On confident outrankings with multiple criteria of uncertain significance

Raymond Bisdorff

Université du Luxembourg
FSTC/ILAS

DA2PL'20014
Paris, 21 November 2014

Motivation

- When modelling preferences following the outranking approach, the sign of the majority margins do sharply distribute validation as well as invalidation of pairwise outranking situations.
- How can we be confident in the resulting outranking digraph, when we acknowledge
  1. the usual imprecise knowledge of criteria significance weights, and
  2. a small majority margin?

Content

Modelling uncertain criteria significances

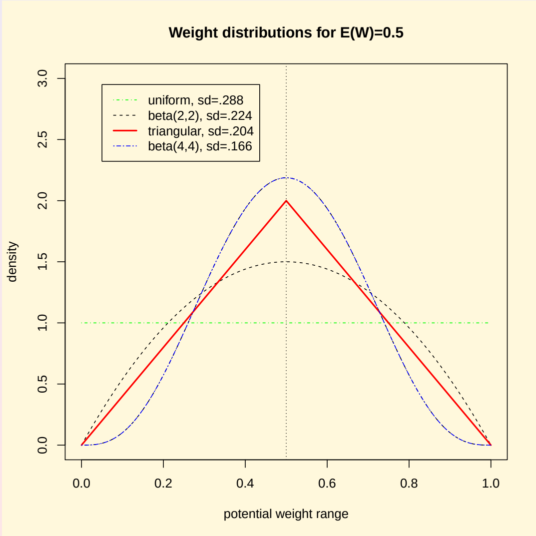
We consider the criterion significance weight to be independent random variable W, distributing the potential significance weight of the given criterion around a mean value $E(W)$ with variance $V(W)$.

1. A continuous *uniform* distribution on the range 0 to $2 * E(W)$. Thus $W \sim \mathcal{U}(0, 2E(W))$ and $V(W) = \frac{1}{3}E(W)^2$;
2. A symmetric *beta*(a, b) distribution with, for instance, parameters $a = 2$ and $b = 2$. Thus, $W \sim \mathcal{B}eta(2, 2) \times 2E(W)$ and $V(W) = \frac{1}{5}E(W)^2$.
3. A symmetric *triangular* distribution on the same range with mode $E(W)$. Thus $W \sim \mathcal{T}r(0, 2E(W), E(W))$ with $V(W) = \frac{1}{6}E(W)^2$;
4. A narrower *beta*(a, b) distribution with for instance parameters $a = 4$ and $b = 4$. Thus $W \sim \mathcal{B}eta(4, 4) \times 2E(W)$, $V(W) = \frac{1}{9}E(W)^2$

Decreasing uncertainty

The four potential uncertainty models all admit the same expected value, $E(W)$.

However, with a respective standard deviation which goes decreasing from $\sqrt{1/3} = 0.58$, to $\sqrt{1/9} = 1/3$ of $|E(W)|$.



Content

Notation

- $A = \{x, y, z, ...\}$: a finite set of n potential decision actions;
- $F = \{1, ..., n\}$, a finite and coherent family of m performance criteria;
- $[0; M_j]$: Performance measurement scale used on criterion j;
- ind_j: Upper-closed indifference threshold;
- pr_j: Lower-closed preference threshold with $0 \leqslant ind_j < pr_j \leqslant M_j$;
- x_j: The marginal performance of any object x on criterion j;
- W_j: The random rational significance weight of criterion j.

Performing marginally *at least as good as*

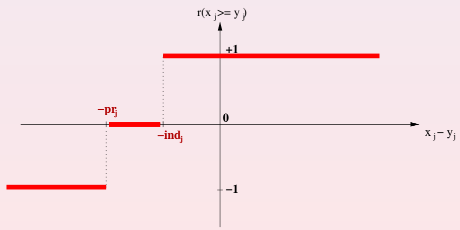
Each criterion j is characterizing a marginal double threshold order $\succcurlyeq_j$ on A in the following way:

$$r(x \succcurlyeq_j y) = \begin{cases} +1 & \text{if} \quad x_j - y_j \geqslant -ind_j \\ -1 & \text{if} \quad x_j - y_j \leqslant -pr_j \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

$+1$ signifies x is *performing at least as good as y* on criterion j,

-1 signifies that x is *not performing at least as good as y* on criterion j.

0 signifies that it is *unclear* whether, on criterion j, x is performing at least as good or not as y.



Performing globally "*at least as good as*"

Each criterion j contributes the random significance W_j of his marginal "*at least as good as*" characterization $r(\succcurlyeq_j)$ to the global characterization $\tilde{r}(\succcurlyeq)$ in the following way:

$$\tilde{r}(x \succcurlyeq y) \;=\; \sum_{j \in F} \left[\, W_j \cdot r(x \succcurlyeq_j y) \,\right] \qquad (2)$$

- $\tilde{r} > 0$ signifies x is *globally performing at least as good as* y,
- $\tilde{r} < 0$ signifies that x is *not globally performing at least as good as* y,
- $\tilde{r} = 0$ signifies that it is *unclear* whether x is globally performing at least as good or not as y.

Likelihood of "*at least as good as*" situations

From the Central Limit Theorem (CLT), we know that $\tilde{r}(x \succcurlyeq y)$ (Eq. 2) leads, with m getting large, to a Gaussian variable Y with:

$$E(Y) \;=\; \sum_j E(W_j) \times r(x \succcurlyeq_j y),$$

$$V(Y) \;=\; \sum_j V(W_j) \times |r(x \succcurlyeq_j y)|.$$

Hence, the bipolar likelihood (lh) of *validation*, respectively *invalidation* of a $(x \succcurlyeq y)$ situation may be assessed as follows:

$$lh(x \succcurlyeq y) \;=\; 2 \times P(Y > 0.0) - 1.0 \;=\; -\,\mathrm{erf}\left(\frac{1}{\sqrt{2}} \frac{-E(Y)}{\sqrt{V(Y)}}\right).$$

The range of $lh(x \succcurlyeq y)$ is $[-1.0; +1.0]$, and $-lh(x \succcurlyeq y) \;=\; lh(x \not\succcurlyeq y)$, i.e. a negative value represents the likelihood of the negated outranking relation. A value $+1.0$ (resp. -1.0) means the outranking situation is certainly validated (resp. invalidated).

Example 1: equi-significant criteria

x and y are evaluated wrt 7 equi-significant criteria;

Four criteria positively support that x outranks y and three criteria support that x does not outrank y.

Suppose $E(W_j) = w$ for $j = 1, ..., 7$;

And $W_j \sim \mathcal{T}r(0, 2w, w)$ for $j = 1, ...7$;

Hence $E(\tilde{r}(x \succcurlyeq y)) = 4w - 3w = w$,

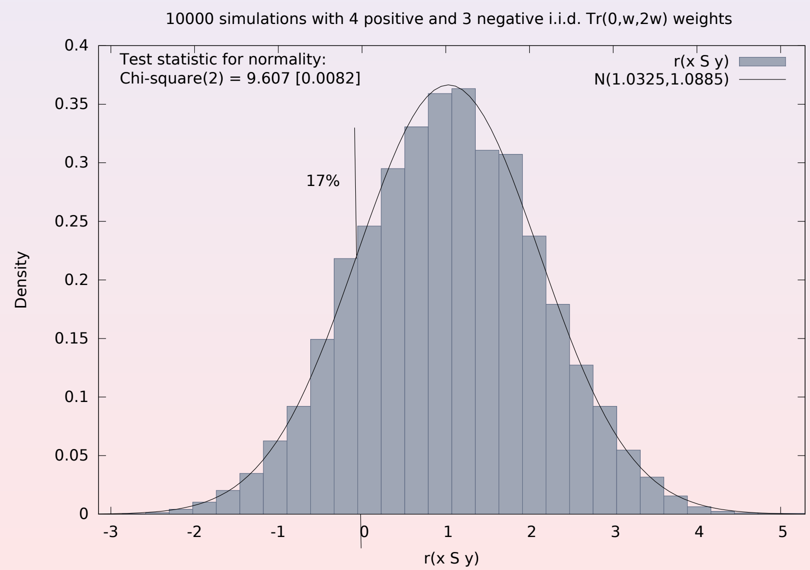
And $V(\tilde{r}(x \succcurlyeq y)) = 7 \times \frac{1}{6} w^2$.

If $w = 1$, $E[\tilde{r}(x \succcurlyeq y)] = 1$ and $sd[\tilde{r}(x \succcurlyeq y)] = 1.08$.

By the CLT, $lh(x \succcurlyeq y) \;=\; 0.66 \;\approx\; 83\%$,

10 000 MC runs confirm $\tilde{r}(x \succcurlyeq y) \rightsquigarrow Y = \mathcal{N}(1.03, 1, 089)$ with $P(Y \leqslant 0) \approx 17\%$.

Example 1 - continue



Example 2: various significance weights

Table : Pairwise comparison of two decision alternatives

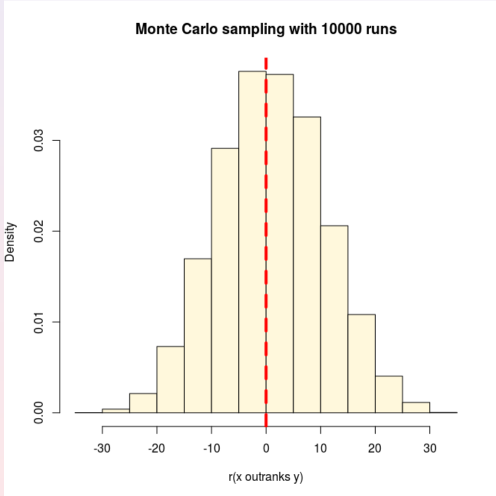
g_j $E(W_j)$	g_1 7	g_2 8	g_3 3	g_4 10	g_5 1	g_6 9	g_7 7
a_1	14.1	71.4	87.9	38.7	26.5	93.0	37.2
a_2	64.0	87.5	67.0	82.2	80.8	80.8	10.6
$a_1 - a_2$	-49.9	-16.1	+20.9	-43.5	-54.3	+12.2	26.5
$r(\succcurlyeq_j)$	-1	0	$+1$	-1	-1	$+1$	$+1$

$$E\left(\tilde{r}(a_1 \succcurlyeq a_2)\right) = \sum_{j=1}^{7} r(a_1 \succcurlyeq_j a_1) \times E(W_j)$$
$$= -7 + 0 + 3 - 10 - 1 + 9 + 7 = +1$$

If now $W_j \sim \mathcal{T}r(0, 2E(W_j), E(W_j))$,
how confident can we be about the actual positiveness of $\tilde{r}(a_1 \succcurlyeq a_2)$?

Example 2 - continue

- $\tilde{r}(a_1 \succcurlyeq a_2) \rightsquigarrow \mathcal{N}(\mu, \sigma)$ with
  $\mu = E(\tilde{r}(a_1 \succcurlyeq a_2)) = +1$
  $\sigma = \sqrt{\sum_i 1/6E(W_i)^2}$
  $= 6.94.$
- $lh(a_1 \succcurlyeq a_2) = +0.114$, hence
  $P(\tilde{r}(a_1 \succcurlyeq a_2) \leqslant 0.0)$
  $= (0.114 + 1.0)/2$
  $\approx 55.7\%.$



Content

The confident outranking relation $\succsim$

From an epistemic point of view, we say that:

1. action x *outranks* action y, denoted $(x \succsim y)$, if
   1.1 a confident majority of criteria validates a global outranking situation between x and y, and
   1.2 no veto is observed on a discordant criterion,
2. action x *does not outrank* action y, denoted $(x \not\succsim y)$, if
   2.1 a confident majority of criteria invalidates a global outranking situation between x and y, and
   2.2 no counter-veto is observed on a concordant criterion.

Considerably better or *worse performing* situations

On a criterion j, we characterize a *considerably less performing* situation, called veto and denoted $\lll_j$, as follows:

$$r(x \lll_j y) = \begin{cases} +1 & \text{if} \quad x_j + v_j \leqslant y_j \\ -1 & \text{if} \quad x_j - v_j \geqslant y_j \\ 0 & \text{otherwise.} \end{cases} \qquad (3)$$

where v_j represents a veto discrimination threshold. A corresponding dual *considerably better performing* situation, called counter-veto and denoted $\ggg_j$, is similarly characterized as:

$$r(x \ggg_j y) = \begin{cases} +1 & \text{if} \quad x_j - v_j \geqslant y_j \\ -1 & \text{if} \quad x_j + v_j \leqslant y_j \\ 0 & \text{otherwise.} \end{cases} \qquad (4)$$

Veto and *counter-veto* situations

A global considerable worst performaning (*veto*) situation, or considerably better perform ing (*counter-veto*) situation is now defined as follows:

$$r(x \lll y) = \circledvee_{j \in F}\, r(x \lll_j y) \qquad (5)$$

$$r(x \ggg y) = \circledvee_{j \in F}\, r(x \ggg_j y) \qquad (6)$$

where $\circledvee$ represents the epistemic polarising (Bisdorff 1997) or symmetric maximum (Grabisch et al. 2009) operator:

$$r \circledvee r' = \begin{cases} \max(r, r') & \text{if} \quad r \geqslant 0 \wedge r' \geqslant 0, \\ \min(r, r') & \text{if} \quad r \leqslant 0 \wedge r' \leqslant 0, \\ 0 & \text{otherwise.} \end{cases} \qquad (7)$$

Characterizing veto and counter-veto situations

1. $r(x \lll y) = 1$ iff there exists a criterion j such that $r(x \lll_j y) = 1$ and there does not exist otherwise any criterion k such that $r(x \ggg_k y) = 1$.
2. Conversely, $r(x \ggg y) = 1$ iff there exists a criterion j such that $r(x \ggg_j y) = 1$ and there does not exist otherwise any criterion k such that $r(x \lll_k y) = 1$.
3. $r(x \ggg y) = 0$ if either we observe no very large performance differences or we observe at the same time, both a very large positive and a very large negative performance difference.

Comment

$r(\not\lll)^{-1}$ *is identical to* $r(\ggg)$.

Polarising the global "*at least as good as*" characteristic

The outranking characteristic $\tilde{r}(\succsim)$ is defined as follows:

$$\tilde{r}(x \succsim y) = \left[\, \tilde{r}(x \succcurlyeq y) \circledvee -r(x \lll y) \,\right]$$

And in particular,

1. $\tilde{r}(x \succsim y) = \tilde{r}(x \succcurlyeq y)$ if no very large positive or negative performance differences are observed,
2. $\tilde{r}(x \succsim y) = 1$ if $\tilde{r}(x \succcurlyeq y) \geqslant 0$ and $r(x \ggg y) = 1$,
3. $r(x \succsim y) = -1$ if $\tilde{r}(x \succcurlyeq y) \leqslant 0$ and $r(x \lll y) = 1$,
4. $\tilde{r}(x \succsim y) = 0$ in all other cases, and especially if conjointly $r(x \ggg y) = 1$ and $r(x \lll y) = 1$.

Confidence level α for outranking situations

By requiring now a certain level α of likelihood for confidently validating all pairwise outranking situations, we may thus enforce the actual confidence we may have in the valued outranking digraph.
For any outranking situation $(x \succsim y)$ we obtain:

$$\hat{r}_\alpha(x \succsim y) \;=\; \begin{cases} E\big[\tilde{r}(x \succsim y)\big] & \text{if } \mathrm{abs}\big(lh(x \succcurlyeq y)\big) \geqslant \alpha, \\ 0 & \text{otherwise.} \end{cases} \qquad (8)$$

If $E(W_j) = w_j$, $E\big[\tilde{r}(x \succsim y)\big]$ equals the corresponding deterministic outranking characteristic $r(x \succsim y)$.
We safely preserve, hence, in our stochastic modelling, all the nice structural properties of the deterministic outranking relation like *weak completeness* and *coduality*.

Example 3: Confident outranking digraph

Table : Random performance tableau

g_i	w_i	a_1	a_2	a_3	a_4	a_5	a_6	a_7
g_1	7	14.1	64.0	73.4	36.4	30.6	85.9	97.8
g_2	8	71.4	87.5	61.9	84.7	60.4	54.5	45.8
g_3	3	87.9	67.0	25.2	34.2	87.3	43.1	30.4
g_4	10	38.7	82.2	94.1	86.1	34.1	97.2	72.2
g_5	1	26.5	80.8	71.9	21.3	56.4	88.1	15.0
g_6	9	93.0	80.8	23.2	57.2	81.4	16.6	93.0
g_7	7	37.2	10.6	64.8	98.9	69.9	24.7	13.6

Thresholds: $ind_i = 10.0$, $pr_i = 20$, and $v_i = 80$ for $i \in F$.

Example 3: Confident outranking digraph

Table : Deterministic credibility of $(x \succsim y)$

$r(\succsim) \times 45$	a_1	a_2	a_3	a_4	a_5	a_6	a_7
a_1	-	+**1**	−**5**	−11	+22	+9	**0**
a_2	+16	-	+21	**0**	+25	+14	+22
a_3	+21	+5	-	−3	+21	+34	+13
a_4	+21	+**45**	+29	-	+19	+19	+**45**
a_5	+28	−7	+10	−5	-	+9	+2
a_6	+6	+5	+31	−3	+7	-	+20
a_7	+**45**	+11	+1	**0**	+15	+13	-

Example 3: Confident outranking digraph

Table : CLT likelihood of the $(x \succcurlyeq y)$ situations

lh	a_1	a_2	a_3	a_4	a_5	a_6	a_7
a_1	-	+.11	−.49	−.89	+1.0	+.76	+.85
a_2	+.98	-	+1.0	+**1.0**	+1.0	+.98	+1.0
a_3	+.99	+.49	-	−.30	+.99	+1.0	+.91
a_4	+.99	+**.49**	+1.0	-	+.99	+1.0	+.91
a_5	+1.0	−.64	+.81	−.49	-	+.76	+.23
a_6	+.66	+.49	+1.0	−.30	+.64	-	+1.0
a_7	+.70	+.91	+.11	+.56	+.97	+.94	-

Example 3: Confident outranking digraph

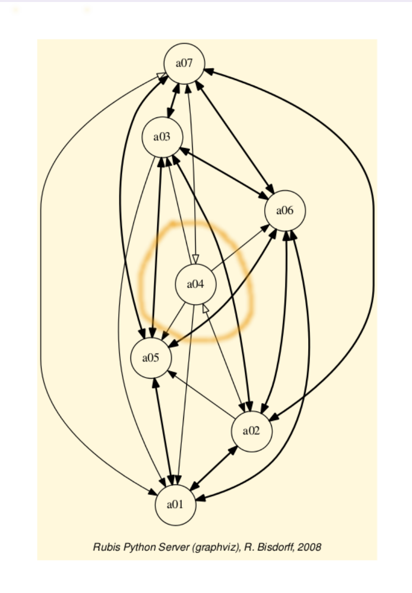
Table : 90% confident outranking situations

$\hat{r}_{90\%}(x \succsim y)$	a_1	a_2	a_3	a_4	a_5	a_6	a_7
a_1	-	**0**(+1)	**0**(−5)	−11	+22	**0**(+9)	0
a_2	+16	-	+21	0	+25	+14	+22
a_3	+21	**0**(+5)	-	**0**(−3)	+21	+34	+13
a_4	+21	**0**(+45)	+29	-	+19	+19	+45
a_5	+28	**0**(−7)	+10	**0**(−5)	-	**0**(+9)	**0**(+2)
a_6	**0**(−7)	**0**(+5)	+31	**0**(−3)	**0**(+7)	-	+20
a_7	**0**(+45)	+11	**0**(+1)	0	+15	+13	-

Content

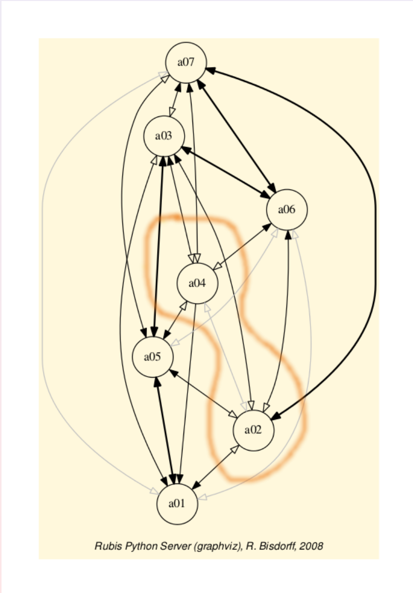
The confident outranking digraph

The deterministic digraph:



Condorcet winner: a_4

The 90% confident digraph:



Weak Condorcet winners: $\{a_4, a_2\}$

Exploiting the confident outranking digraph

Table : Pairwise comparison of alternatives a_4 and a_2

g_j / w_j	g_1 7	g_2 8	g_3 3	g_4 10	g_5 1	g_6 9	g_7 7
a_{4j}	36.5	84.7	34.2	86.1	21.3	57.2	98.9
a_{2j}	60.0	87.5	67.0	82.2	80.8	80.8	10.6
$(a_{4j} - a_{2j})$	-27.5	-2.8	-32.8	+3.8	-59.2	-23.6	+88.8
$r(a_4 \succcurlyeq_j a_2)$	−1	+1	−1	+1	−1	−1	+1
$r(a_4 \lll_j a_2)$	0	0	0	0	0	0	0
$r(a_4 \ggg_j a_2)$	0	0	0	0	0	0	+1

Thresholds: $ind_j = 10.0$, $pr_j = 20$, and $v_j = 80$ for $j \in F$.

$\tilde{r}(a_4 \succcurlyeq a_2) = +5$ and $\tilde{r}(a_4 \succsim a_2) = +45$.
Yet, $lh(a_4 \succcurlyeq a_2) = 0.49 < 0.80$, hence
$\hat{r}_{.80}(a_4 \succsim a_2) = 0$.

Concluding ...

- We illustrate some simple models for tackling uncertain significance weights: uniform, triangular and beta laws.
- Applying the Central Limit Theorem, we are able to compute the actual likelihood of any pairwise *at least as good as* and *not at least as good as* situations.
- This operational result allows to enforce a given confidence level on the corresponding outranking situations.
- On a small illustrative best choice problem, we eventually show the pragmatic decision aid benefit one may expect from exploiting a confident versus a classic deterministic outranking digraph.