Powerful algorithms for decision making under partial prior information and general ambiguity attitudes

Lev Utkin

Department of Computer Science,
St. Petersburg Forest Technical Academy

Thomas Augustin

Department of Statistics; SFB 386
University of Munich

Decision making under incomplete data using the imprecise Dirichlet model

Lev Utkin

Department of Computer Science,
St. Petersburg Forest Technical Academy

Thomas Augustin

Department of Statistics; SFB 386
University of Munich

1. The basic decision problem

- Comprehensive framework
  - * Actions $a_i \in \mathbb{A}$ (treatment; investment)
  - * states of nature $\vartheta_j \in \Theta$ (disease; development of economy)
  - * utility $u(a_i, \vartheta_j) \Longrightarrow$ random variable $\mathbf{u}(a)$
- Find optimal action(s)!
- When everything is finite: utility table

	ϑ_1	$\dots$	ϑ_j	$\dots$	ϑ_m
a_1	$u(a_1, \vartheta_1)$		$\dots$		$u(a_1, \vartheta_m)$
$\vdots$		$\ddots$			
a_i	$\vdots$		$u(a_i, \vartheta_j)$		
$\vdots$				$\ddots$	
a_n	$u(a_n, \vartheta_1)$		$\dots$		$u(a_n, \vartheta_m)$

2. Classical Decision criteria

- Randomized actions: $\lambda(a_i)$ probability to take action a_i

Two classical criteria:

- Bayes optimality
  - $*$ perfect probabilistic knowledge: prior $\pi(\cdot)$ on Θ
  - $*$ maximize expected utility $\mathbb{E}_\pi \mathbf{u}(a) \longrightarrow \max_a$
- Maximin (Wald) optimality
  - $*$ complete ignorance $\Longrightarrow$ focus on the worst state:

$$\min_j \; u(a, \vartheta_j) \longrightarrow \max_a$$

What to do in the case of **partial** prior knowledge?

3. Decision criteria under partial knowledge

- $\mathcal{M}$ convex polyhedron of classical probabilities (e.g. structure of F-probability); $\mathcal{E}(\mathcal{M})$ set of vertices

- $\mathcal{M} = \{\pi(\cdot) | \underline{b}_l \leq \mathbb{E}_\pi f_l \leq \overline{b}_l\} \quad l = 1, \dots, r$

- interval-valued expected utility:

$$\mathbb{E}_{\mathcal{M}}\mathbf{u}(a) := [\underline{\mathbb{E}}_{\mathcal{M}}\mathbf{u}(a),\ \overline{\mathbb{E}}_{\mathcal{M}}\mathbf{u}(a)]$$

$$:= [\inf_{\pi \in \mathcal{M}} \mathbb{E}_\pi \mathbf{u}(a),\ \sup_{\pi \in \mathcal{M}} \mathbb{E}_\pi \mathbf{u}(a)]$$

- axiomatic justifications!

Some Criteria

(Survey: Troffaes (SIPTA-NL, Dec 2004))

- linear ordering
  - a) $\underline{\mathbb{E}}_{\mathcal{M}} u(a) \longrightarrow \max_a$
    - Γ-Maximin
    - Choquet-Integral
  - b) $\eta \cdot \underline{\mathbb{E}}_{\mathcal{M}} u(a) + (1-\eta)\overline{\mathbb{E}}_{\mathcal{M}} u(a) \longrightarrow \max_a$
    - Caution parameter η
- partial ordering
  - c) E-admisibility (Levi)
    - ⋮
    - Maximality (Walley)

4. Calculation of optimal actions

- Far from being straightforward; lack of feasible algorithms
- has hindered large scale applications
- Formulation in terms of linear programming problems also provides theoretical insight.
- Two different situations considered here
  * direct assessment of $\mathcal{M}$ (e.g. by an expert)
  * construction of partial knowledge based on previous observations on Θ (repeated decision problems)

4. a) Pessimistic descision making: Gamma-Maximin

- Bayes and minimax optimality as border cases
- Gamma-Minimax criterion (e.g., Berger (1984[12], Springer), Vidakovic (2000, Insua, D.R., and Ruggeri, F. (eds.))
- Maxmin expected utility model (Gilboa, Schmeidler (1989, Journal of Mathematical Economics)
- MaxEMin (Kofler, Menges (1976)) (cf. also Kofler (1989, Campus) and the references therein)
- maximinity (Walley (1991, Chapman Hall))
- In the case of two-monotonicty: Choquet expected utility (e.g., Chateauneuf, Cohen, Meilijson (1991, Finance))

$$
\begin{aligned}
&\mathbb{E}_{\mathcal{M}}\mathbf{u}(\lambda) \longrightarrow \max_{\lambda} \\
\Longleftrightarrow\ &\min_{\pi\in\mathcal{M}} \sum_{j=1}^{m}\left(\sum_{i=1}^{n} u(a_i,\vartheta_j)\lambda(a_i)\right)\pi(\{\vartheta_j\}) \longrightarrow \max_{\lambda} \\
&\text{subject to } \sum_{i=1}^{n}\lambda(a_i) = 1; \quad \lambda(a_i) \geq 0 \\
\Longleftrightarrow\ & \\
&G \longrightarrow \max \\
&\text{subject to } \sum_{i=1}^{n}\lambda(a_i) = 1; \quad \lambda(a_i) \geq 0 \\
&\sum_{j=1}^{m}\left(\sum_{i=1}^{n} u(a_i,\vartheta_j)\lambda(a_i)\right)\pi(\{\vartheta_j\}) \geq G, \quad \forall \pi \in \mathcal{M}.
\end{aligned} \tag{6}
$$

$\iff$ Augustin (2002, Stat. Pap.), Augustin (2004, ZAMM).

$$G \longrightarrow \max$$

$$\text{subject to} \quad \sum_{i=1}^{n} \lambda(a_i) = 1; \quad \lambda(a_i) \geq 0 \quad \text{and}$$

$$\sum_{j=1}^{m} \left(\sum_{i=1}^{n} u(a_i, \vartheta_j) \lambda(a_i) \right) \pi(\{\vartheta_j\}) \geq G, \quad \forall \pi \in \mathcal{E}(\mathcal{M}).$$

- needs, however, all vertices to be determined in advance
- In case of F-probability: $|\mathcal{E}(\mathcal{M})|$ may be as large as $m!$ (Wallner (2005, ISIPTA))
- considerable simplification in the case of two-monotoncity

Alternative: partial dualization

- $\min_{\pi \in \mathcal{M}} \sum_{j=1}^{m} \left(\sum_{i=1}^{n} u(a_i, \vartheta_j) \lambda(a_i) \right) \pi(\{\vartheta_j\}) \to \max$
  subject to $\lambda \cdot \mathbf{1} = 1$.

- Fix λ, and consider the dual problem of

$$\sum_{j=1}^{m} \left(\sum_{i=1}^{n} u(a_i, \vartheta_j) \lambda(a_i) \right) \pi(\{\vartheta_j\}) \to \min_{\pi \in \mathcal{M}}$$

With $\mathbf{C} = (c_1, ..., c_r)^{\mathrm{T}}, \quad \mathbf{D} = (d_1, ..., d_r)^{\mathrm{T}}$:

$$\max_{c, \mathbf{C}, \mathbf{D}} \left\{ c + \underline{\mathbf{B}}\mathbf{C} - \overline{\mathbf{B}}\mathbf{D} \right\}$$

subject to $c \in \mathbb{R}$, $\mathbf{C}, \mathbf{D} \in \mathbb{R}^r_+$, and

$$c + \mathbf{F}_j (\mathbf{C} - \mathbf{D}) \leq \sum_{i=1}^{n} u(a_i, \vartheta_i) \lambda(a_i), \quad j = 1, ..., m.$$ [1]

[1] Here $c, \mathbf{C}, \mathbf{D}$ are optimization variables such that the variable c corresponds to the constraint $\sum_{j=1}^{m} \pi_j = 1$ in the primal form, c_i corresponds to the constraints $\underline{b}_i \leq E_\pi f_i$ and d_i corresponds to the constraints $E_\pi f_i \leq \overline{b}_i$.

- By the general theory, the values at the optima coincide

$$\min_{\pi \in \mathcal{M}} \sum_{j=1}^{m} \left(\sum_{i=1}^{n} u(a_i, \vartheta_j) \lambda(a_i) \right) \pi(\{\vartheta_j\}) = \max_{c, \mathbf{C}, \mathbf{D}} \left\{ c + \underline{\mathbf{B}} \mathbf{C} - \overline{\mathbf{B}} \mathbf{D} \right\},$$

- Then the additional maximization over λ gives the optimal action:

$$\max_{c, \mathbf{C}, \mathbf{D}, \lambda} \left\{ c + \underline{\mathbf{B}} \mathbf{C} - \overline{\mathbf{B}} \mathbf{D} \right\}$$

subject to $c \in \mathbb{R}$, $\mathbf{C}, \mathbf{D} \in \mathbb{R}^r_+$, $\lambda \cdot \mathbf{1} = 1$ and

$$c + \mathbf{F}_j (\mathbf{C} - \mathbf{D}) \leq \sum_{i=1}^{n} u(a_i; \vartheta_j) \lambda(a_i), \;\; j = 1, ..., m.$$

- Note: single linear programming problem, the vertices are not needed

4 b) Caution parameter η

- More sophisticated representations of interval-valued expected utility to avoid overpessimism
- take additionally into consideration the decision maker's attitude towards ambiguity, e.g.:
- Ellsberg (1961, QJE)
- Jaffray (1989, OR Letters)
- Schubert (1995, IJAR)
- Weichselberger (2001, Physika, Chapter 2.6)
- Weichselberger and Augustin (1998, Galata and Küchenhoff (eds.))

- Criterion:

$$\eta\underline{\mathbb{E}}_{\mathcal{M}}\mathbf{u}(\lambda) + (1-\eta)\overline{\mathbb{E}}_{\mathcal{M}}\mathbf{u}(\lambda) \to \max_{\lambda}$$

- Same tricks can not be applied again: unbounded solutions

- Ensure that in the previous systems some inequalities are equalities $\Rightarrow$ several optimization problems to be solved

- Alternatively, in the approach based on the vertices, consider for every $\widetilde{\pi} \in \mathcal{E}(\mathcal{M})$ the objective function

$$\eta \cdot G + (1-\eta)\sum_{j=1}^{m}\left(\sum_{i=1}^{n} u(a_i, \vartheta_j)\lambda(a_i)\right)\widetilde{\pi}(\vartheta_j) \to \max$$

and maximize over all elements of $\mathcal{E}(\mathcal{M})$

4 c) E-admissibility (and maximality)

- E-admissibility (e.g., Levi (1974, J Phil), Schervish et al. (2003, ISIPTA)):

- Consider all actions that are not everywhere suboptimal :

  $\exists \pi_{a^*} \in \mathcal{M}$ such that a^* is Bayes with respect to π_{a^*}:

$$\sum_{j=1}^{m} u(a^*, \vartheta_j)\pi_{a^*}(\vartheta_j) \geq \sum_{j=1}^{m} u(a, \vartheta_j)\pi_{a^*}(\vartheta_j), \ \forall a \in \mathbb{A}$$

Lemma 1 (Characterization of Bayes actions in classical decision therory) Fix $\pi(\cdot)$ and let $\mathbb{A}_\pi^*$ be the set of all pure Bayes actions with respect to $\pi(\cdot)$, and Λ_π^* the set of all randomized Bayes actions with respect to $\pi(\cdot)$. Then

i) $\mathbb{A}_\pi^* \neq \emptyset$
ii) $\Lambda_\pi^* = \mathrm{conv}(\mathbb{A}_\pi^*)$.[2]

Proof: The task of finding a Bayes action with respect to $\pi(\cdot)$ can be written as a linear programming problem

$$\sum_{j=1}^{m} \left(\sum_{i=1}^{n} u(a_i, \vartheta_j)\, \lambda(a_i) \right) \pi(\vartheta_j) \longrightarrow \max_{\lambda}$$

subject to $\sum_{i=1}^{n} \lambda(a_i) = 1$, and $\lambda(a_i) \geq 0$, for all i.

i) One optimal solution must be attained at a vertex.
ii) Convexity of the set of optimal solutions.

[2] Here every pure action $a_i \in \mathbb{A}$ is identified with the randomized action $\lambda(a) = 1$ if $a = a_i$ and $\lambda(a) = 0$ else, and with the corresponding $(n \times 1)$ vector.

A general algorithm for E-admissibility

- Turn the problem around!
  Now fix the actions!

- For every a_i look at

  $\Pi_i := \{\pi(\cdot) \in \mathcal{M} | \ a_i$ is Bayes action with respect to $\pi(\cdot)\}$
  According to Lemma 1:

$$\Pi_i = \left\{ \pi(\cdot) \in \mathcal{M} \ \middle| \ \sum_{j=1}^{m} u(a_i, \vartheta_j)\, \pi(\vartheta_j) \geq \sum_{j=1}^{m} u(a_l, \vartheta_j)\, \pi(\vartheta_j), \quad \forall\, l = 1, \ldots, n \right\}$$

-
$$\Pi_i = \operatorname{conv}\left(\widetilde{\pi}(\cdot) \in \mathcal{E}(\mathcal{M}) \,\middle|\, \sum_{j=1}^{m} u(a_i, \vartheta_j)\widetilde{\pi}(\vartheta_j) \geq \sum_{j=1}^{m} u(a_l, \vartheta_j)\widetilde{\pi}(\vartheta_j),\ \forall l = 1, \ldots, n\right).$$

- Alternatively, without using $\mathcal{E}(\mathcal{M})$:

$$z \longrightarrow \max_{(\pi^T, z)^T}$$

$$\sum_{j=1}^{m} u(a_i, \vartheta_j)\, \pi(\vartheta_j) \;\geq\; \sum_{j=1}^{m} u(a_l, \vartheta_j)\, \pi(\vartheta_j), \qquad \forall l = 1, ..., n$$

$$\sum_{j=1}^{m} \pi(\vartheta_j) = z\,,\;\; z \leq 1\,, \quad \pi(\vartheta_j) \geq 0\,,\;\; j = 1, \ldots, m\,,$$

$$\underline{b}_l \leq \sum f_l(\vartheta_j)\pi(\vartheta_j) \leq \overline{b}_l\,,\;\; l = 1, \ldots, r\,.$$

- Iff $z = 1$ then $\Pi_i \neq 0$ and a_i is E-admissible

- To determine all E-admissible pure actions: $|\mathbb{A}|$ linear optimization problems have to be solved

- By Lemma 1 ii) adaption possible to calculate *all* E-admissible actions:
  For all $I \subseteq \{1, \ldots, m\}$ check whether there is a prior π under which all a_i, $i \in I$, are simultaneously optimal, i.e. replace (23) by

$$\Pi_I := \left\{ \pi(\cdot) \middle| \sum_{j=1}^{m} u(a_i, \vartheta_j)\, \pi(\vartheta_j) \geq \sum_{j=1}^{m} u(a_l, \vartheta_j)\, \pi(\vartheta_j), \right.$$
$$\left. \forall i \in I,\, l = 1, \ldots, n. \right\}$$

  If Π_I is not empty, then all the elements of $\mathrm{conv}(a_i | i \in I)$ are E-admissible actions.

- If $\Pi_I = \varnothing$ for some I then all index sets $J \supset I$ need not be considered anymore.

maximality

- If Π_i contains π with $\pi(\cdot) > 0$, then a_i is admissible in the classical sense and therefore maximal.

- But if $\mathbb{A}$ is not convex, not all maximal actions are found in that way.

- uniform optimality of a_{i^*}:

  If $\Pi_i = \mathcal{M}$ then $\mathbb{E}_\pi \mathbf{u}(a_{i^*}) \geq \mathbb{E}_\pi \mathbf{u}(a), \quad \forall \pi \in \mathcal{M}, \quad a \in \mathbb{A}.$

  (cp. Weichselberger (2001, Chapter 2.6): structure dominance)

Now (second paper) data on $\theta_1, \ldots \theta_n$

- n_j observations of $\theta_j, \ j = 1, \ldots n$.
- more general: set-valued observations $\subseteq \Theta$
- calculate expected utility based on estimates $\widehat{\pi(\theta_j)}$ resulting from the data

modeling data	naive
regular	relative frequencies
set-valued	empirical belief functions, “random sets”, (e.g. S.Maier (2004, Univ. Munich), Tonon et al. (2000, RESS))

- n_j observations of θ_j, $j = 1, \ldots n$.
- more general: set-valued observations $\subseteq \Theta$
- calculate expected utility based on estimates $\widehat{\pi(\theta_j)}$ obtained from the data

modeling / data	naive
regular	relative frequencies
set valued	empirical belief functions, “random sets”, S.Maier (2004, Univ. Munich)
	but amount of data is not reflected; no difference whether 1 or 10^6 observations

- n_j observations of $\theta_j, \; j = 1, \ldots n$.
- more general: set-valued observations $\subseteq \Theta$
- calculate expected utility based on estimates $\widehat{\pi(\theta_j)}$ obtained from the data

modeling / data	naive	
regular	relative frequencies	IDM
set valued	empirical belief functions, "random sets", S.Maier (2004, Univ. Munich)	extended empirical belief functions, Utkin (2005, FSS)
	but amount of data is not reflected; no difference whether 1 or 10^6 observations	

To calculate optimal actions

- use previous techniques or
- considerable simplications due to the use of the IDM and belief functions: With Möbius inverse $m(\cdot)$

  $$\mathbb{E}\,\mathbf{u}(a) = \Big[\sum_{A \subseteq \Theta}^{m} m(A) \cdot \min_{\theta \in A} u(a,\theta); \quad \sum_{A \subseteq \Theta}^{m} m(A) \cdot \max_{\theta \in A} u(a,\theta)\Big]$$

  Chateauneuf and Jaffray (1989, Math. Social Sc.; Cor.4), Strat (1991, IJAR)
- Leads to a frequency-based Hodges-Lehman criterion
- Be careful when specifying Θ! The embedding principle is not valid in decision theory based on the IDM.

The (Imprecise) Dirichlet Model in decision making

- N multinomial observations on space Ω, Dirichlet prior with parameter S, $\mathbf{t} = (t_1, \ldots t_m)$

- For every $A \subseteq \Omega$ predictive probability
$$P(A|\mathbf{n}, \mathbf{t}, s) = \frac{\sum_{\omega_j \in A} n_j + s \cdot \sum_{\omega_j \in A} t_j}{N + s}$$

- Walley (1996, JRSSB): Consider *all* Dirichlet priors, i.e. vary $\mathbf{t} \in S(1, m)$
$$P(A|\mathbf{n}, \mathbf{t}, s) = \left[\frac{\sum_{\omega_j \in A} n_j}{N + s}; \frac{s + \sum_{\omega_j \in A} n_j}{N + s}\right]$$

- In decision making based on certain value of $\mathbf{t}$

$$E_{\mathbf{t}}\mathbf{u}(\lambda) = \int S(1,m)\sum_{i=1}^{m}\left(u(\lambda,\omega_i)\cdot\pi_i\right)p(\pi)\mathrm{d}\pi$$

$$= \sum_{i=1}^{m} u(\lambda,\omega_i)\cdot\int S(1,m)\pi_i\cdot p(\pi)\mathrm{d}\pi = \sum_{i=1}^{m} u(\lambda,\omega_i)\cdot\mathbb{E}_p\pi_i,$$

where
$$\mathbb{E}_p\pi_i = \frac{n_i + st_i}{N+s}, \qquad (1)$$

finally resulting in
$$\mathbb{E}_{\mathbf{t}}\mathbf{u}(\lambda) = \sum_{i=1}^{m} u(\lambda,\omega_i)\frac{n_i+st_i}{N+s}. \qquad (2)$$

- For the IDM

$$\mathbb{E}\mathbf{u}(\lambda) := \left[\underline{\mathbb{E}}\mathbf{u}(\lambda), \overline{\mathbb{E}}\mathbf{u}(\lambda)\right] := \left[\inf_{\mathbf{t}\in S(1,m)}\mathbb{E}_{\mathbf{t}}\mathbf{u}(\lambda), \sup_{\mathbf{t}\in S(1,m)}\mathbb{E}_{\mathbf{t}}\mathbf{u}(\lambda)\right]$$

Optimal actions in the case of pessimistic decision making

$$\underline{\mathbb{E}}\mathbf{u}(\lambda) \longrightarrow \max_{\lambda}$$

- use previous approaches or:

- for randomized actions solve

$$G \longrightarrow \max_{\lambda}$$

subject to $G \in \mathbb{R}$, $\lambda \cdot \mathbf{1} = 1$, and for $j = 1, ..., m$,

$$G \leq \frac{1}{N+s} \sum_{r=1}^{n} \lambda(a_r) \left(s \cdot u(a_r, \vartheta_j) + \sum_{j=1}^{m} u(a_r, \vartheta_j) \cdot n_j \right) .$$

- for pure actions

$$\left(\sum_{j=1}^{m} u(a_r, \vartheta_j) \cdot n_j + s \cdot \min_{j=1,\ldots m} u(a_r, \theta_j) \right) \longrightarrow \max_r$$

$$\iff \frac{N}{N+s} \cdot (\text{MEU based on } \frac{n_i}{N}) + \frac{S}{N+s} \cdot (\text{Wald criterion})$$

$N \longrightarrow \infty$ maximum expected utility (MEU)

$N = 0$ Wald

Incomplete data

- coarse data, set-valued observations
  make no additional assumptions (like CAR (Heitjan and Rubin (1991, Ann. Stat.), Blumenthal (1968, JASA)))
  $\Longrightarrow$ extended empirical belief functions (Utkin (2005, FSS))

- c_i observations of $A_i \subseteq \Omega$, $i = 1, \ldots M$ such that $\sum_{i=1}^{M} c_i = N$; $\mathbf{c} := (c_1, \ldots, c_M)$

- leads to several IDM's with observations
  $\mathbf{n}^{(k)} = (n_1^{(k)} \ldots, n_m^{(k)})$, $k = 1, \ldots, K$.
  (cp. also de Cooman and Zaffalon (2004, AI), Zaffalon (2002, JSPI))

- for fixed $\mathbf{t}$

$$\underline{P}(A|\mathbf{c}, s) = \frac{\min_k \sum_{\omega_j \in A} n_j^{(k)} + s \cdot \sum_{\omega_j \in A} t_j}{N + s}$$

$$\overline{P}(A|\mathbf{c}, s) = \frac{\max_k \sum_{\omega_j \in A} n_j^{(k)} + s \cdot \sum_{\omega_j \in A} t_j}{N + s}$$

- vary $\mathbf{t} \in S(1, m)$

$$\underline{P}(A|\mathbf{c}, s) = \frac{\sum_{i:A_i \subseteq A} c_i}{N + s}, \quad \overline{P}(A|\mathbf{c}, s) = \frac{\sum_{i:A_i \cap A \neq \varnothing} c_i + s}{N + s}.$$

Relation to empirical belief functions/random sets

- Empirical belief functions: set $m(A_i) = \frac{c_i}{N}$.
- Naive approach does not reflect the sample size,
- leads to $Bel_{emp}(\cdot)$ and $Pl_{emp}(\cdot)$
- Extended empirical belief functions can be written as

$$\underline{P}(A|\mathbf{c}, s) = \frac{N \cdot Bel_{emp}(A)}{N+s}, \quad \overline{P}(A|\mathbf{c}, s) = \frac{N \cdot Pl_{emp}(A) + s}{N+s}$$

with Möbius inverse

$$m^*(A_i) = \frac{c_i}{N+s}; \quad m^*(A_\infty) = \frac{s}{N+s}.$$

Optimal randomized actions (with $J_i := \{j | \omega_j \in A_i\}$)

$$\frac{1}{N+s}\left(s \cdot V_0 + \sum_{k=1}^{M} c_k \cdot V_k\right) \to \max_{\lambda},$$

subject to $V_0, V_i \in \mathbb{R}$, $\lambda \cdot \mathbf{1} = 1$.

$$V_i \leq \sum_{r=1}^{n} u(a_r, \omega_j) \cdot \lambda(a_r), \; i = 1, \ldots M, \; j \in J_i$$

$$V_0 \leq \sum_{r=1}^{n} u(a_r, \omega_j) \cdot \lambda(a_r), \; i = 1, \ldots m.$$

Optimal pure actions

$$\frac{1}{N+s}\left(s \cdot \min_j u(a_r, \omega_j) + \sum_{k=1}^{M} c_k \cdot \min_{\omega_j \in A_k} u(a_r, \omega_j)\right) \longrightarrow \max_{r=1,\ldots n}$$

Concluding remarks

- Other optimality criteria
- Alternative approach:

  incorporate sampling information by considering decision functions (not equivalent under IP (cp. Augustin (2003, ISIPTA), Halpern and Grünwald (2004, UAI), Jaffray (1999, ISIPTA), Seidenfeld (2004, Synthese)))
- Alternative models to learn from multinomial data:

  inference within the frame of Weichselberger's (e.g. 2005, ISIPTA) theory of symmetric probability or circular–$A(n)$–based inference: (Coolen and Augustin (2005, ISIPTA)).