Management of uncertain pairwise comparisons in AHP through probabilistic concepts

J. Benítez[a], S. Carpitella[b], A. Certa[b], J. Izquierdo[a]

[a] *Instituto de Matemática Multidisciplinar (IMM), Universitat Politècnica de València*
[b] *Dipartimento dell'Innovazione Industriale e Digitale (DIID), Universitá degli Studi di Palermo*

Abstract

Fast and judicious decision-making is paramount for the success of many activities and processes. However, various degrees of difficulty may affect the achievement of effective and optimal solutions. Decisions should ideally meet the best trade-off among as many of the involved factors as possible, especially in the case of complex problems. Substantial cognitive and technical skills are indispensable, while not always sufficient, to carry out optimal evaluations. One of the most common causes of wrong decisions derives from uncertainty and vagueness in making forecasts or attributing judgments. The literature shows numerous efforts towards the optimization and modeling of uncertain contexts by means of probabilistic approaches. This paper proposes the use of probability theory to estimate uncertain expert judgments within the framework of the analytic hierarchy process and, more specifically, within a linearization scheme developed by the authors. After describing the necessary probabilistic concepts of interest, the main results are developed. These results can be summarized as using various kinds of random variables with uncertainty embodied in undecided pairwise comparisons. A case study focused on the maintenance management of an industrial water distribution system exemplifies the approach.

Key words: Decision making; uncertainty; probability; industrial management

1 Introduction and literature review

The output of a generic decision-making process consists in selecting and implementing the most appropriate solution to achieve the best level of performance. That decision is expected to be conducted in the most reliable way, seeking to maximize the effects derived from positive factors, and to simultaneously minimize the negative factors.

Broadly speaking, decision-making processes are continuously accomplished in problem-solving contexts to positively contribute to activities in business and in non-business fields. Results, especially in the case of competitive business processes, need to be continuously improved by undertaking faster and better decision-making. This is true, for instance, in the case of investment processes [12], for which advanced management methods should be studied and implemented with the aim of strengthening competitiveness and innovation. With this perspective, various degrees of difficulty may characterize the achievement of an effective solution. Indeed, the most important problems to be resolved are often the most complex as well.

When facing a highly complex problem, making a decision that represents the best trade-off among all the involved factors is not straightforward. Substantial cognitive and technical skills are needed to carry out optimal evaluations [14]. According to [17], one of the most common causes of such complexity derives from uncertainty and vagueness in making forecasts, or in attributing judgments concerning certain aspects of the decision to be made. The author underlines that contradictory conclusions may appear after changing methods and paradigms.

As asserted by Yager and Kreinovich [30], benefits related to a certain decision frequently depend on situations beyond our control, even when rigorous and reliable decision-making procedures are followed. Johnson *et al.* [13] accept that decisions are not often derived from a condition when the evidence is available. In fact, decision-makers may infer the most likely solution while being ignorant about relevant features concerning the problem under analysis.

Regarding this aspect, Shah *et al.* [27] observe that the literature mainly stresses how human judgment usually tends to underestimate the probability of negative consequences, being sometimes unrealistically overoptimistic. However, the authors apply five tests to observe this phenomenon without noticing traces of bias due to a general human tendency to optimism, thus confirming the vast complexity of human cognition. Proper methodologies should support this cognition, especially in the presence of missing information. For example, Soroudi *et al.* [28] face a problem of renewable electricity supply and highlight uncertainties due to the extremely volatile nature of wind power. In particular, they develop the Information Gap Decision Theory to properly handle unknown events. Regarding problems of multi-criteria nature, Pereira *et al.* [23] state the absolute need to formally model uncertainty with the support of a mathematical perspective, in contrast to the traditional and deterministic approach of many multi-criteria methods. In this context, Liu *et al.* [15] suggest undertaking decision-making problems by representing the relative attributes by means of uncertain linguistic variables in terms of fuzzy numbers. They develop a decision support method to solve practical problems with interval probabilities. Yan *et al.* [31] undertake a probabilistic interpretation of weights by implementing a linguistic decision rule through the concepts of random preference and stochastic dominance.

More generally, the literature shows plenty of efforts towards the optimization and modeling of uncertain contexts using various probabilistic approaches. ([8, 16, 19, 32, 33]). In a vast number of real situations and practical problems, it would be more appropriate to speak of “a probably good solution” rather than “the best solution”. As stressed by Biedermann *et al.* [6], a probabilistic approach helps to show decision-maker uncertainty about an unknown quantity or event, even if the personal interpretation of probability cannot be avoided. In this regard, Costello and Watts [7] develop a model to represent how people estimate conditional probabilities. Moreover, Izhakian [11] underlines the factor of ambiguity, whose degree may be interpreted as the volatility of probability. The author proposes a model to deal with uncertain event probabilities.

This paper proposes the use of probability theory to estimate uncertain expert judgments within the framework of the Analytic Hierarchy Process (AHP), [25, 26], an established technique for supporting various types of decision-making processes, as widely recognized in the literature ([9, 18, 22]). Indeed, as shown by Hughes [10], the probability theory fundamentals perfectly fit the properties of the AHP. By using suitable probabilistic concepts herein developed we provide a way to treat uncertainty embodied by pairwise comparisons as expressed through various kinds of random variables.

The research is organized as follows. Section 2 presents a brief review of the linearization process in AHP. Section 3 describes probabilistic concepts of interest in AHP and provides the main results of the paper. Section 4 illustrates a case study focused on maintenance management of an industrial water distribution system. Section 5 closes the work and provides some conclusions.

2 A brief review of the linearization process in AHP

Let $\mathscr{M}_n$ be the set of square $n \times n$ matrices and $\mathscr{M}_n^+$ the subset of $\mathscr{M}_n$ composed of matrices whose entries are positive. A matrix $A = (a_{ij}) \in \mathscr{M}_n^+$ is *reciprocal* when $a_{ij}a_{ji} = 1$ for all indices i, j. A matrix $A \in \mathscr{M}_n^+$ is *consistent* when $a_{ij}a_{jk} = a_{ik}$ for all indices i, j, k. It is evident that any consistent matrix is reciprocal. The trace and the transpose of a matrix A will be denoted by $\mathrm{tr}(A)$ and A^T, respectively. We assume that any vector of $\mathbb{R}^n$ is a column vector.

An important problem in AHP theory is the following: given a reciprocal matrix (which can be easily issued by an expert), find its closest consistent matrix. Consistent matrices in AHP theory are important because they provide a way to rank (also to give weights to) the several alternatives in hand. This is achieved by the Perron vector of the consistent matrix, i.e., the eigenvector associated to $\lambda_{\max}$, where $\lambda_{\max}$ is the largest (in modulus) eigenvalue. The reader is encouraged to consult [25, 26] for a deeper insight of AHP theory.

To define precisely "the closest consistent matrix to another matrix", we previously must give a distance. We define in $\mathscr{M}_n^+$ the distance

$$\mathrm{d}(X, Y) = \|L(X) - L(Y)\|_F, \tag{1}$$

where $L : \mathscr{M}_n^+ \to \mathscr{M}_n$ is the mapping defined as $L(A)_{ij} = \log(a_{ij})$ and $\|\cdot\|_F$ is the Frobenius norm (i.e., $\|A\|_F = \sqrt{\mathrm{tr}(AA^T)}$). The inverse of the mapping L will also play a fundamental role in the sequel; this mapping $E : \mathscr{M}_n \to \mathscr{M}_n^+$ is the component-wise exponential, i.e., $E(A)_{ij} = \exp(a_{ij})$. An obvious consequence from the above definitions is that A is reciprocal if and only if $L(A)$ is skew-Hermitian. We define $\mathscr{L}_n = \{L(A) : A \in \mathscr{M}_n^+,\ A \text{ is consistent}\}$. It can be proved (see [3]) that $\mathscr{L}_n$ is a linear subspace of $\mathscr{M}_n$ whose dimension equals $n-1$. Evidently, $X \in \mathscr{L}_n$ if and only if $E(X)$ is consistent.

The closest consistent matrix to a given reciprocal matrix A (according to the aforementioned distance) is given by $E(p_n(L(A)))$, where $p_n : \mathscr{M}_n \to \mathscr{L}_n$ is the orthogonal projection onto $\mathscr{L}_n$. Of course, when we use the word "orthogonal" in the set $\mathscr{M}_n$, we must specify the considered inner product: for $A, B \in \mathscr{M}_n$, we set $\langle A, B\rangle = \mathrm{tr}(AB^T)$. Of course, $\|A\|_F^2 = \langle A, A\rangle$. We have the following scheme:

$$\mathscr{M}_n^+ \xrightarrow{L} \mathscr{M}_n \xrightarrow{p_n} \mathscr{L}_n \xrightarrow{E} \mathscr{M}_n^+$$

As one can infer from the above diagram, in AHP theory, only projections of matrices of the form $L(A)$, A being reciprocal, are needed. Such projections may be easily obtained from a result provided in [4], where a simple formula for $p_n(M)$ is given for any skew-Hermitian matrix M. Given a consistent matrix $B = (b_{ij}) \in \mathscr{M}_n^+$, its Perron vector $\mathbf{x} = [x_1, \ldots, x_n]^T \in \mathbb{R}^n$ —which, as said, is paramount in AHP theory— must satisfy $x_i/x_j = b_{ij}$ for any pair of indices i, j. The equality $x_i/x_j = b_{ij}$ for arbitrary indices can be written in a shorter way by $\mathbf{x}J(\mathbf{x})^T = B$, where the mapping $J : \mathbb{R}^n \to \mathbb{R}^n$ is given by $(J(\mathbf{x}))_i = 1/x_i$ —defined only when $x_i \neq 0$, $i = 1, \ldots, n$. The next theorem enables finding the priority vector of the closest consistent matrix to a given reciprocal matrix (see [4]).

Theorem 1 *Let $A = (a_{ij}) \in \mathscr{M}_n^+$ be a reciprocal matrix and $\mathbf{x} \in \mathbb{R}^n$ Then $\mathbf{x}J(\mathbf{x})^T = E(p_n(L(A)))$ if and only if there exists $C > 0$ such that $\mathbf{x} = C[x_1, \ldots, x_n]^T$, where $x_i = \sqrt[n]{a_{i1} \cdots a_{in}}$.*

3 AHP and probabilistic related concepts

When an expert has doubts in assigning a specific value to an entry in a reciprocal matrix, then the idea of using random instead of constant entries is suggested. Consequently, we will consider matrices $A = (a_{ij})$ whose components can be random variables. Another use of random variables in AHP can be the following: imagine that two experts express their judgements and thus form two reciprocal matrices, say $A = (a_{ij})$ and $B = (b_{ij})$. If there exists $i \neq j$ with $a_{ij} \neq b_{ij}$, then one can consider a discrete random variable X such that $\mathrm{pr}(X = a_{ij}) = w_A$ and $\mathrm{pr}(X = b_{ij}) = w_B$, where w_A, w_B are the respective weights given to the experts (of course, $0 \leq w_A, w_B \leq 1$, $w_A + w_B = 1$).

A *random reciprocal matrix* is an $n \times n$ matrix $A = (a_{ij})$ whose entries are positive random variables whose expectation and variances are finite and $a_{ij}a_{ji} = 1$, see [29].

Let $B = (b_{ij})$ be the closest consistent matrix to A (in the sense of the distance defined in (1)). What can be said about b_{ij}? And about the priority vector? These questions will be dealt with in this section.

The *expectation* and *variance* of a random variable X will be denoted by $\mathsf{E}(X)$ and $\mathsf{Var}(X)$, respectively. The *covariance* of the random variables X and Y will be denoted by $\mathsf{Cov}(X, Y)$. Throughout this article, when we write E, Var, or Cov we will assume that these numbers are finite. To deal with random reciprocal matrices, in view of Theorem 1, it is plausible that the geometric mean is more natural than the arithmetic mean. Another reason is the following: if $A = (a_{ij})$ is a positive random matrix, since $a_{ij} = 1/a_{ji}$, then it is natural that "mean of (a_{ij}) = 1/mean of (a_{ji})" holds. However, this property does not hold when the mean is the expectation E. Since the function $x \mapsto 1/x$ is convex, then, by Jensen's inequality, one has $\mathsf{E}(X)^{-1} \leq \mathsf{E}(X^{-1})$, and the equality holds if, and only if $\mathsf{Var}(X) = 0$. Therefore, we shall use another kind of expectation, which is defined next.

3.1 The geometric expectation and AHP

Given a positive random variable X, we define the *geometric expectation* by

$$\mathsf{G}(X) = \exp(\mathsf{E}(\log X)).$$

Equivalently, $\log[\mathsf{G}(X)] = \mathsf{E}[\log(X)]$. This expectation has found several applications in economics, see, e.g., [2, 20]. From the very well-known properties of the expectation, one can give the following result.

Theorem 2 *Let X, Y be positive random variables. Then*

(i) $\mathsf{G}(aX^b) = a\,\mathsf{G}(X)^b$, *for constants $a > 0$ and $b \in \mathbb{R}$.*

(ii) $\mathsf{G}(XY) = \mathsf{G}(X)\,\mathsf{G}(Y)$.

In particular, if X is positive, then $\mathsf{G}(X^{-1}) = \mathsf{G}(X)^{-1}$. By Jensen's inequality, since $x \mapsto \log x$ is a concave function, we have $\log[\mathsf{E}(X)] \geq \mathsf{E}[\log X] = \log[\mathsf{G}(X)]$, i.e., $\mathsf{G}(X) \leq \mathsf{E}(X)$, and the inequality becomes an equality if and only if there exists $c \in \mathbb{R}$ such that $\mathrm{pr}(X = c) = 1$.

Theorem 3 *Let $A = (a_{ij})$ be an $n \times n$ reciprocal random matrix. Let $B = (b_{ij})$ the closest consistent matrix in the sense of the distance defined in (1). If $\mathbf{x} = [x_1, \ldots, x_n]^T$ is a random priority vector of the matrix B, then*

$$\mathsf{G}(b_{ij}) = \frac{\mathsf{G}(x_i)}{\mathsf{G}(x_j)}$$

and there exists $C \in \mathbb{R}$ such that

$$\mathsf{G}(x_i) = C\sqrt[n]{\mathsf{G}(a_{i1})\cdots\mathsf{G}(a_{in})}.$$

PROOF: The expression for $\mathsf{G}(x_i)$ follows from Theorems 1 and 2. The expression for $\mathsf{G}(b_{ij})$ follows from $B = \mathbf{x}J(\mathbf{x})^T$ and Theorem 2. □

Observe that in the above theorem there is no need to assume that the judgements in matrix A have to be independent.

3.2 The geometric variance, the geometric covariance and AHP

Measures of deviation from the geometric expected value $\mathsf{G}(X)$ analogous to the variance of X can be defined. For a given positive random variable X, we define the *geometric variance* as follows:

$$\mathsf{Var}_g(X) = \mathsf{Var}(\log X). \tag{2}$$

In some textbooks, the expression $\exp(\mathsf{Var}(\log X))$ can be found as the definition for the geometric variance; however, (2) is easier to handle. Obviously, $\mathsf{Var}_g(X) \geq 0$ and $\mathsf{Var}_g(X) = 0$ if and only if there exists $c \in \mathbb{R}^+$ such that $\mathrm{pr}(X = c) = 1$.

We shall give two examples to show why we will not use the "usual" variance and why we suggest using the geometric variance.

1. Let us consider the following two situations:

   (i) a_{12} is the discrete random variable such that $\mathrm{pr}(a_{12} = 1) = \mathrm{pr}(a_{12} = 2) = 1/2$.

   (ii) b_{12} is the discrete random variable such that $\mathrm{pr}(b_{12} = 8) = \mathrm{pr}(b_{12} = 9) = 1/2$.

   In the first situation, the expert has doubts between "equal importance" and "weak importance" (in [26] one can find the fundamental scale in AHP proposed by Saaty). In the second situation, the expert's doubts are much smaller (his/her doubts vary between "major importance" and "extreme importance").

   However, $\mathsf{Var}(a_{12}) = \mathsf{Var}(b_{12})$ —as one can trivially deduce from the expression $\mathsf{Var}(X + k) = \mathsf{Var}(X)$, where X is a random variable and $k \in \mathbb{R}$ is a constant. This fact is not intuitive since the expert's doubts in the first situation are greater than in the second situation. In contrast, one has $\mathsf{Var}_g(a_{12}) = 0.12011$ and $\mathsf{Var}_g(b_{12}) = 0.00347$.

2. In AHP theory, if $A = (a_{ij})$ is a reciprocal matrix, then $a_{ij} = 1/a_{ji}$. Therefore, it must be intuitive that "variance of $1/X$ = variance of X". However, the "usual variance" does not satisfy this property (a trivial example is the random variable X such that $\mathrm{pr}(X = 1) = \mathrm{pr}(X = 2) = 1/2$). Instead, we will see that the geometric variance does satisfy this property (see item (i) of Theorem 4).

The following is a step further in the same line of definitions. Given two positive random variables X and Y, the *geometric covariance* of X and Y is defined as

$$\mathsf{Cov}_g(X, Y) = \mathsf{Cov}(\log X, \log Y).$$

We next prove several properties of the geometric variance and geometric covariance.

Theorem 4 *Let X and Y be positive random variables.*

(i) $\mathsf{Var}_g(X^r) = r^2 \mathsf{Var}_g(X)$, *where* $r \in \mathbb{R}$ *is a constant.*

(ii) $\mathsf{Var}_g(XY) = \mathsf{Var}_g(X) + \mathsf{Var}_g(Y) + 2\,\mathsf{Cov}_g(X,Y)$.

(iii) *If* X *and* Y *are independent,* $\mathsf{Var}_g(XY) = \mathsf{Var}_g(X) + \mathsf{Var}_g(Y)$.

(iv) *If* $X_1, \dots, X_n$ *and* $Y_1, \dots, Y_m$ *are positive random variables, and* $a_1, \dots, a_n$, $b_1, \dots, b_m$ *are real constants, then* $\mathsf{Cov}_g(\prod_{i=1}^n X_i^{a_i}, \prod_{j=1}^m Y_j^{b_j}) = \sum_{i,j} a_i b_j \,\mathsf{Cov}_g(X_i, Y_j)$.

(v) *If* A *is a positive constant, then* $\mathsf{Var}_g(AX) = \mathsf{Var}_g(X)$ *and* $\mathsf{Cov}_g(A,X) = 0$.

PROOF: (i): We use that if Z is a random variable and $a \in \mathbb{R}$, then $\mathsf{Var}(aZ) = a^2 \mathsf{Var}(Z)$.

$$\mathsf{Var}_g(X^r) = \mathsf{Var}(\log X^r) = \mathsf{Var}(r \log X) = r^2 \mathsf{Var}(\log X) = r^2 \mathsf{Var}_g(X).$$

(ii): By the pevious definitions and known properties of the variance, we have

$$\begin{aligned}\mathsf{Var}_g(XY) &= \mathsf{Var}[\log(XY)] \\ &= \mathsf{Var}(\log X + \log Y) \\ &= \mathsf{Var}(\log X) + \mathsf{Var}(\log Y) + 2\,\mathsf{Cov}(\log X, \log Y) \\ &= \mathsf{Var}_g(X) + \mathsf{Var}_g(Y) + 2\,\mathsf{Cov}_g(X,Y).\end{aligned}$$

(iii): Since X and Y are independent, $\log X$ and $\log Y$ are also independent, hence the covariance of $\log X$ and $\log Y$ is zero. The conclusion follows from the computation made in the proof of (ii).

(iv): It follows from the definition of the geometric covariance and the property

$$\mathsf{Cov}\left(\sum_i a_i X_i, \sum_j b_j Y_j\right) = \sum_{i,j} a_i b_j \,\mathsf{Cov}(X_i, Y_j),$$

which is valid for arbitrary random variables X_i, Y_j and constants a_i, b_j.

(v): Since A is a constant, using the properties of the expectation,

$$\begin{aligned}\mathsf{Cov}_g(A,X) &= \mathsf{Cov}(\log A, \log X) \\ &= \mathsf{E}(\log A \log X) - \mathsf{E}(\log A)\,\mathsf{E}(\log X) = (\log A)\,\mathsf{E}(\log X) - (\log A)\,\mathsf{E}(\log X) = 0.\end{aligned}$$

The theorem is proved. □

Property (ii) above can be generalized by applying the formula of the variance of the sum of n random variables. If $X_1, \dots, X_n$ are positive random variables, then

$$\mathsf{Var}_g(X_1 \cdots X_n) = \sum_{i=1}^n \mathsf{Var}_g(X_i) + 2\sum_{i<j} \mathsf{Cov}_g(X_i, X_j) \tag{3}$$

and if $X_1, \dots, X_n$ are pairwise independent, then

$$\mathsf{Var}_g(X_1 \cdots X_n) = \sum_{i=1}^n \mathsf{Var}_g(X_i).$$

Now we give the geometric variance of the closest consistent matrix to a given random reciprocal matrix.

Theorem 5 *Let $A = (a_{ij})$ be an $n \times n$ reciprocal random matrix. Let $B = (b_{ij})$ be the closest consistent matrix in the sense of the distance defined in (1). If $\mathbf{x} = [x_1, \ldots, x_n]^T$ is a random vector which is a priority vector of the matrix B, then*

$$\mathsf{Var}_g(b_{ij}) = \mathsf{Var}_g(x_i) + \mathsf{Var}_g(x_j) - 2\,\mathsf{Cov}_g(x_i, x_j),$$

$$\mathsf{Cov}_g(b_{ij}, b_{rs}) = \mathsf{Cov}_g(x_i, x_r) - \mathsf{Cov}_g(x_i, x_s) - \mathsf{Cov}_g(x_j, x_r) + \mathsf{Cov}_g(x_j, x_s),$$

$$\mathsf{Var}_g(x_i) = \frac{1}{n^2}\left[\sum_{j=1}^{n} \mathsf{Var}_g(a_{ij}) + 2\sum_{j<k} \mathsf{Cov}_g(a_{ij}, a_{ik})\right],$$

and

$$\mathsf{Cov}_g(x_i, x_j) = \frac{1}{n^2}\sum_{r,s} \mathsf{Cov}_g(a_{ir}, a_{js}). \tag{4}$$

PROOF: Since $b_{ij} = x_i/x_j$, it follows from Theorem 4 that

$$\begin{aligned}\mathsf{Var}_g(b_{ij}) &= \mathsf{Var}_g(x_i x_j^{-1}) \\ &= \mathsf{Var}_g(x_i) + \mathsf{Var}_g(x_j^{-1}) + 2\,\mathsf{Cov}_g(x_i, x_j^{-1}) = \mathsf{Var}_g(x_i) + \mathsf{Var}_g(x_j) - 2\,\mathsf{Cov}_g(x_i, x_j).\end{aligned}$$

In an analogous way, we can prove the expression of $\mathsf{Cov}_g(b_{ij}, b_{rs})$. If, in addition, we use (3) and Theorem 1, the remaining expressions can be similarly proved. □

Given a random reciprocal matrix $A = (a_{ij})$, it is reasonable to assume that a_{ij} are independent for $1 \le i < j \le n$ (see [24]).

Corollary 1 *Under the notation of Theorem 5, if a_{ij} are pairwise independent for $1 \le i < j \le n$, then*

$$\mathsf{Var}_g(x_i) = \frac{1}{n^2}\sum_{j=1}^{n} \mathsf{Var}_g(a_{ij}), \quad i = 1, \ldots, n,$$

and

$$\mathsf{Cov}_g(x_i, x_j) = -\frac{1}{n^2}\mathsf{Var}_g(a_{ij}), \quad i, j = 1, \ldots, n,\ i \neq j.$$

PROOF: By the independence hypothesis, if $\mathsf{Cov}_g(a_{ir}, a_{js}) \neq 0$, then $(i, r) = (j, s)$ or $(i, r) = (s, j)$. The expression for the geometric variance follows from Theorem 5. To complete the proof, if $i \neq j$, then the unique non vanishing term on the right hand side of (4) corresponds to $(i, r) = (s, j)$, which is $\mathsf{Cov}_g(a_{ir}, a_{js}) = \mathsf{Cov}_g(a_{ij}, a_{ji}) = \mathsf{Cov}_g(a_{ij}, a_{ij}^{-1}) = -\mathsf{Cov}_g(a_{ij}, a_{ij}) = -\mathsf{Var}_g(a_{ij})$. □

If $\mathbf{x} = [x_1, \ldots, x_n]^T$ is a vector of random variables, we define the matrix whose (i, j)-entry is $\mathsf{Cov}_g(x_i, x_j)$. This matrix will be named as the *geometric variance-covariance matrix* of $\mathbf{x}$ and denoted from now on by $\Sigma_g(\mathbf{x})$. Notice that $\mathsf{Cov}_g(x_i, x_i) = \mathsf{Var}_g(x_i)$. Observe that the geometric variance of b_{ij} can be computed by using the geometric variance-covariance matrix and Theorem 5. If $\mathbf{d}_{ij}$ denotes the column vector of $\mathbb{R}^n$ whose ith component is 1 and whose jth component is -1, and its remaining components are 0, then $\mathsf{Cov}_g(b_{ij}, b_{rs}) = \mathbf{d}_{ij}^T \Sigma_g(\mathbf{x}) \mathbf{d}_{rs}$.

The importance of the random variables b_{ij} comes from the fact that these random variables are useful to rank the priorities. Recall that if a priority vector of the consistent matrix $B = (b_{ij})$ is $\mathbf{x} = [x_1, \ldots, x_n]^T$, then $b_{ij} = x_i/x_j$. Hence, $b_{ij} > 1$ if and only if $x_i > x_j$ and, thus, $\mathrm{pr}(b_{ij} > 1)$ is the probability of the ith alternative being preferred to the jth alternative. Also, the random variables b_{ij} are useful to rank a complete order of preferences: for example, $x_i > x_j > x_k \iff b_{ij} > 1$ and $b_{jk} > 1$; thus, that rank order can be evaluated by finding $\mathrm{pr}(b_{ij} > 1$ and $b_{jk} > 1)$.

3.3 Chebyshev's inequalities and their applications in AHP

There are basic inequalities in probability theory used to give bounds for certain probabilities. These inequalities are important because they provide useful information about *arbitrary* random variables. Chebyshev's inequality says that the probability that a random variable X is outside the interval $[\mathsf{E}(X)-\varepsilon, \mathsf{E}(X)+\varepsilon]$ is negligible if $\mathsf{Var}(X)/\varepsilon^2$ is small enough. Precisely, we have that for any $\varepsilon > 0$,

$$\mathrm{pr}(|X - \mathsf{E}(X)| \geq \varepsilon) \leq \frac{\mathsf{Var}(X)}{\varepsilon^2}.$$

We give now a similar inequality concerning the geometric expectation and variance.

Theorem 6 *Let X be a positive random variable. For any $u > 0$ one has*

$$\mathrm{pr}(\mathrm{e}^{-u} < X / \mathsf{G}(X) < \mathrm{e}^{u}) \geq 1 - \frac{\mathsf{Var}_g(X)}{u^2}.$$

PROOF: Since log is an increasing function,

$$\begin{aligned}
\mathrm{pr}(\mathrm{e}^{-u} < X / \mathsf{G}(X) < \mathrm{e}^{u}) &= \mathrm{pr}(\mathrm{e}^{-u}\, \mathsf{G}(X) < X < \mathrm{e}^{u}\, \mathsf{G}(X)) \\
&= \mathrm{pr}(-u + \log \mathsf{G}(X) < \log X < u + \log \mathsf{G}(X)) \\
&= \mathrm{pr}\,(|\log X - \log \mathsf{G}(X)| < u) \\
&= \mathrm{pr}\,(|\log X - \mathsf{E}(\log X)| < u) = 1 - \mathrm{pr}\,(|\log X - \mathsf{E}(\log X)| \geq u)\,.
\end{aligned}$$

From Chebyshev's inequality, one has

$$\mathrm{pr}\,(|\log X - \mathsf{E}(\log X)| \geq u) \leq \frac{\mathsf{Var}(\log X)}{u^2} = \frac{\mathsf{Var}_g(X)}{u^2}.$$

Therefore, the conclusion of the theorem follows. □

In [1] it is proven the following two dimensional version of Chebyshev's inequality.

Theorem 7 *Let X and Y be two random variables and $\varepsilon > 0$. Then*

$$\mathrm{pr}(|X - \mu_x| \geq \varepsilon\sigma_x \ \text{ or } \ |Y - \mu_y| \geq \varepsilon\sigma_y) \leq \frac{1 + \sqrt{1-\rho^2}}{\varepsilon^2},$$

where $\mu_x = \mathsf{E}(X)$, $\mu_y = \mathsf{E}(Y)$, $\sigma_x^2 = \mathsf{Var}(X)$, $\sigma_y^2 = \mathsf{Var}(Y)$, and ρ is the correlation between X and Y, i.e., $\rho = \mathsf{Cov}(X,Y)/\sqrt{\mathsf{Var}(X)}\sqrt{\mathsf{Var}(Y)}$.

We include in the appendix the proof of this theorem for the sake of readability. We now give a related theorem (in the context of this paper) that gives bounds for some probabilities.

Theorem 8 *Let X e Y be positive random variables. If $\varepsilon > 0$, then*

$$\mathrm{pr}\left(\mathrm{e}^{-\varepsilon\,\mathsf{Var}_g(X)} < \frac{X}{\mathsf{G}(X)} < \mathrm{e}^{\varepsilon\,\mathsf{Var}_g(X)} \ \text{ and } \ \mathrm{e}^{-\varepsilon\,\mathsf{Var}_g(Y)} < \frac{Y}{\mathsf{G}(Y)} < \mathrm{e}^{\varepsilon\,\mathsf{Var}_g(Y)}\right) \geq 1 - \frac{1+\sqrt{1-\rho^2}}{\varepsilon^2},$$

where ρ is the correlation between $\log X$ and $\log Y$.

PROOF: Let $\omega_x = \mathsf{Var}_g(X)$ and $\omega_y = \mathsf{Var}_g(Y)$. Since $x \mapsto \log x$ is a non decreasing function, then

$$\begin{aligned} e^{-\varepsilon\omega_x} < X/\,\mathsf{G}(X) < e^{\varepsilon\omega_x} &\iff -\varepsilon\omega_x + \mathsf{E}(\log X) < \log X < \varepsilon\omega_x + \mathsf{E}(\log X) \\ &\iff |\log X - \mathsf{E}(\log X)| < \varepsilon\omega_x \end{aligned}$$

and, similarly for $Y/\,\mathsf{G}(Y)$. Therefore,

$$\begin{aligned} &\mathrm{pr}\left(e^{-\varepsilon\omega_x} < \frac{X}{\mathsf{G}(X)} < e^{\varepsilon\omega_x} \text{ and } e^{-\varepsilon\omega_y} < \frac{Y}{\mathsf{G}(Y)} < e^{\varepsilon\omega_y}\right) \\ &\quad = \mathrm{pr}\big(|\log X - \mathsf{E}(\log X)| < \varepsilon\omega_x \text{ and } |\log Y - \mathsf{E}(\log Y)| < \varepsilon\omega_y\big) \\ &\quad = 1 - \mathrm{pr}\big(|\log X - \mathsf{E}(\log X)| \geq \varepsilon\omega_x \text{ or } |\log Y - \mathsf{E}(\log Y)| \geq \varepsilon\omega_y\big). \end{aligned}$$

Recall that $\omega_x = \mathsf{Var}_g(X) = \mathsf{Var}(\log X)$ and $\omega_y = \mathsf{Var}(\log Y)$; hence the conclusion of the theorem follows from Theorem 7. □

Example. Let us consider the following random reciprocal matrix

$$A = \begin{bmatrix} 1 & a_{12} & 2 \\ a_{12}^{-1} & 1 & 3 \\ 1/2 & 1/3 & 1 \end{bmatrix},$$

where a_{12} is a positive random variable. For the sake of conciseness, we denote $\gamma = \mathsf{G}(a_{12})$ and $\omega = \mathsf{Var}_g(a_{12})$. Note that by Theorem 2, one has that $\mathsf{G}(a_{12}^{-1}) = 1/\gamma$. Let $\mathbf{x} = [x_1\ x_2\ x_3]^T$ be the priority vector of the closest consistent matrix to A. By Theorem 3, exists $C > 0$ such that

$$\mathsf{G}(x_1) = C\sqrt[3]{2\gamma}, \qquad \mathsf{G}(x_2) = C\sqrt[3]{3/\gamma}, \qquad \mathsf{G}(x_3) = C\sqrt[3]{1/6}. \tag{5}$$

Let $B = \mathbf{x}J(\mathbf{x})^T = (b_{ij})$ be the closest consistent matrix to A. If G denotes the 3×3 matrix whose (i,j) entry is $\mathsf{G}(b_{ij})$, then by Theorem 3, $\mathsf{G}(b_{ij}) = \mathsf{G}(x_i)/\,\mathsf{G}(x_j)$, hence

$$G = \begin{bmatrix} 1 & \sqrt[3]{2\gamma^2/3} & \sqrt[3]{12\gamma} \\ \sqrt[3]{3/2\gamma^2} & 1 & \sqrt[3]{18/\gamma} \\ \sqrt[3]{1/12\gamma} & \sqrt[3]{\gamma/18} & 1 \end{bmatrix}. \tag{6}$$

By Theorems 4 and 5, one has that

$$\mathsf{Var}_g(x_1) = \frac{1}{9}\mathsf{Var}_g(a_{12}) = \frac{\omega}{9}, \quad \mathsf{Var}_g(x_2) = \frac{1}{9}\mathsf{Var}_g(a_{12}^{-1}) = \frac{1}{9}\mathsf{Var}_g(a_{12}) = \frac{\omega}{9}, \quad \mathsf{Var}_g(x_3) = 0.$$

Now we write the variance-covariance geometric matrix of the random vector $\mathbf{x}$, denoted by $\Sigma_g(\mathbf{x})$. From the previous computations we know the entries of the main diagonal of $\Sigma_g(\mathbf{x})$ because $\mathsf{Cov}_g(x_i, x_i) = \mathsf{Var}_g(x_i)$. By property (v) of Theorem 4, the unique non vanishing term in the left hand side of $\mathsf{Cov}_g(x_1, x_2) = n^{-2}\sum_{r,s}\mathsf{Cov}_g(a_{1r}, a_{2s})$ is $\mathsf{Cov}_g(a_{12}, a_{21})$. But

$$\mathsf{Cov}_g(a_{12}, a_{21}) = \mathsf{Cov}_g(a_{12}, a_{12}^{-1}) = -\mathsf{Cov}_g(a_{12}, a_{12}) = -\mathsf{Var}_g(a_{12}) = -\omega.$$

Since $\Sigma_g(\mathbf{x})$ is symmetric, $\mathsf{Cov}_g(a_{21}, a_{12}) = -\omega$. Finally, since the third row of A is composed of constants, then the third row and the third column of $\Sigma_g(\mathbf{x})$ must be filled with zeroes, because from item (v) of Theorem 4 and Theorem 5,

$$\mathsf{Cov}_g(x_3, x_i) = \frac{1}{3^2}\sum_{r,s}\mathsf{Cov}_g(a_{3r}, a_{is}) = 0.$$

Thus,

$$\Sigma_g(\mathbf{x}) = \frac{1}{9}\begin{bmatrix} \omega & -\omega & 0 \\ -\omega & \omega & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{7}$$

If V is the 3×3 matrix whose (i,j) entry is $\mathsf{Var}_g(b_{ij})$, again by Theorem 5, we have

$$V_{12} = \mathsf{Var}_g(b_{12}) = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \Sigma_g(\mathbf{x}) \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \frac{4\omega}{9}.$$

The remaining entries of V can be similarly computed and we can obtain

$$V = \frac{\omega}{9}\begin{bmatrix} 0 & 4 & 1 \\ 4 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \tag{8}$$

Finally, we will find $\mathsf{Cov}_g(b_{ij}, b_{rs})$ for $1 \le i < j \le n$, $1 \le r < s \le n$, and $(i,j) \ne (r,s)$. By Theorem 5 and (7),

$$\mathsf{Cov}_g(b_{12}, b_{13}) = \mathbf{d}_{12}^T \Sigma_g(\mathbf{x}) \mathbf{d}_{13} = \begin{bmatrix} 1 & -1 & 0 \end{bmatrix} \Sigma_g(\mathbf{x}) \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \frac{2\omega}{9}.$$

Similarly, we obtain $\mathsf{Cov}_g(b_{12}, b_{23}) = -2\omega/9$ and $\mathsf{Cov}_g(b_{13}, b_{23}) = -\omega/9$. Observe that there is no need to compute more covariances because $\mathsf{Cov}_g(X,X) = \mathsf{Var}_g(X)$, $\mathsf{Cov}_g(X, Y^{-1}) = -\mathsf{Cov}_g(X,Y)$, and $\mathsf{Cov}_g(k,X) = 0$ when X, Y are positive random variables and $k \in \mathbb{R}$ is a constant.

We will use Theorem 6 to study the random variable b_{12} (recall that this random variable is the $(1,2)$ entry of B, which is the closest consistent matrix to the given reciprocal matrix A). Let $u > 0$. We know that

$$\mathrm{pr}\left(\mathrm{e}^{-u} \cdot \sqrt[3]{2\gamma^2/3} < b_{12} < \mathrm{e}^{u} \cdot \sqrt[3]{2\gamma^2/3}\right) \ge 1 - \frac{4\omega/9}{u^2}. \tag{9}$$

To fix ideas, let us assume that the expert has no preference between $a_{12} = 5$ or $a_{12} = 6$. Thus, it is natural to say that a_{12} is a random variable such that $\mathrm{pr}(5 \le a_{12} \le 6) = 1$ and $\mathsf{G}(a_{12})$ is the geometric mean of 5 and 6, i.e., $\gamma = \mathsf{G}(a_{12}) = \sqrt{5 \cdot 6} = \sqrt{30} \simeq 5.477$.

To give a value to $\mathsf{Var}_g(a_{12})$, let us consider that the larger the variance of a random variable, the worse the behaviour of X. Moreover, since $\mathrm{pr}(\log 5 \le \log a_{12} \le \log 6) = 1$, then $\mathsf{Var}_g(a_{12}) = \mathsf{Var}(\log a_{12}) \le (\log 6 - \log 5)^2/4 \simeq 0.00831$ (see [5]). We will assume the worst situation: $\omega = \mathsf{Var}_g(a_{12}) = 0.00831$.

We use (5) to get that the random (non normalised) vector $\mathbf{x}$ of priorities satisfies

$$C[\mathsf{G}(x_1) \;\; \mathsf{G}(x_2) \;\; \mathsf{G}(x_3)] \simeq C[2.221 \;\; 0.8182 \;\; 0.5503].$$

The geometric variance-covariance matrix of $\mathbf{x}$ is given in (7). If B is the nearest consistent matrix to A, then the geometric mean of the entries of B is given by (6); especifically, in this example we have

$$G \simeq \begin{bmatrix} 1 & 2.714 & 4.036 \\ 0.3684 & 1 & 1.487 \\ 0.2478 & 0.6727 & 1 \end{bmatrix},$$

and the matrix of the variances $(\mathsf{Var}_g(b_{ij}))$ is given in (8).

We will use Theorem 6 to exemplify about the preference order between the first and the second alternative (the remaining orders can be dealt with analogously). From (9) we have for any $u > 0$ that

$$\mathrm{pr}(2.714 \cdot \mathrm{e}^{-u} \leq b_{12} \leq 2.714 \cdot \mathrm{e}^{u}) \geq 1 - \frac{0.003693}{u^2}.$$

We list some concrete values of u to see the goodness of these the bounds.

Value of u	Interval of b_{12}	Lower bound of the probability
0.7	$[1.347, 5.466]$	0.99246
0.3	$[2.011, 3.664]$	0.95896
0.15	$[2.336, 3.153]$	0.83585

We can see that $\mathrm{pr}(x_1 < x_2) = \mathrm{pr}(x_1 x_2^{-1} < 1) = \mathrm{pr}(b_{12} < 1) < \mathrm{pr}(b_{12} \notin [1.347, 5, 466])$, hence $\mathrm{pr}(x_1 < x_2)$ is very small. What is more, $\mathrm{pr}(x_1 < 2x_2) = \mathrm{pr}(b_{12} < 2) < \mathrm{pr}(b_{12} \notin [2.011, 3.644]) < 1 - 0.95896 \simeq 0.041$, almost negligible.

Now we study the probability of certain preference order, for example, $x_1 < x_2 < x_3$. Observe that $x_1 < x_2 < x_3$ if and only if $x_1 x_2^{-1} < 1$ and $x_2 x_3^{-1} < 1$, i.e., $b_{12} < 1$ and $b_{23} < 1$. By Theorem 8 we have that for all $\varepsilon > 0$ one has

$$\mathrm{pr}\left(\mathrm{e}^{-\varepsilon\omega_{12}} < \frac{b_{12}}{\mathsf{G}(b_{12})} < \mathrm{e}^{\varepsilon\omega_{12}} \text{ and } \mathrm{e}^{-\varepsilon\omega_{23}} < \frac{b_{23}}{\mathsf{G}(b_{23})} < \mathrm{e}^{\varepsilon\omega_{23}} \right) \geq 1 - \frac{1 + \sqrt{1-\rho^2}}{\varepsilon^2}, \tag{10}$$

where $\omega_{12} = \mathsf{Var}_g(b_{12})$, $\omega_{23} = \mathsf{Var}_g(b_{23})$, and ρ is the correlation between $\log(b_{12})$ and $\log(b_{23})$. Since

$$\rho = \frac{\mathsf{Cov}_g(b_{12}, b_{23})}{\sqrt{\mathsf{Var}_g(b_{12})}\sqrt{\mathsf{Var}_g(b_{23})}} = \frac{-2\omega/9}{\sqrt{4\omega/9}\sqrt{\omega/9}} = -1,$$

then we obtain from (10) the following table for several values of ε.

Value of ε	Interval of b_{12}	Interval of b_{23}	Lower bound of the probability
$\varepsilon = 1.5$	$[2.699, 2.729]$	$[1.485, 1.489]$	0.56
$\varepsilon = 2$	$[2.694, 2.734]$	$[1.484, 1.490]$	0.75
$\varepsilon = 3$	$[2.685, 2.745]$	$[1.483, 1.491]$	0.89
$\varepsilon = 5$	$[2.665, 2.765]$	$[1.480, 1.493]$	0.96
$\varepsilon = 10$	$[2.616, 2.817]$	$[1.473, 1.501]$	0.99

As we can see, we get good bounds for these probabilities.

3.4 The log-normal distribution and AHP

We say that the random variable X follows a *log-normal* distribution with parameters μ and σ (denoted as $X \sim \log\mathcal{N}(\mu, \sigma)$) if X is positive and $\log X$ follows a normal distribution such that $\mathsf{E}(\log X) = \mu$ and $\mathsf{Var}(\log X) = \sigma^2$. Evidently,

$$\mathsf{G}(X) = \exp(\mathsf{E}(\log X)) = \mathrm{e}^{\mu}, \qquad \mathsf{Var}_g(X) = \mathsf{Var}(\log X) = \sigma^2$$

The importance in AHP of this distribution lies in the following fact: if $X \sim \log \mathcal{N}(\mu, \sigma)$, then $1/X$ also follows a log-normal distribution. More concretely, $1/X \sim \log \mathcal{N}(-\mu, \sigma)$.

We will use the following two results, which can be found in any textbook dealing with multivariate normal distributions.

Theorem 9 *The random vector $\mathbf{x} \in \mathbb{R}^k$ is multivariate normal if and only if $\mathbf{a}^T \mathbf{x}$ is univariate normal for all $\mathbf{a} \in \mathbb{R}^k$.*

Theorem 10 *If the random variables $X_1, \ldots, X_m$ are independent and if X_i has a normal distribution ($i = 1, \ldots, m$), then $a_1 X_1 + \cdots a_m X_m$ has a normal distribution for arbitrary constants $a_1, \ldots, a_m \in \mathbb{R}$.*

When the judgements are independent and follow a log-normal distribution, we can give the following theorem.

Theorem 11 *Let $A = (a_{ij}) \in \mathcal{M}_n^+$ be a reciprocal random matrix. Assume that a_{ij} are independent for $1 \le i < j \le n$ and $a_{ij} \sim \log \mathcal{N}(\mu_{ij}, \sigma_{ij})$. Let $B = (b_{ij})$ be the closest consistent matrix to A in the sense of the distance defined in (1) and $\mathbf{x} = [x_1, \ldots, x_n]^T$ be a priority vector of B. Then the random vectors $\mathbf{y} = [\log x_1, \ldots, \log x_n]^T$ and*

$$\mathbf{b} = [\log b_{12}, \ldots, \log b_{1n}, \log b_{23}, \ldots, \log b_{2n}, \ldots, \log b_{n-1,n}]^T$$

follow a multivariate normal distribution.

PROOF: We use Theorem 9 to prove that $\mathbf{y}$ has a multivariate normal distribution. Let $\mathbf{a} = [\xi_1, \ldots, \xi_n]^T \in \mathbb{R}^n$. From $x_i = C \sqrt[n]{a_{i1} \cdots a_{in}}$ for some fixed constant $C > 0$, if we denote $l_{ij} = \log(a_{ij})$ for all indices i, j, then

$$\mathbf{a}^T \mathbf{y} = \sum_{i=1}^{n} \xi_i \log x_i = \frac{C}{n} \sum_{i=1}^{n} \xi_i (l_{i1} + \cdots + l_{in}). \tag{11}$$

Since a_{ij} are independent for $1 \le i < j \le n$, from Theorems 9 and 10, the vector

$$\mathbf{l} = [l_{12}, \ldots, l_{1n}, l_{23}, \ldots, l_{2n}, \ldots, l_{n-1,n}]^T \in \mathbb{R}^p$$

(here $p = n(n-1)/2$) has a multivariate normal distribution. In addition, using $l_{ij} = -l_{ji}$, $l_{ii} = 0$, and (11), we can see that there exists $\mathbf{c} \in \mathbb{R}^p$ such that $\mathbf{a}^T \mathbf{y} = \mathbf{c}^T \mathbf{l}$. By Theorem 9, $\mathbf{a}^T \mathbf{y}$ has a univariate normal distribution. Since $\mathbf{a}$ is arbitrary, again by Theorem 9, the random vector $\mathbf{y}$ has a multivariate normal distribution.

Let $\mathbf{d} = [d_{12}, \ldots, d_{n-1,n}]^T \in \mathbb{R}^p$. By using $b_{ij} = x_i / x_j$ we have

$$\mathbf{d}^T \mathbf{b} = \sum_{i<j} d_{ij} \log b_{ij} = \sum_{i<j} d_{ij} (\log x_i - \log x_j) = \frac{C}{n} \sum_{i<j} \sum_{k=1}^{n} d_{ij} (l_{ik} - l_{jk})$$

Using again $l_{rs} = -l_{sr}$ and $l_{rr} = 0$, there exists a vector $\mathbf{e} \in \mathbb{R}^p$ such $\mathbf{d}^T \mathbf{b} = \mathbf{e}^T \mathbf{l}$. A similar argument as before can be used to prove that $\mathbf{b}$ follows a multivariate normal distribution. $\square$

We do not specifiy the parameters of the multivariate distributions of the foregoing theorem as they can be easily found in Theorem 3, Theorem 5, and Corollary 1.

Example. Let us consider the following reciprocal random matrix

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} \\ 1/a_{12} & 1 & 2 \\ 1/a_{13} & 1/2 & 1 \end{bmatrix}.$$

The expert considers $3 \le a_{12} \le 4$ and $4 \le a_{13} \le 5$. Therefore, it is natural to set $\mathsf{G}(a_{12}) = \sqrt{12}$ and $\mathsf{G}(a_{13}) = \sqrt{20}$. The expert assumes that a_{12} and a_{13} follow a log-normal distribution and these variables are independent. To set the geometric variance of a_{12}, several random samples from the log-normal distribution with $\mathsf{G}(a_{12}) = \sqrt{12}$ and $\mathsf{Var}_g(a_{12}) = 0.5^2$ were generated. In Octave, ten samples can be easily obtained by executing `exp(normrnd(log(sqrt(12)),0.5,10,1))`. By performing this, we can observe that there are samples outside $[3,4]$, which is not admissible by the expert, and therefore, we must decrease the variance. After several tries, the expert says that the value of $\mathsf{Var}_g(a_{12}) = 0.05^2$ is adequate. In a similar way, $\mathsf{G}(a_{13}) = \sqrt{20}$ and $\mathsf{Var}_g(a_{13}) = 0.05^2$ will be considered.

We denote $\gamma_{12} = \mathsf{G}(a_{12})$, $\gamma_{13} = \mathsf{G}(a_{13})$, and $\omega = \mathsf{Var}_g(a_{12}) = \mathsf{Var}_g(a_{13})$. Let $B = (b_{ij})$ the consistent matrix closest to A and let $[x_1\ x_2\ x_3]^T$ be a priority vector of B. By Theorem 3, there exists $C > 0$ such that

$$\mathsf{G}(x_1) = C\sqrt[3]{\gamma_{12}\gamma_{13}}, \quad \mathsf{G}(x_2) = C\sqrt[3]{2/\gamma_{12}}, \quad \mathsf{G}(x_3) = C\sqrt[3]{1/(2\gamma_{13})}, \quad \mathsf{G}(b_{ij}) = \mathsf{G}(x_i)/\,\mathsf{G}(x_j).$$

As an example we shall find $\mathrm{pr}(x_1 < 2x_2)$ and $\mathrm{pr}(x_1 < 3x_2 \ \&\ x_1 < 5x_3)$. Observe first that

$$\mathrm{pr}(x_1 < x_2) = \mathrm{pr}(x_1/x_2 < 2) = \mathrm{pr}(b_{12} < 2) = \mathrm{pr}(\log b_{12} < \log 2).$$

By Theorem 11, $\log b_{12}$ follows a normal distribution. To find its parameters, we apply Theorem 3:

$$\mathsf{E}(\log b_{12}) = \log(\mathsf{G}(b_{12})) = \log(\mathsf{G}(x_1)/\,\mathsf{G}(x_2)) = \frac{1}{3}\left[2\log\gamma_{12} + \log\gamma_{13} - \log 2\right] \simeq 1.097. \tag{12}$$

By Theorem 5, one gets $\mathsf{Var}(\log b_{12}) = \mathsf{Var}_g(b_{12}) = \mathsf{Var}_g(x_1) + \mathsf{Var}_g(x_2) - 2\,\mathsf{Cov}_g(x_1, x_2)$. But Corollary 1 leads to

$$\mathsf{Var}_g(x_1) = \frac{1}{9}\left(\mathsf{Var}_g(a_{11}) + \mathsf{Var}_g(a_{12}) + \mathsf{Var}_g(a_{13})\right) = \frac{2\omega}{9},$$

$$\mathsf{Var}_g(x_2) = \frac{1}{9}\left(\mathsf{Var}_g(a_{21}) + \mathsf{Var}_g(a_{22}) + \mathsf{Var}_g(a_{23})\right) = \frac{\omega}{9},$$

and

$$\mathsf{Cov}_g(x_1, x_2) = -\frac{1}{9}\,\mathsf{Var}_g(a_{12}) = -\frac{\omega}{9}.$$

Therefore, $\mathsf{Var}(\log b_{12}) = 5\omega/9 \simeq 0.00139$. Now, it is simple to compute $\mathrm{pr}(\log b_{12} < \log 2)$, obtaining that this probability is approximately 0.

To find $\mathrm{pr}(x_1 < 3x_2 \ \&\ x_1 < x_3) = \mathrm{pr}(b_{12} < 3 \ \&\ b_{13} < 1)$, we need to know the parameters of the joint distribution of (b_{12}, b_{13}). By Theorems 9 and 11, $(\log b_{12}, \log b_{13})$ follows a bivariate normal distribution. The mean of $\log b_{12}$ was computed in (12). Similarly, we have

$$\mathsf{E}(\log b_{13}) = \log(\mathsf{G}(b_{13})) = \log(\mathsf{G}(x_1)/\,\mathsf{G}(x_3)) = \frac{1}{3}\left[\log 2 + \log\gamma_{12} + 2\log\gamma_{13}\right] \simeq 1.644.$$

The covariance matrix of $(\log b_{12}, \log b_{13})$ is

$$\mathbf{\Sigma} = \left[\begin{array}{cc} \mathsf{Var}(\log b_{12}) & \mathsf{Cov}(\log b_{12}, \log b_{13}) \\ \mathsf{Cov}(\log b_{12}, \log b_{13}) & \mathsf{Var}(\log b_{13}) \end{array} \right],$$

which can be computed by using Theorem 5 and Corollary 1. Observe that $\mathsf{Var}(\log b_{12})$ was computed before. Since

$$\begin{aligned}\mathsf{Var}(\log b_{13}) &= \mathsf{Var}_g(b_{13}) = \mathsf{Var}_g(x_1) + \mathsf{Var}_g(x_3) - 2\,\mathsf{Cov}_g(x_1, x_3) \\ &= \frac{1}{9}\big[\big(\mathsf{Var}_g(a_{11}) + \mathsf{Var}_g(a_{12}) + \mathsf{Var}_g(a_{13})\big) + \\ &\quad + \big(\mathsf{Var}_g(a_{31}) + \mathsf{Var}_g(a_{32}) + \mathsf{Var}_g(a_{33})\big) + 2\,\mathsf{Var}_g(a_{13})\big] = \frac{5\omega}{9}\end{aligned}$$

and

$$\begin{aligned}\mathsf{Cov}(\log b_{12}, \log b_{13}) &= \mathsf{Cov}_g(b_{12}, b_{13}) \\ &= \mathsf{Cov}_g(x_1, x_1) - \mathsf{Cov}_g(x_1, x_3) - \mathsf{Cov}_g(x_2, x_1) + \mathsf{Cov}_g(x_2, x_3) \\ &= \mathsf{Var}_g(x_1) + \frac{1}{9}\mathsf{Var}_g(a_{13}) + \frac{1}{9}\mathsf{Var}_g(a_{21}) - \frac{1}{9}\mathsf{Var}_g(a_{23}) = \frac{4\omega}{9},\end{aligned}$$

we get

$$\mathbf{\Sigma} = \frac{\omega}{9}\left[\begin{array}{cc} 5 & 4 \\ 4 & 5 \end{array} \right].$$

Observe that in this example, matrix $\mathbf{\Sigma}$ is not singular.

If $\mathbf{\Sigma}$ were singular, then there would exist constants $\alpha, \beta \in \mathbb{R}$ such that $\log b_{12} = \alpha \log b_{13} + \beta$. This constants can be easily obtained from Theorem 1 and $b_{ij} = x_i/x_j$. In this case (which recall it is not satisfied by the example), one can find $\mathrm{pr}(b_{12} < 3 \;\&\; b_{13} < 5)$.

Finally, to find $\mathrm{pr}(b_{12} < 3 \;\&\; b_{13} < 5) = \mathrm{pr}(\log b_{12} < \log 3 \;\&\; \log b_{13} < 0)$, we will use the Octave program. By executing

```
g12=sqrt(12); g13=sqrt(20);
e1=(2*log(g12)+log(g13)-log(2))/3; % Mean of log(b12)
e2=(2*log(g13)+log(g12)+log(2))/3; % Mean of log(b13)
mu = [e1 e2];
om=0.05^2; % Omega
Sigma= [5 4; 4 5]*om/9; % Covariance matrix of (log b12,log b13)
mvncdf([log(3) log(5)],mu,Sigma) % pr(log b12 < log 3 & log b13 < log 5)
```

we obtain $\mathrm{pr}(b_{12} < 3 \;\&\; b_{13} < 5) \simeq 0.172$.

4 Case study

This case study refers to a manufacturing firm that must decide about implementing one or more of five maintenance actions (MA_1, MA_2, MA_3, MA_4, MA_5) aimed at keeping an industrial water distribution system (IWDS), which feeds the company factories, under suitable operational conditions. Consequently, the aim is to minimize the plant shutdown risk. These actions must be prioritized for the purpose of finding a suitable trade-off between improving the plant condition, while not shouldering the simultaneous implementation of numerous interventions. The AHP

Policy	ID Alternative	Maintenance action description
Preventive	MA_1	Electric pump redundancy
	MA_2	Preliminary supply of "special parts" (such as valves, fittings, and pipes), to make eventual substitution interventions faster
Corrective	MA_3	Intensifying plant flexibility by increasing the number of disconnection points in the water network for closing those parts to be maintained, and avoiding plant shutdown
	MA_4	Creation of water storage, in case of sudden interruption of the water service
Predictive	MA_5	Implementation of a tele-surveillance system for the water feeding, to monitor parameters such as temperature, flow rate, and pressure

Table 1: Description of the maintenance actions to be ranked

technique is applied to obtain the final ranking of actions. These maintenance actions belong to the following categories of maintenance policies: preventive, corrective, and predictive. The description of the actions focused on the IWDS in relation to their policy categories is provided in Table 1.

Those maintenance actions are evaluated by means of four criteria (C_1, C_2, C_3, C_4). The evaluation criteria considered are, respectively: security; cost; productivity; and hygiene.

The first criterion refers to the plant's compliance with the regulations in force. The second criterion regards the cost for implementing an action and facing a possible plant shutdown. The third criterion is related to the fulfilment of production standards and then to the need to keep the system available. Lastly, the fourth criterion evaluates the hygienic conditions for drinking water supply to the personnel and plant sanitation. The hierarchical structure of the problem is represented in Figure 1.

The vector of criteria weights is obtained by involving a decision group, whose components (D_1, D_2, D_3) are assumed to have different weights in the decision process. Table 2 shows the roles of each decision maker and their weights, whereas Table 3 reports their pairwise comparison judgments of the criteria, collected in three random reciprocal matrices.

Decision maker	Role	Weight
D_1	Technician	40%
D_2	Quality manager	35%
D_3	Productivity manager	25%

Table 2: Roles and weights of the decision makers

In formulating their judgements, the experts had doubts in assigning some evaluations. Particularly, experts D_1 and D_3 doubted in expressing a clear opinion about the pairwise comparisons

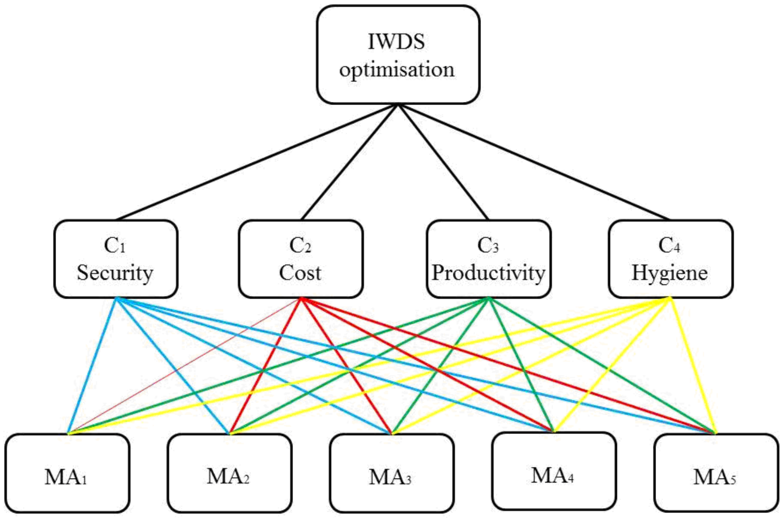

Figure 1: Hierarchical structure

D_1	C_1	C_2	C_3	C_4
C_1	1	5	4	$\boldsymbol{X_1}$
C_2	1/5	1	3	1/5
C_3	1/4	1/3	1	1/5
C_4	$\boldsymbol{X_1^{-1}}$	5	5	1

D_2	C_1	C_2	C_3	C_4
C_1	1	3	3	1
C_2	1/3	1	$\boldsymbol{X_2}$	1/5
C_3	1/3	$\boldsymbol{X_2^{-1}}$	1	1/4
C_4	1	5	4	1

D_3	C_1	C_2	C_3	C_4
C_1	1	1/3	1/6	$\boldsymbol{X_3}$
C_2	3	1	1/3	2
C_3	6	3	1	3
C_4	$\boldsymbol{X_3^{-1}}$	1/2	1/3	1

Table 3: Decision-makers' random reciprocal matrices of criteria evaluations. The first matrix is A_1, the second is A_2, and the third is A_3.

C_1 / C_4, that is to say, between security and hygiene. Specifically, D_1 doubted between the values of 1 and 2, whereas D_3 doubted between 0.20 and 0.25. Moreover, expert D_2 doubted between the values 2 and 3 to be assigned to the pairwise comparison C_2 / C_3, related to the aspects of cost and productivity. For these reasons, we consider three random reciprocal matrices, two of them with random entry a_{14} and the other with random entry a_{23}, in addition to their relative reciprocal entries a_{14}^{-1} and a_{23}^{-1}. These entries are positive random variables. Let A_i be the reciprocal matrix provided by the ith expert and let X_1, X_2, and X_3 be the random variables a_{14}, a_{23}, and a_{14} for the experts D_1, D_2, and D_3, respectively. We assume that these random variables are continuous and uniformly distributed on the aforementioned intervals, specifically, $X_1 \sim \mathscr{U}(1,2)$, $X_2 \sim \mathscr{U}(2,3)$, and $X_3 \sim \mathscr{U}(0.2, 0.25)$. It is simple to check (see the appendix) that

$$\begin{array}{ccc} \text{Random variable} & \text{Geom. Expectation} & \text{Geom. Variance} \\ \hline X_1 & \mathsf{G}(X_1) \simeq 1.472 & \mathsf{Var}_g(X_1) = 0.0391 \\ X_2 & \mathsf{G}(X_2) \simeq 2.483 & \mathsf{Var}_g(X_2) = 0.0136 \\ X_3 & \mathsf{G}(X_3) \simeq 0.225 & \mathsf{Var}_g(X_3) = 0.00414 \end{array} \tag{13}$$

We apply Theorem 3 to calculate the geometric expectations, and Theorems 4 and 5 to obtain the geometric variances and covariances. Let B_i be the closest consistent matrix to A_i and let $\mathbf{x}_i$

be a priority vector of B_i. We have that there exists $C_1 > 0$ such that

$$\mathsf{G}(\mathbf{x}_1) = C_1 \begin{bmatrix} \sqrt[4]{1 \cdot 5 \cdot 4 \cdot \mathsf{G}(X_1)} \\ \sqrt[4]{1/5 \cdot 1 \cdot 3 \cdot 1/5} \\ \sqrt[4]{1/4 \cdot 1/3 \cdot 1 \cdot 1/5} \\ \sqrt[4]{\mathsf{G}(X_1^{-1}) \cdot 5 \cdot 5 \cdot 1} \end{bmatrix} \simeq C_1 \begin{bmatrix} 2.329 \\ 0.5886 \\ 0.3593 \\ 2.030 \end{bmatrix}$$

and analogously,

$$\mathsf{G}(\mathbf{x}_2) \simeq C_2 \, [1.732 \;\; 0.6379 \;\; 0.4280 \;\; 2.115]^T, \quad \mathsf{G}(\mathbf{x}_3) \simeq C_3 \, [0.3342 \;\; 1.189 \;\; 2.711 \;\; 0.9282]^T$$

for some $C_2, C_3 > 0$.

Furthermore, for each decision maker, as shown in subsection 3.1, we can obtain matrices G_i (given by (14), (15), and (16)) representing, respectively, the geometric means for the entries of the consistent matrices that are closer to the given reciprocal random matrices A_i. In other words, the entry (r,s) of G_i is the geometric expectation of the entry (r,s) of B_i. For the next equations, it is worth remembering the mapping $J : \mathbb{R}^n \to \mathbb{R}^n$ defined in Section 2.

$$\mathrm{D}_1 \;\rightarrow\; G_1 = \mathsf{G}(\mathbf{x}_1) J(\mathsf{G}(\mathbf{x}_1))^T = \begin{bmatrix} 1 & 3.9573 & 6.4824 & 1.1472 \\ 0.2527 & 1 & 1.6381 & 0.2899 \\ 0.1543 & 0.6105 & 1 & 0.1770 \\ 0.8717 & 3.4494 & 5.6504 & 1 \end{bmatrix}, \tag{14}$$

$$\mathrm{D}_2 \;\rightarrow\; G_2 = \mathsf{G}(\mathbf{x}_2) J(\mathsf{G}(\mathbf{x}_2))^T = \begin{bmatrix} 1 & 2.7154 & 4.0468 & 0.8190 \\ 0.3683 & 1 & 1.4903 & 0.3016 \\ 0.2471 & 0.6710 & 1 & 0.2024 \\ 1.2209 & 3.3153 & 4.9509 & 1 \end{bmatrix}, \tag{15}$$

and

$$\mathrm{D}_3 \;\rightarrow\; G_3 = \mathsf{G}(\mathbf{x}_3) J(\mathsf{G}(\mathbf{x}_3))^T = \begin{bmatrix} 1 & 0.2810 & 0.1233 & 0.3601 \\ 3.5584 & 1 & 0.4387 & 1.2812 \\ 8.1114 & 2.2796 & 1 & 2.9205 \\ 2.7774 & 0.7805 & 0.3424 & 1 \end{bmatrix}. \tag{16}$$

The resemblance of these figures with the respective original judgments is very noticeable.

We will compute the matrices of variances, one for each decision maker. We denote by $\omega_i = \mathsf{Var}_g(X_i)$, values that are computed in (13). For decision maker D_1,

$$\mathsf{Var}_g(\mathbf{x}_1) = \begin{bmatrix} \mathsf{Var}_g(\sqrt[4]{20X_1}) \\ \mathsf{Var}_g(\sqrt[4]{3/25}) \\ \mathsf{Var}_g(\sqrt[4]{1/60}) \\ \mathsf{Var}_g(\sqrt[4]{25/X_1}) \end{bmatrix} = \begin{bmatrix} \omega_1/16 \\ 0 \\ 0 \\ \omega_1/16 \end{bmatrix},$$

and analogously, for D_2 and D_3,

$$\mathsf{Var}_g(\mathbf{x}_2) = \left[0 \;\; \frac{\omega_2}{16} \;\; \frac{\omega_2}{16} \;\; 0\right]^T, \qquad \mathsf{Var}_g(\mathbf{x}_3) = \left[\frac{\omega_3}{16} \;\; 0 \;\; 0 \;\; \frac{\omega_3}{16}\right]^T.$$

Let $\Sigma_g(\mathbf{x}_i)$ be the geometric variance-covariance matrix of the random vector $\mathbf{x}_i$. By doing similar computations as in the example of the subsection 3.3,

$$\Sigma_g(\mathbf{x}_1) = \frac{\omega_1}{16}\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}, \qquad \Sigma_g(\mathbf{x}_2) = \frac{\omega_2}{16}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

and

$$\Sigma_g(\mathbf{x}_3) = \frac{\omega_3}{16}\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}.$$

Finally, if we denote by V_i the matrix whose (r,s) entry is the geometric variance of the (r,s) entry of B_i, then again by performing similar computations as in the example of subsection 3.3,

$$V_1 = \frac{\omega_1}{16}\begin{bmatrix} 0 & 1 & 1 & 4 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 4 & 1 & 1 & 0 \end{bmatrix}, \quad V_2 = \frac{\omega_2}{16}\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 4 & 1 \\ 1 & 4 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \quad V_3 = \frac{\omega_3}{16}\begin{bmatrix} 0 & 1 & 1 & 4 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 4 & 1 & 1 & 0 \end{bmatrix}.$$

By considering $\omega_1 = 0.0391$, $\omega_2 = 0.0136$, and $\omega_3 = 0.00414$, (these values were shown in (13)) and using some specific values of u (as in the numerical example of subsection 3.3), we can use Theorem 6 to calculate the lower bounds of the probability for each considered variable. This is shown in Table 4. The probabilities that the considered variables do not belong to the indicated intervals are almost negligible. For example, the probability that X_2 (corresponding to b_{23} for expert D_2) does not belong to the interval $[1.1041,\ 2.0117]$ is lower than $1 - 0.962 = 0.0377$. This confirms the goodness of the evaluations.

Note that, although the study has been performed only for those variables originally introducing randomness in the original matrices A_i, similar calculations should be performed for all the random entries of matrices B_i that can be identified by the non-vanishing positions of the corresponding matrices V_i.

After having shared results with the decision-makers, who agreed with the final composition of the three matrices shown in (14), (15), and (16), their entries are aggregated in a single matrix (17) using the geometric mean. The corresponding priority vector is given in the last column of (17).

	C_1	C_2	C_3	C_4	Priorities
C_1	1	1.791	2.041	0.763	29.77%
C_2	0.558	1	1.140	0.426	16.63%
C_3	0.490	0.877	1	0.374	14.59%
C_4	1.310	2.346	2.675	1	39.01%

(17)

Table 5 gives the evaluations of the problem alternatives related to the considered criteria. The last two columns, respectively, give the local priorities, given by their corresponding Perron vectors, and the values of the consistency ratios CR. In particular, the consistency of the judgment is verified, because the CR values do not surpass the threshold of 0.1 (see, for example, [25, 26]).

On the basis of criteria priorities, the global score for each alternative has been obtained by applying the weighted sum of the respective local priorities, and the final ranking is shown in (18).

Reference random matrix	Random variable of the closest consistent matrix	Value of u	Interval of variable	Lower bound of the probability
		0.7	[0.5697, 2.3103]	0.980
A_1	b_{14}	0.3	[0.8499, 1.5449]	0.891
		0.15	[0.9874, 1.3329]	0.565
		0.7	[0.7400, 3.0011]	0.993
A_2	b_{23}	0.3	[1.1041, 2.0117]	0.962
		0.15	[1.2827, 1.7315]	0.849
		0.7	[0.1789, 0.7251]	0.998
A_3	b_{14}	0.3	[0.2667, 0.4860]	0.988
		0.15	[0.3099, 0.4183]	0.954

Table 4: Lower bounds of the probability

Position	Alternative	Score
1^{st}	MA_5	0.4424
2^{nd}	MA_1	0.2248
3^{rd}	MA_3	0.1254
4^{th}	MA_4	0.1130
5^{th}	MA_2	0.0944

(18)

The ranking gives the prioritization values for the five maintenance actions starting from the MA_5 alternative, which corresponds to the predictive maintenance policy. Moreover, it is interesting to note that the corrective policies (MA_3, MA_4 and MA_2) have no relevant priorities in minimizing the plant shutdown risk, and the relative interventions may be postponed.

5 Conclusions

Decision-making in problem-solving contexts seeks to select and implement the most appropriate solution with the aim of optimizing the generally broad range of available possibilities. The most common difficulty is represented by a condition of uncertainty, in which decision-makers may be immersed in the task of attributing their evaluations and making suitable selections by facing various factors or criteria. We claim that a probabilistic approach can be considered a good support for this type of situation. For this reason, the use of the probability theory is herein proposed in integration with the Analytic Hierarchy Process, which is one of the most widespread methods used to carry out decision-making processes.

The case in which an expert or a group of decision-makers have doubts in assigning crisp judgments, but can provide probabilistic values, is considered. In this context, the pairwise comparison matrices of AHP are treated as random reciprocal matrices with one or more random

C_1	MA_1	MA_2	MA_3	MA_4	MA_5	Local priorities	CR
MA_1	1	5	4	2	1/3	0.2383	
MA_2	1/5	1	1	1/3	1/6	0.0579	
MA_3	1/4	1	1	1/3	1/3	0.0755	0.0748
MA_4	1/2	3	3	1	1/6	0.1387	
MA_5	3	6	3	6	1	0.4896	

C_2	MA_1	MA_2	MA_3	MA_4	MA_5	Local priorities	CR
MA_1	1	1/3	1/2	1/4	7	0.2283	
MA_2	3	1	2	1	9	0.2897	
MA_3	2	1/2	1	2	7	0.1747	0.2855
MA_4	4	1	1/2	1	9	0.2843	
MA_5	1/7	1/9	1/7	1/9	1	0.0230	

C_3	MA_1	MA_2	MA_3	MA_4	MA_5	Local priorities	CR
MA_1	1	6	5	4	1/4	0.2672	
MA_2	1/6	1	1/2	1/2	1/7	0.0461	
MA_3	1/5	2	1	3	1/5	0.1011	0.0838
MA_4	1/4	2	1/3	1	1/6	0.0640	
MA_5	4	7	5	6	1	0.5217	

C_4	MA_1	MA_2	MA_3	MA_4	MA_5	Local priorities	CR
MA_1	1	7	3	7	1/5	0.2449	
MA_2	1/7	1	1/4	1	1/7	0.0430	
MA_3	1/3	4	1	3	1/5	0.1143	0.0809
MA_4	1/7	1	1/3	1	1/7	0.0448	
MA_5	5	7	5	7	1	0.5530	

Table 5: Evaluation of alternatives respect to the criteria, local priorities and CR value

entries, which are random positive variables that capture expert uncertainty. We have developed the necessary theory to handle AHP-based decisions under the umbrella of the probability theory. In addition, lower bounds of probability in terms of confidence intervals for the various variables involved are estimated. The obtained results confirm the goodness of proposal.

A case study focused on a decision-making process to be undertaken in a manufacturing firm is approached and resolved as a real-world case-study. Specifically, the AHP technique is applied to prioritize five maintenance actions tailored to the industrial water distribution system feeding the industrial plants of the firm. The aim is to pursue technological innovation and structure a long-term strategy of maintenance for the organization.

Acknowledgments

The use of English has been supervised by John Rawlins, a qualified member of the UK Institute of Translation and Interpreting.

References

[1] E. Alirio Valencia, Y.A. Poveda, C.A. Escudero Salcedo, *Desigualdad de Chebyshev bidimensional.* Scientia Et Technica, Universidad Tecnológica de Pereira, Pereira, Colombia, 51 (2012), 242-246.

[2] M. Bean. Probability: The Science of Uncertainty with Applications to Investments, Insurance, and Engineering

[3] J. Benítez, X. Delgado-Galván, J. Izquierdo, R. Pérez-Garcia, *Achieving matrix consistency in AHP through linearization.* Applied Mathematical Modelling, 35, (2011) 4449-4457.

[4] J. Benítez, J. Izquierdo, R. Pérez-García, E. Ramos-Martínez *A simple formula to find the closest consistent matrix to a reciprocal matrix.* Applied Mathematical Modelling, 38, (2014) 3968-3974.

[5] R. Bhatia, C. Davis. *A Better Bound on the Variance.* The American Mathematical Monthly, 107 (4), (2000) 353-357.

[6] A. Biedermann, S. Bozza, F. Taroni, C. Aitken. *The consequences of understanding expert probability reporting as a decision.* Science & Justice, 57 (1), (2017) 80-85.

[7] F. Costello, P. Watts, *People's conditional probability judgments follow probability theory (plus noise).* Cognitive Psychology, 89, (2016) 106-33.

[8] P.H. Giang, *Decision making under uncertainty comprising complete ignorance and probability.* International Journal of Approximate Reasoning, 62, (2015) 27-45.

[9] C. Huang, P. Chu, Y. Chiang, *A fuzzy AHP application in government-sponsored R&D project selection.* Omega, 36 (6), (2008) 1038-1052.

[10] W.R. Hughes, *A statistical framework for strategic decision making with AHP: Probability assessment and Bayesian revision.* Omega, 37 (2), (2009) 463-470.

[11] Y. Izhakian, *Expected utility with uncertain probabilities theory.* Journal of Mathematical Economics, 69, (2017) 91-103.

[12] A.I. Jabbarova, *Solution for the Investment Decision Making Problem through Interval Probabilities.* Procedia Computer Science, 102, (2016) 465-468.

[13] S.G.B. Johnson, G. Rajeev-Kumar, F.C. Keil, *Sense-making under ignorance.* Cognitive Psychology, 89, (2016) 39-70.

[14] M. Karanik, L. Wanderer, J.A. Gomez-Ruiz, J.I. Pelaez, *Reconstruction methods for AHP pairwise matrices: How reliable are they?* Applied Mathematics and Computation, 279, (2016) 103-124.

[15] P. Liu, F. Jin, X. Zhang, Y. Su, M. Wang, *Research on the multi-attribute decision-making under risk with interval probability based on prospect theory and the uncertain linguistic variables.* Knowledge-Based Systems, 24 (4), (2011) 554-561.

[16] Z. Magyar, L. Garbai, A. Jasper, *Risk-based determination of heat demand for central and district heating by a probability theory approach.* Energy and Buildings, 110, (2016) 387-395.

[17] M.A. Matos, *Decision under risk as a multicriteria problem.* European Journal of Operational Research, 181 (3), (2007) 1516-1529.

[18] M.G. Melon, P.A. Beltran, M.C.G. Cruz, *An AHP-based evaluation procedure for innovative educational projects: a face-to-face vs. computer-mediated case study.* Omega, 36 (5), (2008) 754-765.

[19] L. Narens, *An introduction to lattice based probability theories.* Journal of Mathematical Psychology, 74, (2016) 66-81.

[20] Marc Paolella. Fundamental Probability: A Computational Approach

[21] A. Papoulis, Probability, random variables, stochastic processes

[22] F. Y. Partovi, *An analytical model for locating facilities strategically.* Omega, 34, (2006) 41-55.

[23] J.G. Pereira Jr, P.Y. Ekel, R.M. Palhares, R.O Parreiras, *On multicriteria decision making under conditions of uncertainty.* Information Sciences, 324, (2015) 44-59.

[24] E.S. Rosenbloom, *A probabilistic interpretation of the final rankings in AHP.* European Journal of Operational Research, 96, (1996) 371-378.

[25] T.L. Saaty, The analytic hierarchy process. McGraw-Hill, New York, NY. (1980).

[26] T.L. Saaty, *Relative measurement and its generalization in decision making. Why pairwise comparisons are central in mathematics for the measurement of intangible factors. The analytic hierarchy/network process.* Revista de la Real Academia de Ciencias Serie A: Matemáticas 102 (2), (2008) 251-318.

[27] P. Shah, A.J.L. Harris, G. Bird, C. Catmur, U. Hahn, *A pessimistic view of optimistic belief updating.* Cognitive Psychology, 90, (2016) 71-127.

[28] A. Soroudi, A. Rabiee, A. Keane, *Information gap decision theory approach to deal with wind power uncertainty in unit commitment.* Electric Power Systems Research, 145, (2017) 137-148.

[29] L.G. Vargas, *Reciprocal matrices with random coefficient.* Mathematical Modelling 3, (1982) 69-81.

[30] R.R. Yager, V. Kreinovich, *Decision making under interval probabilities.* International Journal of Approximate Reasoning, 22 (3), (1999) 195-215.

[31] H.-B. Yan, T. Ma, V.-N. Huynh, *On qualitative multi-attribute group decision making and its consensus measure: A probability based perspective.* Omega, 70, (2017) 94-117.

[32] Z.-W. Yu, J.-F. Mao, *Probability analysis of train-track-bridge interactions using a random wheel/rail contact model.* Engineering Structures, 144, (2017) 120-138.

[33] H. Zhang, L. Ha, Q. Li, M. Beer, *Imprecise probability analysis of steel structures subject to atmospheric corrosion.* Structural Safety, 67, (2017) 62-69.

Appendix

For the sake of general readability, in this appendix we provide the proof of Theorem 7 presented in [1], a source which is not in English. The following lemma is needed.

Lemma 1 *If X and Y are random variable such that* $\mathsf{E}(X) = \mathsf{E}(Y) = 0$ *and* $\mathsf{Var}(X) = \mathsf{Var}(Y) = 1$*, then* $\mathsf{E}(\max\{X^2, Y^2\}) \leq 1 + \sqrt{1-\rho^2}$*, where ρ is the correlation between X and Y.*

PROOF: By the hypotheses,

$$\rho = \frac{\mathsf{Cov}(X,Y)}{\sqrt{\mathsf{Var}(X)}\sqrt{\mathsf{Var}(Y)}} = \mathsf{E}(XY) - \mathsf{E}(X)\,\mathsf{E}(Y) = \mathsf{E}(XY).$$

Furthermore, $1 = \mathsf{Var}(X) = \mathsf{E}(X^2) - \mathsf{E}(X)^2 = \mathsf{E}(X^2)$, and $1 = \mathsf{E}(Y^2)$. Since $2\max\{X^2, Y^2\} = |X^2 - Y^2| + X^2 + Y^2$, we have

$$2\,\mathsf{E}(\max\{X^2, Y^2\}) = \mathsf{E}(|X^2 - Y^2|) + \mathsf{E}(X^2) + \mathsf{E}(Y^2) = \mathsf{E}(|X^2 - Y^2|) + 2,$$

Therefore, it is enough to prove $\mathsf{E}(|X^2 - Y^2|) \leq 2\sqrt{1 - \mathsf{E}(XY)^2}$.

Now we use the Cauchy-Schwartz inequality for random variables, i.e., if U and V are random variables, then $[\mathsf{E}\,|UV|]^2 \leq \mathsf{E}(U^2)\,\mathsf{E}(V^2)$.

$$\begin{aligned}
\left[\mathsf{E}\,|X^2 - Y^2|\right]^2 &= [\mathsf{E}(|X+Y| \cdot |X-Y|)]^2 \\
&\leq \mathsf{E}(X^2 + Y^2 + 2XY)\,\mathsf{E}(X^2 + Y^2 - 2XY) \\
&= [\mathsf{E}(X^2) + \mathsf{E}(Y^2) + 2\,\mathsf{E}(XY)][\mathsf{E}(X^2) + \mathsf{E}(Y^2) - 2\,\mathsf{E}(XY)] \\
&= (2 + 2\rho)(2 - 2\rho) = 4(1 - \rho^2).
\end{aligned}$$

Therefore, $\mathsf{E}(|X^2 - Y^2|) \leq 2\sqrt{1-\rho^2}$. $\square$

PROOF OF THEOREM 7: Let $Z = \max\{X, Y\}$. It is clear that $Z^2 = \max\{X^2, Y^2\}$. Now we have, by Markov's inequality and the previous lemma,

$$\begin{aligned}
\mathrm{pr}(|X - \mu_x| \geq \varepsilon\sigma_x \ \text{ or } \ |Y - \mu_y| \geq \varepsilon\sigma_y) &= \mathrm{pr}(|X - \mu_x|/\sigma_x \geq \varepsilon \ \text{ or } \ |Y - \mu_y|/\sigma_y \geq \varepsilon) \\
&= \mathrm{pr}(\max\{|X - \mu_x|^2/\sigma_x^2, |Y - \mu_y|^2/\sigma_y^2\} \geq \varepsilon^2) \\
&\leq \frac{1}{\varepsilon^2}\,\mathsf{E}(\max\{|X - \mu_x|^2/\sigma_x^2, |Y - \mu_y|^2/\sigma_y^2\}) \\
&\leq \frac{1 + \sqrt{1-\rho^2}}{\varepsilon^2}. \quad \square
\end{aligned}$$

Next, we shall compute the geometric expectation and variance of a continuous uniformly distribution X on the interval $[a, b]$. Since $\log(\mathsf{G}(X)) = \mathsf{E}(\log X)$ and $\mathsf{Var}_g(X) = \mathsf{Var}(\log X)$, it is convenient to study the distribution of $Y = \log(X)$. Let F_Y be the cumulative distribution function of Y.

$$F_Y(y) = \mathrm{pr}(Y \leq y) = \mathrm{pr}(\log X \leq y) = \mathrm{pr}(X \leq \mathrm{e}^y) = \begin{cases} 0 & \text{if } y < \log a, \\ (\mathrm{e}^y - a)/(b - a) & \text{if } \log a \leq y \leq \log b, \\ 1 & \text{if } \log b < y. \end{cases}$$

By differentiating we get the density funtion of Y:

$$f_Y(y) = \begin{cases} \mathrm{e}^y/(b-a) & \text{if } y \in [\log a, \log b], \\ 0 & \text{if } y \notin [\log a, \log b]. \end{cases}$$

From $\mathsf{E}(Y) = \int_{\mathbb{R}} y f_Y(y)\,\mathrm{d}y$ we can get $\mathsf{G}(X) = \exp(\mathsf{E}(Y))$. From $\mathsf{E}(Y^2) = \int_{\mathbb{R}} y^2 f_Y(y)\,\mathrm{d}y$ and $\mathsf{Var}(Y) = \mathsf{E}(Y^2) - \mathsf{E}(Y)^2$ we get the variance of $Y = \log X$.