Supplements to the Article

“Data-Based Decisions under Imprecise Probability and Least Favorable Models”

Robert Hable

1 Introduction

The present text contains supplements to the article “Data-Based Decisions under Imprecise Probability and Least Favorable Models”([Hable (2007)]). The notation is clompetely adopted from [Hable (2007)].

2 Proof of Theorem 4.1

Let Θ be a finite index set. Let π be a prior distribution on $(\Theta, 2^{\Theta})$ so that $\pi_\theta := \pi[I_{\{\theta\}}] > 0 \;\; \forall\, \theta \in \Theta$. Let $(p_\theta)_{\theta\in\Theta}$ be a precise model on $(\mathcal{X}, \mathcal{A})$ and $(\overline{Q}_\theta)_{\theta\in\Theta}$ an imprecise model on $(\mathcal{Y}, \mathcal{B})$ where $(\mathcal{M}_\theta)_{\theta\in\Theta}$ is the corresponding family of structures. Let $s^{(p_\theta)_\theta}$ be the standard measure of $(p_\theta)_{\theta\in\Theta}$ and $\overline{S}$ the standard upper expectation of $(\overline{Q}_\theta)_{\theta\in\Theta}$ on $(\mathcal{U}, \mathcal{C})$.
Let Ψ be the set of all functions $k \in \mathcal{L}_\infty(\mathcal{U}, \mathcal{C})$ such that there is some decision space $(\mathbb{D}, \mathcal{D})$ and a loss function $(W_\theta)_{\theta\in\Theta} \subset \mathcal{L}_\infty(\mathbb{D}, \mathcal{D})$ where $k(u) = \inf_{\tau\in\mathbb{D}} \sum_{\theta\in\Theta} n\pi_\theta W_\theta(\tau)\iota_\theta(u) \;\; \forall\, u \in \mathcal{U}$.
[Hable (2007), Theorem 4.1] is the analogon to [Buja (1984), Theorem 7.1]. The arguments in the main steps of the proof are similar to that one given in [Buja (1984)].

Theorem 4.1 *The following statements are equivalent:*

(a) $(p_\theta)_{\theta\in\Theta}$ *is worst-case-sufficient for* $(\overline{Q}_\theta)_{\theta\in\Theta}$.

(b) $s^{(p_\theta)_\theta}[k] \;\le\; \overline{S}[k] \qquad \forall\, k \in \Psi$

(c) *For every finite decision space* $(\mathbb{D}, \mathcal{D})$ *and every loss function* $(W_\theta)_{\theta\in\Theta} \subset \mathcal{L}_\infty(\mathbb{D}, \mathcal{D})$,

$$\inf_{\rho\in\mathcal{T}(\mathcal{X},\mathbb{D})} R\big((p_\theta)_\theta, \rho, (W_\theta)_\theta\big) \;\le\; \inf_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big)$$

(d) *For every decision space* $(\mathbb{D}, \mathcal{D})$ *and every loss function* $(W_\theta)_{\theta\in\Theta} \subset \mathcal{L}_\infty(\mathbb{D}, \mathcal{D})$,

$$\inf_{\rho\in\mathcal{T}(\mathcal{X},\mathbb{D})} R\big((p_\theta)_\theta, \rho, (W_\theta)_\theta\big) \;\leq\; \inf_{\sigma\in\mathcal{T}(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big)$$

Proof: The proof has the following structure: (a)⇔(d), (d)⇔(c), (d)⇔(b)

(a)⇒(d): This is a direct consequence of [Hable (2007), Lemma 8.2].

(a)⇐(d): Put $\mathbb{D} = \mathcal{Y}$ and $\sigma_0(\mu) = \mu \;\; \forall\, \mu \in \mathrm{ba}(\mathcal{Y}, \mathcal{B})$. Then (d) implies that for all $(g_\theta)_{\theta\in\Theta} \subset \mathcal{L}_\infty(\mathcal{Y}, \mathcal{B})$,

$$\inf_{T\in\mathcal{T}(\mathcal{X},\mathcal{Y})} R\big((p_\theta)_\theta, T, (g_\theta)_\theta\big) \leq R\big((\overline{Q}_\theta)_\theta, \sigma_0, (g_\theta)_\theta\big)$$

This may be rewritten as

$$\inf_{T\in\mathcal{T}(\mathcal{X},\mathcal{Y})} \sum_{\theta\in\Theta} \pi_\theta \Big(T(p_\theta)[g_\theta] - \overline{Q}_\theta[g_\theta]\Big) \leq 0$$

for every $(g_\theta)_\theta \subset \mathcal{L}_\infty(\mathcal{Y}, \mathcal{B})$. Put $\Gamma\big(T, (g_\theta)_\theta\big) := \sum_{\theta\in\Theta} \pi_\theta \big(T(p_\theta)[g_\theta] - \overline{Q}_\theta[g_\theta]\big)$. Then,

$$\sup_{(g_\theta)_\theta\subset\mathcal{L}_\infty(\mathcal{Y},\mathcal{B})} \; \inf_{T\in\mathcal{T}(\mathcal{X},\mathcal{Y})} \Gamma\big(T, (g_\theta)_\theta\big) \;\leq\; 0 \tag{1}$$

$\mathcal{T}(\mathcal{X}, \mathcal{Y})$ is compact, $T \mapsto \Gamma\big(T, (g_\theta)_\theta\big)$ is continuous and konvex, $(g_\theta)_\theta \mapsto \Gamma\big(T, (g_\theta)_\theta\big)$ is concave. So, the minimax theorem [Fan (1953, Theorem 2] and (1) yield

$$\inf_{T\in\mathcal{T}(\mathcal{X},\mathcal{Y})} \; \sup_{(g_\theta)_\theta\subset\mathcal{L}_\infty(\mathcal{Y},\mathcal{B})} \Gamma\big(T, (g_\theta)_\theta\big) \;\leq\; 0$$

Compactness of $\mathcal{T}(\mathcal{X}, \mathcal{Y})$ and lower semicontinuity of

$$T \;\mapsto\; \sup_{(g_\theta)_\theta\subset\mathcal{L}_\infty(\mathcal{Y},\mathcal{B})} \Gamma\big(T, (g_\theta)_\theta\big)$$

imply the existence of some $T_0 \in \mathcal{T}(\mathcal{X}, \mathcal{Y})$ so that

$$\sup_{(g_\theta)_\theta\subset\mathcal{L}_\infty(\mathcal{Y},\mathcal{B})} \Gamma\big(T_0, (g_\theta)_\theta\big) \;\leq\; 0 \tag{2}$$

(cf. [McShane et al. (1959), Theorem 3.7]). Since $\pi_\theta > 0 \;\; \forall\, \theta \in \Theta$, it follows from (2) that

$$T_0(p_\theta)[g_\theta] \;\leq\; \overline{Q}[g_\theta] \qquad \forall\, g_\theta \in \mathcal{L}_\infty(\mathcal{Y}, \mathcal{B})$$

for every $\theta \in \Theta$.

(d)⇒(c): This is obvious.

(d)⇐(c): Let $\sigma : \mu \mapsto \kappa^*(\mu)$ be a restricted randomisation from $\mathcal{Y}$ to $\mathbb{D}$ where

$$\kappa^*(\mu)[g] = \mu\Big[\sum_{t\in D} g(t)\alpha_t\Big]$$

and D is a finite subset of $\mathbb{D}$. $(D, 2^D)$ may be regarded as a finite decision space and σ may be regarded as an element of $\mathcal{T}(\mathcal{Y}, D)$. Then, (c) implies

$$\inf_{\hat\rho\in\mathcal{T}(\mathcal{X},D)} R\big((p_\theta)_\theta, \hat\rho, (W_\theta)_\theta\big) \leq R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big) \tag{3}$$

Since every element of $\mathcal{T}_r(\mathcal{X}, D)$ may be regarded as an element of $\mathcal{T}_r(\mathcal{X}, \mathbb{D})$, [Hable (2007), Proposition 3.1] implies

$$\begin{aligned} \inf_{\rho\in\mathcal{T}(\mathcal{X},\mathbb{D})} R\big((p_\theta)_\theta, \rho, (W_\theta)_\theta\big) \;\leq& \\ \leq& \inf_{\hat\rho\in\mathcal{T}(\mathcal{X},D)} R\big((p_\theta)_\theta, \hat\rho, (W_\theta)_\theta\big) \end{aligned} \tag{4}$$

Hence, (according to [Hable (2007), Proposition 3.3]

$$\begin{aligned} \inf_{\rho\in\mathcal{T}(\mathcal{X},\mathbb{D})} R\big((p_\theta)_\theta, \rho, (W_\theta)_\theta\big) \;\leq& \\ \overset{(4),(3)}{\leq}& \inf_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big) \;= \\ =& \inf_{\sigma\in\mathcal{T}(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big) \end{aligned}$$

(d)⇔(b): This is a direct consequence of [Hable (2007), Theorem 3.2] and [Hable (2007), Theorem 3.4] □

3 Proof of Lemma 5.2

Let $(\overline{Q}_\theta)_{\theta\in\Theta}$ be an imprecise model on $(\mathcal{Y}, \mathcal{B})$ where $(\mathcal{M}_\theta)_{\theta\in\Theta}$ is the corresponding family of structures. For $\mathcal{F} \in (\mathcal{M}_\theta)_{\theta\in\Theta}$, put

$$\Phi_{\mathcal{F}} := \big\{ h \in \mathcal{L}_\infty(\mathcal{U}, \mathcal{C}) \;\big|\; s^{\mathcal{F}}[h] = \overline{S}[h] \big\}$$

where $s^{\mathcal{F}}$ is the standard measure of $\mathcal{F}$ and $\overline{S}$ is the standard upper expectation of $(\overline{Q}_\theta)_{\theta\in\Theta}$ on $(\mathcal{U}, \mathcal{C})$.

Lemma 5.2 *$\Phi_{\mathcal{F}}$ is a norm-closed convex cone in $\mathcal{L}_\infty(\mathcal{U}, \mathcal{C})$.*

Proof:

- For $h \in \Phi_{\mathcal{F}}$ and $c \in [0, \infty)$,

$$\overline{S}[ch] = c\overline{S}[h] = cs^{\mathcal{F}}[h] = s^{\mathcal{F}}[ch]$$

- For $h_1, h_2 \in \Phi_{\mathcal{F}}$,

$$\begin{aligned} \overline{S}[h_1 + h_2] \;\leq&\; \overline{S}[h_1] + \overline{S}[h_2] = s^{\mathcal{F}}[h_1] + s^{\mathcal{F}}[h_2] = \\ =&\; s^{\mathcal{F}}[h_1 + h_2] \leq \overline{S}[h_1 + h_2] \end{aligned}$$

- For $(h_m)_{m\in\mathbb{N}} \subset \Phi_{\mathcal{F}}$, $\lim_m \|h_m - h\| = 0$ and $h \in \mathcal{L}_\infty(\mathcal{U}, \mathcal{C})$,

$$\begin{aligned} \overline{S}[h] \;\leq&\; \limsup_m \big(\overline{S}[h_m] + \overline{S}[h - h_m] \big) = \\ =&\; \limsup_m s^{\mathcal{F}}[h_m] = s^{\mathcal{F}}[h] \end{aligned}$$

  i.e. $s^{\mathcal{F}}[h] = \overline{S}[h]$.

□

4 Proof of Lemma 8.3

Let Θ be a finite index set. Let π be a prior distribution on $(\Theta, 2^\Theta)$ and $\pi_\theta := \pi[I_{\{\theta\}}] \quad \forall\theta \in \Theta$. Let $(p_\theta)_{\theta\in\Theta}$ be a precise model on $(\mathcal{X}, \mathcal{A})$ and $(\overline{Q}_\theta)_{\theta\in\Theta}$ an imprecise model on $(\mathcal{Y}, \mathcal{B})$ where $(\mathcal{M}_\theta)_{\theta\in\Theta}$ is the corresponding family of structures.
Let $(\mathbb{D}, \mathcal{D})$ be a decision space with loss function $(W_\theta)_{\theta\in\Theta} \subset \mathcal{L}_\infty(\mathbb{D}, \mathcal{D})$

Lemma 8.3

(a) $$\inf_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big) \;=\; \sup_{(q_\theta)_\theta\in(\mathcal{M}_\theta)_\theta} \inf_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} R\big((q_\theta)_\theta, \sigma, (W_\theta)_\theta\big)$$

(b) $$\inf_{\sigma\in\mathcal{T}(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big) \;=\; \sup_{(q_\theta)_\theta\in(\mathcal{M}_\theta)_\theta} \inf_{\sigma\in\mathcal{T}(\mathcal{Y},\mathbb{D})} R\big((q_\theta)_\theta, \sigma, (W_\theta)_\theta\big)$$

Proof:

(a) $\prod_{\theta\in\Theta}\mathcal{M}_\theta$ is a compact Hausdorff space (cf. [Hable (2007), Theorem 2.1] and [Dunford et al. (1957), Lemma V.3.3, Lemma I.8.2 and Theorem I.8.5]. For every $\sigma \in \mathcal{T}_r(\mathcal{Y}, \mathbb{D})$ there is some $\kappa : \mathcal{L}_\infty(\mathcal{Y}, \mathcal{B}) \to \mathcal{L}_\infty(\mathcal{X}, \mathcal{A})$ so that $\sigma(\mu)[g] = \mu[\kappa(g)]$ for every $g \in \mathcal{L}_\infty(\mathcal{Y}, \mathcal{B}),\ \mu \in \mathrm{ba}(\mathcal{Y}, \mathcal{B})$. Hence,

$$\mathcal{M}_\theta \;\to\; \mathbb{R}\,, \qquad q_\theta \;\mapsto\; \sigma(q_\theta)[W_\theta]$$

is continuous for every $\theta \in \Theta$ and this implies continuity of the map

$$(q_\theta)_\theta \;\mapsto\; -\sum_{\theta\in\Theta} \pi_\theta\sigma(q_\theta)[W_\theta] =: \Gamma\big((q_\theta)_\theta, \sigma\big)$$

on $\prod_{\theta\in\Theta}\mathcal{M}_\theta$ for every $\sigma \in \mathcal{T}_r(\mathcal{Y}, \mathbb{D})$. $(q_\theta)_\theta \mapsto \Gamma\big((q_\theta)_\theta, \sigma\big)$ is convex on $\prod_{\theta\in\Theta}\mathcal{M}_\theta$ for every $\sigma \in \mathcal{T}_r(\mathcal{Y}, \mathbb{D})$ and $\sigma \mapsto \Gamma\big((q_\theta)_\theta, \sigma\big)$ is concave on $\mathcal{T}_r(\mathcal{Y}, \mathbb{D})$ for every $(q_\theta)_\theta \in \prod_{\theta\in\Theta}\mathcal{M}_\theta$.
Then, the minimax theorem [Fan (1953, Theorem 2] yields

$$\begin{aligned}
\inf_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big) \;&=\; -\sup_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} \inf_{(q_\theta)_\theta\in(\mathcal{M}_\theta)_\theta} \Gamma\big((q_\theta)_\theta, \sigma\big) = \\
&= -\inf_{(q_\theta)_\theta\in(\mathcal{M}_\theta)_\theta} \sup_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} \Gamma\big((q_\theta)_\theta, \sigma\big) = \\
&= \sup_{(q_\theta)_\theta\in(\mathcal{M}_\theta)_\theta} \inf_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} R\big((q_\theta)_\theta, \sigma, (W_\theta)_\theta\big)
\end{aligned}$$

(b) [Hable (2007), Proposition 3.1] and part (a) of the present lemma yield

$$\begin{aligned}
&\inf_{\sigma\in\mathcal{T}(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big) \geq \\
&\quad\geq \sup_{(q_\theta)_\theta\in(\mathcal{M}_\theta)_\theta} \inf_{\sigma\in\mathcal{T}(\mathcal{Y},\mathbb{D})} R\big((q_\theta)_\theta, \sigma, (W_\theta)_\theta\big) = \\
&\quad= \sup_{(q_\theta)_\theta\in(\mathcal{M}_\theta)_\theta} \inf_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} R\big((q_\theta)_\theta, \sigma, (W_\theta)_\theta\big) = \\
&\quad\overset{(a)}{=} \inf_{\sigma\in\mathcal{T}_r(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big) \geq \inf_{\sigma\in\mathcal{T}(\mathcal{Y},\mathbb{D})} R\big((\overline{Q}_\theta)_\theta, \sigma, (W_\theta)_\theta\big)
\end{aligned}$$

□

References

[Buja (1984)] Buja, A. (1984). *Simultaneously Least Favorable Experiments, Part I: Upper Standard Functionals and Sufficiency.* Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 65, 367–384.

[Dunford et al. (1957)] Dunford, N., Schwartz, J.T. (1957). *Linear Operators, Part I: General Theory.* Wiley-Interscience, New-York.

[Fan (1953] Fan, K. (1953). *Minimax theorems.* Proceedings of the National Academy of Sciences of the United States of America 39, 42–47.

[Hable (2007)] Hable, R. (2007). *Data-Based Decisions under Imprecise Probability and Least Favorable Models.* Submitted to: 5th International Symposium on Imprecise Probabilities and Their Applications, Prague.

[McShane et al. (1959)] McShane, E.J., Botts, T.A. (1959). *Reals analysis.* Van Nostrand, Princeton.