# Decision Making based on Sampled Disease Occurrence in Animal Herds

Michael Höhle[1,2] and Erik Jørgensen[2]

[1] Department of Animal Science and Animal Health, Royal Veterinary and Agricultural University, Grønnegårdsvej 3, 1870 Frb. C, Denmark
hoehle@dina.dk
[2] Department of Animal Breeding and Genetics, Danish Institute of Agricultural Sciences, Research Centre Foulum, PO Box 50, 8830 Tjele, Denmark
Erik.Jorgensen@agrsci.dk

**Abstract.** To make qualified decisions when extrapolating results from a survey sample with imprecise tests requires careful handling of uncertainty. Both the imprecise test and uncertainty introduced by the sampling have to be taken into account in order to act optimally. This paper formulates an influence diagram with discrete and continuous nodes to handle an example typical for animal production: a veterinarian who – as part of a biosecurity program – has to decide whether to treat a herd of animals after inspecting a small fraction of them.
Our aim is to investigate the robustness of the obtained strategy by performing a two-way sensitivity analysis with respect to the proportion of false positives and false negatives of the test. Output of the analysis is a treatment map illustrating how the chosen strategy varies according to variation in these proportions. The map helps to investigate whether a certain variation is acceptable or if the test procedure has to be standardized in order to reduce variation. Objective of the paper is to be an appetizer to work more with the issues raised in obtaining a practical solution.

## 1 Introduction

Traditional survey sampling as e.g. in [1] is concerned with establishing the proportion of individuals having a specific characteristic in a population. This is done by extrapolating results from a sample to the entire population. In the traditional case, investigation of each individual in the sample will reveal its true state, i.e. as either having the property or not. In many practical applications such precise answers are not available – the test is imprecise thus introducing both false negatives and false positives. An example from the veterinarian field is the use of a diagnostic test to determine the disease prevalence of a herd. The task of establishing the disease status of a herd is typical for biosecurity programs, e.g. for salmonella in pigs or Johne's disease in cattle [2, 3]. Similar examples are found in clinical decision making or when testing for GM-seeds in seed lots [4, 5]. Estimates on disease prevalence, $0 \leq p \leq 1$, need to take the

*sensitivity* and *specificity* of the diagnostic test into account, i.e. respectively the fractions of diseased and non-diseased cases correctly diagnosed by the test. In practical situations these fractions can be hard and resource demanding to establish for a test method. Even worse, they are also open for a great deal of variation. For example when different veterinarians have to determine herd prevalence of e.g. pneumonia or diarrhea in a section of slaughter pigs [4].

If the same test is performed in all cases, an investigation could be performed to establish the sensitivity and specificity $(Se, Sp)$ of the test procedure. With the uncertainty in $p$ due to sampling taken into account a biosecurity program could recommend the following treatment strategy: Treat all animals in the section at cost $C_D$ if $p$ is above some threshold $T$ and do nothing if $p \leq T$. Based on the true prevalence and treatment chosen at the current time stage a reward is given. The aim is to choose the threshold maximizing the expected reward. Even though we find the optimal $T$, and recommend the strategy to all veterinarians, we would not take into account the variability in $(Se, Sp)$ due to each veterinarian making his own subjective clinical diagnosis for every investigated individual. Assume the specific veterinarian has a true (but unknown) setup of $(Se + \delta, Sp + \epsilon)$. If he follows the threshold based on $(Se, Sp)$ he might not achieve the maximum expected utility because his uncertainty in $p$ is of a different magnitude and shape.

Current biosecurity programs, e.g. the voluntary herd status program against Johne's disease [2], operate with fixed point estimates on sensitivity and specificity of the diagnostic test. To assess the impact of the above variability a proper sensitivity analysis should therefore be an integral part of the modeling. Methods such as one-way and two-way sensitivity analysis, tornado, rainbow diagrams, etc., provide valuable insights about implementational robustness of an optimal strategy [6]. Another approach would be to quantify uncertainty on sensitivity and specificity by distributions [7, 8]. Our interest, however, is the decision analytic dimension of the problem: How does variability affect a biosecurity program that assume fixed point estimates on sensitivity and specificity? How large deviations are allowed before the recommended strategy is suboptimal.

The following sections will show how the above considerations boil down to performing a two-way sensitivity analysis for an influence diagram with both discrete and continuous nodes. How to perform analytically sensitivity analysis in Bayesian networks is already well established [9, 10] whereas the matter is more complicated in influence diagrams. Here, especially sequential decision problems quickly become intractable to handle [11, 10]. As the above treatment considerations only contain a single decision, analytical calculations are tractable up to certain herd sizes, e.g. using Maple [12]. Solution of the diagram can also be done numerically using Gibbs sampling, where sensitivity analysis becomes a matter of performing many point-wise evaluations. For small herds both analytic and numeric solutions can be applied to verify correctness, while the numeric approach is the only tractable method once herd size become large.

The structure of this article is as follows. Section 2 describes how the clinical treatment example can be formulated as an influence diagram. Hereafter,

Section 3 describes how to calculate expected utilities in this model in order to select the best decision alternative. Robustness of these decisions to variation of the diagnostic test sensitivity and specificity is illustrated in Section 4. Finally, a discussion of the obtained results is given.

## 2 Influence Diagram Formulation

This section introduces the notation used to describe the decision problem. Let the herd be of size $N$ from which a simple random sample of size $n$ is drawn. The aim of the investigation is to determine the proportion of diseased animals, i.e. $p = d/N$ with $d$ being the number of sick in the population. We assume that the true number of diseased individuals in the sample, $D^+$, is obtained by drawing a sample of size $n$ without replacement from the population. In this case $D^+$ follows the *hypergeometric distribution* with parameters $N, d, n$. If sampling is with replacement or an infinite population can be assumed, $D^+$ is a sample from the *binomial distribution* with herd prevalence $p \in [0, 1]$. Also, if $n$ is small compared to both $d$ and $N - d$ the binomial distribution is a good approximation to the hypergeometric distribution. Such approximations are necessary because computations with the hypergeometric distribution quickly become intractable [1]. In the following, only binomial sampling is considered. The number of test positives, $T^+$, is then given as a sum of two binomial distributions with fixed values of the sensitivity, $Se$, and specificity, $Sp$, of the diagnostic test as parameters. Note that our interest is in the fixed value $Se$ and $Sp$ situation; otherwise a natural way to quantify uncertainty on the two variables would be by e.g. a beta distribution as in [7, 8]. Figure 1 illustrates the above as a graphical model using notation from [13]. By specifying a graphical model we obtain a clear overview of the dependence structure of the variables. Furthermore, the decision part is easily specified using influence diagram notation for which software would exist to solve at least a discretized version of the problem.

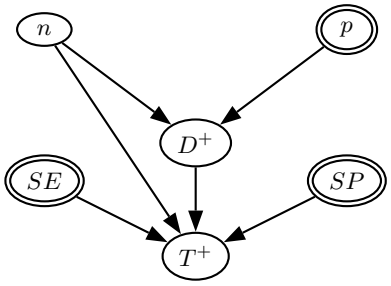


**Fig. 1.** Graphical model illustrating how the number of test positives, $T^+$, is obtained by sampling with replacement introducing both false positives and false negatives. Double lined nodes indicate continuous nodes, however, the $Se$ and $Sp$ distributions will be trivial in our application.

The above distributional explanations are expressed as

$$D^+ \sim \text{Bin}(n, p),$$
$$T^+ \sim \text{Bin}(D^+, Se) + \text{Bin}(n - D^+, 1 - Sp).$$

Inference for the herd prevalence can be formulated in the Bayesian context as follows. Given $\{n, T^+, Se, Sp\}$, what is the posterior distribution on $p$? A typical application would be to use this distribution to calculate a posterior mean for $p$ together with a credibility interval. This estimate could then be used by the veterinarian to determine whether a herd should be classified as disease free [7, 8].

Classical survey sampling would be concerned with how large to choose $n$ in order to get a certain confidence in $p$. Our focus is, however, on the application of the prevalence estimate, namely a decision to apply a treatment reducing prevalence. Going back the the herd context, a veterinarian typically has to decide between two decision alternatives: Either treat all animals in the herd, e.g. by adding antibiotics to the water supply, or do nothing. Whether to apply treatment is decided by the observed number of test positives. In order to decide which treatment to use, it is necessary to model how the disease prevalence will develop with time and how treatment influences it. A reward is given based on the disease prevalence which reflects the price of animals being sick. Figure 2 extends the graphical model from Fig. 1 with decision and utility nodes (see [14]) making it an influence diagram.

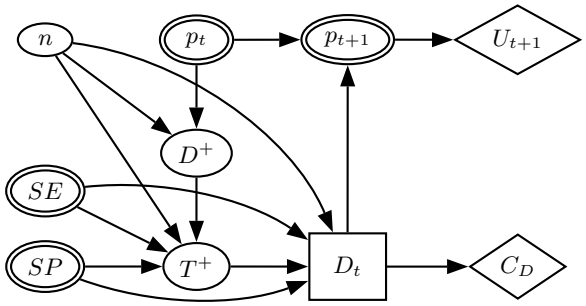


**Fig. 2.** Influence diagram describing the treatment strategy based on the number of animals tested positive.

Here, the $D_t$ node is the treat decision with states *treat all* (ta) and *do nothing* (dn). Furthermore, $p_{t+1}$ is the new prevalence[3], $C_D$ a utility node reflecting the cost of the treatment, and $U_{t+1}$ a utility node indicating the cost of disease as a function of the new prevalence. The transition probability between the two

[3] Basically, the situation could be handled without introducing a $p_{t+1}$ node by simply integrating the disease development into the utility function. But our choice is conceptual clearer.

prevalences is given by

$$P(p_{t+1}|p_t, D_t) = \begin{cases} k_3 p_t \text{ if } D_t = \text{ta} \\ p_t \quad \text{otherwise} \end{cases}$$

To illustrate the principle, simple proportional reduction in case of treating, i.e. $0 \leq k_3 \leq 1$ and preservation of status quo in case of not treating, is used. This ignores that an infectious disease would spread within the population if nothing is done. Modeling such a characteristic could although easily be done using e.g. a logistic model.

Economic preference is modeled with the two utility nodes $C_D$ and $U_{t+1}$. Typically, costs can be established on a per animal basis, which requires knowledge of the the number of animals, $N$, in the herd to make calculations realistic. A possible specification of the two utility functions could then be as follows.

$$\begin{aligned} C(D_t) &= -k_1 N \, I(D_t = \text{treat all}) \\ U_{t+1}(p_{t+1}) &= -k_2(p_{t+1} N), \end{aligned}$$

where $I$ is an indicator function.

To solve the decision scenario of Fig. 2 it is necessary to find the decision alternative for $D_t$, which given evidence $e = \{n, T^+, Se, Sp\}$, yields the highest expected utility. Because we are using a continuous representation of $p$, standard Bayesian Network software for solving the influence diagram of Fig. 2 is not directly applicable. Instead both an analytic solution method in Maple [12] and a simulation based using WinBugs [15] are investigated. For small herds the analytic approach is doable and allows us to verify how good an approximation the sampling approach is in this situation. Advantage of the analytic implementation is also that we can use the capabilities of Maple when performing sensitivity analysis.

## 3 Derivation of the expected utility

In order to calculate the required expected utility given $e = \{n, T^+, Se, Sp\}$ we need to calculate the posterior distribution $P(p_{t+1}|e)$, which again requires calculation of $P(p_t|e)$. As already mentioned, only the binomial case is considered. To calculate $P(p_t|e)$ we exploit the standard result, see e.g. [16], that

$$P(T^+ = x|\ldots) = \binom{n}{x} \Big[ \, pSe + (1-p)(1-Sp) \Big]^x \Big[ \, p(1-Se) + (1-p)Sp \Big]^{n-x}.$$

If expert information exist on the prevalence of the herd this is easily integrated using prior distributions. If nothing is known, a uniform prior distribution for $p$ is sufficient. Bayes Rule is exploited to obtain the posterior distribution

$$P(p|T^+, n, Se, Sp) \propto P(T^+|p, n, Se, Sp) P(p|n, Se, Sp).$$

To ensure that the above distribution is proper it is necessary to find an expression for the normalization constant $P(T^+|n, Se, Sp)$. Normally in a Bayesian

analysis proportionality of the posterior is sufficient, but, as $P(T^+|n, Se, Sp)$ depends on $Se$ and $Sp$, calculating it becomes a concern in the latter sensitivity analysis.

Continuing our calculations we observe that $p_{t+1}$ is just a functional transformation of $p_t$ when $D_t = ta$, i.e. we can use the standard rule for transformation of random variables to calculate the posterior $P(p_{t+1}|e, D_t = ta)$. If $D_t = dn$ no transformation is needed. Regarding $U_{t+1}$ as a random variable its distribution can be obtained in the same way as for $p_{t+1}$ by exploiting the above rule. Given an observed number of test positives, $T^+ = x$, the expected utility of the treat and no-treat alternatives can now be calculated as

$$\begin{aligned}\mathrm{EU}(D_t = ta) &= \mathrm{E}\left[U_{t+1}(p_{t+1}, D_t = ta)\right] + C_D(ta),\\ \mathrm{EU}(D_t = dn) &= \mathrm{E}\left[U_{t+1}(p_{t+1}, D_t = dn))\right] + C_D(dn).\end{aligned}$$

The above has been implemented in Maple yielding functions of $Se, Sp$. To evaluate the approximation of a simulation based approach the model was also formulated in WinBugs [15], which uses Gibbs sampling to calculate the expected utility. Figure 3 shows the posterior distribution of $U_{t+1}$ obtained from Gibbs Sampling (using 10,000 samples after a burn-in of 1,000) and the analytical distribution of $EU(D_t = ta)$ in a pseudo realistic setup of $Se = 0.8, Sp = 0.6, n = 5, T^+ = 2, k_3 = \frac{1}{2}, k_2 = -20, k_1 = -1, N = 100$. In the figure the analytical expected utility (obtained by integrating the density between the worst case -1100 and best case -100) is $-475.1$. The numeric mean (obtained as empirical mean of the samples) is $-481.2$. In the case $D_t = dn$ we obtain values of $-750.4$ and $-762.5$, respectively. Hence, in the chosen setup we decide to treat all animals. Note that the WinBugs approach is much easier to implement and solve than the analytic approach and appears to be a good approximation. However, it lacks the power of being able to describe the expected utility as function of sensitivity and specificity.

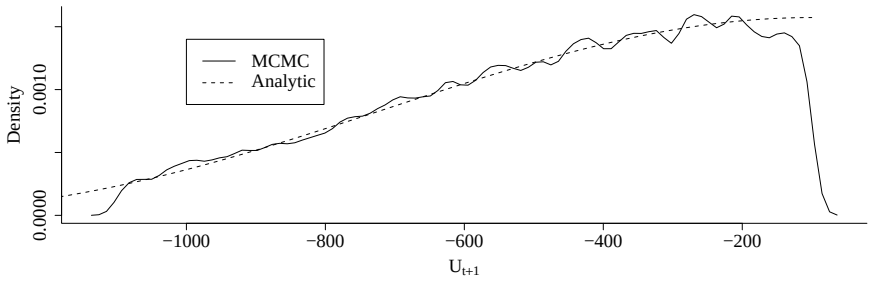


**Fig. 3.** Comparison of the posterior density $P(U_{t+1}|\ldots)$ calculated analytically and numerically; the x-axis is obtained utility while the y-axis is the corresponding density. The MCMC density is obtained by kernel smoothing the posterior samples obtained from WinBugs. The deviations at the end points are partly due to the kernel smoother and partly due to problems of the Gibbs sampler to investigate these areas.

The desired strategy for $D_t$ is now obtained by investigating the expected utility for both the $ta$ and $dn$ alternative for all $0 \leq T^+ \leq n$, Empirical investigations show that for this strategy there will exist a unique threshold $T$, s.t.

$$\arg\max_{d_t \in D_t} \mathrm{EU}(D_t = d_t | T^+ = x) = \begin{cases} dn \text{ if } 0 \leq x < T \\ ta \text{ if } T \leq x \leq n \end{cases}$$

That is, with the chosen specification of utilities and transitions, any strategy for $D_t$ can be compactly represented by the minimum number of test positives necessary before all animals will be treated. In the setup used in Fig 3 we obtain $T = 0$, i.e. we trivially treat no matter the number of observed test-positives.

Assuming $n, Sp$, and $Se$ to be fixed, the expected value of a strategy $s$ for $D_t$ is given as

$$\mathrm{EU}(s) = \sum_{x=0}^{n} P(T^+ = x | n, Se, Sp)\, \mathrm{EU}(D_t | T^+ = x, n, Se, Sp),$$

where $\mathrm{EU}_s(D_t|\ldots)$ denotes the expected utility obtained for $D_t$ when choosing the decision dictated by $s(x)$.

## 4 Sensitivity Analysis

In realistic situations, the sensitivity and specificity of the test are either unknown or subject to a great deal of variation. If we e.g. recommend a fixed threshold to all veterinarians investigating diarrhea in pig herds, the large variation in the two parameters between veterinarians would be ignored. A way to investigate a strategy's robustness towards variations in sensitivity and specificity is to find out how the best decision alternative changes with variation in $Se$ and $Sp$. Here, the analytical representation in Maple is of advantage because we immediately have the expected utility as a function of $Se$ and $Sp$. This is not possible using a simulation approach, instead the influence diagram would have to be solved for a grid of $Se$ and $Sp$ combinations.

Continuing with the values from the veterinarian example, but changing the sample size $n$ to 10 and increasing the price of a treatment to $k_1 = -2$, gives a more interesting example. Figure 4 shows the line of indifference, i.e. the solution of

$$f(Se, Sp) = \mathrm{EU}(D_t = ta | Se, Sp) - \mathrm{EU}(D_t = dn | Se, Sp) = 0.$$

To investigate the robustness of the decision using the sensitivity and specificity configuration $p = (Se', Sp')$ it might be worth to investigate how much $p$ can change before a different decision is made. This is equivalent to finding the distance to the intersection line, i.e.

$$r_c = \mathrm{dist}(p, l), \quad \text{where} \quad l = \{(Se, Sp) \mid f(Se, Sp) = 0\}$$

also known as the *radius of change* or radius of the *safe-ball*, see [11, 10]. The higher this radius the more robust the specific policy is against variations. Also,

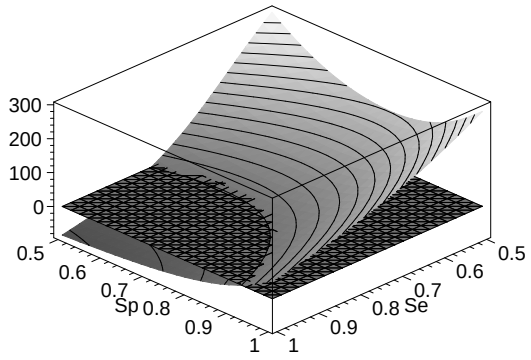


**Fig. 4.** Indifference between the two decision alternatives occurs on the line $f(Se, Sp) = 0$, The figure shows the intersection of $f(Se, Sp)$ with $z = 0$, i.e. the z-axis is the difference in utility between the two strategies. To the left of the intersection line, *dn* is selected, to the right *ta*.

the difference in expected utility between the two alternatives evaluated at specific points tells us about the benefit of getting the $(Se, Sp)$ correctly estimated.

To get a better overview of the variation in the strategy we can illustrate the obtained $T$-values as a function of $Se$ and $Sp$ – a two-dimensional analogue of a rainbow diagram. Figure 5 shows this *treatment map* in case 30 of the herd's 100 individuals are investigated.

For a fixed sensitivity above 0.6, $T$ is higher for specificities near 0.5 than those near 1. This might be surprising because a heuristic like "the higher the test quality the higher the number of positive tests before we react" feels natural. But such a heuristic neglects that a good test also results in fewer test-positives, because fewer are erroneously classified as positives. Looking at the figure also reveals that the radius of change for $T$ will be quite low due to the high variation of $T$ values over the parameter space. Again, this underlines the fact that care should be taken when sensitivity and specificity varies.

## 5 Discussion

Generation of treatment maps illustrating the sensitivity for varying probabilities is a strong tool helping to provide insight into the decision scenario. Calculations, where the expected utility function is given as an analytical function of $Se$ and $Sp$ works until samples sizes of 30-40. Hereafter, Maple is not capable of dealing with the generated polynomials anymore. By fixing $(Se, Sp)$ and calculating its values on a grid much higher $n$ can be achieved – either in Maple or by using Gibbs sampling in WinBugs.

Estimation of the constants $k_1, k_2$, and $k_3$ for a specific decision problem is problematic; guesstimates, small scale experiments, and sensitivity analysis could be employed. Once a reasonable single time-slice model is established, extension

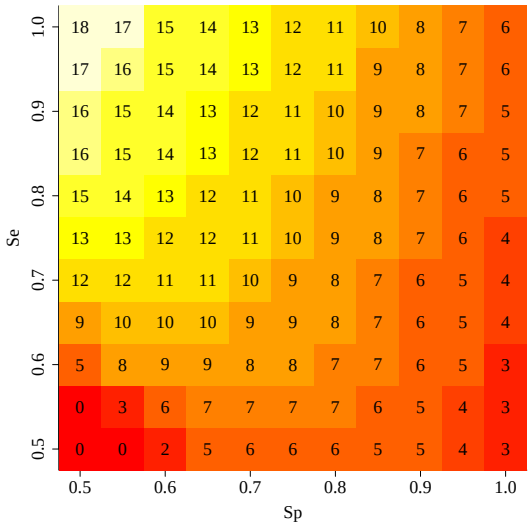


**Fig. 5.** Threshold $T$ as a function of $(Se, Sp)$ – a so called treatment map. Calculated by evaluating the analytical expression for a grid layout of $(Se, Sp)$ configurations.

to the more realistic case with additional time-slices is desirable. Biosecurity programs are often a temporal matter, where diagnosis and treatment are made repetitively. Limited memory strategies as in [17] might be necessary to obtain a tractable solution of the influence diagram. Despite such approximation our approach to sensitivity analysis would not scale up very well in respect to additional decisions; even Gibbs sampling would only be feasible for a small number of decisions.

To establish how large a sample size $n$ to choose in order to make an optimal decision about treatment would require conversion of $n$ in Fig. 2 into a decision node together with a cost of performing the diagnostic test. An analytical computation quickly becomes intractable here because $n$ is part of the exponent of $P(T^+|\ldots)$. Solving the influence diagram with the two sequential decisions would have to be done by numerical methods such as forward Monte Carlo sampling or Markov chain Monte Carlo sampling as described in [18, 19].

All these above mentioned problems would arise, in case one tries to evaluate and revise e.g. the current Danish Salmonella treatment strategy [3], which currently is taking neither uncertainty from imprecise tests nor any variability in sensitivity and specificity into account. This paper is merely an appetizer to work more intensively with the issues raised to get a practical solution.

## References

1. Barnett, V.: Sample survey : Principles and methods. Edward Arnold (1991)

2. Anonymous: U.S Voluntary Johne's Disease Herd Status Program for Cattle. Technical report, United States Animal Health Association (1998)
3. Alban, L., Stege, H., Dahl, J.: The new classification system for slaughter pig-herds in the danish salmonella surveillance-and-control program. Preventive Veterinary Medicine **1659** (2001) 1–14 Submitted for publication.
4. Baadsgaard, N.P.: Development of Clinical Monitoring Methods in Pig Health Management. PhD thesis, Department of Clinical Studies, The Royal Veterinary and Agricultural University and Department of Animal Health and Welfare, Research Center Foulum (2001)
5. Kristensen, K.: A collection of some statistical issues to consider when testing for GM seeds in conventional seed lots. Technical report, Biometry Research Unit, Danish Institute of Agricultural Sciences (2001)
6. Clement, R.: Making Hard Decisions: An Introduction to Decision Analysis. Duxbury Press (1996)
7. Johnson, W., Su, C.L., Gardner, I.: Sample size calculations for surveys to substantiate freedom of populations from infectious agents. (2002) Submitted to Biometrics.
8. Hanson, T., Johnson, W., Gardner, I., Georgiadis, M.: Determining the infection status of a herd. Journal of Agricultural, Biological, and Environmental Statistics (2003) In press.
9. Coupé, V.M.H., van der Gaag, L.C.: Practicable sensitivity analysis of Bayesian belief networks. Technical Report UU-CS-1998-10, Utrecht University, Department of Computer Science (1998)
10. Nielsen, T.D., Jensen, F.V.: Sensitivity analysis in influence diagrams. IEEE Transactions on Systems Man and Cybernetics (2001) Submitted for publication.
11. Höhle, M., Kristiansen, B.: Sensitivity analysis in Bayesian networks and influence diagrams. `http://www.dina.dk/~hoehle/pubs/sensitivity.pdf` (1998)
12. Waterloo Maple Inc.: Maple 6.02. (2001)
13. Lauritzen, S.L.: Graphical Models. Oxford University Press (1996)
14. Jensen, F.V.: Bayesian Networks and Decision Graphs. Statistics for Engineering and Information Science. Springer (2001)
15. Spiegelhalter, D., Thomas, A., Best, N.: WinBUGS Version 1.2 User Manual. MRC Biostatistics Unit. (1999)
16. Cameron, A., Baldock, F.: A new probability formula for surveys to substantiate freedom from disease. Prev. Vet. Medicine **34** (1998) 1–17
17. Lauritzen, S.L., Nilsson, D.: Representing and solving decision problems with limited information. Management Science **47** (2001) 1235–51
18. Charnes, J., Shenoy, P.: A forward Monte Carlo method for solving influence diagrams using local computation. Working paper No. 273, School of Business, University of Kansas (2000)
19. Bielza, C., Müller, P., Insua, D.: Decision analysis by augmented probability simulation. Management Science **45** (1999) 995–1007