Measuring and testing the ordinal correlation between valued outranking relations

Raymond Bisdorff

Université du Luxembourg
FSTC/ILAS

DA2PL - Mons, November 15-16, 2012

Motivation

- Fitness measures for posteriors in inverse multiple criteria preference analysis,
- Fine-tuning meta-heuristics for multiple criteria based clustering,
- Comparing multiple criteria rankings with rules like Kemeny's, Kohler's, the PROMETHEE net flows rule, or, more recently, Dias-Lamboray's prudent leximin rule.

Kendall's rank correlation τ measure

Let R_1 and R_2 be two binary relations defined on the same finite set X of dimension n.

Let $C = \#\{\, (x, y) \in X^2 \;:\; x \neq y \text{ and } ((x\,\mathrm{R}_1\,y) \Leftrightarrow (x\,\mathrm{R}_2\,y))\,\}$ denote the number of concordant non reflexive relational situations we observe.

$$\tau(\mathrm{R}_1, \mathrm{R}_2) \quad := \quad 2 \times \frac{C}{n(n-1)} \; - \; 1$$

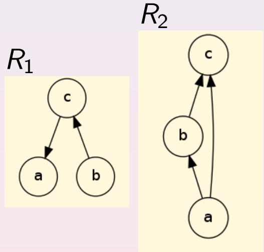
The τ correlation measure – continue

Comment

- *Unanimously concordant relations (100% equivalent situations) are matched to a correlation index of value* $+1.0$,
- *50% concordance between the relations (50% equivalent and 50% not equivalent situations) is matched to a zero-valued correlation index, and*
- *Unanimously discordant relations (100% non equivalent situations) are matched to a correlation index of value:* -1.0.

Example

- R_1, R_2 defined on $X = \{a, b, c\}$
- $R_1 = \{(b, c), (c, a)\}$
- $R_2 = \{(a, b), (a, c), (b, c)\}$
- Concordant pairs: $\{(b, a), (b, c), (c, b)\}$
- Discordant pairs: $\{(a, b), (a, c), (c, a)\}$



$$\tau(\mathrm{R}_1, \mathrm{R}_2) \quad = \quad 2 \times \frac{3}{6} - 1 \quad = \quad 0.0$$

r-valued relations of order n

- Let R_1 and R_2 be two binary relations defined on the same finite set X of dimension n and characterized via a bipolar characteristic function r taking values in the rational interval $[-1.0; 1.0]$.

  We call such relations, for short, *r-valued* and *of order n*.
- The r-valuation supports the following semantics:
  1. $r(x \,\mathrm{R}\, y) = \pm 1.0$ signifies that the relational situation $x \,\mathrm{R}\, y$ is certainly valid ($+1.0$), resp. invalid (-1.0);
  2. $r(x \,\mathrm{R}\, y) > 0.0$ signifies that the relational situation $x \,\mathrm{R}\, y$ is more valid than invalid;
  3. $r(x \,\mathrm{R}\, y) < 0.0$ signifies that the relational situation $x \,\mathrm{R}\, y$ is more invalid than valid;
  4. $r(x \,\mathrm{R}\, y) = 0.0$ signifies that the relational situation $x \,\mathrm{R}\, y$ is indeterminate, i.e. neither valid, nor invalid.

Logical r-valued operators

$$\begin{aligned}
r(\neg(x \,\mathrm{R}\, y)) &= -\, r(x \,\mathrm{R}\, y) \\
r\big((x \,\mathrm{R}_1\, y) \wedge (x \,\mathrm{R}_2\, y)\big) &= \min\big(r(x \,\mathrm{R}_1\, y), r(x \,\mathrm{R}_2\, y)\big), \\
r\big((x \,\mathrm{R}_1\, y) \vee (x \,\mathrm{R}_2\, y)\big) &= \max\big(r(x \,\mathrm{R}_1\, y), r(x \,\mathrm{R}_2\, y)\big).
\end{aligned}$$

$$\begin{aligned}
& r\big(\, (x \,\mathrm{R}_1\, y) \Leftrightarrow (x \,\mathrm{R}_2\, y) \,\big) \\
&\quad = r\Big[\big(\neg(x \,\mathrm{R}_1\, y) \vee (x \,\mathrm{R}_2\, y)\big) \wedge \big(\neg(x \,\mathrm{R}_2\, y) \vee (x \,\mathrm{R}_1\, y)\big)\Big] \\
&\quad = \min\Big[\max\big(-\, r(x \,\mathrm{R}_1\, y), r(x \,\mathrm{R}_2\, y)\big), \\
&\qquad\qquad \max\big(r(-x \,\mathrm{R}_2\, y), r(x \,\mathrm{R}_1\, y)\big)\Big]
\end{aligned}$$

Determinateness of r-valued relations

The determinateness of an r-valued relation R of order n, denoted $d(\mathrm{R})$, is defined as follows:

$$d(\mathrm{R}) \;:=\; \frac{\sum_{(x,y) \in X^2}^{x \neq y} \mathrm{abs}(r(x \,\mathrm{R}\, y))}{n(n-1)}$$

Comment

- *A crisp – a completely ± 1-valued – relation shows a determinateness degree of 1, whereas*
- *an indeterminate – a completely 0-valued – relation shows a determinateness degree of 0.*

A useful result

The equivalence of two r-valued relational situations verifies the following

Property

Let R_1 and R_2 be any two r-valued relations defined on the same set X. For all (x, y) in X^2, we have:

$$r\big((x\,\mathrm{R}_1\,y) \Leftrightarrow (x\,\mathrm{R}_2\,y)\big) = \pm \min\big(\,\mathrm{abs}(r(x\,\mathrm{R}_1\,y)), \mathrm{abs}(r(x\,\mathrm{R}_2\,y))\,\big)$$

Correlations between r-valued relations

The r-valued ordinal correlation τ between two r-valued relations R_1 and R_2, defined on a same set X, is formulated as follows:

$$\tau(\mathrm{R}_1, \mathrm{R}_2) := \frac{\sum_{x \neq y} r\big((x\,\mathrm{R}_1\,y) \Leftrightarrow (x\,\mathrm{R}_2\,y)\big)}{\sum_{x \neq y} \min\Big[\mathrm{abs}\big(r(x\,\mathrm{R}_1\,y)\big), \mathrm{abs}\big(r(x\,\mathrm{R}_2\,y)\big)\Big]}$$

Comment

- *In the crisp case, following Kendall, we divide the sum of pairwise equivalences by $n(n-1)$.*
- *If we would proceed this way in the valued case, the resulting measure would integrate a mixture of both the ordinal correlation as well as the actual determinateness of the equivalence observed between the considered r-valued relations.*
- *To factor out both these effects we take, instead, as denominator the maximum possible sum of r-valued equivalences we could potentially observe when both r-valued relations would show completely concordant relational situations.*

Example

Table: Examples of randomly valued relations

$r(x\,\mathrm{R}_1\,y)$	a	b	c
a	–	+0.68	+0.35
b	−0.94	–	+0.80
c	−1.00	+0.36	–

$r(x\,\mathrm{R}_2\,y)$	a	b	c
a	–	−0.32	+0.58
b	−0.14	–	+0.75
c	−1.00	+0.08	–

Example – continue

Table: r-valued equivalence between R_1 and R_2

$r(x\,\mathrm{R}_1\,y \Leftrightarrow x\,\mathrm{R}_2\,y)$	a	b	c
a	–	−0.32	+0.35
b	+0.14	–	+0.75
c	+1.00	+0.08	–

$$\tau(\mathrm{R}_1, \mathrm{R}_2) = \frac{-0.32 + 0.35 + 0.14 + 0.75 + 1.00 + 0.08}{+0.32 + 0.35 + 0.14 + 0.75 + 1.00 + 0.08} = \frac{0.200}{0.264} = +0.7575 = \frac{0.200}{6} \div 0.44$$

$$d(\mathrm{R}_1 \Leftrightarrow \mathrm{R}_2) = \frac{0.264}{6} = 0.44$$

Properties of the ordinal correlation measure

Let R_1 and R_2 be two r-valued binary relations defined on a same set X:

1. If R_1 and R_2 show a same, respectively an opposite, orientation, $\tau(\mathrm{R}_1, \mathrm{R}_2)$ equals $+1.0$, respectively -1.0, independently of their equivalence determinateness $d((\mathrm{R}_1 \Leftrightarrow \mathrm{R}_2))$.
2. If $\neg\mathrm{R}$ and Я denote resp. the negation and the converse of relation R, we may notice that:

$$\begin{aligned} \tau(\mathrm{R}_1, \mathrm{R}_2) &= \tau(\mathrm{R}_2, \mathrm{R}_1) \\ \tau(\neg\mathrm{R}_1, \mathrm{R}_2) &= -\tau(\mathrm{R}_1, \mathrm{R}_2) \\ \tau(\text{Я}_1, \text{Я}_2) &= \tau(\mathrm{R}_1, \mathrm{R}_2) \\ \tau(\neg\text{Я}_1, \neg\text{Я}_2) &= \tau(\mathrm{R}_1, \mathrm{R}_2) \end{aligned}$$

Correlation between r-valued relations of order n

- To each non reflexive pair (x, y) in X are associated two uniform random floats: $r(x\,\mathrm{R}_1\,y)$ and $r(x\,\mathrm{R}_2\,y)$.
- $r(x\,\mathrm{R}_1\,y \Leftrightarrow x\,\mathrm{R}_2\,y) \sim \triangle(-1, 1, 0)$, $\mu_e = 0$, $\sigma_e = \sqrt{3/18}$
- $d(x\,\mathrm{R}_1\,y \Leftrightarrow x\,\mathrm{R}_2\,y) \sim \triangle(0, 1, 0)$: $\mu_d = 1/3$, $\sigma_d = \sqrt{1/18}$.
- $\mathcal{T}_n(\mathrm{R}_1, \mathrm{R}_2) \rightsquigarrow \mathcal{N}\left(\frac{\mu_e}{\mu_d}, \frac{\sigma_e}{\mu_d}\frac{1}{\sqrt{n(n-1)}}\right)$
- $\hat{\mu}_{\mathcal{T}_n} \approx 0.0$
- $\hat{\sigma}_{\mathcal{T}_n} \approx \frac{3\sqrt{3/18}}{\sqrt{n(n-1)}}$

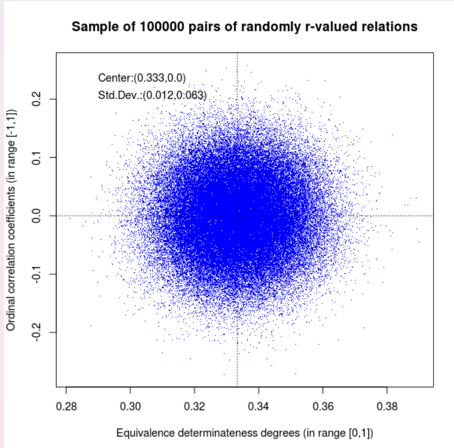


Table: Summary Statistics, for 100000 pairs of randomly r-valued relations

$d(\mathrm{R}_1, \mathrm{R}_2)$	$\overline{d}$	$\hat{\sigma}_d$	$\hat{\sigma}_d\sqrt{n(n-1)}$	Conf. 90%	Conf. 99%
$n = 5$	0.3333	0.0527	0.23568	±0.0866	±0.1355
$n = 10$	0.3334	0.0249	0.23622	±0.0406	±0.0645
$n = 15$	0.3333	0.0162	0.23476	±0.0266	±0.0418
$n = 20$	0.3333	0.0121	0.23587	±0.0202	±0.0276
$n = 30$	0.3333	0.0080	0.23597	±0.0132	±0.0207
$n = 50$	0.3333	0.0048	0.23758	±0.0078	±0.0121
$\tau(\mathrm{R}_1, \mathrm{R}_2)$	$\overline{\tau}$	$\hat{\sigma}_\tau$	$\hat{\sigma}_\tau\sqrt{n(n-1)}$	Conf. 90%	Conf. 99%
$n = 5$	0.0003	0.2731	1.22134	±0.4500	±0.6766
$n = 10$	0.0000	0.1289	1.22285	±0.2181	±0.3291
$n = 15$	0.0000	0.0842	1.22017	±0.1386	±0.2156
$n = 20$	0.0000	0.0621	1.21055	±0.1035	±0.1425
$n = 30$	0.0000	0.0414	1.22113	±0.0681	±0.1064
$n = 50$	0.0000	0.0247	1.22259	±0.0406	±0.0636

$\sigma_d = \sqrt{1/18} = 0.23570$, $\frac{\sigma_e}{\mu_d} = 3\sqrt{3/18} = 1.22474$.

Random weakly complete r-valued relations

- R is weakly complete if for all $(x, y) \in X$, $r(x\,\mathrm{R}\,y) < 0$ implies $r(x\,\mathrm{R}\,y) \geq 0$.
- Each link is, either a double, or a single forward or backward link, with equal probability 1/3.
- $\hat{\mu}_\tau = +0.111$.
- A weakness degree of 1.0 (resp. 0.0) gives a constant correlation measure of 1.0 (resp. 0.0).

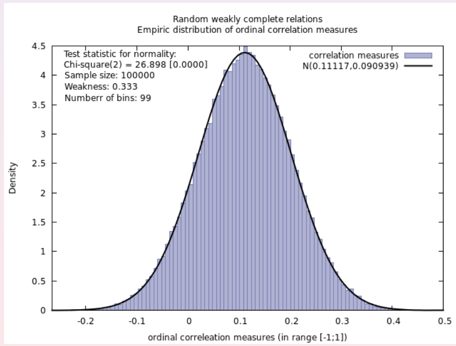


Table: Summary Statistics, for 100000 pairs of random weakly (1/3) complete relations

$d(R_1, R_2)$	$\hat{\mu}_d$	$\hat{\sigma}_d$	$\hat{\sigma}_d\sqrt{n(n-1)}$	Conf. 90%		Conf. 99%	
$n = 5$	0.33344	0.05268	0.23568	0.24920	0.42208	0.20493	0.47489
$n = 10$	0.33316	0.02490	0.23622	0.29262	0.37433	0.27069	0.39861
$n = 15$	0.33316	0.01634	0.23476	0.30646	0.36025	0.29164	0.37578
$n = 20$	0.33320	0.01209	0.23587	0.31342	0.35315	0.30262	0.36486
$n = 30$	0.33316	0.00799	0.23597	0.32006	0.34630	0.31269	0.35406
$n = 50$	0.33318	0.00477	0.23758	0.32537	0.34102	0.32099	0.34558
$\tau(R_1, R_2)$	$\hat{\mu}_\tau$	$\hat{\sigma}_\tau$	$\hat{\sigma}_\tau\sqrt{n(n-1)}$	Conf. 90%		Conf. 99%	
$n = 5$	0.1112	0.3032	1.3560	−0.3981	+0.6039	−0.6559	+0.8229
$n = 10$	0.1113	0.1592	1.5103	−0.1537	+0.3713	−0.3019	+0.5082
$n = 15$	0.1112	0.1138	1.6491	−0.0767	+0.2978	−0.1810	+0.3978
$n = 20$	0.1112	0.0909	1.7720	−0.0395	+0.2604	−0.1231	+0.3399
$n = 30$	0.1116	0.0681	2.0087	−0.0003	+0.2234	−0.0631	+0.2869
$n = 50$	0.1112	0.0484	2.3957	+0.0313	+0.1905	−0.0132	+0.2348

$\sigma_d = \sqrt{1/18} = 0.23570$, $\frac{\sigma_e}{\mu_d} = 3\sqrt{3/18} = 1.22474$.

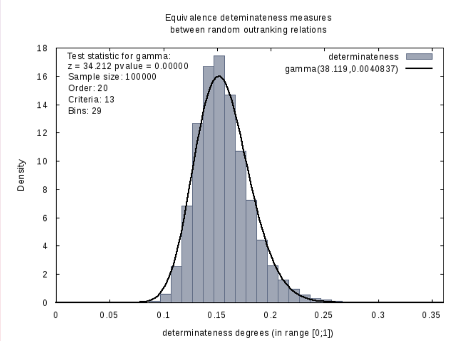
A model of random outranking relations

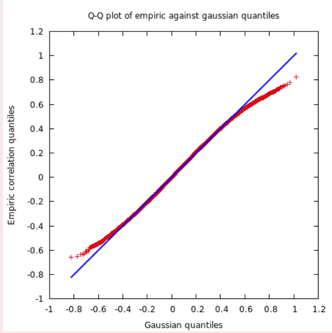
- Three types of decision actions: *cheap*, *neutral* and *expensive* ones with an equal proportion of 1/3.
- Two types of weighted criteria: *cost* criteria to be *minimized*, and *benefit* criteria to be *maximized*; in the proportions 1/3 respectively 2/3.
- Random performances on each type of criteria , either from an ordinal scale [0; 10], or from a cardinal scale [0.0; 100.0], following a parametric triangular law of mode: 30% performance for cheap, 50% for neutral, and 70% performance for expensive decision actions, with constant probability repartition 0.5 on each side of the respective mode.
- Cost criteria use mostly cardinal scales (3/4), whereas benefit criteria use mostly ordinal scales (2/3).
- The sum of weights of the cost criteria always equals the sum of weights of the benefit criteria.
- On cardinal criteria, both of cost or of benefit type, we observe following constant preference discrimination quantiles: 5% indifferent situations, 90% strict preference situations 90%, and 5% veto situation.
- We call this random model of r-valued relations for short random *CB*-outranking relations.

Table: Summary Statistics, for 100000 pairs of random *CB*-outranking relations

$d(R_1, R_2)$	$\hat{\mu}_d$	$\hat{d}_{50\%}$	$\hat{\sigma}_d$	Conf. 90%		Conf. 99%	
$n = 5, c = 3$	0.3259	0.3250	0.1131	0.1500	0.5255	0.0750	0.6333
$n = 10, c = 7$	0.2207	0.2165	0.0482	0.1494	0.3072	0.1204	0.3681
$n = 15, c = 9$	0.1910	0.1867	0.0362	0.1399	0.2577	0.1196	0.3102
$n = 20, c = 13$	0.1557	0.1527	0.0252	0.1203	0.2013	0.1053	0.2435
$n = 30, c = 21$	0.1372	0.1357	0.0174	0.1120	0.1674	0.1002	0.1989
$\tau(R_1, R_2)$	$\hat{\mu}_\tau$	$\hat{\tau}_{50\%}$	$\hat{\sigma}_\tau$	Conf. 90%		Conf. 99%	
$n = 5, c = 3$	0.0378	0.0345	0.5145	−0.7929	+0.8610	−1.0000	+1.0000
$n = 10, c = 7$	0.0629	0.0644	0.3037	−0.4420	+0.5560	−0.6483	+0.7467
$n = 15, c = 9$	0.0727	0.0761	0.2417	−0.3323	+0.4667	−0.5206	+0.6354
$n = 20, c = 13$	0.0984	0.1017	0.2085	−0.2492	+0.4383	−0.4224	+0.5904
$n = 30, c = 21$	0.1239	0.1272	0.1712	−0.1639	+0.4007	−0.3162	+0.5339

Correlation between pairs of random CB-outrankings





Application

Table: Example CB-outranking relation R_1 ($n = 10$, $c = 7$, $d(R_1) = 0.397$)

R_1	1	2	3	4	5	6	7	8	9	10
1	–	+0.14	+0.43	−0.14	+0.29	+0.14	+0.43	−0.14	±0.00	+0.43
2	+0.43	–	+0.43	−0.43	+0.43	+0.14	+0.14	+0.43	+0.14	+0.14
3	−0.43	−0.43	–	−0.71	+0.43	+0.00	−0.43	−1.00	−1.00	−0.14
4	**+0.14**	**+0.71**	**+1.00**	–	**+0.71**	**+0.43**	**+0.71**	**+0.43**	**+0.14**	**+0.57**
5	+0.14	−0.43	−0.43	−0.71	–	−0.71	−1.00	+0.14	−1.00	−0.43
6	−0.14	−0.14	+1.00	−0.43	+0.71	–	−0.14	−0.14	+0.14	−0.43
7	+0.14	+0.14	+0.43	−0.43	+1.00	+0.14	–	+0.14	+0.43	+0.29
8	+0.43	−0.14	+1.00	−0.43	+0.43	+0.14	−0.14	–	+0.43	−0.14
9	+1.00	−0.14	+1.00	−0.14	+1.00	+0.43	+0.14	−0.14	–	+0.14
10	−0.43	−0.14	+0.43	+0.14	+0.43	+0.43	+0.29	+0.14	−0.14	–

Assessing different ranking rules:

1. Kemeny ranking: $Ke = [4, 2, 7, 8, 9, 1, 10, 6, 3, 5]$, $\tau(R_1, Ke) = +0.888$,
2. Net flow scores: $Nf = [4, 9, 2, 7, 8, 10, 1, 6, 3, 5]$, $\tau(R_1, Nf) = +0.776$,
3. Kohler ranking: $Ko = [4, 2, 8, 10, 9, 6, 1, 7, 3, 5]$, $\tau(R_1, Ko) = +0.776$,
4. Ranked pairs (leximin): $Rp = [4, 2, 8, 9, 1, 7, 10, 6, 3, 5]$, $\tau(R_1, Rp) = +0.872$.

Conclusion

- We have consistently generalized Kendall's rank correlation measure τ to r-valued binary relations via a corresponding r-valued logical equivalence measure.
- The so extended ordinal correlation measure, besides remaining identical to Kendall's measure in the case of completely determined linear orders, shows interesting properties like its independence with the actual determinateness degree of the r-valued equivalence.
- Empirical confidence intervals for different models of random r-valued relations, like weakly complete and, more particularly, r-valued outranking relations are elaborated.