} \right| = \mathbf{???}, \quad \left| \frac{\sigma^2 - \widetilde{\sigma}^2}{\widetilde{\sigma}^2} \right| = \mathbf{???} \tag{5}$$

§ Fractional-error info-gap model of uncertainty:

$$\mathcal{U}(h) = \left\{ (\gamma, \sigma^2) : \ \left| \frac{\gamma - \widetilde{\gamma}}{\widetilde{\gamma}} \right| \le h, \ \ \left| \frac{\sigma^2 - \widetilde{\sigma}^2}{\widetilde{\sigma}^2} \right| \le h, \ \sigma^2 \ge 0 \right\}, \quad h \ge 0$$

- $h =$ **unbounded horizon of uncertainty.**

  No known worst case.
- **Non-probabilistic deep uncertainty.**

§ **Least-squares estimate of slope:** $\widetilde{s} = \widetilde{\gamma}/\widetilde{\sigma}^2$.

Actual value: $s = \gamma/\sigma^2$.

§

§ **Least-squares estimate of slope:** $\widetilde{s} = \widetilde{\gamma}/\widetilde{\sigma}^2$**.**

Actual value: $s = \gamma/\sigma^2$**.**

§ **Satisficing performance requirement:**

$$|s(\gamma, \sigma^2) - \widetilde{s}| \leq R_{\mathrm{C}} \tag{6}$$

§

§ **Least-squares estimate of slope:** $\widetilde{s} = \widetilde{\gamma}/\widetilde{\sigma}^2$.

Actual value: $s = \gamma/\sigma^2$.

§ **Satisficing performance requirement:**

$$|s(\gamma, \sigma^2) - \widetilde{s}| \le R_{\mathrm{C}} \tag{7}$$

§ **Robustness of LS estimate** $\widetilde{s}$**:**

- **Maximum tolerable uncertainty.**
-

§ **Least-squares estimate of slope: $\widetilde{s} = \widetilde{\gamma}/\widetilde{\sigma}^2$.**

Actual value: $s = \gamma/\sigma^2$.

§ **Satisficing performance requirement:**

$$|s(\gamma, \sigma^2) - \widetilde{s}| \leq R_{\mathrm{C}} \tag{8}$$

§ **Robustness of LS estimate $\widetilde{s}$:**

- **Maximum tolerable uncertainty.**
- **Max horizon of uncertainty in moments at which $\widetilde{s}$ errs no more than R_{C}:**

$$\widehat{h}(\widetilde{s}, R_{\mathrm{C}}) = \max\left\{h : \left(\max_{\gamma,\sigma^2 \in \mathcal{U}(h)} |s(\gamma, \sigma^2) - \widetilde{s}|\right) \leq R_{\mathrm{C}}\right\} \tag{9}$$

-

§ **Least-squares estimate of slope:** $\widetilde{s} = \widetilde{\gamma}/\widetilde{\sigma}^2$.

Actual value: $s = \gamma/\sigma^2$.

§ **Satisficing performance requirement:**

$$|s(\gamma, \sigma^2) - \widetilde{s}| \le R_{\rm C} \tag{10}$$

§ **Robustness of LS estimate $\widetilde{s}$:**

- **Maximum tolerable uncertainty.**
- **Max horizon of uncertainty in moments at which $\widetilde{s}$ errs no more than $R_{\rm C}$:**

$$\widehat{h}(\widetilde{s}, R_{\rm C}) = \max\left\{h : \ \left(\max_{\gamma,\sigma^2\in\mathcal{U}(h)} |s(\gamma, \sigma^2) - \widetilde{s}|\right) \le R_{\rm C}\right\} \tag{11}$$

- **Satisfice the error.**
- **Maximize the robustness to info-gaps.**
-

§ **Least-squares estimate of slope:** $\widetilde{s} = \widetilde{\gamma}/\widetilde{\sigma}^2$.

Actual value: $s = \gamma/\sigma^2$.

§ **Satisficing performance requirement:**

$$|s(\gamma, \sigma^2) - \widetilde{s}| \le R_{\mathrm{C}} \tag{12}$$

§ **Robustness of LS estimate** $\widetilde{s}$**:**

- **Maximum tolerable uncertainty.**
- **Max horizon of uncertainty in moments at which** $\widetilde{s}$ **errs no more than** R_{C}**:**

$$\widehat{h}(\widetilde{s}, R_{\mathrm{C}}) = \max\left\{ h : \left(\max_{\gamma, \sigma^2 \in \mathcal{U}(h)} |s(\gamma, \sigma^2) - \widetilde{s}| \right) \le R_{\mathrm{C}} \right\} \tag{13}$$

- **Satisfice the error.**
- **Maximize the robustness to info-gaps.**
- **Don't try to minimize the error.**

1.2.3 *Derivation of the Robustness Function*

§ Definition of robustness:

$$\widehat{h}(\widetilde{s}, R_{\mathrm{C}}) = \max\left\{h : \left(\max_{\gamma,\sigma^2 \in \mathcal{U}(h)} \left|s(\gamma,\sigma^2) - \widetilde{s}\right|\right) \leq R_{\mathrm{C}}\right\} \qquad (14)$$

§

§ Definition of robustness:

$$\widehat{h}(\widetilde{s}, R_{\rm C}) = \max\left\{h:\ \left(\max_{\gamma,\sigma^2 \in \mathcal{U}(h)} \left|s(\gamma, \sigma^2) - \widetilde{s}\right|\right) \le R_{\rm C}\right\} \qquad (15)$$

§ Derivation of robustness:

- $m(h)$ = **inner maximum in eq.(15).**
-

§ Definition of robustness:

$$\widehat{h}(\widetilde{s}, R_{\mathrm{C}}) = \max\left\{h : \left(\max_{\gamma,\sigma^2\in\mathcal{U}(h)} |s(\gamma,\sigma^2) - \widetilde{s}|\right) \le R_{\mathrm{C}}\right\} \qquad (16)$$

§ Derivation of robustness:

- $m(h) =$ **inner maximum in eq.(16).**
- **Slope:** $s = \gamma/\sigma^2$**.**
-

§ Definition of robustness:

$$\widehat{h}(\widetilde{s}, R_{\rm C}) = \max\left\{h:\ \left(\max_{\gamma,\sigma^2\in\mathcal{U}(h)} \left|s(\gamma,\sigma^2) - \widetilde{s}\right|\right) \le R_{\rm C}\right\} \qquad (17)$$

§ Derivation of robustness:

- $m(h) =$ **inner maximum in eq.(17).**
- **Slope:** $s = \gamma/\sigma^2$.
- $m(h)$ **occurs at** $\gamma = (1+h)\widetilde{\gamma}$, $\sigma^2 = (1-h)^+\widetilde{\sigma}^2$.
-

§ Definition of robustness:

$$\widehat{h}(\widetilde{s}, R_{\rm C}) = \max\left\{h: \ \left(\max_{\gamma,\sigma^2\in\mathcal{U}(h)} |s(\gamma,\sigma^2)-\widetilde{s}|\right) \le R_{\rm C}\right\} \qquad (18)$$

§ Derivation of robustness:

- $m(h) =$ **inner maximum in eq.(18).**
- **Slope:** $s = \gamma/\sigma^2$.
- $m(h)$ **occurs at** $\gamma = (1+h)\widetilde{\gamma}$, $\sigma^2 = (1-h)^+\widetilde{\sigma}^2$.
- **Thus, for** $h \le 1$**:**

$$m(h) = \left|\frac{(1+h)\widetilde{\gamma}}{(1-h)\widetilde{\sigma}^2} - \frac{\widetilde{\gamma}}{\widetilde{\sigma}^2}\right| \qquad (19)$$

$$= \left(\frac{1+h}{1-h} - 1\right)\left|\frac{\widetilde{\gamma}}{\widetilde{\sigma}^2}\right| \qquad (20)$$

$$= \frac{2h}{1-h}|\widetilde{s}| \qquad (21)$$

- **Recall:**

$$m(h) = \frac{2h}{1-h}|\widetilde{s}| \tag{22}$$

- **Equate $m(h) = R_{\mathrm{C}}$ and solve for h (recall $\widetilde{s} < 0$):**

$$\frac{2h}{1-h} = -\frac{R_{\mathrm{C}}}{\widetilde{s}} = \rho \ \textbf{(definition)} \quad \Longrightarrow \quad \widehat{h} = \frac{\rho}{2+\rho} \ \ (\leq 1) \tag{23}$$

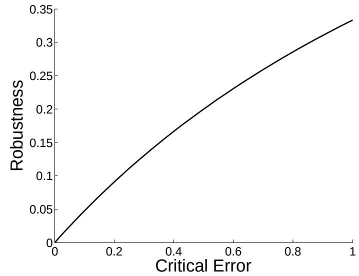


Figure 21: Robustness of estimated slope, $\widehat{h}(\widetilde{s}, \rho)$, vs. critical error, ρ. Eq.(24).

§ Robustness of LS estimate $\widetilde{s}$ (fig. 21):

$$\widehat{h}(\widetilde{s}, \rho) = \frac{\rho}{2+\rho}, \qquad \rho = -R_{\mathrm{C}}/\widetilde{s} \tag{24}$$

Recall: $\widetilde{s} < 0$ **so** $\rho > 0$**.**

•

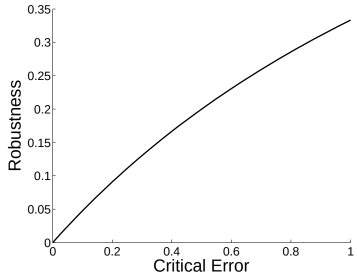


Figure 22: Robustness of estimated slope, $\widehat{h}(\widetilde{s}, \rho)$, vs. critical error, ρ. Eq.(25).

§ Robustness of LS estimate $\widetilde{s}$ (fig. 22):

$$\widehat{h}(\widetilde{s}, \rho) = \frac{\rho}{2+\rho}, \qquad \rho = -R_{\mathrm{C}}/\widetilde{s} \tag{25}$$

Recall: $\widetilde{s} < 0$ so $\rho > 0$.

- **Zeroing: Zero error has zero robustness.**
-

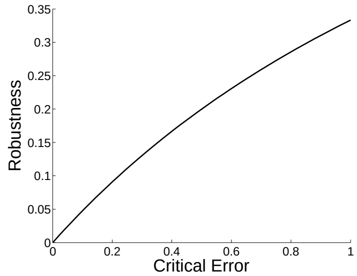


Figure 23: Robustness of estimated slope, $\widehat{h}(\widetilde{s}, \rho)$, vs. critical error, ρ. Eq.(26).

§ Robustness of LS estimate $\widetilde{s}$ (fig. 23):

$$\widehat{h}(\widetilde{s}, \rho) = \frac{\rho}{2+\rho}, \qquad \rho = -R_{\mathrm{C}}/\widetilde{s} \tag{26}$$

Recall: $\widetilde{s} < 0$ **so** $\rho > 0$**.**

- **Zeroing: Zero error has zero robustness.**
- **Trade-off: robustness goes up (good) as allowed estimation error, ρ, goes up (bad).**
-

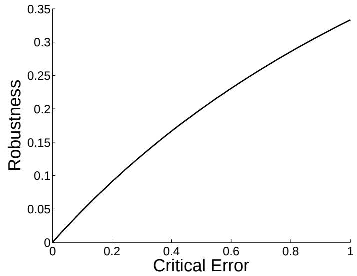


Figure 24: Robustness of estimated slope, $\widehat{h}(\widetilde{s}, \rho)$, vs. critical error, ρ. Eq.(27).

§ Robustness of LS estimate $\widetilde{s}$ (fig. 24):

$$\widehat{h}(\widetilde{s}, \rho) = \frac{\rho}{2+\rho}, \qquad \rho = -R_{\rm C}/\widetilde{s} \tag{27}$$

Recall: $\widetilde{s} < 0$ so $\rho > 0$.

- **Zeroing: Zero error has zero robustness.**
- **Trade-off: robustness goes up (good) as allowed estimation error, ρ, goes up (bad).**
- **Examples (fig. 24):**
  $\rho = 0.0$, $\widehat{h} = 0$. $\rho = 0.3$, $\widehat{h} = 0.13$. $\rho = 0.6$, $\widehat{h} = 0.23$.

1.2.4 *Can We Do Better? Crossing Robustness Curves*

§ **Can we do better than the LS estimate?**

§ **Yes, but . . . .**

§ Estimates of slope:

- $\widetilde{s}$ = **LS estimate, with robustness** $\widehat{h}(\widetilde{s}, R_{\mathrm{C}})$.
-

§ Estimates of slope:

- $\widetilde{s}$ = **LS estimate, with robustness** $\widehat{h}(\widetilde{s}, R_{\mathrm{C}})$.
- s_{e}= **any estimate, with robustness** $\widehat{h}(s_{\mathrm{e}}, R_{\mathrm{C}})$.
-

§ Estimates of slope:

- $\widetilde{s}$ = **LS estimate, with robustness** $\widehat{h}(\widetilde{s}, R_{\mathrm{C}})$.
- s_{e}= **any estimate, with robustness** $\widehat{h}(s_{\mathrm{e}}, R_{\mathrm{C}})$.
- **Definitions:** $\zeta = s_{\mathrm{e}}/\widetilde{s}$, $\rho = -R_{\mathrm{C}}/\widetilde{s}$. **(Recall:** $\widetilde{s} < 0$.**)**
-

§ Estimates of slope:

- $\widetilde{s}$ = **LS estimate, with robustness** $\widehat{h}(\widetilde{s}, R_{\mathrm{C}})$.
- s_{e}= **any estimate, with robustness** $\widehat{h}(s_{\mathrm{e}}, R_{\mathrm{C}})$.
- **Definitions:** $\zeta = s_{\mathrm{e}}/\widetilde{s}$, $\rho = -R_{\mathrm{C}}/\widetilde{s}$. **(Recall:** $\widetilde{s} < 0$.**)**
- **Robustness of** s_{e}**, in analogy to eq.(14), p.39:**

$$\widehat{h}(s_{\mathrm{e}}, R_{\mathrm{C}}) = \max\left\{h:\ \left(\max_{\gamma,\sigma^2\in\mathcal{U}(h)} |s(\gamma,\sigma^2) - s_{\mathrm{e}}|\right) \le R_{\mathrm{C}}\right\} \qquad (28)$$

-

§ Estimates of slope:

- $\widetilde{s}$ **= LS estimate, with robustness** $\widehat{h}(\widetilde{s}, R_{\mathrm{C}})$**.**
- s_{e}**= any estimate, with robustness** $\widehat{h}(s_{\mathrm{e}}, R_{\mathrm{C}})$**.**
- **Definitions:** $\zeta = s_{\mathrm{e}}/\widetilde{s}$**,** $\rho = -R_{\mathrm{C}}/\widetilde{s}$**. (Recall:** $\widetilde{s} < 0$**.)**
- **Robustness of** s_{e}**, in analogy to eq.(14), p.39:**

$$\widehat{h}(s_{\mathrm{e}}, R_{\mathrm{C}}) = \max\left\{h : \left(\max_{\gamma,\sigma^2 \in \mathcal{U}(h)} \left|s(\gamma, \sigma^2) - s_{\mathrm{e}}\right|\right) \le R_{\mathrm{C}}\right\} \quad (29)$$

- **Let** $m(h)$ **denote the inner maximum:**

$$m(h) = \max_{\gamma,\sigma^2 \in \mathcal{U}(h)} \left|\frac{\gamma}{\sigma^2} - s_{\mathrm{e}}\right| \quad (30)$$

-

§ Estimates of slope:

- $\widetilde{s}$ = **LS estimate, with robustness** $\widehat{h}(\widetilde{s}, R_{\rm C})$.
- $s_{\rm e}$= **any estimate, with robustness** $\widehat{h}(s_{\rm e}, R_{\rm C})$.
- **Definitions:** $\zeta = s_{\rm e}/\widetilde{s}$, $\quad \rho = -R_{\rm C}/\widetilde{s}$. **(Recall:** $\widetilde{s} < 0$.**)**
- **Robustness of** $s_{\rm e}$**, in analogy to eq.(14), p.39:**

$$\widehat{h}(s_{\rm e}, R_{\rm C}) = \max \left\{ h : \ \left(\max_{\gamma, \sigma^2 \in \mathcal{U}(h)} |s(\gamma, \sigma^2) - s_{\rm e}| \right) \leq R_{\rm C} \right\} \tag{31}$$

- **Let** $m(h)$ **denote the inner maximum:**

$$m(h) = \max_{\gamma, \sigma^2 \in \mathcal{U}(h)} \left| \frac{\gamma}{\sigma^2} - s_{\rm e} \right| \tag{32}$$

- $m(h)$ **occurs at one of the following (for** $h \leq 1$**):**

$$\textbf{Either:} \quad \gamma = (1+h)\widetilde{\gamma}, \quad \sigma^2 = (1-h)\widetilde{\sigma}^2 \tag{33}$$

$$\textbf{Or:} \quad \gamma = (1-h)\widetilde{\gamma}, \quad \sigma^2 = (1+h)\widetilde{\sigma}^2 \tag{34}$$

- $m(h)$ **occurs at one of the following (for** $h \le 1$**):**

$$\textbf{Either:} \quad \gamma = (1+h)\widetilde{\gamma}, \quad \sigma^2 = (1-h)\widetilde{\sigma}^2 \tag{35}$$

$$\textbf{Or:} \quad \gamma = (1-h)\widetilde{\gamma}, \quad \sigma^2 = (1+h)\widetilde{\sigma}^2 \tag{36}$$

-

- $m(h)$ **occurs at one of the following (for** $h \leq 1$**):**

$$\textbf{Either:}\quad \gamma = (1+h)\widetilde{\gamma}, \quad \sigma^2 = (1-h)\widetilde{\sigma}^2 \tag{37}$$

$$\textbf{Or:}\quad \gamma = (1-h)\widetilde{\gamma}, \quad \sigma^2 = (1+h)\widetilde{\sigma}^2 \tag{38}$$

- **Denote the corresponding** $m(h)$**'s:**

$$m_1(h) = \left|\frac{(1+h)\widetilde{\gamma}}{(1-h)\widetilde{\sigma}^2} - s_{\mathrm{e}}\right| \tag{39}$$

$$m_2(h) = \left|\frac{(1-h)\widetilde{\gamma}}{(1+h)\widetilde{\sigma}^2} - s_{\mathrm{e}}\right| \tag{40}$$

-

- $m(h)$ **occurs at one of the following (for** $h \leq 1$**):**

$$\textbf{Either:} \quad \gamma = (1+h)\widetilde{\gamma}, \quad \sigma^2 = (1-h)\widetilde{\sigma}^2 \tag{41}$$

$$\textbf{Or:} \quad \gamma = (1-h)\widetilde{\gamma}, \quad \sigma^2 = (1+h)\widetilde{\sigma}^2 \tag{42}$$

- **Denote the corresponding** $m(h)$**'s:**

$$m_1(h) = \left| \frac{(1+h)\widetilde{\gamma}}{(1-h)\widetilde{\sigma}^2} - s_{\mathrm{e}} \right| \tag{43}$$

$$m_2(h) = \left| \frac{(1-h)\widetilde{\gamma}}{(1+h)\widetilde{\sigma}^2} - s_{\mathrm{e}} \right| \tag{44}$$

- $m(h)$ **is the greater of these two alternatives:**

$$m(h) = \max[m_1(h),\ m_2(h)] \tag{45}$$

The maximum depends on the value of h**.**

- **After some algebra, recalling $\rho = -R_{\rm C}/\widetilde{s}$, and equating $m(h) = R_{\rm C}$, the robustness function is:**

$$\widehat{h}(s_{\rm e}, \rho) = \begin{cases} \dfrac{\rho + \zeta - 1}{\rho + \zeta + 1} & \textbf{if } \rho^2 \geq \zeta^2 - 1 \textbf{ and } \rho \geq 1 - \zeta \\[2ex] \dfrac{\rho - \zeta + 1}{-\rho + \zeta + 1} & \textbf{if } \rho^2 \leq \zeta^2 - 1 \textbf{ and } \rho \geq \zeta - 1 \end{cases} \tag{46}$$

$\widehat{h}(s_{\rm e}, \rho)$ **is zero otherwise. Note $\widehat{h} \leq 1$.**

- **Eq.(46) includes eq.(24) as a special case, when $\zeta = 1$.**
- **Figures coming!**

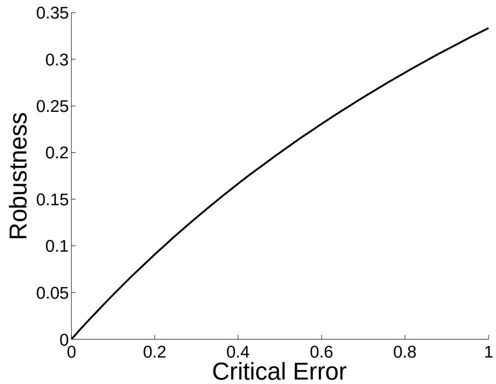


Figure 25: $\widehat{h}(\widetilde{s}, \rho)$ vs. ρ.

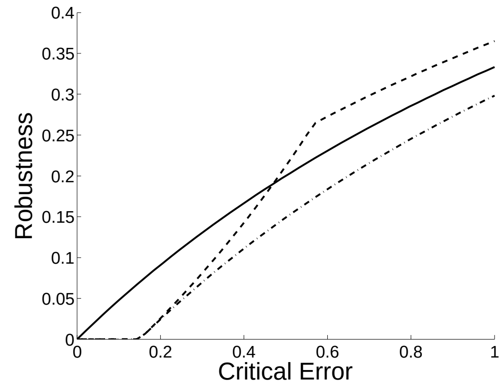


Figure 26: $\widehat{h}(s_{\rm e}, \rho)$ vs. ρ. $\zeta = 1$ (solid), 1.15 (dash), 0.85 (dot-dash).

- **Left fig: Robustness of LS estimate: $s_{\rm e} = \widetilde{s}$.**
-

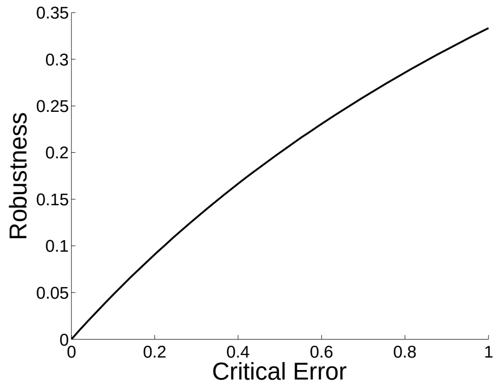


Figure 27: $\widehat{h}(\widetilde{s}, \rho)$ vs. ρ.

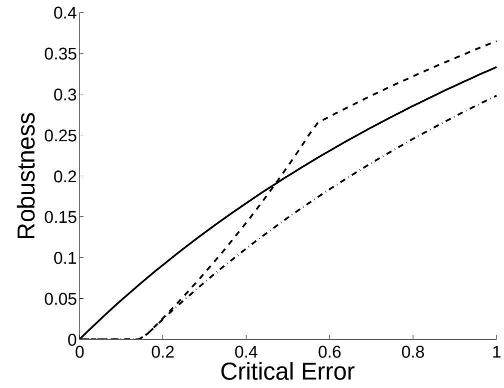


Figure 28: $\widehat{h}(s_{\rm e}, \rho)$ vs. ρ. $\zeta = 1$ (solid), 1.15 (dash), 0.85 (dot-dash).

- **Left fig: Robustness of LS estimate: $s_{\rm e} = \widetilde{s}$.**
- **Right fig:**
  - **Solid: Robustness of LS estimate: $s_{\rm e} = \widetilde{s}$.**
  -

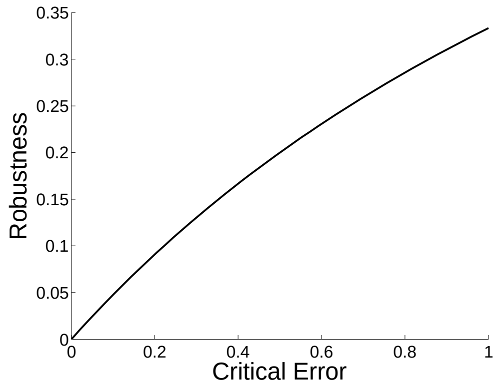


Figure 29: $\widehat{h}(\widetilde{s}, \rho)$ vs. ρ.

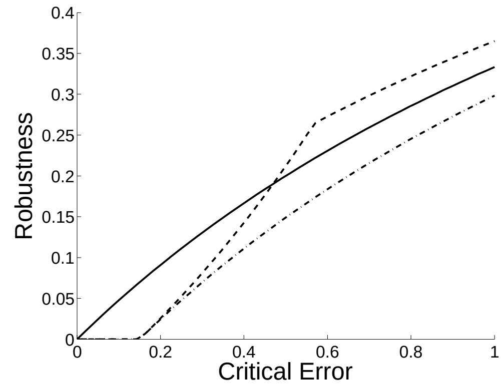


Figure 30: $\widehat{h}(s_{\rm e}, \rho)$ vs. ρ. $\zeta = 1$ (solid), 1.15 (dash), 0.85 (dot-dash).

- **Left fig: Robustness of LS estimate: $s_{\rm e} = \widetilde{s}$.**
- **Right fig:**
  - **Solid: Robustness of LS estimate: $s_{\rm e} = \widetilde{s}$.**
  - **Dot-dash (lower): Robustness of $s_{\rm e} = 0.85\widetilde{s}$. Robust dominated. Not preferred.**
  -

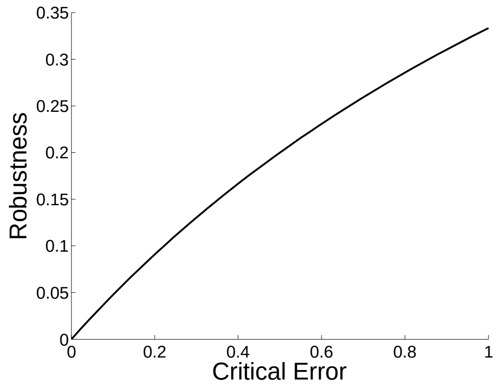


Figure 31: $\widehat{h}(\widetilde{s}, \rho)$ vs. ρ.

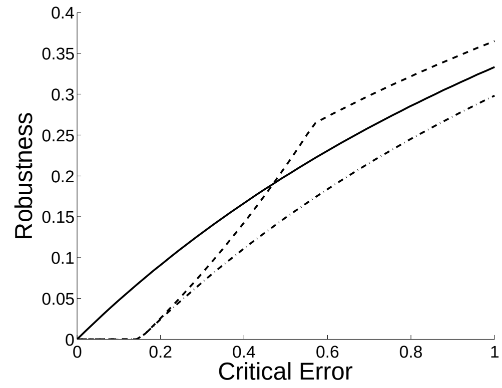


Figure 32: $\widehat{h}(s_{\rm e}, \rho)$ vs. ρ. $\zeta = 1$ (solid), 1.15 (dash), 0.85 (dot-dash).

- **Left fig: Robustness of LS estimate: $s_{\rm e} = \widetilde{s}$.**
- **Right fig:**
  - **Solid: Robustness of LS estimate: $s_{\rm e} = \widetilde{s}$.**
  - **Dot-dash (lower): Robustness of $s_{\rm e} = 0.85\widetilde{s}$. Robust dominated. Not preferred.**
  - **Dash (crossing): Robustness of $s_{\rm e} = 1.15\widetilde{s}$. Preference reversal. May be preferred.**

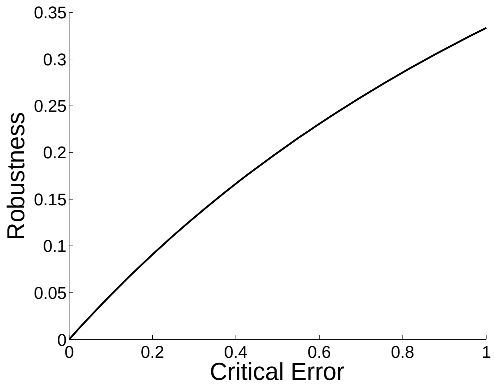


Figure 33: $\widehat{h}(\widetilde{s}, \rho)$ vs. ρ.

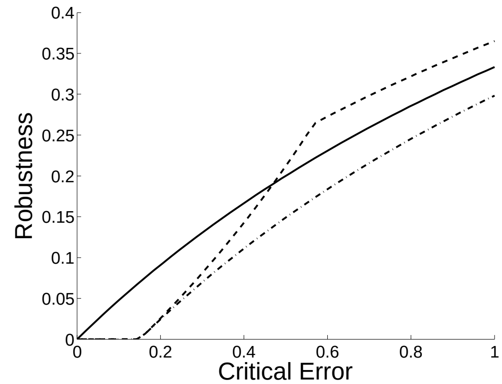


Figure 34: $\widehat{h}(s_{\rm e}, \rho)$ vs. ρ. $\zeta = 1$ (solid), 1.15 (dash), 0.85 (dot-dash).

§ Can we do better than least-squares?

§

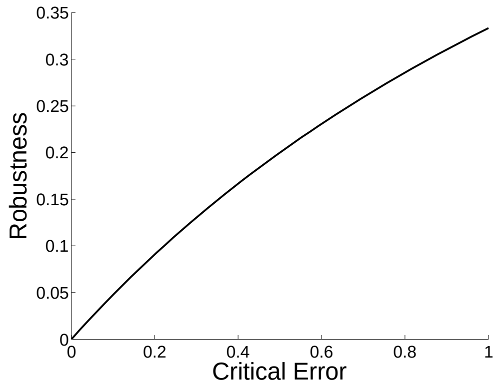


Figure 35: $\widehat{h}(\widetilde{s}, \rho)$ vs. ρ.

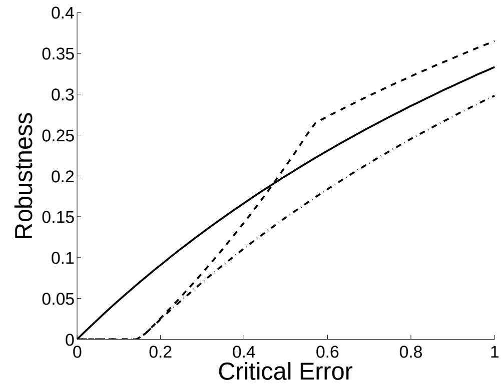


Figure 36: $\widehat{h}(s_{\rm e}, \rho)$ vs. ρ. $\zeta = 1$ (solid), 1.15 (dash), 0.85 (dot-dash).

§ **Can we do better than least-squares?**

§ **Yes, but at a price:**

Crossing robustness curves represent potential for preference reversal: non-LS more robust than LS at positive error.

∼∼∼

2 *Info-Gap Forecasting: 1-D Example*

“Prediction is always difficult, especially of the future.”

Robert Storm Petersen, Danish journalist.

[2] lectures talks lib forecast03one-d.tex 23.5.2023

2.1 *Introduction*

§ **The generic task:**

Given historical data, predict the future state.

§

§ The generic task:

Given historical data, predict the future state.

§ The challenges:

- **Data deficient, uncertain, erroneous.**
-

§ **The generic task:**

Given historical data, predict the future state.

§ **The challenges:**

- **Data deficient, uncertain, erroneous.**
- **Processes complex, poorly understood.**

§

§ **The generic task:**

Given historical data, predict the future state.

§ **The challenges:**

- **Data deficient, uncertain, erroneous.**
- **Processes complex, poorly understood.**

§ **Challenging forecasts:**

- **Prevalence of emerging new disease.**
- **Effectiveness of new drug or technology.**
- **Economic forecast of CPI or GDP.**
- **Duration until next outbreak of disease, terror, war, etc.**

2.2 *Formulation*

§ **State variable (e.g. # infected at time t): λ_t**

§

§ **State variable (e.g. # infected at time t): λ_t**

§ **Evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{47}$$

§

§ **State variable (e.g. # infected at time t): λ_t**

§ **Evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{48}$$

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

§

§ **State variable (e.g. # infected at time t): λ_t**

§ **Evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{49}$$

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

§ **Contextual understanding: λ_t seems to drift upwards.**

§

§ **State variable (e.g. # infected at time t): λ_t**

§ **Evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{50}$$

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

§ **Contextual understanding: λ_t seems to drift upwards.**

§ **Info-gap model for deep uncertainty about the future:**

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_t, t > T : \ 0 \le \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \le h \right\}, \quad h \ge 0 \tag{51}$$

- **Unbounded family of sets.**
- **No known worst case.**
- **No known probability distribution.**

§ Unknown true evolution in time:

$$y_t = \lambda_t y_{t-1} \tag{52}$$

§

§ **Unknown true evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{53}$$

§ **Historically estimated forecast:**

$$y_t^{\mathrm{h}} = \widetilde{\lambda}_t y_{t-1}^{\mathrm{h}} \tag{54}$$

§

§ **Unknown true evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{55}$$

§ **Historically estimated forecast:**

$$y_t^{\mathrm{h}} = \widetilde{\lambda}_t y_{t-1}^{\mathrm{h}} \tag{56}$$

§ **Slope-adjusted forecast:**

$$y_t^{\mathrm{s}} = \ell y_{t-1}^{\mathrm{s}} \tag{57}$$

§

§ **Unknown true evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{58}$$

§ **Historically estimated forecast:**

$$y_t^{\mathrm{h}} = \widetilde{\lambda}_t y_{t-1}^{\mathrm{h}} \tag{59}$$

§ **Slope-adjusted forecast:**

$$y_t^{\mathrm{s}} = \ell y_{t-1}^{\mathrm{s}} \tag{60}$$

§ **Use contextual understanding to choose $\ell \geq \widetilde{\lambda}$.**

§

§ **Unknown true evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{61}$$

§ **Historically estimated forecast:**

$$y_t^{\mathrm{h}} = \widetilde{\lambda}_t y_{t-1}^{\mathrm{h}} \tag{62}$$

§ **Slope-adjusted forecast:**

$$y_t^{\mathrm{s}} = \ell y_{t-1}^{\mathrm{s}} \tag{63}$$

§ **Use contextual understanding to choose $\ell \geq \widetilde{\lambda}$.**

§ **Robust satisficing:**

- **Satisfice the forecast error:**

$$\left| y_{T+k}^{\mathrm{s}} - y_{T+k} \right| \leq \varepsilon_{\mathrm{c}} \tag{64}$$

-

§ **Unknown true evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{65}$$

§ **Historically estimated forecast:**

$$y_t^{\mathrm{h}} = \widetilde{\lambda}_t y_{t-1}^{\mathrm{h}} \tag{66}$$

§ **Slope-adjusted forecast:**

$$y_t^{\mathrm{s}} = \ell y_{t-1}^{\mathrm{s}} \tag{67}$$

§ **Use contextual understanding to choose $\ell \geq \widetilde{\lambda}$.**

§ **Robust satisficing:**

- **Satisfice the forecast error:**

$$\left| y_{T+k}^{\mathrm{s}} - y_{T+k} \right| \leq \varepsilon_{\mathrm{c}} \tag{68}$$

- **Maximize robustness to future surprise.**
-

§ **Unknown true evolution in time:**

$$y_t = \lambda_t y_{t-1} \tag{69}$$

§ **Historically estimated forecast:**

$$y_t^{\mathrm{h}} = \widetilde{\lambda}_t y_{t-1}^{\mathrm{h}} \tag{70}$$

§ **Slope-adjusted forecast:**

$$y_t^{\mathrm{s}} = \ell y_{t-1}^{\mathrm{s}} \tag{71}$$

§ **Use contextual understanding to choose $\ell \geq \widetilde{\lambda}$.**

§ **Robust satisficing:**

- **Satisfice the forecast error:**

$$\left| y_{T+k}^{\mathrm{s}} - y_{T+k} \right| \leq \varepsilon_{\mathrm{c}} \tag{72}$$

- **Maximize robustness to future surprise.**
- **Don't try to minimize the forecast error.**

§ **Recall the info-gap model for uncertain λ_t:**

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_t, t > T : \; 0 \leq \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \leq h \right\}, \quad h \geq 0 \qquad (73)$$

§

§ **Recall the info-gap model for uncertain λ_t:**

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_t, t > T : \ 0 \leq \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \leq h \right\}, \quad h \geq 0 \tag{74}$$

§ **Robustness of forecast ℓ:**

- **Maximum acceptable error in estimate, $\widetilde{\lambda}$.**
-

§ **Recall the info-gap model for uncertain λ_t:**

$$\mathcal{U}(h, \widetilde{\lambda}) = \left\{ \lambda_t, t > T : \ 0 \le \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \le h \right\}, \quad h \ge 0 \tag{75}$$

§ **Robustness of forecast ℓ:**

- **Maximum acceptable error in estimate, $\widetilde{\lambda}$.**
- **Max h up to which all λ_{T+i} in $\mathcal{U}(h, \widetilde{\lambda})$ satisfice forecast error at ε_c:**

$$\widehat{h}_s(\ell, \varepsilon_c) = \max \left\{ h : \ \left(\max_{\substack{\lambda_{T+i} \in \mathcal{U}(h, \widetilde{\lambda}) \\ i=1, \ldots, k}} \left| y^s_{T+k} - y_{T+k} \right| \right) \le \varepsilon_c \right\} \tag{76}$$

§

§ **Recall the info-gap model for uncertain λ_t:**

$$\mathcal{U}(h,\widetilde{\lambda}) = \left\{ \lambda_t, t > T: \ 0 \le \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \le h \right\}, \quad h \ge 0 \tag{77}$$

§ **Robustness of forecast ℓ:**

- **Maximum acceptable error in estimate, $\widetilde{\lambda}$.**
- **Max h up to which all λ_{T+i} in $\mathcal{U}(h,\widetilde{\lambda})$ satisfice forecast error at $\varepsilon_{\rm c}$:**

$$\widehat{h}_{\rm s}(\ell,\varepsilon_{\rm c}) = \max\left\{ h: \ \left(\max_{\substack{\lambda_{T+i}\in\mathcal{U}(h,\widetilde{\lambda}) \\ i=1,\,\ldots,\,k}} \left| y^{\rm s}_{T+k} - y_{T+k} \right| \right) \le \varepsilon_{\rm c} \right\} \tag{78}$$

§ **Robust-satisficing preference:** $\ell \succ \ell'$ **if** $\widehat{h}_{\rm s}(\ell,\varepsilon_{\rm c}) > \widehat{h}_{\rm s}(\ell',\varepsilon_{\rm c})$

-

§ **Recall the info-gap model for uncertain λ_t:**

$$\mathcal{U}(h,\widetilde{\lambda}) = \left\{ \lambda_t, t > T : \ 0 \le \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \le h \right\}, \quad h \ge 0 \tag{79}$$

§ **Robustness of forecast ℓ:**

- **Maximum acceptable error in estimate, $\widetilde{\lambda}$.**
- **Max h up to which all λ_{T+i} in $\mathcal{U}(h,\widetilde{\lambda})$ satisfice forecast error at $\varepsilon_{\rm c}$:**

$$\widehat{h}_{\rm s}(\ell,\varepsilon_{\rm c}) = \max\left\{ h : \ \left(\max_{\substack{\lambda_{T+i}\in\mathcal{U}(h,\widetilde{\lambda}) \\ i=1,\ \ldots,\ k}} \left| y^{\rm s}_{T+k} - y_{T+k} \right| \right) \le \varepsilon_{\rm c} \right\} \tag{80}$$

§ **Robust-satisficing preference: $\ell \succ \ell'$ if $\widehat{h}_{\rm s}(\ell,\varepsilon_{\rm c}) > \widehat{h}_{\rm s}(\ell',\varepsilon_{\rm c})$**

- **Satisfice error; maximize robustness.**
-

§ **Recall the info-gap model for uncertain λ_t:**

$$\mathcal{U}(h,\widetilde{\lambda}) = \left\{ \lambda_t, t > T : \ 0 \le \frac{\lambda_t - \widetilde{\lambda}}{\widetilde{\lambda}} \le h \right\}, \quad h \ge 0 \tag{81}$$

§ **Robustness of forecast ℓ:**

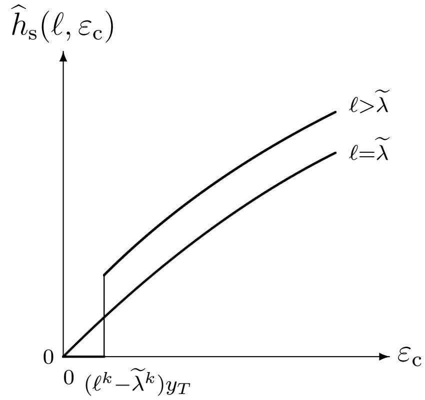
- **Maximum acceptable error in estimate, $\widetilde{\lambda}$.**
- **Max h up to which all λ_{T+i} in $\mathcal{U}(h,\widetilde{\lambda})$ satisfice forecast error at ε_c:**

$$\widehat{h}_s(\ell, \varepsilon_c) = \max \left\{ h : \ \left(\max_{\substack{\lambda_{T+i} \in \mathcal{U}(h,\widetilde{\lambda}) \\ i=1, \ldots, k}} \left| y^s_{T+k} - y_{T+k} \right| \right) \le \varepsilon_c \right\} \tag{82}$$

§ **Robust-satisficing preference: $\ell \succ \ell'$ if $\widehat{h}_s(\ell, \varepsilon_c) > \widehat{h}_s(\ell', \varepsilon_c)$**

- **Satisfice error; maximize robustness.**
- **Don't try to minimize error.**

Robustness



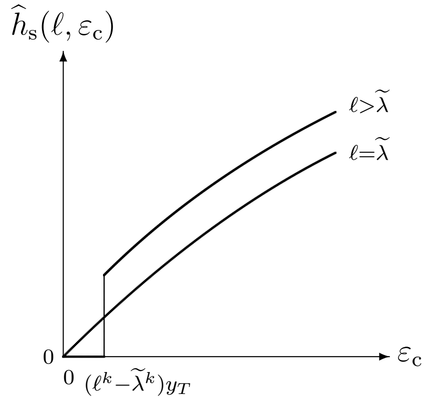
Forecast error

§ **Trade off:**

robustness up (good) $\Longleftrightarrow$ forecast error down (bad).

§

Robustness



Forecast error

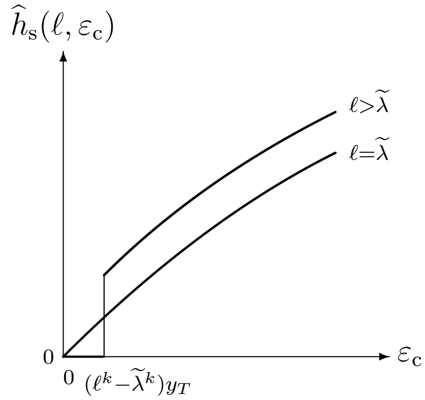
§ **Trade off:**

robustness up (good) $\Longleftrightarrow$ forecast error down (bad).

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

§

Robustness



Forecast error

§ **Trade off:**

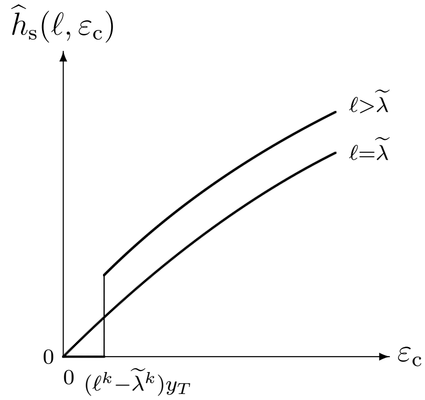
robustness up (good) $\Longleftrightarrow$ forecast error down (bad).

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

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

- **Preference reversal.**
-

Robustness



Forecast error

§ **Trade off:**

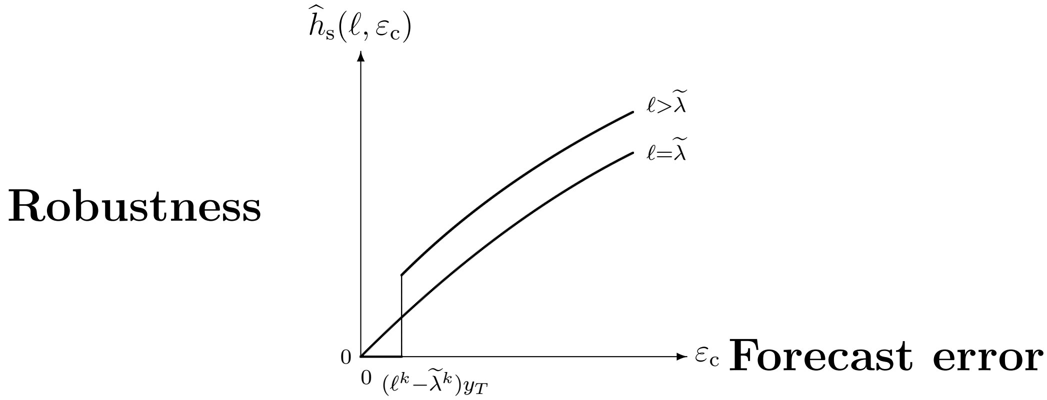
robustness up (good) $\Longleftrightarrow$ forecast error down (bad).

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

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

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

§



§ **Trade off:**

robustness up (good) $\Longleftrightarrow$ forecast error down (bad).

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

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

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

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

2.3 *Numerical example*

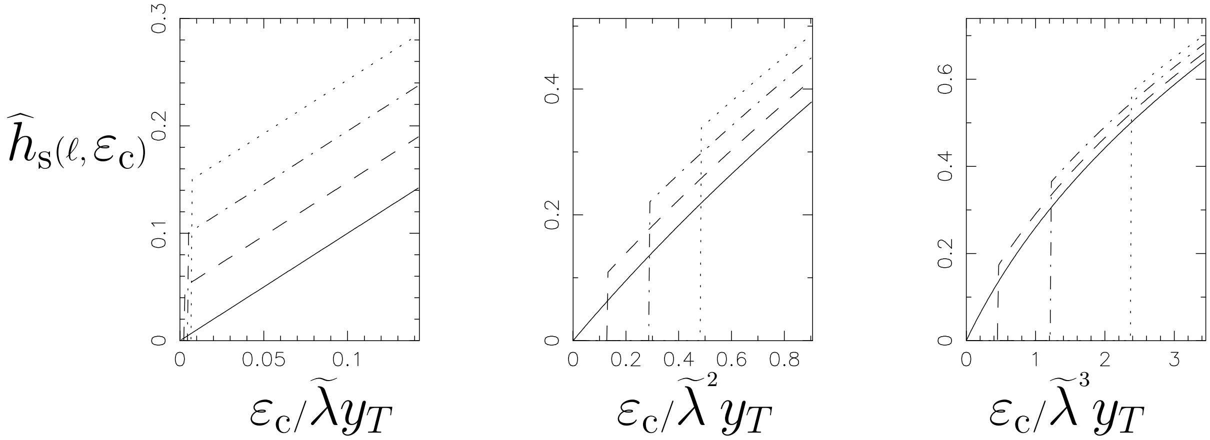


Figure 37: Robustness vs. normalized forecast error for $\ell = 1.05$, 1.1, 1.15, 1.2 from bottom to top curve. $\widetilde{\lambda} = 1.05$, $y_T = 1$. $k = 1$ (left), 2(mid), 3(right).

- **ℓ increases from bottom to top curve.**
-

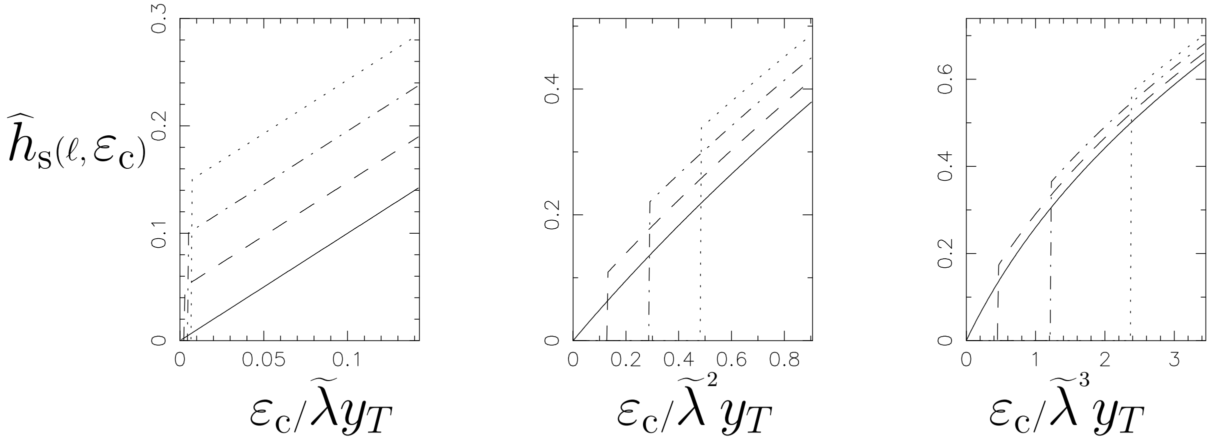


Figure 38: Robustness vs. normalized forecast error for $\ell = 1.05$, 1.1, 1.15, 1.2 from bottom to top curve. $\widetilde{\lambda} = 1.05$, $y_T = 1$. $k = 1$ (left), 2(mid), 3(right).

- **ℓ increases from bottom to top curve.**
- **Preference reversal at all times, $k = 1, 2, 3$.**
-

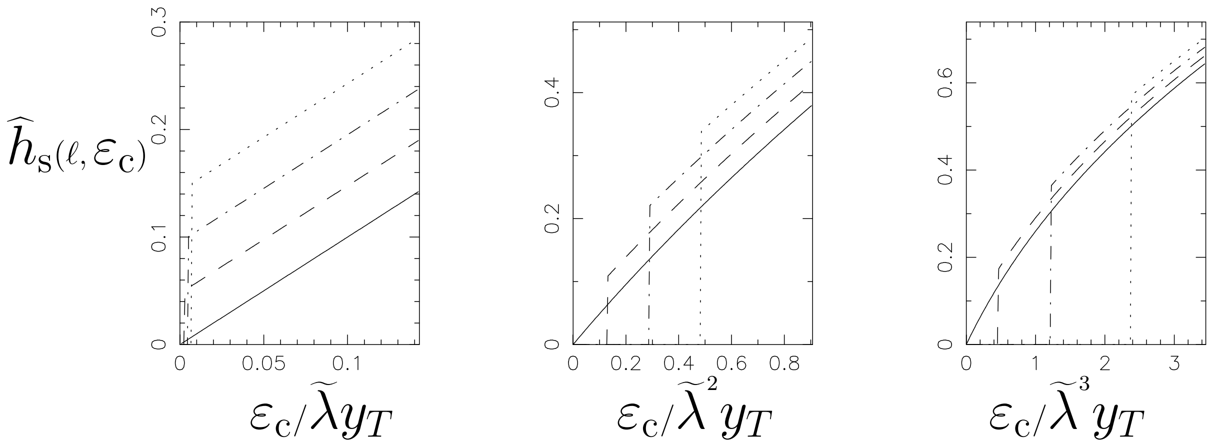


Figure 39: Robustness vs. normalized forecast error for $\ell = 1.05$, 1.1, 1.15, 1.2 from bottom to top curve. $\widetilde{\lambda} = 1.05$, $y_T = 1$. $k = 1$ (left), 2(mid), 3(right).

- **ℓ increases from bottom to top curve.**
- **Preference reversal at all times, $k = 1, 2, 3$.**
- **Robustness premium decreases with time.**
-

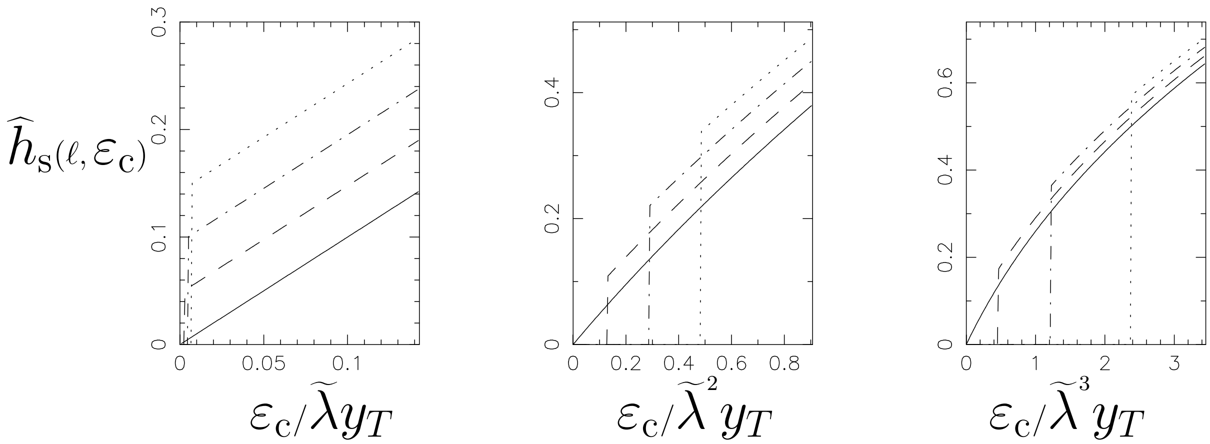


Figure 40: Robustness vs. normalized forecast error for $\ell = 1.05$, 1.1, 1.15, 1.2 from bottom to top curve. $\widetilde{\lambda} = 1.05$, $y_T = 1$. $k = 1$ (left), 2(mid), 3(right).

- **ℓ increases from bottom to top curve.**
- **Preference reversal at all times, $k = 1, 2, 3$.**
- **Robustness premium decreases with time.**
- **Reversal-ε_{c} increases with time.**

2.4 *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.**

§

§ **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:**

$$\mathcal{Y}(\ell) = \{\lambda_{T,k} \in D : \ |y^{\mathrm{s}}_{T+k}(\ell) - y_{T+k}| \le \varepsilon_{\mathrm{c}}\} \qquad (83)$$

§

§ **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:**

$$\mathcal{Y}(\ell) = \{\lambda_{T,k} \in D : \ |y^{\mathrm{s}}_{T+k}(\ell) - y_{T+k}| \le \varepsilon_{\mathrm{c}}\} \tag{84}$$

§ **Probability of success:**

$$P_{\mathrm{s}}(\ell) = F[\mathcal{Y}(\ell)] \tag{85}$$

§ Goal:

Choose ℓ to maximize probability of success.

§

§ **Goal:**

Choose ℓ to maximize probability of success.

§ **Problem:**

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

§

§ Goal:

Choose ℓ to maximize probability of success.

§ Problem:

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

§ Solution:

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

§ Theorem.

Probability of successful forecast, $P_{\rm s}(\ell)$,

increases with

increasing info-gap robustness, $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c})$.

If:

$$\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c}) > \widehat{h}_{\rm s}(\ell', \varepsilon_{\rm c}) > 0 \tag{86}$$

Then:

$$P_{\rm s}(\ell) \ge P_{\rm s}(\ell') \tag{87}$$

§

§ **Theorem.**

Probability of successful forecast, $P_{\rm s}(\ell)$,
increases with
increasing info-gap robustness, $\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c})$.

If:

$$\widehat{h}_{\rm s}(\ell, \varepsilon_{\rm c}) > \widehat{h}_{\rm s}(\ell', \varepsilon_{\rm c}) > 0 \tag{88}$$

Then:

$$P_{\rm s}(\ell) \geq P_{\rm s}(\ell') \tag{89}$$

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

One can maximize probability of success without knowing the probability distribution.

2.5 *Conclusion*

§ Forecasters do better if they robust-satisfice.

§ Robust-satisficing is not a last resort.
It is strategically advantageous.

Any Questions?