FH AACHEN
UNIVERSITY OF APPLIED SCIENCES

Fachhochschule Aachen
Campus Jülich

Stochastic methods for uncertain performance matrices in Multi-Criteria Decision Analysis with an application to the transition process towards sustainable mobility

Bachelor's Thesis

Luis Hasenauer

Faculty of
Medizintechnik und Technomathematik
Study program Angewandte Mathematik und Informatik B.Sc.

Jülich, August 8, 2023

Abstract

In many situations, decision-makers face complex challenges that require robust and effective decision-making processes. Multi Criteria Decision Analysis (MCDA) has become a valuable technique to assist decision-makers in weighing multiple competing criteria and making or analyzing decisions. MCDA methods highly rely on a performance matrix, which captures the evaluations of alternatives against various criteria. However, in many cases these methods use fixed performance matrices even though real-world decision problems often involve uncertainties in the evaluations that can significantly impact the accuracy and reliability of the performance matrix.

This bachelor thesis aims to investigate and propose solutions to address the challenges posed by uncertainties in MCDA. Here, uncertainties are modeled in a probabilistic matter and are included in the MCDA methods SAW and PROMETHEE II. Furthermore, the normal distribution and the uniform distribution are addressed specifically to model these uncertainties. On the other hand, a more general but approximate approach which uses Monte-Carlo-Simulation is developed and discussed in terms of convergence.

The presented methods are applied, discussed and visualized in various numerical experiments with different performance matrices and MCDA methods which show the need for a reliable way of uncertainty handling. Lastly, they are incorporated in a Bayesian approach on a specific example about sustainable mobility.

Contents

List of abbreviations

MCDA Multi Criteria Decision Analysis
SAW Simple Additive Weighting
PROMETHEE II Preference Ranking Organisation Method for Enrichment Evaluation II
PDF probability density function
CDF cumulative distribution function
CLT central limit theorem
MCMC Markov-Chain-Monte-Carlo

List of Figures

List of Tables

1 Methods to face uncertainties

In the context of MCDA, the expression $\mathbb{P}(f(\tilde{P}, w) = r|w)$ captures the probability of a specific ranking of alternatives $r \in A^d$ being valid, given a weight vector $w \in S_{n-1}$, when considering a possibly non-fixed matrix $\tilde{P} \in \mathbb{R}^{m \times n}$. Let the ranking be constructed in a way that all consecutive alternatives contained in it are considered superior to each other and all alternatives which are not contained in the ranking (denoted as $\tilde{r}$) are inferior to the alternative ranked last in r. Now, if the MCDA method has form $f(\tilde{P}, w) = argsort(\varphi)$, like in section **??**, equation (**??**) can be applied and we obtain:

$$\mathbb{P}(f(\tilde{P}, w) = r|w) = \mathbb{P}\Big(\langle P_{r_2}^T - P_{r_1}^T, w\rangle \leq 0, \ldots, \langle P_{r_d}^T - P_{r_{d-1}}^T, w\rangle \leq 0, \\ \langle P_{\tilde{r}_1}^T - P_{r_d}^T, w\rangle \leq 0, \ldots, \langle P_{\tilde{r}_{n-d}}^T - P_{r_d}^T, w\rangle \leq 0|w\Big)$$

Here, the terms $\langle P_{r_2}^T - P_{r_1}^T, w\rangle \leq 0, \ldots, \langle P_{r_d}^T - P_{r_{d-1}}^T, w\rangle \leq 0$ in $\mathbb{P}(f(\tilde{P}, w) = r|w)$ ensure that the performance of each alternative in the ranking r is higher or equal to the performance of the alternative ranked below it. This captures the preference order specified by the ranking r for the given weights w. Additionally, the terms $\langle P_{\tilde{r}_1}^T - P_{r_d}^T, w\rangle \leq 0, \ldots, \langle P_{\tilde{r}_{n-d}}^T - P_{r_d}^T, w\rangle \leq 0$ account for the performance differences between the alternatives in $\tilde{r}$ and the alternative ranked last in r. These terms ensure that the performance of the alternatives in $\tilde{r}$ is less than or equal to the performance of the alternative ranked last in r.

1.1 Fixed performance matrix

Let $\tilde{P}$ be a fixed performance matrix and P be a matrix obtained through an arbitrary MCDA method of form $\varphi = Pw$. Furthermore, let $H_{ij} = \left\{w \mid \langle P_j^T - P_i^T, w\rangle \leq 0\right\}$ represent a half space indicating the preference of alternative i over alternative j based on their performance values. The probability of this occurring can be represented by an indicator function $\mathbb{P}(\langle P_{r_j}^T - P_{r_i}^T, w\rangle \leq 0|w) = \mathbf{1}_{H_{r_i r_j}}(w)$. Moreover, the joint probability $\mathbb{P}(f(\tilde{P}, w) = r|w)$ of a weight being contained in all halfspaces is calculated using a product of indicator functions:

$$\mathbb{P}(f(\tilde{P}, w) = r|w) = \prod_{i=1}^{d-1} \mathbf{1}_{H_{r_i r_{i+1}}}(w) \prod_{i=1}^{n-d} \mathbf{1}_{H_{r_d \tilde{r}_i}}(w).$$

The subsequent product of indicator functions, $\prod_{i=1}^{d-1} \mathbf{1}_{H_{r_i r_{i+1}}}(w)$, represents the condition that the weights must lie in the half spaces corresponding to the preferences between consecutive alternatives in the ranking r. Similarly, the product $\prod_{i=1}^{n-d} \mathbf{1}_{H_{r_d \tilde{r}_i}}(w)$ ensures that the weights satisfy the preference conditions between the last alternative in the ranking r and the remaining alternatives not included in r, hence $\tilde{r}$.

With $W_r := S_{n-1} \bigcap_{i=1}^{d-1} H_{r_i r_{i+1}} \bigcap_{i=1}^{n-d} H_{r_d \tilde{r}_i}$, this can be simplified to $\mathbb{P}(f(\tilde{P}, w) = r|w) = \mathbf{1}_{W_r}(w)$

for $w \in S_{n-1}$ where W_r is a convex bounded polytope containing all weight vectors that result in the desired ranking [1].

1.2 Monte-Carlo-Simulation

The Monte-Carlo-Simulation provides a powerful approach to approximate $\mathbb{P}(f(\tilde{P}, w) = r|w)$ when considering a variable performance matrix and an arbitrary MCDA-method of form $\varphi = Pw$. In this context, the performance of alternatives is uncertain, and we aim to estimate the likelihood that a particular ranking holds true based on the given weights and the range of possible performance outcomes.

The Monte-Carlo approximation, denoted as $\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N$, allows us to calculate an estimate of the probability by generating N independent samples of the matrix P, $(P_1, \ldots, P_N)$, and evaluating the conditions that define the ranking. By aggregating the results from these samples and with W_r^i and H_{jk}^i being the respective sets from section 1.1 with $P = P_i$, we can obtain an approximation of the true probability $\mathbb{P}(f(\tilde{P}, w) = r|w)$:

$$\begin{aligned}\mathbb{P}(f(\tilde{P}, w) = r|w) &\overset{!}{\approx} \hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N \\ &= \frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{W_r^i}(w) \\ &= \frac{1}{N}\sum_{i=1}^{N}\left[\prod_{j=1}^{d-1} \mathbf{1}_{H_{r_j r_{j+1}}^i}(w) \prod_{j=1}^{n-d} \mathbf{1}_{H_{r_d \bar{r}_j}^i}(w)\right].\end{aligned}$$

To show that this Monte Carlo approximation is consistent, we need to show that it converges in expectation to the true probability $\mathbb{P}(f(\tilde{P}, w) = r|w)$ as the number of Monte Carlo samples goes to infinity. The expected value of this approximation can be computed as:

$$\mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} \mathbf{1}_{W_r^i}(w)\right] = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\left[\mathbf{1}_{W_r^i}(w)\right].$$

Since each Monte Carlo sample P_i is independently generated, the indicator function $\mathbf{1}_{W_r^i}(w)$ is an unbiased estimator of $\mathbb{P}(f(\tilde{P}, w) = r|w)$. Therefore, we have:

$$\frac{1}{N}\sum_{i=1}^{N} \mathbb{E}\left[\mathbf{1}_{W_r^i}(w)\right] = \frac{1}{N}\sum_{i=1}^{N} \mathbb{P}(f(\tilde{P}, w) = r|w) = \frac{1}{N} \cdot N \cdot \mathbb{P}(f(\tilde{P}, w) = r|w) = \mathbb{P}(f(\tilde{P}, w) = r|w).$$

Thus, the expected value of the Monte Carlo approximation is equal to the true probability $\mathbb{P}(f(\tilde{P}, w) = r|w)$. Since the mean almost surely converges to the expected value due to the strong law of large numbers, this demonstrates the consistency of the approximation.

Note that the consistency holds under the assumption that the Monte Carlo samples P_i are generated independently and according to the distribution that captures the uncertainty in the performance matrix $\tilde{P}$. Additionally, the convergence is in expectation, meaning that there can still be some variance or fluctuations in the estimated probabilities for a finite number of samples.

To asses convergence, one can use Chebyshev's inequality for a confidence interval [2]:

$$\mathbb{P}\left(\left| \hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N - \mathbb{P}(f(\tilde{P}, w) = r|w) \right| \leq \alpha \cdot \sigma \left[\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N \right] \right) \geq 1 - \frac{1}{\alpha^2}$$

where $\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N$ is the Monte Carlo approximation, $\mathbb{P}(f(\tilde{P}, w) = r|w)$ is the true probability, $\sigma \left[\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N \right]$ is the standard deviation of $\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N$, and α is a constant that determines the width of the confidence interval.

The maximum difference with at least probability $1 - \frac{1}{\alpha^2}$ between the Monte Carlo approximation and the true probability, denoted as $\Delta\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N$, is given by:

$$\Delta\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N = \alpha \cdot \sigma \left[\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N \right] = \alpha \cdot \sigma \left[\frac{1}{N} \sum_{i=1}^{N} \mathbf{1}_{W_r^i}(w) \right] = \alpha \frac{\sigma \left[\mathbf{1}_{W_r}(w) \right]}{\sqrt{N}}.$$

Since the standard deviation of $\mathbf{1}_{W_r}(w)$, which is Bernoulli distributed, and therefore:

$$\sigma[\mathbf{1}_{W_r}(w)] = \sqrt{\mathbb{P}(f(\tilde{P}, w) = r|w) \left(1 - \mathbb{P}(f(\tilde{P}, w) = r|w) \right)},$$

is unknown, one can use the empirical standard deviation:

$$s_N = \sqrt{\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N \left(1 - \hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N \right)},$$

as an approximation. To determine the termination criterion for convergence, we can set a threshold, ϵ, and check whether the ratio of $\Delta\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N$ to $\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N$ is less than ϵ [3], i.e.:

$$\frac{\Delta\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N}{\hat{\mathbb{P}}(f(\tilde{P}, w) = r|w)_N} < \epsilon.$$

If this condition is met, we can consider the Monte Carlo approximation to have converged within the desired tolerance. Additionally, one should add a threshold, n_{min}, of minimum iterations since, otherwise, only observing the desired ranking in the first iterations would

always result in a termination. The reason for this is that in this case:

$$\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N = \frac{1}{N}\sum_{i=1}^{N}\mathbf{1}_{W_r^i}(w) = \frac{1}{N}\sum_{i=1}^{N}1 = \frac{1}{N}N = 1.$$

Therefore, $s_N = \sqrt{\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N\left(1-\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N\right)} = 1(1-1) = 0$ and consequently:

$$\frac{\Delta\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N}{\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N} = \frac{\alpha\frac{s_N}{\sqrt{N}}}{\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N} = \frac{\alpha\frac{0}{\sqrt{N}}}{1} = 0 < \epsilon \quad \forall\epsilon > 0.$$

This approach provides a measure of confidence in the accuracy of the approximation, as the probability of the true value falling within the confidence interval increases as α becomes larger. By adjusting the values of α and ϵ, we can control the trade-off between accuracy and computational cost.

In pseudocode, an implementation could look the following:

Algorithm 1: Monte-Carlo-Simulation for MCDA

Input: w, r, α, ϵ, n_{min}

$n \leftarrow 0$;

$X \leftarrow$ `empty list`;

repeat

- $n \leftarrow n+1$;
- $\tilde{P} \leftarrow$ `sample from distribution`;
- $P \leftarrow P$ `from` $\tilde{P}$;
- $\varphi \leftarrow Pw$;
- $r_n \leftarrow$ `argsort(`φ`)`;
- **if** `rankingsMatch(`r_n, r`)` **then**
  - `addToList(`X, 1`)`;
- **else**
  - `addToList(`X, 0`)`;
- **end**
- $\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N \leftarrow$ `mean(`X`)`;
- $s_N \leftarrow$ `sqrt(`$\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N(1-\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N)$`)`;
- $\Delta\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N = \alpha\frac{s_N}{\sqrt{N}}$;

until $\frac{\Delta\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N}{\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N} < \epsilon$ *and* $n > n_{min}$;

return $\hat{\mathbb{P}}(f(\tilde{P},w)=r|w)_N$

1.3 Normally distributed performance matrix

A reasonable choice to model uncertainty in the performance matrix could be a normal distribution. Therefore, let $P \in \mathbb{R}^{m \times n}$ be equal to the performance matrix, hence $P = \tilde{P}$ which is a special case of Simple Additive Weighting (SAW), where each element P_{ij} is independently drawn from a normal distribution $\mathcal{N}(\mu_{ij}, \sigma_{ij}^2)$. Consequently, the rows of P can be regarded as multivariate normal distributions with a diagonal covariance matrix. This covariance matrix Σ_i for each row i takes the form:

$$\Sigma_i = \begin{pmatrix} \sigma_{i0}^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_{in}^2 \end{pmatrix}.$$

Additionally, the mean vector μ_i for each row i can be represented as:

$$\mu_i = (\mu_{io}, \ldots, \mu_{in})^T.$$

Consequently, the transposed row vector P_i^T follows a multivariate normal distribution $\mathcal{N}_n(\mu_i, \Sigma_i)$.

An alternative i is considered superior to an alternative j if $\langle P_j^T - P_i^T, w \rangle \leq 0$. By focusing on the inner part of the dot product, we find that $P_j^T - P_i^T$ follows a multivariate normal distribution $\mathcal{N}_n(\mu_j - \mu_i, \Sigma_j + \Sigma_i)$. When considering a fixed weight vector w, it can be deduced that $\langle P_j^T - P_i^T, w \rangle$ follows a normal distribution $\mathcal{N}(w^T(\mu_j - \mu_i), w^T(\Sigma_j + \Sigma_i)w)$ [4].

Let $r \in \mathbb{R}^d$ represent the given ranking, and $\tilde{r} \in \mathbb{R}^{n-d}$ denote all alternatives not included in r. In order to satisfy all $n-1$ conditions specified by the ranking simultaneously, it is necessary to consider the joint distribution of these halfspaces. This joint distribution follows a multivariate normal distribution $\mathcal{N}_{n-1}(\hat{\mu}, \hat{\Sigma})$, where the means are given by:

$$\hat{\mu_i} = \begin{cases} w^T(\mu_{r_{i+1}} - \mu_{r_i}) & , 1 \leq i < d \\ w^T(\mu_{\tilde{r}_{i-d+1}} - \mu_{r_d}) & , d \leq i < n \end{cases}$$

The covariances, due to the independence of the entries of P, are given by:

$$\mathrm{Cov}\left[\langle P_j^T - P_i^T, w \rangle, \langle P_l^T - P_k^T, w \rangle\right] = \begin{cases} 0 & , j \neq l \neq i \neq k \\ -\langle w^{\circ 2}, \sigma_i^2 \rangle & , j \neq l = i \neq k \\ \langle w^{\circ 2}, \sigma_i^2 \rangle & , j \neq l \neq i = k \\ \langle w^{\circ 2}, \sigma_i^2 + \sigma_j^2 \rangle & , j = l \neq i = k \end{cases} \tag{1.1}$$

with $w^{\circ 2}$ being the Hadamard power which denotes that $\tilde{w} = w^{\circ 2} \Leftrightarrow \tilde{w}_i = w_i^2 \ \forall 1 \leq i \leq n$. Furthermore, it is important to note that $\mathrm{Cov}[X, Y] = \mathrm{Cov}[Y, X]$ and all other cases of indices are impossible if the ranking is constructed according to the presented way.

The covariance between $\langle P_j - P_i, w\rangle$ and $\langle P_k - P_j, w\rangle$ with $i \neq j \neq k$, which corresponds to the scenario where alternative j is inferior to alternative i but superior to alternative k, can be proven as follows:

$$\begin{aligned}
\mathrm{Cov}\Big[\langle P_j - P_i,\ w\rangle, \langle P_k - P_j, w\rangle\Big] &= \mathrm{Cov}\left[\sum_{l=1}^{n}(P_{jl} - P_{il})w_l,\ \sum_{q=1}^{n}(P_{kq} - P_{jq})w_l\right] \\
&= \sum_{l=1}^{n}\sum_{q=1}^{n} w_l w_q \mathrm{Cov}\Big[P_{jl} - P_{il},\ P_{kq} - P_{jq}\Big] \\
&= \sum_{l=1}^{n}\sum_{q=1}^{n} w_l w_q \Big(\mathrm{Cov}\Big[P_{jl},\ P_{kq}\Big] - \mathrm{Cov}\Big[P_{jl},\ P_{jq}\Big] \\
&\qquad - \mathrm{Cov}\Big[P_{il},\ P_{kq}\Big] + \mathrm{Cov}\Big[P_{il},\ P_{jq}\Big]\Big)
\end{aligned}$$

Since all entries of P are assumed to be independent, their covariances between each other will be zero. Hence, for $l \neq q$ we only obtain zero summands and consequently:

$$\begin{aligned}
&\sum_{l=1}^{n}\sum_{q=1}^{n} w_l w_q \Big(\mathrm{Cov}\Big[P_{jl},\ P_{kq}\Big] - \mathrm{Cov}\Big[P_{jl},\ P_{jq}\Big] - \mathrm{Cov}\Big[P_{il},\ P_{kq}\Big] + \mathrm{Cov}\Big[P_{il},\ P_{jq}\Big]\Big) \\
&= \sum_{l=1}^{n} w_l^2 \Big(\mathrm{Cov}\Big[P_{jl},\ P_{kl}\Big] - \mathrm{Cov}\Big[P_{jl},\ P_{jl}\Big] - \mathrm{Cov}\,[P_{il},\ P_{kl}] + \mathrm{Cov}\Big[P_{il},\ P_{jl}\Big]\Big).
\end{aligned}$$

Additionally, since $i \neq j \neq k$, the random variables P_{il}, P_{jl} and P_{kl} will have a covariance of zero. Therefore, the covariance of $\langle P_j - P_i,\ w\rangle$ and $\langle P_k - P_j, w\rangle$ simplifies to:

$$\begin{aligned}
&\sum_{l=1}^{n} w_l^2 \Big(\mathrm{Cov}\Big[P_{jl},\ P_{kl}\Big] - \mathrm{Cov}\Big[P_{jl},\ P_{jl}\Big] - \mathrm{Cov}\,[P_{il},\ P_{kl}] + \mathrm{Cov}\Big[P_{il},\ P_{jl}\Big]\Big) \\
&= -\sum_{l=1}^{n} w_l^2 \mathrm{Cov}\Big[P_{jl},\ P_{jl}\Big] \\
&= -\sum_{l=1}^{n} w_l^2 \sigma_{jl}^2 \\
&= -\langle w^{\circ 2}, \sigma_j^2\rangle.
\end{aligned}$$

The proof for the covariance between $\langle P_j - P_i, w\rangle$ and $\langle P_k - P_i, w\rangle$ is analogous. Furthermore, when considering the covariance of a variable with itself, it simplifies to the variance $w^T(\Sigma_j + \Sigma_i)w = \langle w^{\circ 2}, \sigma_i^2 + \sigma_j^2\rangle$ since Σ_i and Σ_j are diagonal matrices.

Table 1: The covariance matrix $\hat{\Sigma}$. Here, $r_i \geq r_j$ is short for the random variable $\langle P_{r_j}^T - P_{r_i}^T, w \rangle$.

	$r_1 \geq r_2$	$r_2 \geq r_3$	...	$r_{d-1} \geq r_d$	$r_d \geq r_1^C$	$r_d \geq r_2^C$	...	$r_d \geq r_{n-d}^C$
$r_1 \geq r_2$	$\langle w^{\circ 2}, \sigma_{r_1}^2 + \sigma_{r_2}^2 \rangle$	$\langle w^{\circ 2}, \sigma_{r_2}^2 \rangle$	0	0	0	0	…	0
$r_2 \geq r_3$	$\langle w^{\circ 2}, \sigma_{r_2}^2 \rangle$	$\langle w^{\circ 2}, \sigma_{r_2}^2 + \sigma_{r_3}^2 \rangle$	$\ddots$	0	0	0	…	0
$\vdots$	0	$\ddots$	$\ddots$	$\ddots$	0	0	…	0
$r_d \leq r_{d-1}$	0	0	$\ddots$	$\langle w^{\circ 2}, \sigma_{r_{d-1}}^2 + \sigma_{r_d}^2 \rangle$	$\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	$\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	…	$\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$
$r_d \geq r_1^C$	0	0	0	$\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	$\langle w^{\circ 2}, \sigma_{r_d}^2 + \sigma_{r_1^C}^2 \rangle$	$-\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	…	$-\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$
$r_d \geq r_2^C$	0	0	0	$\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	$-\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	$\langle w^{\circ 2}, \sigma_{r_d}^2 + \sigma_{r_2^C}^2 \rangle$	$\ddots$	$-\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\ddots$	$-\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$
$r_d \geq r_{n-d}^C$	0	0	0	$\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	$-\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	$-\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	$-\langle w^{\circ 2}, \sigma_{r_d}^2 \rangle$	$\langle w^{\circ 2}, \sigma_{r_d}^2 + \sigma_{r_{n-d}^C}^2 \rangle$

Finally, it follows that $\mathbb{P}(f(\tilde{P}, w) = r | w)$ can be calculated using the multivariate normal CDF which does not have a closed form.

1.4 Uniformly distributed performance matrix

Another feasible choice to model uncertainty in the performances is the uniform distribution. Hence, let us consider a scenario where the performances are assumed to be independent and uniformly distributed. We denote this as $P_{ij} \sim \mathcal{U}(a_{ij}, b_{ij})$, where $a_{ij} < b_{ij}$ for all $1 \leq i \leq m$ and $1 \leq j \leq n$. The dot product of a row with a weight vector w can be expressed as the sum of independent uniform random variables: $\langle P_j^T, w \rangle = \sum_{k=1}^n P_{jk} w_k$, where $P_{jk} w_k \sim \mathcal{U}(w_k a_{jk}, w_k b_{jk})$ [5]. The condition for alternative i being superior to alternative j can also be represented as a sum of independent uniform random variables:

$$\begin{aligned} \langle P_j - P_i, w \rangle &\leq 0 \\ \langle P_j^T, w \rangle - \langle P_i^T, w \rangle &\leq 0 \\ \sum_{k=1}^n P_{jk} w_k - P_{ik} w_k &\leq 0. \end{aligned}$$

The individual summands will follow a uniform distribution, where $-P_{ik}w_k \sim \mathcal{U}(-w_k b_{ik}, -w_k a_{ik})$ and $P_{jk}w_k$ distributed as described previously. It is worth noting that the distribution of a sum of non-identically independent uniform random variables is known [6]. However, in this case, there is no available information regarding the joint probability between different sums which would be needed to compute the probability $\mathbb{P}(f(\tilde{P}, w) = r|w)$ for the whole ranking. Therefore, we can analyze the individual sums, but we do not have information about their joint behavior.

Since the distribution of $\langle P_j - P_i, w\rangle$ would be easier to handle if it was normally distributed, it can be tried to approximate it in this way. This approximation is made feasible by the Lyapunov central limit theorem (CLT).

Lyapunov CLT *Let $X_1, X_2, \ldots, X_n$ be a sequence of independent random variables with finite means $\mu_k = E[X_k]$ and finite variances $\sigma_k^2 = Var(X_k)$. If for some $\delta > 0$, the following condition holds:*

$$\lim_{n\to\infty} \frac{1}{(\sigma_1^2 + \ldots + \sigma_n^2)^{\frac{2+\delta}{2}}} \sum_{k=1}^{n} \mathbb{E}\left[|X_k - \mu_k|^{2+\delta}\right] = 0,$$

then:

$$\lim_{n\to\infty} \frac{1}{s_n} \sum_{k=1}^{n} (X_k - \mu_k) \xrightarrow{d} \mathcal{N}(0,1),$$

with $s_n^2 = \sum_{i=1}^{n} \sigma_i^2$.

We want to apply this theorem for $X_k = w_k(P_{jk} - P_{ik})$ and therefore $\mu_k = \frac{1}{2} w_k(b_{jk} + a_{jk} - b_{ik} - a_{ik})$ and $s_n^2 = \frac{1}{12} \sum_{k=1}^{n} w_k^2 \left((b_{jk} - a_{jk})^2 + (b_{ik} - a_{ik})^2\right)$.

Since proving the Lyapunov condition for $\langle P_j - P_i, w\rangle$ would extend the limit of this bachelor's thesis, we investigate its asymptotic normality empirically.

To do so, 1000 samples of P_i and P_j are generated where each sample is independently drawn from a uniform distribution $P_{kl} \sim \mathcal{U}(a_{kl}, b_{kl})\ \forall 1 \leq l \leq n,\ k \in \{i, j\}$. Furthermore, the corresponding intervals are subject to the conditions $0 \leq a_{kl} < 1$ and $b_{kl} - a_{kl} \leq 1$.

Subsequently, for each sampled pair of rows P_i and P_j, the random variable $Y = \frac{1}{s_n} \sum_{k=1}^{n} w_k(P_{jk} - P_{ik}) - \mu_k$ is calculated. If the CLT would apply, this random variable Y would converge in distribution to a standard normal distribution. For the sake of investigating this, the samples of Y are tested for normality [7]. This is tested for different numbers of criteria between 10 and

100 and for two sets of weights.

The first set contains 100 weights which are uniformly distributed on the standard simplex S_{n-1}. Specifically, each weight is generated by sampling n i.i.d. samples from an exponential distribution and then scaling them by their sum. This results in the weights following a uniform distribution on S_{n-1} [8]. Since it is unlikely that a CLT holds with extreme weights on, i.e. corners of S_{n-1}, a more centered set of weights is also taken into consideration. Here, the idea is that all summands have a relatively equal influence on their weighted average or rather, there is no summand that significantly determines the result of $\langle P_j - P_i, w \rangle$ alone. Consequently, the second set contains weights which are more concentrated towards the center of S_{n-1} and are generated by sampling n i.i.d. samples from a standard uniform distribution which are then scaled by their sum.

We want to analyze the corresponding p-values of the normality test. Therefore, the average p-value of all weight vectors and the percentage of rejected normality tests using a significance level of 5% are visualized with respect to the number of criteria. Ideally, the percentage of rejected tests should approach the significance level as n approaches infinity.

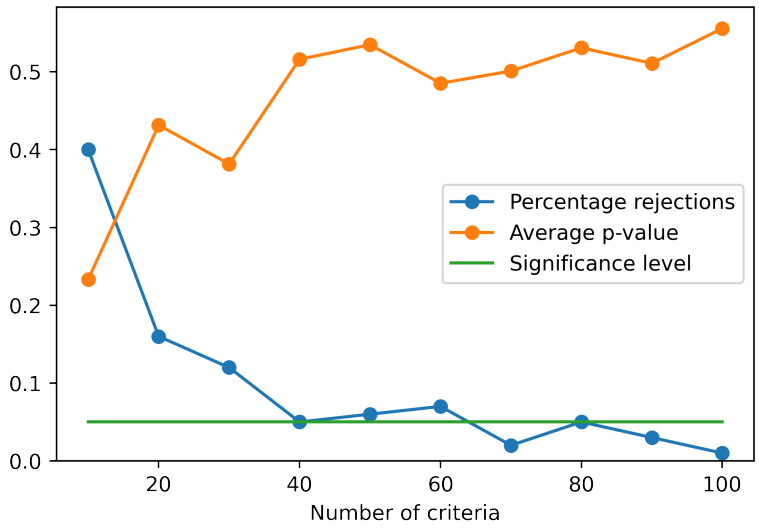

Figure 1: Results of normality test with uniformly distributed weights on S_{n-1}.

In Figure 1, it can be seen that the number of rejected null hypotheses seems to shrink as the number of criteria grows which is a good sign for assuming asymptotic normality. Furthermore the average p-value is sufficiently far away from the significance level.

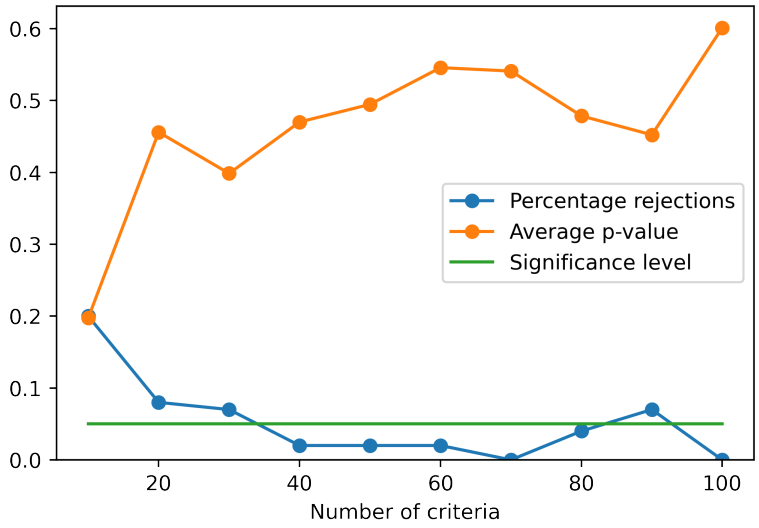

Figure 2: Results of normality test with weights closer to the center of S_{n-1}.

As expected, the results with the second sets of weights are even better than with the first set of weights. The percentage of rejected null hypotheses is really close to the significance level of 5% for almost every test (see Figure 2).

Since the CLT applies for $n \to \infty$, the test is repeated for different numbers of criteria between 100 and 1000. We observe promising results for both sets of weights (see Figures 3 & 4).

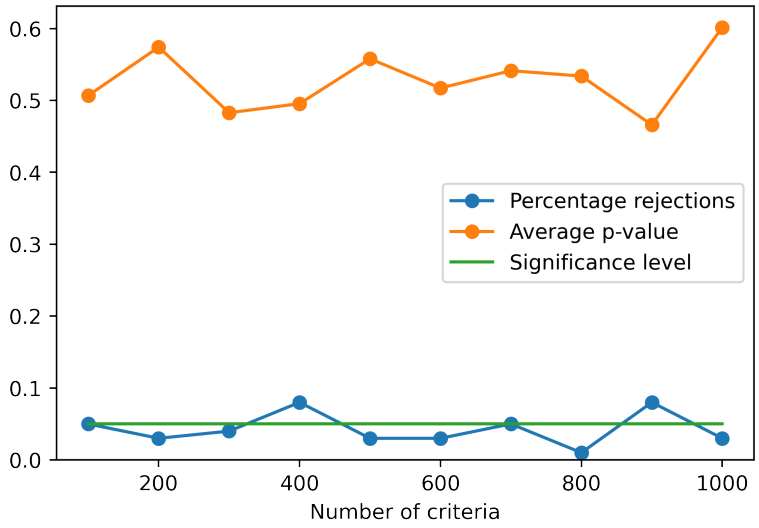

Figure 3: Results of normality test with uniformly distributed weights on S_{n-1}.

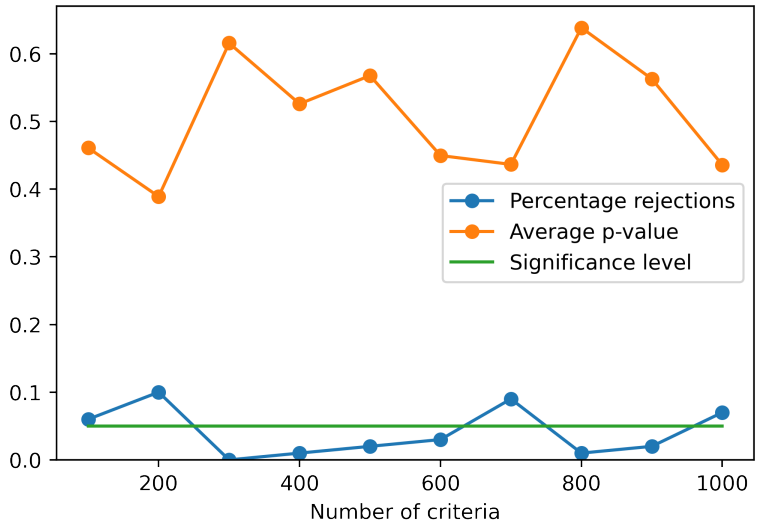

Figure 4: Results of normality test with weights closer to the center of S_{n-1}.

In summary, the empirical studies suggest that the CLT applies for the random variable $X_k = w_k(P_{jk} - P_{ik})$. Furthermore, the results were also reasonable if the number of criteria is less than 100 which is a more realistic number when applying MCDA. In our case, where $n < 100$, we take:

$$\sum_{k=1}^{n} w_k(P_{jk} - P_{ik}) \sim \mathcal{N}\left(\sum_{k=1}^{n} \mu_k,\ s_n^2\right),$$

with $\mu_k = \frac{1}{2} w_k(b_{jk} + a_{jk} - b_{ik} - a_{ik})$ and $s_n^2 = \frac{1}{12} \sum_{k=1}^{n} w_k^2 \left((b_{jk} - a_{jk})^2 + (b_{ik} - a_{ik})^2\right)$.

Lastly, it needs to be noted that this approximation may become worse if the weights move away from the center of S_{n-1} and therefore Monte-Carlo-Simulation might be a better approach in some cases if computation time is no significant factor.

1.5 Simple Additive Weighting

When using the SAW method, it is necessary to scale the criteria to comparable units. Hence, using the scaling as presented in **??**, will result in P_{SAW} loosing its original distribution and therefore $\mathbb{P}(f(\tilde{P}, w) = r|w)$ needs to be computed by using Monte-Carlo-Simulation.

To preserve the original distribution of $\tilde{P}$ in P_{SAW}, alternatively, the scaling can be approximated by dividing each element of the performance matrix by the sum of the expected values

of the corresponding column, which are assumed to be greater than zero when using SAW:

$$P_{SAW\,ij} = \frac{\tilde{P}_{ij}}{\sum_{k=1}^{m} \mathbb{E}\left[\tilde{P}_{kj}\right]}.$$

Consequently, after scaling, the sum of the expected values in each column of the scaled matrix P will be one:

$$\mathbb{E}\left[\sum_{i=1}^{m} P_{SAW\,ij}\right] = \sum_{i=1}^{m} \mathbb{E}\left[P_{SAW\,ij}\right] = \sum_{i=1}^{m} \mathbb{E}\left[\frac{\tilde{P}_{ij}}{\sum_{k=1}^{m} \mathbb{E}\left[\tilde{P}_{kj}\right]}\right] = \frac{\sum_{i=1}^{m} \mathbb{E}\left[\tilde{P}_{ij}\right]}{\sum_{k=1}^{m} \mathbb{E}\left[\tilde{P}_{kj}\right]} = 1.$$

To incorporate benefit and cost to the criteria, criteria associated to cost can be set to their negative values since this punishes high values and benefits low values which is the intended behavior.

Furthermore, if we assume that the elements of $\tilde{P}$ are normally distributed, which was discussed in section 1.3, we can express the distribution of $P_{SAW\,ij}$ as follows:

$$P_{SAW\,ij} \sim \mathcal{N}\left(\frac{\mu_{ij}}{\sum_{k=1}^{m} \mu_{kj}}, \frac{\sigma_{ij}^2}{(\sum_{k=1}^{m} \mu_{kj})^2}\right).$$

Here, μ_{ij} represents the mean and σ_{ij}^2 represents the variance associated with $\tilde{P}_{ij}$. Similarly, if we assume a uniform distribution, as in section 1.4, the distribution of $P_{SAW\,ij}$ can be expressed as:

$$P_{SAW\,ij} \sim \mathcal{U}\left(\frac{a_{ij}}{\sum_{k=1}^{m} \mu_{kj}}, \frac{b_{ij}}{\sum_{k=1}^{m} \mu_{kj}}\right). \tag{1.2}$$

In this case, a_{ij} and b_{ij} represent the lower and upper bounds of the uniform distribution, respectively. If one would try to approximate this by a normal distribution under the assumption that a CLT holds, the tests in section 1.4 used the conditions $\frac{a_{ij}}{\sum_{k=1}^{m} \mu_{kj}} < 1$ and $\frac{b_{ij}}{\sum_{k=1}^{m} \mu_{kj}} - \frac{a_{ij}}{\sum_{k=1}^{m} \mu_{kj}} \leq 1$. For the first condition, it can be obtained that:

$$\frac{a_{ij}}{\sum_{k=1}^{m} \mu_{kj}} < \frac{a_{ij}}{\mu_{ij}} = \frac{a_{ij}}{\frac{a_{ij}+b_{ij}}{2}} < \frac{a_{ij}}{a_{ij}} = 1.$$

Furthermore, the second condition results in:

$$\begin{aligned} & \frac{b_{ij}}{\sum_{k=1}^{m} \mu_{kj}} - \frac{a_{ij}}{\sum_{k=1}^{m} \mu_{kj}} \leq 1 \\ \Leftrightarrow \quad & \frac{b_{ij}}{\sum_{k=1}^{m} \mu_{kj}} \leq 1 + \frac{a_{ij}}{\sum_{k=1}^{m} \mu_{kj}} \end{aligned}$$

$$\Leftrightarrow \qquad b_{ij} \leq \sum_{k=1}^{m} \mu_{kj} + a_{ij}.$$

Therefore, the upper bound b_{ij} of each entry $\tilde{P}_{ij}$ needs to be smaller than the sum of means of criterion j and its lower bound a_{ij}. This condition will most likely be fulfilled if the number of alternatives is reasonably large.

These considerations highlight the flexibility of the SAW method in handling uncertain performance matrices, allowing for different scaling approaches and distribution assumptions.

1.6 Preference Ranking Organisation Method for Enrichment Evaluation II

While utilizing Preference Ranking Organisation Method for Enrichment Evaluation II (PROMETHEE II), the entries of P_{PROM2}, as in (**??**), are defined by:

$$P_{PROM2_{ij}} = \frac{1}{m-1} \sum_{k=1}^{m} \Big[g_j(\tilde{P}_{ij} - \tilde{P}_{kj}) - g_j(\tilde{P}_{kj} - \tilde{P}_{ij}) \Big].$$

Inserting this equation into the probability of an alternative i being preferred over an alternative j results in:

$$\begin{aligned}
&\mathbb{P}\Big(\langle P^T_{PROM2_j} - P^T_{PROM2_i}, w \rangle \leq 0 \Big| w \Big) \\
&= \mathbb{P}\left(\sum_{l=1}^{n} w_l (P_{PROM2_{jl}} - P_{PROM2_{il}}) \leq 0 \Big| w \right) \\
&= \mathbb{P}\left(\frac{1}{m-1} \sum_{l=1}^{n} w_l \sum_{k=1}^{m} \Big[g_l(\tilde{P}_{jl} - \tilde{P}_{kl}) - g_l(\tilde{P}_{kl} - \tilde{P}_{jl}) - g_l(\tilde{P}_{il} - \tilde{P}_{kl}) + g_l(\tilde{P}_{kl} - \tilde{P}_{il}) \Big] \leq 0 \Big| w \right) \\
&= \mathbb{P}\left(\sum_{l=1}^{n} w_l \sum_{k=1}^{m} \Big[g_l(\tilde{P}_{jl} - \tilde{P}_{kl}) - g_l(\tilde{P}_{kl} - \tilde{P}_{jl}) - g_l(\tilde{P}_{il} - \tilde{P}_{kl}) + g_l(\tilde{P}_{kl} - \tilde{P}_{il}) \Big] \leq 0 \Big| w \right).
\end{aligned}$$

To investigate this probability, let us begin with the pairwise differences $\tilde{P}_{ij} - \tilde{P}_{kl}$. Assuming the performances are normally distributed with $\tilde{P}_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma^2_{ij})$ and $\tilde{P}_{kl} \sim \mathcal{N}(\mu_{kl}, \sigma^2_{kl})$, it follows that $\tilde{P}_{ij} - \tilde{P}_{kl} \sim \mathcal{N}(\mu_{ij} - \mu_{kl},\ \sigma^2_{ij} + \sigma^2_{kl})$.

On the other hand, if the performances are uniformly distributed with $\tilde{P}_{ij} \sim \mathcal{U}(a_{ij}, b_{ij})$ and $\tilde{P}_{kl} \sim \mathcal{U}(a_{kl}, b_{kl})$, further investigations need to be done. The difference $\tilde{P}_{ij} - \tilde{P}_{kl}$ can be written as $\tilde{P}_{ij} + (-\tilde{P}_{kl})$ where $-\tilde{P}_{kl} \sim \mathcal{U}(-b_{kl}, -a_{kl})$. Therefore, their distribution can be found by calculating their convolution [9, Theorem 5.5.1].

To calculate the convolution of two uniform random variables, let $X \sim \mathcal{U}(a,b)$, $Y \sim \mathcal{U}(c,d)$

and $Z = X + Y$. The convolution of their respective densities is then defined as:

$$\begin{aligned}
f_Z(z) &= (f_X * f_Y)(z) \\
&= \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx \\
&= \int_{-\infty}^{\infty} \frac{\mathbf{1}_{(a,b)}(x)}{b-a} \cdot \frac{\mathbf{1}_{(c,d)}(z-x)}{d-c} dx \\
&= \int_{a}^{b} \frac{1}{b-a} \cdot \frac{\mathbf{1}_{(c,d)}(z-x)}{d-c} dx \\
&= \int_{\max(a,z-d)}^{\min(b,z-c)} \frac{1}{b-a} \cdot \frac{1}{d-c} dx \\
&= \frac{\min(b,z-c)}{(b-a)(d-c)} - \frac{\max(a,z-d)}{(b-a)(d-c)} \\
&= \frac{\min(b,z-c) - \max(a,z-d)}{(b-a)(d-c)}.
\end{aligned}$$

To eliminate the $\min$ and $\max$ from this PDF, let us investigate $\min(b, z-c)$ first. We observe that:

$$\min(b,z-c) = \begin{cases} z-c & a+c \le z \le b+c \\ b & b+c \le z \le b+d \end{cases}.$$

Additionally, for $\max(a, z-d)$ it can be seen that:

$$\max(a,z-d) = \begin{cases} a & a+c \le z \le a+d \\ z-d & a+d \le z \le b+d \end{cases}.$$

To resolve $\min(b,z-c) - \max(a,z-d)$, these intervals need to be intersected. For the sake of this, let us assume that $a+d \le b+c$, which can always be achieved by swapping the respective values of X and Y. Consequently, $f_Z(z)$ can be written as:

$$f_Z(z) = \begin{cases} 0 & , z < a+c \\ \dfrac{z-c-a}{(b-a)(d-c)} & , a+c \le z \le a+d \\ \dfrac{1}{b-a} & , a+d \le z \le b+c \\ \dfrac{-z+b+d}{(b-a)(d-c)} & , b+c \le z \le b+d \\ 0 & , b+d < z \end{cases}.$$

Hence, by integration, the CDF is defined as:

$$F_Z(z) = \begin{cases} 0 & ,z < a+c \\ \dfrac{0.5z^2 - (a+c)z + 0.5(a+c)^2}{(b-a)(d-c)} & ,a+c \le z \le a+d \\ \dfrac{0.5(a+d)^2 - (a+c)(a+d) + 0.5(a+c)^2}{(b-a)(d-c)} + \dfrac{z-(a+d)}{b-a} & ,a+d \le z \le b+c \\ \left[\dfrac{0.5(a+d)^2 - (a+c)(a+d) + 0.5(a+c)^2}{(b-a)(d-c)} + \dfrac{(b+c)-(a+d)}{b-a}\right. & \\ \left. + \dfrac{-0.5z^2 + (b+d)z + 0.5(b+c)^2 - (b+d)(b+c)}{(b-a)(d-c)}\right] & ,b+c \le z \le b+d \\ 1 & ,b+d < z \end{cases},$$

which can be simplified to:

$$F_Z(z) = \begin{cases} 0 & ,z < a+c \\ \dfrac{0.5z^2 - (a+c)z + 0.5(a+c)^2}{(b-a)(d-c)} & ,a+c \le z \le a+d \\ \dfrac{0.5(d-c)}{b-a} + \dfrac{z-(a+d)}{b-a} & ,a+d \le z \le b+c \\ \dfrac{-a+b+0.5(c-d)}{b-a} + \dfrac{-0.5\left(z^2-(b+c)^2\right) + (b+d)(z-(b+d))}{(b-a)(d-c)} & ,b+c \le z \le b+d \\ 1 & ,b+d < z \end{cases}.$$

Therefore, the CDF of $\tilde{P}_{ij} - \tilde{P}_{kl}$ will be a piecewise linear or quadratic function on the interval $(a_{ij} - b_{kl},\ b_{ij} - a_{kl})$.

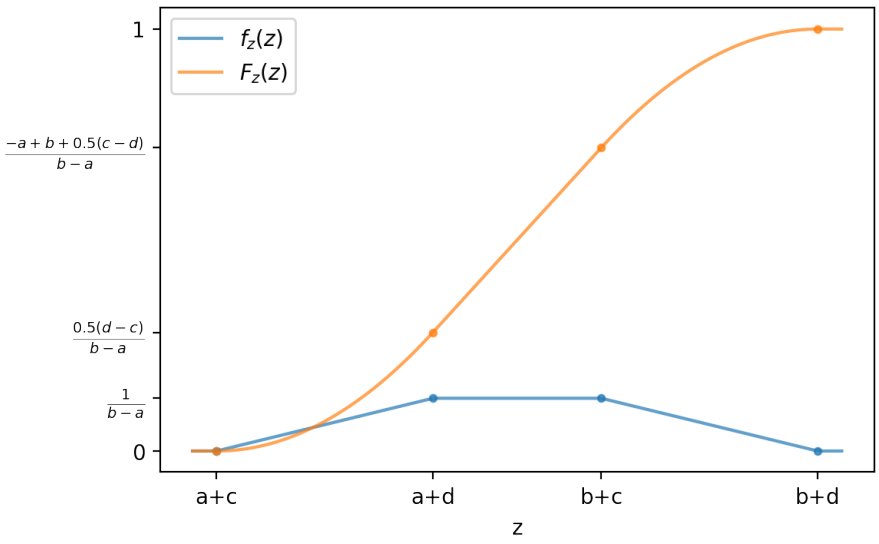

Figure 5: PDF and CDF of the sum of two uniformly distributed random variables.

To make further statements about $\mathbb{P}\left(\langle P^T_{PROM2_j} - P^T_{PROM2_i}, w\rangle \leq 0 \middle| w\right)$, the preference functions needs to be addressed.

Lemma: The probability of an alternative i being preferred over an alternative j in PROMETHEE II, denoted as $\mathbb{P}\left(\langle P^T_{PROM2_j} - P^T_{PROM2_i}, w\rangle \leq 0 \middle| w\right)$, is piecewise constant when the preference functions have a discrete codomain.

Proof: To investigate this probability, let us first reconsider the pairwise differences $\tilde{P}_{ij} - \tilde{P}_{kl}$. As shown before, if the performances are normally distributed, $\tilde{P}_{ij} \sim \mathcal{N}(\mu_{ij}, \sigma^2_{ij})$, and $\tilde{P}_{kl} \sim \mathcal{N}(\mu_{kl}, \sigma^2_{kl})$, then $\tilde{P}_{ij} - \tilde{P}_{kl} \sim \mathcal{N}(\mu_{ij} - \mu_{kl},\ \sigma^2_{ij} + \sigma^2_{kl})$. On the other hand, if the performances are uniformly distributed, $\tilde{P}_{ij} \sim \mathcal{U}(a_{ij}, b_{ij})$ and $\tilde{P}_{kl} \sim \mathcal{U}(a_{kl}, b_{kl})$, then the distribution of $\tilde{P}_{ij} - \tilde{P}_{kl}$ can be obtained by calculating their convolution, as previously shown.

However, since we are considering preference functions with a discrete codomain, the probability of $g(\tilde{P}_{ij} - \tilde{P}_{kl})$ taking a particular value, e.g. (0, 0.5 or 1), will be constant, independent of the specific distribution of $\tilde{P}_{ij} - \tilde{P}_{kl}$ and determined by the corresponding CDF.

If preference function one (**??**) is used, we observe that:

$$\mathbb{P}\left(g\left(\tilde{P}_{ik} - \tilde{P}_{jk}\right) = 0\right) = \begin{cases} 1 & , i = j \\ \mathbb{P}\left(\tilde{P}_{ik} - \tilde{P}_{jk} \leq 0\right) & , i \neq j \end{cases},$$

and $\mathbb{P}\left(g\left(\tilde{P}_{ik} - \tilde{P}_{jk}\right) = 1\right) = 1 - \mathbb{P}\left(g\left(\tilde{P}_{ik} - \tilde{P}_{jk}\right) = 0\right)$.

For preference function two (**??**), it can be seen that:

$$\mathbb{P}\left(g\left(\tilde{P}_{ik} - \tilde{P}_{jk}\right) = 0\right) = \begin{cases} 1 & , i = j \text{ and } q \leq 0 \\ 0 & , i = j \text{ and } q > 0 \\ \mathbb{P}\left(\tilde{P}_{ik} - \tilde{P}_{jk} \leq q\right) & , i \neq j \end{cases},$$

and also $\mathbb{P}\left(g\left(\tilde{P}_{ik} - \tilde{P}_{jk}\right) = 1\right) = 1 - \mathbb{P}\left(g\left(\tilde{P}_{ik} - \tilde{P}_{jk}\right) = 0\right)$.

Finally, using the fourth preference function (**??**), which is the last function with a discrete codomain, the probability can be calculated as:

$$\mathbb{P}\left(g\left(\tilde{P}_{ik} - \tilde{P}_{jk}\right) = 0\right) = \begin{cases} 1 & , i = j \text{ and } q \leq 0 \\ 0 & , i = j \text{ and } q > 0 \\ \mathbb{P}\left(\tilde{P}_{ik} - \tilde{P}_{jk} \leq q\right) & , i \neq j \end{cases}.$$

Furthermore, the probability of observing $\frac{1}{2}$ is given by:

$$\mathbb{P}\left(g\left(\tilde{P}_{ik}-\tilde{P}_{jk}\right)=\frac{1}{2}\right)=\begin{cases}1 & ,i=j \text{ and } q<0\leq p\\ 0 & ,i=j \text{ and } (q\geq 0 \text{ or } p<0)\\ \mathbb{P}\left(\tilde{P}_{ik}-\tilde{P}_{jk}\leq p\right)-\mathbb{P}\left(\tilde{P}_{ik}-\tilde{P}_{jk}<q\right) & ,i\neq j\end{cases},$$

and, lastly, the probability of obtaining 1 will be:

$$\mathbb{P}\left(g\left(\tilde{P}_{ik}-\tilde{P}_{jk}\right)=1\right)=\begin{cases}1 & ,i=j \text{ and } p<0\\ 0 & ,i=j \text{ and } p\geq 0\\ \mathbb{P}\left(\tilde{P}_{ik}-\tilde{P}_{jk}<q\right) & ,i\neq j\end{cases}.$$

Therefore, under the assumption that the preference functions have a discrete codomain, one could calculate the probabilities of all combinations of rows P_i and P_j which result in $\langle P^T_{PROM2_j}-P^T_{PROM2_i},w\rangle\leq 0$ and join their probability to obtain $\mathbb{P}(\langle P^T_{PROM2_j}-P^T_{PROM2_i},w\rangle\leq 0\big|w)$. It follows that only the number of possible combinations for P_i and P_j is dependent on w but not the associated probabilities, hence $\mathbb{P}(\langle P^T_{PROM2_j}-P^T_{PROM2_i},w\rangle\leq 0\big|w)$ will be piecewise constant. The specific probabilities associated with the pairwise differences, P_j-P_i, do not affect the piecewise constant property of the overall probability. □

Corollary: If PROMETHEE II is used and all preference functions have a discrete codomain, $\mathbb{P}(f(\tilde{P},w)=r|w)$ will be piecewise constant.

Proof of corollary: Instead of considering all combinations of two rows, the previous idea can be extended to all combinations of the whole matrix since $\mathbb{P}(f(\tilde{P},w)=r|w)$ just implies $\langle P^T_{PROM2_j}-P^T_{PROM2_i},w\rangle\leq 0$ for $n-1$ pairs of rows simultaneously. □

Concluding, since the number of combinations rises rapidly, as the number of alternatives and criteria increases, this approach might not be feasible for large performance matrices and Monte-Carlo-Simulation could have a better trade-off between accuracy and computational time. Additionally, this approach is not applicable for preference functions with a continuous codomain and therefore not suitable for a wide variety of possible preference functions.

References

[1] HASENAUER, Luis: Bayesian statistics in the context of Multi Criteria Decision Analysis. (2023)

[2] BALLIO, Francesco ; GUADAGNINI, Alberto: Convergence assessment of numerical Monte Carlo simulations in groundwater hydrology. In: *Water Resources Research* 40 (2004), No. 4. http://dx.doi.org/https://doi.org/10.1029/2003WR002876. – DOI https://doi.org/10.1029/2003WR002876

[3] WIEN, TU: *Konvergenzverhalten, Abschätzung des Simulationsfehlers*. https://www.iue.tuwien.ac.at/phd/harlander/node31.html, 2023

[4] SIEGRIST, Kyle: *The Multivariate Normal Distribution*. https://stats.libretexts.org/Bookshelves/Probability_Theory/Probability_Mathematical_Statistics_and_Stochastic_Processes_(Siegrist)/05%3A_Special_Distributions/5.07%3A_The_Multivariate_Normal_Distribution, 2022

[5] ROSS, Kevin: *Transformations of random variables*. https://bookdown.org/kevin_davisross/probsim-book/transformations-of-random-variables.html, 2022

[6] BRADLEY, David M. ; GUPTA, Ramesh C.: On the distribution of the sum of n non-identically distributed uniform random variables. In: *Ann Inst Stat Math* 54 (2002), No. 3, p. 689–700

[7] SCIPY: *Normaltest*. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.normaltest.html, 2005

[8] VENKATASUBRAMANIAN, Suresh: *Sampling from the simplex*. http://blog.geomblog.org/2005/10/sampling-from-simplex.html, 2005

[9] TSUN, Alex: *Multiple Random Variables*. https://courses.cs.washington.edu/courses/cse312/20su/files/student_drive/5.5.pdf, 2023