Towards Uncertainty Analysis of Bayesian Networks

Linda C. van der Gaag, Remco Kuijper, Yuri M. van Geffen, Jordi L. Vermeulen

Department of Information and Computing Sciences, Faculty of Sciences,
Utrecht University, Utrecht, The Netherlands

Abstract

To study the effects of inaccuracies in the parameter probabilities of a Bayesian network, often a sensitivity analysis is performed. In such an analysis, one or more parameter probabilities are varied systematically, by means of which their functional relationship with an output probability of interest is established. For reasons of computational complexity and difficulty of interpretation, sensitivity analysis of a Bayesian network is restricted to a single parameter, or to two parameter probabilities at most. From the results of such restricted analyses however, it is not easily predicted how inaccuracies in multiple parameter probabilities will interact and jointly affect the output probability of interest. Another general technique for investigating the effects of parameter inaccuracy is to perform an uncertainty analysis. Taking a sampling approach, this technique is less prone to computational problems than sensitivity analysis is. There being little experience as yet with uncertainty analysis of Bayesian networks, we re-consider this technique for studying the effects of inaccuracies in a network's parameter probabilities and provide some insights for the interpretation of the results obtained.

1 Introduction

Bayesian networks are widely accepted in artificial-intelligence research as intuitively appealing, valuable representations of uncertainty, and are nowadays being realised in a range of application domains. Experience shows however, that networks of realistic size are not always easily developed. Constructing a network with the help of domain experts can be especially difficult and costly: where building the graphical structure is considered doable in general, obtaining assessments for all parameter probabilities involved tends to make Bayesian-network construction a daunting engineering task [5]. Since directly available probabilistic information, from literature or from data, often is not amenable to encoding in a network's parameters, the majority of required probabilities need typically be assessed by experts. From cognitive-science research the problems encountered when eliciting probabilities from people in general are widely known [9]; also the more tailored setting of eliciting probabilities for a real-world Bayesian network from its domain experts tends to reveal the well-known problems of bias and poor calibration [6]. As a consequence of the difficulty of obtaining well-calibrated assessments, most parameter probabilities for a Bayesian network will include at least some inaccuracies.

Before a Bayesian network can be employed for problem solving in the real-world setting for which it was developed, the robustness of its output need be established, that is, the extent to which the inaccuracies in its parameter probabilities can affect an output probability of interest. Henrion and his co-workers were among the first researchers to investigate the robustness of a network's output in view of parameter inaccuracy [8, 12]. They performed a range of experiments in which they studied the effects of adding random noise to all parameter probabilities. The experiments were run on a large number of problem cases for three networks of relatively simple topology, and the effects on diagnostic performance were studied. The experimental findings from this particular setting strongly suggested that the output of a Bayesian network is quite insensitive to the inaccuracies in its parameter probabilities.

After the first results from investigating output robustness were published, researchers continued to study the effects of parameter variation on the probabilities of interest computed from a Bayesian

network. The research focused almost exclusively on a systematic approach in which a single parameter probability is varied in a step-wise fashion and in which a shift in the parameter's original value is related explicitly to a shift in the computed output probability; this approach to studying the effects of parameter variation is known as (*one-way*) *sensitivity analysis* of Bayesian networks. The research efforts resulted in important insights in output robustness of Bayesian networks and in a range of practical computational methods. Sensitivity analyses of real-world networks have moreover revealed by now that the output of a Bayesian network *can* show very high sensitivities to the values of the parameter probabilities involved.

While research so far focused on methods for one-way sensitivity analysis of Bayesian networks, the insights gained and methods developed can in essence be generalised to analyses in which multiple parameter probabilities are varied simultaneously. Such higher-order analyses would be particularly useful for uncovering interaction effects among the parameters under study. Straightforward extension of the existing methods for one-way sensitivity analysis however, would result in a technique with a highly impractical runtime. The results of higher-order sensitivity analyses moreover are known to be particularly hard to visualise and interpret, even for small numbers of parameters [7].

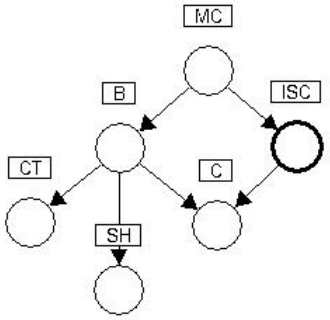
For mathematical models in general, also another approach has been proposed for analysing output robustness in view of parameter inaccuracy, which is called *uncertainty analysis*. Rather than a systematic approach to parameter variation, uncertainty analysis takes a sampling approach in which the values for a model's parameters are drawn from a pre-specified distribution. While a sensitivity analysis serves to detail the functional relationship between one or more model parameters and the output of interest, an uncertainty analysis reveals just the distribution of output values without any explicit reference to parameter values. Performing an uncertainty analysis involves high computational costs if tight error bounds need to be guaranteed for the established distribution of output values. Given its sampling nature however, uncertainty analysis has an any-time property, and sufficiently insightful results can be obtained with a reasonable number of iterations. The research by Henrion and his co-workers referenced above, actually built upon this sampling approach for studying output robustness of Bayesian networks.

In this paper, we re-consider uncertainty analysis as a technique for studying the effects of inaccuracies in the parameter probabilities of a Bayesian network, and address the interpretation of its results more specifically. For a small example network, we compare to this end the results obtained from one-way and two-way sensitivity analyses with those from matching uncertainty analyses for a restricted number of parameter probabilities. The paper is organised as follows. In Section 2, we briefly review Bayesian networks and introduce our running example. In Section 3, we illustrate one-way and two-way sensitivity analysis for the example network. Section 4 introduces uncertainty analysis of Bayesian networks in general. In Section 5 we re-visit our example network and compare the results from the different types of analysis. The paper ends in Section 6 with our concluding observations.

2 Bayesian networks

A Bayesian network is a representation of a joint probability distribution over a set of random variables, capturing domain knowledge along with the uncertainties involved. It consists of a graphical part and an associated numerical part. The graphical part takes the form of an acyclic directed graph, or digraph for short. Each node in this digraph represents a domain variable that takes its value from a finite set of discrete values. For ease of exposition, we restrict the discussion in this paper to binary variables, taking one of the values *true* and *false*. If a variable V has the value *true*, we write v; the notation $\bar{v}$ is used to indicate that $V = false$. When referring to either of the two values for V, we write v'. The arcs in the digraph represent influential relationships among the represented variables; absence of an arc between two variables means that these variables do not influence each other directly. More formally, the set of arcs of the digraph has an interpretation of independence through the d-separation criterion [11]. The strengths of the relationships between the variables are described by probability distributions: for each variable V, conditional distributions $p(V \mid \pi(V))$ over its values are specified, conditioned on the possible value combinations for its set of parents $\pi(V)$ in the digraph. The probabilities defining these distributions are referred to as the network's *parameter probabilities*. A Bayesian network allows the computation of any probability of interest over its variables [10, 11]. In the sequel, we will explicitly distinguish computed probabilities, written as $\Pr$, from parameter probabilities, denoted by p.

In this paper, we consider the *Brain tumour* network from Figure 1 for our running example. This small Bayesian network captures some (fictitious and incomplete) medical knowledge, adapted from



$p(mc) = 0.20$

$p(b \mid mc) = 0.20$
$p(b \mid \overline{mc}) = 0.05$

$p(c \mid b, isc) = 0.80$
$p(c \mid \overline{b}, isc) = 0.80$
$p(c \mid b, \overline{isc}) = 0.80$
$p(c \mid \overline{b}, \overline{isc}) = 0.05$

$p(isc \mid mc) = 0.80$
$p(isc \mid \overline{mc}) = 0.20$

$p(ct \mid b) = 0.95$
$p(ct \mid \overline{b}) = 0.10$

$p(sh \mid b) = 0.80$
$p(sh \mid \overline{b}) = 0.60$

Figure 1: The *Brain tumour* network.

[2]. It describes the medical problems associated with metastatic cancer for an arbitrary patient in oncology. Metastatic cancer (modelled by the variable *MC*) may lead to the development of a brain tumour (*B*) and may give rise to an increased level of serum calcium (*ISC*). The presence of a brain tumour can be established from a CT scan (*CT*); severe headaches (*SH*) may also be indicative of the presence of a brain tumour. A brain tumour or an increased level of serum calcium are both likely to cause a patient to fall into a coma (*C*). The digraph modelling the relationships among the six variables is shown on the left of Figure 1; the parameter probabilities associated with the digraph are shown on the right. The probabilities specified for the variable ISC for example, express that knowing whether or not metastatic cancer is present has a considerable influence on the probability of finding an increased level of serum calcium in an arbitrary patient. On the other hand, severe headaches are expressed as being quite common in both patients with and without a brain tumour.

3 Sensitivity analysis of Bayesian networks

Sensitivity analysis is a general technique for investigating the effects of varying the parameters of a mathematical model on the model's output. The analysis amounts to systematically varying the values of one or more parameters and recording the output for each value combination. Different types of sensitivity analysis are distinguished, dependent of the number of parameters that are varied in the analysis. The most common type of sensitivity analysis performed in practice is a one-way sensitivity analysis in which a single parameter is varied. Two-way analyses are often restricted to pairs of parameters of specific interest, for example because strong interaction effects on the output are expected. The restriction to a limited number of parameter pairs has its origin in the computational burden involved. Higher-order sensitivity analyses are hardly ever conducted in practice, not just because of their impractically high runtime, but also because their results are particularly hard to visualise and interpret [7].

When applied to a Bayesian network, sensitivity analysis entails studying the effects of varying the network's parameter probabilities on the computed output of interest. Research efforts have so far focused on the effects of varying a single parameter probability. Such a one-way analysis amounts to establishing, for a specific output probability of interest, the function that expresses this probability in terms of the parameter being varied [3]. The resulting sensitivity function has a highly constrained form: it is a quotient of two functions that are linear in the parameter probability under study [1, 4]. More formally, we consider a probability of interest $\Pr(a' \mid e)$, where a' is a specific value of a variable A of interest and e denotes the available (possibly compound) evidence, and a parameter probability $x = p(b' \mid \pi)$, where b' is a value of a variable B and π is a combination of values for the parents of B. The sensitivity function $f_{\Pr(a'|e)}(x)$ that expresses $\Pr(a' \mid e)$ in x has the following general form:

$$f_{\Pr(a'|e)}(x) = \frac{c_1 \cdot x + c_2}{c_3 \cdot x + c_4}$$

where the constants c_i are built from the values of the network's non-varied parameters. The numerator of the quotient expresses the joint probability $\Pr(a', e)$ as a function of x; its denominator describes $\Pr(e)$ in terms of x. From the property that the joint probability distribution $\Pr$ factorises in the network's parameter distributions and building upon the property of marginalisation, it follows that both $\Pr(a', e)$ and $\Pr(e)$ can be written as a sum of products of parameter probabilities, one of which is x.

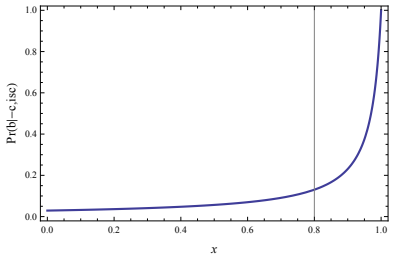


Figure 2: The sensitivity function describing the effect of varying the value of the parameter probability $x = p(c \mid \overline{b}, isc)$ from 0 to 1, on the output probability $\Pr(b \mid \overline{c}, isc)$ of the *Brain tumour* network.

We illustrate the general form of a one-way sensitivity function by studying, for the *Brain tumour* network, the function that describes the effects of varying the parameter probability $x = p(c \mid \overline{b}, isc)$ on the output probability $\Pr(b \mid \overline{c}, isc)$. The sensitivity function $f_{\Pr(b|\overline{c},isc)}(x)$ is established as

$$f_{\Pr(b|\overline{c}\,isc)}(x) = \frac{-0.03}{x - 1.03}$$

and is plotted in Figure 2. From the figure we read that varying the parameter x to smaller values than the originally specified one (indicated in the figure by a vertical line) has very little effect on the output probability and that variation to larger values can result in a considerable increase in the probability of interest. To explain this finding, we recall that the parameter x pertains to the probability of falling into a coma for a patient who does not have a brain tumour but does have an increased level of serum calcium. Now, if this parameter is increased to a higher value, then in a patient with an increased level of serum calcium without having fallen into a coma, the presence of a brain tumour must become more likely.

In a two-way sensitivity analysis of a Bayesian network, two parameter probabilities are varied simultaneously. For a specific output probability of interest, the analysis amounts to establishing the function that expresses this probability in terms of the two parameters under study [3]. The resulting sensitivity function again has a highly constrained functional form, and in fact is a quotient of two bi-linear functions, that is, of two functions that are linear in each parameter separately [1, 4]. More formally, we consider the sensitivity function $f_{\Pr(a'|e)}(x, y)$ that expresses the probability of interest $\Pr(a' \mid e)$ with a' and e as before, as a function of the parameter probabilities $x = p(b' \mid \pi_B)$ and $y = p(d' \mid \pi_D)$, where b' and d' are values of the variables B and D, and π_B and π_D are value combinations for the parents of B and of D respectively. The function has the following general form:

$$f_{\Pr(a'|e)}(x, y) = \frac{c_1 \cdot x \cdot y + c_2 \cdot x + c_3 \cdot y + c_4}{c_5 \cdot x \cdot y + c_6 \cdot x + c_7 \cdot y + c_8}$$

where the constants c_i are again built from the values of the network's non-varied parameters.

We illustrate the general form of a two-way sensitivity function by considering again our example *Brain tumour* network. The sensitivity function $f_{\Pr(c)}(x, y)$ that describes the effects of simultaneous variation of the parameter probabilities $x = p(b \mid mc)$ and $y = p(isc \mid mc)$ on the prior output probability $\Pr(c)$ of an arbitrary patient in oncology falling into a coma, is established to be

$$f_{\Pr(c)}(x, y) = 0.374 + 0.15 \cdot x \cdot y - 0.15 \cdot x - 0.15 \cdot y$$

We would like to note that since this sensitivity function pertains to a prior probability of interest, and hence lacks a term referring to evidence, it is a bi-linear function rather than a quotient of two such functions. The function is plotted in three dimensions in Figure 3 on the left. The plot on the right of the figure also shows the function, this time in two dimensions by contour lines. These contour lines connect the combinations of values for the two parameter probabilities under study that result in the same value for the output probability $\Pr(c)$. The distance between two contour lines indicates the variation necessary in the two parameters to shift the probability of interest from one contour line to another. If the contour lines are very close to one another therefore, a small variation in the parameter probabilities suffices to have a strong effect on the output probability; if, in contrast, the contour lines are further apart, then the probability of interest is not very sensitive to variation of the two parameters. In the two-dimensional plot of the sensitivity function, the distances between the contour lines differ, which

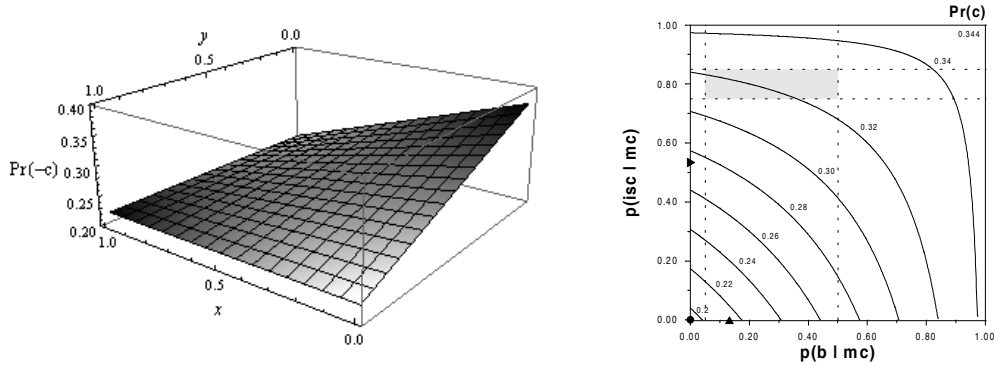


Figure 3: The two-way sensitivity function describing the effect of varying the parameter probabilities $x = p(b \mid mc)$ and $y = p(isc \mid mc)$ on the prior output probability $\Pr(c)$ of the *Brain tumour* network; the function is shown in three dimensions (*left*) and in two dimensions by contour lines (*right*).

indicates that varying the values of the parameter probabilities simultaneously has an interaction effect on the probability of interest in addition to the effects of their separate variation; this interaction effect originates from the synergistic joint influence of the variables B and ISC on the variable C described by the set of probability distributions $\Pr(C \mid B, ISC)$. We observe that the contour lines are closer to one another in the lower left part of the plot than in the upper right part. Varying the two parameter probabilities $p(b \mid mc)$ and $p(isc \mid mc)$ to quite small values therefore, will have a stronger effect on the probability of interest than varying them to the higher value range. We note that these observations are confirmed by the three-dimensional plot on the left of Figure 3.

4 Uncertainty analysis

Robustness of the output of a mathematical model in general can be studied by performing several one-way and two-way sensitivity analyses, yet can also be investigated by conducting an uncertainty analysis. Both techniques allow an investigation of the effects of varying a model's parameters on its output. While the result of a sensitivity analysis details the functional relationship between one or more parameters and the output of interest, the result of an uncertainty analysis shows the distribution of output values without any explicit reference to parameter values. The basic idea of an uncertainty analysis is taking a sampling approach to studying robustness. For this purpose, each parameter of the model under study is associated with a sampling distribution that describes the parameter's plausible values; most commonly, Gaussian distributions are used for these sampling distributions. In each sampling iteration, for every parameter, a value is drawn from its associated distribution. The result of a single iteration is an instantiation of the model under study, from which the output of interest is computed. By recording the computed values from multiple iterations, a distribution over the output is obtained.

When applying uncertainty analysis to a Bayesian network, the network's parameter probabilities are associated with sampling distributions. With each parameter probability x, we associate to this end a Gaussian distribution having the original value of x for its mean μ_x; the variance σ_x of the distribution is chosen so as to define the plausible range of values for x. We will presently return to the choice of Gaussian distributions for the analysis. Now, in each sampling iteration, for every parameter probability, a value is drawn from the associated distribution. The parameter values obtained are entered into the conditional probability distributions for the Bayesian network under study. From the thus instantiated network, the output probability of interest is computed. By recording the computed output values over multiple iterations, a distribution for the output probability of interest is obtained. Figure 4 shows the result of an uncertainty analysis of the *Brain tumour* network for the output probability $\Pr(b \mid \overline{c}, isc)$; for all parameter probabilities, a Gaussian distribution with $\sigma = 0.1$ was used.

For uncertainty analysis in general, most commonly Gaussian distributions are used for describing the ranges of plausible values for a model's parameters. For Bayesian networks however, using Gaussian distributions for sampling is not without problems. From a Gaussian distribution for a parameter probability with an original value very close to 0 for example, many values smaller than 0 will be drawn. Simply discarding these as impossible values for the parameter under study, will result in a bias

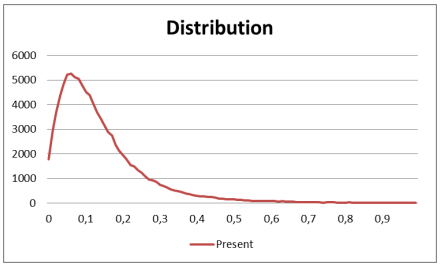


Figure 4: The distribution of values for the output probability $\Pr(b \mid \overline{c}, isc)$ resulting from uncertainty analysis of the *Brain tumour* network, using Gaussian distributions with $\sigma = 0.1$ for all parameters.

favouring the higher parameter values, which in turn may bias the distribution of values for the output probability of interest resulting from the analysis. Henrion and his co-workers suggest for this reason that the variance be added to the log-odds form after transforming the value of a parameter probability. The log-odds transform can result in a bimodal distribution with peaks at 0 and 1 however, and may therefore result in a sampling distribution that no longer has the intended meaning for the parameter probability under study. Another option is to use a symmetrically truncated Gaussian distribution, which by definition assigns equal weight to values smaller than and values larger than the chosen mean. Using such truncated sampling distributions for just the parameter probabilities with extreme original values however, may equally introduce a bias in the results of the analysis since for these parameters a necessarily small variance is assumed. In our experiments with the *Brain tumour* network, using the log-odds transform and using symmetrically truncated Gaussian distributions for sampling yielded quite similar results upon uncertainty analysis. Since the example network does not include any truely extreme probabilities, this finding may not hold for real-world Bayesian networks in general however.

5 Interpreting the results from an uncertainty analysis

While research in sensitivity analysis of Bayesian networks has yielded important fundamental insights for relating the results from such an analysis to output robustness, there is little experience as yet with performing uncertainty analyses of Bayesian networks and, hence, hardly any experience with interpreting the results obtained. To gain some insight in the interpretation of the output distributions from an uncertainty analysis in terms of exhibited sensitivities, we performed uncertainty analyses of the *Brain tumour* network matching the sensitivity analyses from Section 3, and compared the results obtained.

We begin by studying, for our example network, the effect of varying the single parameter probability $x = p(c \mid \overline{b}, isc)$ on the output probability $\Pr(b \mid \overline{c}, isc)$. We recall that Figure 2 plots the one-way sensitivity function which expresses this output in the parameter under study. The function shows that variation of x to smaller values than its originally specified value 0.8, has little effect on the output probability; variation to higher parameter values on the other hand, has a relatively strong effect on the output, as evidenced by the function's increasingly growing derivative for the higher-value range. For the parameter x and the output probability of interest, we now also performed an uncertainty analysis. For this purpose, we associated with the parameter x a (symmetrically truncated) Gaussian distribution with 0.8 for its mean μ_x and with a variance equal to $\sigma_x = 0.2$. The result of this analysis is depicted in Figure 5 on the left. From the figure we read that the output distribution has a somewhat larger variance to the right, that is, for output values larger than the mode. This finding is readily explained by studying the results of the matching one-way sensitivity analysis. The larger variance to the right of the mode from the uncertainty analysis indicates that varying the parameter under study in a particular direction (in this example: to the range of higher parameter values) will give higher output values with a more substantial variance than varying in the other direction gives smaller output values. Informally spoken, the more gradual slope of the output distribution to the right of the mode indicates that the output probability of interest is sensitive to inaccuracies in the parameter under study, and may thus actually have a higher value than the originally established one. To further support our explanation, we conducted another uncertainty analysis for the same parameter and output probabilities, this time using a sampling distribution with a variance equal to 0.1. Figure 5 shows the resulting output distribution on the right. We observe that the difference in variance to the left and to the right of the distribution's mode is now not

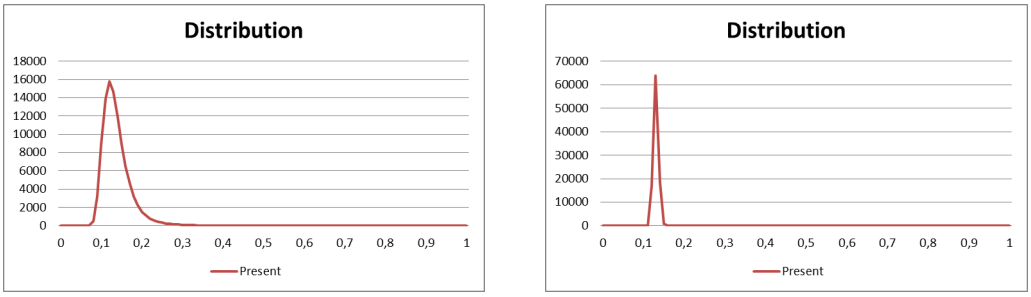


Figure 5: The distribution of values of the output probability $\Pr(b \mid \overline{c}, isc)$ resulting from an uncertainty analysis of the *Brain tumour* network for the parameter $x = p(c \mid \overline{b}, isc)$, using a Gaussian distribution with $\mu_x = 0.8$ and with $\sigma_x = 0.2$ (*left*) and $\sigma_x = 0.1$ (*right*) respectively.

as marked as in the output distribution obtained from using a variance of $\sigma_x = 0.2$. From the matching sensitivity function, we observe that the high sensitivity of the output to the parameter probability under study is quite prominent only for parameter values higher than 0.9. Upon uncertainty analysis using a sampling distribution with $\mu_x = 0.8$ and $\sigma_x = 0.1$ therefore, few samples will be drawn from this range of very high parameter values, thereby explaining the diminished effect on the output distribution.

We now further investigate, for our example network, the effect, on the prior probability $\Pr(c)$, of simultaneously varying the two parameter probabilities $x = p(b \mid mc)$ and $y = p(isc \mid mc)$. We recall that Figure 3 plots the two-way sensitivity function that results from systematic variation of these parameters. The function shows that varying the two probabilities has an interaction effect on the output of interest, albeit a weak one: varying both parameters to smaller values has a slightly increasing diminishing effect on the output. For the output of interest and the two parameters x and y, we now also performed an uncertainty analysis. To this end, we associated with both parameter probabilities a Gaussian distribution with the original values 0.2 and 0.8 for the means μ_x and μ_y respectively, and with variance $\sigma_x = \sigma_y = 0.2$. Figure 6 plots the resulting output distribution. The figure reveals a slightly larger variance for output values smaller than the mode of the distribution. This observation is readily explained from our earlier findings from the matching two-way sensitivity function. The larger variance of output values to the left of the mode indicates that varying the value combination for the two parameter probabilities in a particular direction within the parameter space (in this example: to the range of smaller values for both parameters) will give slightly decreasing output values.

The above considerations show that comparing the results from the two sensitivity analyses from Section 3 with those obtained from matching uncertainty analyses, yields quite consistent findings. We now consider again the uncertainty analysis of the *Brain tumour* network in which all parameters are varied simultaneously, the result of which was shown in Figure 4. We observe that the variance of the output distribution is markedly larger to the right of the distribution's mode than to the left. Based on our considerations above, we cautiously conclude that the output probability of interest shows quite some sensitivity to inaccuracies in the network's parameter probabilities. More specifically, the output probability of interest will adopt a considerably higher value than the originally established one

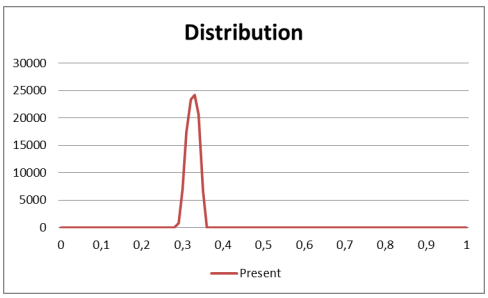


Figure 6: The distribution for the output probability $\Pr(c)$ resulting from an uncertainty analysis of our network for the parameters $p(b \mid mc)$ and $p(isc \mid mc)$, using Gaussian distributions with $\sigma = 0.2$.

if the combination of parameter values is varied in a specific direction in the joint parameter space. Unfortunately, the uncertainty analysis does not reveal in which direction this effect will arise.

6 Conclusions

Real-world application of a Bayesian network requires insight in the extent to which inaccuracies in its parameter probabilities can affect its output. While research in one-way sensitivity analysis of Bayesian networks has resulted in fundamental insights in the sensitivity of a network's output probabilities to the value of a single parameter, there is little insight as yet in the patterns of interaction exhibited by multiple parameters in their joint effect on an output probability of interest. In this paper we focused on uncertainty analysis as an alternative type of analysis that allows studying the effects of simultaneous variation of multiple parameter probabilities by taking a sampling approach. By means of experiments on a small fictitious network, we illustrated some insights in the interpretation of the results from such an uncertainty analysis. For a deeper understanding, further experimentation is required in addition to a more fundamental study, for example of the way in which properties of the sampling distribution influence the results obtained. With the considerations presented in this paper, we hope to initiate further study which will ultimately result in practicable uncertainty analysis of Bayesian networks.

Acknowledgments. We would like to thank Frank Dignum for facilitating the reported research.

References

[1] E. Castillo, J.M. Gutiérrez, A.S. Hadi (1997). Sensitivity analysis in discrete Bayesian networks. *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 27, pp. 412 – 423.

[2] G.F. Cooper (1984). *NESTOR: a Computer-based Medical Diagnostic Aid that Integrates Causal and Probabilistic Knowledge*. Report HPP-84-48, Stanford University.

[3] V.M.H. Coupé, L.C. van der Gaag (1998). Practicable sensitivity analysis of Bayesian belief networks. In: M. Hušková, P. Lachout, J.A. Víšek (editors). *Prague Stochastics '98*, Union of Czech Mathematicians and Physicists, pp. 81 – 86.

[4] V.M.H. Coupé, L.C. van der Gaag (2002). Properties of sensitivity analysis of Bayesian belief networks. *Annals of Mathematics and Artificial Intelligence*, vol. 36, pp. 323 – 356.

[5] M.J. Druzdzel, L.C. van der Gaag (2000). Building probabilistic networks: "Where do the numbers come from ?". *IEEE Transactions on Knowledge and Data Engineering*, vol. 12, pp. 481 – 486.

[6] L.C. van der Gaag, S. Renooij, H.J.M. Schijf, A.R. Elbers, W.L. Loeffen (2012). Experiences with eliciting probabilities from multiple experts. In: S. Greco, B. Bouchon-Meunier, G. Coletti, M. Fedrizzi, B. Matarazzo, R.R. Yager (editors). *Advances in Computational Intelligence*, CCIS 299, Springer, Berlin, pp. 151 – 160.

[7] J.D.F. Habbema, P.M.M. Bossuyt, D.W.J. Dippel (1990). Analysing clinical decision analyses. *Statistics in Medicine*, vol. 9, pp. 1229 – 1242.

[8] M. Henrion, M. Pradhan, B. Del Favero, K. Huang, P. O'Rorke (1996). Why is diagnosis using belief networks insensitive to imprecision in probabilities. In: E. Horvitz, F.V. Jensen (editors). *Twelfth Conference on Uncertainty in Artificial Intelligence*, Portland, pp. 307 – 314.

[9] D. Kahneman, P. Slovic, A. Tversky (1982). *Judgment under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge.

[10] S.L. Lauritzen, D.J. Spiegelhalter (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B*, vol. 50, pp. 157 – 224.

[11] J. Pearl (1988). *Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference*. Morgan Kaufmann, Palo Alto.

[12] M. Pradhan, M. Henrion, G. Provan, B. Del Favero, K. Huang (1996). The sensitivity of belief networks to imprecise probabilities: An experimental investigation. *Artificial Intelligence Journal*, vol. 84, pp. 357 – 397.