SENSITIVITY ANALYSIS ON A CHANCE NODE WITH MORE THAN TWO BRANCHES

In *Medical Decision Making*, **19**, 499-502

Peter P. Sendi, MD and Robert T. Clemen, PhD

Running title: Sensitivity analysis and multi-branch chance nodes

Internal Medicine Outpatient Department, University of Basel, Basel, Switzerland (PPS); Fuqua School of Business, Duke University, Durham, North Carolina, USA (RTC).

Current address for correspondence: Dr. med. P. Sendi, Medizinische Universitäts-Poliklinik, Kantonsspital Basel, Petersgraben 4, CH-4031 Basel, Switzerland, Phone: ++41 61 265 50 12, Fax: ++41 61 265 46 04, e-mail sendi@datacomm.ch

ABSTRACT

Sensitivity analysis is an essential part of decision analysis. The literature on medical decision analysis suggests the use of two-branch chance nodes in decision trees to avoid logical inconsistencies during sensitivity analysis. We show that the two-branch decomposition is not appropriate for sensitivity analysis when multiple outcomes from a single state cannot be disentangled into a sensible sequence of events. We recommend retaining the natural structure of the tree and propose two sensitivity-analysis methods for use on chance nodes with three or more branches.

Keywords: Sensitivity analysis, decision analysis, decision tree, probability

INTRODUCTION

Sensitivity analysis is a crucial step in decision analysis. It allows the analyst to answer questions such as “Which variables are important in this problem?” or “If this variable changes by some amount, does the optimal choice change?” or “What is the threshold value of this variable?” In particular, sensitivity analysis of probabilities allows the construction of graphs showing what strategies are optimal for different probabilities or combinations of probabilities.[1]

Some writers advocate the use of only two-branch chance nodes in their decision trees to avoid logical inconsistencies during sensitivity analysis.[2] Of course, it is natural to use two branches at those chance nodes with dichotomous outcomes. However, investigators may sometimes need to model situations involving more than two outcomes, and the natural approach would be to include chance nodes with more than two branches. In these situations, the suggested two-branch decomposition may work poorly for sensitivity analysis, especially when multiple outcomes from a single state cannot be respecified in a sensible sequence of events. First, the original tree may have been created in terms of marginal (i.e., unconditional) probabilities, but the two-branch decomposition must perforce be specified in terms of both marginal and conditional probabilities. Thus, each new sensitivity analysis of a marginal probability requires a complete restructuring of the tree, as we will show below. Furthermore, the results of the analysis may be difficult to interpret because the two-branch decomposition may require construction of conditioning events that are in an arbitrary order and hence make little intuitive sense. Finally, the two-branch decomposition may prove less useful for communication with policy makers. For these reasons, this tutorial proposes viable methods for performing sensitivity analysis on the probabilities in chance nodes with three or more branches.

THE PROBLEM STATED

In this section, we discuss in more detail the concerns raised above using an example from AIDS research. For demonstration purposes, all notation and figures conform to the computer program DATA (TreeAge Software, Version 3.0.17, Williamstown, MA). The same analysis can be done with most other popular decision-tree programs [e.g., Decision Maker (Pratt Medical Group, Boston, MA) or SMLTREE (Hollenberg JP, Roslyn, NY)], although implementation details may vary somewhat.

HIV-infected individuals face the risk of falling into one of three different categories according to their pattern of disease progression (Fig. 1a): rapid, intermediate, or late progressors.[3] For example, we might model these as having life expectancies of 2, 10, or 30 years, respectively. When using the tree structure in Figure 1a, sensitivity analysis of a marginal probability may lead to situations where the sum of the three probabilities exceeds 1.[2] This happens, for example, when p1 is varied over a range that exceeds 1-p2, assuming p3 has been defined as the complementary probability 1-p1-p2. Fortunately, most software packages alert the user that a logical error has occurred.

To circumvent this problem, the two-branch decomposition has been the recommended solution in medical decision making. [2] Figure 1b shows such a decomposition with rapid progression as the conditioning event. In this model, the value of p1 = P(Rapid progression) can be varied between 0 and 1. Difficulties arise when we move to sensitivity analysis of marginal or unconditional probability p2 = P(Intermediate progression). The two-branch decomposition approach sets up the tree so that the probabilities of events on the second node are conditional probabilities; in Figure 1b, they are conditioned on the event ‘not rapid.’ Thus, in Figure 1b we have probability P(Intermediate progressor | Not Rapid) = p2/(p2+p3). (The conditional probability can be derived using Bayes Theorem. In this particular case, the calculations amount to normalizing probabilities p2 and p3 to sum to one.) Analyzing p2 in this model is not straightforward. To see the problem, note that we can indeed vary P(Intermediate progressor | Not Rapid) from 0 to 1, but as we do, p2 can vary only from 0 to 1-p1, because p1 remains fixed. The same problem exists for any other probabilities beyond the root node of the tree. Thus, with sequential two-branch chance nodes, we cannot immediately perform a sensitivity analysis over the entire range between 0 and 1 for any marginal probability except the one at the root node.

It is possible, of course, to rearrange the branches so that the root node is for a different event. For example, we could set the “Intermediate progressor” branch as the conditioning one. This allows us to vary the unconditional probability p2 between 0 and 1 as we did for p1 above. Restructuring the chance events to study the marginal probabilities successively is always possible, but may be quite complicated in large trees. Method 1 below provides an algebraic approach that does not require restructuring the tree. Also, we might be interested in analyzing the relationships among the probabilities to answer questions like, “What if the marginal probability of being a rapid progressor is three times the marginal probability of being an intermediate progressor, and nine times the marginal probability of being a late progressor?” The tree structure displayed in Figure 1b is not appropriate in this case because the probabilities on the “Intermediate” and “Late” branches are conditional probabilities. Method 2 below provides a way to answer such questions.

METHOD 1

Instead of expressing the conditioning event as a preceding branch in the tree structure, Method 1 incorporates the conditioning event in the definition of variables for a three-branch chance node. Given the three-branch node in Figure 1a, assuming the conditioning event is whether the patient is a rapid progressor, the intermediate progressor branch should be assigned (1-p1)*p2/(p2+p3), and the late progressor branch (1-p1)*p3/(p2+p3). Conditional probability p2/(p2+p3) serves as a “weighting coefficient” to determine how the complementary probability (1-p1) is divided up among the immediate and late progressor branches. With this specification, it is straightforward to perform sensitivity analysis on p1.

We can do the same for p2, writing p1 = (1-p2)*p1/(p1+p3) and p3 = (1-p2)*p3/p1+p3). Similarly, we can analyze p3 by replacing p1 by (1-p3)*p1/(p1+p2) and p2 by (1-p3)*p2/(p1+p2). Thus, although we have not physically restructured the tree, for each marginal probability we have a set of algebraic expressions that is fully consistent with such restructuring.

Instead of manually changing the specification each time sensitivity analysis is performed on another marginal probability, we automate this process using the choose function in DATA. The syntax of the *choose* function is *choose* (*index*; *value*1; *value*2; … ; *value*n), and the function returns a value from the list as determined by the index. For example, choose (2; 100; 200; 300) returns 200. To automate the sensitivity analysis, we define four new variables, *a*, *b*, *c* and *x*. The variable *a* is assigned to the upper branch, *b* to the middle branch, and *c* to the lower branch of a three-branch chance node as in Figure 2a; these variables are defined at the root of this chance node. By specifying *x*, we determine the "active" probability specification and hence the marginal probability for which sensitivity analysis can be performed. Then the appropriate probability for each branch is determined by a separate *choose* function as shown in Figure 2a. That is, set *x* equal to 1, 2, or 3 to run sensitivity analysis on p1, p2, or p3, respectively. This approach can be extended to chance nodes with more than three branches. The same may be accomplished with other decision-tree programs, although specific details may vary.

In some cases we need to perform a two-way sensitivity analysis for two marginal probabilities simultaneously. Of course, the sum of the two cannot exceed 1. For example, if we wanted to perform a two-way sensitivity analysis of p1 and p2, write p3 = 1-p1-p2. Now vary p1 between 0 and 1 and p2 between 0 and 1-p1.

For chance nodes n>3 branches, write pi = (1-p1-p2) pi/(p3+ … + pn) for branches 3, … n. See Figure 2b for an example with four branches. In this specification, pi/(p3+ … + pn) is the conditional probability of following branch i given that neither branch 1 nor 2 was followed. On multiplying the conditional probability by 1-p1-p2, we obtain the marginal probability p*i*. By writing p*i* as indicated above, we preserve the relative values of p3 through p*n* even while varying p1 and p2. We could use the *choose* function the same way as described above to automate the procedure for performing two-way sensitivity analyses over any or all combinations of the marginal probabilities.

METHOD 2

The major drawback of Method 1 is that it preserves the relative likelihood of those events whose probabilities are not being subjected to sensitivity analysis; i.e., when performing sensitivity analysis of p*i*, all other probabilities bear the same relative values, regardless of the value of p*i*. For example, suppose that Branch 2 is twice as likely as Branch 3, and we conduct a sensitivity analysis of p1. With Method 1, no matter what value of p1 is applied, Branch 2 will always be twice as likely as Branch 3.

Method 2 provides a flexible approach for performing sensitivity analysis on one or more probabilities, including investigating the effects of relaxing the assumption of constant relative likelihood. Suppose we have the situation in Figure 1a. Define a variable y = w1+w2+w3 at the root of the chance node, and express the probabilities of the three branches as p1n = w1/y for the upper branch, p2n = w2/y for the middle branch, and p3n = w3/y for the lower branch (Figure 3). Regardless of w1, w2, and w3, probabilities p1n, p2n, and p3n will always sum to 1. It is important to note that w1, w2, and w3 are not necessarily probabilities. These three values represent relative likelihood in contrast to p1n, p2n, and p3n, which represent probabilities.

In a one-way sensitivity analysis of w1, p1n changes relative to p2n and p3n, but p2n and p3n retain the same relative values. For example, suppose w1 = 3, w2 = 1, and w3 = 6. With this setup, one-way sensitivity analysis on w1 will result in changes in p1n, but p3n will always be six times as great as p2n. To explore the effect of changes in the relative likelihoods, the analyst can easily change the relative values of the w's, either by exploring a set of discrete scenarios in a "what-if" analysis or by performing a multi-way sensitivity analysis to changes in more than one of the w's.

As with Method 1, Method 2 is easily extended to situations with more than three branches. For branch i, write p*i* = w*i*/(w1 + ... + w*n*). Now any subset of the w's can be subjected to sensitivity analysis.

Note that with Method 2 there is no way to force any of the probabilities to 1 without setting the other probabilities to zero. Very large relative values for w*i* can be used to place arbitrarily small probabilities on the other branches. For example, with w1 = 500,000, w2 = 2, w3 = 1, the lower branch will still have half the probability of the middle branch, even though p2n and p3n are close to zero.

CONCLUSION

The practice of structuring chance events as a sequence of two-branch chance nodes can actually complicate rather than simplify sensitivity analysis of marginal probabilities. We advocate retaining the natural structure of the decision tree and using the methods described here for sensitivity analysis of the probabilities. Our methods provide easy, intuitive, and yet rigorous ways to perform the required sensitivity analysis.

REFERENCES

1. Clemen RT. Making Hard Decisions: An Introduction to Decision analysis, 2nd Ed. Belmont, CA: Duxbury Press, 1996.

2. Detsky AS, Naglie G, Krahn MD, Redelmeier DA, Naimark D. Primer on medical decision analysis: Part 2 - Building a tree. Med Decis Making. 1997;17:126-35.

3. Report of the NIH Panel to Define Principles of Therapy of HIV Infection and Guidelines for the Use of Antiretroviral Agents in HIV-Infected Adults and Adolescents. MMWR. 1998; 47:1-42.

ACKNOWLEDGMENTS

We are grateful to Dr. Heiner C. Bucher for his critiques on earlier drafts of this article. We thank Dr. Arthur S. Elstein, Dr. Bruce E. Hillner, and two anonymous reviewers of *Medical Decision Making* for their constructive and helpful comments.

Figure Legends

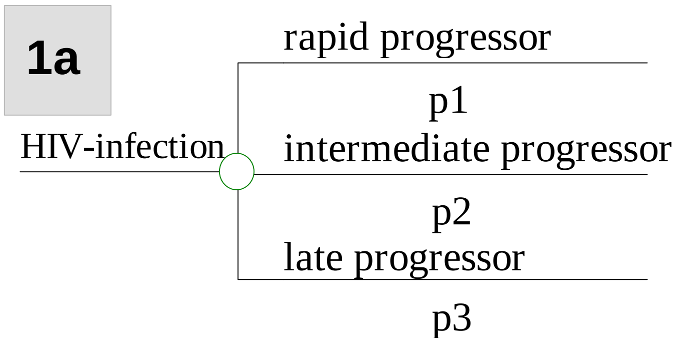
Figure 1. Event nodes showing two representations of the HIV infection problem.

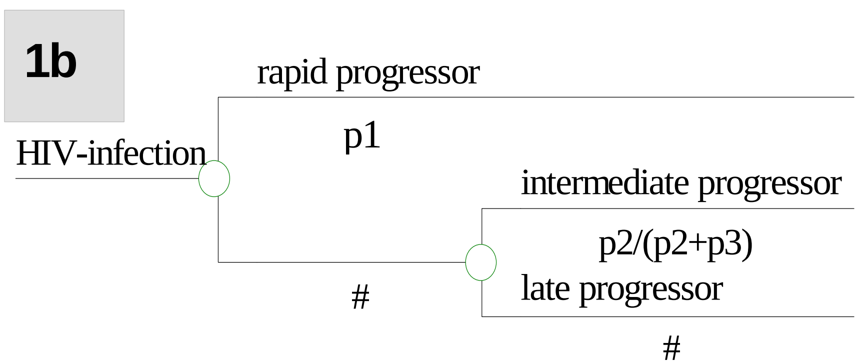
Figure 1a shows the basic problem and Figure 1b a decomposition using two-branch chance nodes. As used in DATA, the symbol # denotes the complementary value of the respective probability (first branch 1-p1, second branch 1-p2/(p2+p3)).

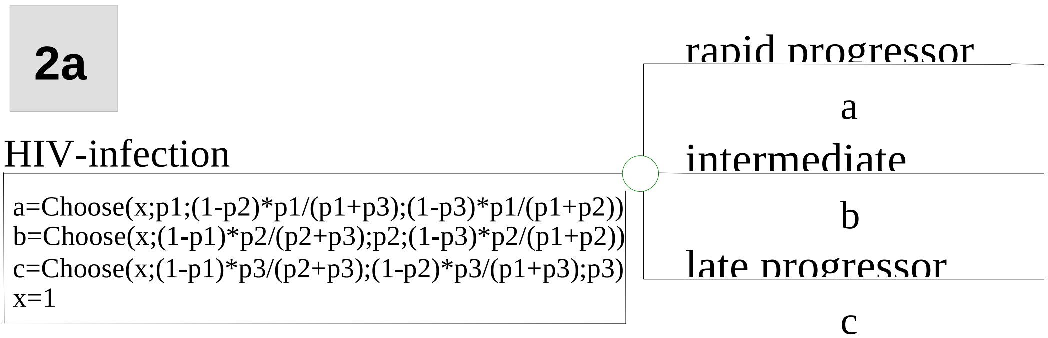
Figure 2. Chance nodes showing variable definitions for Method 1.

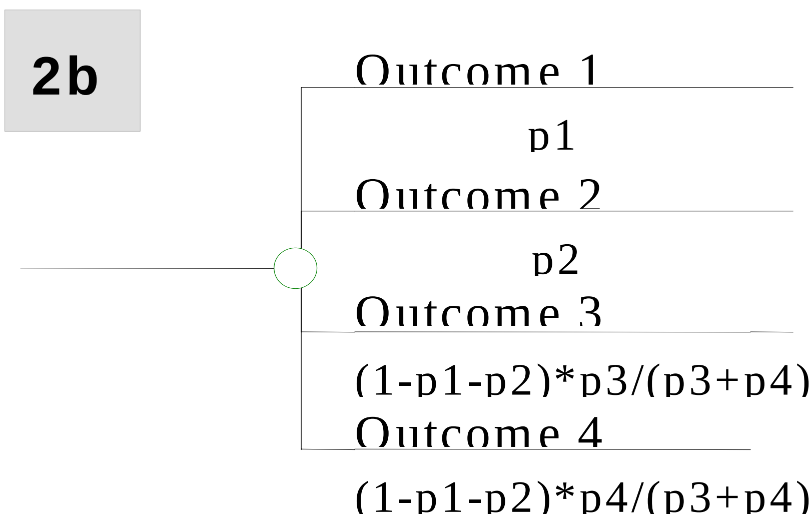
Figure 2a shows variable definitions in DATA for performing one-way sensitivity analyses; setting *x*= 1, 2, or 3 determines whether the sensitivity analysis is performed for p1, p2, or p3, respectively. (Other programs may require slightly different specifications. Figure 2b shows the variables for two-way sensitivity analysis of p1 and p2 in a four-branch chance event.

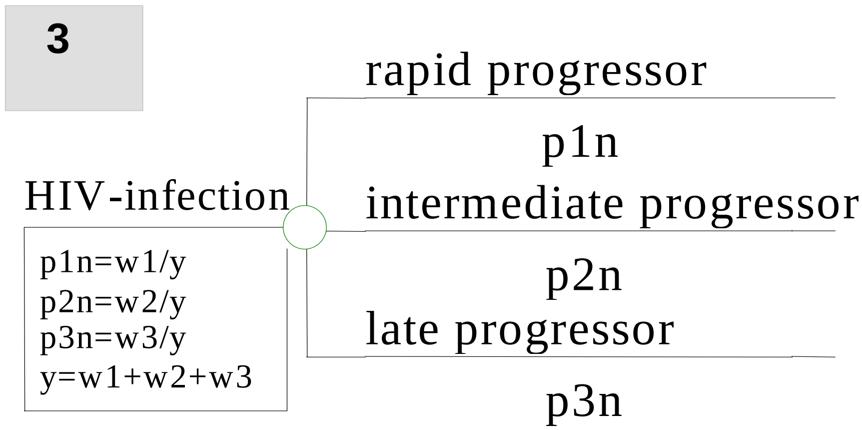
Figure 3. Chance node demonstrating variable definitions for Method 2.


1a
rapid progressor
p1
HIV-infection
intermediate progressor
p2
late progressor
p3


1b
rapid progressor
p1
HIV-infection
intermediate progressor
p2/(p2+p3)
late progressor
#
#


2a
HIV-infection
a=Choose(x;p1;(1-p2)*p1/(p1+p3);(1-p3)*p1/(p1+p2))
b=Choose(x;(1-p1)*p2/(p2+p3);p2;(1-p3)*p2/(p1+p2))
c=Choose(x;(1-p1)*p3/(p2+p3);(1-p2)*p3/(p1+p3);p3)
x=1
rapid progressor
a
intermediate
b
late progressor
c


2b
Outcome 1
p1
Outcome 2
p2
Outcome 3
(1-p1-p2)*p3/(p3+p4)
Outcome 4
(1-p1-p2)*p4/(p3+p4)


3
rapid progressor
p1n
HIV-infection
intermediate progressor
p1n=w1/y
p2n=w2/y
p3n=w3/y
y=w1+w2+w3
p2n
late progressor
p3n