Graphical Models for Decision Making

Lesson 5, week 14, class 27

Juan A Fdez del Pozo (CIG-UPM)

December 13, 2015

Contens

Outline of Contens

Overview about decision model examples

- Model script: helicopter.r
- Model script: underground.r
- Model script: server.r
- Model script: feb09/10.r

Gómez, M., C. Bielza, J. A. Fernandez del Pozo, and S. Ríos-Insua, "A Graphical Decision-Theoretic Model for Neonatal Jaundice", Medical Decision Making, vol. 27, no. 3, pp. 250-265, 2007

Talk IctNeo

- dss-seminar-2014.pdf
- Model script: ictneo2HGM.r
- Set the model on GeNie:
  - source("ictneo2HGM.r");
  - dump.netG(ictneo2HGM,"ictneo2HGM") → network-ictneo2HGM.xdls (GeNie)

Bielza, C., J. A. Fernandez del Pozo, and P. Lucas, "Explaining Clinical Decisions by Extracting Regularity Patterns", Decision Support Systems, vol. 44, pp. 397-408

Talk NHL

- clusterkbm2lcpts.pdf
- talk-kbm2l-cpt.pdf
- Model script: nhlv1.r
- Set the model on GeNie:
  - source("nhlv1.r");
  - dump.netG(nhlv1,"nhlv1") $\rightarrow$ network-nhlv1.xdls (GeNie)

Outline of Contens

Learnig and simulation Bayesian Networks

Use some of the following datasets, do the analysis and write a report in pdf format to upload in moodle

- Data: dataset-Z6a.csv, dataset-Z6b.csv; for each dataset:
- Define 6 variables $Z_i, i = 1 \ldots 6$, estimate the probability distribution
- Write using IdR format a BN with marginal nodes and evaluate it.
- Learn a BN using IdR Ibn() function, try it with diferent parameters
- Learn a BN using GeNie, try it with diferent parameters and algorithms
- Evaluate a BN using the exact algorithm and the snb() function
- Export each model to GeNie
- Analize and evaluate each model using GeNie and show instances of evidence

Check relevance of probabilistic influence

Learnig and simulation Bayesian Networks

Use some of the following datasets, do the analysis and write a report in pdf format to upload in moodle

- Data: dataset-P20a.csv, dataset-P24b.csv, dataset-P14a.csv, dataset-P14a.csv; for each dataset:
- Define 20/24 variables $P_i, i = 1 \ldots 20/P_i, i = 1 \ldots 24$,
- Learn a BN for each dataset using ldR format
- Evaluate each BN using the exact algorithm and the snb() function
- Export each model to GeNie
- Learn a BN using GeNie, try it with diferent parameters and algorithms
- Analize and evaluate each model using GeNie and show instances of evidence

Check relevance of probabilistic influence

Use some of the following models, do the analysis and write a report in pdf format to upload in moodle

Decision Models Definition and Evaluation

- Models: feb2009.r, feb2010.r, itneo2HGM.r, nhl1.r, underground.r, server.h, helicopter.r
- Evaluate and show the optimal decision alternative tables
- Export to GeNie and evaluate the models
- Show some instances of evidence

The aim of this assignment is to build up experience in developing Influence Diagrams for realistic problems. Given the Non-Hodgkin lymphoma problem, from the clinical domain:

The problem

Non-Hodgkin lymphoma of the stomach, gastric NHL for short, is a relatively uncommon malignant disorder, accounting for about 5% of tumours of the stomach. Until recently, the cause of gastric NHL was unknown; it is now generally believed that the main factor in the development of this disease is a chronic infection with the bacterium Helicobacter pylori. This has had a major effect on treatment practice. Whereas originally, as in most cancers, treatment consisted of surgery (total or partial removal of the stomach), chemotherapy, radiotherapy or a combination of two or three of these, there is now also a place for antibiotics. Only 10 years ago, no medical doctor would have believed you when you had said that cancer can be treated by antibiotics. So, the impact of these recent findings has been dramatic. Now, the selection of treatment for gastric NHL is a complicated process, because only part of the patient findings necessary for therapy selection may be known at a particular stage of the disease, and knowledge of adverse reactions to particular treatments in patient groups may influence treatment selection significantly. This explains why the Netherlands Cancer Institute in Amsterdam considered developing a influence diagram. It was hoped that a influence diagram of gastric NHL might help doctors in the prescription of optimal treatment of a patient. The network discussed here is still in prototype stage; further development needs to take place in order to introduce it in actual clinical practice.

The aim of this assignment is to build up experience in developing Influence Diagrams for realistic problems. Given the NHL model, the nhlv1.r:

The network

The gastric NHL influence diagram only incorporates variables that are widely used by clinicians in choosing the appropriate therapy for patients. The relevance of most of these variables is supported by literature on prognostic factors in gastric NHL. First, the information used in the clinical management of primary gastric NHL was subdivided in pretreatment information, i.e. information that is required for treatment selection, treatment information, i.e. the various treatment alternatives, and posttreatment information, i.e. side effects, and early and long-term treatment results for the disease. The most important pretreatment variables in the table are the variable 'clinical stage', which expresses severity of the disease according to a common clinical classification, and histological classification, which stands for the assessment by a pathologist of tumour tissue obtained from a biopsy. Various treatments are in use for gastric NHL such as chemotherapy, radiotherapy, and a combination of these two, which has been represented as the single variable 'ct&rtschedule' with possible values: chemotherapy (CT), radiotherapy (RT), chemotherapy followed by radiotherapy (CT-next-RT), and neither chemotherapy nor radiotherapy (none). Furthermore, surgery is a therapy with is modelled by the variable 'surgery' with possible values: 'curative', 'palliative' or 'none', where curative surgery means total or partial resection of the stomach with the complete removal of tumour mass. Finally, prescription of antibiotics is also possible. The most important posttreatment variables are the variable 'early result', being the endoscopically verified result of the treatment, six to eight weeks after treatment (possible outcomes are: complete remission – i.e. tumour cells are no longer detectable –, partial remission – some tumour cells are detectable –, no change or progressive disease), and the variable '5-year result', which represents the patient either or not surviving five years following treatment.

Graphical Models for Decision Making

Practical Assignment

N (Patient)	Age	Gender	Clinical Stage	Histological Classification	Bulky Disease	Helicobacter Pylori	Clinical Presentation
1	61	m	I	high	non-bulky	−	gastric perforation
2	39	m	I	high	bulky	−	none
3	64	f	I	low	non-bulky	+	none
4	63	m	II1	high	non-bulky	+	gastric obstruction
5	77	m	II1	low	non-bulky	+	none
6	82	f	II1	high	bulky	−	none
7	46	m	II1	high	non-bulky	+	gastric hemorrhage
8	60	m	I	high	non-bulky	+	none
9	47	m	I	high	bulky	+	none
10	67	f	IV	high	non-bulky	−	none
11	73	m	IV	high	bulky	−	none

Stage: Clinical stage according to the Ann Arbor classification of NHL by Musshoff
Grade: histological MALT classification
HP: H. pylori present (+) or absent (−)

Exercises

→ For each of these patients, try to determine the treatment which yields the best results.

→ A typical clinical research question might be whether there is a difference between those patients who live shorter and those that live longer than 5 years following treatment. Use the network to answer this question.

→ Is age also a factor that affects the results of this question, i.e. can different patient groups be distinguished?

¿Remarks and Questions?

IDA 2015

12/01/2015, CIG