Risk and Environmental Modelling
Delft Institute of Applied Mathematics

Quantification of non-parametric continuous BBNs with expert judgment

Master Thesis

Iwona Danuta Jagielska

Delft University of Technology, The Netherlands
July 2007

Members of the Committee

Chairperson of Graduate Committee:
Prof. dr. Roger M. Cooke

Graduate Committee:

Prof. dr. Roger M. Cooke	Delft University of Technology, the Netherlands Faculty of Electrical Engineering, Mathematics and Computer Science
Prof. dr. Ben J.M. Ale	Delft University of Technology, the Netherlands Faculty of Technology Policy and Management
Prof. Jolanta Misiewicz	University of Zielona Gora, Poland Faculty of Mathematics, Informatics and Econometrics
Dr. Dorota Kurowicka	Delft University of Technology, the Netherlands Faculty of Electrical Engineering, Mathematics and Computer Science
Ir. Oswaldo Morales Napoles	Delft University of Technology, the Netherlands Faculty of Electrical Engineering, Mathematics and Computer Science

Abstract

The aviation is considered as one of the safest ways of traveling in the wide world today, however accidents and incidents still occur. If we would wish to trace ways in which an accident occurred, we would see a chain of simple errors that led to the event. As human plays an important role in the air transport system, their contribution in this chain is quite essential.

In the following thesis the Maintenance Performance Model is developed as a part of Causal Model for Air Transport Safety project. The motivation for this project is the need to a better understanding of air transport accidents, so that efforts to improve safety can be made as effective as possible. The Bayesian Belief Network is used as a modelling tool while constructing CATS model and therefore the Maintenance Performance Model. Mainly the new approach to continuous BBNs [33, 37] is used. To quantify the BBN certain amount of information are needed, i.e. the strength of dependence relations among variables in the model and marginal distribution. In the absence of data expert judgment can be used for this quantification. Dependence elicitation requires to calculate multivariate normal probabilities (by integrating appropriate multivariate normal density).

In order to find a suitable algorithm for calculating the dependence relations, in this thesis we compare two types of integration procedures for the numerical computation of multivariate normal probabilities. The fist one is based on Quasi-Monte Carlo, while the second on the Adaptive Multidimensional Integration Routine. The best algorithm in our view is used to quantify the Maintenance Performance Model.

Acknowledgments

The two years at Delft University of Technology, The Netherlands, have been a valuable and unforgettable experience. I would like to thank Prof. Roger M. Cooke for giving me the opportunity to take part in the MSc program "Risk and Environmental Modelling" and Prof. Jolanta Misiewicz for encouraging me to take a part in this program.

I wish to express a special gratitude to my supervisors Dr. Dorota Kurowicka, Prof. Ben J.M. Ale and Prof. Roger M. Cooke for giving me opportunity of working on this thesis. I would like to thank Ir. Oswaldo Morales Napoles, for supervising my work, his time, help and patient during the whole work.

I would also like to thank my colleagues from TUDelft for the time we have had together during these two years. I particular I am grateful to Lidka and Mariya - without their friendship and encouragement I would not have finished my study and this thesis. Also many thanks to Joanna, Karolina, Ulas, and all my colleagues for being here together and for every kindness.

Last but not the least; I am grateful to my family for their love, support and belief in me.

To my parents

Contents

Chapter 1

Introduction

The aviation is considered as one of the safest ways of traveling in the wide world today, however accidents and incidents still occur. If we would wish to trace ways in which an accident occurred, we would see a chain of simple errors that led to the event. As human plays an important role in the air transport system, their contribution in this chain is quite essential.

Human error in maintenance can impact on safety and performance of the whole aviation industry in a number of ways. Poor repairs, for example, can increase the amount of malfunctions which can further influence the risks associated with equipment failure and personal accidents. A maintenance operator who is motivated, well trained, under no time pressure, given the correct information, and working with equipment which has been designed to be maintenance friendly, will likely complete all specified maintenance tasks. However, when some of these requirements are not fulfilled, it becomes more likely that the maintenance work will not be fully completed. As a result, equipment can become poorly maintained causing reduced reliability/availability or simply direct damage to the airplane. These consequences can increase the safety risk to the maintenance operator, other employees and the public.

In this thesis the Maintenance Performance Model is developed as a part of CATS project; where CATS stands for Causal Model for Air Transport Safety. The CATS project is built by the consortium consisting of different companies, such as: National Aerospace Laboratory (NLR), Delft University of Technology, Det Norske Veritas (DNV) together with White Queen. The motivation for this project is the need to a better understanding of air transport accidents, so that efforts to improve safety can be made as effective as possible. The Bayesian Belief Network is used as a modelling tool while constructing CATS model and therefore the Maintenance Performance Model. Mainly the new approach to continuous BBNs [33, 37] is used. To quantify the BBN certain amount of information are needed, i.e. the strength of dependence relations among variables in the model. In the absence of data the expert judgment procedure can be used for this purpose, however, it requires to calculate multivariate normal probabilities (by integrating appropriate multivariate normal density).

The numerical computation of a multivariate normal probability is often a difficult problem and has received considerable attention in the literature [29, 35]. Many different numerical methods have been proposed (see, e.g., [21, 22, 1, 2, 3]). Moreover, for one- and two-dimensional problems various reliable and efficient software are available for computing (see, e.g., [20, 19] for one dimension problems and [13, 10] for two dimensional problems). The main difference between those numerical methods is based on the integration region that they use and by the different numerical schemes which they use in calculations.

In order to find a suitable algorithm for calculating the dependence relations, in this thesis we compare two types of integration procedures for the numerical computation of multivariate normal probabilities. The fist one is based on Quasi-Monte Carlo, while the second on the Adaptive Multidimensional Integration Routine. The best algorithm in our view will be used to quantify the Maintenance Performance Model.

1.1 Goal of the thesis

The main purpose of this thesis is

1. to quantify the dependencies in the continuous BBNs with use of experts judgment, and

2. to apply obtained information while building the Maintenance Performance Model.

1.2 Outline of the thesis

This thesis describes the results of a graduation project, carried out as part of the Master of Science Program in Applied Mathematics at Delft University of Technology.

The following master thesis consist of two main parts:

- description and discussion about a case study in which BBN is applied for analysis of maintenance human performance;
- discussion about different algorithms to calculate multivariate normal probabilities which can be used to obtain conditional rank correlations from conditional probabilities of exceedence provided by experts.

Chapter 2, the methodology, briefly describes techniques used in modelling task. Mainly, it describes Bayesian Belief Networks and way of quantifying it. The procedure of assessing dependencies from the conditional probabilities of exceedence provided by experts is presented. The application of BBNs to model the maintenance performance is presented and discussed in chapter 3. Chapter 4 can be seen as the separate part of this thesis where discussion and comparison between different algorithms used to calculate multivariate normal probabilities is presented. Chapter 5 deals with the description of software UniExp used in the elicitation of dependencies among variables in the BBN. Finally, conclusions and recommendations are presented. In the Appendices reader can find some additional information such as basic definitions used in this thesis, or questionnaire protocol used in elicitation of expert opinion.

Chapter 2

Methodology

This chapter provides main information about the tools used later on in this thesis. Only the most significant definition (as the one of BBNs) and description (process of assessing conditional rank correlations from conditional probabilities of exceedence) are included here. We assume that the reader is familiar with notions such as rank correlation, normal distribution, copula, etc. However to make this thesis more self contained we provide basic definitions in Appendix A.

2.1 Bayesian Belief Network

Belief networks, also called Bayesian belief networks (BBNs), are graphical tools used to represent a high-dimensional probability distribution. They are convenient tools for making inferences about uncertain states when limited information is available. Belief nets are often used for making diagnoses, with applications to both medical science as well as various engineering disciplines, in particular to emergency planning.

BBN is directed acyclic graph in which nodes represent the variables of interest, and arrows (directed edges) between these nodes indicate dependence relations [6, 12]. The requirement that the graph be acyclic means that there is no directed path through the graph that returns to its own starting point. The arrows drawn between nodes represent qualitative influences which must be quantified. In order to describe the network the following parameters have to be specified: prior probability of all root nodes and conditional probability functions of all other nodes given predecessors.

The BBNs with only discrete random variables have recently become a very popular tool in modelling risk and reliability. They can be seen as the simplest case of probabilistic networks. Building and working with discrete BBNs is very efficient as long as one is not forced to quantify complex BBNs. A high assessment burden of discrete BBNs is often caused by the discretization of continuous variables, i.e. if a child node has 6 parents, and each node is discretized to take 4 possible values, then the conditional probability table for the child node contains 16384 entries (16384 conditional probabilities must be assessed). This assessment burden can only be reduced by a drastic discretization of the nodes, or simplification of the model.

One would wish to be able to handle models in which some or all random variables can take values in a continuous range. Until recently, continuous BBNs were restricted to the joint normal distribution [6, 12] for which the 'influence' of the parents on a child was interpreted as partial regression coefficients when the child was regressed on the parents. To fully describe that kind of BBNs for each normal variable, the unconditional mean and (by assumption constant) conditional variance must be assessed together with a conditional regression coefficient for each arc. The main disadvantage of this approach is that assessing partial regression coefficients may be unintuitive, especially if the variables must first undergo transformation to joint normal. Moreover adding or deleting variables requires re-assessing previously assessed partial regression coefficients.

In this thesis the new approach to continuous BBNs [33, 37] is used. Since no form of joint distribution is assumed this approach is called *non-parametric BBNs*. In this procedure nodes are associated with arbitrary continuous invertible distributions and arcs with (conditional) rank correlations, which are realized by the chosen copula. The number of (conditional) rank correlations is equal to the number of arcs in the BBN. Quite essential is the fact that those conditional rank correlations are algebraically independent and hence they can take any value between (-1,1). This approach is also general and allows traceable and defendable quantification methods, but it comes at a price: these BBNs must be evaluated by Monte Carlo simulation.

We explain techniques used in quantifying those BBNs on the example of Controlled Flight into Terrain CFIT (figure 2.1) Model [5]. Here node *Missed Approach Execution* is a child node of each of the parents *Fuel Weight, Visibility, Crew Alertness, Speed Deviation, Mean Cross Wind* and *Separation in Air*. In chapter 3 we use techniques explained below to quantify Maintenance Performance Model. CFIT is part of the CATS[1] project the same as the Maintenance Performance Model.

[1]Causal Model for Air Transport Safety

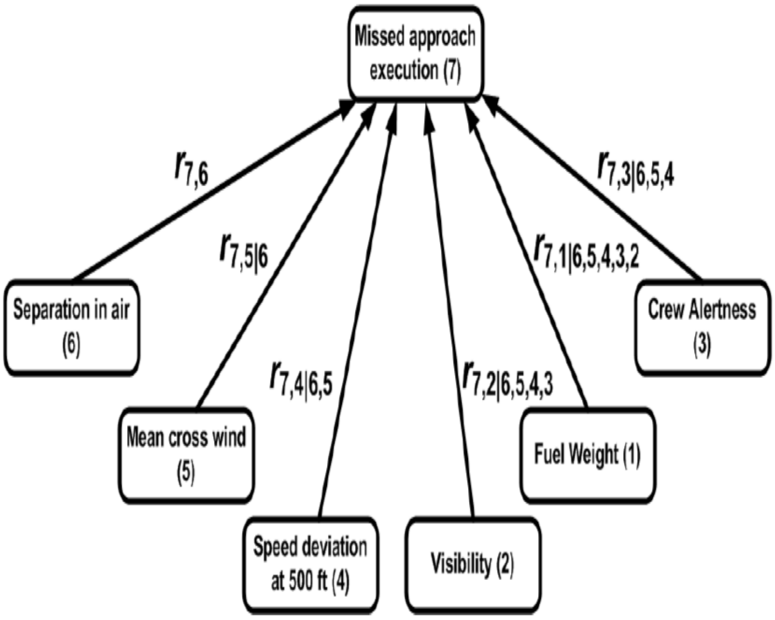

Figure 2.1: Missed Approach Model Structure.

We see in Figure 2.1 that the number of *Missed Approaches* is influenced by variables such as e.g. *Separation in air*, *Mean cross wind* etc. The explanation of the nodes in the CFIT model in Figure 2.1 together with sources of data are provided below:

1. **Fuel Weight**: Measured in kilograms and is the remaining fuel at arrival based on data for 172 flights of a Boeing 737 at Schiphol airport.

2. **Visibility**: Measured in meters and is based on a sample of 27 million observations over Europe.

3. **Crew Alertness**: Measured by the Stanford Sleepiness Scale in an increasing scale from 1 to 7, where 1 signifies "feeling active and vital; wide awake" and 7 stands for "almost in reverie; sleep onset soon; struggle to remain awake" the distribution used for this study comes from field studies by the Aviation Medicine Group of TNO Human Factors in 1,295 flights.

4. **Speed Deviation** at 500 ft: Deviation from bug at 500 ft. The data comes from 13,753 approaches of a major European airline.

5. **Mean Cross Wind**: Usually expressed as a combination of speed (in knots) and direction (compass course) of the wind at any direction not favorable for the aircraft, the cross wind distribution comes from 380,000 takeoffs and landings conducted on three large European airports.

6. **Separation in Air**: Longitudinal distance (in nautical miles) between the landing aircraft and the preceding aircraft in the approach path. The distribution was retrieved from a sample size of 2,382 landings at Schiphol airport.

7. **Missed Approach Execution**: Number of missed Approach Execution per 100,000 flights at Schiphol airport. The Expectation of the number of missed approach executions divided by 100,000 would be an estimate of the probability of executing a missed approach maneuver. No data is available for quantification.

To quantify the BBN we need to specify all one-dimensional marginal distributions associated with the nodes of the BBN. In the absence of data as in case for node *Missed Approach Execution* the expert judgment procedure can be used. In this context, we briefly present in section 2.2.1 the Expert Judgment methodology which can be used to quantify marginal distributions. Next, the information about required rank correlation has to be provided. The following protocol is applied [8] in order to determine which (conditional) correlations are necessary.

1. The the sampling order for the nodes needs to be defined, such that all ancestors of a node i appear before i in the ordering. We choose the sampling order 1, 2, 3, 4, 5, 6, 7 for the BBN structure. Of course this sampling order in not unique and we could have chosen a different order; as long as nodes 1–6 appears before node 7 they can be interchanged freely.

2. Using previously specified sampling order we can write the complete factorization of the joint distribution as:

$$\begin{aligned} P(1,\ldots,7) &= P(1)P(2|\underline{1})P(3|\underline{21})P(4|\underline{321}) \\ &= P(5|\underline{4321})P(6|\underline{54321})P(7|654321) \end{aligned} \tag{2.1}$$

The underscored nodes are the nodes which are not parents of the conditioned variable and thus are not necessary in sampling of the conditioned variable. If we drop the underscored variables, we obtain the standard factorization for the BBN given as follows [12]:

$$P(1,\ldots,7) = \prod_{i=1}^{7} P(X_i|pa(X_i)) \tag{2.2}$$

where $pa(X_i)$ denotes the parents of variable X_i (variable X_i is associated with node i).

3. For each term i with parents (non-underscored variables) $i_1 \ldots i_{pa(i)}$ in (2.1), associate the arc $i_{pa(i)-k} \rightarrow\ i$ with the conditional rank correlation

$$\begin{cases} r(i, i_{pa(i)}), & k = 0 \\ r(i, i_{pa(i)-k}|i_{pa(i)}, \ldots, i_{pa(i)-k+1}), & 1 \leq k \leq pa(i) - 1. \end{cases} \tag{2.3}$$

where the assignment is vacuous if $\{i_1 \cdots_{pa(i)}\} \neq \emptyset$. Assigning conditional rank correlations for $i\ =\ 1, \ldots, 7$, every arc in the BBN is assigned a conditional rank correlation between parent and child.

Hence, the (conditional) rank correlations that need to be assigned to the edges of BBS are: $\{r_{76}, r_{75|6}, r_{74|56}, r_{73|456}, r_{72|3456}, r_{71|23456}\}$ as already marked in figure (2.1). In Section 2.2.2, we present the procedure how the values of the required conditional and unconditional rank correlations are obtained.

2.2 Expert judgment

In many fields of science and technology, some parameters used for modeling are not known and cannot be estimated from experiments or observation. Then we can use experts' knowledge about a particular topic to estimate the required information. Roughly speaking, expert judgment is recognized as another type of scientific data [26].

In this thesis expert judgment procedure is used for two purposes:

1. to provide information about unknown marginal distribution for variable of interests and

2. to obtain information about the strength of dependence relation between variables in our model.

In this project in order to estimate the parameters (required marginal distributions) we have used the so-called Classical Model for experts judgments[2].

2.2.1 Marginal distributions

In order to obtain information about the marginal distribution required in the CFIT model, the Classical Method of expert judgment was used. It has been introduced by [26] and applied in many risk and reliability studies [28, 27]. The term "classical" indicates the analogy between the concepts of this model and classical hypothesis testing. The classical model constructs the weighted combination of expert' probability assessments, where experts were asked to provide their subjective probability distribution in the form of a number of specified quantiles. Weights are derived based on the performance measures and satisfy the strictly proper scoring rule. There exist two quantitative measures of performance of experts - calibration and information (or informativeness), which are assessed based on experts' estimates on so-called *seed* or *calibration* variables. These are variables from experts field whose true values are unknown to experts when giving their opinions, but whose values are known post hoc. Roughly, *calibration* measures the statistical likelihood that a set of experimental results correspond, in a statistical sense, with the experts assessments. *Information* represents a degree to which a distributions provided by experts are concentrated. For more details, see [26].

[2] Only one expert was used in elicitation.

2.2.2 Dependence relations

For full quantification of the CFIT model the information about the dependence relations is still required. Therefore the (conditional) rank correlations associated with edges need to be determined. In general it is possible that those required conditional correlations are not constant so their values may depend on the conditioning variables. This would significantly complicate calculations, however. We will therefore restrict our study only to constant conditional rank correlations. In section 2.1 we shown that rank correlation required for quantification are: $\{r_{76}, r_{75|6}, r_{74|56}, r_{73|456}, r_{72|3456}, r_{71|23456}\}$.

As in the case of marginal distributions we could obtain the information about dependencies from the data. However for that purpose we would need to have access to reliable and applicable historical data which is in suitable format for us. For this reason most studies make use of some form of expert judgment. Conditional rank correlations are not elicited directly. In practice the conditional probabilities of exceedence are more frequently asked. From them we can retrieved the desired rank correlations by assuming a copula. There are great advantages of using joint normal copula due to its useful properties, such as know relationship between rank and product moment correlation and the fact that partial and conditional correlations are equal [8]. In general any copula that represents (conditional) independence as zero (conditional) rank correlation can be used. In this thesis normal copula is used. We show now how to ask for exceedence probabilities and how to retrieve rank correlations from information provided by experts, following explanation is based on [30].

In description below we use following notation:

- X_i is univariate random variables associated with i node in BBN,
- x_i refer to the realization of random variable X_i,
- F_i denotes the cumulative distribution function of the i-th node hence the variable X_i.

To elicit the rank correlation r_{76} between variables X_7 and X_6 , we ask the expert the following question:

Suppose that the variable *Separation in Air* was observed to be above its median value. What is your probability that the *Missed Approach Execution* would also lie about its median value?

which can be written as $P(X_7 \geq x_{7_{50}}|\ X_6 \geq x_{6_{50}})^3$. By answering to this question expert is providing us his estimate for the following probability:

$$P(F_{X_7} \geq 0.5|\ F_{X_6} \geq 0.5). \tag{2.4}$$

Since we can transform $F_{X_i}(X_i)$ to standard normal variables by applying the following transformation $Y_i = \Phi^{-1}(F_{X_i}(X_i))$ i=6, 7, the probability (2.4) can be written now:

$$P(Y_7 \geq 0|\ Y_6 \geq 0). \tag{2.5}$$

Since rank correlation is invariant under strictly increasing transformation, the problem of finding the required rank correlation can be now treated as the problem of finding correlation r_{76} for joint normal distribution for which the conditional probability is (2.5).

The standard bivariate normal distribution of variables Y_7 and Y_6 has a density function which depends on their product moment correlation ρ_{76}:

$$f(y_7, y_6) = f(\rho_{76}) = \frac{1}{2\pi\sqrt{1-\rho_{76}^2}} \exp\left(-\frac{y_2^2 - 2\rho_{76}y_2y_1 + y_1^2}{2(1-\rho_{76}^2)}\right) \tag{2.6}$$

where ρ_{76} is a parameter in [-1, 1], and $(Y_6, Y_7) \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_{76} \\ \rho_{76} & 1 \end{bmatrix}\right)$.

The conditional probability becomes:

$$\begin{aligned} P(Y_7 \geq 0|\ Y_6 \geq 0) &= \frac{P(Y_7 \geq 0,\ Y_6 \geq 0)}{P(Y_6 \geq 0)} = \frac{P(Y_7 \geq 0,\ Y_6 \geq 0)}{1/2} \\ &= 2P(Y_7 \geq 0,\ Y_6 \geq 0) \\ &= 2\int_0^\infty \int_0^\infty f(y_7, y_6)\, dy_7\, dy_6 \end{aligned} \tag{2.7}$$

We can first find ρ_{76} and then using Pearson transformation [24] obtain corresponding value of r_{76}. Where the Person transformation is [24]:

$$\rho_{76} = 2\sin\left(\frac{\pi}{6} r_{76}\right)$$

The main disadvantage of joint normal distribution is that it does not have a close analytical form for joint normal distribution and therefore the integral in (2.7) needs to be evaluated numerically. The relationship between rank correlation r_{76} and the conditional probability from (2.7) is presented in figure (2.2).

---

[3]Here the median value is used, but in practice we could use any other quantile value. However, then the resulting rank correlation is dependent on the choice of copula.

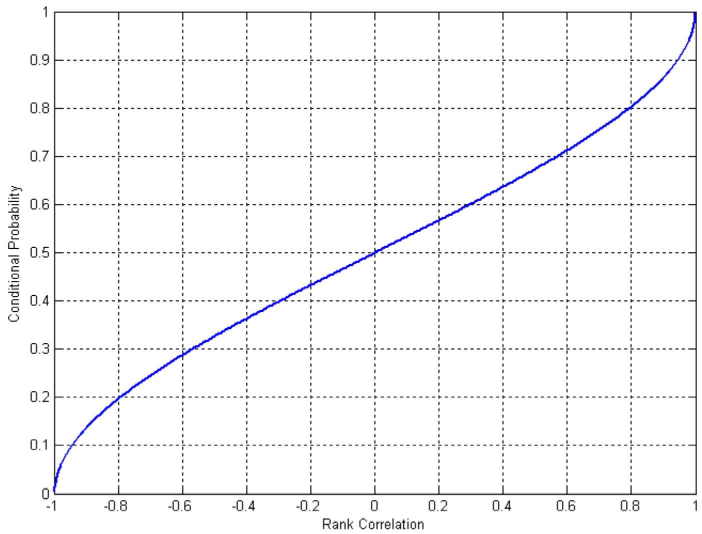

Figure 2.2: Conditional Probability $P(Y_7 \geq 0|\ Y_6 \geq 0)$ versus r_{76}.

When we know the value of $P(Y_7 \geq 0|Y_6 \geq 0)$ the corresponding value of rank correlation r_{76} can be read from this figure. We should notice here that:

- if $P(Y_7 \geq 0|Y_6 \geq 0) = 0$ then $\rho_{76} =$ -1;
- if $P(Y_7 \geq 0|Y_6 \geq 0) = 1$ then $\rho_{76} = 1$;
- if $P(Y_7 \geq 0|Y_6 \geq 0) = 0.5$ then $\rho_{76} = 0$;

When an expert stated that the value of $P(X_7 \geq x_{7_{50}}|\ X_6 \geq x_{6_{50}}) = 0.7$, then from figure (2.2) we read that $r_{76} = 0.57$. At this point we can already notice that to evaluate unconditional rank correlation between two random variables we need to have fast and accurate algorithm for numerical computation of double integral.

In the next step rank correlation between *Missed Approach Execution* and *Mean Cross Wind* conditioning on value of *Separation in Air* must be assessed. The expert is asked to state the value of $P(X_7 \geq x_{7_{50}}|\ X_5 \geq x_{5_{50}},\ X_6 \geq x_{6_{50}})$ as follows:

Suppose that not only the variable *Separation in Air* but also variable *Mean Cross Wind* are observed to be above their median

values. What is your probability that in this case the *Missed Approach Execution* would also lie about its median value?

After transformation of variables $X_7,\ X_6,\ X_5$ to normals $Y_7,\ Y_6,\ Y_5$ we get that the following probability is provided:

$$P(Y_7 \geq 0|\ Y_5 \geq 0,\ Y_6 \geq 0). \tag{2.8}$$

We can notice that in our example (figure 2.1) variables $X_5,\ X_6$ are independent and therefore also $Y_5,\ Y_6$ are independent. Using vector notation $y_{765} = (y_7,\ y_6,\ y_5)\ \in \Re^3$, we can write the trivariate joint density function as:

$$f(y_{765}) = \frac{1}{\sqrt{8\pi^3|\Sigma_{765}|}} \exp[-\frac{1}{2}(y_{765})\Sigma_{765}^{-1}(y_{765})'] \tag{2.9}$$

where $\Sigma_{765} = \begin{bmatrix} 1 & \rho_{76} & \rho_{75} \\ \rho_{76} & 1 & \rho_{65} \\ \rho_{75} & \rho_{65} & 1 \end{bmatrix} = \begin{bmatrix} 1 & \rho_{76} & \rho_{75} \\ \rho_{76} & 1 & 0 \\ \rho_{75} & 0 & 1 \end{bmatrix}$ is covariance matrix, with determinant $|\Sigma_{765}| = 1 - \rho_{76}^2 - \rho_{75}^2$. The value of correlation ρ_{76} was assessed by expert in previous step.

Relationship between ρ_{75} and probability obtained from experts is

$$\begin{aligned} P(Y_7 \geq 0|\ Y_5 \geq 0,\ Y_6 \geq 0) &= \frac{P(Y_7 \geq 0,\ Y_5 \geq 0,\ Y_6 \geq 0)}{P(Y_5 \geq 0,\ Y_6 \geq 0)} = \\ &= 4\int_0^\infty \int_0^\infty \int_0^\infty f(y_{765}).\, dy_7\, dy_6\, dy_5 \end{aligned} \tag{2.10}$$

We can determine the relationship between $\rho_{75|6}$ and conditional probability (2.10) benefiting from useful properties of normal copula. Knowing values of $\rho_{76},\ \rho_{75}$ and ρ_{65} we can calculate partial correlation $\rho_{75|6}$ by [15]:

$$\rho_{75;6} = \frac{\rho_{75} - \rho_{76} \cdot \rho_{65}}{\sqrt{1-\rho_{76}^2}\sqrt{1-\rho_{65}^2}}. \tag{2.11}$$

Since for joint normal distribution partial and conditional correlations are equal, hence $\rho_{75;6} = \rho_{75|6}$ and with Pearson transformation we can obtain $r_{75|6}$. Figure (2.3) shows relationship between $r_{75|6}$ and conditional probability (2.10), where previously assess $r_{76} = 0.57$.

We should notice here that possible values for probability $P(Y_7 \geq 0|\ Y_5 \geq 0,\ Y_6 \geq 0)$ are now restricted to interval $[0.4,\ 1]$. We can explain this

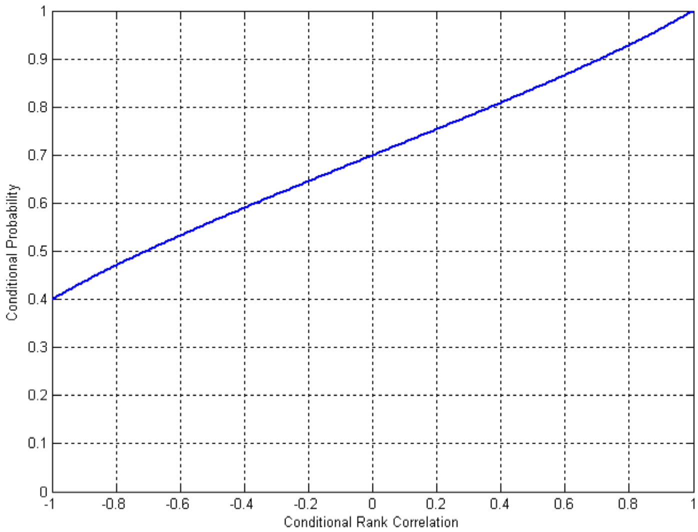

Figure 2.3: Conditional Probability $P(Y_7 \geq 0 |\ Y_5 \geq 0,\ Y_6 \geq 0)$ versus $r_{75|6}$.

fact in the following way. While expert provides his estimate for $P(X_7 \geq x_{7_{50}} |\ X_6 \geq x_{6_{50}})$ he/she describes how much information (variability) of X_7 is explained by X_6. Therefore the $P(X_7 \geq x_{7_{50}} |\ X_5 \geq x_{5_{50}},\ X_6 \geq x_{6_{50}})$ can only explaine remaining part of information about X_7 that was not explained already by X_6. If expert would think that variables X_7 and X_6 are independent, so $P(X_7 \geq x_{7_{50}} |\ X_6 \geq x_{6_{50}}) = 0.5$ then the possible values for $P(X_7 \geq x_{7_{50}} |\ X_5 \geq x_{5_{50}},\ X_6 \geq x_{6_{50}})$ would be again the whole interval [0, 1].

One can read from figure (2.3) that when an expert provides his assessment $P(X_7 \geq x_{7_{50}} |\ X_5 \geq x_{5_{50}},\ X_6 \geq x_{6_{50}}) = 0.5$ the corresponding conditional rank correlation $r_{75|6}$ is equal to -0.71.

For higher order conditional rank correlations we apply similar routine to the one just described. At each step the expert is asked to state the value of corresponding conditional probability where one additional conditioning variable is added. The relationship between the required conditional rank correlation and the conditional probability is found by integrating appropriate normal distribution.

One difficulty of using the approach to obtain conditional rank correlation for BBNs can be noticed. Namely, fast and accurate algorithm to calculate multivariate normal probabilities is needed. In order to calculate the conditional rank correlation where we condition on n random variables, the $n+2$ dimensional normal cumulative distribution needs to be calculated. This requirement is a motivation for our survey between different accessible algorithms which can be seen in chapter 4 of this thesis.

2.3 Updating

In the previous section the (conditional) rank correlations required to sample the BBN structure of the CFIT model shown in Figure 2.1 were obtained. Marginal distributions of each variable are derived from data or as described in Section 2.2.1 from expert judgment. As we have mentioned at the beginning of this chapter these information together with the copula assumption are sufficient to determine the whole joint distribution [33].

The main use of BBNs is in decision support, and in particular updating on the basis of possible observations. In the following section we aim to update our beliefs in the CFIT model given some observations. If for instance, new policies are proposed to be implemented, updating the BBN structure allows us to evaluate the impact of such policies on our variable(s) of interest.
In [18] techniques of dealing efficiently with the joint distribution when evidence becomes available (updating the BBN) are discussed. In this thesis the Normal Copula Vine Approach, which is implemented in a software application, called "UniNet"[4] is used. Since in this method all the calculation are performed on a joint normal vine [37], any conditional distribution can be computed analytically (which also will be normal with appropriate mean and variance). Finding the conditional distribution of the corresponding original variables is just a matter of transforming it back using the inverse distribution function of these variables and the standard normal distribution function. This simulation is performed very fast. Figure 2.4 presents CFIT model with assigned rank correlations and marginal distributions. At the bottom of each histogram the expectation (and standard deviation after the $\pm$) of each variable is shown.

The expectation of the number of missed approach executions divided by 100,000 would be an estimate of the probability of executing a missed approach maneuver. When there is no evidence available this probability is equal to 0.00179. Let us suppose that we know before the flight that the crew did not have enough sleep and moreover we observed that *separation*

[4]UniNet has been developed for the Project commissioned by the Dutch Ministry of Transport. Currently is still under development.

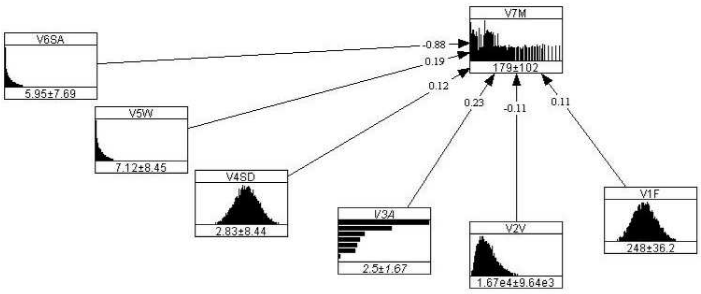

Figure 2.4: MissedAppoach Model Structure.

in air between the landing aircraft and the preceding aircraft is equal to 2 nautical miles (figure 2.5).

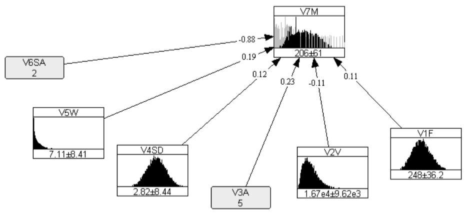

Figure 2.5: Conditional Missed Approach Model Structure.

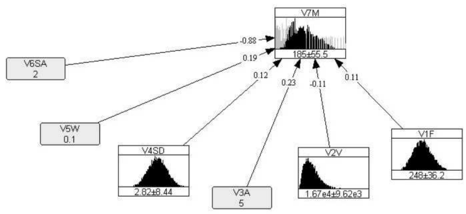

Figure 2.6: Conditional Missed Approach Model Structure.

The probability of executing a missed approach maneuver increases in this case to 0.0026. Let us assume additionally that *Mean Cross Wind* is equal to 0.1 Kt (figure 2.6). We can notice that this requirement “improves” situation and the expectation of the number of missed approach executions on 100,000 flights is equal to 185.

Similar type of analysis will be performed for the Maintenance Performance Model.

Chapter 3

Maintenance Performance Model

3.1 Overview

Human error in maintenance can impact on safety and performance in a number of ways. Poor repairs, for example, can increase the amount of breakdowns which can increase the risk associated with equipment failure and personal accidents. The aircraft maintenance human factors is said to be one of the last "frontiers" that can have significant impact on aviation safety. Industry statistics show that human error contributes to nearly 80%[1] of airline accidents and incidents, as illustrated on figure 3.1 . This figure includes all aspects of human factors such as operations, maintenance, and air traffic control.

Because aviation systems are continually improving, the aircraft devices seldom fail. Humans, rather than equipment, are more likely to be at the root cause of an accident or incident. Therefore, the best opportunity for safety improvement is to understand and manage the human factors that pose safety risks.

In the next sections we build the model for human performance in maintenance aviation industry. We try to recognize the most influential factors for the maintenance crew and propose a way of quantifying them. For the quantification part we turn our attention for the data available and in case

[1] "*Improving the Continued Airworthiness of Civil Aircraft*", A Strategy for the FAAs Aircraft Certification Service, National Research Council, 1988, *http://books.nap.edu/html/airworthiness/*

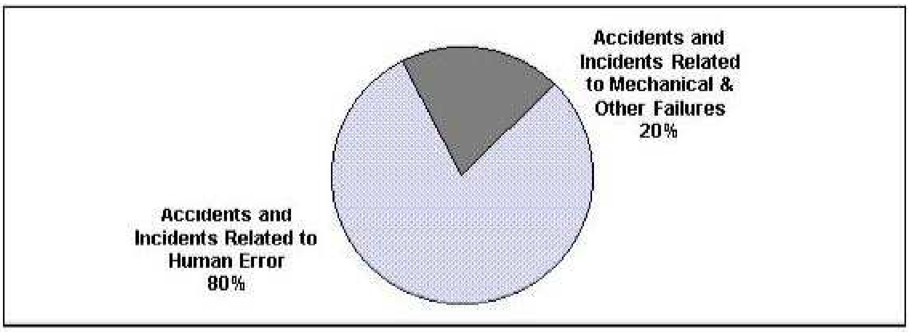

Figure 3.1: Accident & Incident Related to Human Error vs. Mechanical & Other Failures,*http://amelia.db.erau.edu/hfami/StrategicProgramPlan.pdf.*

of absence of data to expert judgment.

3.2 Building a model

Human behavior is difficult, if not impossible to predict. However, there are some common factors which influence this behavior. With the use of BBNs we are able to graphically represent the fact that 'elements' may influence and react with each other. Below we present the most important in our opinion features which influence the maintenance crew with a short description clarifying our choice. Moreover the one dimensional marginal distribution for each node are assessed. A combination of techniques can be used for this quantification. The use of historical data is often preferred, however, not always available. Expert Judgment is used in these cases. To ensure objectivity and transparency, it is essential that the expert opinions are obtained via a traceable, structured process ([34]). To complete the Maintenance Performance model the quantification of rank correlation for edges is also presented.

The first task while building the Maintenance Performance Model is to point out the most influential factors on human behavior. Moreover the dependence relations among them needs to be specified. On figure 3.2 the structure of our model is presented. Explanation why factors like: *alertness*, *working condition*, *communication*, *aircraft generation*, *job training* and *experience* are taken into account can be found in section 3.2.1.

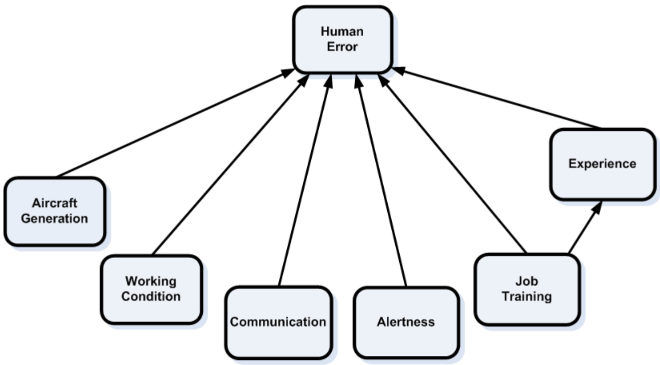

Figure 3.2: Maintenance Human Performance Model - structure.

3.2.1 Model variables

Alertness

Historically, sleep has been viewed as a state when the human organism is turned off. Scientific researchers have defined that sleep is a complex, active physiological state which is essential to human survival. Like human requirements for food and water, sleep is a vital physiological need. When an individual is deprived of food and water, the brain provides specific signals-hunger and thirst to drive the individual to meet these basic physiological needs. Similarly, when deprived of sleep, the physiological response is sleepiness.

How much sleep does an individual need? An individual requires the amount of sleep necessary to achieve full alertness and the highest level of functioning during working hours. There is a range of individual sleep needs and, though most adults will require about 8 hours of sleep, some people need 6 hours while others require 10 hours to feel wide awake and function at their peak level during wakefulness.

As some of the effects of not enough sleep we can consider: reduced reaction time, impaired short-term memory, decreased vigilance, reduced motivation, increased irritability, and an increase in the numbers of errors made, among others. Sleep deprivation is not the only cause of decreasing of alertness state. Among other causes we can distinguish i.e. too strong lighting or too much noise. However, time-on-duty and time-since-awake are common criteria researchers use to measure alertness (also called fatigue) based on their unambiguous definitions. In our project, alertness is defined as the total number of hours slept by the mechanic from the maintenance crew.

Data about hours of sleep of aircraft mechanics is obtained from [23]. Sleep data was collected using Actiwatch measurement equipment, which participants were asked to wear 7 days a week, 24 hours a day, for two weeks.

Shift	**N**^{2}	**Minimum**	**Maximum**	**Mean**
Day	30	3:24	6:38	5:06
Afternoon	19	2:40	6:31	5:00
Night	12	4:01	6:09	5:04
All	**61**	**2:40**	**6:38**	**5:05**

Table 3.1: Summary of Sleep Data.

A minimally informative distribution is fitted to the this data assuming the 3 point estimates (minimum, median and maximum hours of sleep of **all participants**) represent the 5th 50th and 95th percentiles of a continuous distribution function 3.3.

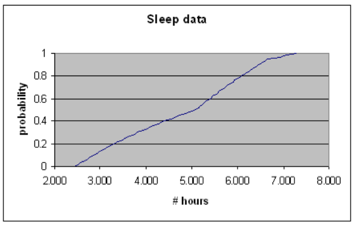

Figure 3.3: Cumulative distribution function for Alertness.

Experience

Usually the maintenance worker works in a specific aviation company ”for years”. However this years of experience between the crew members can be different. Mechanics just after concluding their studies do not possess yet such a useful practical experience as a person with 20 years of professional experience. From the other side his/her theoretical knowledge can be wider and includes most of latest innovations.

We define *Experience* as the total number of years worked by the mechanic in the aviation industry. Data on experience of aircraft mechanics is obtained from a Bureau of Labor Statistics[3]. Data about years of experience is presented in table 3.2.1.

Experience in years	**% of crew**
3 or less years	22.8%
4 – 9 years	28.5%
10 – 19 years	16.2 %
more than 20 years	32.5%
Median	9.4 years

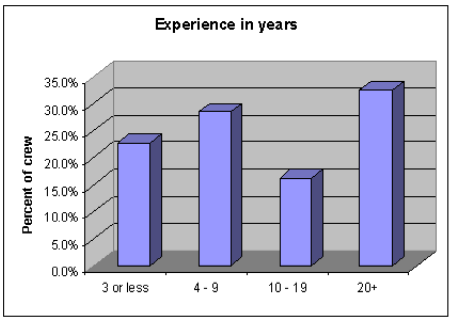

Figure 3.4: Experience of aircraft mechanics in years.

By looking at characteristic of data about years of *Experience* we can noticed that 67.5% of the whole population of workers have less than 20 years of experience. We can noticed that highest % of crew has more than 20 years of experience. To describe this data we could use four valued discrete distribution, however this approach is not so successful later while conditioning. Since we could only condition on intervals 0-3, 4-9, 10-19 or >20, and not on single values. Therefore instead of using discrete distribution to describe this data, we assume the continuous gamma distribution[4], which allows us late to conditionize on single values.

Aircraft generation

Nowadays in the aviation industry we can see different types of aircraft. New aircraft are becoming more sophisticated and complex, since they rely on composite construction and integrated electronics. However, aviation or-

[3]The BLS is an independent national statistical agency that collects, processes, analyzes, and disseminates essential statistical data to the American public, the U.S. Congress, and other Federal agencies, State and local governments, etc.

[4]The density of the gamma distribution is as follows:

$$g(x;\alpha,\beta) = x^{\alpha-1}\frac{\beta^{\alpha}\, e^{-\beta\, x}}{\Gamma(\alpha)} \text{ for } x > 0.$$

To describe the distribution of experience the gamma distribution with parameters $\alpha = 1.2,\ \beta = 0.098$ was used.

ganizations are still keeping older aircraft in service. Therefore the maintenance job for some type of aircraft can be easier than for others. From one side, we can suspect that older types of aircraft are easier to maintain that the newest one, since they contain a little bit less complicated systems in them. Some of the subsystems did not exist in earlier models, but exist now and have to be checked and repaired as well. However, from the other side while maintaining latest types of aircrafts worker can rely on quite accurate computer systems of measurements which makes the maintenance task easier.

Data on aircraft generation is obtained from "Shiphol Statistical Annual Review 2000-2002". This data was sorted by a scale from 1 to 4 where 4 is the most recent generation of aircrafts, and 1 denoted the oldest generation of aircrafts nowadays used. The detailed specification which types of the aircraft belongs to which generation can be found in Appendix B.

Type of the aircraft	**participation**
1	0.092%
2	5.908%
3	91%
4	3%

Table 3.2: Summary of Aircraft Generation data.

Working conditions

Aviation maintenance working conditions has many features in common with other industries. The physical facilities in which aviation technicians work, however, are unique. No other industry uses aircraft hangars as its primary work-site. The primary reason for using hangars is obvious, of course. Aviation maintenance technicians work on airplanes, and hangars are needed to shelter aircraft and workers. However there are a number of situations when the aircraft needs to be maintained outside the hangar. Probably when there are small repairs to perform, or due to time constrains. As the part of working conditions we could also distinguish if the worker need to deal with heavy parts being moved around, with rotating machinery, with toxic or hazardous materials, or with work locations that is above the ground. Although the above factors are considered of significant influence for the working condition node, the suitable definition combining all these effects appeared to be difficult to find or quantify. Therefore the only distinguish is made whether the work is performed outside or inside the hangar. The marginal distribution for this node is retrieved by expert judgment.

Figure 3.5: C-17 hangar facility at McGuire Air Force Base, New Jersey, April 5, 2006

Job training

Aviation maintenance has changed over the years. Newer aircraft contain materials, power plants, and electronic subsystems that did not exist in earlier models, and the number of older aircraft has increased. Technicians use more and more sophisticated equipment and procedures. It is understandable that to match these increasing requirements the workmen need to update their certain maintenance-related skills all the time. Moreover, aviation maintenance workers must maintain the skills and knowledge required to keep a wide variety of both new and old aircraft flying. It is reasonable to assume that there is interrelation between number of years of experience and the average number of job trainings.

For this node the average number of job trainings per year is used as formal definition for the quantification. In the absence of data we conduct the expert judgment elicitation.

Communication

Most of mechanics working in the aviation industry deal with the situation when the maintenance tasks often span for more than one shift. This requires information to be passed from one shift to the next. Shift turnover and therefore the information transfer can be a source of many errors, mostly due to the fact that information transfer can take various ways. For example, during a shift turnover there can be a meeting in the maintenance hangar. In this case the communication between one and another shift is through person-to-person communication. In this situation speech can be seen as a tool of communication. Information can be left in form of paper notes in workcards, or through the computer display equipment. At the extreme situation we can look when no informations are shared between shifts. Since the population which we consider refers to aircraft mechanics in western countries, situation in which no informations are shared between the shifts is hardly possible. In our project we define *communication* as the current messages transfer procedure in use, where we distinguish 2 categories:

1. only paper notes,
2. paper notes with oral feedback.

Also in this case expert judgment is used as the source for marginal distribution function.

The last variable is *human error*, which is our primary goal of interest.

Human Error

Formal definition of this node is: number of maintenance human errors that might lead to hazardous situations per 10,000 maintenance tasks. In this definition the word *task* is used instead operation, since one task can consist of few operations, i.e. replacement of the particular electrical device can be divided into following steps: unscrew the old device; disconnecting and removing old device; connecting and putting in place new device; screw new devise. In this thesis we are mostly interested in number of human error in the context of performed tasks rather than operations. Currently, we do not have any data about the number or probability of human error among maintenance crew and marginal distribution for this node is obtained through expert judgment procedure where the expert was asked to assess his/her 5%, 50%, and 95% quantile of their cumulative distribution function. However in the future those information will be obtained from data provided by DNV[5].

Table 3.3 presents the nodes considered in the Maintenance Performance Model together with their formal definitions and source of data for marginal distribution.

[5]Det Norske Veritas

Node	Definition	Unit	#	Source for marginal distribution
Human Error	Number of maintenance human errors that might lead to hazardous situations per 10,000 maintenance tasks	Number of errors	7	Expert Judgment
Job training	Average number of training per year	Number of trainings	1	Expert Judgment
Alertness	Average number of hours an aircraft mechanic sleeps per day	Number of hours	2	Data
Communication	Current information transfer procedure in use, distinguishing: 1. only paper notes, 2. paper notes with oral feedback	1 – 2	3	Expert Judgment
Experience	Average number of years a person worked as aircraft mechanic	Number of years	4	Data
Working Conditions	Average number of maintenance operations needed to be performed 1.out-side / 2. inside the hangar per 10,000 maintenance operations	1 – 2	5	Expert Judgment
Aircraft Generation	Aircraft generation in scale from 1 to 4 where 4 is the most recent generation of aircrafts	1 – 4	6	Data

Table 3.3: Variables used in Maintenance Performance Model.

Having defined the model variables, their (inter)relations are defined in the model structure as in Figure 3.6. Each variable identified in 3.3 is a node in the model structure.

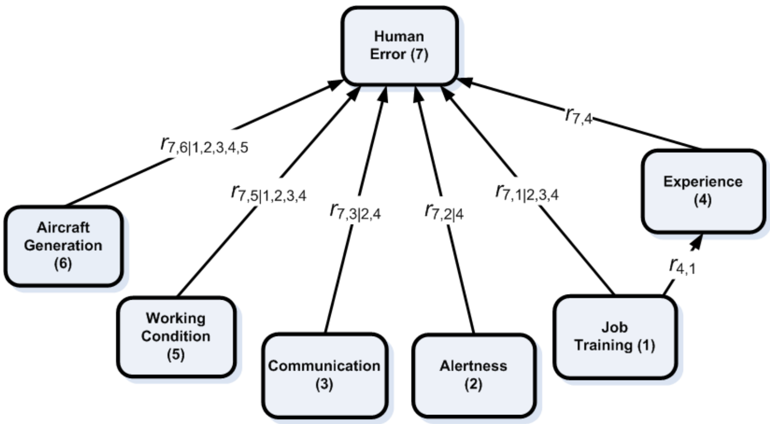

Figure 3.6: Maintenance Performance Model.

3.2.2 Quantification of the model

To quantify the model the marginal distributions need to be obtained together with (conditional) rank correlations specified by edges. In the table 3.3 we can see that for only three variables the data for marginal distribution is available. In case of variables: *Working Condition*, *Communication*, *Job Training* and *Human Error* the expert judgment elicitation protocol needs to be performed. Moreover the conditional rank correlations that are represented by arrows in the model must be determined from the conditional probabilities of exeedence provided by expert. The following order of the variables for elicitation is assumed: 4, 2, 3, 1, 5, 6.

A prototype elicitation using the method previously described (section 2.2.1, 2.2.2) was performed at NLR[6] on June 14$^{\text{th}}$ with a single expert who is a researcher for NLR. Expert were asked in total 11 questions, which 4 of them were used to assign marginal distribution to variables via classical method of expert judgment. To capture his uncertainty we asked expert to specify his 5%, 50% and 95% quantile of uncertainty distribution for each variable of interest. Later, to build expert distribution we find minimum information distribution with respect to background measure satisfying experts quantiles. The last 7 questions corresponds to conditional probabilities from which we obtained corresponding conditional rank correlations as described in section (2.2.2). Elicitation protocol used for this purpose can be found in Appendix C. In this questionnaire, it is assumed that the population considered refers to aircraft mechanics in western countries. In case of airplanes we consider the population of Western-built large aircrafts currently flying in commercials operations worldwide. Moreover, while asking about conditional probabilities expert was provide with bound for conditional probabilities only after his assessment was obtained to check correctness.

The results from the elicitation of dependencies are summarized in Table 3.4. Those numbers were calculated with software tool UniExp[7].

[6] National Aerospace Laboratory

[7] Description of the software UniExp can be found in Chapter 5.

Value of interest	**Index**	**P**	r
$r_{4,1}$	4:Exp, 1:JobT	0.4	-0.3
$r_{7,4}$	7:HE, 4:Exp	0.45	-0.15
$r_{7,2\|4}$	7:HE, 2:Alert ‖ 4:Exp	0.3	-0.44
$r_{7,3\|2,4}$	7:HE, 3:Com ‖ 2:Alert, 4:Exp	0.295	-0.15
$r_{7,1\|2,3,4}$	7:HE, 1:JobT ‖ 2:Alert, 3:Com, 4:Exp	0.2	-0.32
$r_{7,5\|1,2,3,4}$	7:HE, 5:WCond ‖ 1:JobT, 2:Alert, 3:Com, 4:Exp	0.18	-0.31
$r_{7,6\|1,2,3,4,5}$	7:HE, 6:AG ‖ 1:JobT, 2:Alert, 3:Com, 4:Exp, 5:WCond	0.179	-0.03

Table 3.4: Results from elicitation of dependencies for Maintenance Performance Model.

Therefore obtaining all required information (marginal distributions and conditional rank correlation) the Maintenance Performance Model can be quantified. In figure 3.7 we can see how (conditional) rank correlations are assigned to each arc of the BBN, together with marginal distribution to each node.

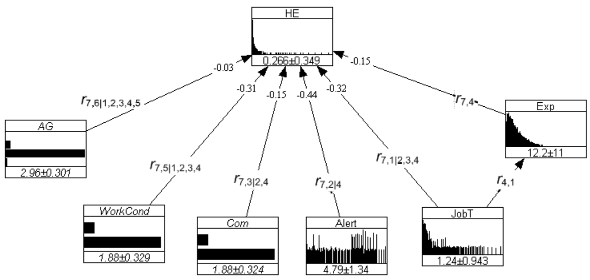

Figure 3.7: Maintenance Performance Model - quantified.

In ([33]) is shown that rank correlation specification on a BBN plus copula determines the whole joint distribution. In the following section we analyzes updating the conditional probability of the Human Error given some observations on certain variables.

3.3 Analysis of the model

In the previous section the (conditional) rank correlations required to sample the BBN structure of the Maintenance Performance model shown in Figure 3.7 were obtained. In the following section we aim to update our beliefs in the Maintenance Performance model given some observations. If for instance, new policies are proposed to be implemented, updating the BBN structure allows us to evaluate the impact of such policies on our variable(s) of interest. Software application, called UniNet is used for simulations.

In table 3.3 the nonzero unconditional rank correlations are presented.

Unconditional Rank Correlation						
	JobT	Alert	Com	Exp	WorkCond	AG
HE	-0.216	-0.434	-0.132	-0.15	-0.235	-0.0232

	JobT
Exp	-0.3

Table 3.5: Unconditional Rank Correlations for MPM.

We can notice that all of the model variables are negatively correlated with variable *Human error*. The negative rank correlation can be understood roughly as a degree to which one variable takes high values when the other takes low values, therefore while we conditionize on hight values (above median) of *Job Training*, *Alertness*, etc., we observe that the estimated number of *Human Errors* is decreasing.

Moreover, from table 3.3 can be noticed that the variables *Job training* and *Experience* are also negatively correlated. This can be explained in the following way: expert believes that new employee, which just entered the company and does not have many years of experience, will receive more trainings than the person which works already for many years. Therefore the average number of training for new employee is higher than for "older" employee.

In Figure 3.8 we can see the Maintenance Performance Model with assigned conditional rank correlations and marginal distributions. At the bottom of each histogram the expectation (and standard deviation after the $\pm$) of each variable is shown. The expectation of the number of human errors divided by 10,000 would be an estimate of the probability of human error per tasks. When there is no evidence available this probability is equal to $2.66 * 10^{-5}$.

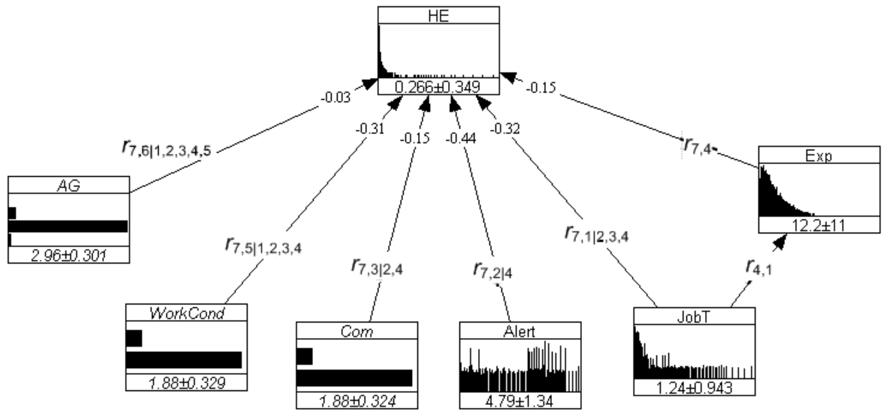

Figure 3.8: Maintenance Performance Model - quantified.

From table 3.3 we can noticed that the highest rank correlation in the absolute value is between the variables *Human error* and *Alertness* and is equal to -0.434. We can observe how low values of the *Alertness*, which means that the maintenance worker feels asleep and is not fully awake, influence the total chance of making mistake (committing error) on figure 3.9.

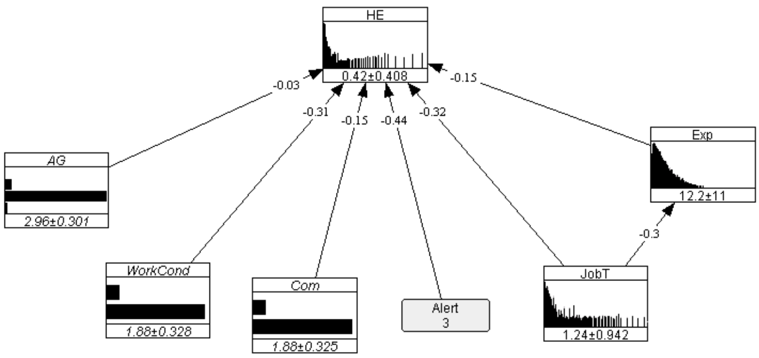

Figure 3.9: Conditional distributions of $X_7|Alert = 3$.

In Figure 3.9, the conditioning is performed on the value of the *Alertness* equal to 3 hours, which is equal to its 13 quantile. Such low amount of sleep can highly influence the behavior of worker and the amount of possible mistakes he/she commits. The expectation of the number of human errors increased in this situation from $2.66 \cdot 10^{-5}$ to $4.2 \cdot 10^{-5}$.

We can observe on figure 3.10 how previously obtained expected number of human error changes, when we now consider that the work is performed inside the hangar (which is more probable), together with the fact that worker is well trained.

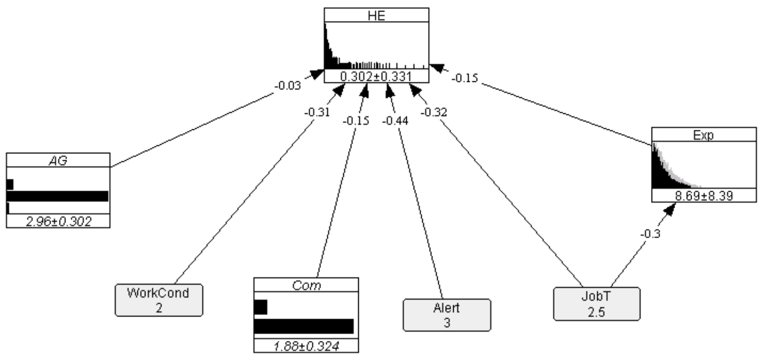

Figure 3.10: Conditional distributions of $X_7|Alert = 3$, $JobT = 2.5$, $WorkCond = inside$.

We should remained in this point that the variables *Job Training* and *Experience* are negatively correlated, and conditioning on high values of one variable effect that the other takes low values. The unconditional distribution of *Experience* is shown in grey behind the black histogram representing the conditional distribution of $Exp|JobT = 2.5$. When turning our attention to variable *human error* we can notice that in this situation the conditional mean is equal to 0.302 with standard derivation 0.331 human errors per 10,000 maintenance tasks. This is still a worst results than in the situation where no conditioning is performed, figure 3.8.

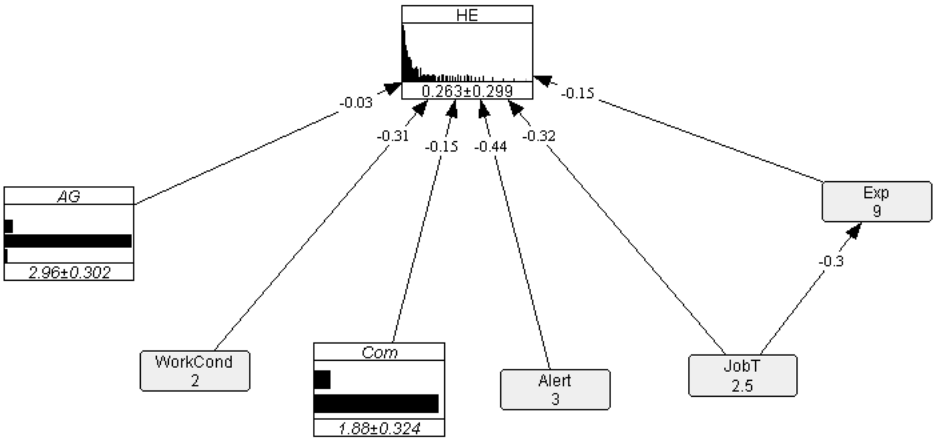

Figure 3.11: Conditional distributions of $X_7|Alert = 3$, $JobT = 2.5$, $WorkCond = inside$, $Exp = 9$.

To obtained the expected number of human errors similar to the one when no conditioning is performed, we would required additional information that worker has at least 9 years of experience.

By this simple example we can see how strong influence the variable *Alertness* has. To neglect the effect of insufficient sleep three additional information are required, namely that the task needs to be performed inside the hangar, that mechanic is well trained and moreover he/she has at least 9 years of experience.

Now, the joint distribution of the Maintenance Performance model is shown in Figure 3.12. Similar way of conditioning as previously described can be performed also in this situation. We can conditionalize this whole joint distribution on low values of the *Alertness* as presented in Figure 3.13. This visual representation allows us to observe the effect of this conditionalization on the whole joint distribution.

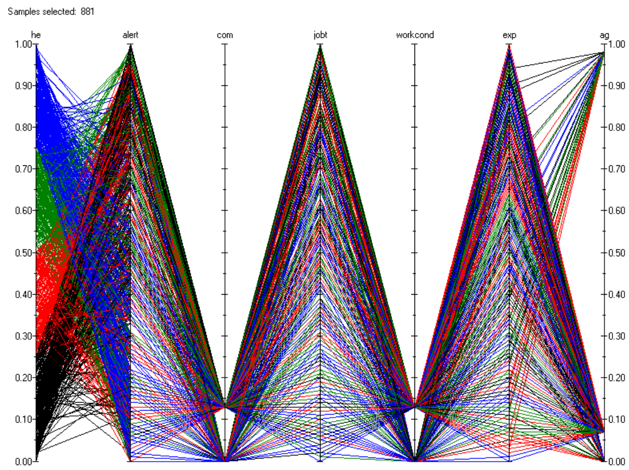

Figure 3.12: Cobweb plot of the joint distribution of the Maintenance Performance model.

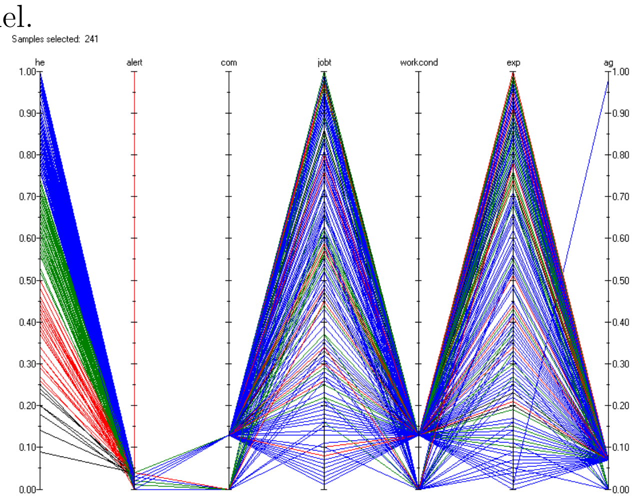

Figure 3.13: Cobweb plot of the Maintenance Performance conditionalized in low values of the Alertness.

We should state here that conditionalization is a "rough" way of carrying out a sensitivity analysis. In this way we simply evaluate the importance of the variables by just "guessing" or conditionalizing on different values of the variables, which are relevant in our criterion or client's criterion. However, the appropriate way of investigating the importance of some variable(s) for the maintenance performance is to calculate the correlation ratio[8]. This is accomplished below.

To carry out the sensitivity analysis several statistical and sensitivity measures are obtained by using the Sensitivity Analysis program [9] as part of Unicorn[9], based on 10^4 samples derived from a BBN created in UniNet.

The "predicted variables" are those whose behavior we want to explain in terms of other variables, called the "base variables". Here we are interested in the variable *Human error*, and we want to see how this variable depends on the variables *JobT*, *Exp*, *Alert*, *Com*, *WorkCond* and *AG*.

Table shows the sensitivity indices and statistics obtained by relating the *Human error* to each variable. These include: The product moment correlation, the Spearman rank correlation, the regression coefficient, the correlation ratio and the partial correlation coefficient.

Id	Predicted variable	Base variable	Product moment correlation	Rank correlation	Regresion coefficient	Correlation ratio	Partial corelation coefficient
1	HE	JobT	-0.1723	-0.2119	-0.0637	0.0326	-0.2371
2	HE	Alert	-0.3799	-0.4324	-0.0988	0.1539	-0.3951
3	HE	Com	-0.0711	-0.0842	-0.0765	0.0051	-0.0871
4	HE	Exp	-0.1356	-0.1681	-0.0043	0.0231	-0.2023
5	HE	WorkCond	-0.163	-0.173	-0.1727	0.0266	-0.1759
6	HE	AG	-0.0045	-0.0048	-0.0052	0.0002	0.0028

Table 3.6: Sensitivity indices for the predicted variable *HE* and a given base variable.

Each row in Table 3.6 shows the sensitivity indices for a given base variable, for the predicted variable *HE*. Note that *Alert* has the highest correlation ratio to the *Human Error*; while *AG* has the smallest.

[8] Refer to [8] for mathematical details concerning this sensitivity indices. In the Appendix A the definition of correlation ratio is presented as well.

[9] Unicorn (**Un**certainty Analysis with **Cor**relatio**n**s tool) developed at the Department of Mathematics of Delft University of Technology, The Netherlands.

In the BBN for the Maintenance Performance model the variable *Alertness*, which represents the average number of hours an aircraft mechanic sleeps per day is considered the most representative variable to explain the number of maintenance human errors that might lead to hazardous situations *Human error*. The generation of the aircraft on which maintenance task is performed is the considered as the less representative variable for *Human error*. The correlations ratio for variables *Job Training*, *Experience* and *Working Condition* are similar, therefore they effect on *Human Error* is comparable.

Chapter 4

Multivariate Normal Probability

Multiple integrals occur in a great variety of practical application. Very often there are just two or three integration variables (in which case the integral is said to be two-dimensional or three-dimensional). For low dimensional problems of this kind, the traditional method of integration are generally satisfactory and widely discuss in literature, see e.g. [14]. However there is increasingly a demand for cost-effective methods for high-dimensional integrals. There are often used in problems as, for example, in the quantum mechanics of atoms, molecules, or nuclei, or the theory of probability.

The need to perform fast calculations of multivariate normal probabilities arises also in process of quantifying the CATS model. As mentioned in previous sections *non-parametric* continuous BBN is used there as a modelling tool. In section 2.2.2 process of obtaining information about the dependence relations from the conditional probabilities of exeedence provided by experts is presented. In order to calculate those rank correlations the multivariate normal cumulative distribution needs to be calculated.

A common problem in many statistic computations is that of finding the probability of a multivariate normal random variable over a rectangle. We begin with the formal definition of the problem.

Let an n-dimensional rectangle $(\mathbf{a},\mathbf{b})$=$(a_1, b_1) \times (a_2, b_2) \times (a_n, b_n)$, and a symmetric n x n positive define matrix Σ and a n-dimensional vector μ be given. The problem is then to find

$$F(\mathbf{a},\mathbf{b}) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} \frac{1}{\sqrt{|\Sigma|(2\pi)^n}} e^{-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)}\, d\mathbf{x} \tag{4.1}$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n)$.

Some components of $\mathbf{a}$ and $\mathbf{b}$ may be infinite, and if $\mathbf{a} = (-\infty, \ldots, -\infty)$, then $F(\mathbf{a},\ \mathbf{b})$ is the value of the distribution function $F(\mathbf{b})$. Without loss of generality we assume further that Σ is a correlation matrix with ones on the diagonal, that is diag(Σ) = (1,. . . , 1) and $\mu = 0$ (we can standardize each component by subtracting its mean and dividing by its standard deviation [8]).

In the next section we describe several methods for calculating the multivariate normal probabilities.

4.1 Genz's transformation

The main idea of this method is to transform the original integral into an integral over a unit hyper-cube $[0,\ 1]^n$ by a sequence of elementary transformations.

The sequence begins with the Cholesky decomposition transformation $\mathbf{x} = \mathbf{C}\mathbf{y}$, where C is the Cholesky factor of Σ, that is, C is lower triangular matrix and $\Sigma = CC'$.

Now

$$\mathbf{x}'\boldsymbol{\Sigma}^{-\mathbf{1}}\mathbf{x} = (y'C')((C')^{-1}C^{-1})(Cy) = \mathbf{y}'\mathbf{y}$$

and

$$\mathrm{d\ x} = |C| d\ \mathbf{y} = |\Sigma|^{1/2}\mathrm{d}\mathbf{y}.$$

Moreover, the inequality $\mathbf{a} \le \mathbf{x} = \mathbf{C}\mathbf{y} \le \mathbf{b}$ is decomposed into the following set of inequalities:

$$a_1 \le y_1 \le b_1$$
$$\left(a_i - \sum_{j=1}^{i-1} c_{ij} y_j\right) / c_{ii} \le y_i \le \left(b_i - \sum_{j=1}^{i-1} c_{ij} y_j\right) / c_{ii}, i = 2, \ldots, n.$$

The integral (4.1) is therefore transformed into the iterated integral

$$\begin{aligned} F(\mathbf{a},\mathbf{b}) &= \int_{a_1}^{b_1} \frac{1}{\sqrt{2\pi}} e^{-y_1^2/2} \int_{a_2'(y_1)}^{b_2'(y_1)} \frac{1}{\sqrt{2\pi}} e^{-y_2^2/2} \dots \\ &\dots \int_{a_n'(y_1,\dots,y_{n-1})}^{b_n'(y_1,\dots,y_{n-1})} \frac{1}{\sqrt{2\pi}} e^{-y_n^2/2}\, dy_n, \dots\, dy_1 \end{aligned} \tag{4.2}$$

where

$$a_i'(y_1,\dots,y_{i-1}) = \left(a_i - \sum_{j=1}^{i-1} c_{ij} y_j\right) / c_{ii},$$

and

$$b_i'(y_1,\dots,y_{i-1}) = \left(b_i - \sum_{j=1}^{i-1} c_{ij} y_j\right) / c_{ii}.$$

Next, each y_i is transformed separately using $y_i = \Phi^{-1}(z_i)$, where $\Phi(t) = \int_{-\infty}^{t} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt$ is the standard univariate normal distribution function.

After these transformation, $\mathbf{F(a,b)}$ becomes

$$F(\mathbf{a},\mathbf{b}) = \int_{g_1}^{h_1} \int_{g_2(z_1)}^{h_2(z_1)} \dots \int_{g_n'(z_1,\dots,z_{n-1})}^{h_n'(z_1,\dots,y_{z-1})} dz_n, \dots\, dz_1 \tag{4.3}$$

where one defines

$g_1 = \Phi(a_1)$,
$h_1 = \Phi(b_1)$,
$g_2(z_1) = \Phi((a_2 - c_{12}\Phi^{-1}(z_1))/c_{11})$,
$h_2(z_1) = \Phi((b_2 - c_{12}\Phi^{-1}(z_1))/c_{11})$,

and for general $i = 3, \dots, n$,

$g_i(z_1,\dots,z_{i-1}) = \Phi((a_i - \sum_{j=1}^{i-1} c_{ij}\Phi^{-1}(z_j))/c_{ii})$,
$h_i(z_1,\dots,z_{i-1}) = \Phi((b_i - \sum_{j=1}^{i-1} c_{ij}\Phi^{-1}(z_j))/c_{ii})$.

The integrand in the form (4.3) is much simpler than the original integrand (4.1). The integration region is more complicated, however, and cannot be handled directly with standard numerical multiple integration algorithms. Therefore, the transformation $z_i = g_i + w_i(h_i - g_i)$ standardizes the integration region, giving

$$
\begin{aligned}
F(\mathbf{a},\mathbf{b}) &= (h_1 - g_1)\int_0^1 (h_2(w_1) - g_2(w_1))\int_0^1 \ldots \\
&= \cdots \int_0^1 (h_n(w_1,\ldots,w_{n-1}) - g_n(w_1,\ldots,w_{n-1}))\, dw_{n-1}\ldots\, dw_1 = \\
&= \underbrace{\int_0^1\int_0^1 \ldots \int_0^1}_{n-1\ integrals} f(\mathbf{w})\mathbf{dw} \qquad (4.4)
\end{aligned}
$$

where
$g_i(w_1,\ldots,w_{i-1}) = \Phi((a_i - \sum_{j=1}^{i-1} c_{ij}\Phi^{-1}(g_j + w_j(h_j - g_j)))/c_{ii})$,
and
$h_i(z_1,\ldots,z_{i-1}) = \Phi((b_i - \sum_{j=1}^{i-1} c_{ij}\Phi^{-1}(g_j + w_j(h_j - g_j)))/c_{ii})$.

With the sequence of transformation described here $\mathbf{F(a, b)}$ becomes now an integral of $f(w) = (h_1 - g_1)\cdot\cdots\cdot(h_n - g_n)$ over the (n - 1)-dimensional unit hypercube (since each integrand $h_i - g_i$ is independent of w_i and the innermost integral $\int_0^1 dw_n = 1$ can be ignored).

This sequence of transformations has also forced a priority ordering on the integration variables. The w_1 variable is the most important one, because all of the integrand factors f_i for $i > 1$, depend on it. The w_2 variable is the next most important one, and so on. Due to Genz [1] the method works best if the components are pre-sorted, so that the innermost integration has the most "weight". This can be done statically (i.e., by sorting the components once in increasing order of $\Phi(b_i) - \Phi(a_i)$) or dynamically.

Both of the two algorithms described below use Quasi-Monte Carlo method for computation, however the way of choosing quasi-random integration point set is different. Therefore we will first describe the main idea of Quasi-Monte Carlo method.

4.1.1 Quasi-Monte Carlo method

Monte Carlo (MC) methods are numerical techniques that exploit the stochastic properties of a problem to find a solution. The basic idea consists in drawing random samples from the relevant variables, and then using these samples to derive the required inferences. Below the idea of choosing random sample is explored in more details.

The problem is to approximate the integral of a function f as the average of the function evaluated at a set of points $w_1, \ldots, w_N$.

$$If := \int_{[0,1]^s} f(x)\,dx. \tag{4.5}$$

In the Regular Monte Carlo we let $w_1, \ldots, w_N$ be realizations of s-dimensional random vector W uniformly and independently distributed on the s-dimensional unit cube $[0,1]^s$ (thus each w_j is a vector of s elements). Then the arithmetic average

$$\overline{F}_{MC} = \frac{1}{N}\sum_{j=1}^{N} f(w_j) \tag{4.6}$$

approximates the $E[f(x)] = \int_{[0,1]^s} f(\mathbf{x})\,d\mathbf{x}$. Using the law of large numbers, it is possible to show that this estimate asymptotically converges to the *If* [17].

The key property of the random sequence is its uniformity, so that all subsequences are well spread throughout the cube. This idea has lead to the suggestion to use other than random sequences which are more uniformly distributed and produce better results. Such sequences are called quasi-random or low discrepancy sequences.

One of the most known examples of low discrepancy sequences in one-dimensional case is sequence $\frac{j}{N}$ where *j=1,...,N*, which also provides the approximation of the one dimensional integral *If* (4.5) by so-called one-dimensional rectangle rule

$$R_N f = \frac{1}{N}\sum_{j=1}^{N} f\left(\frac{j}{N}\right). \tag{4.7}$$

One can be now interested in generalizing the rectangle rule to higher dimensions. The most obvious generalization is the product-rectangle rule [17], but this rule is not cost-effective if the dimension s is high. This rule requires N^s integrand evaluation, and the rate of convergence can be very slow. In order to calculate approximation of the integral (4.4) many authors have proposed and examined different types of rules, all of which share the common name *quadrature rules*.

The oldest interesting generalization to s-dimensions of the rectangle rule is so-called *method of good lattice points*. These rules were developed independently by Korobov [32, 31], and Hlawka [11]. In this method the approximation of the integral is of the form:

$$If \approx \frac{1}{N}\sum_{j=1}^{N} f\left(\left\{\frac{j}{N}z\right\}\right) \tag{4.8}$$

where N is a priori chosen number of integration points and z is carefully selected an s-dimensional integer vector, and the braces around a vector indicate that each component is to be replaced by its fractional part, i.e., $\{(x_1,\ldots,x_s)\} = (x_1 \ mod\ 1,\ldots,\ x_s\ mod\ 1)$. It was shown that there exists z depending on N (where N satisfies certain arithmetic restriction, such as being a prime or a product of two primes). The method of good lattice points was investigated in a series of papers by Zaremba [36], and Niederreiter [16]. Two drawbacks are that *good lattice rules* are hard to obtain (choice of z) in high dimensions and that error is hard to estimate [16].

Since there is no easily computable and realistic error estimate for (4.7), Cranley and Petterson [7] proposed to randomize the quadrature rule (4.7) for which the error estimate can be computed. In this randomization procedure all of the quadrature points of a lattice rule (4.7) are "shifted" such as:

$$F_N(w) = \frac{1}{N}\sum_{j=1}^{N} f\left(\left\{\frac{j}{N}z + w\right\}\right). \tag{4.9}$$

where $\mathbf{w} \in \Re^s$ a random vector.

In order to estimate the variance of (4.8) we need to take a random sample of size q from w, where the distribution $\mathrm{F}(\mathbf{w})$ is assumed to be the multivariate uniform distribution. Then the composition of rules,

$$\begin{aligned}\overline{F} &= \frac{1}{q}\sum_{k=1}^{q} F_N(w_k) \\ &= \frac{1}{Nq}\sum_{k=1}^{q}\sum_{j=1}^{N} f\left(\left\{\frac{j}{N}z + w_k\right\}\right).\end{aligned} \tag{4.10}$$

is an unbiased estimator of (4.8). In the terminology of Cranley and Perrerson [7] (4.8) is called a *stochastic family of quadrature rules*. Calculation of this approximation required Nq function evaluation. The standard error of this approximation is given by $\sigma = \varsigma/\sqrt{q}$, where ς^2 is the variance of $F_N(w)$ (4.8). In general we do not know what the true value of σ is. Thus we use the estimate of this standard error given by [7]:

$$\sigma_{\overline{F}} = \sqrt{\frac{\sum_{k=1}^{q} \left(F_N(w_k) - \overline{F}\right)^2}{q(q-1)}}. \tag{4.11}$$

Usually the simulation size q is chosen to be very small (e.g 10 - 20), N denotes number of points used, and z is "strategically" chosen lattice vector. Braces around vectors again indicate that each components has to be replaced by its fractional part. Moreover $w_1, \cdots, w_N$ denote $[0,1]^s$ - uniform random vector.

For more details concerning quadrature rules in general and the approach presented above the reader is referred to book of Sloan and Joe [17].

In both algorithms presented below Genz ([1, 2]) propose, to evaluate function f from (4.9) at points

$$\left|2\left\{\frac{j}{N}z + w_k\right\} - 1\right| \tag{4.12}$$

instead on

$$\left(\left\{\frac{j}{N}z + w_k\right\}\right)$$

which appears to increase the accuracy of the lattice rule methods (4.11) (since quasi-Monte Carlo rules have better convergence properties for periodic integrand). Below I present description of two algorithms written by Genz, both implemented in Matlab environmental.

4.1.2 Algorithm I

In the method of good lattice points described above, one drawback which can be found is that whenever we change N we have to search for a new sequence of nodes in z. It is therefore desirable to have an infinite sequence available whose first N terms provide the N nodes for the numerical integration, and which yields a smaller error term than the method of low discrepancy sequences. Niederlander in [16] proposes and justifies the following choice of lattice vector z:

$$z_i = 2^{i/(s+1)} \text{ where } 1 \leq i \leq s. \tag{4.13}$$

To calculate the multivariate normal probability $F(\mathbf{a}, \mathbf{b})$ with lower integration limit column vector **a**, upper integration limit column vector **b**, and correlation matrix Σ, **Algorithm I** first applies the Genz's transformation described in section (4.1) and then uses a randomized quasi-random rule (4.9 and 4.11) with N points. The lattice vector z used for calculations is of the form (4.13). In the quasi-random rule we need to specify two numbers q and N. Therefore, first we will focus on the relationship between them and the time of calculations as well as how they influence the error of estimated probability.

In our investigation we use q equal 8, 14, together with different values of simulation size N. Later we will see that it is more effective to increase the simulation size N than q, since we can obtain better approximation of the multivariate normal probability for higher dimensions. However, the cost of obtaining better results is that the time of calculation increases. For each dimension $s = 3, 7, 12, 18$ we calculated the probability P ($a_i \leq X_i \leq b_i,\ i = 1, \cdots, s$) where $\mathbf{X} = [X_1,\ X_2,\ \cdots, X_s]$ has multivariate normal distribution with mean zero and randomly generated correlation matrix Σ.

In the first test we used the special form of constant correlation matrices. These matrices are defined in the following way: $\sigma_{i,j} = \rho$, for $i \neq j$ and $\sigma_{i,i} = 1$, where $\rho \in [0, 1]$. That means that in each correlation matrix all off-diagonal elements are equal. In further description those kind of correlation matrices are simply called *constant correlation matrices.* Example of four dimensional matrix is presented below:

$$\begin{pmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{pmatrix}$$

where $\rho \in [0, 1]$.
When we assume further that the elements of lower integration limit vector a are all equal to $-\infty$, then for such a problems the multivariate normal probabilities (4.1) can be computed using [38, p.193]:

$$F(b) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} \prod_{i=1}^{s} \Phi((b_i + \sqrt{\rho}t)/\sqrt{1-\rho})\, dt. \tag{4.14}$$

where $b \in \Re^s$ is upper integration limit vector.

Values for these probabilities are gathered in tables in [38] which we used for comparison with results from our algorithms. In simulation for each dimension s, we used sample of 56 different covariance matrices, with random ρ values, chosen uniformly from [0,1]. Moreover, the vector $\mathbf{b}=(b_1,\ldots,b_s)$ of upper limits were also random, chosen in way that $b_1 = b_2 = \cdots = b_s$, where $b_i \in [-2, 3.5]$ $i = 1, 2, \ldots, s$. Below we present tables with average estimated absolute error Ev_N for different values of N together with and the average running time (in seconds) Tv_N of algorithm. Value of average estimated absolute error was calculated as a mean of differences between true value of probability provided by [38] and approximation provided by our algorithm for all matrices. Value q was chosen equal to 8. We expect that with increasing number N, the average error of estimation should decrease, since we take more points into consideration.

s	Ev_{200}	Tv_{200}	Ev_{700}	Tv_{700}	Ev_{1500}	Tv_{1500}
3	2.595e-006	0.0198	2.014e-006	0.0547	1.382e-006	0.0913
7	1.207e-004	0.0593	3.536e-005	0.1449	2.693e-005	0.2730
12	1.398e-004	0.1111	4.982e-005	0.2675	2.831e-005	0.5096
18	2.872e-004	0.2051	1.246e-004	0.4694	7.553e-005	0.9754

Table 4.1: Constant covariance matrix, q = 8

In the table above we see that indeed the average error of estimation decreases while we increase number of samples N. Moreover this conclusion is true for all dimensions. Also we can say that **Algorithm I** is relatively fast, since to calculate the 18 dimensional integral it needs around 1 second. We can observe that while we increase number of points N from 700 to 1500 the average time of calculation almost doubled.

In the next table we provide result obtained with simulation size q equal 14. Covariance matrix and the integration limits remain unchanged.

s	Ev_{200}	Tv_{200}	Ev_{700}	Tv_{700}	Ev_{1500}	Tv_{1500}
3	9.818e-007	0.0366	7.930e-007	0.0839	1.103e-008	0.1570
7	9.036e-005	0.0997	2.643e-005	0.2400	2.211e-005	0.4596
12	1.157e-004	0.1845	3.500e-005	0.4372	2.414e-005	0.8697
18	2.695e-004	0.3097	9.071e-005	0.8629	5.786e-005	1.4928

Table 4.2: Constant covariance matrix , q = 14

Comparing result in Table 4.1 and 4.2 we see that the average error of estimation for small dimension $s = 3$ improvement about one decimal point. For $N = 1500$ this improvement is even about 2 decimal points. However this kind of impressive change cannot be observed for higher dimensions. For 12 and 18 dimensional integrals we see almost no improvement between average error of estimation for different q. Comparison between average time of calculation leads to the following conclusion: change of q from 8 to 14 leads to increase by ≈ 1.5 times the average time of the calculation. For the constant covariance matrices we see that increasing number q to 14 do not improve our results significantly comparing to the time which is used to calculation.

In the second test we used the 50 random covariance matrices. Word *random* should highlight that all off-diagonal elements are chosen randomly from [0,1] and do not have to be equal any more. The integration limits for these tests were chosen as in the previous tests. Since now we do not know the exact solution of the probability $\mathbf{F(b)}$ to compare, different measure of error estimation needs to be defined in this case. The **Algorithm I** returns the error estimate $3\sigma_{\overline{F}}$ (4.10) in order to provide an approximate confidence level of 99% for the algorithm. For each dimension the mean of those estimation is calculated and denoted as E_N, where N is number of points used to calculations. Tables 4.3, 4.4 shows the results of these tests.

s	E_{200}	Tv_{200}	E_{700}	Tv_{700}	E_{1500}	Tv_{1500}
3	3.6647e-004	0.0245	1.6447e-004	0.0558	7.7632e-005	0.1087
7	8.9856e-004	0.0688	4.9838e-004	0.1670	3.8259e-004	0.3150
12	1.2337e-003	0.1325	6.2360e-004	0.3141	3.9564e-004	0.6048
18	9.2182e-004	0.2259	4.0458e-004	0.5230	1.9139e-004	1.0150

Table 4.3: Random covariance matrix, q = 8

s	E_{200}	Tv_{200}	E_{700}	Tv_{700}	E_{1500}	Tv_{1500}
3	3.2762e-004	0.0365	1.3326e-004	0.0825	7.2297e-005	0.1575
7	7.3185e-004	0.1020	3.2482e-004	0.2440	2.6696e-004	0.4683
12	1.1204e-003	0.1957	5.5812e-004	0.4547	3.9030e-004	0.9940
18	1.1449e-004	0.3083	4.9851e-005	0.7232	3.0841e-005	1.4291

Table 4.4: Random covariance matrix, q = 14

In both tables we can see that time needed to estimate 18-dimensional integral is around 1/1.5 second. For dimensions 3, 7 and 12 we see satisfactory results in both tables. The average error of estimations increases while the dimension increases. Moreover together with the increasing number of N, the error gets smaller. Surprisingly both tables (4.3) and (4.4) provide for higher dimension (18) smaller average error then for integral of 12 dimension. It would mean that while the dimension of the integral increases our algorithm gives better and better result. Therefore, as a conclusion from this test, we can say that error provide by the **Algorithm I** is probably not a good measure in case of high dimensional integrals.

Since, from all four tables above we cannot conclude significant influence of number q to the average error. It is more effective to increase the simulation size N than q, since we can obtain better approximation of the multivariate normal probability for higher dimensions. Therefore for later calculation we will keep q constant and equal to 10.

4.1.3 Algorithm II

Algorithm II is different from the previous one in choice of quasi-random points. The most common choice of the lattice vector z (the one which was also used here) was proposed by Korobov [31], and is of the form:

$$z = (1,\ h,\ h^2\ mod\ N, \ldots, h^{s-1}\ mod\ N),\ 1 \leq h \leq \lfloor N/2 \rfloor.$$

where s is the dimension of the integral and N is the number of points used for calculation. There are many possibilities of choosing value h, and the one which was used in **Algorithm II** was proposed by Breckers [25] by minimized the bounds of error associated with approximation. More details and also some examples of vectors z (where z depends now on possible number of points N as well as on dimension of the integral s) can be found in [17, p.217-221].

In **Algorithm II** neither number N nor q can be specified by the user. The defined number for q is 25, and the number of points N depends on the estimated error. First **Algorithm II** calculates the approximation of the integral with N=n points, together with estimated error. If the provided error is too big ($> 10^{-4}$ for $s > 4$, and $> 10^{-7}$ for $s \leq 4$), then the number n is increased and the whole calculation is repeated for new N. **Algorithm II** stops when either the estimated error is smaller than $< 10^{-4}$ or the maximal number N has been reached ($N_{max} = 10000000$). Since for 2, and 3 dimensional integral, **Algorithm II** provides the estimation of the error of the order 10^{-7}, therefore we start with more interesting case with dimensions ≥ 4. Similarly as in the previous section we first tested constant and later randomly chosen covariance matrices. In the first table we see the average absolute error Ev for different dimensions s together with the average running time (in seconds) Tv of algorithm. The average absolute error Ev was calculated in the same matter as in case for algorithm I. In the second table E denotes the mean of errors provided by algorithm for each dimension.

s	Ev	Tv
4	1.5893e-005	0.2199
7	1.6428e-005	2.9325
12	1.7143e-005	5.7100
18	1.5178e-005	18.4213

Table 4.5: Constant covariate matrices

s	E	Tv
4	2.4039e-005	0.1994
7	7.6323e-005	7.8097
12	3.5599e-005	32.6454
18	8.0138e-005	56.3332

Table 4.6: Random covariate matrices

From the table 4.5 we can conclude that indeed the error between true solution and algorithm approximation is of order 10^{-5}. The same is true when we look at table 4.6 and error provided by the algorithm. However, this algorithm is much slower than **Algorithm I** and to compute the 18 dimensional integral with constant covariance matrices it needs ≈ 18 seconds, while for random covariance matrices around 56 second. Although, if we would look closer for calculation time we would see that for most of the matrices with dimension 12 and 18 algorithm needs much less than 32 and 56 seconds. Such a large average time of calculation is caused by few cases in which algorithm failed to calculate results with error smaller than requested even with maximal number of N. For them the time of calculation is around 200 seconds.

Both algorithms presented above will be compared with the standard numerical integration algorithms implemented in Maple. In the next sections reader can find description of two new methods together with the numerical comparison.

4.2 Maple software

Maple is one of most wide known and used software packages for symbolic calculations. It provides functions for numerical integration based on various rules of integration. The user can specify what kind of rule should be used for calculations. Below we present a small description of rules used for numerical integration.

4.2.1 Algorithm III

An Adaptive Multidimensional Integration Routine is the third integration algorithm which we want to investigate. The general algorithm of this routine is based on successive subdivisions of integration region, where each subdivision is used to provide a better approximation of integrand. These subdivisions are designed to dynamically concentrate the computational work in the subregions where the integrand *f(x)* is most irregular, and therefore adapt to the behavior of the integrand. This method is described in more details in [4].

With this algorithm, the subdivision strategy involves dividing (at each step) a subregion with largest estimated error. This subregion is halved, in the direction (chosen from some finite set of directions) where the integrand is most irregular. This algorithm uses possible subdivision directions that are parallel to the coordinate axes. The basic integration rule, that is used for the subregion adaptive algorithms take the form:

$$If = \int_H f(x), dx = \sum_{i=1}^{N} w_i f(\mathbf{x}_i) \tag{4.15}$$

where $H \in \Re^s$ is the region of integration, x_i the evaluation points and w_i the corresponding weights, $i = 1, \ldots, N$. The points $\mathbf{x}_j$ and weights w_j are chosen to make the rule exact for all polynomials of degree d or smaller, for some fixed d and s, with N small. The rules that are used in practical calculations typically have degrees in the range 5-13. The basic error estimate used for subregion adaptive integration algorithms is a difference of two integration rules. In **Algorithm III** 9 degree rule is used to approximation and difference of 7 and 9 degree rules to estimate the error.

4.2.2 Algorithm IV

Last method which we are going to investigate and compare is just a small modification of the **Algorithm III**. In this method first some techniques of symbolic analysis are used to simplify calculation and to deal with the singularities. Later this modified problem is pass to *An Adaptive Multidimensional Integration Routine*. In case if routine fails, which mean that routine does not provide results with requested by user accuracy, then the problem is treated via nested 1-D integration[1]. For one dimension integration, following methods are implemented: Clenshaw-Curtis quadrature method, adaptive double-exponential method, adaptive Gaussian quadrature method, adaptive sinc quadrature method and adaptive Newton-Cotes method.

4.3 Numerical comparison

This section present comparison of methods described so far. We will try to identify factors that might influence computation time and precision. In our investigation we will mainly focus on 4 and 7 dimensional integrals of multivariate standard normal density function. For simplicity we assume that elements of upper (lower) integration limit are equal to each other. For both dimensions $s = 4,\ 7$, we investigate two types of integration regions: $[-1,\ 1]^s$ and $[-0.5, \infty]^s$. The requested order of estimated error is 10^{-5} in case of **Algorithm II**, while for **Algorithm III** and **IV** its 4 or 6 digits. Moreover since in **Algorithm I** user needs to specify number of points used to calculation, we used 700 for dimension 4 and 1500 for dimension 7 (choice is motivated by read literature). For comparison different covariance matrices will be taken into account, depending of their determinant and eigenvalues.

[1] In case of multivariate normal integration all of the 1-D integration routines implemented by Maple failed to perform calculation in case of $s \geq 4$ dimensional integrals, regardless of covariance matrix and integration limits.

The standard notation used in tables below is:

- T_{A_i} - time of calculation of algorithm i,
- P_{A_i} - probability obtained by algorithm i,
- E_{A_i} - estimated error provide by algorithm i,
- dig - order of requested accuracy in digits.

4.3.1 Four dimensional integrals

First choice of covariance matrix is the identity matrix, where determinant and all eigenvalues are equal 1. Table 4.7 presents time of calculation and probability provided by all four algorithms. Moreover in case of **Algorithm III** and **IV** we make distinction when the required accuracy is 4 or 6 digits. In case of infinite region of integration those two methods do not provide precise value of estimated error. The only information which we have is that error is of order 10^{-4} or 10^{-6} as we wanted. Generally, if division of integration region is applied those methods do not provide the estimate of error.

limit	T_{A_1}	P_{A_1}	T_{A_2}	P_{A_2}	dig.	T_{A_3}	P_{A_3}	T_{A_4}	P_{A_4}
$[-1,1]^4$	0.10	0.217217	0.06	0.217217	4	0.02	0.2172	0.02	0.2172
					6	0.05	0.217217	0.03	0.217217
	E_{A_1}=0		E_{A_2}=1.98e-17		4	E_{A_3}=0.3e-4		E_{A_4}=0.3e-4	
					6	E_{A_3}=0.6e-6		E_{A_4}=0.6e-6	
$[-0.5,\infty]^4$	0.11	0.228599	0.08	0.228599	4	5.6	0.2286	4.8	0.2286
					6	20.5	0.228599	19.5	0.228599
	E_{A_1}=0		E_{A_2}=3.96e-17		4	–		–	
					6	–		–	

Table 4.7: Comparison of algorithms for matrix 1.

On the first glance, for integration region $[-1,\ 1]^4$ all four algorithms are extremely fast and their time of calculation if comparable. All four methods need much less than one second to obtain results, moreover in this simple case they return the same value for estimated probability. Now we turn our attention to estimated errors and compare times of calculations taking them into account. We can notice that algorithms III and IV required similar amount of time to obtain result with estimated error which is three times higher then for algorithms I and II. This can be seen as not promising fact for future comparison. When we change the integration region to $[-0.5,\infty]^4$ we can notice rapid increase of time for calculation for algorithms III and IV. To obtain accuracy of 6 digits they need approximately 4 times more time then for accuracy 4 digits. Also, since the division of integration region was applied the estimation of the error is not provided by these algorithms. Time for calculations for algorithms I and II have not changed significantly comparing to finite integration region. As in previous case the estimated probability is the same for all four algorithms. As the interesting fact we need to point out that Algorithm I returned value 0 as the estimation error. This can suggest that inside the algorithm the truncation of the error is applied.

Working with identity matrix for multivariate normal distribution can be seen as the most easy case, therefore our next choice is Matrix 2 with determinant equal to 0.5271. in figure 4.1 we see eigenvalues of this covariance matrix, together with matrix.

Determinant: 0.5271
Eigenvalues
0.3001
1.0007
1.0921
1.6071

$$\begin{pmatrix} 1 & -0.6003 & 0.0570 & 0.2125 \\ -0.6003 & 1 & 0.0643 & 0.1267 \\ 0.0570 & 0.0643 & 1 & -0.0668 \\ 0.2125 & 0.1267 & -0.0668 & 1 \end{pmatrix}$$

Figure 4.1: Matrix 2.

limit	T_{A_1}	P_{A_1}	T_{A_2}	P_{A_2}	dig.	T_{A_3}	P_{A_3}	T_{A_4}	P_{A_4}
$[-1,1]^4$	0.10	0.246965	0.06	0.246997	4	0.03	0.2470	0.03	0.2470
					6	0.09	0.246991	0.06	.246991
	E_{A_1}=1.26e-4		E_{A_2}=8.98e-5		4	E_{A_3}=0.6e-4		E_{A_4}=0.6e-4	
					6	E_{A_3}=0.2e-5		E_{A_4}=0.2e-5	
$[-0.5,\infty]^4$	0.09	0.225862	0.17	0.226050	4	7.7	0.2261	7.4	0.2261
					6	36.5	0.226052	35.6	0.226052
	E_{A_1}=5.16e-4		E_{A_2}=8.86e-5		4	–		–	
					6	–		–	

Table 4.8: Comparison of algorithms for Matrix 2.

For the integration region $[-1,1]^4$ all four methods required similar amount of computational time as for identity covariance matrix. Slight increase can be noticed only in case of algorithms III and IV, however those are still satisfactory results. The first method provides now the estimation of the error of order 10^{-4}, while the second method requested order 10^{-5}. When we look at the infinite integration region we observe increase of computational time for all methods beside first one. The most significant increase can be noticed for algorithms III and IV while we requested 6 digits of accuracy. In that case time of calculation is around 35 seconds. Algorithms I and II still provide results in less than one second. Probability provided by these methods is now slightly different due to different algorithms used to calculations, however this difference is not large.

Third matrix used to comparison has determinant ≈ 0.1 and eigenvalues shown below. Since this matrix has relatively small determinant we can expect worst results than in previous two cases.

Determinant: 0.1099
Eigenvalues
0.0636
0.7161
1.1862
2.0341

$$\begin{pmatrix} 1 & -0.6036 & 0.2233 & -0.0843 \\ -0.6036 & 1 & -0.1963 & -0.5487 \\ 0.2233 & -0.1963 & 1 & 0.5194 \\ -0.0843 & -0.5487 & 0.5194 & 1 \end{pmatrix}$$

Figure 4.2: Matrix 3.

limit	T_{A_1}	P_{A_1}	T_{A_2}	P_{A_2}	dig.	T_{A_3}	P_{A_3}	T_{A_4}	P_{A_4}
$[-1,1]^4$	0.09	0.294206	0.37	0.294273	4	0.40	0.2942	0.39	0.2942
					6	1.43	0.294228	1.35	0.294228
	E_{A_1}=2.60e-4		E_{A_2}=9.96e-5		4	E_{A_3}=0.2e-3		E_{A_4}=0.2e-3	
					6	E_{A_3}=0.2e-5		E_{A_4}=0.2e-5	
$[-0.5,\infty]^4$	0.10	0.184998	0.49	0.184948	4	15.7	0.1849	14.8	0.1849
					6	87.6	0.184945	84.4	0.184945
	E_{A_1}=1.23e-4		E_{A_2}=9.03e-5		4	–		–	
					6	–		–	

Table 4.9: Comparison of algorithms for Matrix 3.

Time of calculation for the first method remained the same as for matrices 1 and 2. However this is not the case for rest algorithms. For finite integration region second, third and fourth methods required approximately similar amount of time to obtain comparable results (accuracy for algorithms II can be almost treated as of order 10^{-4}). When we requested six digits of accuracy for method III and IV, computational time increased to more than one second. For invite integration region all of those thee methods perform much slower than for the second matrix. Time which they used for calculation has almost doubled what can be seen comparing tables 4.8 and 4.9. While the second method still needs less than one second to obtain results, methods three and four require about 15 seconds to obtain comparable results. For accuracy of six digits the user needs to wait more than one minute for the outcome. The best performance of the first method is obvious.

The last covariance matrix for which we compare our four algorithms is almost singular matrix with determinant ≈ 0.0275. Matrix together with its eigenvalues is shown below.

Determinant: 0.0275
Eigenvalues
0.0134
0.9553
1.1261
1.9052

$$\begin{pmatrix} 1 & -0.2803 & 0.870 & 0.2125 \\ -0.2803 & 1 & 0.0643 & 0.1267 \\ 0.870 & 0.0643 & 1 & -0.0668 \\ 0.2125 & 0.1267 & -0.0668 & 1 \end{pmatrix}$$

Figure 4.3: Matrix 4.

limit	T_{A_1}	P_{A_1}	T_{A_2}	P_{A_2}	dig.	T_{A_3}	P_{A_3}	T_{A_4}	P_{A_4}
$[-1,1]^4$	0.09	0.295772	1.74	0.295669	4	2.44	0.2957	1.97	0.2957
					6	9.83	0.295691	9.69	0.295691
	E_{A_1}=2.24e-4		E_{A_2}=6.696e-5		4	E_{A_3}=0.1e-3		E_{A_4}=0.1e-3	
					6	E_{A_3}=0.2e-5		E_{A_4}=0.2e-5	
$[-0.5,\infty]^4$	0.10	0.312291	1.57	0.312337	4	342.8	0.3124	311.3	0.3124
					6	–		–	
	E_{A_1}=8.3e-4		E_{A_2}=5.99e-5		4	–		–	
					6	–		–	

Table 4.10: Comparison of algorithms for matrix 4.

Almost singular covariance matrix can be seen as the most difficult case for calculation. For infinite integration region the second and the third method requires more than 300 second, which is more than 5 minutes, to perform calculation for requested accuracy of 4 digits. In case of 6 digits accuracy calculation were stopped after 700 seconds (more than 11 minutes). In this case, second algorithm required more than one second for calculations. Moreover we can observe that in this extreme case for the second method computational time is no longer dependent on the integration region, since for both of them we obtain comparable results. Relation between number of quasi random points used in first algorithms and computational time is now obvious.

4.3.2 Seven dimensional integrals

The same methodology of comparing was used for seven dimensional integral resulting in the following tables. Number of quasi random points in first algorithm is now changed and equal to 1500. We start the comparison with identity covariance matrix (matrix 5).

limit	T_{A_1}	P_{A_1}	T_{A_2}	P_{A_2}	dig.	T_{A_3}	P_{A_3}	T_{A_4}	P_{A_4}
$[-1,1]^7$	0.33	0.069113	0.12	0.069113	4	0.15	0.06911	0.11	0.06911
					6	0.81	0.0691135	0.73	0.0691135
	E_{A_1}=0		E_{A_2}=9.96e-18		4	E_{A_3}=0.3e-4		E_{A_4}=0.3e-4	
					6	E_{A_3}=0.4e-6		E_{A_4}=0.4e-6	
$[-0.5,\infty]^7$	0.32	0.075575	0.24	0.075575					
	E_{A_1}=0		E_{A_2}=0						

Table 4.11: Comparison of algorithms for matrix 5.

Table 4.11 presents time of calculations and probability provided by all four algorithms. Since in first method 1500 points are used computational time increase from 0.1 to around 0.3 second. As we have already observed for four dimension integral time for calculation for this method depends only on number of used points. Therefore approximately similar time will be observed for other covariance matrices. In case of infinite integration region third and fourth method needed a long time for calculation. In all of the cases for seven dimensional integration, algorithms did not yield results in less than 700 seconds and they were stopped. As already mentioned identity covariance matrix can be seen as the most easy case from computational point of view. When we look at the estimated error provided by algorithm I and II in table 4.11 we can conclude that they are quite accurate. Calculation time is also good. Here the second method is much better for both finite and infinite region of integration. Algorithms III and IV required similar amount of time to obtain results in case of 4 digits accuracy and finite integration region as the second algorithm. However the estimated error which they provided is three times higher than the one provide by the second method.

Next matrix used for comparison has determinant ≈ 0.5 and eigenvalues shown below.

Determinant: 0.489

Eigenvalues
0.4156
0.6404
0.7262
0.9092
1.0513
1.5552
1.7022

$$\begin{pmatrix} 1 & -0.0240 & 0.1653 & -0.0973 & -0.0974 & -0.0205 & -0.0957 \\ -0.0240 & 1 & 0.0289 & -0.0310 & -0.0672 & 0.3566 & 0.3214 \\ 0.1653 & 0.0289 & 1 & -0.1573 & -0.3348 & -0.2292 & 0.0505 \\ -0.0973 & -0.0310 & -0.1573 & 1 & 0.2576 & -0.1655 & -0.0695 \\ -0.0974 & -0.0672 & -0.3348 & 0.2576 & 1 & 0.2650 & 0.0032 \\ -0.0205 & 0.3566 & -0.2292 & -0.1655 & 0.2650 & 1 & 0.1583 \\ -0.0957 & 0.3214 & 0.0505 & -0.0695 & 0.0032 & 0.1583 & 1 \end{pmatrix}$$

Figure 4.4: Matrix 6.

Numerical results are visible in the following table.

limit	T_{A_1}	P_{A_1}	T_{A_2}	P_{A_2}	dig.	T_{A_3}	P_{A_3}	T_{A_4}	P_{A_4}
$[-1,1]^7$	0.33	0.081753	0.10	0.081771	4	0.42	0.08176	0.41	0.08176
					6	5.83	0.0817573	5.73	0.0817573
	E_{A_1}=1.06e-5		E_{A_2}=6.71e-5		4	E_{A_3}=0.4e-4		E_{A_4}=0.4e-4	
					6	E_{A_3}=0.4e-6		E_{A_4}=0.4e-6	
$[-0.5,\infty]^7$	0.31	0.082866	0.39	0.082900					
	E_{A_1}=4.83e-5		E_{A_2}=9.6e-5						

Table 4.12: Comparison of algorithms for matrix 6.

For the integration region $[-1,1]^7$ we can notice quite large difference between calculation times for all four methods. For the first and second algorithm the computation time almost have not changed comparing with the case with identity covariance matrix. However, estimated error is now of order 10^{-5} and not of order 10^{-17} as in previous case. For 4 digits of accuracy algorithms III and IV still provide results in less than one second, which is not the case for requested six digits. While we request additionally this two digits of accuracy time of computation increases drastically more than 10 times. When we look at the half-infinite integration region we observe comparable results.

The last covariance matrix for which we compare our four algorithms is almost singular matrix with determinant ≈ 0.11. Matrix together with its eigenvalues is shown below.

Determinant: 0.1102

Eigenvalues
0.1593
0.2374
0.8231
0.8937
1.1138
1.8532
1.9195

$$\begin{pmatrix} 1 & -0.0740 & 0.1653 & -0.1973 & -0.0974 & -0.2205 & -0.0957 \\ -0.0740 & 1 & 0.3289 & -0.0310 & -0.0672 & 0.5566 & 0.3214 \\ 0.1653 & 0.3289 & 1 & -0.1573 & -0.5348 & -0.2292 & 0.2505 \\ -0.1973 & -0.0310 & -0.1573 & 1 & 0.2576 & -0.1655 & -0.3695 \\ -0.0974 & -0.0672 & -0.5348 & 0.2576 & 1 & 0.2650 & 0.2032 \\ -0.2205 & 0.5566 & -0.2292 & -0.1655 & 0.2650 & 1 & 0.1583 \\ -0.0957 & 0.3214 & 0.2505 & -0.3695 & 0.2032 & 0.1583 & 1 \end{pmatrix}$$

Figure 4.5: Matrix 7.

Numerical results are visible in the following table.

limit	T_{A_1}	P_{A_1}	T_{A_2}	P_{A_2}	dig.	T_{A_3}	P_{A_3}	T_{A_4}	P_{A_4}
$[-1,1]^7$	0.33	0.103841	0.64	0.103829	4	3.77	0.1038	3.547	0.1038
					6	125.88	0.103801	119.766	0.103801
	E_{A_1}=1.3e-4		E_{A_2}=9.1e-5		4	E_{A_3}=0.5e-4		E_{A_4}=0.5e-4	
					6	E_{A_3}=0.5e-6		E_{A_4}=0.5e-6	
$[-0.5,\infty]^7$	0.31	0.087535	4.90	0.087537					
	E_{A_1}=1.6e-4		E_{A_2}=8.8e-5						

Table 4.13: Comparison of algorithms for matrix 7.

For the finite integration region and requested four digits of accuracy algorithm III and IV need approximately 3.5 second to perform calculation. In case of six digits of accuracy this time increased drastically to around 120 seconds (two minutes). Time which second algorithm needs to obtain accuracy of 10^{-5} is less than one second for $[-1,1]^7$, while for $[-0.5,\infty]^7$ it is almost 5 seconds.

From the provided numerical results the following conclusions can be made:

- In Algorithm I user needs to specify number of quasi random points used to calculation. There is no control of provided error of estimation. Moreover, time of calculation depends on the number of quasi random points, not of covariance matrix or integration region.

- Time of calculation for Algorithm II is comparable or sometimes grater than for Algorithm I. It depends on covariance matrix used in calculation together with integration region. Number of quasi random points depends on requested accuracy on solution.

- Algorithms III and IV are unpractical for large scale applications since they require long time for numerical calculations. Time for hypercube $[-0.5,\infty]^7$ is higher than 700seconds. Moreover when the procedure of subdivision of integration region is applied, algorithm do not provide the exact estimated error.

- Probabilities provided by all four methods are comparable.

Based on obtained results we decided that the second algorithm is the best among the tested ones. This can be justified as follows. Assigning constant number of quasi random points used in first method for each dimension, provided us with calculation time which is insensitive for changes of covariance matrix. Moreover first method do not provided any kind of control of obtained error of approximation. Those two features (sensitivity to covariance matrix and provided error) makes second algorithm more preferable in our view.

Chapter 5

Elicitation Software Tool - UniExp

5.1 Introduction

In this chapter we present the software tool UniExp which was created to support the elicitation of conditional rank correlations associated with the edges of a Bayesian Belief Network. The name of this package can be seen as the connections of two words: UNcertainty and EXPert - what can be explained as using expert knowledge to quantify uncertainty about dependence relations. The method of eliciting dependencies in the form of unconditional and conditional rank correlations through conditional probabilities of exceedence which is used in this software was discussed in section 2.2.2. The second algorithm presented in chapter 4 is implemented for calculations of multivariate standard normal probabilities. Here the brief overview of the software is presented.

5.2 Example

Let we suppose that we want to assess required rank correlations for the following BBN (figure 5.1).

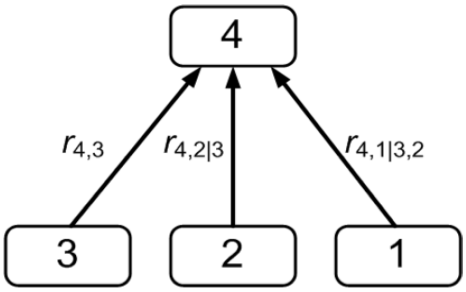

Figure 5.1: A simple example of BBN on 4 Variables.

The graphical structure of this BBN needs to be created by specifying names of random variables and connections between them (Section 5.3). Eliciting conditional rank correlations associated with the edges of a BBN from provided by expert conditional probabilities of exceedence is describe in section 5.4.

5.3 Creating graphical structure of BBN

The creation of input variables and connections between them is the first step in a UniExp project. To specify the number of nodes in BBN, a user needs to fill in a integer bigger than zero in the window, as it is shown below (figure 5.2). Later names of those nodes must be specified, see figure 5.3. The most important assumption used here is that *the names of the parents have to occur before the names of the children*, which is equal of having a topological order of the nodes. In case if names are not provided by user, the standard names are assigned.

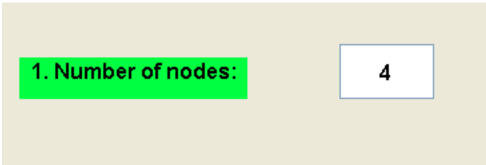

Figure 5.2: Input - number of nodes.

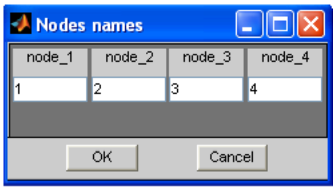

Figure 5.3: Input - nodes names.

To describe connections between nodes, the parent-child relations for each variable need to be specified. It is done in two steps, first by providing the number of parents for each node (figure 5.4), and later by specifying names of those parents (figure 5.5).

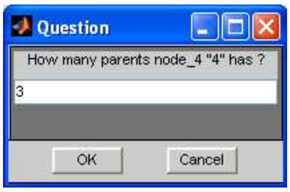

Figure 5.4: Input - assigning number of parents.

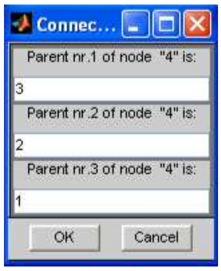

Figure 5.5: Input - connection window.

The order in which parents are written is quite essential since it determines which unconditional and conditional rank correlation needs to be assessed. From figure 5.5 we can read that the ranks correlations which will be calculated in the next step are: r_{43}, $r_{42|3}$ and $r_{41|32}$.

Alternatively names and connections can be imported from a files created elsewhere.

5.4 Assessing conditional rank correlations

The main part of the window "Rank_Rxv_zy" (figure 5.6) is concerned with calculating conditional rank correlations between parents and children in the BBN from excedeence probabilities provided by experts. Since the node *4* has three parents, three boxes for assessing conditional rank correlation will appear. Bounds for the first conditional probability P1 are filled in by the program immediately after pressing "Start" button. Moreover the bounds for unconditional rank correlation are shown (figure 5.6). The value of expert estimate for P1 needs to be included in the box between upper and lower bound[1] (figure 5.7). When the estimate of P1 is introduced, the calculation is initiated by pressing “Rxy” button.

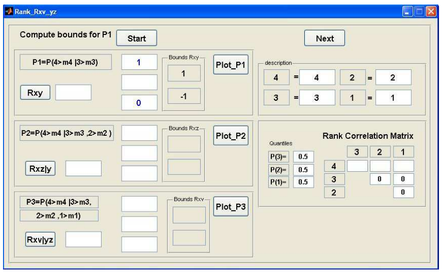

Figure 5.6: “Rank_Rxv_yz” window.

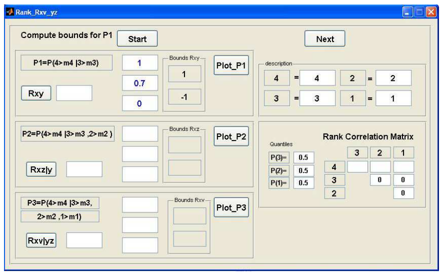

Figure 5.7: Initialized “Rank_Rxz_y” window.

After few seconds, the calculation of the rank correlation between variables *4* and *3* which satisfies the user’s conditional probability assessment is found. Moreover the bounds for next conditional probability and for unconditional rank correlation are calculated and filled in (figure 5.8).

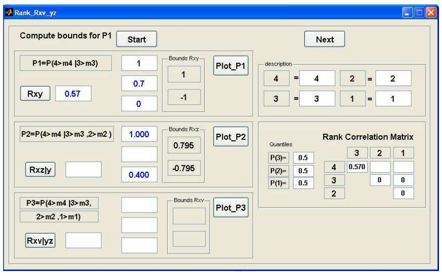

Figure 5.8: “Rank_Rxv_yz” window.

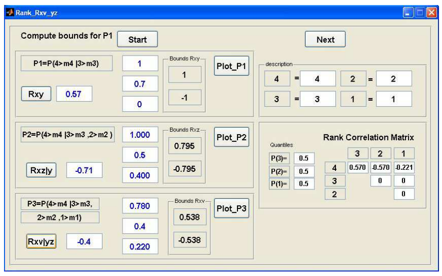

Figure 5.9: Initialized “Rank_Rxz_y” window.

After proving estimates for all required conditional probabilities and therefore calculating all conditional rank correlations associated with the edges of a Bayesian Belief Network (figure 5.9) program is finished.

[1]This estimate cannot be equal to upper or lower limit.

5.5 Output

The main goal of this software is to calculate conditional rank correlation assigned to the edges of BBN from excedeence probabilities specified by experts. Therefore, after the program is successfully finished, the user can find information about rank correlations in the file: *RankCorrelationValues.txt*. Moreover, in the file: *RankCorrelationMatrix.txt* the whole unconditional rank correlation matrix for the BBN can be found.

Chapter 6

Conclusions

This chapter summarizes the essential elements of the quantification of non-parametric Bayesian Belief Network described in this thesis on the example of the Maintenance Performance model, and presents the valuable findings that have been obtained during the study.

The following fundamental problems have been formulated and solved during the research:

- A model for Maintenance Performance was created in the form of a non-parametric BBN as a part of Causal Model for Air Transport Safety;
- quantification of the dependence relations in the Maintenance Performance model with use of experts judgment, together with the quantification of marginal distributions (based on data or expert judgment) was performed;
- the fast and accurate algorithm to calculate multivariate normal probabilities needed for quantification of dependence relations in the BBN was found.

Since human error/mistake in maintenance can impact on safety and performance in a number of ways, in this thesis a demonstration causal model for maintenance performance as a part of CATS model for aviation safety has been successfully developed. The construction of the model required a number of distinct steps to be carried out. First of all the most influential features in the maintenance crew behavior needed to be point out. Each of those variables needed to be described with unambiguous definitions, moreover the source of data for marginal distribution needed to be assign. A combination of techniques was used for quantification of marginal distributions. In cases where no data was available the Expert Judgment procedure was used. The following factors were taken into account while building the model:

1. Job Training - average number of training per year;
2. Alertness - average number of hours an aircraft mechanic sleeps per day;
3. Communication - current information transfer procedure in use distinguishing: 1. only paper notes, and 2. paper notes with oral feedback;
4. Experience - average number of years a person worked as aircraft mechanic
5. Working Condition - average number of maintenance operations needed to be performed 1.out-side / 2. inside the hangar per 10,000 maintenance operations;
6. Aircraft Generation - aircraft generation in scale from 1 to 4 where 4 is the most recent generation of aircrafts

To full quantify the Maintenance Performance model the information about the dependence relations was still required. These information was obtained from experts as well.

To obtain information about the strength of dependence relation between variables the information about required rank correlation has to be provided. In practice the conditional rank correlations are not elicited directly, the conditional probabilities of exceedence are more frequently asked. From them we can retrieve the desired rank correlations by assuming a copula. In our approach we assumed the joint normal copula due to its useful properties, such as know relationship between rank and product moment correlation and the fact that partial and conditional correlations are equal [8]. Since the relationship between required conditional rank correlation and the conditional probability is found by integrating appropriate normal distribution, the fast and accurate algorithm to calculate multivariate normal probabilities was needed. For that purpose different accessible algorithms were compared. We compared four algorithms, which belong to two groups due to the integration procedures which they use for the numerical computation of multivariate normal probabilities. The fist one is based on Quasi-Monte Carlo, while the second on the Adaptive Multidimensional Integration Routine.

Our computational study shows that the computation time depends on the problems dimension, the correlation structure, as well as on the required accuracy. The results reported in this thesis suggest that it is possible to reliably compute accurate multivariate normal probabilities for practical problems within as many as ten variables, in a few seconds or less. The best algorithm in our view was used to quantify the Maintenance Performance Model.

Moreover, the software tool UniExp which was created to support the elicitation of conditional rank correlations associated with the edges of a Bayesian Belief Network was presented.

In the future, when more data will become available to validate the Maintenance Performance model (marginal distribution or/and dependence relations), the similar sensitivity analysis to the one presented in this thesis can be performed. In this way we can assess the risk related to human error of the maintenance crew.

Bibliography

[1] Genz A. Numerical computation of multivariate normal probabilities. *J. Comput. Graph. Stat.*, 1:141–150, 1992.

[2] Genz A. and Bretz F. Numerical computation of multivariate t-probabilities with application to power calculation of multiple contrasts. *Journal of Statistical Computation and Simulation*, 63:361–378, 1999.

[3] Genz A. and Bretz F. Comparison of methods for the computation of multivariate t probabilities. *J.Comp.Graph.Stat.*, 11:950–971, 2002.

[4] Espelid T.O. Bernstsen J. and Genz A. Algorithm 698: An adaptive multidimensionsional integration routine for a vector of integrals. *ACM Transactions on Mathematical Software*, 17:452–456, 1991.

[5] Ale B.J.M. Causal model of air transport safety. *Risk Centre, Technical University of Delft*, 2006.

[6] et al. Cowell R.G. *Probabilitic Networks and Expert Systems.* Springer, New York, 1999.

[7] R. Cranley and T. N. L. Patterson. Randomization of number theoretic methods for multiple integration. *SIAM J. Numer. Anal.*, 13:904–914, 1976.

[8] Kurowicka D. and Cooke R.M. *Uncertainty Analysis with High Dimensional Dependence Modeling.* Wiley, New York, 2006.

[9] Lewandowski D. Current version of sensitivity analysis: v.2.01.10. *Department of Mathematics of Delft University of Technology*, 2007.

[10] Wesolowsky G. O. Drezner Z. On the computation of the bivariate normal integral. *Journal of Stat. Comput. and Simul.*, 35:101107, 1990.

[11] Hlawka E. and T. N. L. Patterson. Zur angengherten berechnung mehrfacher integrale. *Monatsh. Math.*, 66:140–151, 1962.

[12] Jensen F.V. *An Introduction to Bayesian Networks.* University College London Press, London, 1996.

[13] Donnelly T. G. Algorithm 462: Bivariate normal distribution. *Communications of the ACM*, 16:638, 1973.

[14] Evans G. *Practical Numerical Integration.* John Wiley & Sons Inc, West Sussex, England, 1993.

[15] Yule G. and Kendall M. *An introduction to the theory of statistics Analysis.* Charles Griffin and Co., Belmont, California., 14th edition, 1965.

[16] Niederreiter H. On a number-theoretical integration method. *Aequationes Mathematicae*, 8:304–311, 1972.

[17] Sloan I. H. and Joe S. *Lattice Methods for Multiple Integration.* Oxford University Press, Oxford, 1994.

[18] Kurowicka D. Hanea A.M. and Cooke R.M. Hybrid method for quantifying and analyzing bayesian belief nets. *Quality and Reliability Engineering International*, 22:709–729, 2005.

[19] et al. Hart J. F. *Computer Approximations.* New York: Wiley, 1968.

[20] Joyce S. A. Hill I. D. Algorithm 304: Normal curve integral. *Communications of the ACM*, 1967.

[21] Deák I. Three digit accurate multiple normal probabilities. *Numerische Mathematik*, 35:369380, 1980.

[22] Schervish M. J. Multivariate normal probabilities with error bound. *Applied Statistics*, 33:8187, 1984.

[23] Hall S. Johnson W., Mason F. and Watson J. Evaluation of aviation maintenance working environment, fatigue, and human performance. *FAA Report. Federal Aviation Administration.*, 2001.

[24] Pearson K. Mathematical contributions to the theory of evolution. *Biometric.*, Series I:1–34, 1904.

[25] Beckers M. Numerical integration in high dimensions. unpublished Ph.D. thesis., 1992.

[26] Cooke R. M. *Experts in Uncertainty.* Oxford University Press, 1991.

[27] Cooke R. M. and Slijkhuis K.A. *Expert Judgment in the Uncertainty Analysis of Dike Ring Failure Frequency.* Delft University of Technology, Delft/Bouwdienst, Utrecht, The Netherlands, 1999.

[28] Cooke R. M. and Goossens L.H.J. Procedures guide for structured expert judgment in accident consequence modeling. *Radiation Protection Dosimetry*, 90(3):303–309, 1996.

[29] Kendall M.G. Proof of relations connected with the tetrachoric series and its generalizations. *Biometrika*, 32:196198, 1941.

[30] Kurowicka D. Morales O. and Roelen A. Eliciting conditional and unconditional rank correlation from conditional probabilities. *Reliability Engineering and System Safety*, 2007.

[31] Korobov N.M. The approximate calculation of multiple integrals using number-theoretic methods. *Dokl. Akad. Nauk SSSR*, 116:1062–1065, 1957.

[32] Korobov N.M. The approximate calculation of multiple integrals. *Ibid*, 124:1207–1210, 1959.

[33] Cooke R.M. and Kurowicka D. Distribution - free continuous bayesian belief nets. *Proc.Mathematical Methods in Reliability Conference*, 2004.

[34] Cooke R.M. and Goossens L.H.J. Procedure guide for structured expert judgment in accident consequence modelling. *Radiation Protection Dosimetry*, 90(3):303–311, 2000.

[35] Gupta S. S. Probability integrals of multivariate normal and multivariate t. *Annals of Mathematical Statistics*, 34:792828, 1963.

[36] Zaremba S.K. Good lattice points, discrepancy and numerical integration. *Ann.Mat.Pura Appl.*, 73:293–318, 1966.

[37] Bedford T.J. and Cooke R.M. Vines a new graphical model for dependent random variables. *The Annals of Statistics*, 32(4):1031–1068, 2002.

[38] Y.L. Tong. *The multivariate normal distribution.* Springer-Verlag, New York, 1990.

Appendix A

Basic Definitions

Through the whole thesis we assumed that the reader is familiar with some definitions from statistics and probability field. However to make this thesis more self contained we provide those definitions here.

Definition A.1 ***Product Moment Correlation***
The correlation $\rho_{X,Y}$ between two random variables X *and* Y *with expected values μ_X and μ_Y respectively, and standard deviations σ_X and σ_Y^2 is defined as:*

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y}. \tag{A.1}$$

Remark
The correlation is the measure which takes values from [-1,1] and is associated with the strength of the relationship between two variables. If correlation coefficient ρ takes value 1 - then it means that there is a perfect linear relationship, in case when $\rho =$ -1 the perfect linear relationship is negative, but when $\rho = 0$ - then there is no relationship between variables.

Definition A.2 ***Sperman's Rank Correlation***
For random variables X and Y with cumulative distribution functions F_X and F_Y respectively, ***Spearman's rank correlation*** *is defined as:*

$$r(X,Y) = \rho(F_X(X), F_Y(Y)). \tag{A.2}$$

Another way of defining rank correlation is by introducing the so called population version:

$$r = 3(P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]) \tag{A.3}$$

where (X_1Y_1), (X_2, Y_2) are two independent identically distributed random vectors.

Remark The most important properties of rank correlation r are: Spearman's rho always exists, and does not depend on marginal distributions; it is invariant under non-linear strictly increasing transformations ; if $r(X, Y) = 1$ then there exists a strictly increasing function $G : \Re \to \Re$ such that $X = G(Y)$.

Definition A.3 ***Partial Correlation***
Let X_i be random variables with zero mean and standard deviations σ_i, $i = 1, \ldots, n$, and let the numbers $b_{12;3,\ldots,n}$, $\ldots$, $b_{1n;2,\ldots,n-1}$, minimize the following expected value:

$$E((X_1 - b_{12;3,\ldots,n}X_2 - \cdots - b_{1n;2,\ldots,n-1}X_n)^2). \tag{A.4}$$

Then partial correlation is defined as [8]:

$$\rho_{12;3,\ldots,n} = sgn(b_{12;3,\ldots,n})\sqrt{b_{12;3,\ldots,n}b_{21;3,\ldots,n}}. \tag{A.5}$$

Partial correlations can be computed from correlations with the following recursive formula [8]:

$$\rho_{12;3,\ldots,n} = \frac{\rho_{12;3,\ldots,n-1} - \rho_{1n;3,\ldots,n-1}\rho_{2,n;3,\ldots,n-1}}{\sqrt{1-\rho^2_{1n;3,\ldots,n-1}}\sqrt{1-\rho^2_{2,n;3,\ldots,n-1}}}. \tag{A.6}$$

Let us take the partition (X_i, X_j, X_a) of the n-dimensional random vector $X = (X_1, X_2, \ldots, X_n)$, where $a = 1, 2, \ldots, n \backslash i, j$, so the vector X_a is a vector composed of all the other variables except X_i and X_j.

Definition A.4 ***Conditional Correlation***
The conditional correlation of X_i and X_j given X_a denoted by $\rho_{X_iX_j|X_a}$ (or simply $\rho_{ij|a}$) is a product moment correlation of X_i and X_j conditioned on X_a with respect to the conditional distribution between X_i and X_j conditioned on X_a.

$$\rho_{X_iX_j|X_a} = \rho X_i|X_a, X_j|X_a = \frac{E(X_iX_j|X_a) - E(X_i|X_a)E(X_j|X_a)}{\sigma(X_i|X_a)\sigma(X_j|X_a)}. \tag{A.7}$$

Definition A.5 ***Correlation Ratio***
For random variables $G, X_1, \ldots, X_n$, the correlation ratio of G with X_i is:

$$CR(G, X_i) = \frac{Var(E(G|X_i))}{Var(G)}. \tag{A.8}$$

Remark Unlike correlation, the correlation ratio is not symmetric; that is $CR(G, X) \neq CR(X, G)$. The correlation ratio is treated as a measure of the relationship between the statistical dispersion within individual categories and the dispersion across the whole population or sample.

Definition A.6 ***Quantiles***
Lets take random variable X. *A* p *quantile is such an* q *that*

$$P(X \leq q) = p. \tag{A.9}$$

Theorem A.1 ***Sklar's Theorem***
Let H *denote* n*-dimensional distribution function with marginal distributions $F_1, \ldots, F_n$. Then there exists the* n*-dimensional copula* C *such that for all $(x_1, \ldots, x_n)$:*

$$H(x_1, \ldots, x_n) = C(F(x_1), \ldots, F(x_n)). \tag{A.10}$$

If all marginals are continuous then the copula is unique. The converse of the above statement is also true.

Definition A.7 ***Copula***
Let $F_1^{-1}, \ldots, F_n^{-1}$ denote the inverses of marginal distributions, then for every $(u_1, \ldots, u_n)$ there exists unique copula C *such that:*

$$C(u_1, \ldots, u_n) = H(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)). \tag{A.11}$$

denote n*-dimensional distribution function with marginal distributions $F_1, \ldots, F_n$. Then there exists the* n*-dimensional copula* C *such that for all $(x_1, \ldots, x_n)$:*

From this proposition we know that given any marginals and a copula we can construct a joint distribution. The copula density contains all information about dependence in the random vector.

Definition A.8 ***Normal Copula***
The Φ_Σ is the multivariate normal cumulative distribution function with correlation matrix Σ if the distribution function of normal copula is given by:

$$C_N(u_1, \ldots, u_N) = \Phi_{\Sigma^N}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_N)). \tag{A.12}$$

Definition A.9 ***Bivariate Normal Distribution***
(a) A two-dimensional random variable $\mathbf{X} = (\mathrm{X}_1, \mathrm{X}_2)'$ *is said to have a nonsingular bivariate normal distribution if its density function is of the form*

$$f\left(x; \mu, \sigma^2\right) = \frac{1}{2\pi|\Sigma|^{1/2}} e^{-(x-\mu)'\Sigma^{-1}(x-\mu)}, \mathbf{x} \in \Re^2, \tag{A.13}$$

where

$$\mu = (\mu_1, \mu_2), \ \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix},$$

$\sigma_i^2 > 0\,(\mathrm{i} = 1, 2),$ *and* $|\sigma_{12}| < \sigma_1\sigma_2$.

(b) $\mathbf{X}$ *is said to have a singular bivariate normal distribution if there exist real numbers* $\sigma_1, \sigma_2, \mu_1, \mu_2$ *such that* $\mathbf{X}$ *and* $(\sigma_1 Z + \mu_1, \sigma_2 Z + \mu_2)'$ *are identically distributed, where* Z *has an N(0,1) distribution. The* μ *and* $\mathbf{\Sigma}$ *are called, respectively, the mean vector and the covariance matrix of the bivariate normal distribution.*

Definition A.10 ***Positive Definite Matrix***
An $n \times n$ *symmetric matrix* Σ *is said to be positive definite if* $\mathbf{c}'\mathbf{\Sigma}\mathbf{c} \geq 0$ *holds for all real vectors* $\mathbf{c}$, *and equality holds only for* $\mathbf{c} = \mathbf{0}$.

Definition A.11 ***Multivariate Normal Distribution***
An n*-dimensional random variable* $\mathbf{X}$ *with mean vector* μ *and positive definite covariance matrix* Σ *is said to have a multivariate normal distribution if the density function of* $\mathbf{X}$ *is of the form*

$$f(x; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n|\Sigma|}} e^{-(x-\mu)'\Sigma^{-1}(x-\mu)/2}. \tag{A.14}$$

Appendix B

Aircraft Generation

- First generation: Antonov 12/ 24/ 26, BAE HS 748, Belfast Boeing 707F, Canadair CL 64, DC6A, DC8-62MC, DC8-63F, DC8F, Falcon 20, FK/FA 27-200, FK27-500, Fokker 27-500, Hercules M382, L188F, Merlin III, Metro III, Short 360.
- Second generation: Airbus A300/ A300-4P/ A300B4/ A300F, Antonov 74, BAC 1-11-500, Beech 1900C/ 1900D, Boeing 727-200/ 727-200ADV/ 727F/ 737-200ADVP/ 737-200P/ 747-100F/ 747-100P/ 747-200F/ 747-200MC/ 747-200P/ 747-300MC/ 747-300P/ 747SP, DC10-10P, DC10-30P, DC10-40, DC10F, DC9-20, DC9-30P, DC9-40, DC9-50, Dornier 228-200, EMB-110, EMB-120, ilyushin 86, Merlin IV, Tristar 1, Tristar 100/ 500, Tupolev 134/ 154, YAK42.
- Third generation: Airbus 330-300/ A300-6P/ A310-2P/ A310-3P/ A310F/ A319/ A320-1/ A320-2/ A321/ A330/ A330-300/ A340-211/ A340-311/ A340-313, Antonov 124, ATR42 300/ 500, ATR72-202, BAE 146-100/ 146-200P/ 146-300/ 146-RJ85/ ATP/ 146, Boeing 717/ 737-300/ 737-400/ 737-500/ 737-600/ 737-700/ 737-800/ 737-800W/ 737-900/ 747-400F/ 747-400MC/ 747-400P/ 757-200F/ 757-200MC/ 757-200P/ 757-300/ 767-200/ 767-300/ 767-400/ 777-200, Bombardier CL604 Canadair CL600/ CL600/ 604/ RJ700/ 100/ 200ER, Cessna 650, Dash 8-100/200/ 8-300/ 8-400/ DO328JET, Dornier 328-100/ 328JET, EMB 135/ 145, Fokker 100/ 50/ 70/ MD11/ 11F/ 81/ 82/ 83/ 87/ 88/ 90-30, Saab 2000/ SF340.
- Fourth generation: Airbus A319/ A320-1/ A320-2/ A321/ A330/ A330-300/ A340-211/ A340-311/ A340-313, Boeing 777-200.

Appendix C

Elicitation protocol

Introduction

Thank you for agreeing to serve as an expert in this expert judgment exercise, which provides the quantitative assessment of the relations between influential factors in the Maintenance Human Error Model.

The main purpose of this project is to obtain information on the probability distribution of variables (such as experience, fatigue, massages transfer, etc.) influencing risk associated with human error in the Maintenance Human Error Model. Moreover, information about the dependence relation between these variables will be assessed.

Elicitation Method

Eliciting Marginal distribution

The method used in this elicitation is called structured expert judgment. For each variable of interest, the probability distribution is elicited from experts' uncertainty estimates via a structured protocol of questions in section 3.1.

The information about the probability distribution of variables will be elicited via three numbers:

- The median value of the distribution, i.e. if you would have 100 samples of the variable value, then 50 of them should be below and 50 should be higher that the median value.
- The 5% percentile value (also called 5% quantile), which can be interpreted as: it would surprise you if more than 5 out of 100 samples have

a value *lower* than this value.

- The 95% percentile value (also called 95% quantile), which can be interpreted as: it would surprise you if more than 5 out of 100 samples have a value *higher* than this value.

Later, in order to create a probability distribution for each variable, the minimum informative distribution is if fitted to this quantiles.

Elicitation of Dependencies between variables

The information on the dependencies between certain variables will be elicited via "conditional probabilities", i.e we ask for the probability that variable X is above its median value, given the condition that variable Y is above its median value:

$$P(X \geq x_{50}|\ Y \geq y_{50}).$$

The main methodology of this procedure will be explained on the example below. Let us consider the population of employees in a certain organization, and assume that:

- their median age is 40 years,
- their yearly gross salary is around 45,000 euros,
- the median number of projects worth more than 150,000 euros in which the employee has participated is 5.

Let us now look more closely on the 20,000 samples from the population of employees. For each person from this sample file we can plot their age and salary as shown on picture (C.1). Now, let consider the part of the population of employees which are older than the median value, i.e. older than 40 years. If we expect that their yearly gross salary on average is higher than the median of the whole population, we expect a positive correlation between age and salary.

We can write this in the form of a question as:
Suppose we take the 10,000 employees older than 40 years (number of points in D+B). How many of them do you expect to have a higher salary than 45,000 euros (number of points in B)?

The answer to this question would be approximately $\frac{B}{D+B}$. Since the dependence relations are not always clear, therefore answering to this question

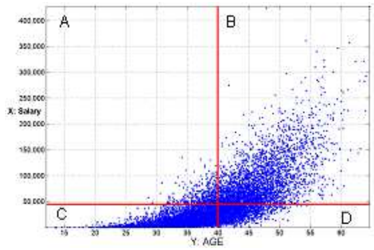

Figure C.1: Hypothetical joint distribution of employee age and salary (original units). Rank correlation = 0.8

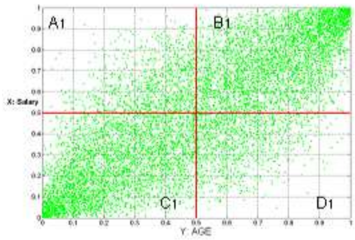

Figure C.2: Hypothetical joint distribution of employee age and salary (ranks). Rank correlation = 0.8

only by looking at picture (C.1) would be rather difficult. Instead of taking figure (C.1) we can plot the rank of each employee's age and salary in its distribution (C.2). We notice there the positive rank correlation since all samples lay along the diagonal. To distinguish how hight the positive correlation is we look at the width of the spread of the diagonal. The example of perfect negative and positive correlations are shown in figures (C.3) and (C.4) respectively. Rank correlation can be understood roughly as a degree to which two random variables take high or low values together. Negative rank correlation, as a degree to which one variable takes high values when the other takes low values. More examples are available in Appendix.

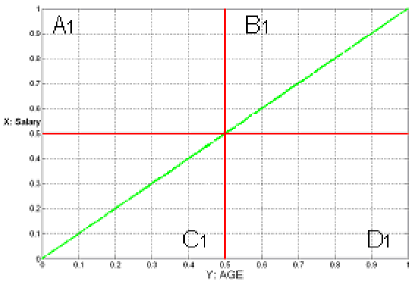

Figure C.3: Perfect positive dependence. Rank correlation = 1.0

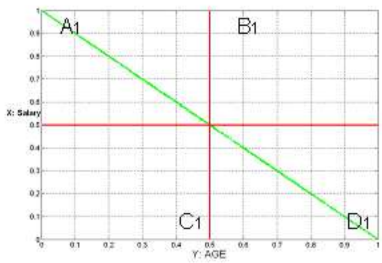

Figure C.4: Perfect negative dependence. Rank correlation = -1.0

By looking at the figure (C.2), if you think that in B1 there are 7,000 points then your answer to our question would be:

$$P(Salary \geq \; x_{50} = \; 45,000| \; Age \; \geq \; y_{50} = \; 40) \; \approx \frac{B}{B+D} = \; 0.7$$

In the next step we would ask about:
From our sample file look at the population of employees older than 40 years and which have worked in more than 5 projects worth more than 150,000 euros. How many of them do you expect to have a higher salary than 45,000 euros?

From the previous answer we have that 7,000 out 10,000 have a salary higher that 45,000 euros. We are wondering now how this number would change considering that all 10,000 employees have also worked in more than 5 projects worth more than 150,000 euros. In fact we are asking for an estimate of:

$$P(Salary \geq\ x_{50} =\ 45,000|\ Age\ \geq\ y_{50} =\ 40\ and\ \sharp Projects\ \geq\ z_{50} =\ 5)$$

Confidentiality

The names and quantification of the experts will not be published in the final report, however the individual expert assessment and all information relevant will. The names of the experts will not be directly identified with individual assessments.

Elicitation

In this questionnaire, it is assumed that the population considered refers to aircraft mechanics in western countries. In case of airplanes we consider the population of Western-built large aircrafts (>5,700 kg Maximum Take-off Weight) currently flying in commercials operations worldwide. All questions (unless explicitly stated) will be based on this populations.

We would like to obtain your view on the maintenance human error, in order to obtain insight into the impact of individual factors on the overall risk of maintenance incident/accident. To capture your uncertainty, in all questions from this section, we will ask you to provide the 5%, 50% and 95% quantiles for the following uncertain distribution.

Marginal Distribution

Dependence Information

Table C.1 presents the nodes considered in the Maintenance Human Error model together with their formal definitions and source for data for marginal distribution.

Q1	Human Error
Consider 10,000 maintenance tasks taken randomly from the population. On how many of these tasks would you expect a human error would occur that might lead to a hazardous situation?	

______	______	______
5%	50%	95%

Q2	Working Condition
Consider 10,000 maintenance tasks taken randomly from the population. How many of these tasks (percentage of them) would have to be performed out-side the hangar?	

______	______	______
5%	50%	95%

Q3	Comunication
Consider 100 reports which need to be pass to the next shift. How many of these messages/reports would be pass to the next shift only through the paper notes (i.e. workcards)?	

______	______	______
5%	50%	95%

Q4	Job Trainings
What is the average number of trainings/courses that a person in your position would receive per one year?	

______	______	______
5%	50%	95%

Node	Definition	Unit	Label	♯	Source for marginal distribution
Human Error	Number of maintenance human errors that might lead to a hazardous situation per 10,000 maintenance tasks	Number of errors	HE	7	Data / Expert Judgment
Job training	Average number of training in one year	Number of training	JobT	1	Expert Judgment
Experience	Number of year working as aircraft mechanic	Number of years	Exp	2	BL Statistics
Alertness	Average number of hours slept by the mechanic per day	Number of hours	Alert	3	Data
Communication	Current information transfer procedure in use, distinguishing: 1. only paper notes, 2. paper notes with oral feedback.	1 – 2	Com	4	Expert Judgment
Working Conditions	Average number of maintenance operations which needs to be performed 1.out-side / 2.inside the hangar per 10,000 maintenance operations	1 – 2	WCond	5	Expert Judgment
Aircraft Generation	Aircraft generation in scale from 1 to 4 where 4 is the most recent generation of aircrafts	1 – 4	AG	6	Schiphol Data

Table C.1: Variables used in Maintenance Human Error model.

<table>
<tr><td>Q5</td><td>$P_1 = P(Exp > 9.4 \mid JobT > JobT_{50})$</td></tr>
<tr><td colspan="2">Suppose that out of 20,000 aircraft mechanics, 10,000 were selected for which the average number of training per 1 year is above the median value.

What percent of these 10,000 mechanics will have their years of experience above the median value (≈ 9.4 years) ?</td></tr>
<tr><td>P_1</td><td>$r_{4,1}$</td></tr>
<tr><td>______________</td><td>______________</td></tr>
</table>

<table>
<tr><td>Q6</td><td>$P_1 = P(HE > HE_{50} \mid Exp > 9.4)$</td></tr>
<tr><td colspan="2">Suppose that out of 640,000 aircraft mechanics 320,000 were selected that have been working more that the median number of years (≈9.4 years).

What percent of these 320,000 mechanics will commit more than the median number of errors per 10,000 maintenance tasks?</td></tr>
<tr><td>P_1</td><td>$r_{7,4}$</td></tr>
<tr><td>______________</td><td>______________</td></tr>
</table>

<table>
<tr><td>Q7</td><td>$P_2 = P(HE > HE_{50} \mid Exp > 9.4,\ Alert > Alert_{50})$</td></tr>
<tr><td colspan="2">Suppose that out of above mentioned 320,000 aircraft mechanics that have been working more that the median number of years (≈9.4 years), 160,000 mechanics are selected for which their alertness is above the median value (≈5 hours).

What percent of these 160,000 mechanics will commit more than the median number of errors per 10,000 maintenance tasks?

The remaining rank correlation should be at most __________. What percentage of this “remaining dependence” should be explained by the variable Alertness?</td></tr>
<tr><td>P_2</td><td>$r_{7,2|4}$</td></tr>
<tr><td>______________</td><td>______________</td></tr>
</table>

<table>
<tr><td>Q8</td><td>$P_3 = P(HE > HE_{50} \mid Exp > 9.4,\ Alert > Alert_{50},\ Com > Com_{50})$</td></tr>
<tr><td colspan="2">Suppose that out of above mentioned 160,000 aircraft mechanics that have been working more that the median number of years (≈9.4 years), and their alertness is above the median value (≈5 hours), 80,000 are selected that have obtained messages/reports from the previous shift by paper notes together with oral feedback.

What percent of these 80,000 mechanics will commit more than the median number of errors per 10,000 maintenance tasks?

The remaining rank correlation should be at most ___________. What percentage of this “remaining dependence” should be explained by the variable Communication?</td></tr>
<tr><td>P_3</td><td>$r_{7,3|2,4}$</td></tr>
<tr><td>_____________</td><td>_____________</td></tr>
</table>

<table>
<tr><td>Q9</td><td>$P_4 = P(HE > HE_{50} \mid Exp > 9.4,\ Alert > Alert_{50},\ Com > Com_{50},\ JobT > JobT_{50})$</td></tr>
<tr><td colspan="2">Suppose that out of above mentioned 80,000 aircraft mechanics that have been working more that the median number of years (≈9.4 years), and their alertness is above the median value (≈5 hours), and have obtained messages/reports from the previous shift by paper notes together with oral feedback, 40,000 mechanics are selected for which number of job training per one year is above the median value (♯ of training).

What percent of these 40,000 mechanics will commit more than the median number of errors per 10,000 maintenance tasks?

The remaining rank correlation should be at most ___________. What percentage of this “remaining dependence” should be explained by the variable Job Training?</td></tr>
<tr><td>P_4</td><td>$r_{7,1|2,3,4}$</td></tr>
<tr><td>_____________</td><td>_____________</td></tr>
</table>

<table>
<tr><td>Q10</td><td>$P_5 = P(HE > HE_{50} \mid Exp > 9.4,\ Alert > Alert_{50},\ Com > Com_{50},\ JobT > JobT_{50},\ WCond = inside)$</td></tr>
<tr><td colspan="2">Suppose that out of above mentioned 40,000 aircraft mechanics that have been working more that the median number of years (≈9 years), and their alertness is above the median value (≈5 hours), and have obtained messages/reports from the previous shift by paper notes together with oral feedback, and for which number of job training per one year is above the median value, 20,000 mechanics are selected that are working in-side the hangar.

What percent of these 20,000 mechanics will commit more than the median number of errors per 10,000 maintenance tasks?

The remaining rank correlation should be at most ___________. What percentage of this "remaining dependence" should be explained by the variable Working Condition?</td></tr>
<tr><td>P_5</td><td>$r_{7,5|1,2,3,4}$</td></tr>
<tr><td>_____________</td><td>_____________</td></tr>
</table>

<table>
<tr><td>Q11</td><td>$P_6 = P(HE > HE_{50} \mid Exp > 9.4,\ Alert > Alert_{50},\ Com > Com_{50},\ JobT > JobT_{50},\ WCond = inside,\ AG \geq 3)$</td></tr>
<tr><td colspan="2">Suppose that out of above mentioned 20,000 aircraft mechanics that have been working more that the median number of years (≈9 years), and their alertness is above the median value (≈5 hours), and have obtained messages/reports from the previous shift by paper notes together with oral feedback, and for which number of job training per one year is above the median value, and are working in-side the hangar, 10,000 mechanics are selected that works with latest (3, 4) generation of aircraft.

What percent of these 10,000 mechanics will commit more than the median number of errors per 10,000 maintenance tasks?

The remaining rank correlation should be at most ___________. What percentage of this "remaining dependence" should be explained by the variable Aircraft Generation?</td></tr>
<tr><td>P_6</td><td>$r_{7,6|1,2,3,4,5}$</td></tr>
<tr><td>_____________</td><td>_____________</td></tr>
</table>