Value of Information Analysis in Spatial Models

Jo Eidsvik

Jo.eidsvik@ntnu.no

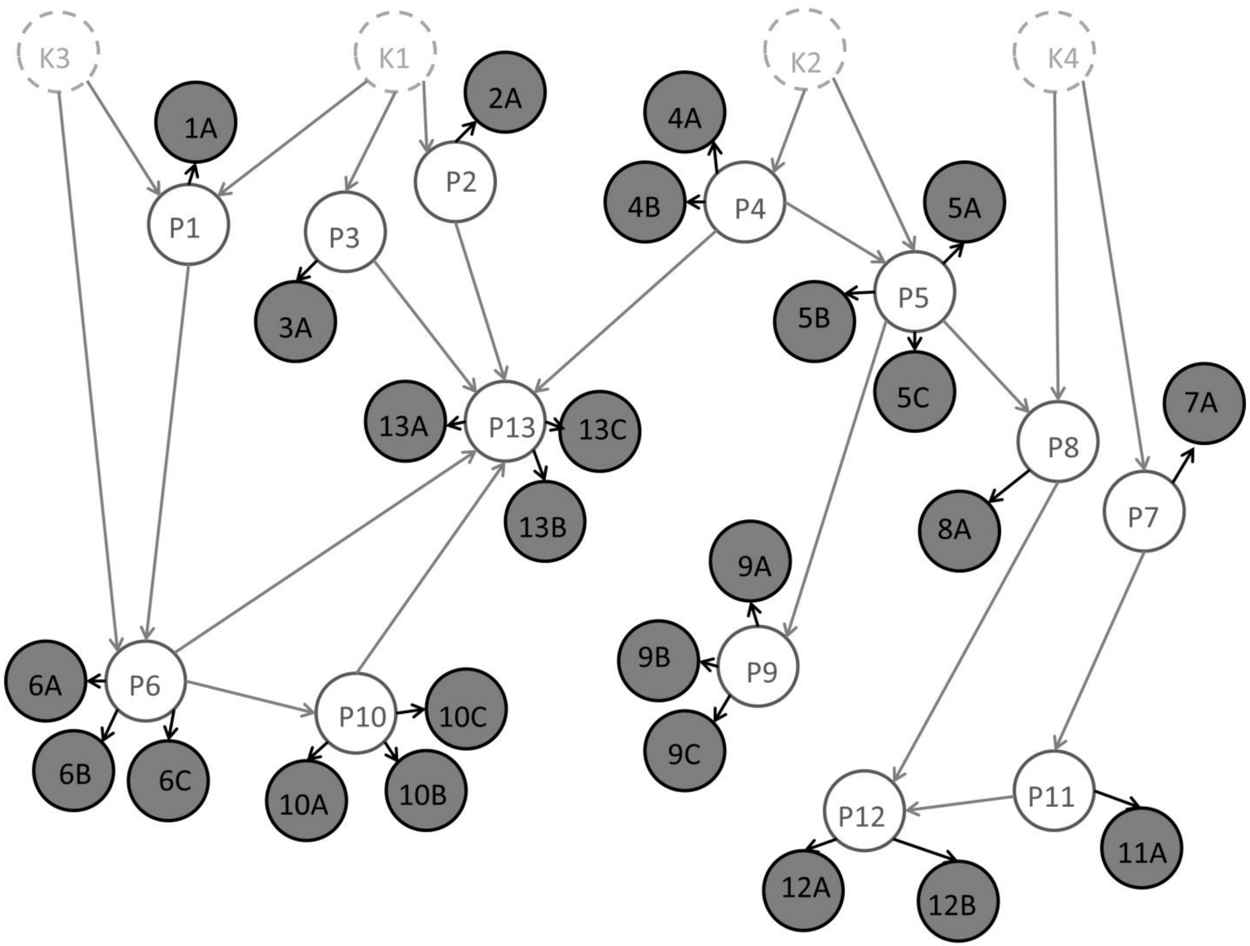
Plan for course

Time	Topic
Monday	Introduction and motivating examples
	Elementary decision analysis and the value of information
Tuesday	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for dependent models
Wednesday	Spatial statistics, spatial design of experiments
	Value of information analysis in spatial decision situations
Thursday	Examples of value of information analysis in Earth sciences
	Computational aspects
Friday	Sequential decisions and sequential information gathering
	Examples from mining and oceanography

Every day: Small exercise half-way, and computer project at the end.

Dependence? Does it matter?

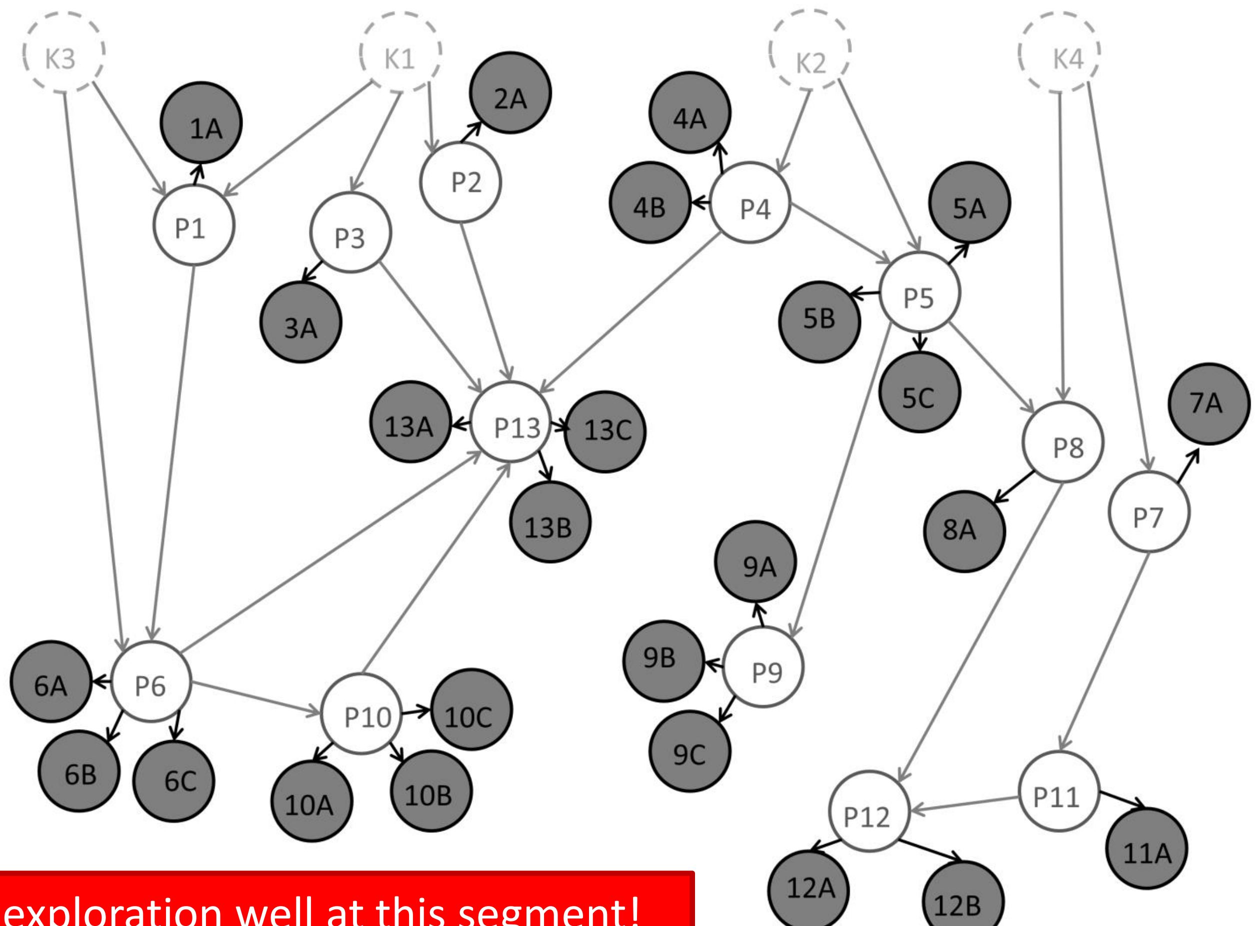
Gray nodes are petroleum reservoir segments where the company aims to develop profitable amounts of oil and gas.



Martinelli, G., Eidsvik, J., Hauge, R., and Førland, M.D., 2011, Bayesian networks for prospect analysis in the North Sea, *AAPG Bulletin*, 95, 1423-1442.

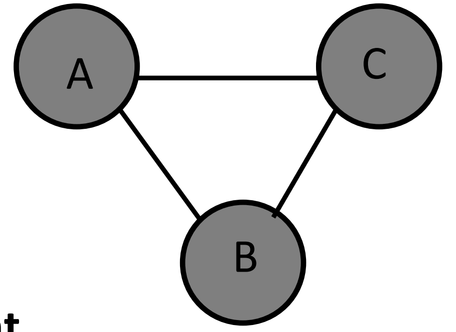
Dependence? Does it matter?

Gray nodes are petroleum reservoir segments where the company aims to develop profitable amounts of oil and gas.



Drill the exploration well at this segment! The value of information is largest.

Joint modeling of multiple variables



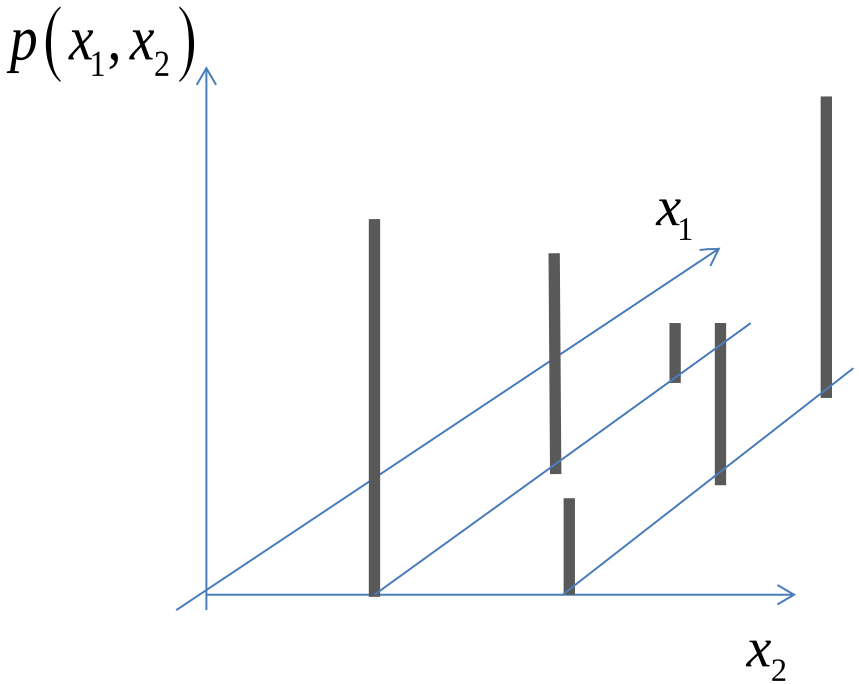
Spatial variables are often not independent!

To study if dependence matter, we need to model the **joint** properties of uncertainties.

- What is the probability that variable A is 1 and, at the same time, variable B is 1 ?
- What is the probability that variable C is 0, and both A and B are 1 ?

Joint pdf

$$p(\boldsymbol{x}) = p(x_1, \ldots, x_n)$$



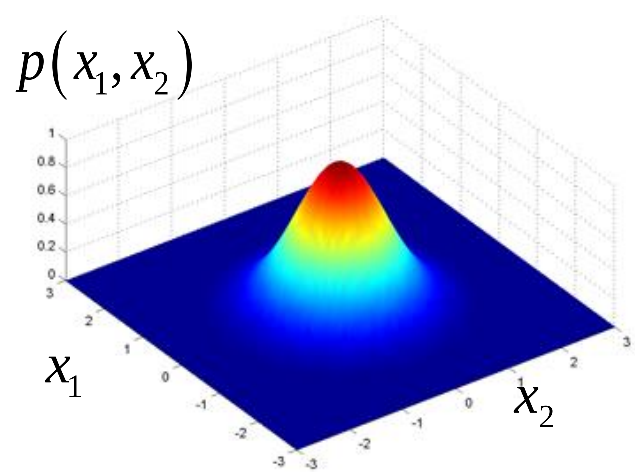
Discrete sample space:

$$p(\boldsymbol{x}) \geq 0, \qquad \boldsymbol{x} \in \Omega,$$
$$\sum_{x_1 \in \Omega_1} \cdots \sum_{x_n \in \Omega_n} p(\boldsymbol{x}) = 1.$$

Probability mass function (pdf)

Continuous sample space:

$$p(\boldsymbol{x}) \geq 0, \qquad \boldsymbol{x} \in \Omega,$$
$$\int_{x_1 \in \Omega_1} \cdots \int_{x_n \in \Omega_n} p(\boldsymbol{x}) dx_1 \ldots dx_n = 1.$$



Probability density function (pdf)

Multivariate statistical models

The joint probability mass or density function (**pdf**) defines all probabilistic aspects of the distribution!

$$p(\boldsymbol{x}) = N(\mathbf{0}, \boldsymbol{\Sigma}), \qquad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

$$\downarrow$$

$$E(\boldsymbol{x}) = \boldsymbol{\mu} = \int \boldsymbol{x} p(\boldsymbol{x}) d\boldsymbol{x},$$

$$Var(\boldsymbol{x}) = \boldsymbol{\Sigma} = \int (\boldsymbol{x} - \boldsymbol{\mu})(\boldsymbol{x} - \boldsymbol{\mu})^t \, p(\boldsymbol{x}) d\boldsymbol{x},$$

$$E(f(\boldsymbol{x})) = \int f(\boldsymbol{x}) p(\boldsymbol{x}) d\boldsymbol{x}.$$

Marginal and conditional probability

$$\boldsymbol{x} = (\boldsymbol{x}_{\mathbb{K}}, \boldsymbol{x}_{\mathbb{L}})$$

$$p(\boldsymbol{x}_{\mathbb{K}}) = \int p(\boldsymbol{x})\, d\boldsymbol{x}_{\mathbb{L}}$$

Marginalization in joint pdf.

$$p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}}) = \frac{p(\boldsymbol{x})}{p(\boldsymbol{x}_{\mathbb{L}})} = \frac{p(\boldsymbol{x})}{\int p(\boldsymbol{x})\, d\boldsymbol{x}_{\mathbb{K}}}$$

Conditioning in joint pdf.

Conditional mean and variance

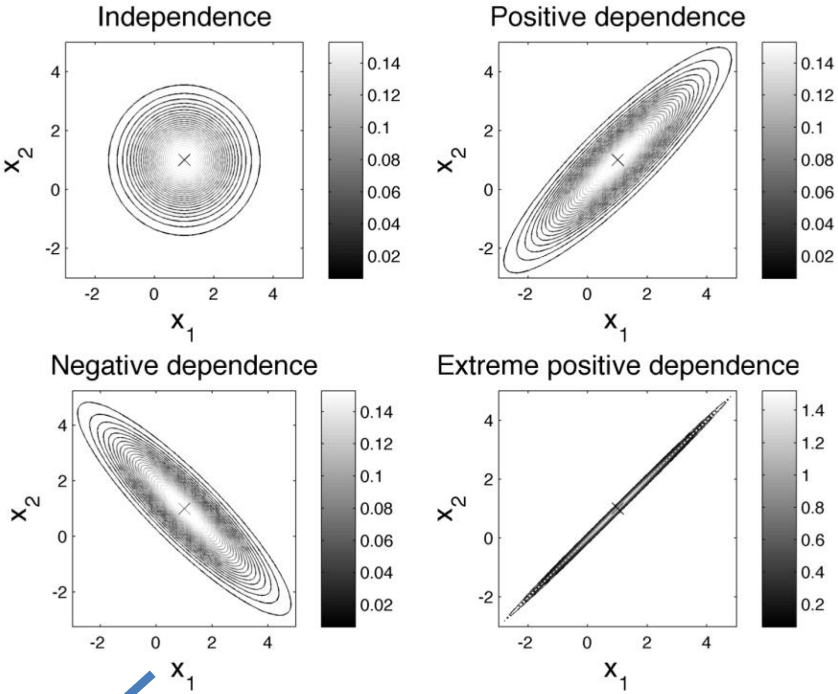
$$E(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}}) = \int \boldsymbol{x}_{\mathbb{K}}\, p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}})\, d\boldsymbol{x}_{\mathbb{K}},$$

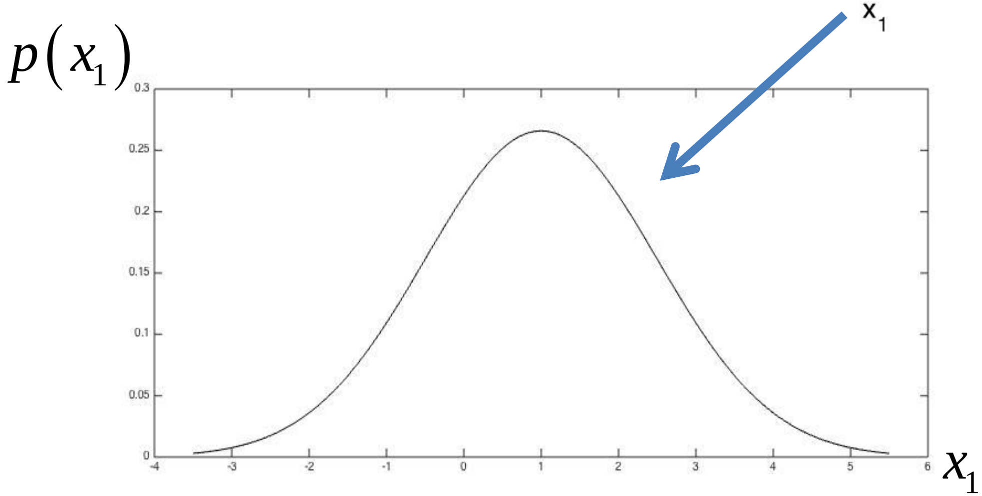
$$Var(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}}) = \int \left(\boldsymbol{x}_{\mathbb{K}} - E(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}})\right)\left(\boldsymbol{x}_{\mathbb{K}} - E(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}})\right)^t p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}})\, d\boldsymbol{x}_{\mathbb{K}}.$$

Marginalization

$$\boldsymbol{x} = (\boldsymbol{x}_{\mathbb{K}}, \boldsymbol{x}_{\mathbb{L}})$$

$$p(\boldsymbol{x}_{\mathbb{K}}) = \int p(\boldsymbol{x})\, d\boldsymbol{x}_{\mathbb{L}}$$



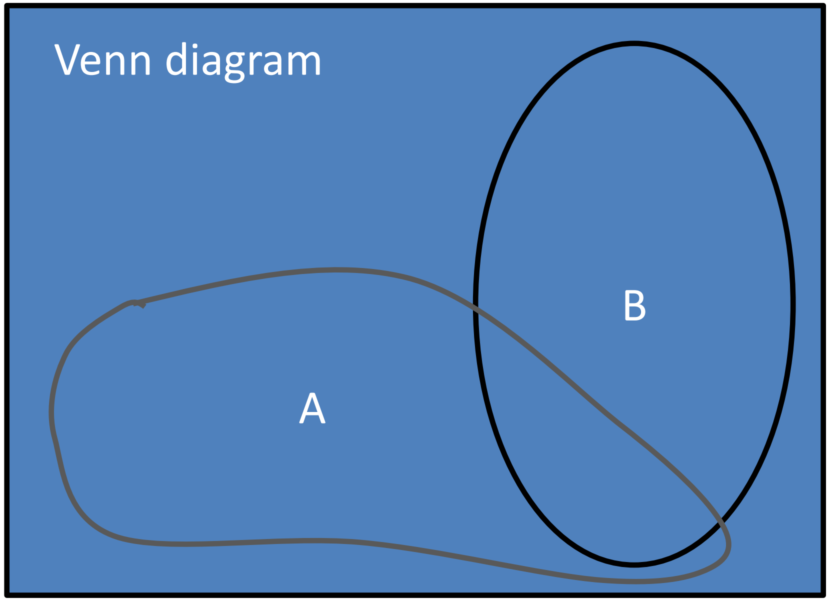


Conditional probability

$$p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}}) = \frac{p(\boldsymbol{x})}{p(\boldsymbol{x}_{\mathbb{L}})} = \frac{p(\boldsymbol{x})}{\int p(\boldsymbol{x})\, d\boldsymbol{x}_{\mathbb{K}}}$$

$$p(A \mid B) = \frac{Area(A \cap B)}{Area(B)}$$

$$B = (A \cap B) \cup (A^C \cap B)$$



Conditional probability

$$p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}}) = \frac{p(\boldsymbol{x})}{p(\boldsymbol{x}_{\mathbb{L}})} = \frac{p(\boldsymbol{x})}{\int p(\boldsymbol{x})\, d\boldsymbol{x}_{\mathbb{K}}}$$

$$p(\boldsymbol{x}_{\mathbb{K}}) \neq p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}})$$

$$p(\boldsymbol{x}_{\mathbb{K}}) = p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}})$$

Independence!

Must hold for all outcomes and
for all subsets!
Unrealistic in most applications!

Modeling by conditional probability

The **joint** pdf can be difficult to model directly.
Instead we can build the joint pdf from **conditional** distributions.

$$p(\boldsymbol{x}) = p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}})\, p(\boldsymbol{x}_{\mathbb{L}})$$

$$p(\boldsymbol{x}) = p(x_1)\, p(x_2 \mid x_1) \ldots p(x_n \mid x_{n-1}, \ldots, x_1)$$

Holds for any ordering of variables.

Modeling by conditional probability

Modeling by conditionals is done by conditional statements, not joint assessment:

- What is likely to happen for variable K when variable L is 1?
- What is the probability of variable C being 1 when variables A and B are both 0?

Such statements might be easier to specify,
and can more easily be derived from physical principles.

$$p(\boldsymbol{x}) = p(\boldsymbol{x}_{\mathbb{K}} \mid \boldsymbol{x}_{\mathbb{L}})\, p(\boldsymbol{x}_{\mathbb{L}})$$

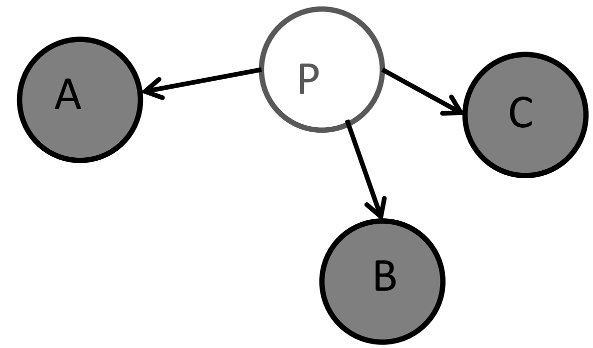
Modeling by conditional probability

$$p(\boldsymbol{x}) = p(x_1)\, p(x_2 \mid x_1) \ldots p(x_n \mid x_{n-1}, \ldots, x_1)$$

Holds for any ordering of variables. Some conditioning variables can often be skipped.
Conditional independence in modeling.
This simplifies modeling and interpretation! And computing!

Modeling by conditional probability

Conditional independence:



$$p\left(x_A, x_B, x_C \mid x_P\right) = \prod_{i \in \{A,B,C\}} p\left(x_i \mid x_P\right)$$

- What is the chance of success at B, when there is success at parent P?
- What is the chance of success at B, when there is failure at parent P?

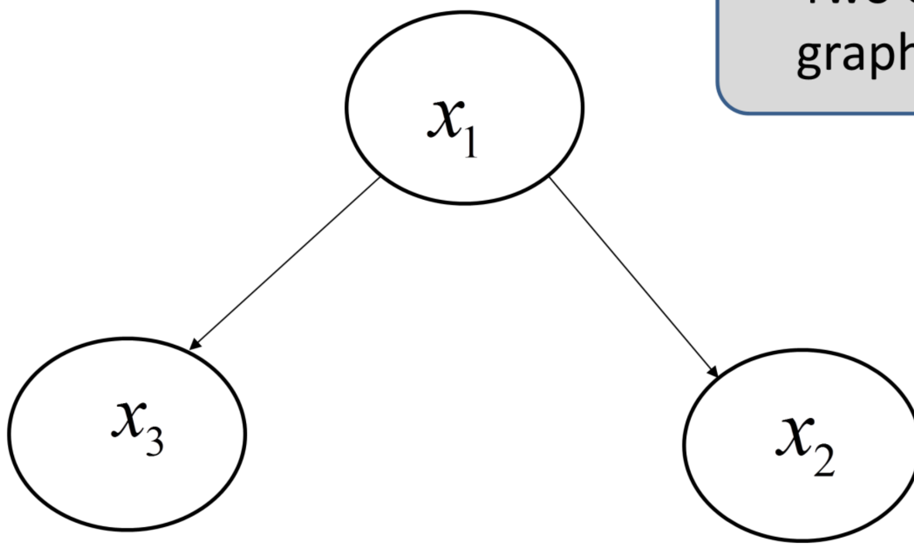
$$p\left(x_B = 1 \mid x_P = 1\right) = 0.9$$

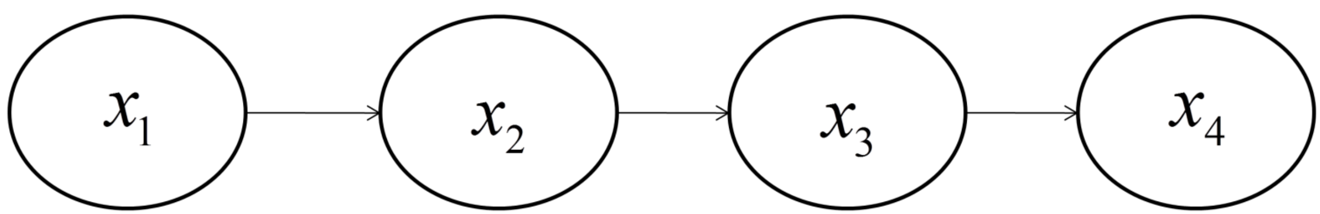
$$p\left(x_B = 1 \mid x_P = 0\right) = 0$$

Must set up models for all nodes, using marginals for root nodes, and conditionals for all nodes with edges.

Bayesian networks and Markov chains

Two examples of graphical models

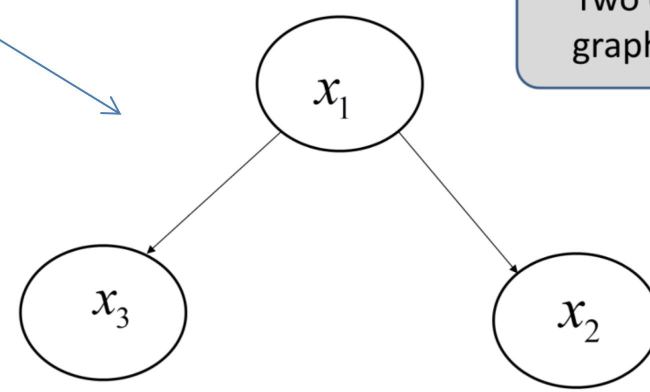


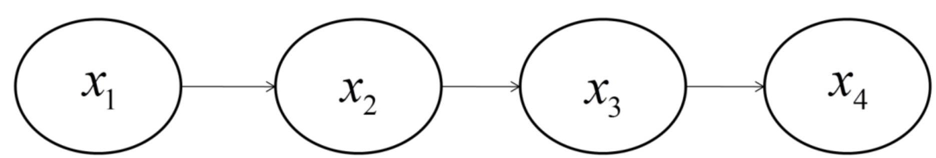


Bayesian networks and Markov chains

$$p(\boldsymbol{x}) = p(x_1, x_2, x_3) = p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_1)$$

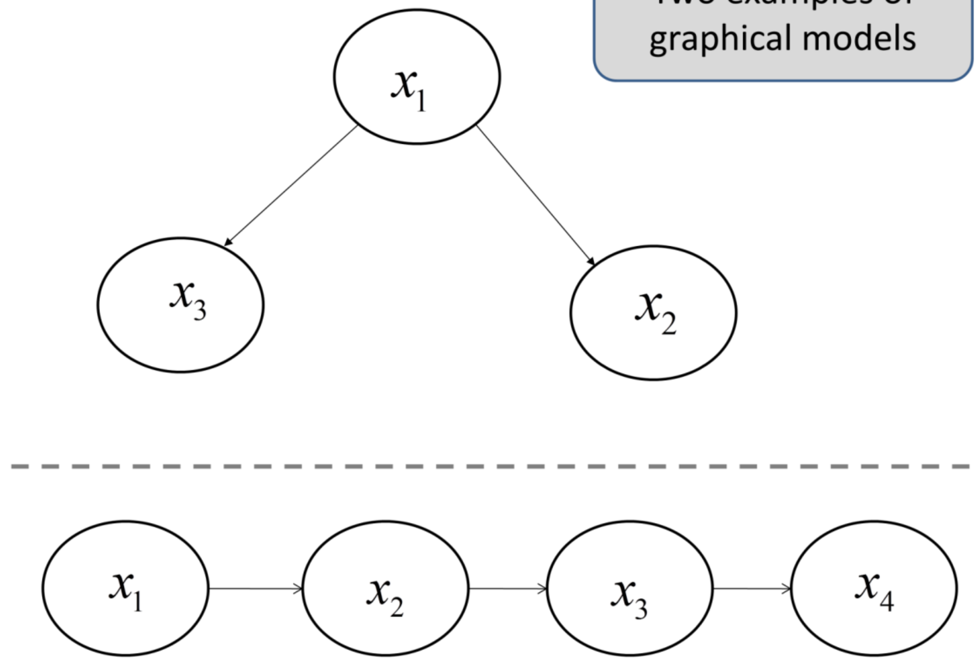
Two examples of graphical models





Bayesian networks and Markov chains

Two examples of graphical models



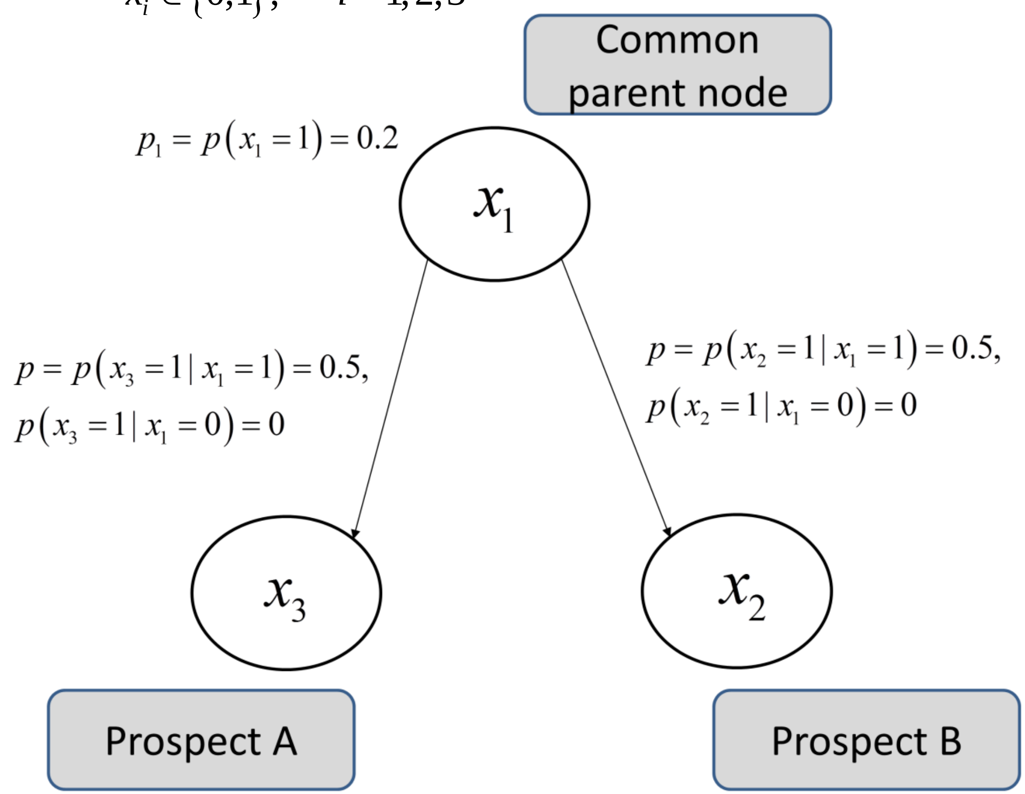
$$p(\boldsymbol{x}) = p(x_1, x_2, x_3, x_4) = p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_2)\, p(x_4 \mid x_3)$$

Bivariate petroleum prospects example

Conditional independence between prospect A and B, given outcome of parent!

Similar network models have been used in medicine/genetics, and testing for heritable diseases.

$x_i \in \{0,1\}, \quad i = 1,2,3$

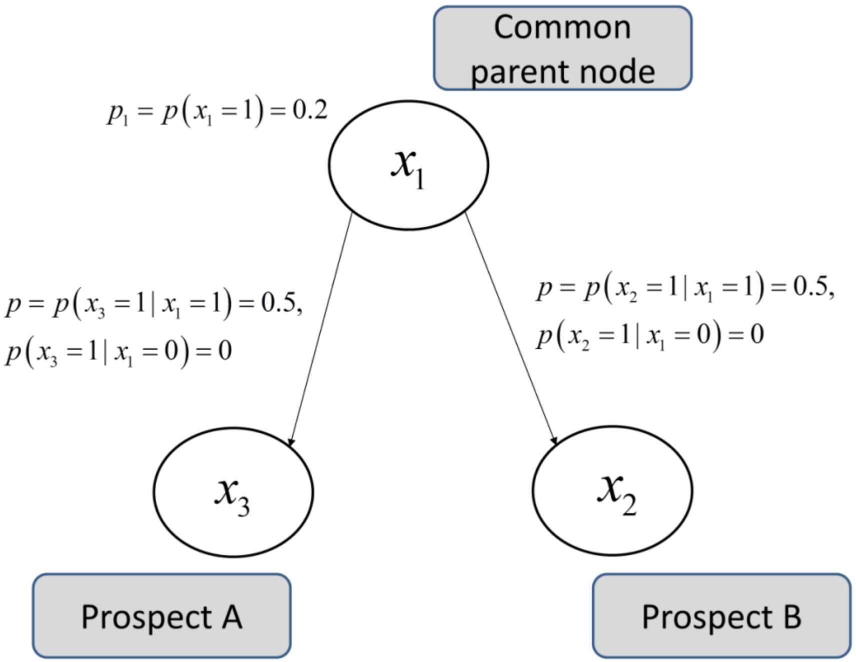


Exercise: Bivariate petroleum prospects example

Exercise:

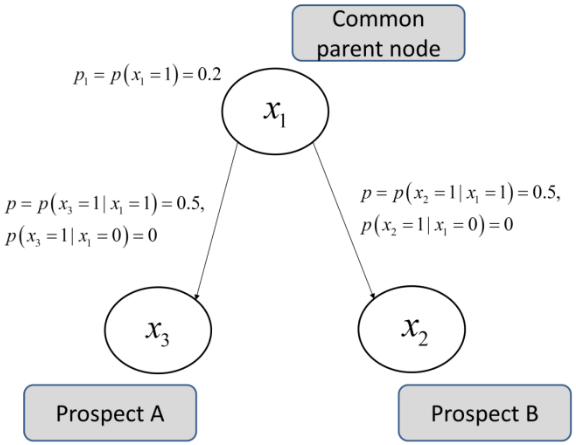
1. **Compute the conditional probability at prospect A, when one knows the success or failure outcome of prospect B.**
2. **Compare with marginal probability.**

$$x_i \in \{0,1\}, \quad i = 1,2,3$$

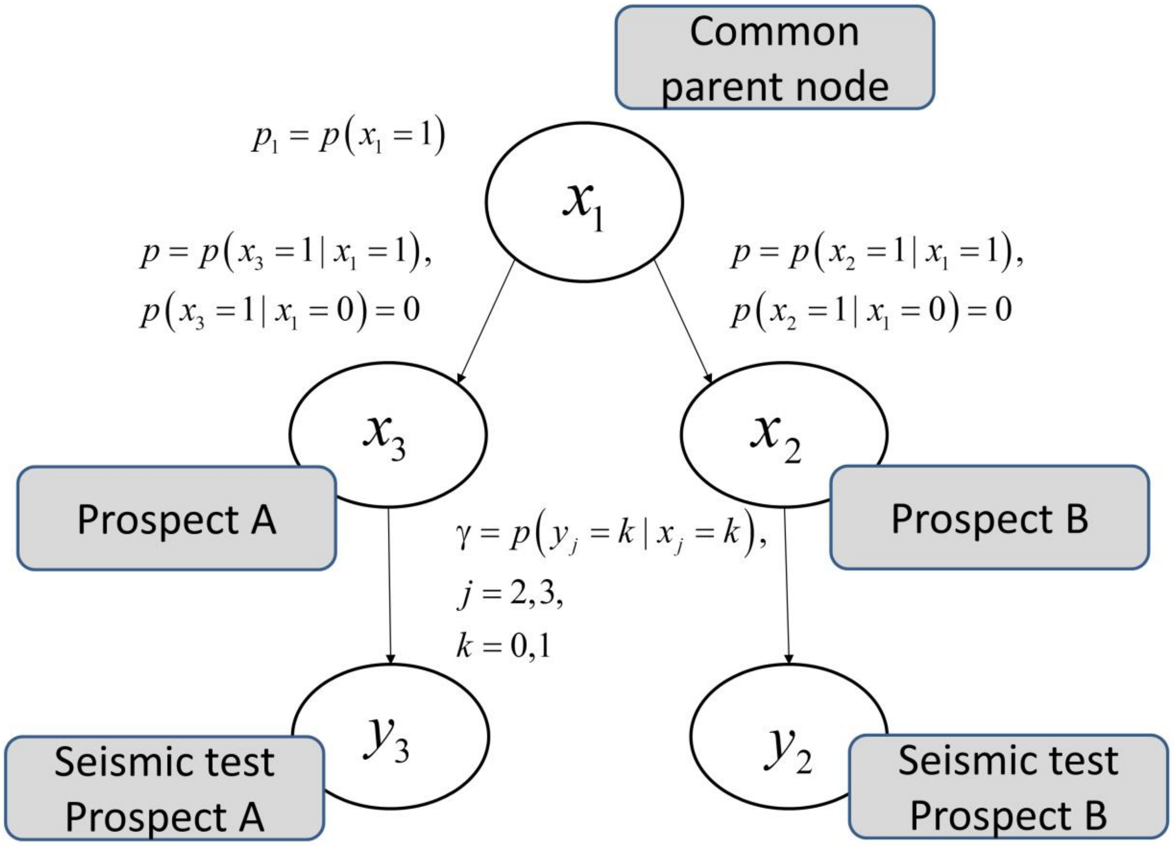


Bivariate petroleum prospects example

Joint	Failure prospect B	Success prospect B	Marginal probability
Failure prospect A	0.85	0.05	0.9
Success prospect A	0.05	0.05	0.1
Marginal probability	0.9	0.1	1



Example - Bivariate petroleum prospects



Collect seismic data :VOI - Should data be collected at both prospects, or just one of them? Partial or total? Imperfect or perfect?

Bivariate petroleum prospects

Need to frame the **decision situation**:

- Can one freely select (profitable) prospects, or must both be selected.
- Does value decouple?
- Can one do sequential selection?

Need to study **information gathering** options:

- Imperfect (seismic), or perfect (well data)?
- Can one test both prospects, or only one (total or partial)?
- Can one perform sequential testing?

Bivariate petroleum prospects

Need to frame the **decision situation**:

- Can one freely select (profitable) prospects, or must both be selected. Free selection.
- Does value decouple? Yes, no communication between prospects.
- Can one do sequential selection? Non-sequential.

Need to study **information gathering** options:

- Imperfect (seismic), or perfect (well data)? Study both.
- Can one test both prospects, or only one (total or partial)? Study both.
- Can one perform sequential testing? Not done here.

Bivariate prospects example - perfect

Assume we can freely select (develop) prospects, if profitable.

$$\mathrm{Rev}_1 = \mathrm{Rev}_2 = \mathrm{Rev} = 3$$

$$PV = \sum_{i \in \{A,B\}} \max\left\{0, \mathrm{Rev} \cdot p\left(x_i = 1\right) - \mathrm{Cost}\right\}$$

$$= 2\max\left\{0, 0.3 - \mathrm{Cost}\right\}$$

Total clairvoyant information

$$PoV\left(\boldsymbol{x}\right) = \sum_{i \in \{A,B\}} p\left(x_i = 1\right) \cdot \max\left\{0, \mathrm{Rev} - \mathrm{Cost}\right\}$$

$$= 0.2\max\left\{0, 3 - \mathrm{Cost}\right\}$$

$$VOI\left(\boldsymbol{x}\right) = PoV\left(\boldsymbol{x}\right) - PV$$

Bivariate prospects example - perfect

Assume we can freely select (develop) prospects, if profitable.

$$\text{Rev}_1 = \text{Rev}_2 = \text{Rev} = 3$$

$$PV = \sum_{i \in \{A,B\}} \max\left\{0, \text{Rev} \cdot p\left(x_i = 1\right) - \text{Cost}\right\}$$

$$= 2\max\left\{0, 0.3 - \text{Cost}\right\}$$

Partial clairvoyant information

$$PoV\left(x_A\right) = p\left(x_A = 1\right) \cdot \max\left\{0, 3 - \text{Cost}\right\}$$

$$+\sum_{l} p\left(x_A = l\right) \cdot \max\left\{0, \text{Rev} \cdot p\left(x_B = 1 \mid x_A = l\right) - \text{Cost}\right\}$$

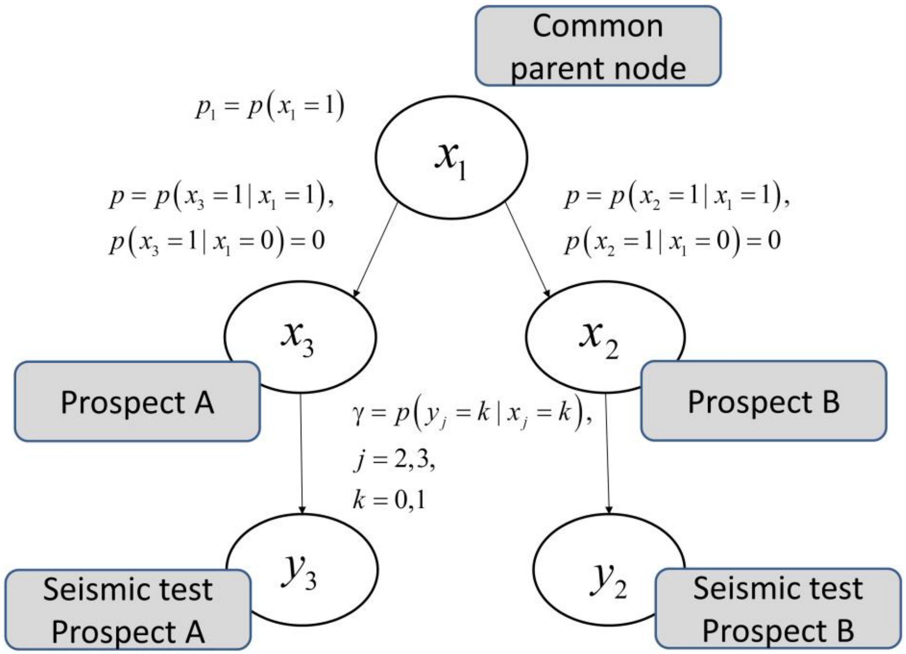
$$= 0.1 \cdot \max\left\{0, 3 - \text{Cost}\right\} + 0.1 \cdot \max\left\{0, \text{Rev} \cdot 0.5 - \text{Cost}\right\}$$

$$+0.9 \cdot \max\left\{0, 3 \cdot 0.055 - \text{Cost}\right\}$$

Bivariate prospects example - imperfect

Define sensitivity of seismic test (imperfect):

$$p\left(y_j = k \mid x_j = k\right) = \gamma = 0.9, \qquad k = 1,2$$



Bivariate prospects example - imperfect

Assume we can freely select (develop) prospects, if profitable.

$$\text{Rev}_1 = \text{Rev}_2 = \text{Rev} = 3$$

$$PV = \sum_{i \in \{A,B\}} \max\left\{0, \text{Rev} \cdot p\left(x_i = 1\right) - \text{Cost}\right\}$$

$$= 2\max\left\{0, 0.3 - \text{Cost}\right\}$$

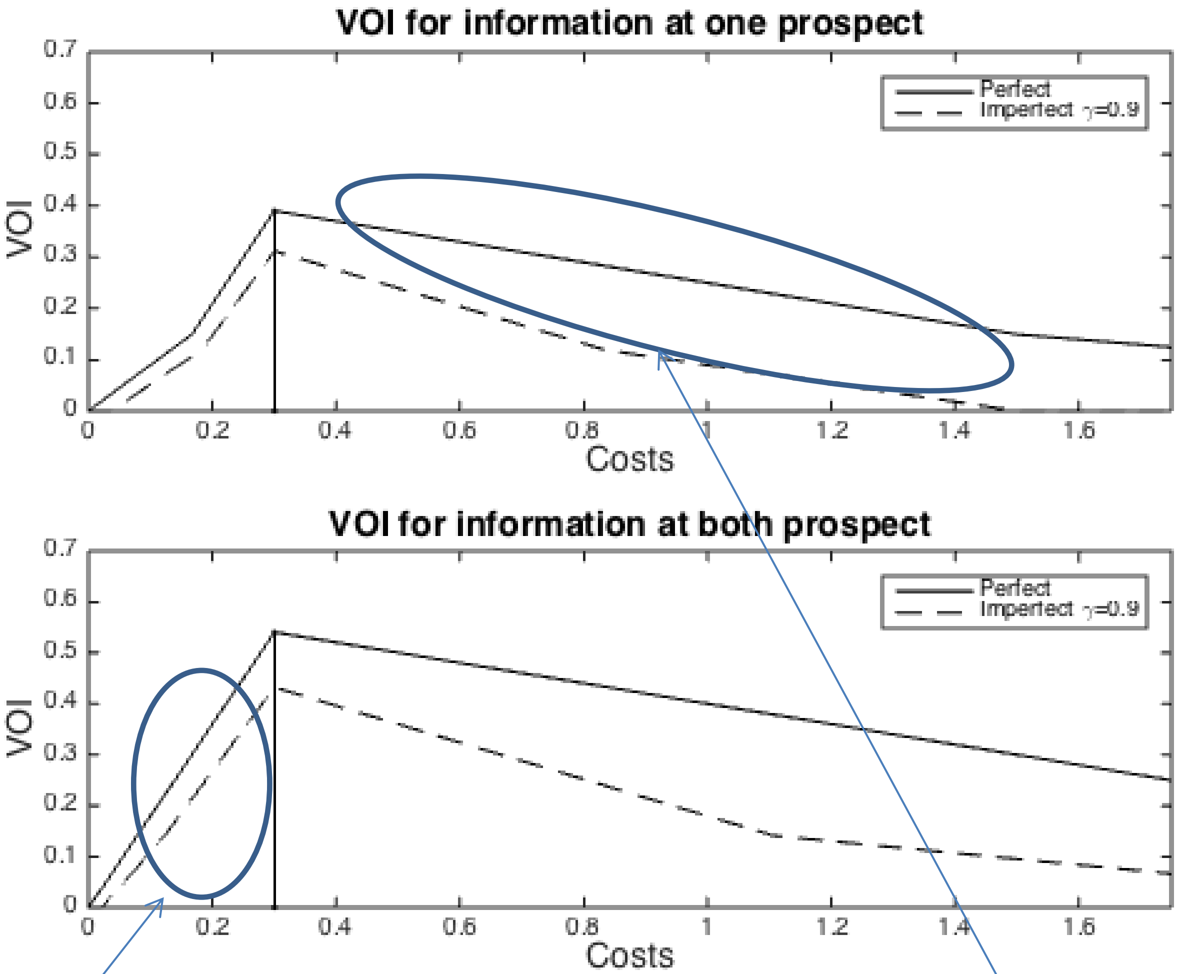
Total imperfect information

$$PoV\left(\boldsymbol{y}\right) = \sum_{\boldsymbol{y}} \sum_{i \in \{A,B\}} \max\left\{0, \text{Rev}\, p\left(x_i = 1 \mid \boldsymbol{y}\right) - \text{Cost}\right\} p\left(\boldsymbol{y}\right)$$

$$VOI\left(\boldsymbol{y}\right) = PoV\left(\boldsymbol{y}\right) - PV$$

Can also purchase imperfect partial information i.e. about one of the prospects?

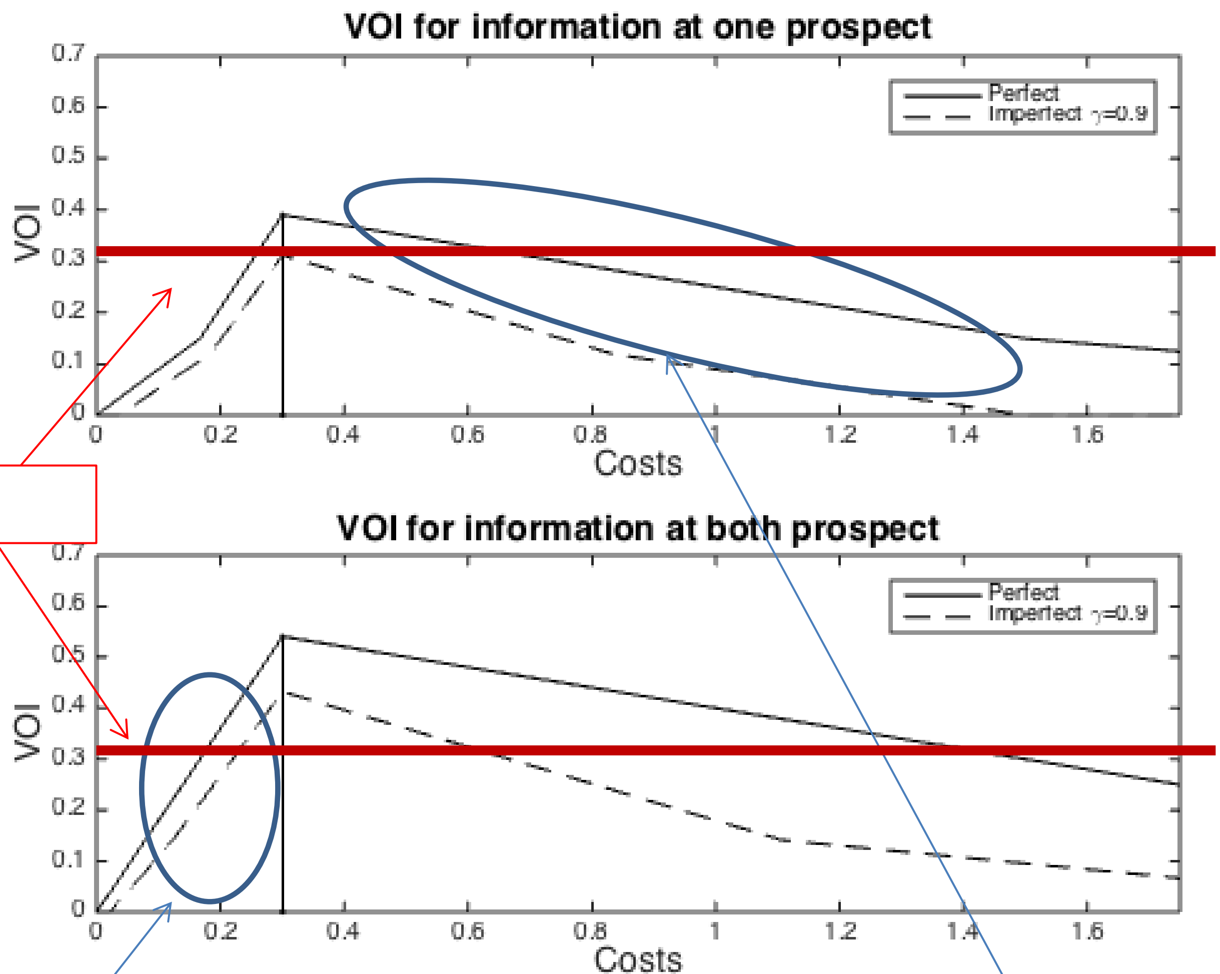
VOI for bivariate prospects example



Imperfect total better then partial perfect.

Partial perfect is better than imperfect total.

VOI for bivariate prospects example



Price of test is 0.3

Imperfect total better then partial perfect.

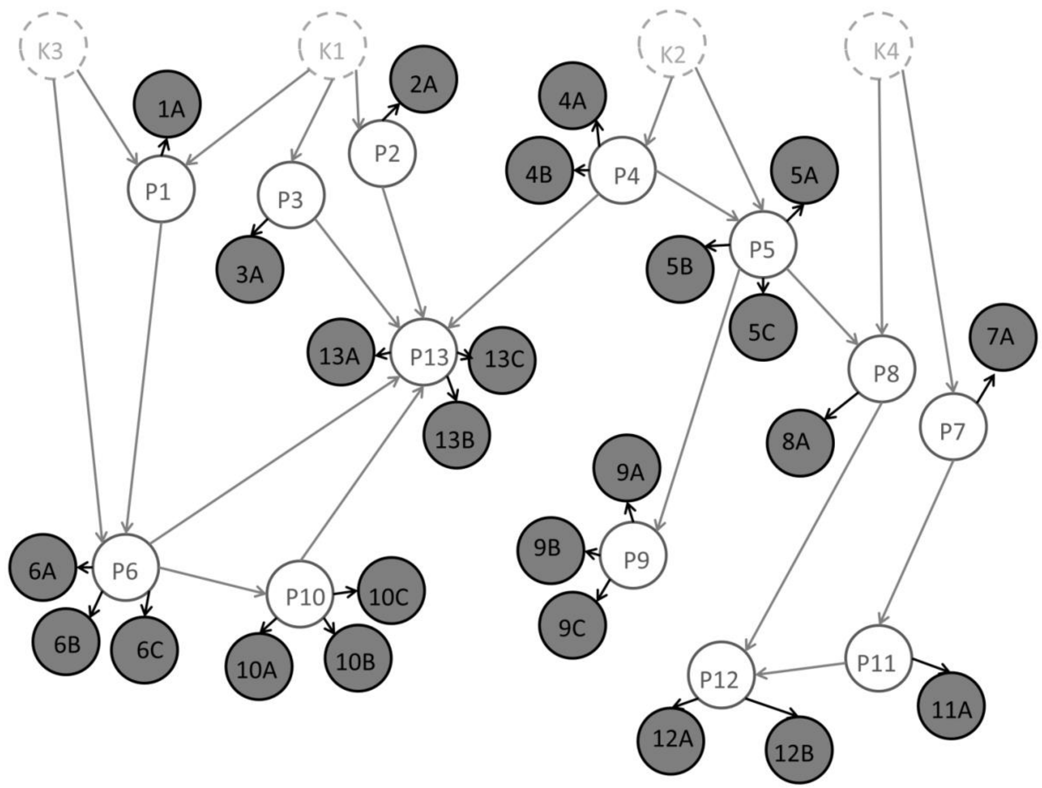
Partial perfect is better than imperfect total.

Insight in VOI – Bivariate prospects

- VOI of partial testing is always less than total testing, with same accuracy.
- Total imperfect test can give less VOI than a partial perfect test. Difference depends on the accuracy, prior mean for values, and correlation in spatial model.
- VOI is small for low costs (easy to start development) and for high cost (easy to avoid development). We do not need more data in these cases. We can make decisions right away.

Larger networks - computation

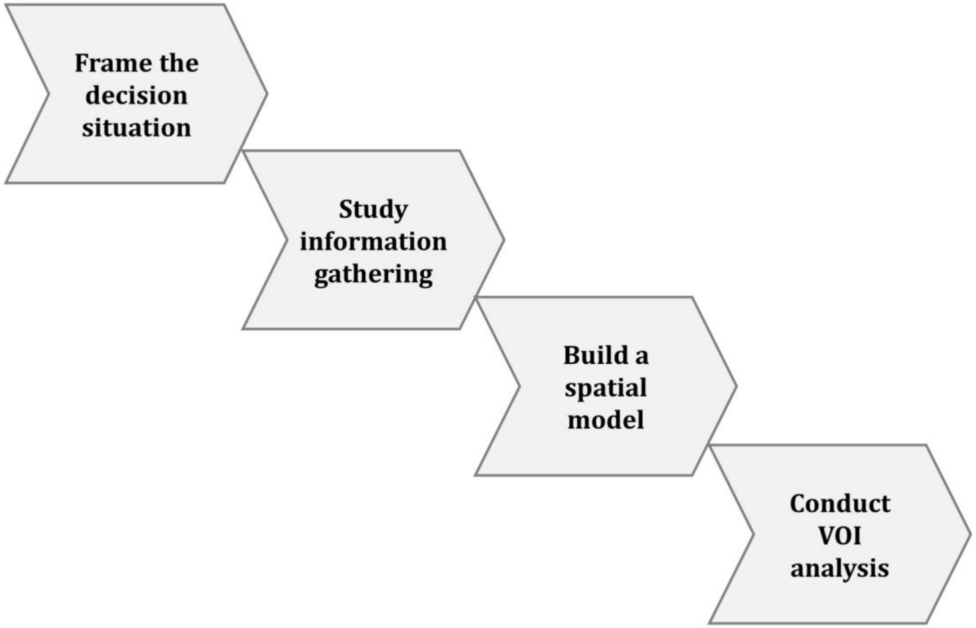
Algorithms have been developed for efficient marginalization, conditioning.



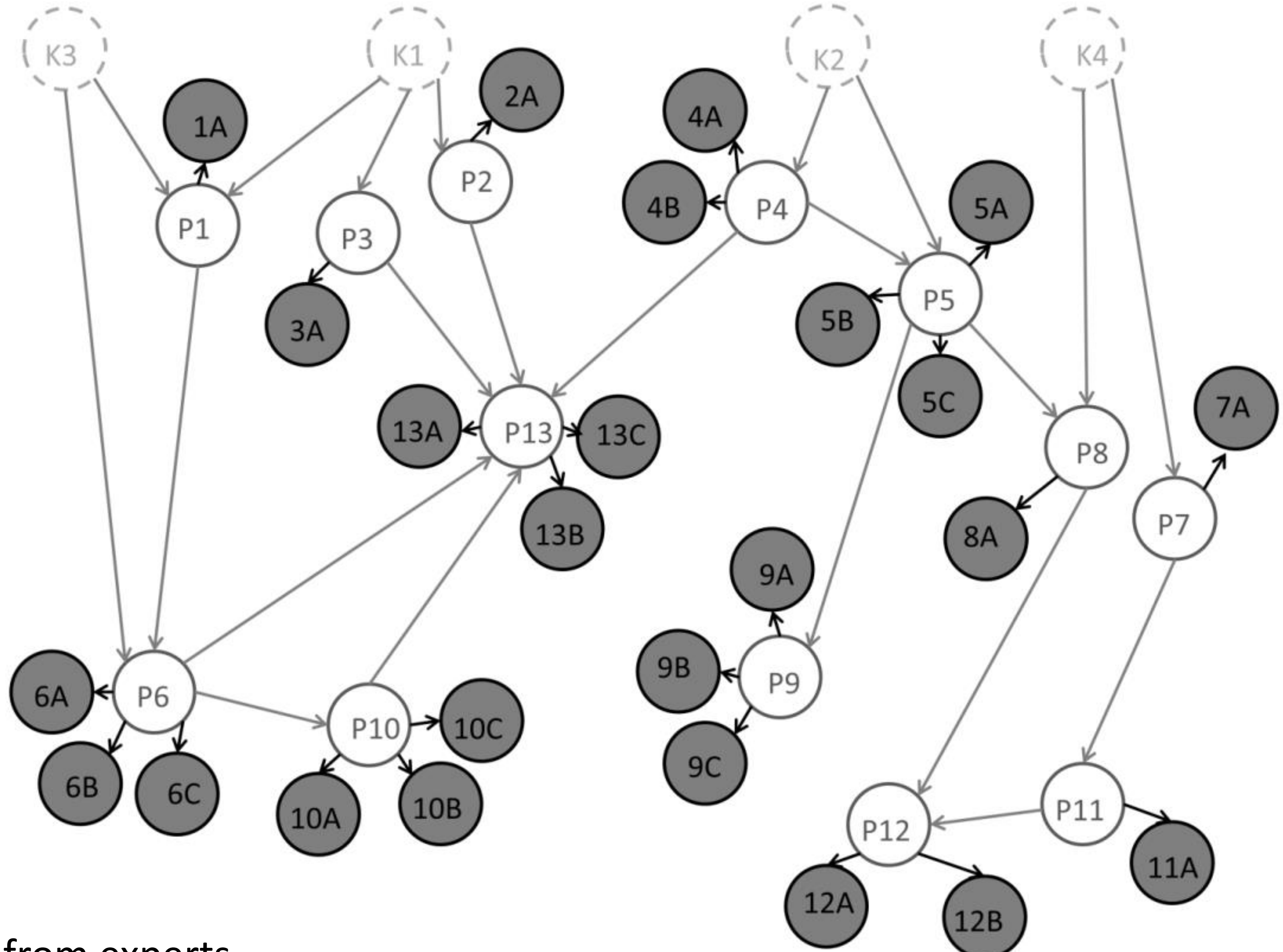
Martinelli, G., Eidsvik, J., Hauge, R., and Førland, M.D., 2011, Bayesian networks for prospect analysis in the North Sea, *AAPG Bulletin*, 95, 1423-1442.

VOI workflow

- Develop prospects separately. Shared costs for segments within one prospect.
- Gather information by exploration drilling. One or two wells. No opportunities for adaptive testing.
- Model is a Bayesian network model elicited from expert geologists in this area.
- VOI analysis done by exact computations for Bayesian networks (Junction tree algorithm – efficient marginalization and conditioning).



Bayesian network , Kitchens

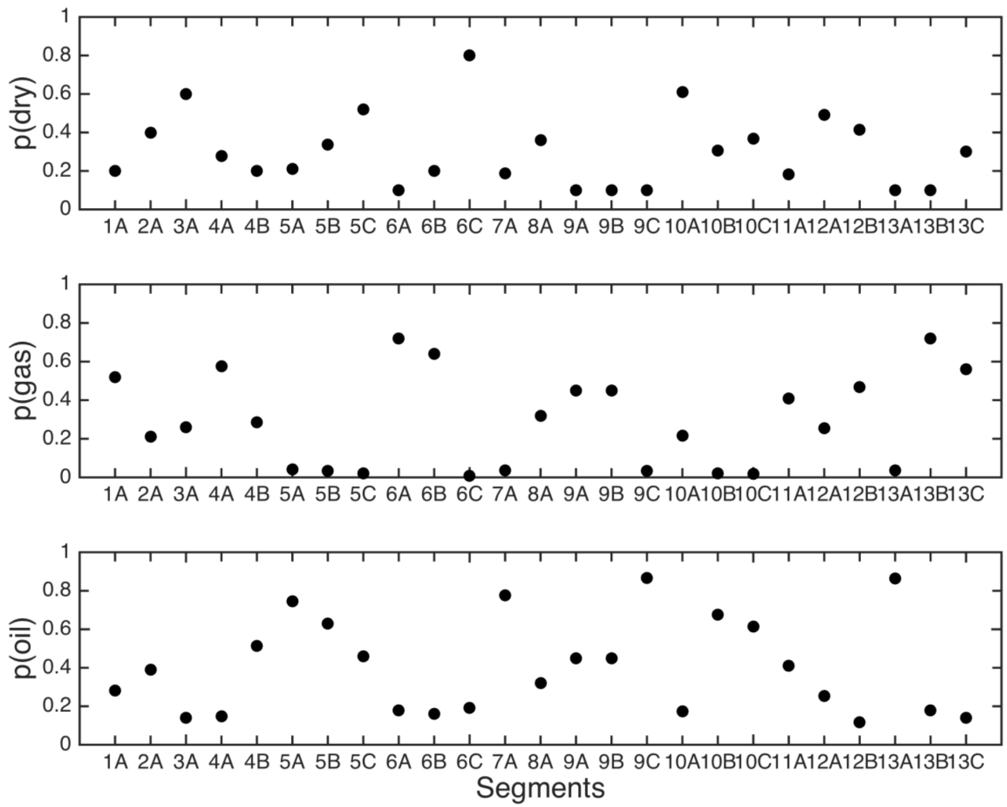


Model elicited from experts.
Migration from kitchens.
Local failure probability of migration.

Prior marginal probabilities

Three possible classes at all nodes:

- Dry
- Gas
- Oil



Prior values

Development fixed cost. Infrastructure at prospect r.

$$PV = \sum_{r=1}^{13} \max\left\{0, \sum_{i \in \mathrm{Pr}} IV(x_i) - DFC\right\}$$

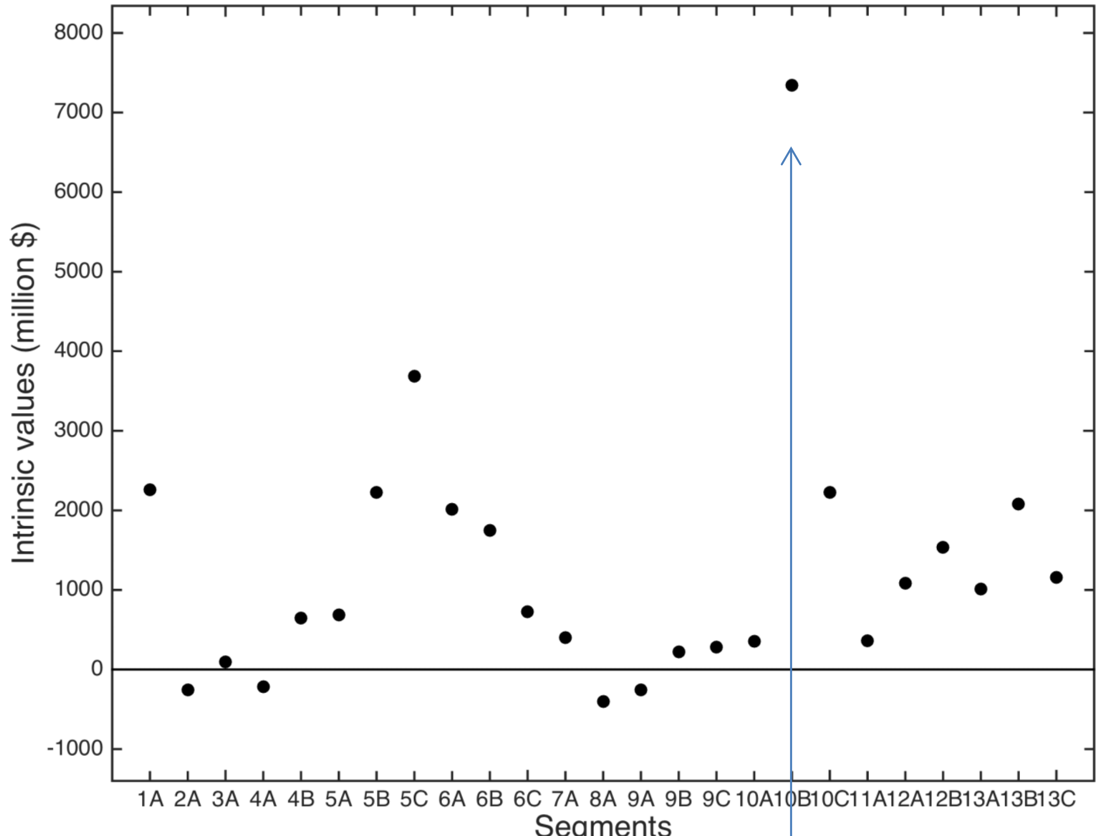
$$IV(x_i) = \sum_{k=1}^{3} \Big(\mathrm{Rev}_{i,k}\, p(x_i = k) - \mathrm{Cost}_{i,k}\, p(x_i = k)\Big) - \mathrm{Cost}_{i,0}$$

Revenues of oil/gas, 0 otherwise.

Cost if dry, 0 otherwise.

Cost of drilling segment i.

Values



Most lucrative. But might not be most informative.

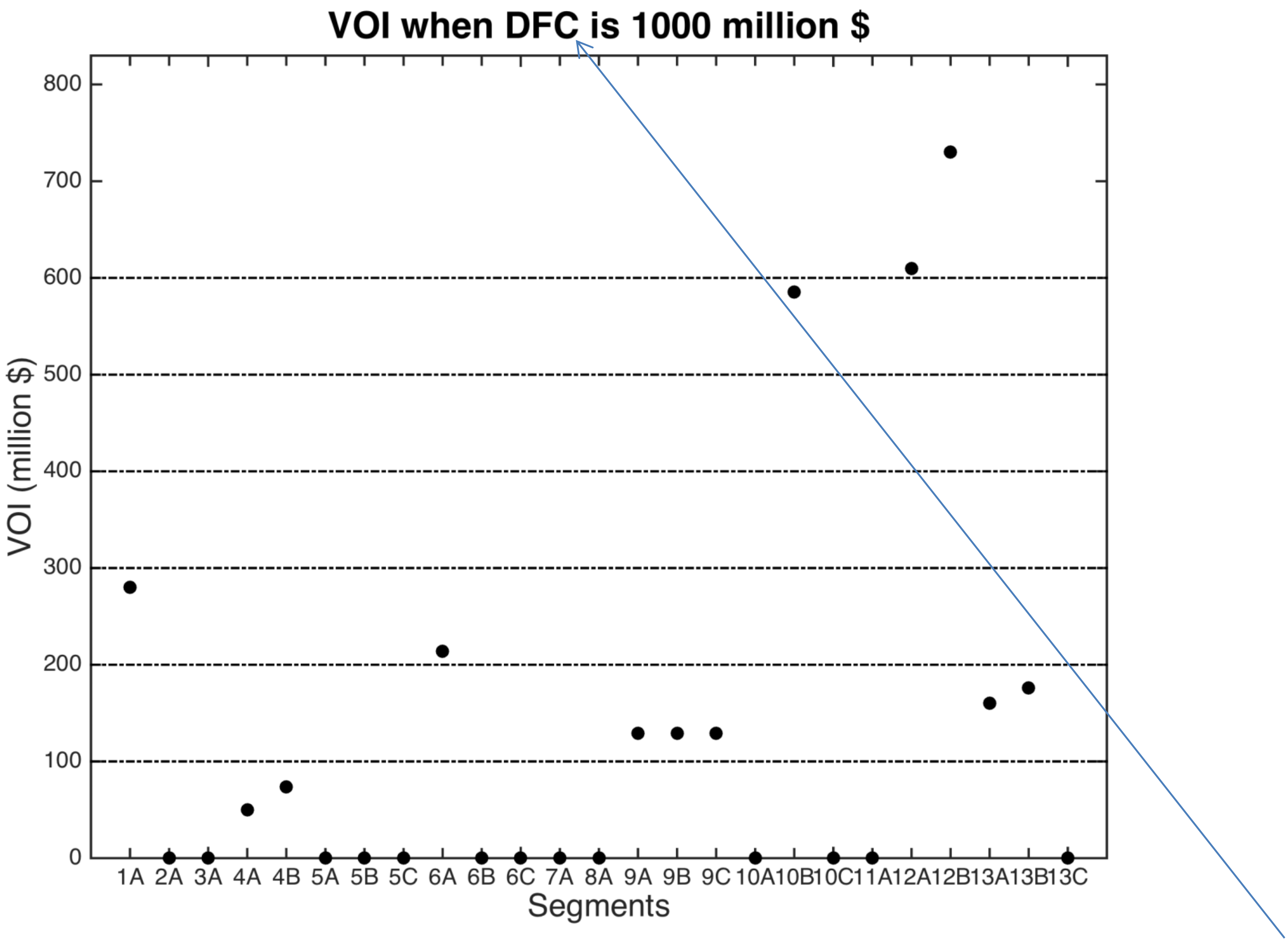
Posterior values and VOI

$$PoV\left(x_{\mathbb{K}}\right)=\sum_{l=1}^{3}\sum_{r=1}^{13}\max\left\{0,\sum_{i\in\Pr}IV\left(x_i \mid x_{\mathbb{K}}=l\right)-DFC\right\}p\left(x_{\mathbb{K}}=l\right)$$

$$VOI\left(x_{\mathbb{K}}\right)=PoV\left(x_{\mathbb{K}}\right)-PV$$

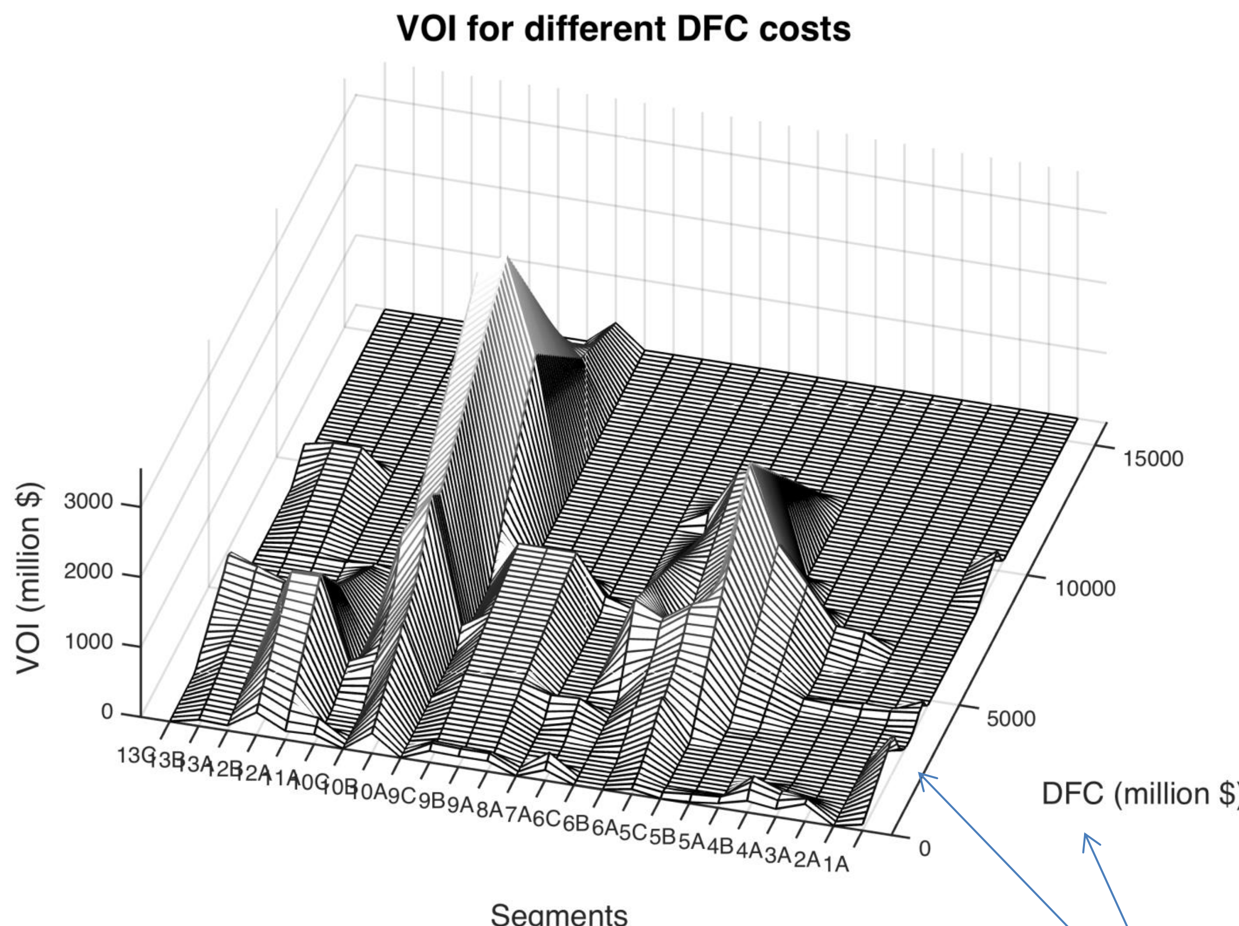
Data acquired at single well.

VOI single wells



Development fixed cost.

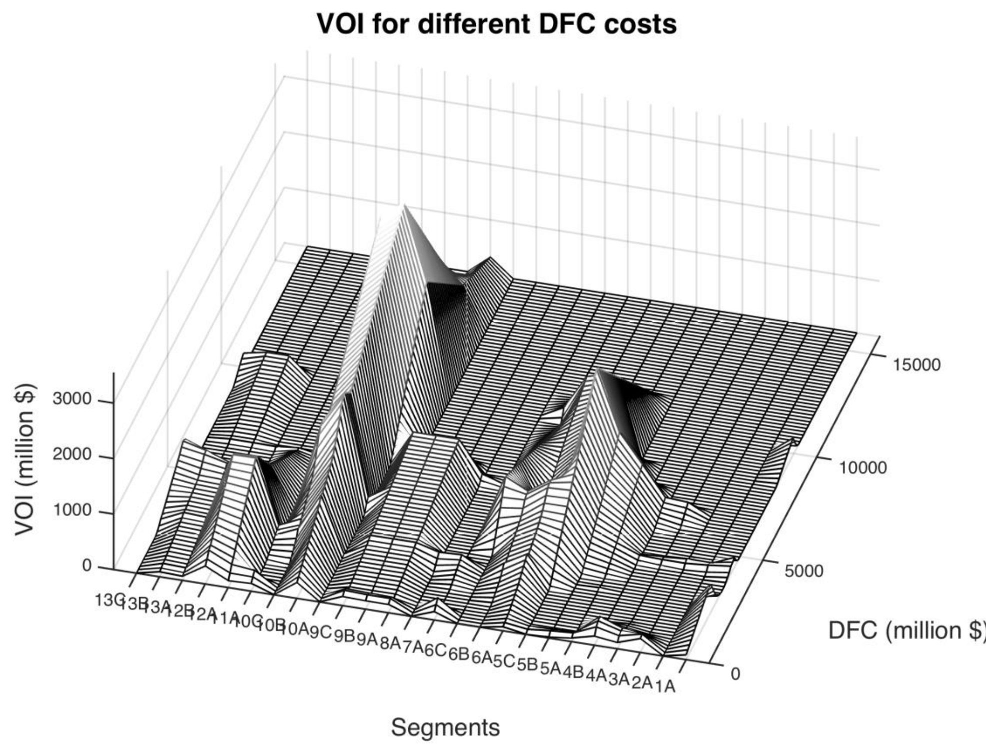
VOI for different costs



Development fixed cost.

VOI for different costs

- For each segment VOI starts at 0 (for small costs), grows to larger values, and decreases to 0 (for large costs).
- VOI is smooth for segments belonging to the same prospect. Correlation and shared costs.
- VOI can be multimodal as a function of cost, because the information influences neighboring segments, at which we are indifferent at other costs.



Take home from this exercise:

- VOI is not largest at the most lucrative prospects.
- VOI is largest where more data are likely to help us make better decisions.
- VOI also depends on whether the data gathering can influence neighboring segments – data propagate in the Bayesian network model.

- Compare with price? Or compare different data gathering opportunities, and provide a basis for discussion.

Never break the chain - Markov models

Markov chains are special graphs, defined by initial probabilities and transition matrices.

$$p(\boldsymbol{x}) = p(x_1, x_2, \ldots, x_n) = p(x_1)\, p(x_2 \mid x_1) \ldots p(x_n \mid x_{n-1})$$

$$p(x_1 = k), \qquad k = 1, \ldots, d$$

$$p(x_{i+1} = l \mid x_i = k) = P(k, l), \qquad k, l = 1, \ldots, d$$

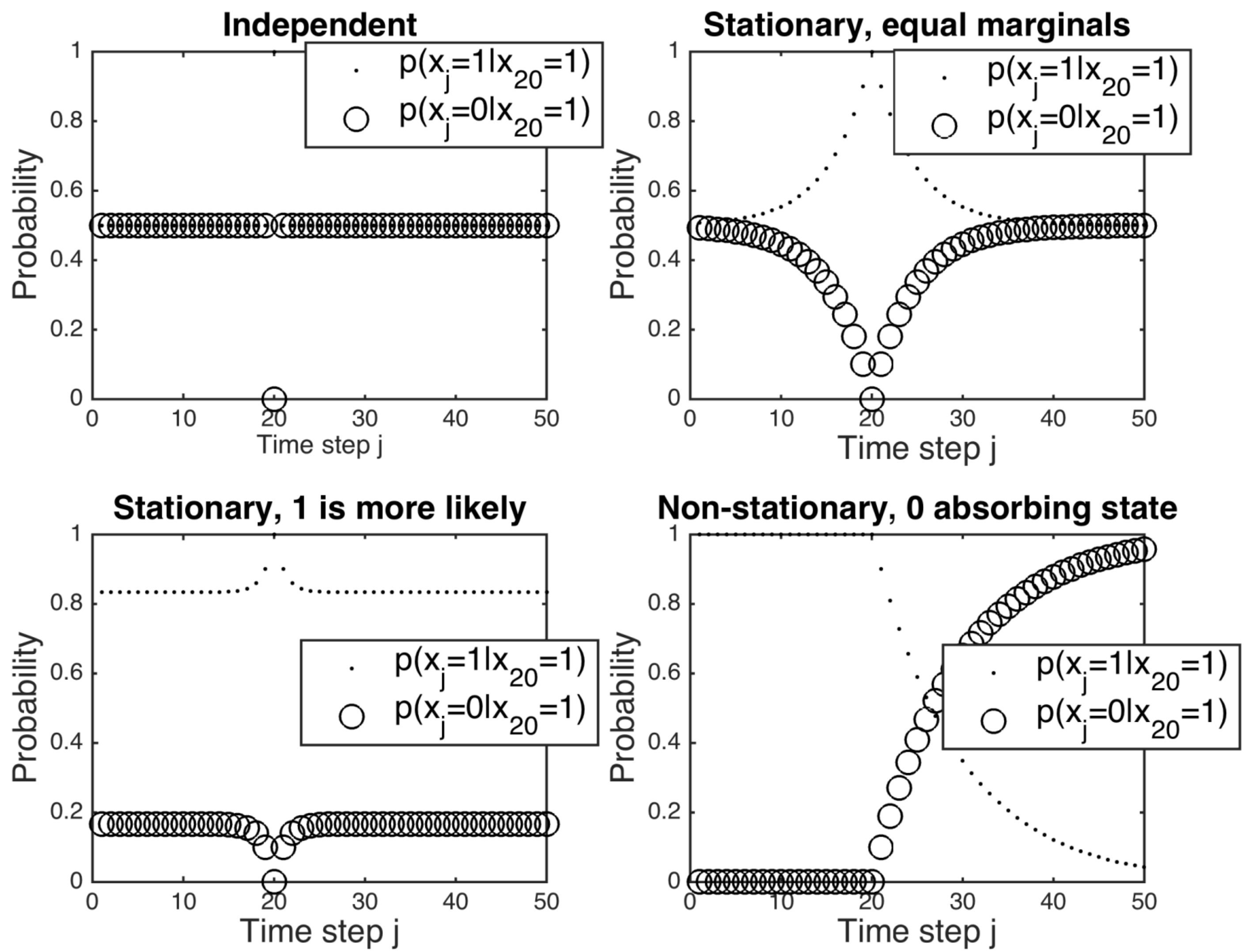
$d = 2$

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \qquad P = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \qquad P = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix} \qquad P = \begin{bmatrix} 1 & 0 \\ 0.1 & 0.9 \end{bmatrix}$$

Independence

Absorbing

Markov chains (given perfect information)



Never break the chain - Hidden Markov model

Latent variable is a Markov chain. This forms the prior model.
Data are measured, imperfectly, at all or some of the variables.

- Commonly used for speech recognition, translation, cleaning signals, etc.

Modeling by conditional probability

Prior:

$$p(\boldsymbol{x}) = p(x_1)\,p(x_2 \mid x_1)\ldots p(x_n \mid x_{n-1},\ldots,x_1)$$

Likelihood:

$$p(\boldsymbol{y} \mid \boldsymbol{x}) = \prod_j p(y_j \mid x_j)$$

Conditionally independent data. Data measures the local properties:
When the latent variable at that location is known, there is nothing to add by knowing other variables.

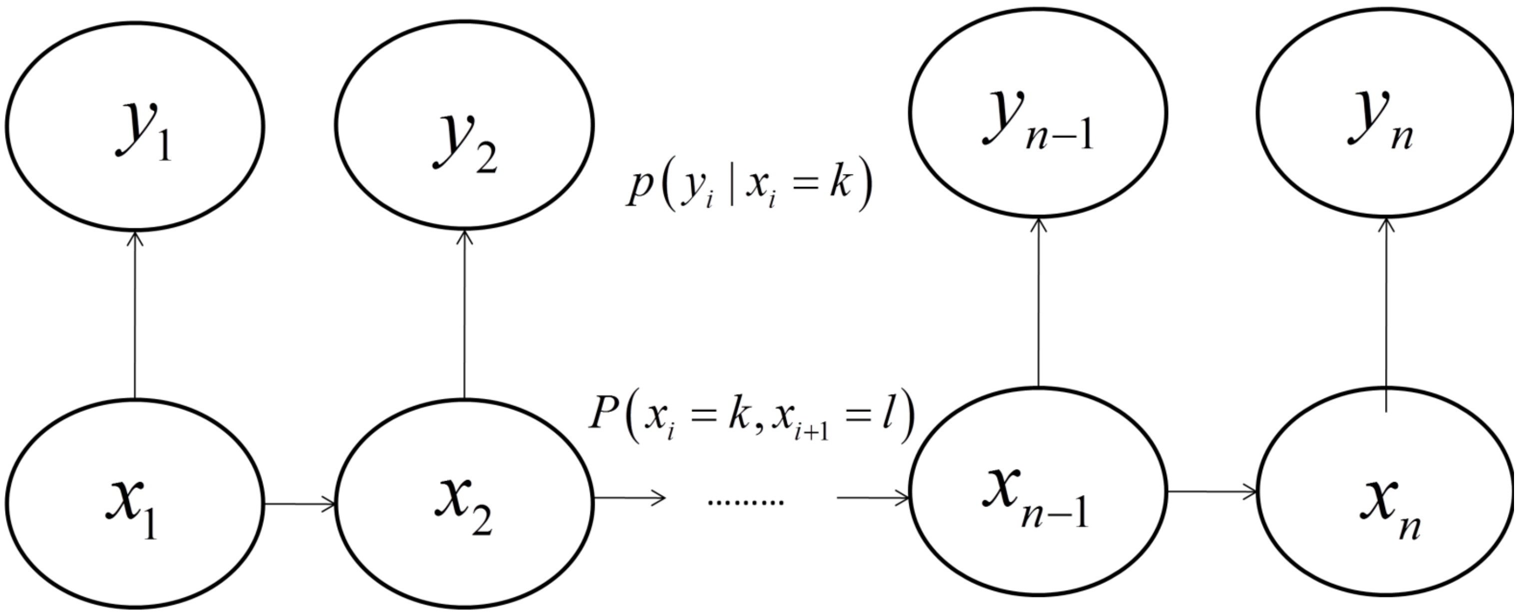
Hidden Markov model

Goal is:

$$p(\boldsymbol{x} \mid \boldsymbol{y}) = \frac{p(\boldsymbol{y} \mid \boldsymbol{x})\, p(\boldsymbol{x})}{p(\boldsymbol{y})}$$

Hidden Markov model (HMM)

Imperfect observations.
Conditionally independent



Latent distinction of interest. Markov chain

Avalanche decisions and sensors

Suppose that parts along a road or railroad are at risk of **avalanche**.
- One can remove risk by cost.
- If it is not removed, the repair cost depends on the unknown risk class.

Data, typically putting out sensors, can help classify the risk class and hence improve the decisions made at different locations.

Avalanche decisions - risk analysis

n=50 identified locations, at risk of **avalanche**.
At every location one can remove risk by cost 10.
If it is not removed, the repair cost depends on the unknown risk class:

$$C_j, \qquad j \in \{1,2,3,4\},$$

$$\mathrm{C}_1 = 0, \mathrm{C}_2 = 5, \mathrm{C}_3 = 20, \mathrm{C}_4 = 40,$$

Decision maker can secure, or not, at each location. The decisions are based on the minimization of expected costs.

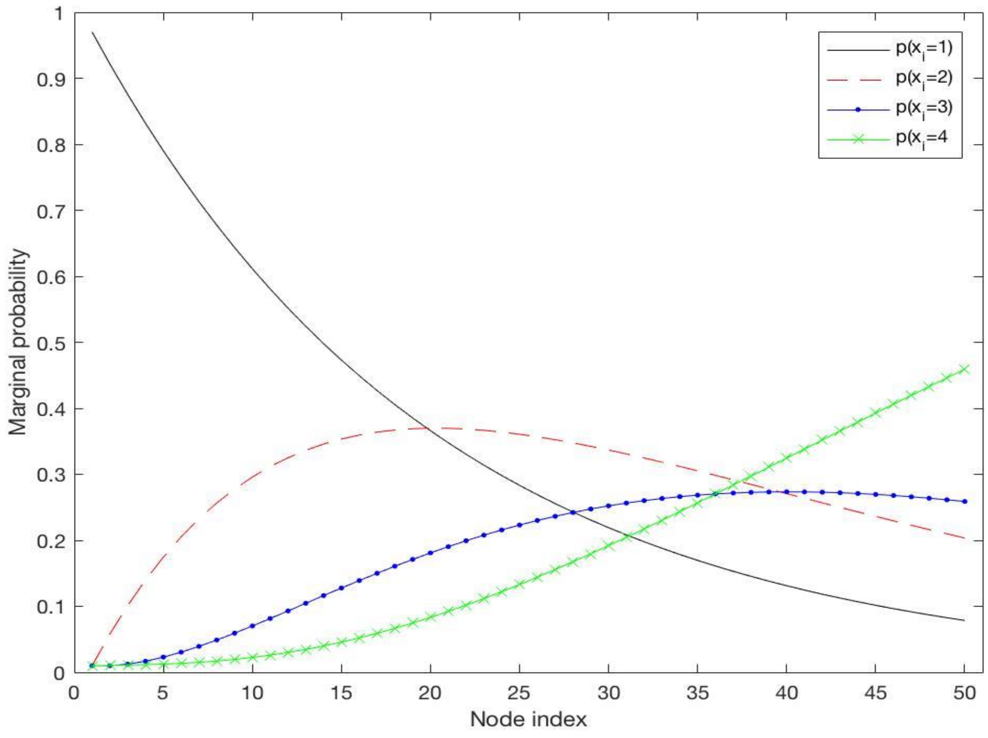
Prior value:

$$PV = \sum_{i=1}^{50} \max\left\{-10, -\sum_{j=1}^{4} C_j p\left(x_i = j\right)\right\}$$

Never break the chain - Prior model

$$x_i \in \{1,2,3,4\}, \qquad i = 1,\dots,50,$$
$$p(x_1 = 1) = 0.97,$$
$$p(x_1 = j) = 0.01, \qquad j = 2,3,4,$$

$$P = \begin{bmatrix} 0.95 & 0.05 & 0 & 0 \\ 0 & 0.95 & 0.05 & 0 \\ 0 & 0 & 0.95 & 0.05 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Never break the chain - Data model

Sensors at some or all of the n=50 identified locations.
Likelihood model:

$$p\left(y_i \mid x_i\right) = N\left(x_i, \tau_i^2\right), \qquad i = 1, \ldots, 50.$$

Posterior value:

$$PoV\left(\boldsymbol{y}\right) = \sum_{\boldsymbol{y}} \sum_{i=1}^{50} \max\left\{-10, -\sum_{j=1}^{4} C_j p\left(x_i = j \mid \boldsymbol{y}\right)\right\} p\left(\boldsymbol{y}\right)$$

Site is measured by sensor: $\tau_i^2 = 1$

Site is not measured by sensor: $\tau_i^2 = 1000^2$

This is a trick to make the forward-backward algorithm easier to implement.

Forward-backward algorithm

Recursive forward step (prediction and updating):

$$p\left(x_i = k \mid y_1,\ldots,y_{i-1}\right) = \sum_{j=1}^{d} p\left(x_{i-1} = j, x_i = k \mid y_1,\ldots,y_{i-1}\right)$$

$$= \sum_{j=1}^{d} p\left(x_i = k \mid x_{i-1} = j\right) p\left(x_{i-1} = j \mid y_1,\ldots,y_{i-1}\right),$$

$$p\left(x_i = k \mid y_1,\ldots,y_i\right) = \frac{p\left(x_i = k, y_i \mid y_1,\ldots,y_{i-1}\right)}{p\left(y_i \mid y_1,\ldots,y_{i-1}\right)}$$

$$= \frac{p\left(y_i \mid x_i = k\right) p\left(x_i = k \mid y_1,\ldots,y_{i-1}\right)}{p\left(y_i \mid y_1,\ldots,y_{i-1}\right)},$$

$d = 4$

Forward-backward algorithm

Recursive backward step :

$$p\left(x_i = k \mid y_1, \ldots, y_n\right) = \sum_{l=1}^{d} p\left(x_i = k \mid y_1, \ldots, y_{i+1}, x_{i+1} = l\right) p\left(x_{i+1} = l \mid y_1, \ldots, y_n\right)$$

$$p\left(x_i = k \mid y_1, \ldots, y_{i+1}, x_{i+1}\right) = \frac{p\left(x_i = k, x_{i+1}, y_{i+1} \mid y_1, \ldots, y_i\right)}{p\left(y_{i+1} \mid x_{i+1}, y_1, \ldots, y_i\right) p\left(x_{i+1} \mid y_1, \ldots, y_i\right)}$$

$$= \frac{p\left(x_{i+1} \mid x_i = k\right) p\left(x_i = k \mid y_1, \ldots, y_i\right)}{p\left(x_{i+1} \mid y_1, \ldots, y_i\right)}.$$

$d = 4$

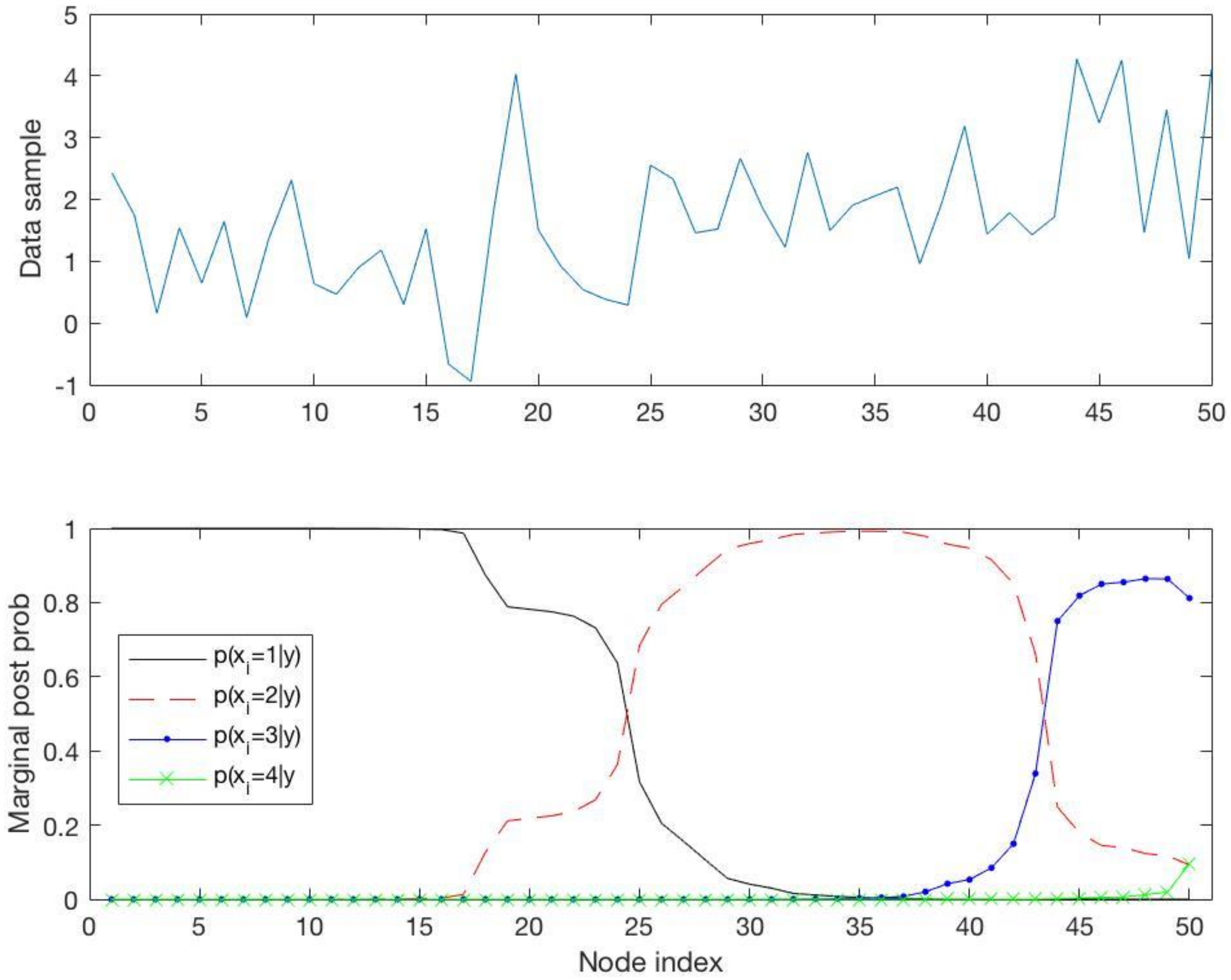
Monte Carlo approximation for VOI

$$PoV(\boldsymbol{y}) = \frac{1}{B}\sum_{b=1}^{B}\sum_{i=1}^{50}\max\left\{-10, -\sum_{j=1}^{4} C_j p\left(x_i = j \mid \boldsymbol{y}^b\right)\right\}$$

1. Sample B Markov chain variables, and next sample B data vectors.
   - The samples will be from the marginal distribution for the data.
2. Compute the marginal posterior probabilities by forward-backward algorithm for each data sample.

Average the results to approximate posterior value and VOI.

Results – one realization



Results – different tests

All sites	Only 10 first	Only 11-20	Only 21-30	Only 31-40	Only 41-50
126	36	69	87	91	82

Only every second (5 measurements) gives VOI=83.

Partial tests can be very valuable! Especially if they are done in interesting subsets of the domain.

Plan for course

Time	Topic
Monday	Introduction and motivating examples
	Elementary decision analysis and the value of information
Tuesday	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for dependent models
Wednesday	Spatial statistics, spatial design of experiments
	Value of information analysis in spatial decision situations
Thursday	Examples of value of information analysis in Earth sciences
	Computational aspects
Friday	Sequential decisions and sequential information gathering
	Examples from mining and oceanography

Every day: Small exercise half-way, and computer project at the end.

Project 2 : Markovian risk of avalanche

Implement a Markov chain example for avalanche risk 1 or 2, with VOI analysis.
At every location one can remove risk by cost 10. If it is not removed, the repair cost depends on the unknown risk class:

$$C_1 = 0, C_2 = 20,$$

- Compute marginal probabilities for the following Markov chain with initial state probability and transition probability:

$$x_i \in \{1,2\}, \qquad i = 1,\ldots,50,$$

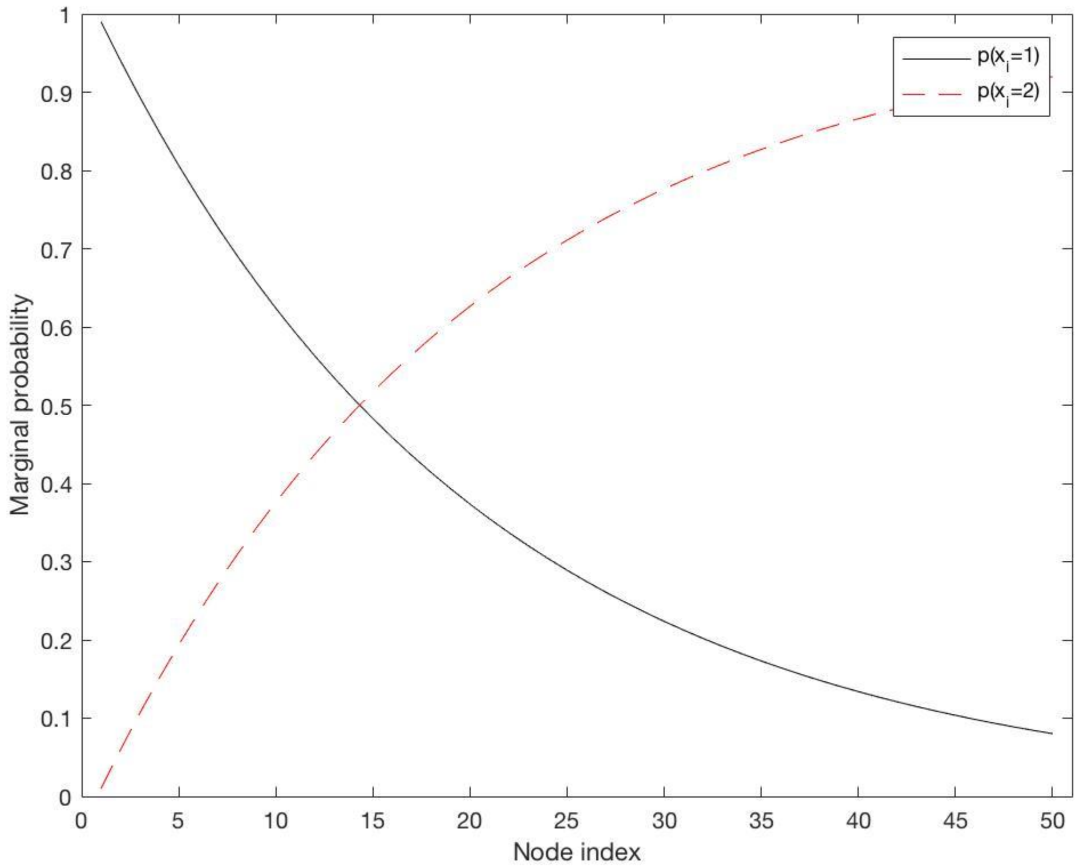
$$p(x_1 = 1) = 0.99, \qquad P = \begin{bmatrix} 0.95 & 0.05 \\ 0 & 1 \end{bmatrix}$$

$$p(x_1 = 2) = 0.01,$$

- Condition on risk-class 1 or 2 (perfect information) at node 20. Compute conditional probabilities. Compare with marginal probabilities.
- Compute the prior value. Compute the posterior value and VOI for different single locations **perfect** observations. What is the best place to survey?

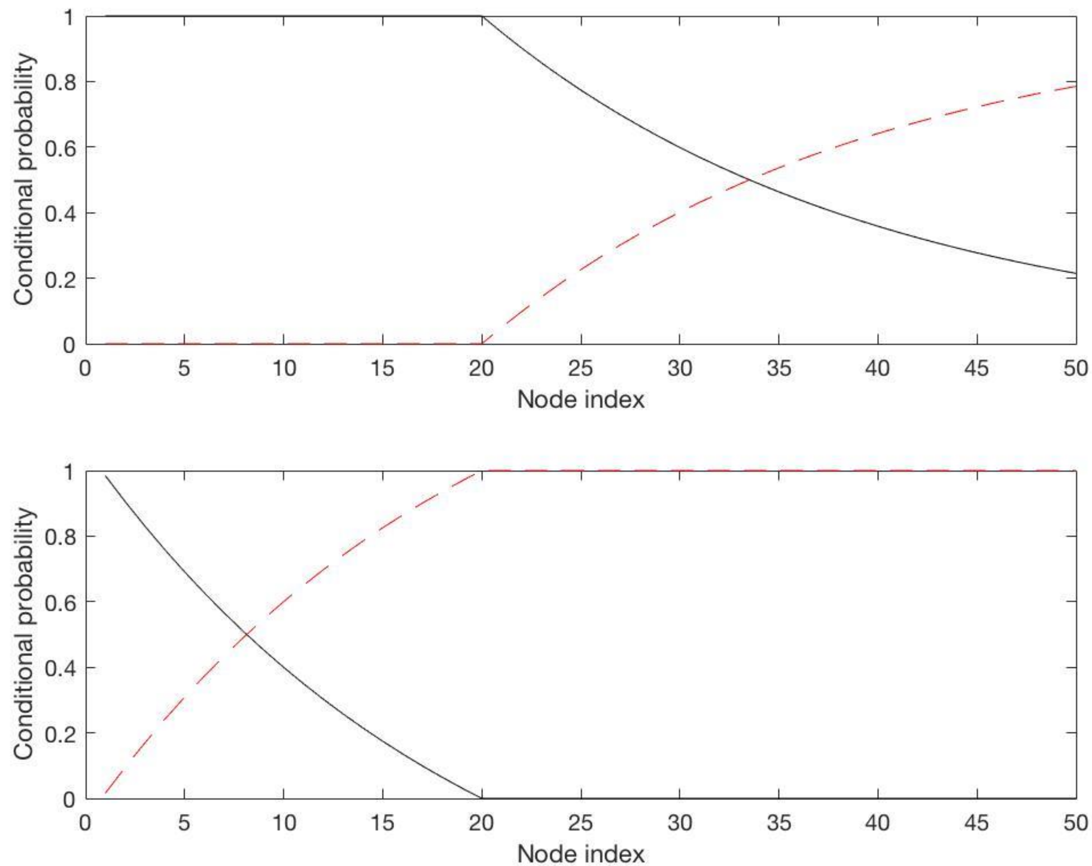
Results – marginals

$$p\left(x_{i+1}=l\right)=\sum_{k=1}^{2} p\left(x_i=k\right) p(x_{i+1}=l \mid x_i=k), \qquad l=1,2, \qquad i=1,\ldots,n-1$$



Results – conditionals (forward)

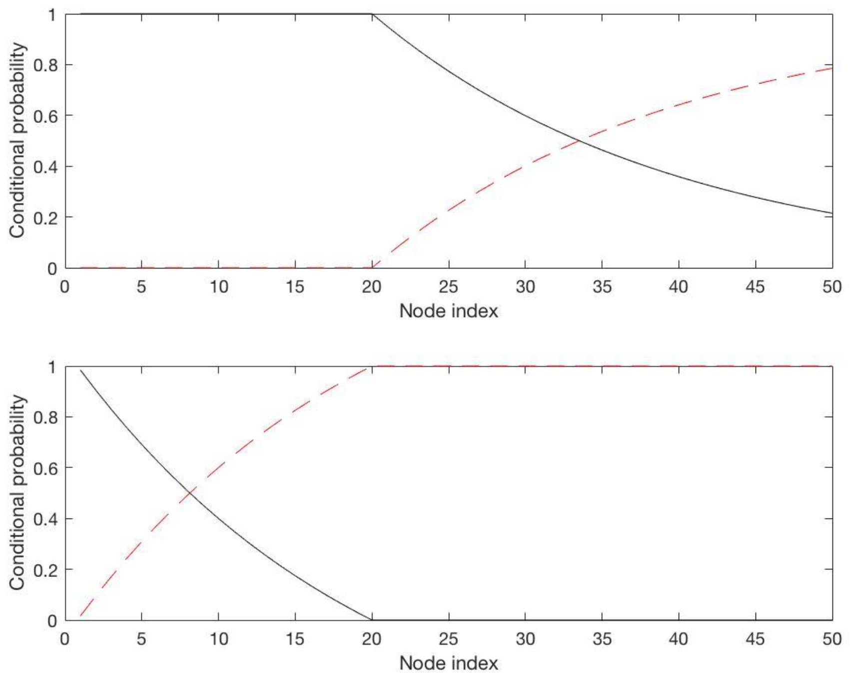
$$p\left(x_i = k \mid x_j = l\right) = \sum_{q=1}^{2} p\left(x_i = k, x_{i-1} = q \mid x_j = l\right) = \sum_{q=1}^{2} P(q,k)\, p\left(x_{i-1} = q \mid x_j = l\right)$$



Results – conditionals (backward)

$$p\left(x_i = k \mid x_{i+1},\ldots,x_n\right) = p\left(x_i = k \mid x_{i+1} = l\right) = \frac{p\left(x_i = k, x_{i+1} = l\right)}{p\left(x_{i+1} = l\right)} = \frac{P\left(k,l\right)p\left(x_i = k\right)}{p\left(x_{i+1} = l\right)}$$

$$p\left(x_i = k \mid x_j = l\right) = \sum_{q=1}^{2} p\left(x_i = k \mid x_{i+1} = q\right) p\left(x_{i+1} = q \mid x_j = l\right)$$



Results – VOI

$$PV = \sum_{i=1}^{50} \max\left\{-\mathrm{Cost}, -\mathrm{Cost1}\cdot p\left(x_i = 1\right) - \mathrm{Cost2}\cdot p\left(x_i = 2\right)\right\}$$

$$PoV\left(x_j\right) = \sum_{l=1}^{2}\sum_{i=1}^{50} \max\left\{-\mathrm{Cost}, -\mathrm{Cost1}\cdot p\left(x_i = 1 \mid x_j = l\right) - \mathrm{Cost2}\cdot p\left(x_i = 2 \mid x_j = l\right)\right\} p\left(x_j = l\right)$$

Cost=10,

Cost1=0,

Cost2=20