Value of Information in the Earth Sciences

Jo Eidsvik,
Department of Mathematical Sciences, NTNU, NORWAY

My background:

Professor of Statistics at NTNU in Trondheim, NORWAY.

Education:

- MSc in Applied Mathematics, Univ of Oslo
- PhD in Statistics, NTNU

Work experience:

- Norwegian Defense Research Establishment
- Statoil (Equinor)

Research interests:

- Spatio-temporal statistics,
- Computational statistics,
- Geosciences applications
- Design of experiments,
- Decision analysis, value of information,

Plan for course

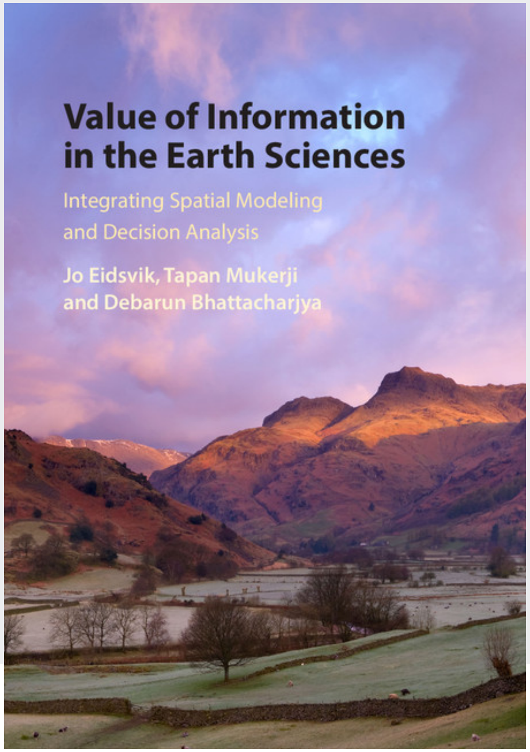
Time	Topic
Monday	**Introduction and motivating examples**
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Tuesday	Examples of value of information analysis in Earth sciences
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.

Material

Relevant background reading :

- Eidsvik, J., Mukerji, T. and Bhattacharjya, D., Value of information in the Earth sciences, Cambridge University Press, 2015.



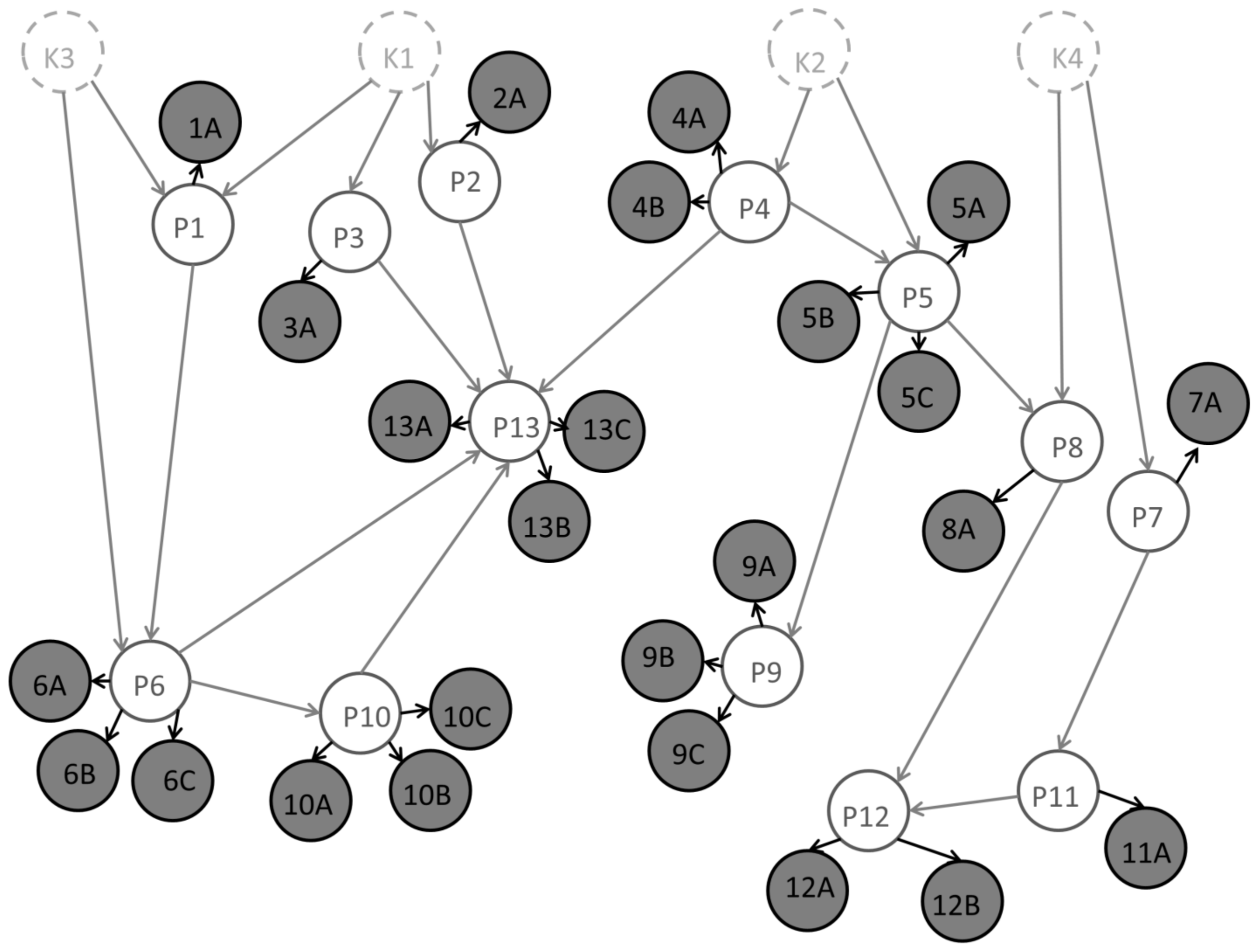
- Howard R.A. and Abbas, A.E., Foundations of decision analysis, Pearson, 2015.
- Many spatial statistics books:
  - Cressie and Wikle (2011),
  Chiles and Delfiner (2012),
  Banerjee et al. (2014),
  Pyrcz and Deutsch (2014),
  etc.

Motivating VOI examples:

- Integration of spatial modeling and decision analysis.
- Collect data to resolve uncertainties and make informed decisions.

Motivation
(a petroleum exploration example)

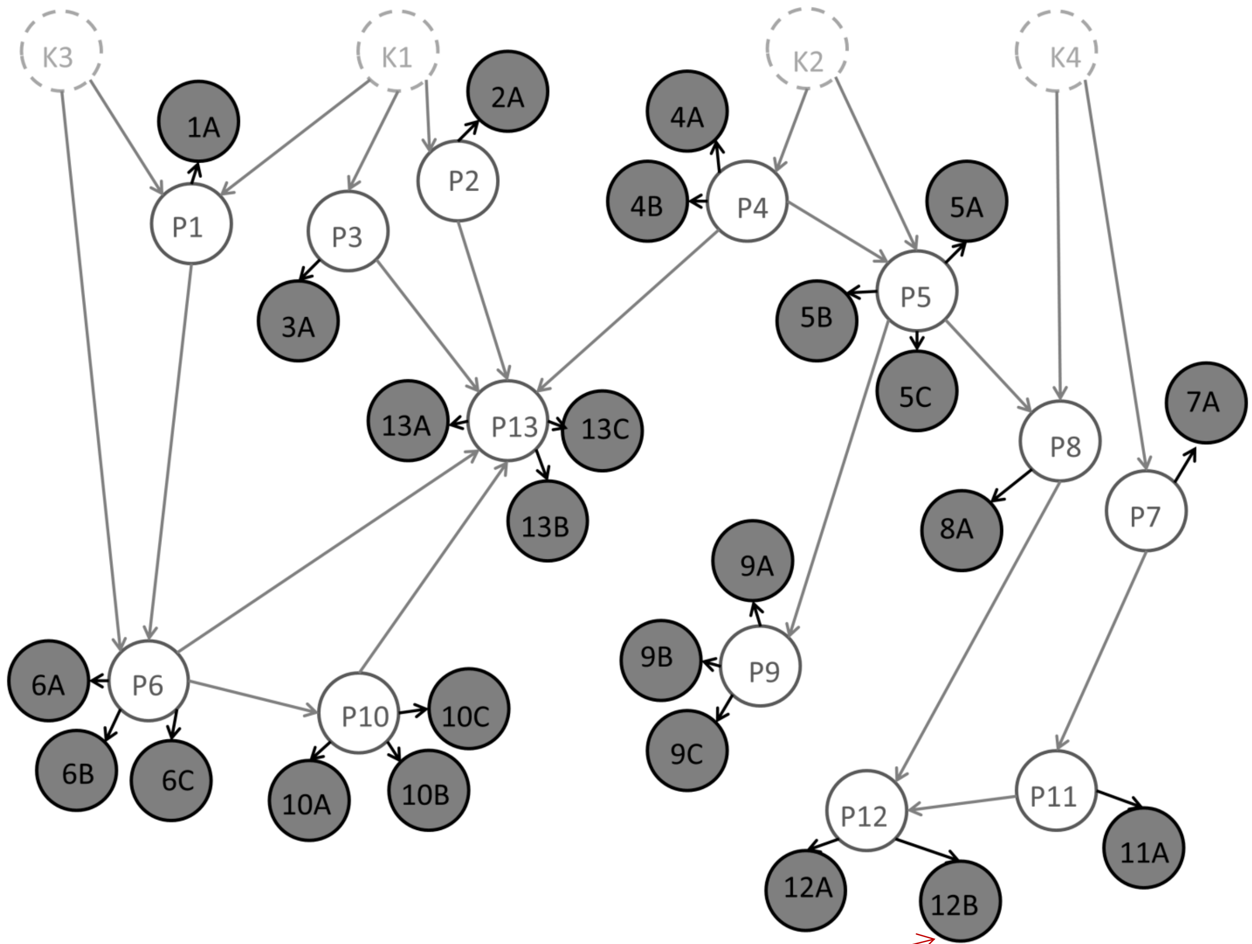
Gray nodes are petroleum reservoir segments where the company aims to develop profitable amounts of oil and gas.



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

Motivation
(a petroleum exploration example)

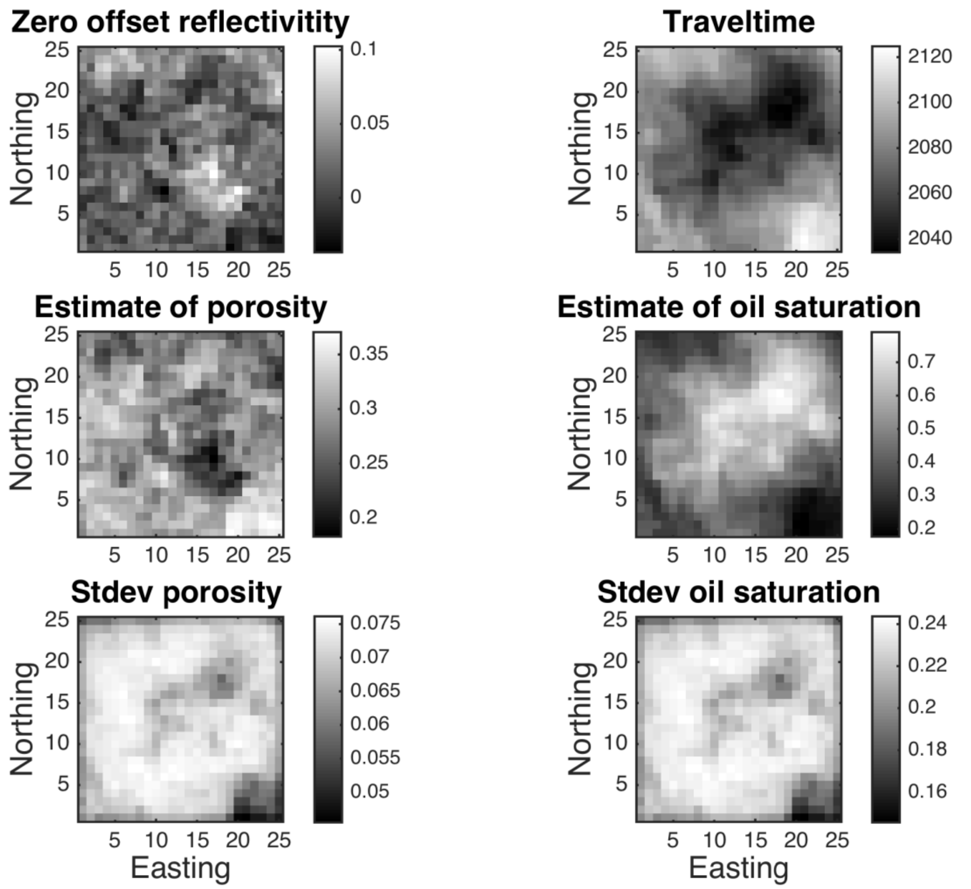
Gray nodes are petroleum reservoir segments where the company aims to develop profitable amounts of oil and gas.



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

Motivation
(a petroleum development example)

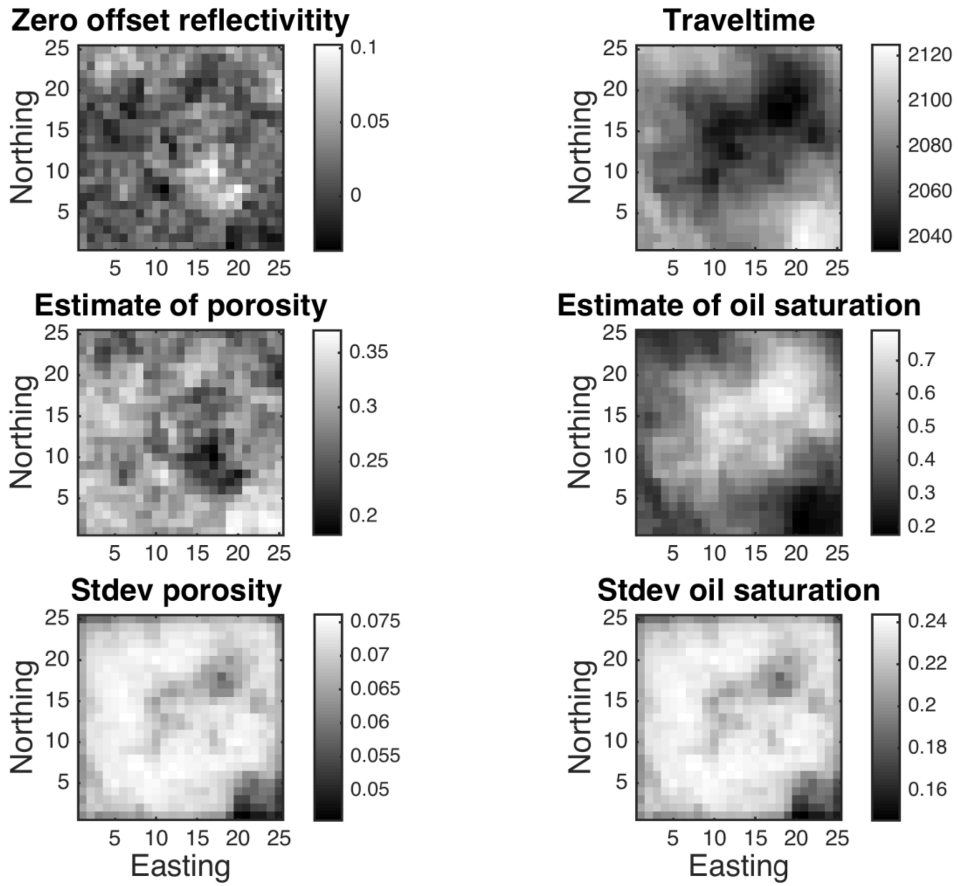
Reservoir predictions from post-stack seismic data!



Eidsvik, J., Bhattacharjya, D. and Mukerji, T., 2008, Value of information of seismic amplitude and CSEM resistivity, *Geophysics*.

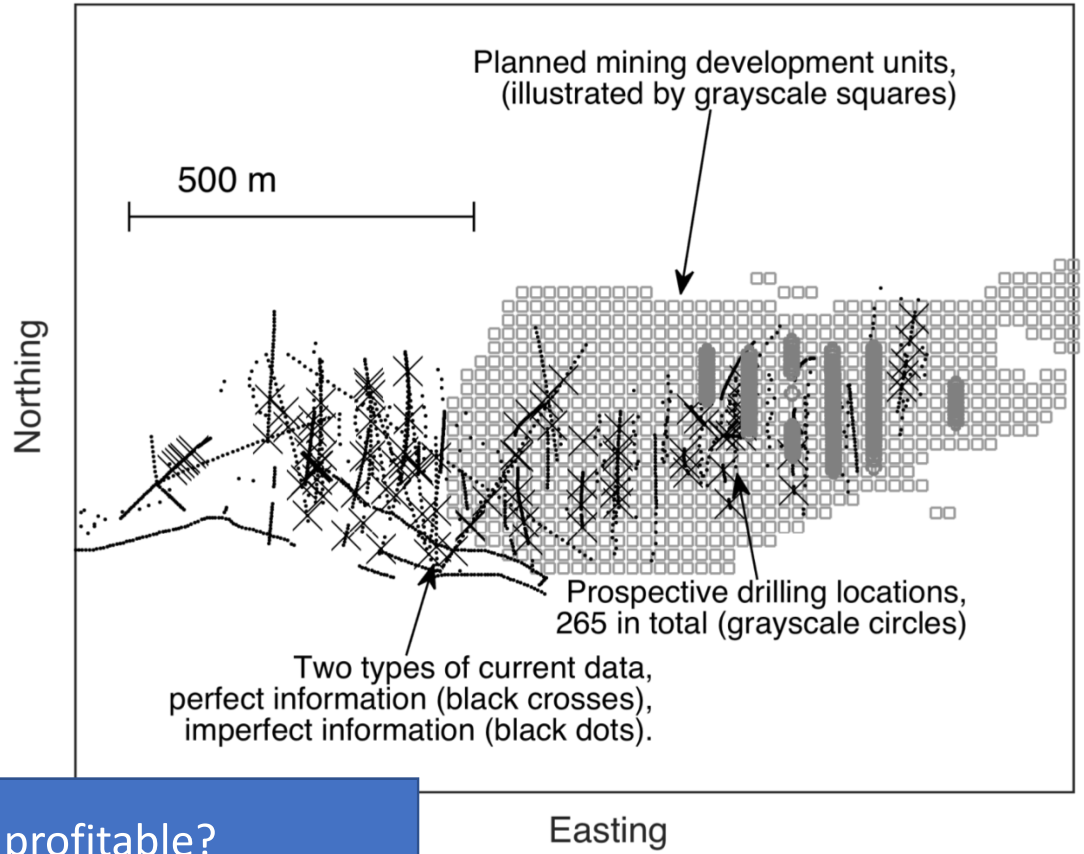
Motivation
(a petroleum development example)

Reservoir predictions from post-stack seismic data!



Process pre-stack seismic data, or electromagnetic data?

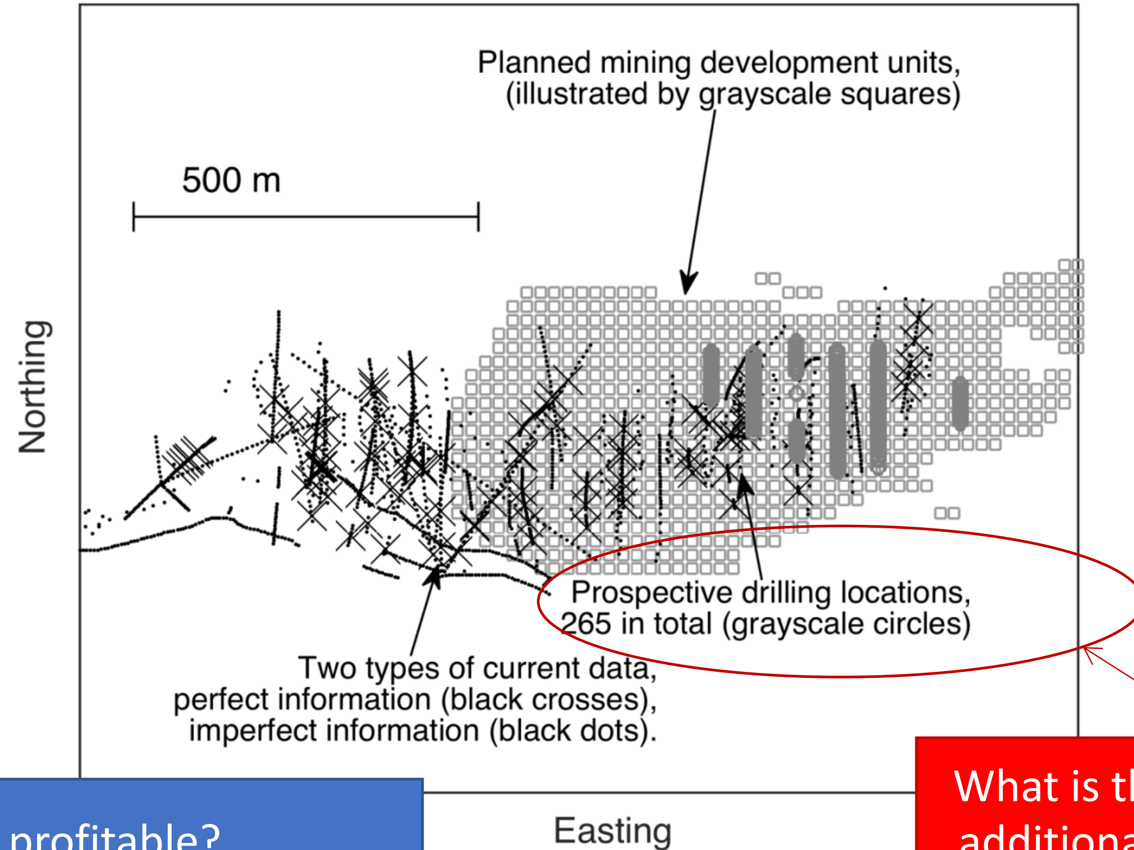
Motivation (an oxide mining example)



Is mining profitable?

Eidsvik, J. and Ellefmo, S.L., 2013, The value of information in mineral exploration within a multi-Gaussian framework, *Mathematical Geosciences*.

Motivation (an oxide mining example)

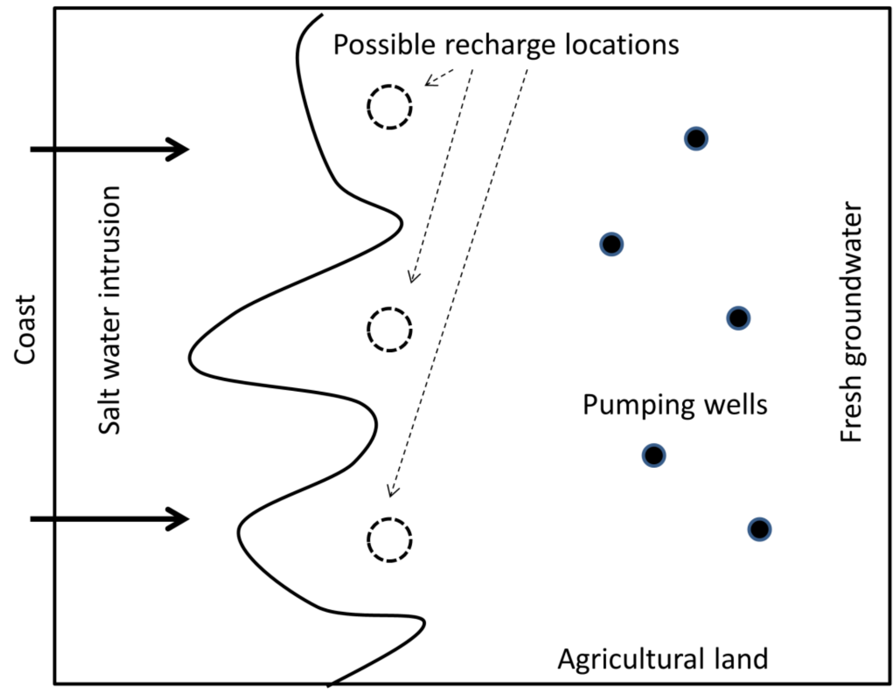


Is mining profitable?

What is the value of this additional information?

Motivation
(a groundwater example)

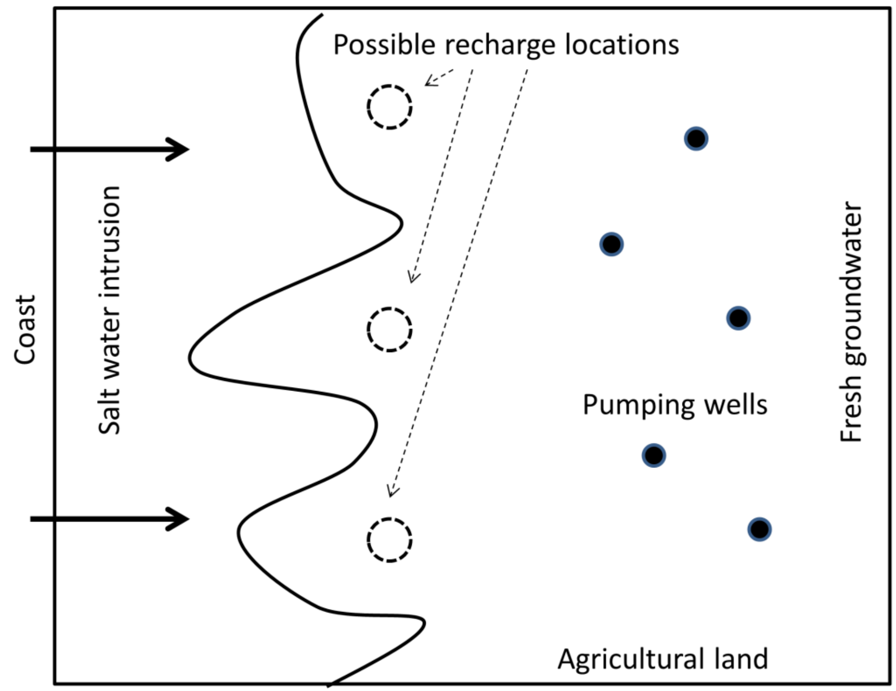
Which recharge location is better to prevent salt water intrusion?



Trainor-Guitton, W.J., Caers, J. and Mukerji, T., 2011, A methodology for establishing a data reliability measure for value of spatial information problems, *Mathematical Geosciences*.

Motivation
(a groundwater example)

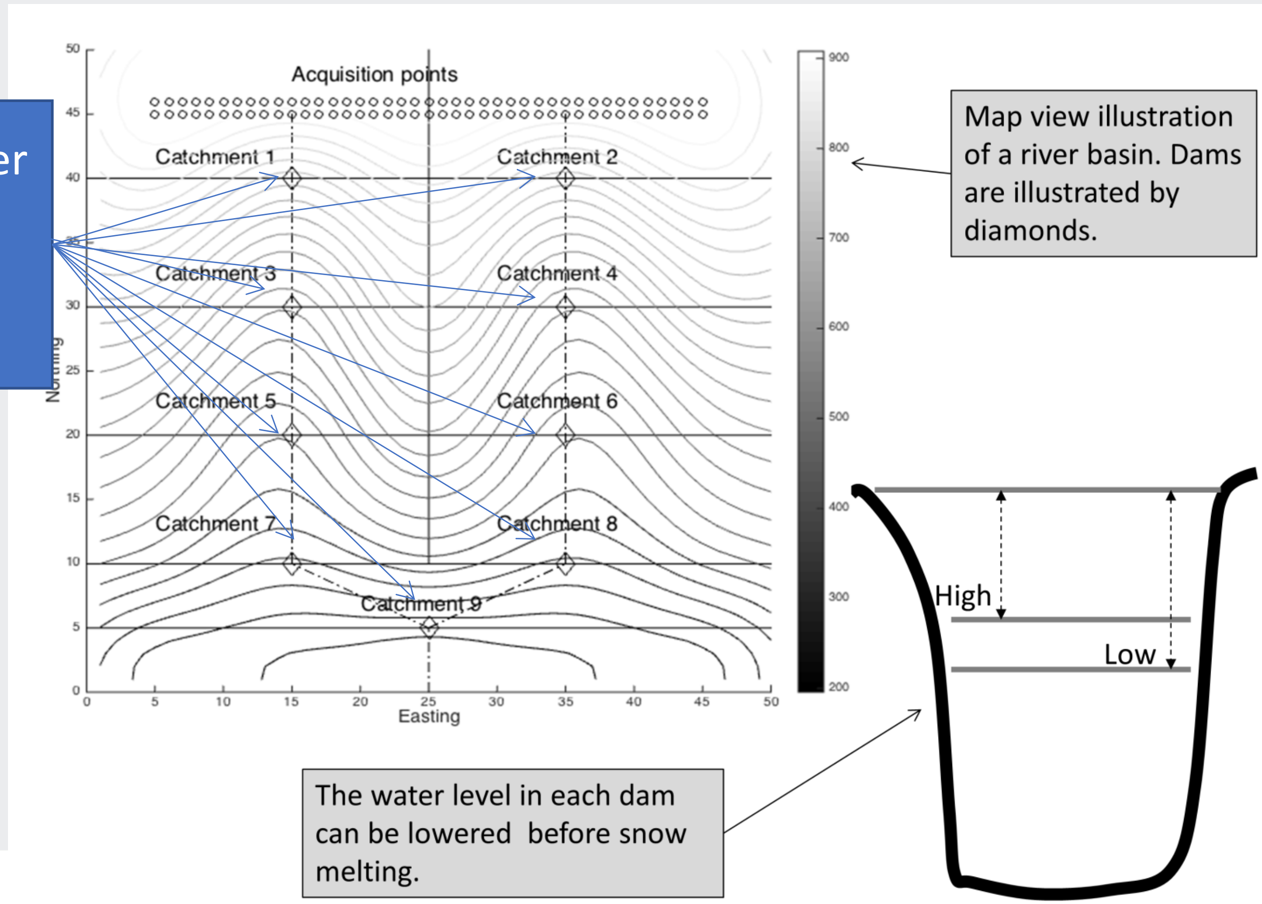
Which recharge location is better to prevent salt water intrusion?



Is it worthwhile to acquire electromagnetic data before making the decision about recharge?

Motivation
(a hydropower example)

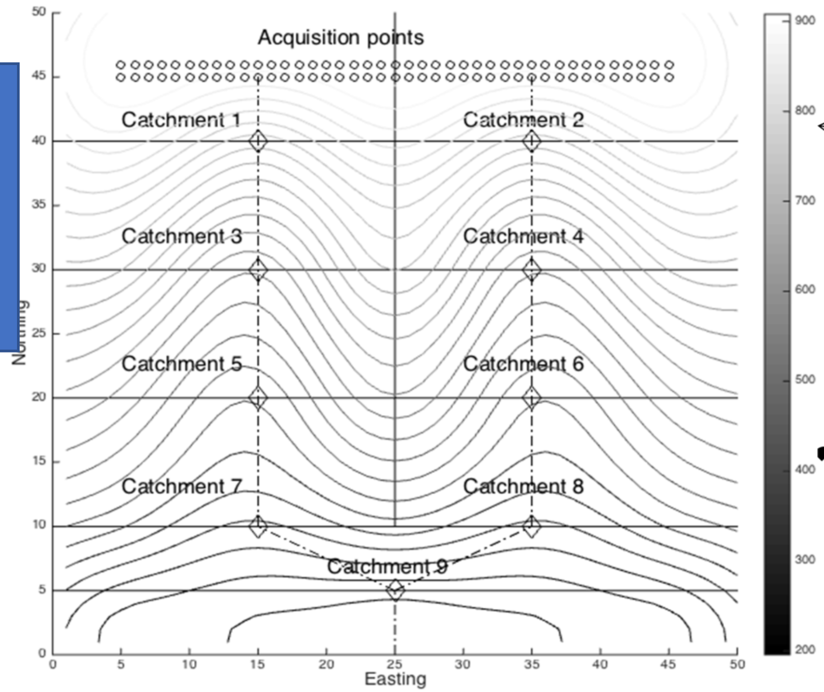
Adjusting water levels in 9 hydropower dams!



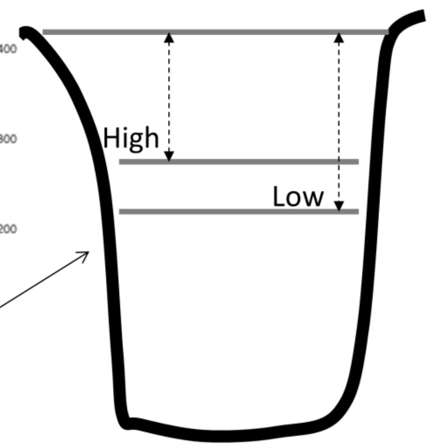
Ødegård, H., Eidsvik, J. and Fleten, S.E., 2017, Value of information analysis of snow measurements for the scheduling of hydropower production, *Energy Systems*.

Motivation
(a hydropower example)

Adjusting water levels in dams!



Map view illustration of a river basin. Dams are illustrated by diamonds.



The water level in each dam can be lowered before snow melting.

Acquire snow measurements?

Other applications

- Environmental – how monitor where pollutants are, to minimize risk or damage.
- Robotics - where should drone (UAV) or submarine (AUV) go to collect valuable data?
- Industry reliability – how to allocate sensors to 'best' monitor state of system?
- Internet of things – which sensors should be active now?



Which data are valuable?

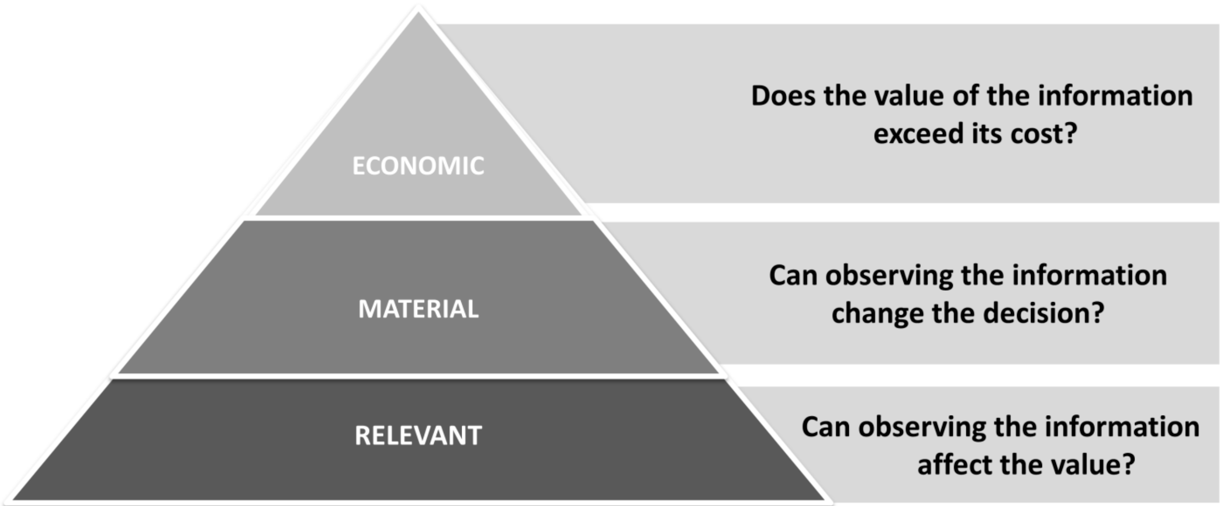
Five Vs of big data:

- Volume
- Variety
- Velocity
- Veracity
- **Value**

We must acquire and process data that has value!
There is often a clear question that one aims to answer, and data should help us.

Value of information (VOI)

In many Earth science applications we consider purchasing more data before making difficult decisions under uncertainty. The value of information (VOI) is useful for quantifying the value of the data, before it is acquired and processed.



This pyramid of conditions - VOI is different from other information criteria (entropy, variance, prediction error, etc.)

Information gathering

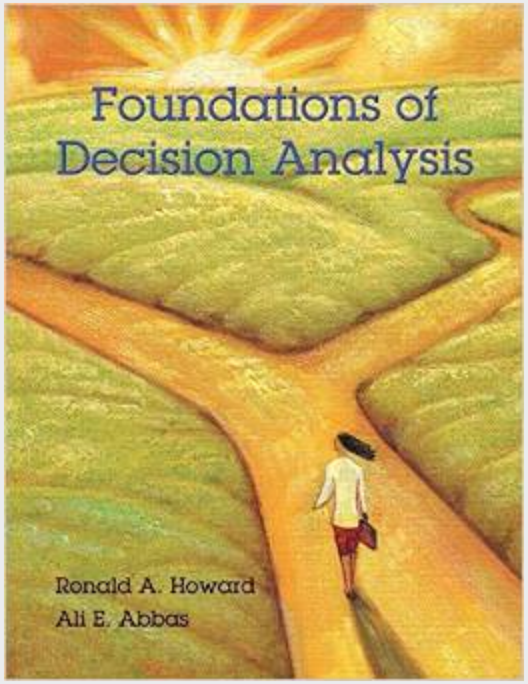
Why do we gather data?

To make better decisions!
To answer some kind of questions!
Reject or strengthen hypotheses!

We will use a **decision theoretic** perspective, but the methods are easily adapted to other criteria or value functions (later in course).

Decision analysis (DA)

Decision analysis attempts to guide a decision maker to clarity of action in dealing with a situation where one or more decisions are to be made, typically in the face of uncertainty.



Howard, R.A. and Abbas, A., 2015, *Foundations of Decision Analysis*, Prentice Hall.

Framing a decision situation

Rules of actional thought. (Howard and Abbas, 2015)

- Frame your decision situation to address the decision makers true concerns.
- Base decisions on maximum expected utility.

'…systematic and repeated violations of these principles will result in inferior long-term consequences of actions and a diminishes quality of life…'

(Edwards et al., 2007, Advances in decision analysis: From foundations to applications, Cambridge University Press.)

Pirate example
(For motivating decision analysis and VOI)

Pirate example

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).

?

Pirate makes decision based on preferences and maximum **utility** or **value**!

- Digging cost.
- Revenues if he finds the treasure .

Pirate example

$$a \in \{0,1\}$$

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).

$$x \in \{0,1\}$$

?

Pirate makes decision based on preferences and maximum **utility** or **value**!

- Digging cost.
- Revenues if he finds the treasure .

$$\max_{a \in \{0,1\}} \{E(v(x,a))\}$$

EAGE EDUCATION

Mathematics of decision situation:

- **Alternatives**

$$a \in \{0,1\} = A$$

- **Uncertainties (probability distribution)**

$$x \in \{0,1\} = \Omega \qquad p(x=1) = 0.01$$

- **Values**

$$v = v(x,a)$$

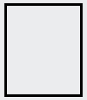
$$v(x=0, a=1) = -10000 \qquad v(x=1, a=1) = 100000 \qquad v(x, a=0) = 0$$

- **Maximize expected value**

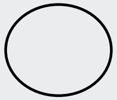
$$a^* = \arg\max_{a \in A}\{E(v(x,a))\}$$

Decision trees

A way of structuring and illustrating a decision situation.



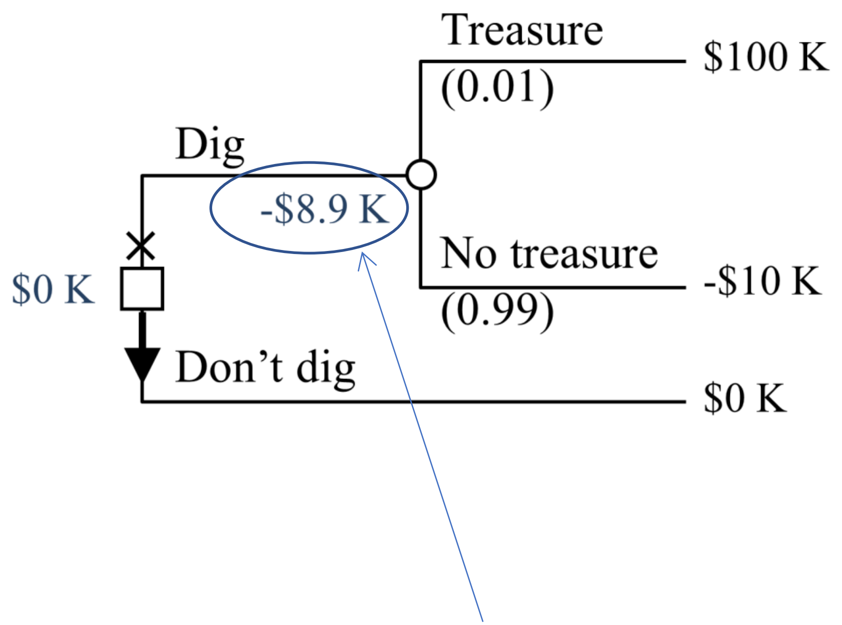
- Squares represent decisions
- Circles represent uncertainties



- Probabilities and values are shown by numbers.

- Arrows indicate the optimal decision.

Pirate's decision situation



$$E\left(u\left(v_{dig}\right)\right)=E\left(v_{dig}\right)=0.01(100000)+0.99(-10000)=-8900$$

Pirate example

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).

- Pirate can collect **data** before making the decision, if the experiment is worth its price!

- Perfect information. Clairvoyant!

- Imperfect information. Detector!

Value of information (VOI)

- VOI analysis is used to compare the additional value of making informed decisions with the price of the information.
- If the VOI exceeds the price, the decision maker should purchase the data.

VOI=Posterior value – Prior value

VOI – Pirate considers clairvoyant

$$PV = 0 = \$0K$$

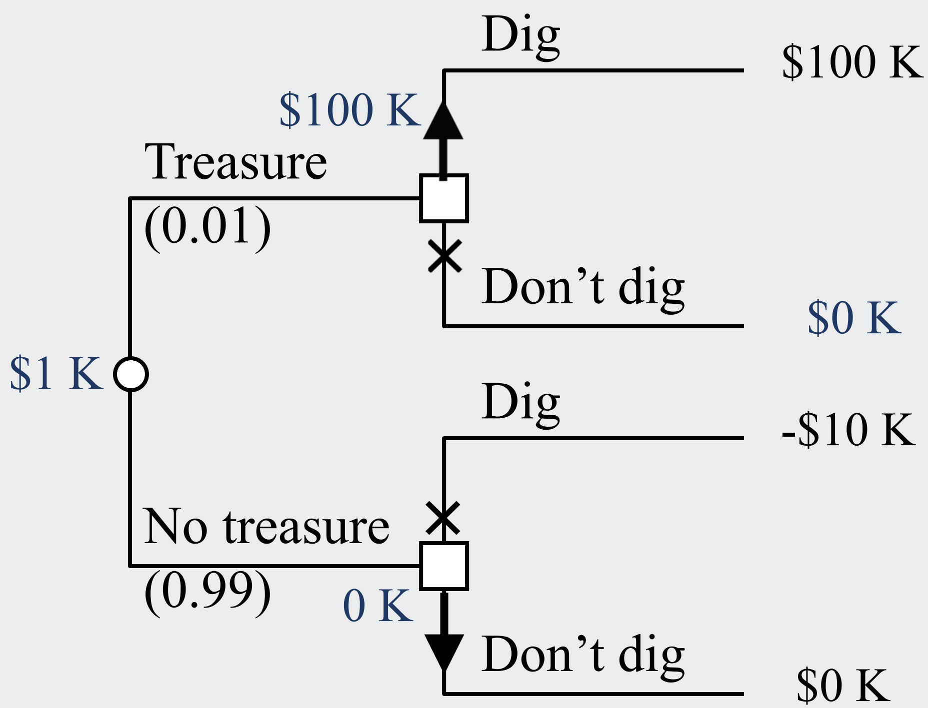
$$PoV(x) = \sum_{x} \max_{a \in A} \{v(x,a)\} \, p(x)$$

$$= \left(0.01 \cdot \max\{0,100\}\right) + \left(0.99 \cdot \max\{0,-10\}\right) = \$1K$$

$$VoI(x) = PoV(x) - PV = 1 - 0 = \$1K$$

Conclusion: Consult clairvoyant if (s)he charges less than $1000.

PoV – decision tree, perfect information



Pirate example - detector

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).
- Pirate can collect **data** before making the decision, if the experiment is worth its price!

Pirate makes decision based on preferences and maximum expected **value**!

- Digging cost.
- Revenues if he finds the treasure .

Pirate example - detector

$a \in \{0,1\}$

- **Pirate example**: A pirate must decide whether to dig for a treasure, or not. The treasure is absent or present (uncertainty).

- Pirate can collect **data** with a detector before making the decision, if this experiment is worth its price!

$x \in \{0,1\}$

$y \in \{0,1\}$

Pirate makes decision based on preferences and maximum expected **value**!

- Digging cost.
- Revenues if he finds the treasure .

$$\max_{a\in\{0,1\}} \{E(v(x,a)|y)\}$$

Detector experiment

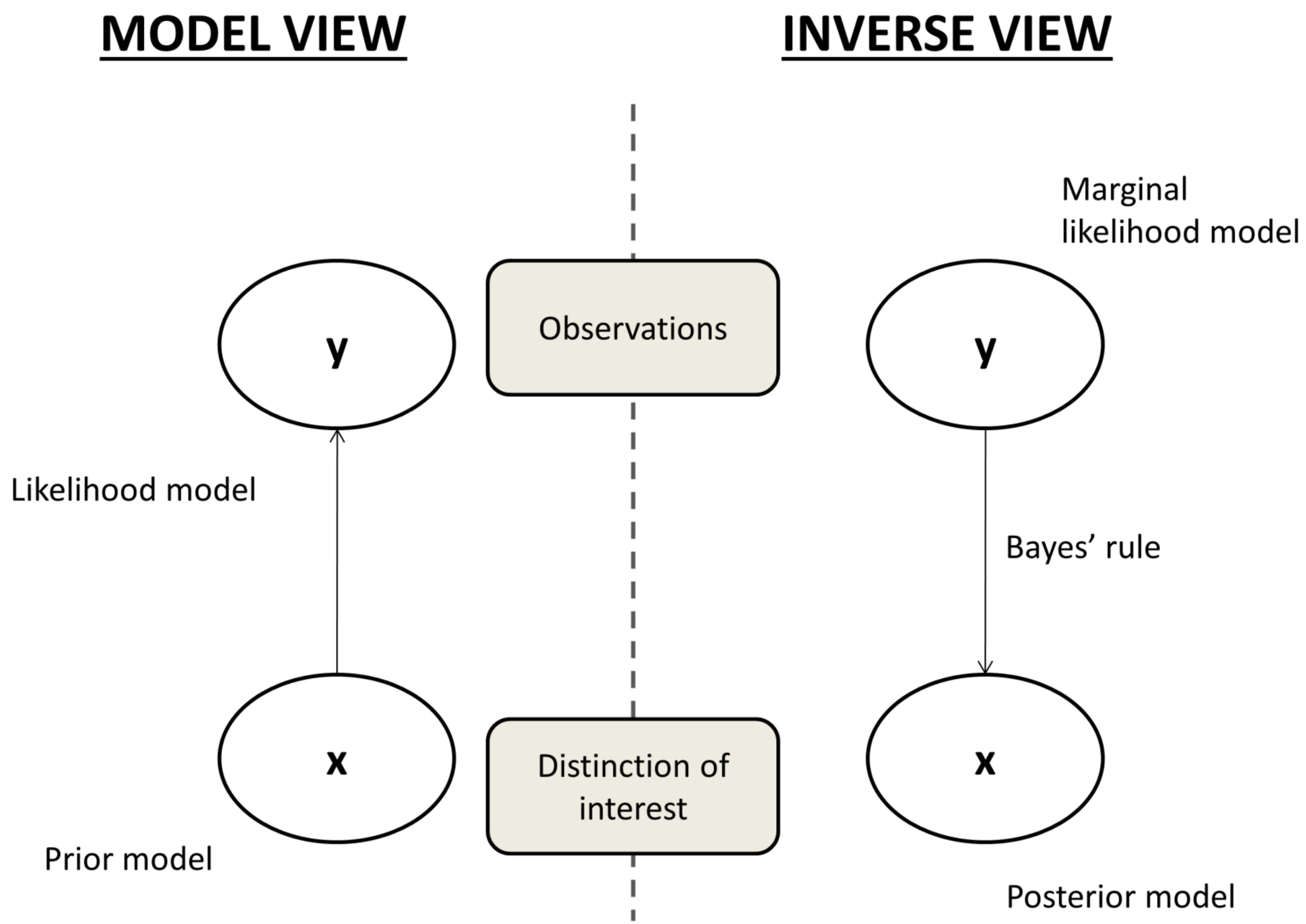
Accuracy of test:

$$p(y=0 \mid x=0) = p(y=1 \mid x=1) = 0.95$$

Should the pirate pay to do a detector experiment?

Does the VOI of this experiment exceed the price of the test?

Bayes rule - Detector experiment



Bayes rule - Detector experiment

Likelihood:

$$p(y=0\,|\,x=0) = p(y=1\,|\,x=1) = 0.95$$

Marginal likelihood:

$$p\left(y=1\right) = p\left(y=1\,|\,x=0\right)p\left(x=0\right) + p\left(y=1\,|\,x=1\right)p\left(x=1\right)$$
$$= 0.05\cdot 0.99 + 0.95\cdot 0.01 = 0.06$$

Posterior: $$p(x=1\,|\,y=1) = \frac{p(y=1\,|\,x=1)\,p(x=1)}{p(y=1)} = \frac{0.95\cdot 0.01}{0.06} \approx 0.16 = 16/100.$$

$$p(x=1\,|\,y=0) = \frac{p(y=0\,|\,x=1)\,p(x=1)}{p(y=0)} = \frac{0.05\cdot 0.01}{0.94} \approx 0.0005 = 5/10000.$$

EAGE EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

VOI – Pirate considers detector test

$$PoV(y) = \sum_{y} \max_{a \in A} \left\{ E\left(v(x,a) \mid y\right) \right\} p(y)$$

$$= \left(0.06 \cdot \max\left\{0, (100 \cdot 0.16) + (-10 \cdot 0.84)\right\}\right)$$

$$+ \left(0.94 \cdot \max\left\{0, (100 \cdot 0.0005) + (-10 \cdot 0.9995)\right\}\right)$$

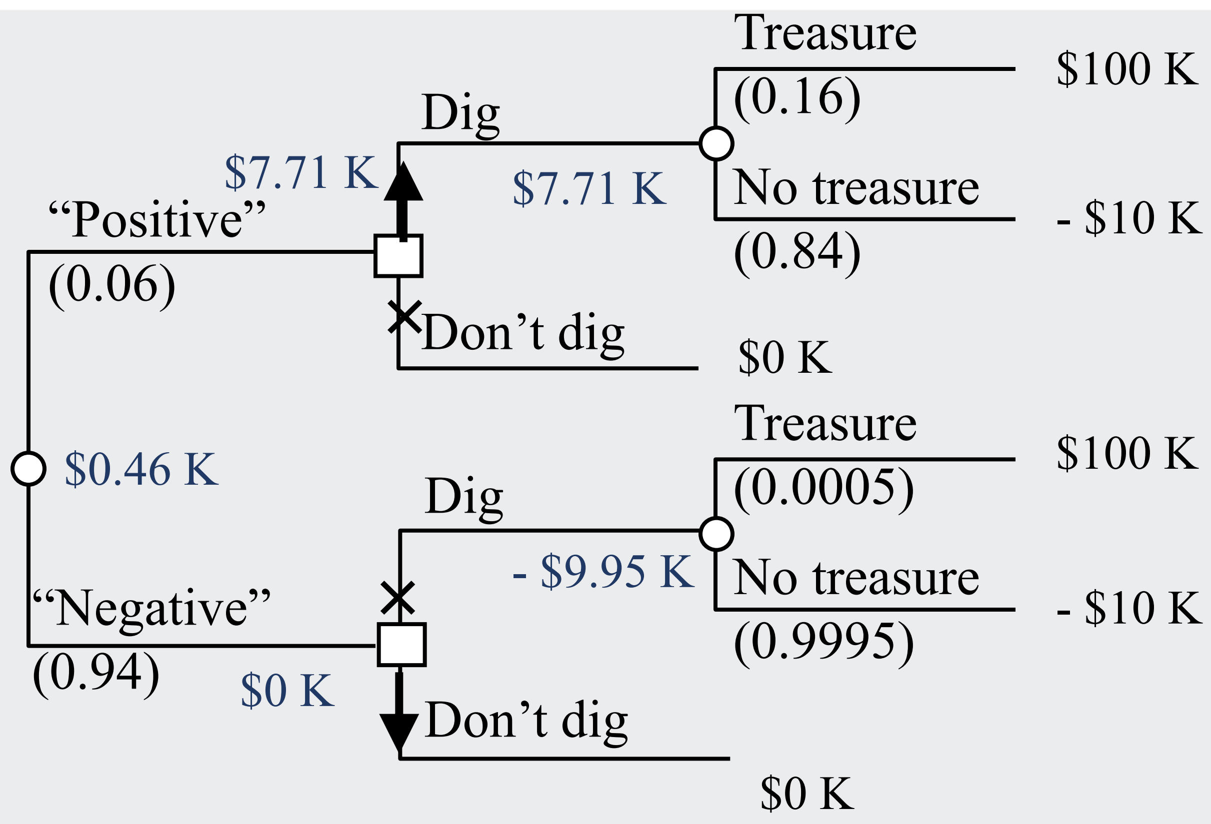
$$= \left(0.06 \cdot \max\left\{0, 7.71\right\}\right) + \left(0.94 \cdot \max\left\{0, -9.95\right\}\right) = \$0.46K.$$

$$VoI(y) = PoV(y) - PV = 0.46 - 0 = \$0.46K$$

Conclusion: Purchase detector testing if its price is less than $460.

EAGE EDUCATION

PoV - imperfect information

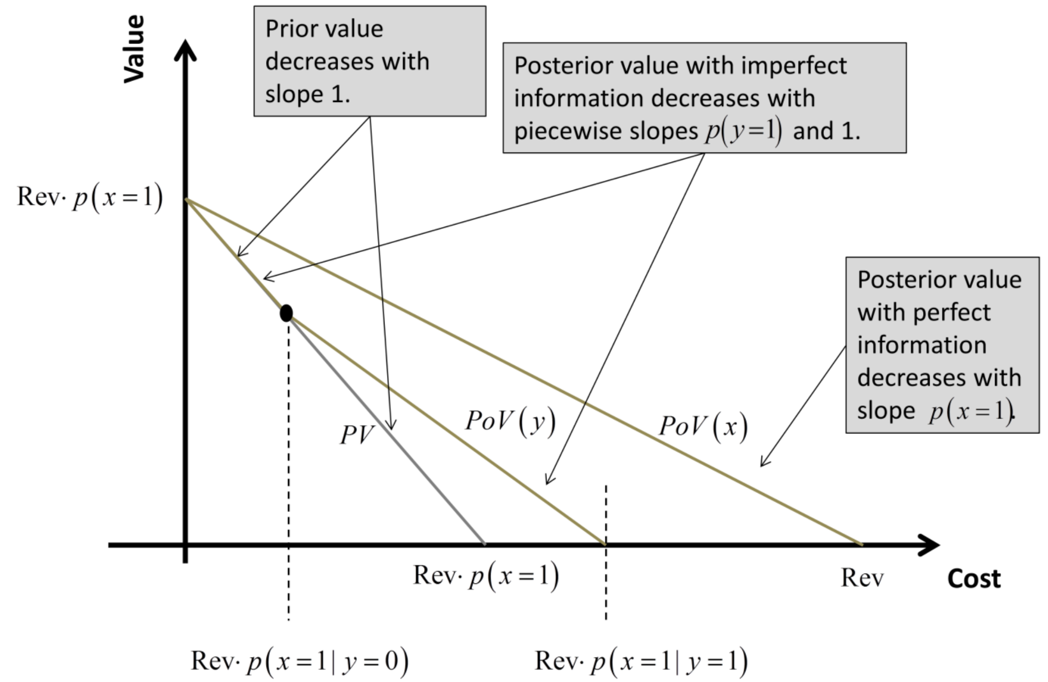


PV and PoV vs Digging Cost

$$PV = \max\{0, \text{Rev}\cdot p(x=1) - \text{Cost}\}$$

$$PoV(x) = \max\{0, \text{Rev} - \text{Cost}\}\, p(x=1)$$

$$PoV(y) = \sum_y \max\{0, \text{Rev}\cdot p(x=1 \mid y) - \text{Cost}\}\, p(y)$$



EAGE
EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

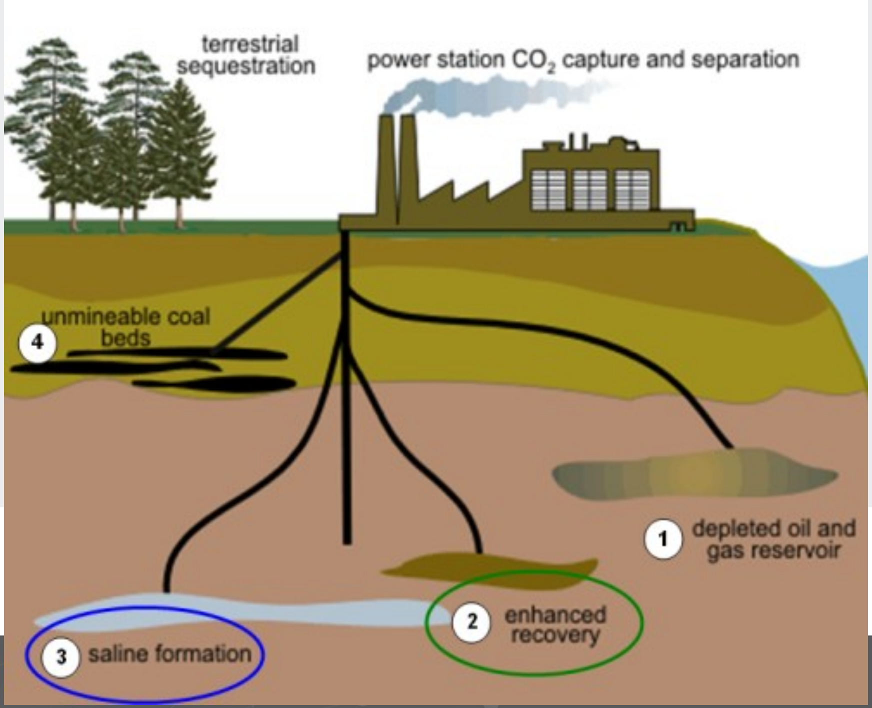
Plan for course

Time	Topic
Monday	Introduction and motivating examples
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Tuesday	Examples of value of information analysis in Earth sciences
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.

Problem: CO_2 sequestration

CO_2 is sequestered to reduce carbon emission in the athmosphere and defer global warming.
Geological sequestration involves pumping CO_2 in subsurface layers, where it will remain, unless it leaks to the surface.



VOI for CO2 sequestration

The decision maker can proceed with CO2 injection or suspend sequestration. The latter incurs a tax of 80 monetary units. The former only has a cost of injection equal to 30 monetary units, but the injected CO2 may leak (x=1). If leakage occurs, there will be a fine of 60 monetary units (i.e. a cost of 90 in total). Decision maker is risk neutral.

$$p(x=1)=0.3 \qquad p(x=0)=0.7$$

Data: Geophysical experiment, with binary outcome, indicating whether the formation is leaking or not.

$$p(y=0 \mid x=0)=0.95 \qquad p(y=1 \mid x=1)=0.9$$

Problem:

1. **Compute the VOI of perfect information.**
2. **Compute conditional probabilities, expected values and the VOI of geophysical data.**

(MATLAB)

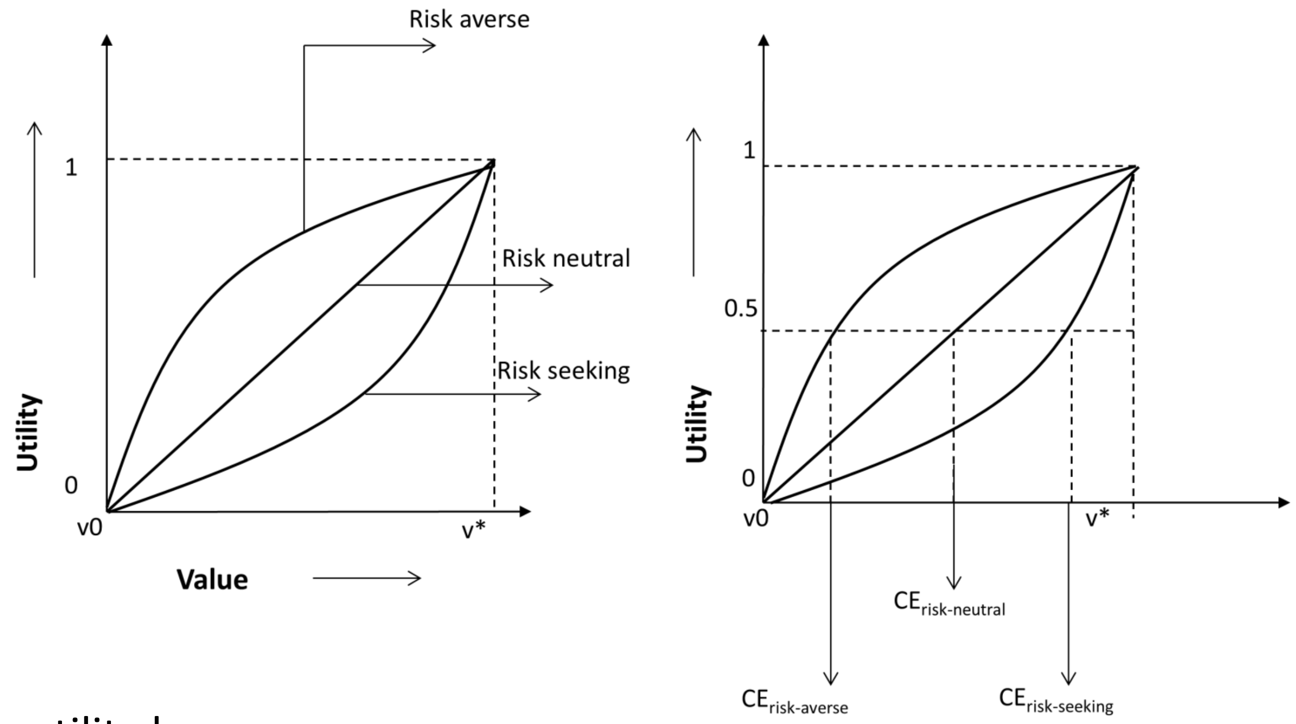
Value of information (VOI) - More general formulation

- VOI analysis is used to compare the additional value of making informed decisions with the price of the information.

- If the VOI exceeds the price, the decision maker should purchase the data.

VOI=Posterior value – Prior value

Risk and utility functions

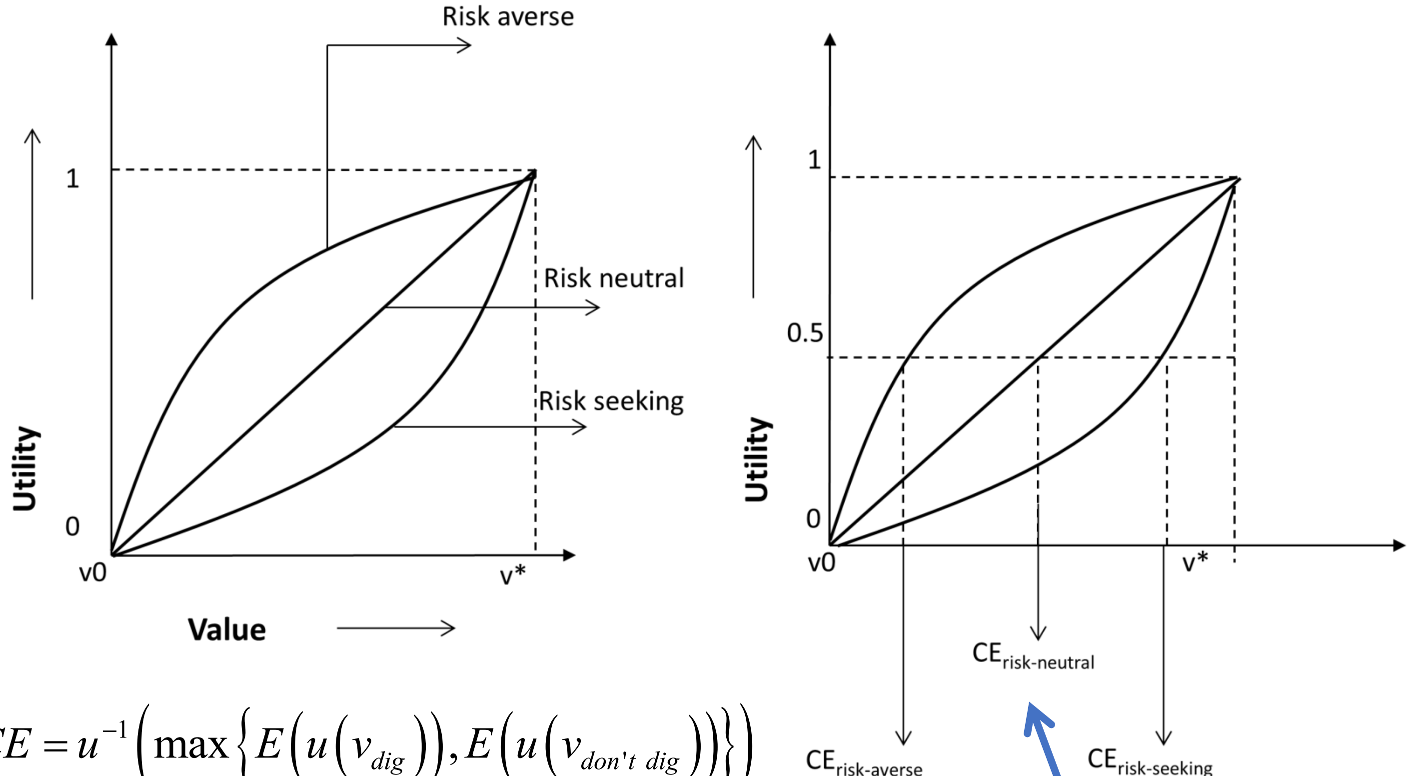


Exponential and linear utility have constant risk aversion coefficient:

$$\gamma = -\frac{u''(v)}{u'(v)}$$

Certain equivalents (CE)

Utilities are mathematical. The certain equivalent is a measure of how much a situation is worth to the decision maker. (It is measured in value).



$$CE = u^{-1}\left(\max\left\{E\left(u\left(v_{dig}\right)\right), E\left(u\left(v_{don't\ dig}\right)\right)\right\}\right)$$

What is the value of indifference? How much would the owner of a lottery be willing to sell it for?

VOI - Clairvoyance

Price P of experiment makes the equality.

$$\sum_x \max_{a \in A}\left\{v(x,a) - P\right\} p(x) = \max_{a \in A}\left\{E\left(v(x,a)\right)\right\}$$

$$\rightarrow P = VOI = \sum_x \max_{a \in A}\left\{v(x,a)\right\} p(x) - \max_{a \in A}\left\{E\left(v(x,a)\right)\right\}$$

VOI=Posterior value – Prior value

Assuming risk neutral decision maker!

VOI- Imperfect

Price of indifference.

$$\sum_y \max_{a \in A}\left\{E\left(v(x,a) - P \mid y\right)\right\} p(y) = \max_{a \in A}\left\{E\left(v(x,a)\right)\right\}$$

$$\sum_y \max_{a \in A}\left\{E\left(v(x,a) - P \mid y\right)\right\} p(y) = \max_{a \in A}\left\{E\left(v(x,a)\right)\right\}$$

$$\rightarrow P = VOI = \sum_y \max_{a \in A}\left\{E\left(v(x,a) \mid y\right)\right\} p(y) - \max_{a \in A}\left\{E\left(v(x,a)\right)\right\}$$

VOI=Posterior value – Prior value

Assuming risk neutral decision maker!

Properties of VOI

a) VOI is always positive

- Data allow better, informed decisions.

$$\max\left\{0, \sum_i v_i\right\} \le \sum_i \max\{0, v_i\}$$

b) If value is in monetary units ,VOI is in monetary units.

c) Data should be purchased if VOI > Price of experiment P.

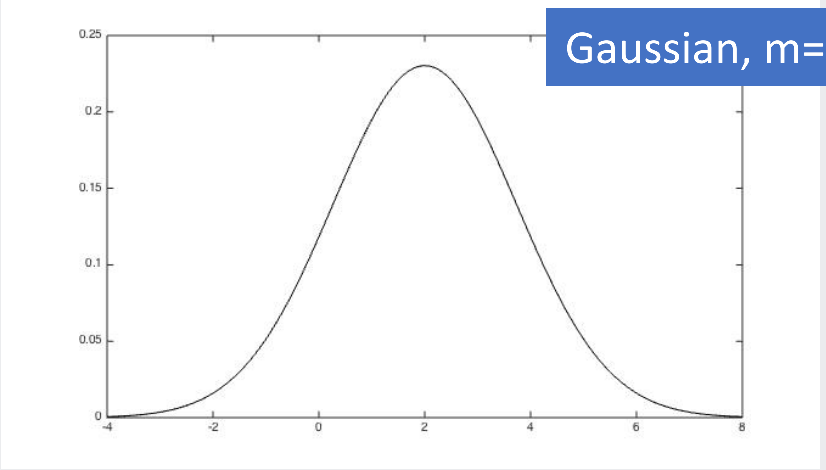
d) VOI of clairvoyance is an upper bound for any imperfect information gathering scheme.

e) When we compare different experiments, we purchase the one with largest VOI compared with the price:

$$\arg\max\{VOI_1 - P_1, VOI_2 - P_2\}$$

Gaussian model for profits

$$p(x) = \frac{1}{\sqrt{2\pi r^2}} \exp\left(-\frac{(x-m)^2}{2r^2} \right)$$



Uncertain profits of a project is Gaussian distributed.

VOI for Gaussian

Uncertain project profit is Gaussian distributed. Invest or not?
The decision maker asks a clairvoyant for perfect information, if the VOI is larger than her price.

$$VOI(x) = \text{PosteriorValue}(x) - \text{PriorValue}$$

$$PV = \max\{0, E(x)\}, \qquad E(x) = m$$

$$PoV(x) = E(\max\{0, x\}) = \int \max\{0, x\}\, p(x)\, dx$$

VOI for Gaussian

Result:

$$E\left(\max\{0,x\}\right)=\int \max\{0,x\}\,p(x)\,dx=\int_0^{\infty} x p(x)\,dx=\int_{-m/r}^{\infty}(m+rz)\phi(z)\,dz$$

$$=m\int_{-m/r}^{\infty}\phi(z)\,dz+r\int_{-m/r}^{\infty}z\phi(z)\,dz=m\left(1-\Phi\left(-m/r\right)\right)+r\phi\left(-m/r\right)$$

$$=m\Phi\left(m/r\right)+r\phi\left(m/r\right),$$

VOI for Gaussian

Gaussian cdf

Gaussian pdf

Result:

$$VOI(x) = m\Phi\left(\frac{m}{r}\right) + r\phi\left(\frac{m}{r}\right) - \max\{0, m\}$$

The analytical form facilitates computing, and it eases the study of VOI properties as a function of the parameters.

$$m = 0,$$

$$VOI(x) = r\phi(0) = \frac{r}{\sqrt{2\pi}}$$

The more uncertain, the more valuable is information.

Problem: VOI for Gaussian

Gaussian cdf

Gaussian pdf

Result:

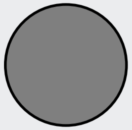
$$VOI(x) = m\Phi\left(m/r\right) + r\phi\left(m/r\right) - \max\{0, m\}$$

Problem:

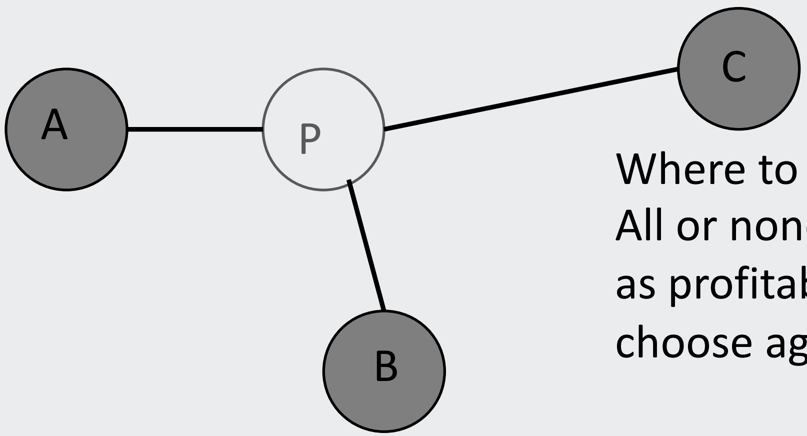
1. **Set mean to 0. Compute the VOI. Does it depend on variance?**
2. **Set variance 1. Compute the VOI. Does it depend on mean?**
3. **Plot VOI as a function of mean and variance.**

(MATLAB)

What if several projects / treasures?



What if several projects / treasures?



Where to invest?
All or none? Free to choose as many as profitable? One at a time, then choose again?

Where should one collect data? All or none? One only? Or two? One first, then maybe another?

VOI and Earth sciences

- **Alternatives are spatial**, often with high flexibiliy in selection of sites, control rates, intervention, excavation opportunities, harvesting, etc.

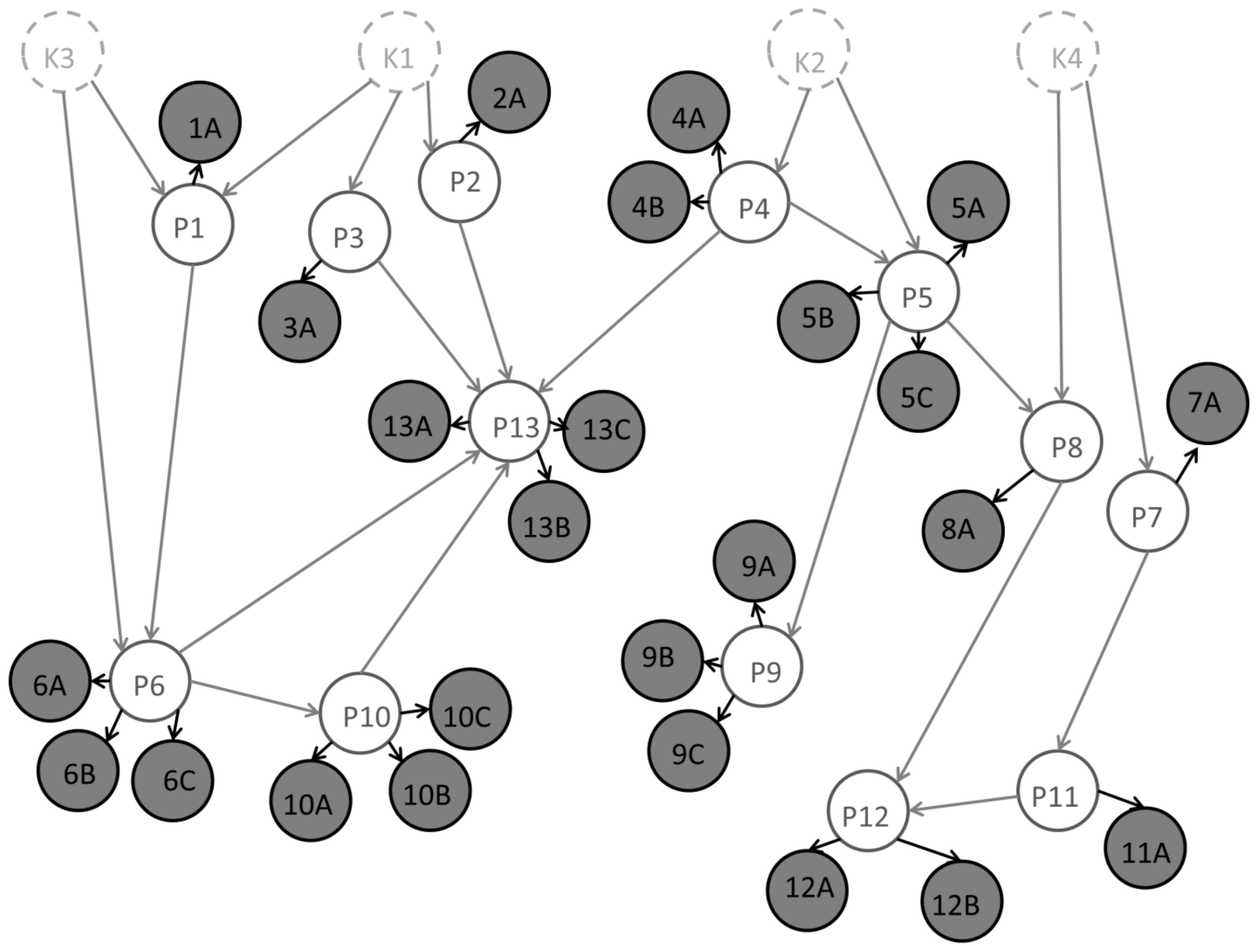
- **Uncertainties are spatial**, with multi-variable interactions . Often both discrete and continuous.

- **Value function is spatial**, typically involving coupled features, say through differential equations. It can be defined by «physics» as well as economic attributes.

- **Data are spatial**. There are plenty opportunities for partial, total testing and a variety of tests (surveys, monitoring sensors, electromagnetic data, , etc.)

Dependence? Does it matter?

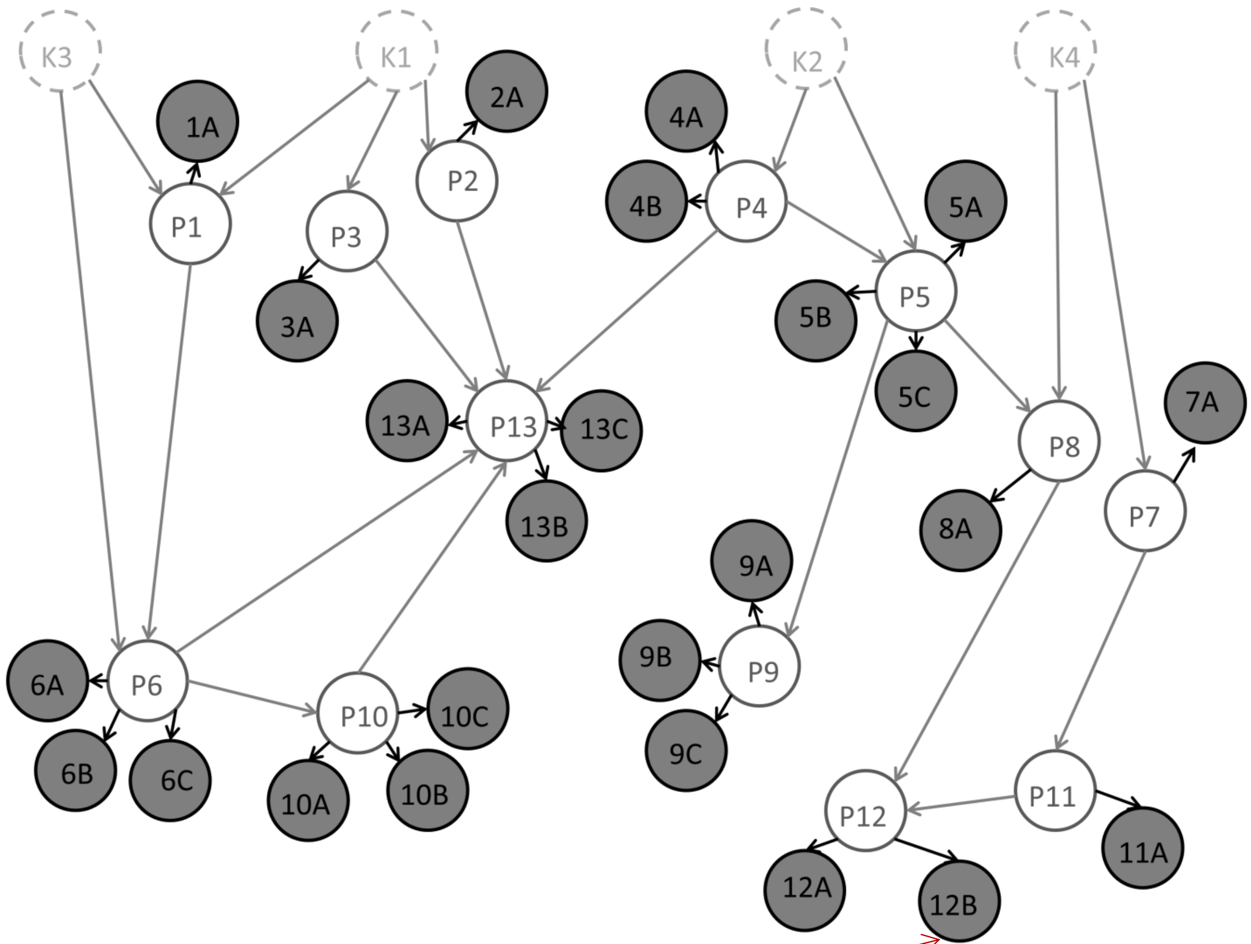
Gray nodes are petroleum reservoir segments where the company aims to develop profitable amounts of oil and gas.



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.



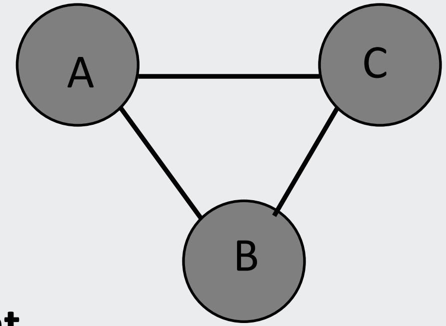
Drill the exploration well at this segment! The value of information is largest.

Plan for course

Time	Topic
Monday	Introduction and motivating examples
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Tuesday	Examples of value of information analysis in Earth sciences
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.

Joint modeling of multiple variables



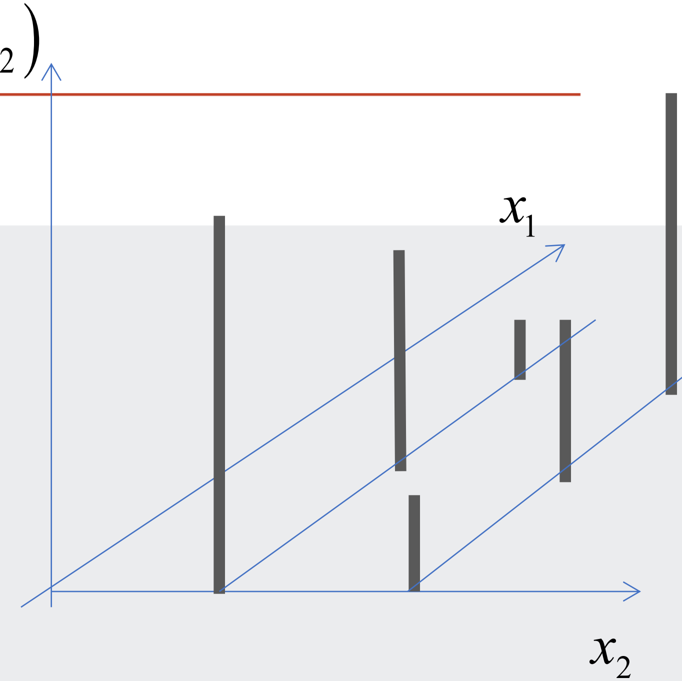
Spatial variables are often not independent!

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:

$$p(\boldsymbol{x}) \geq 0, \qquad \boldsymbol{x} \in \Omega,$$

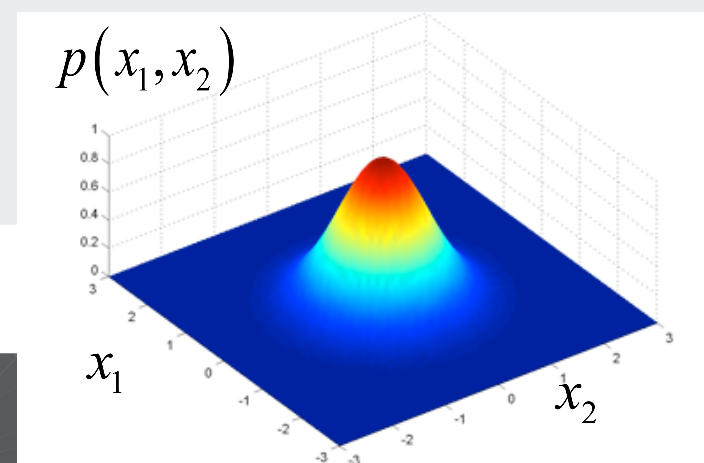
$$\sum_{x_1 \in \Omega_1} \cdots \sum_{x_n \in \Omega_n} p(\boldsymbol{x}) = 1.$$

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)

EAGE
EUROPEAN ASSOCIATION OF GEOSCIENTISTS & ENGINEERS

Multivariate statistical models

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

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

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

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

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

Marginal and conditional probability

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

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

Marginalization in joint pdf.

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

Conditioning in joint pdf.

Conditional mean and variance

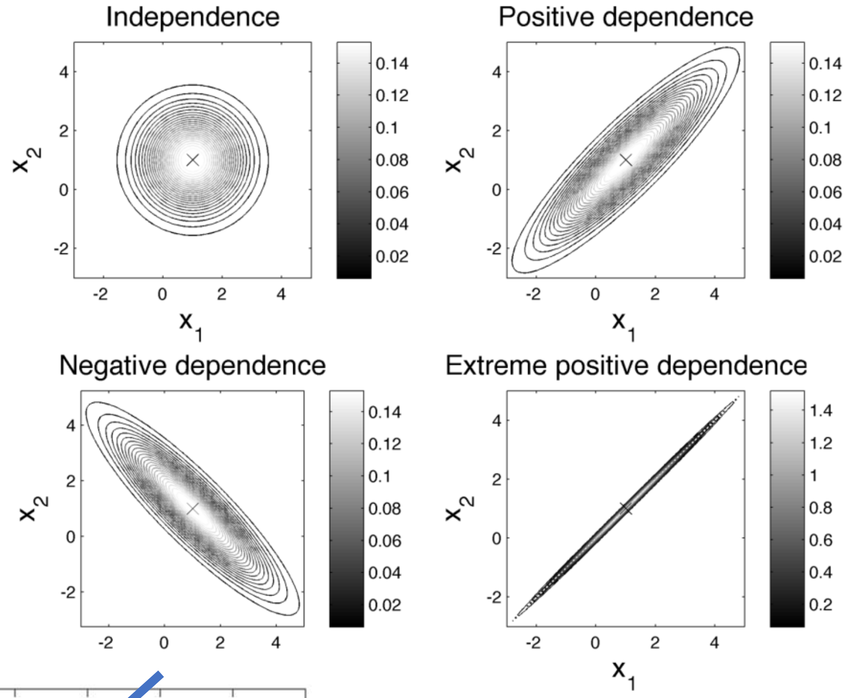
$$E(\boldsymbol{x}_{\mathrm{K}} \mid \boldsymbol{x}_{\mathrm{L}}) = \int \boldsymbol{x}_{\mathrm{K}}\, p(\boldsymbol{x}_{\mathrm{K}} \mid \boldsymbol{x}_{\mathrm{L}})\, d\boldsymbol{x}_{\mathrm{K}},$$

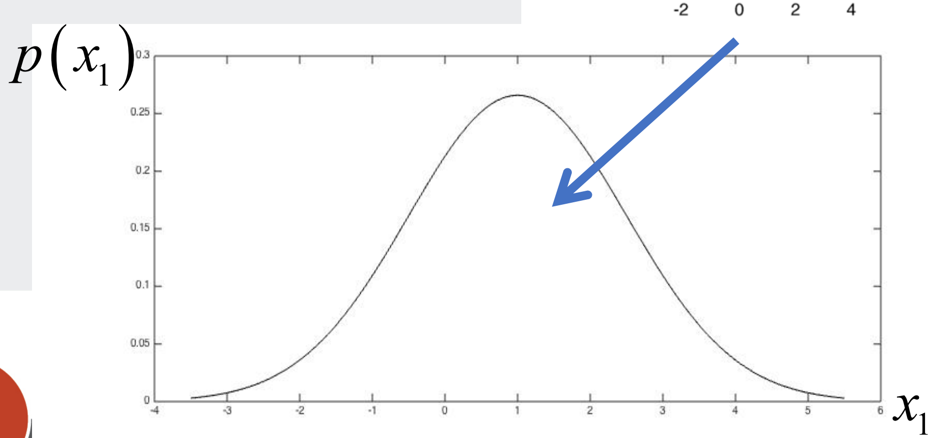
$$Var(\boldsymbol{x}_{\mathrm{K}} \mid \boldsymbol{x}_{\mathrm{L}}) = \int \left(\boldsymbol{x}_{\mathrm{K}} - E(\boldsymbol{x}_{\mathrm{K}} \mid \boldsymbol{x}_{\mathrm{L}})\right)\left(\boldsymbol{x}_{\mathrm{K}} - E(\boldsymbol{x}_{\mathrm{K}} \mid \boldsymbol{x}_{\mathrm{L}})\right)^{t} p(\boldsymbol{x}_{\mathrm{K}}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

Marginalization

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

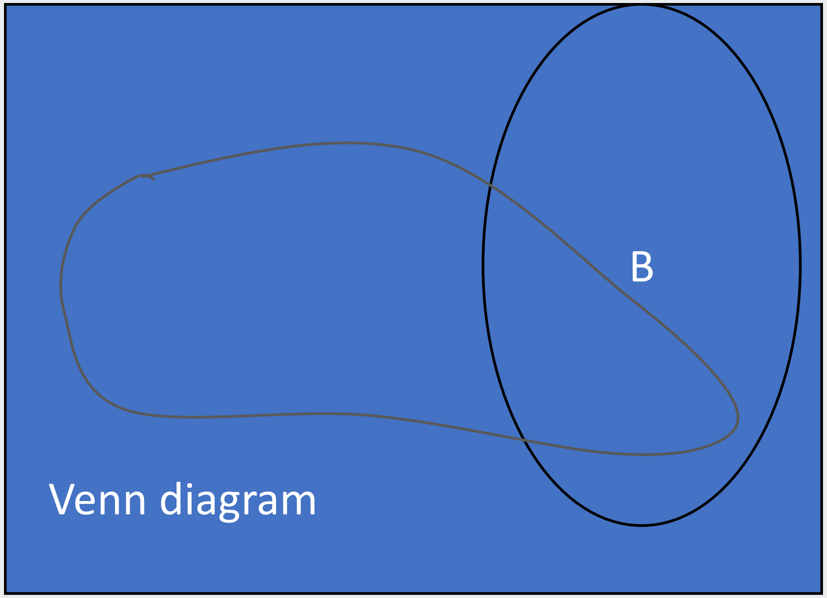
$$p(\boldsymbol{x}_{\mathrm{K}}) = \int p(\boldsymbol{x})\, d\boldsymbol{x}_{\mathrm{L}}$$





Conditional probability

$$p(\boldsymbol{x}_K \mid \boldsymbol{x}_L) = \frac{p(\boldsymbol{x})}{p(\boldsymbol{x}_L)} = \frac{p(\boldsymbol{x})}{\int p(\boldsymbol{x})\, d\boldsymbol{x}_K}$$



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

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

Conditional probability

$$p(\boldsymbol{x}_K \mid \boldsymbol{x}_L) = \frac{p(\boldsymbol{x})}{p(\boldsymbol{x}_L)} = \frac{p(\boldsymbol{x})}{\int p(\boldsymbol{x})\, d\boldsymbol{x}_K}$$

$$p(\boldsymbol{x}_K) \neq p(\boldsymbol{x}_K \mid \boldsymbol{x}_L)$$

$$p(\boldsymbol{x}_K) = p(\boldsymbol{x}_K \mid \boldsymbol{x}_L)$$

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.

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

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

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.

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

Modeling by conditional probability

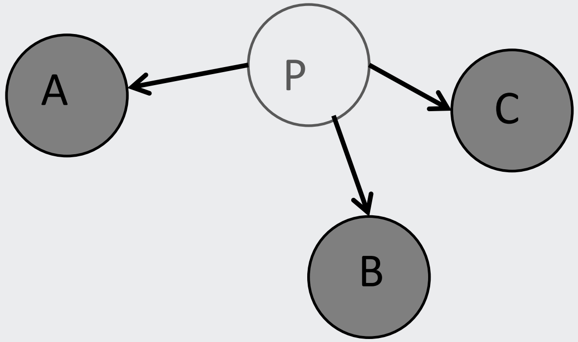
$$p(\boldsymbol{x}) = p(x_1)\,p(x_2 \mid x_1)\ldots p(x_n \mid x_{n-1},\ldots,x_1)$$

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:

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



- 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?

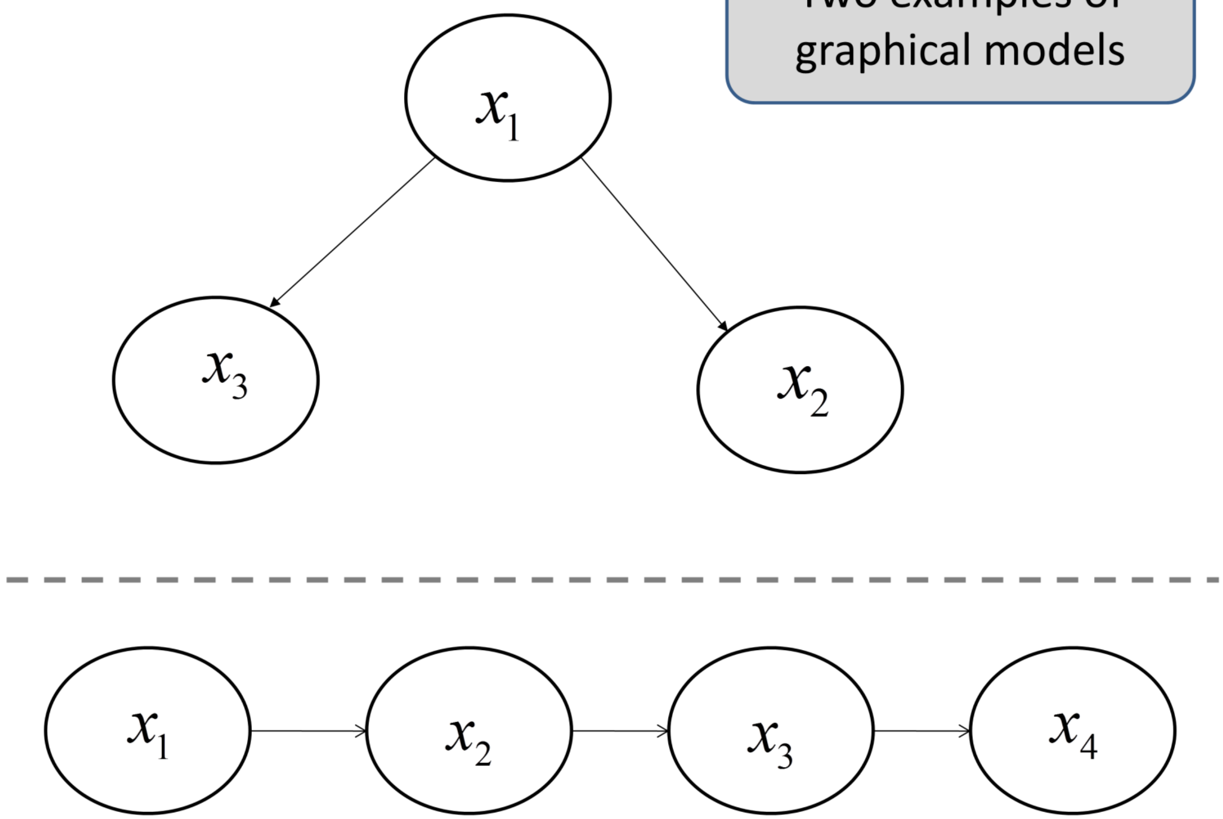
$$p(x_B = 1 \mid x_P = 1) = 0.9$$

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

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

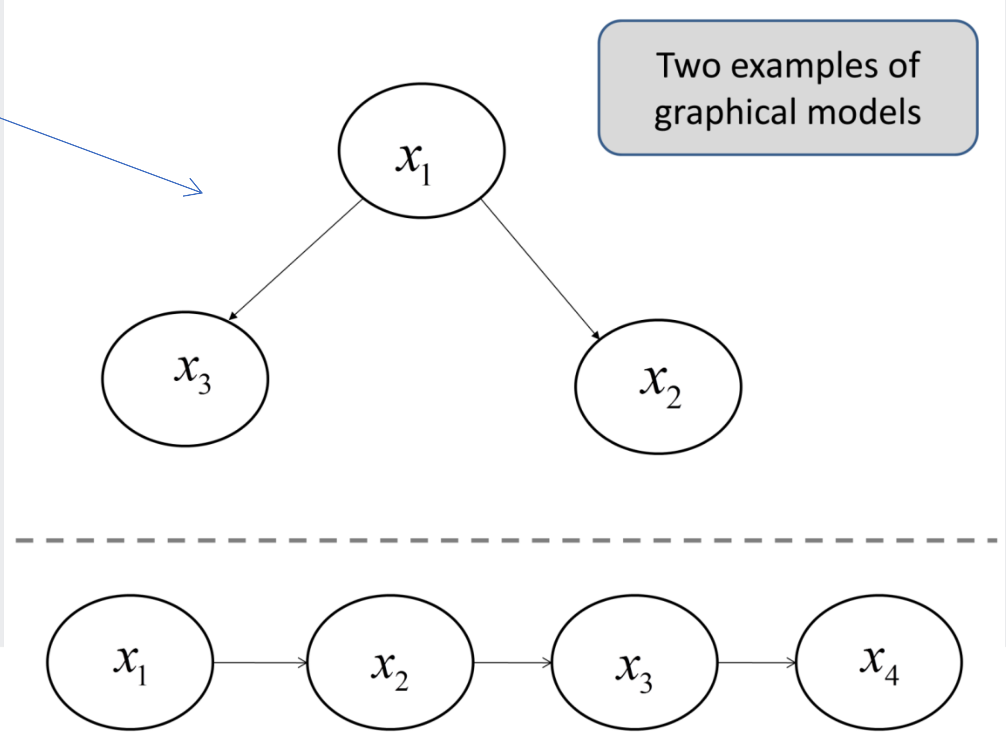
Bayesian networks and Markov chains

Two examples of graphical models



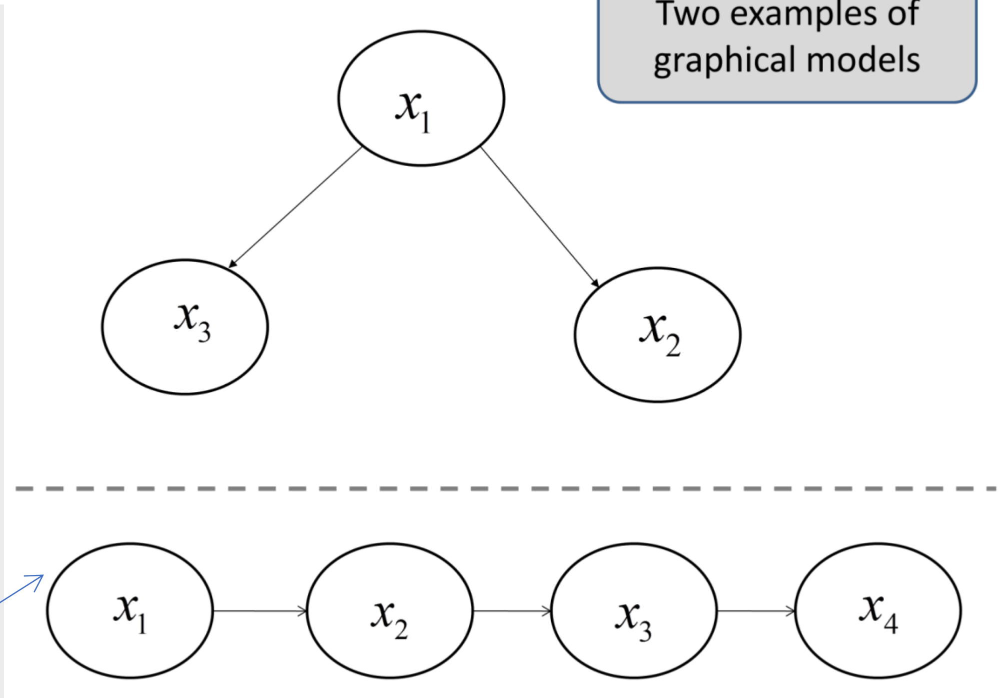
Bayesian networks and Markov chains

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



Bayesian networks and Markov chains

Two examples of graphical models

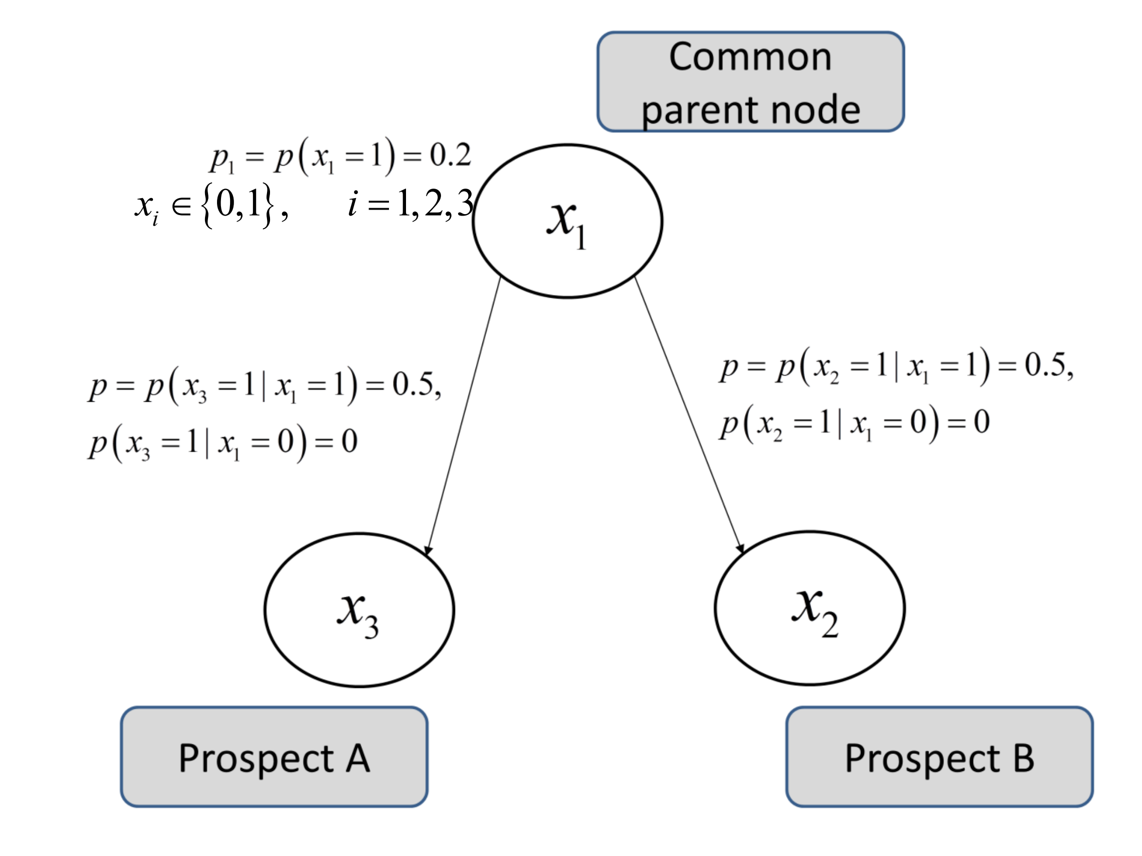


$$p(\boldsymbol{x}) = p(x_1, x_2, x_3, x_4) = p(x_1)\,p(x_2 \mid x_1)\,p(x_3 \mid x_2)\,p(x_4 \mid x_3)$$

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.

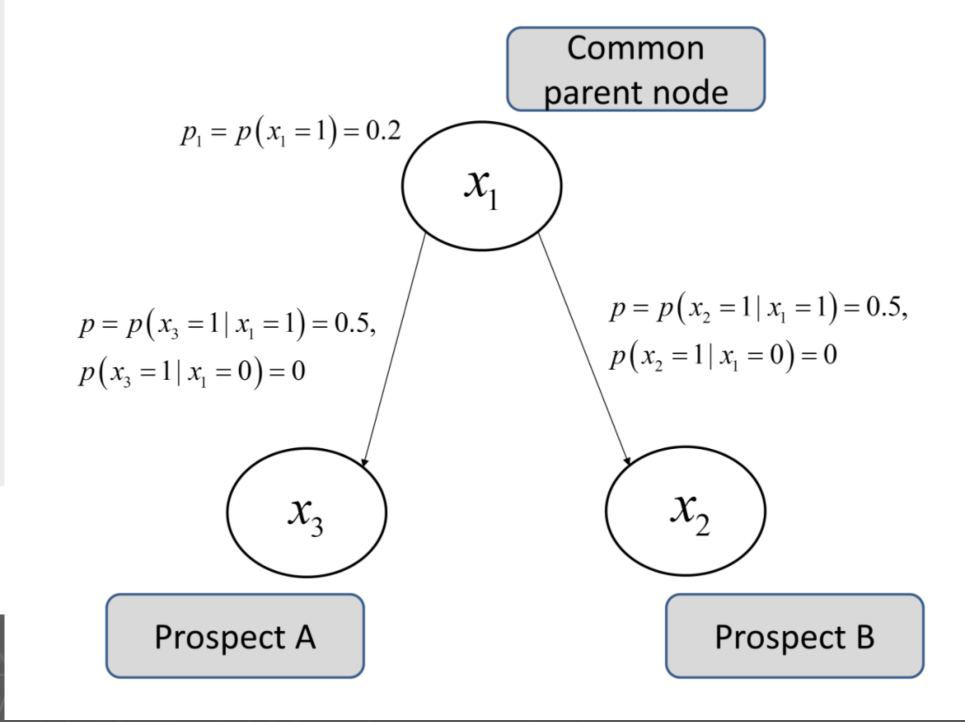


Problem: Bivariate petroleum prospects example

Problem:

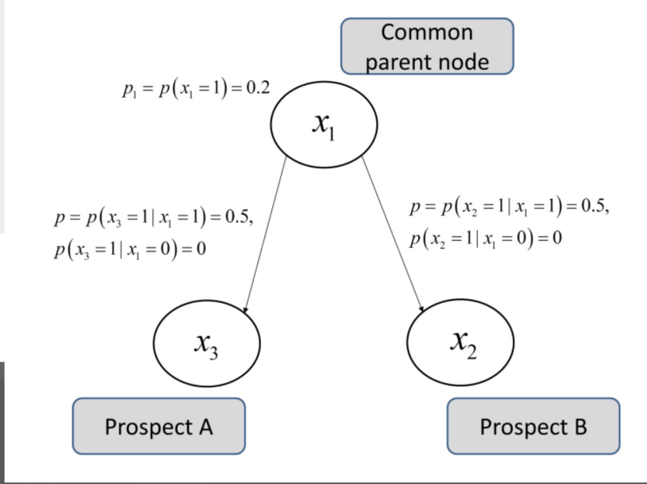
1. **Compute the conditional probability at prospect A, when one knows the success or failure outcome of prospect B.**
2. **Compare with marginal probability.**

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



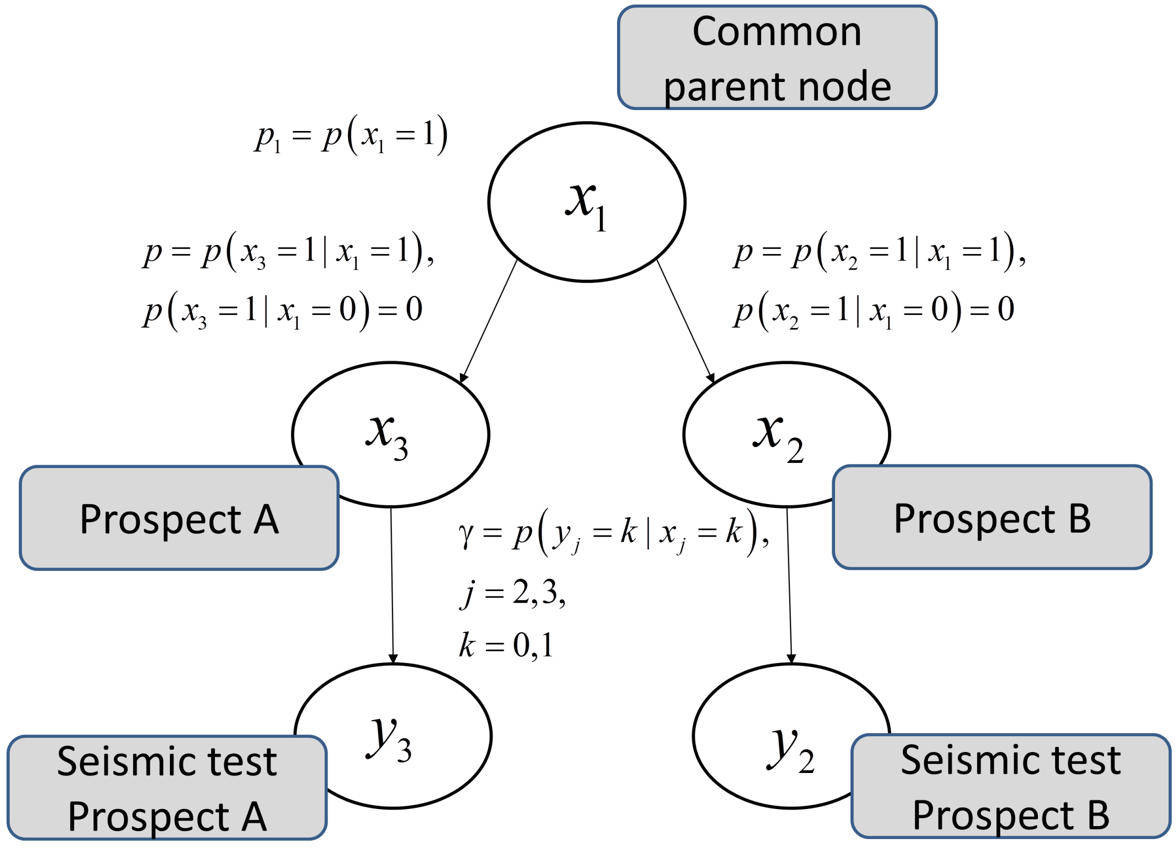
Bivariate petroleum prospects example

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



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.

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

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

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

Total clairvoyant information

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

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

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

EAGE EDUCATION

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

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

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

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

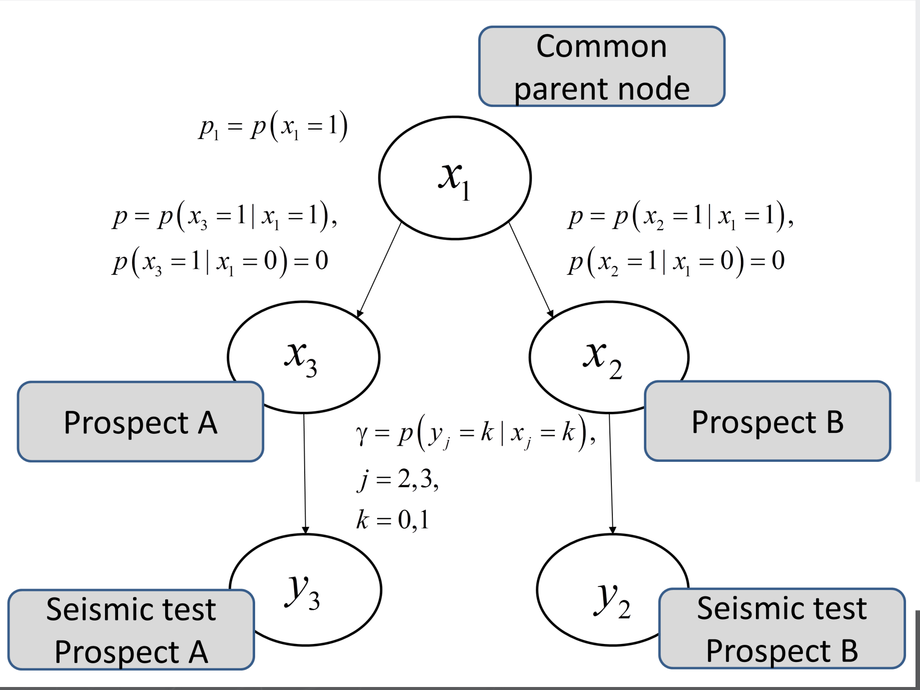
$$+ 0.9 \cdot \max\left\{0, 3 \cdot 0.055 - \text{Cost}\right\}$$

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.

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

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

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

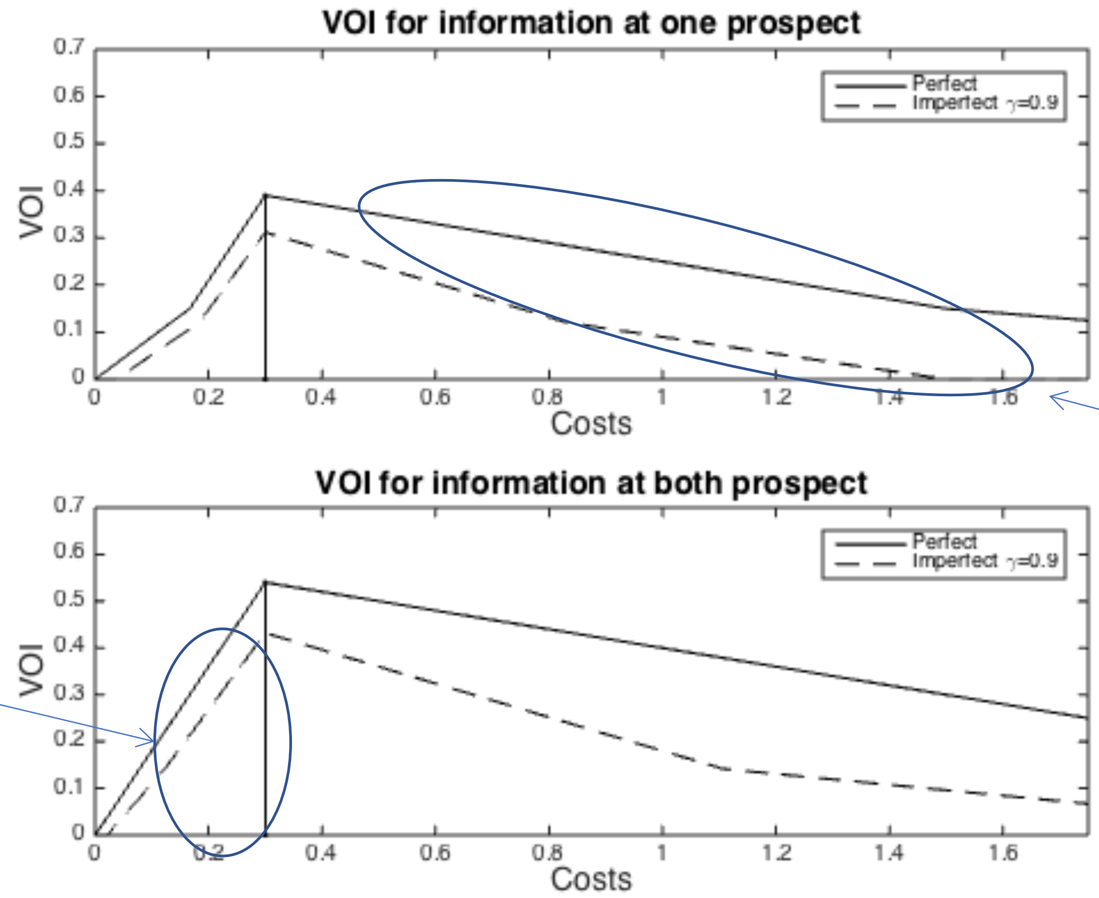
Total imperfect information

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

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

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

VOI for bivariate prospects example

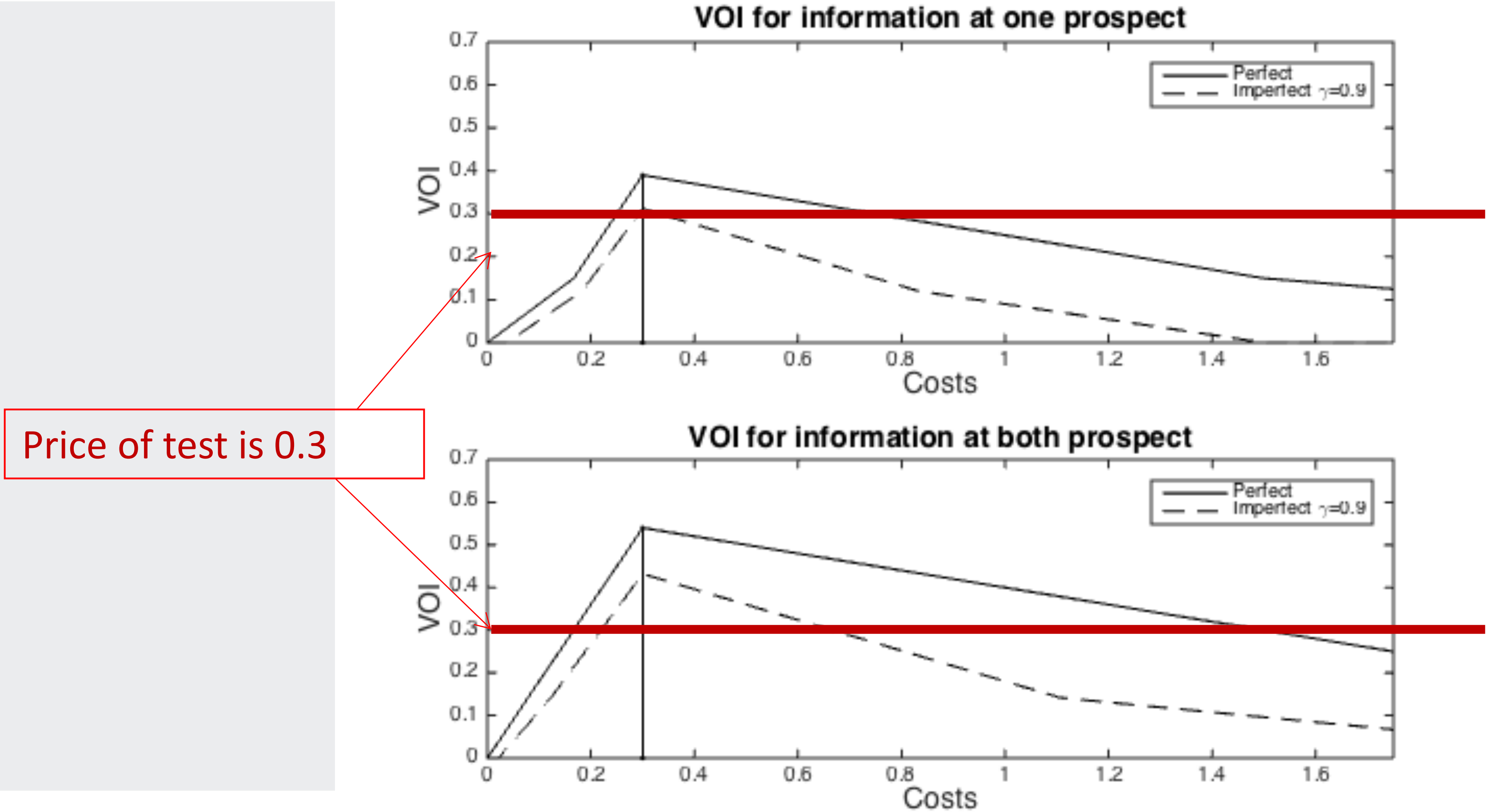


Partial perfect is better than imperfect total.

Imperfect total better then partial perfect.

VOI for bivariate prospects example

.



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.

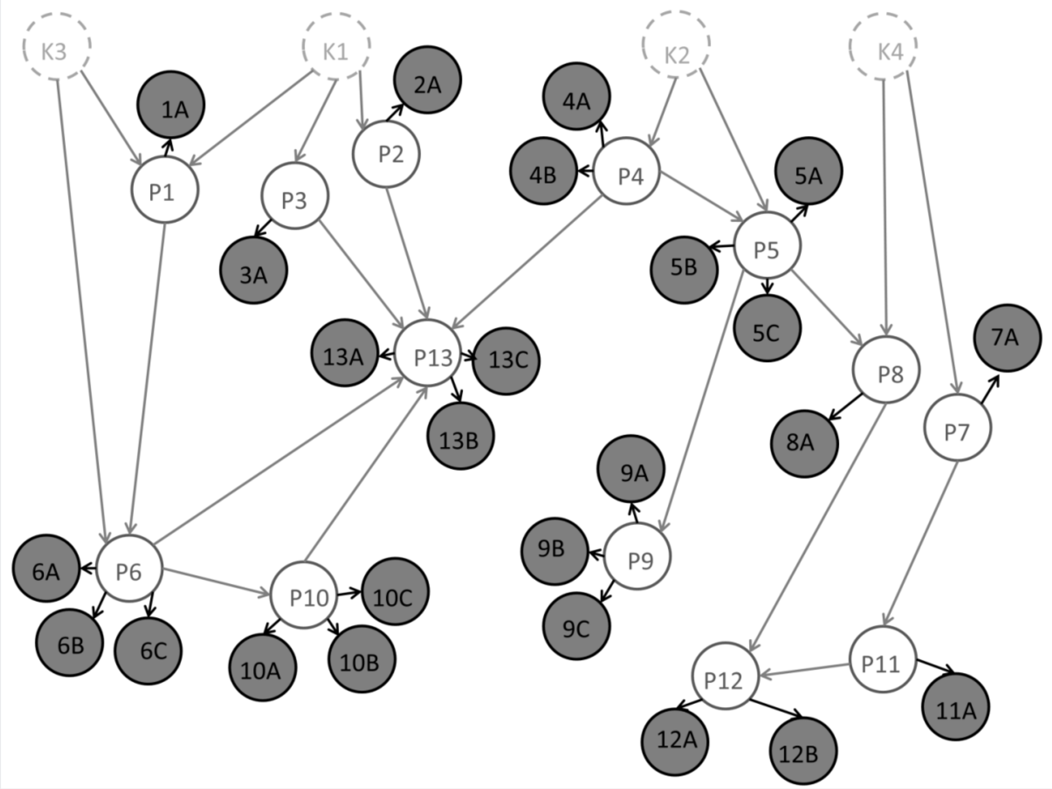
Plan for course

Time	Topic
Monday	Introduction and motivating examples
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Tuesday	Examples of value of information analysis in Earth sciences
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.

Larger networks - computation

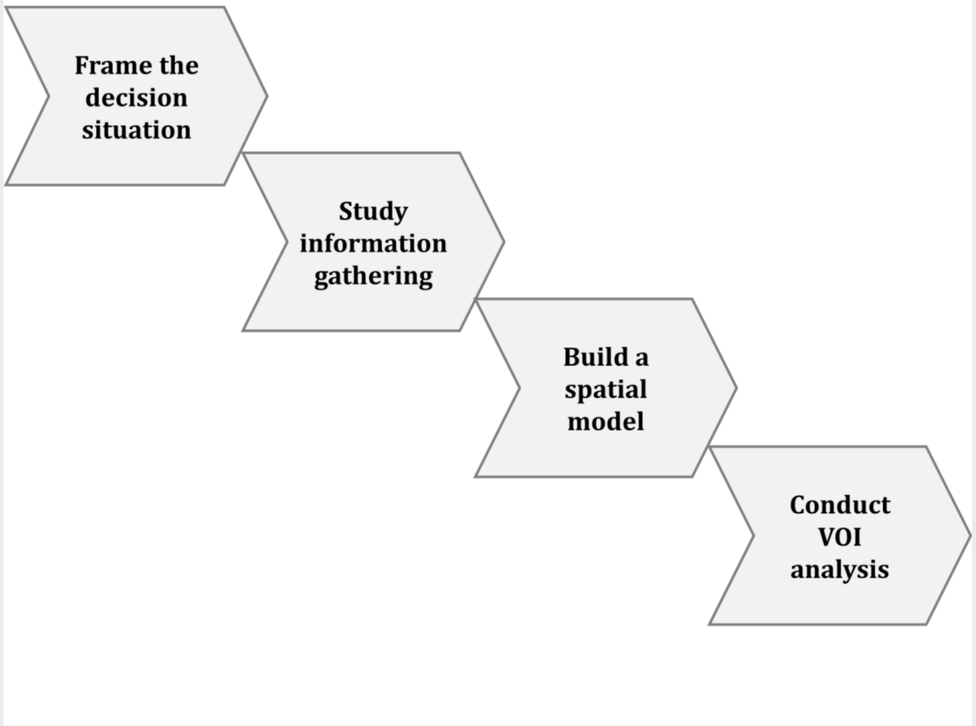
Algorithms have been developed for efficient marginalization, conditioning.



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).

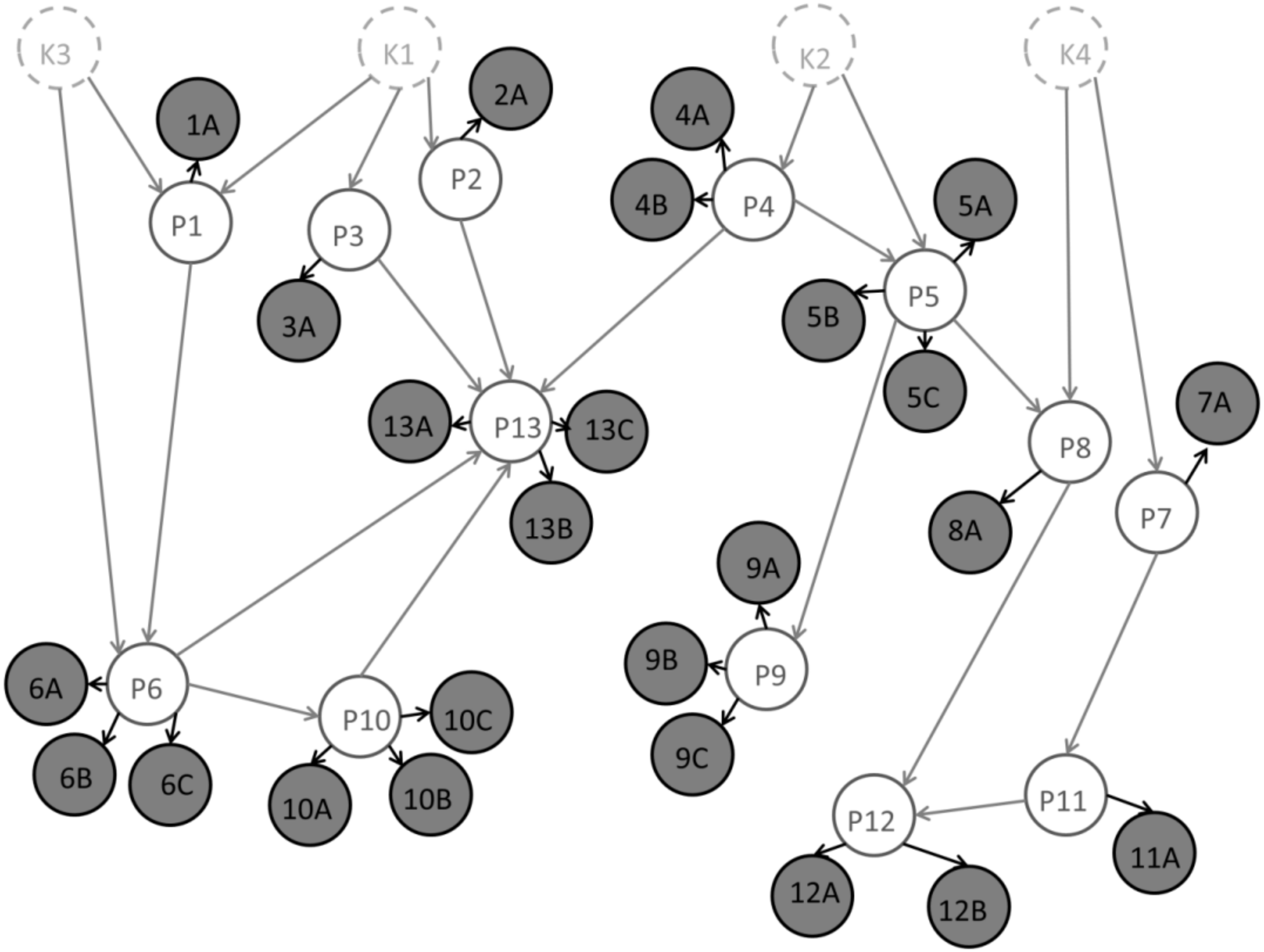


Bayesian networks, Kitchens

.

Model elicited from experts.

Migration from kitchens. Local failure probability of migration.

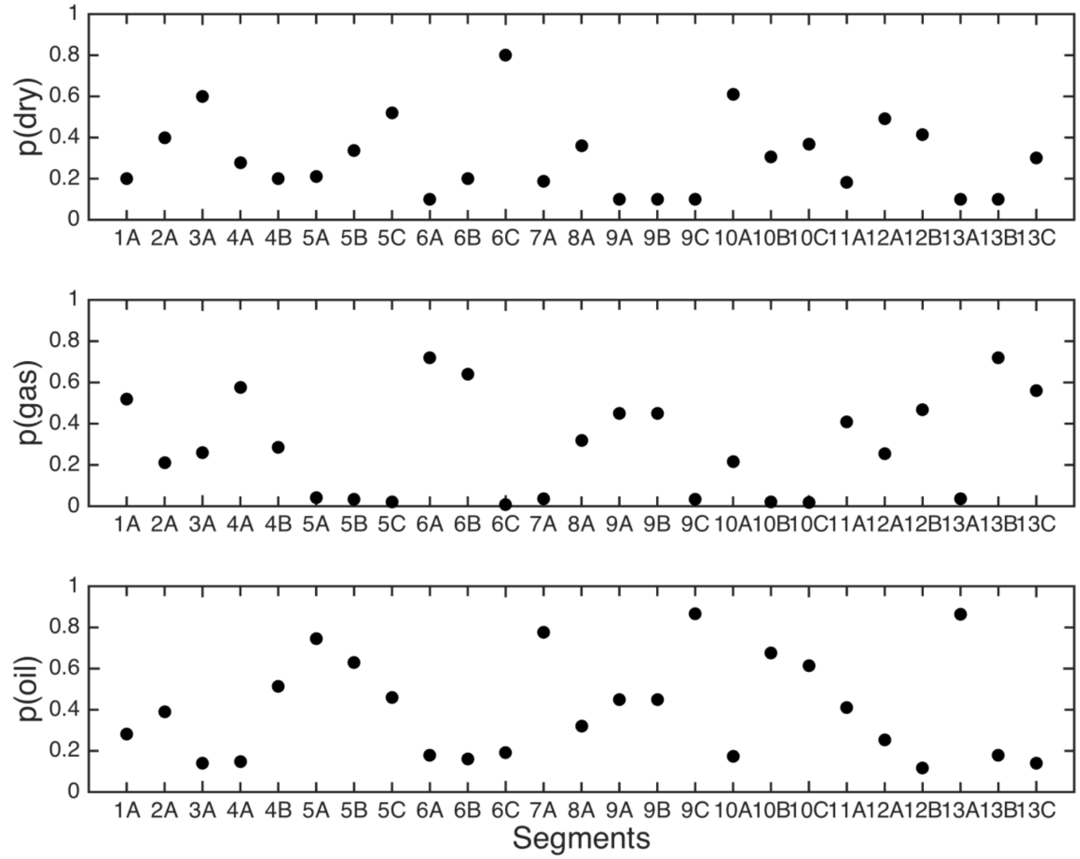


Prior marginal probabilities

.

Three possible classes at all nodes:

- Dry
- Gas
- Oil



Prior values

.

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

Development fixed cost. Infrastructure at prospect r.

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

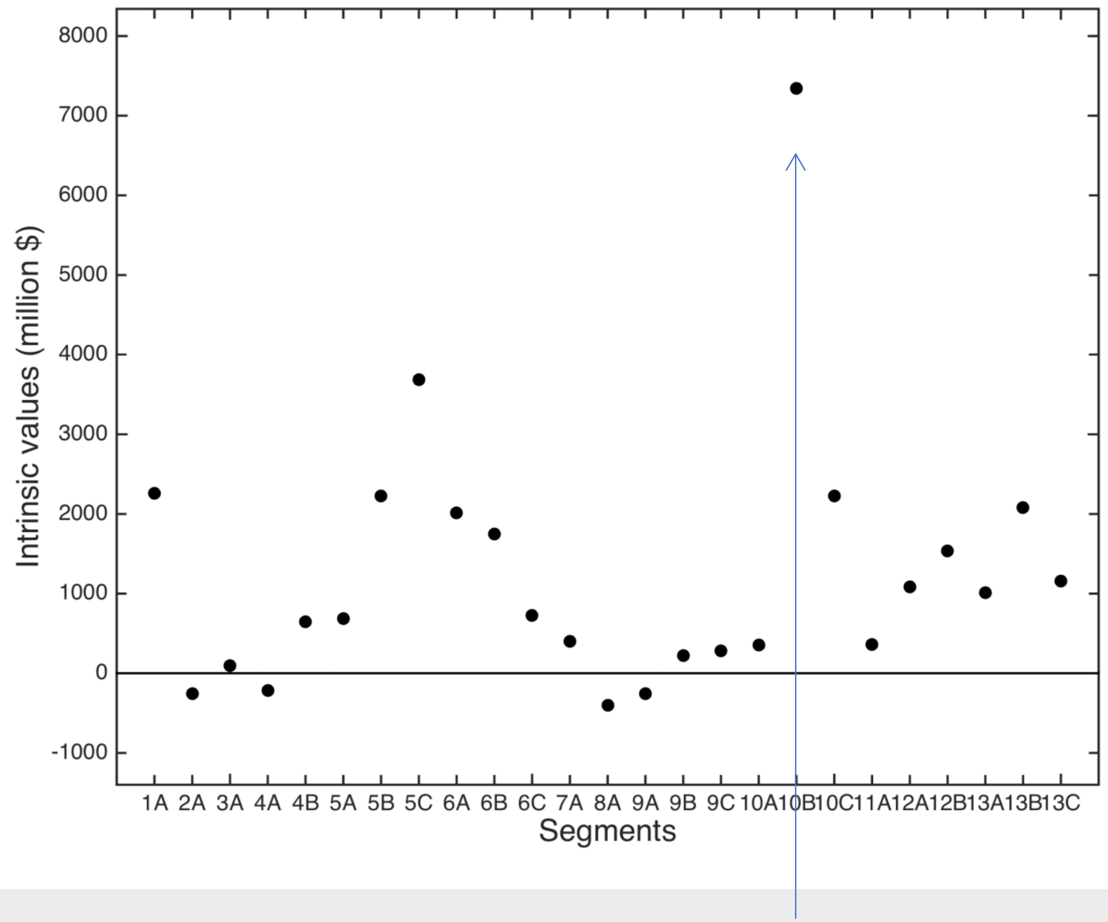
Cost if dry, 0 otherwise.

Revenues of oil/gas, 0 otherwise.

Cost of drilling segment i.

Values

.



Most lucrative. But might not be most informative.

Posterior values and VOI

.

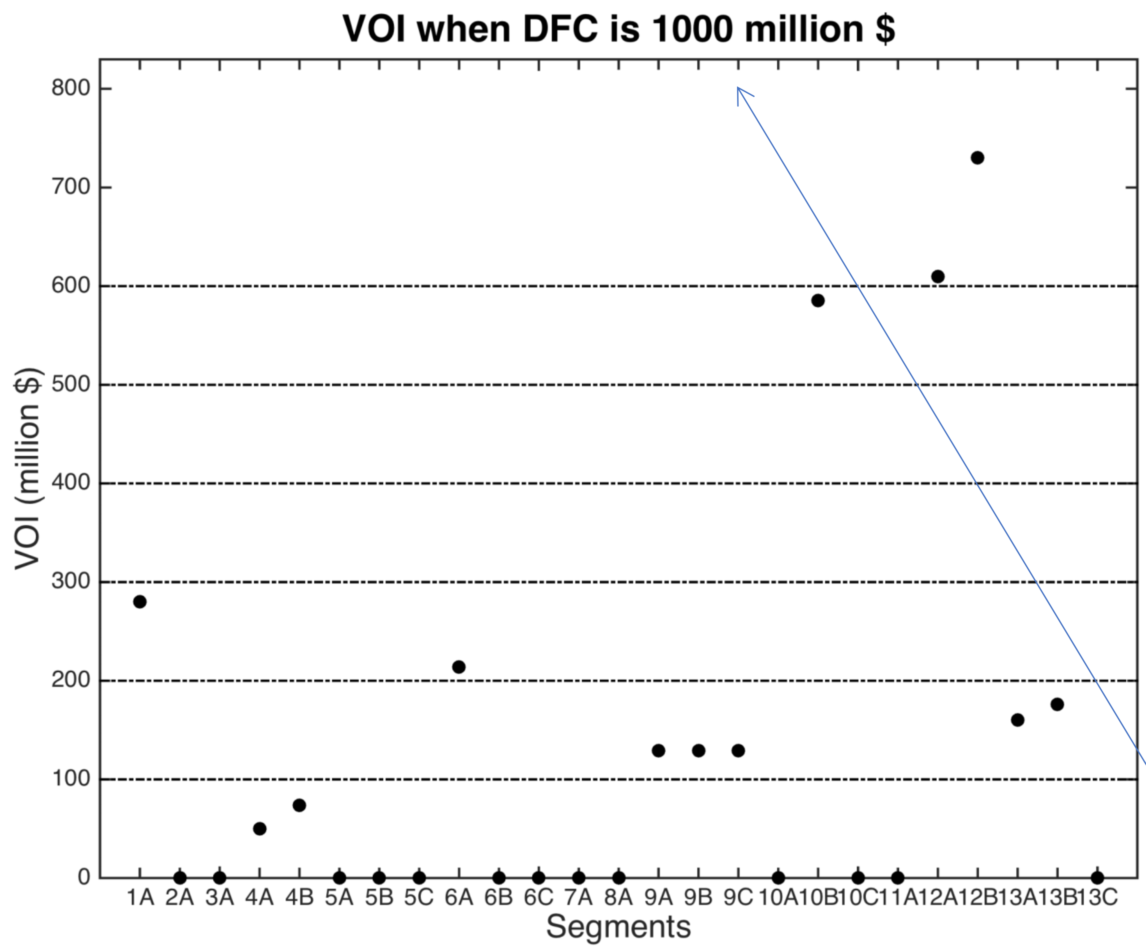
$$PoV(x_K) = \sum_{l=1}^{3} \sum_{r=1}^{13} \max\left\{0, \sum_{i \in \Pr} IV(x_i \mid x_K = l) - DFC\right\} p(x_K = l)$$

$$VOI(x_K) = PoV(x_K) - PV$$

Data acquired at single well.

VOI single wells

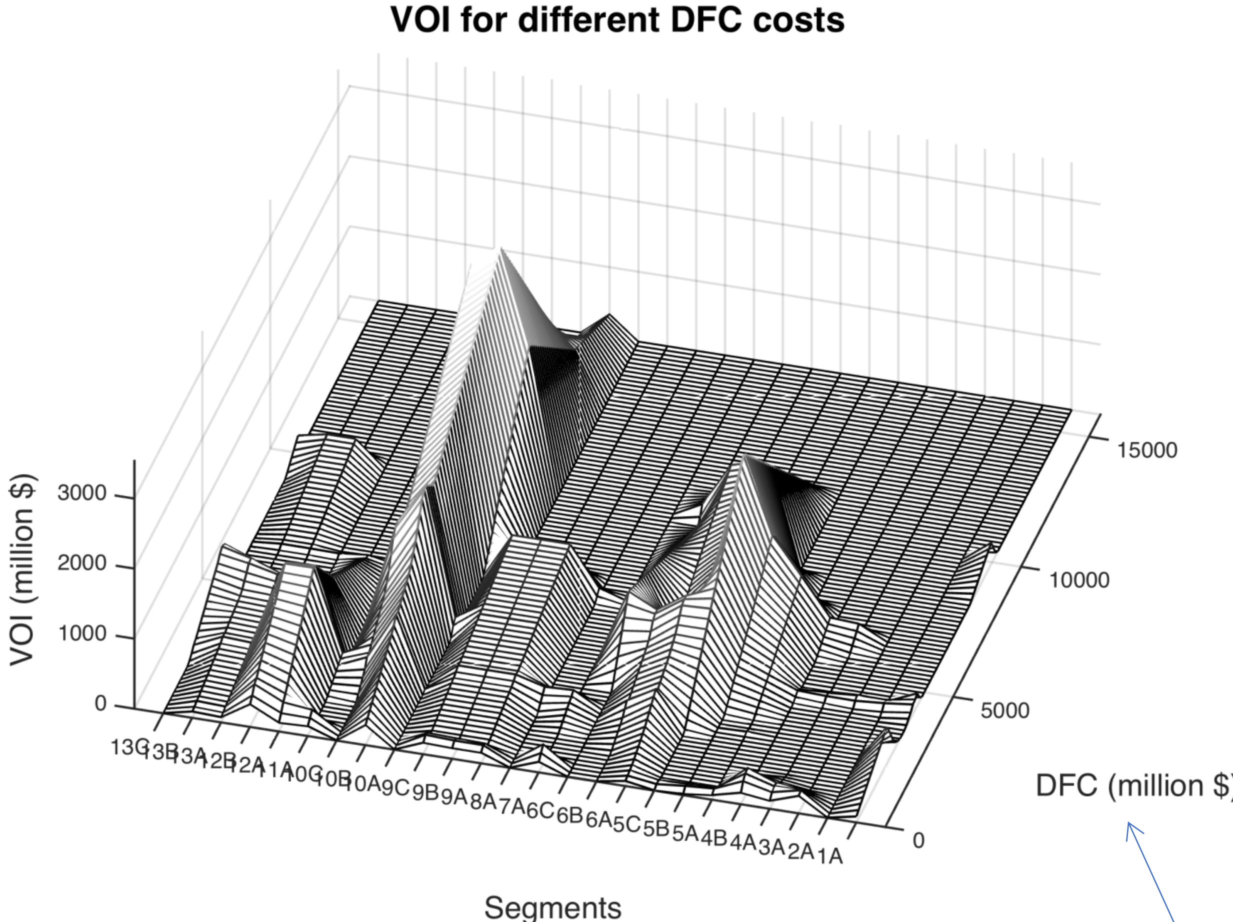
.



Development fixed cost.

VOI for different costs

.

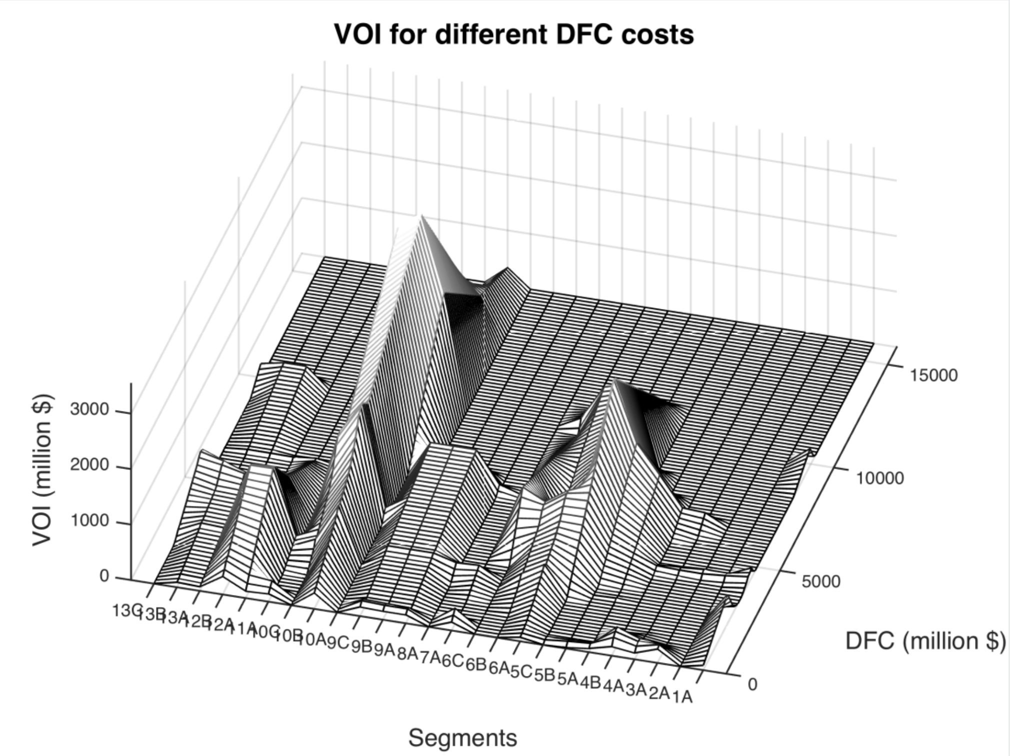


Development fixed cost.

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 example:

.

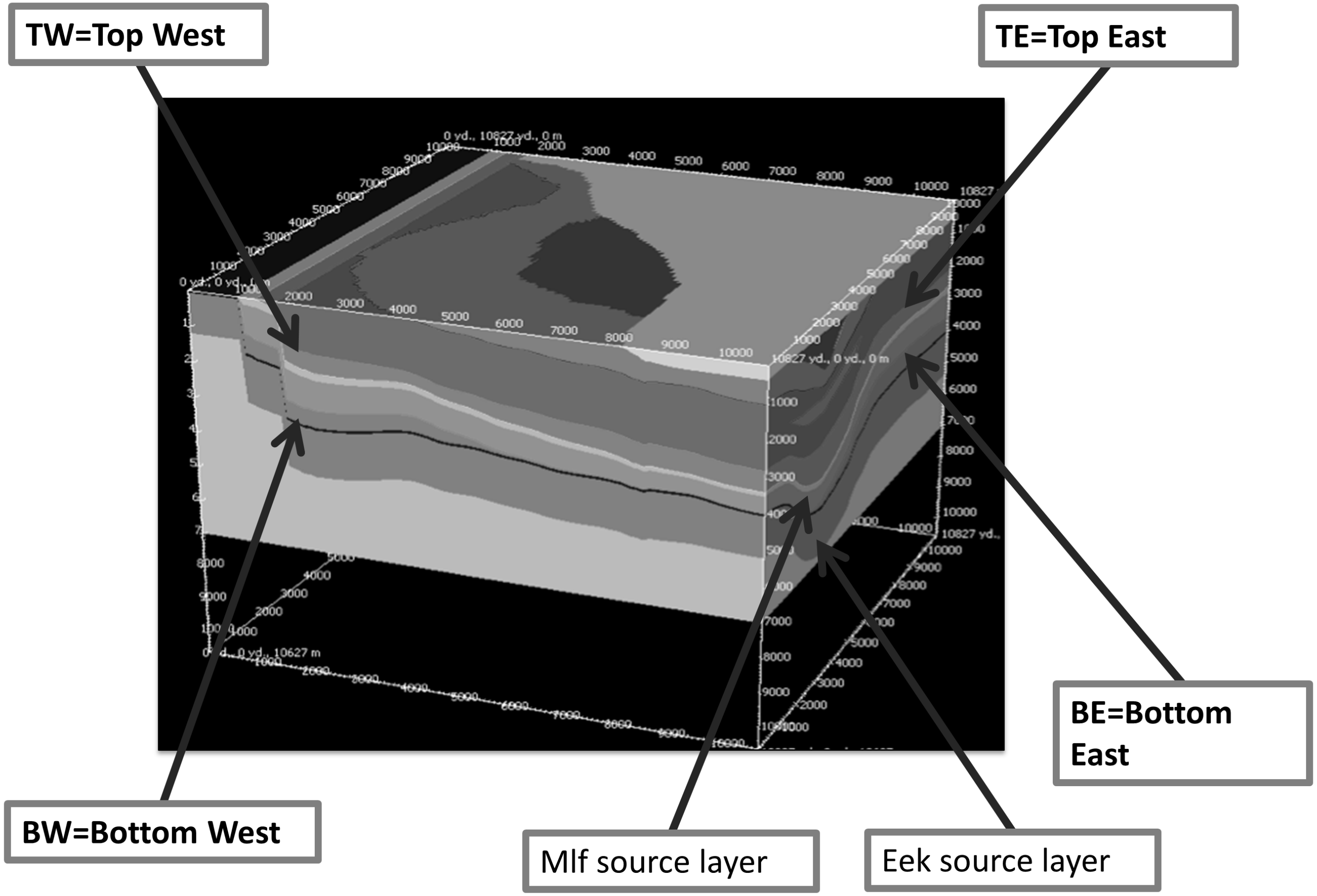
- VOI is not largest at the most lucrative prospects.
- VOI is largest where more data are likely to help us make better decisions.
- VOI also depends on whether the data gathering can influence neighboring segments – data propagate in the Bayesian network model.

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

Training BN models:

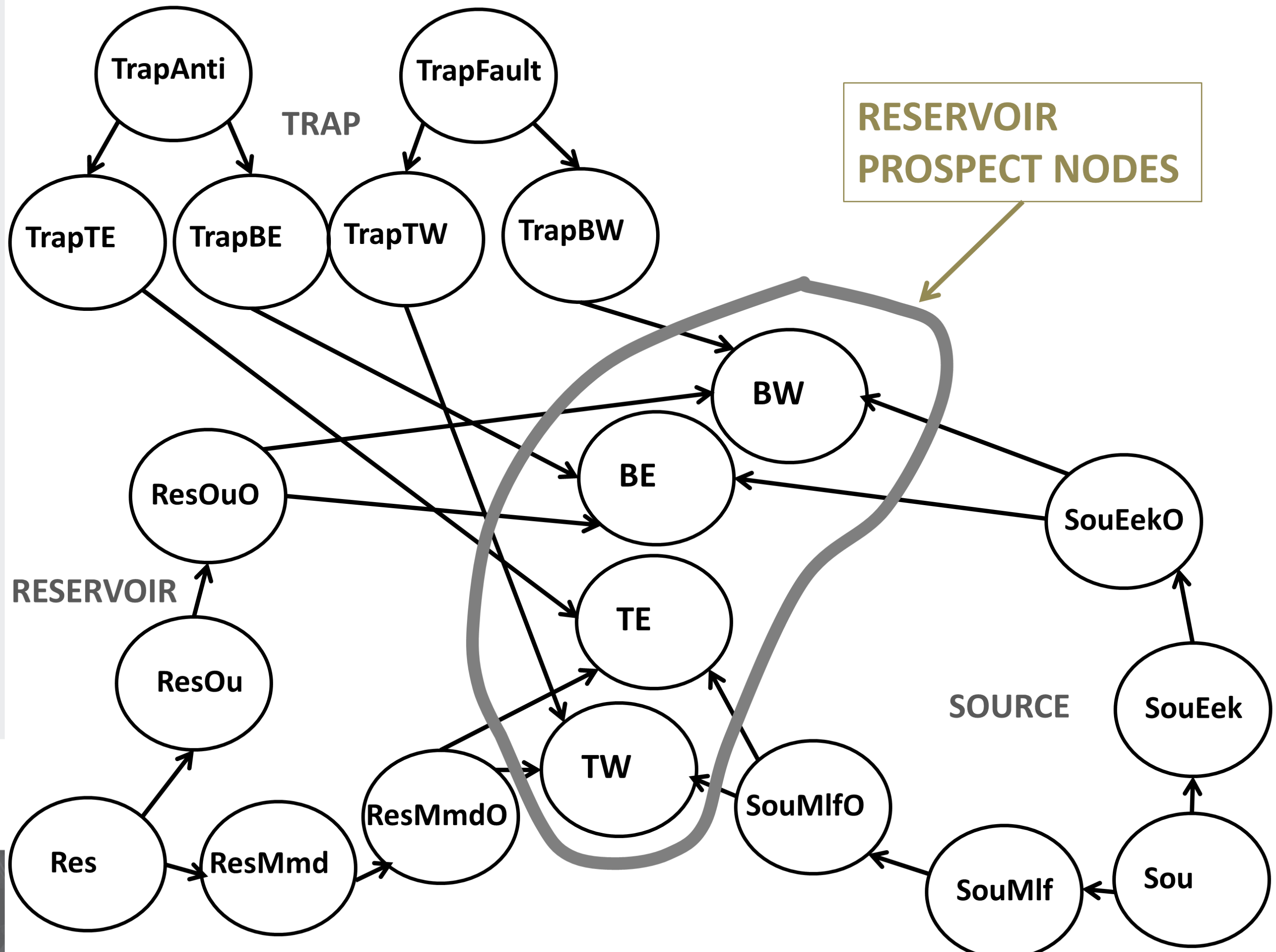
.

Simulated example using Petromod.



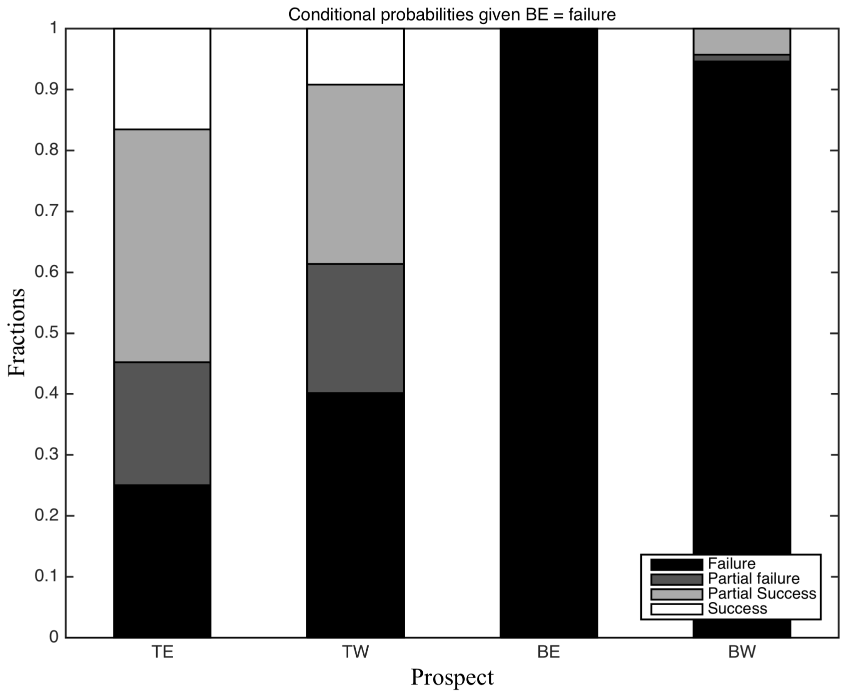
Training BN models:

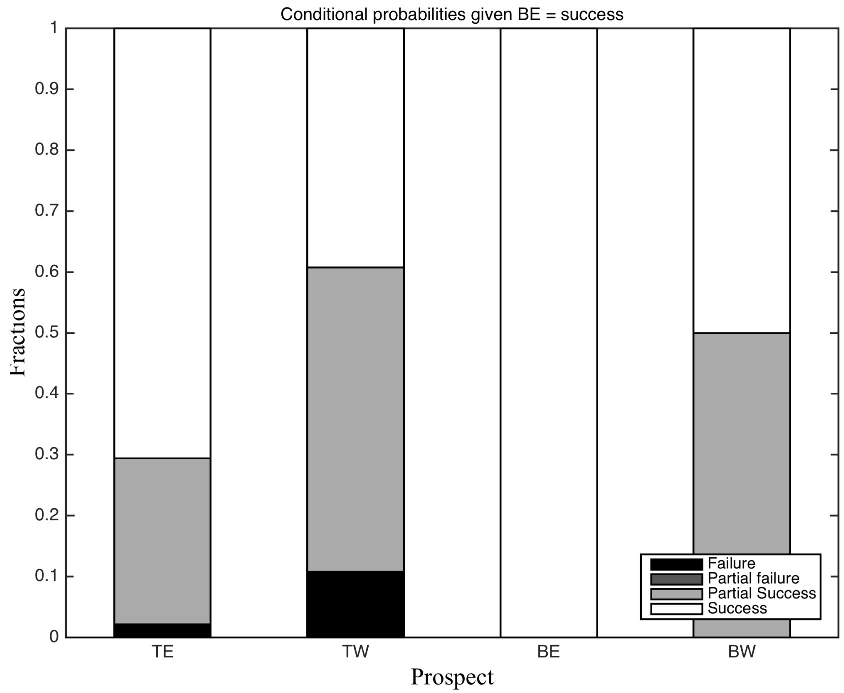
Build network from hydrocarbon generation and accumlations over several runs.



Conditional probabilities:

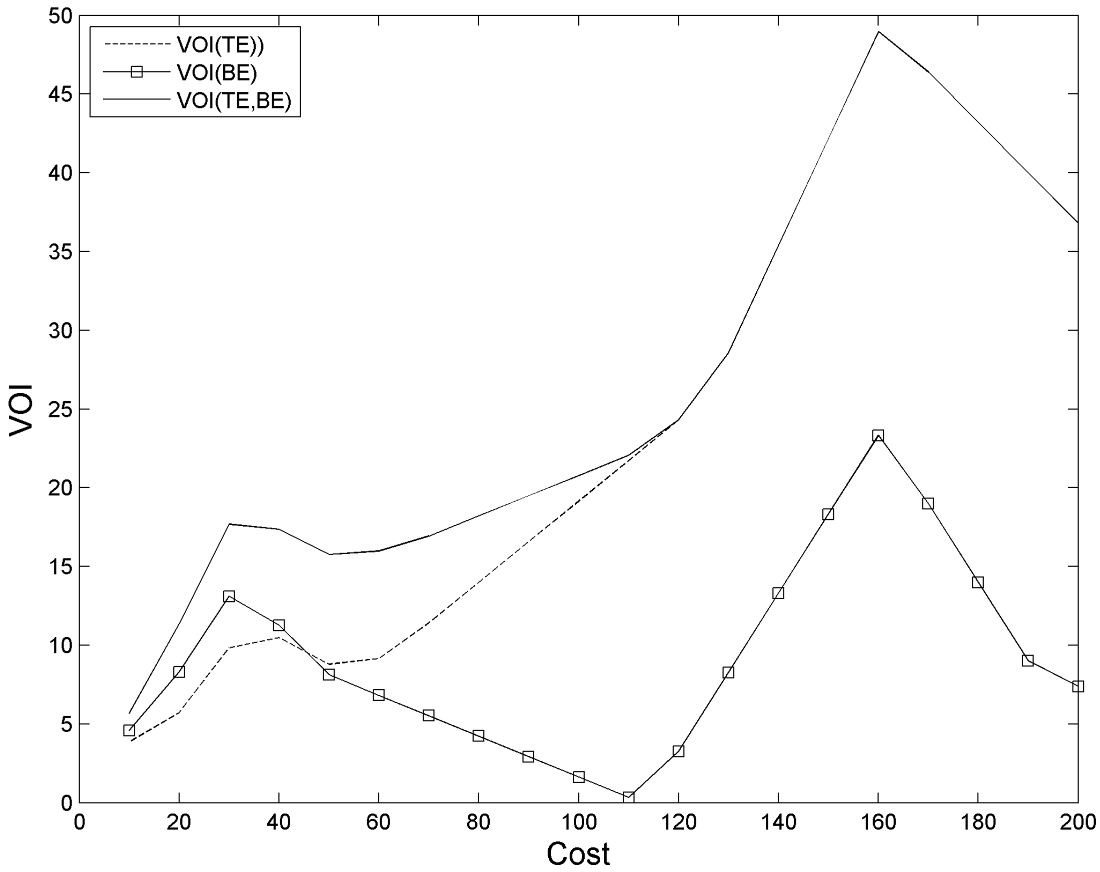
.





VOI results:

.



VOI(TE))
VOI(BE)
VOI(TE,BE)
VOI
Cost
0
5
10
15
20
25
30
35
40
45
50
0
20
40
60
80
100
120
140
160
180
200

Never break the chain - Markov models

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

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

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

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

$$d = 2$$

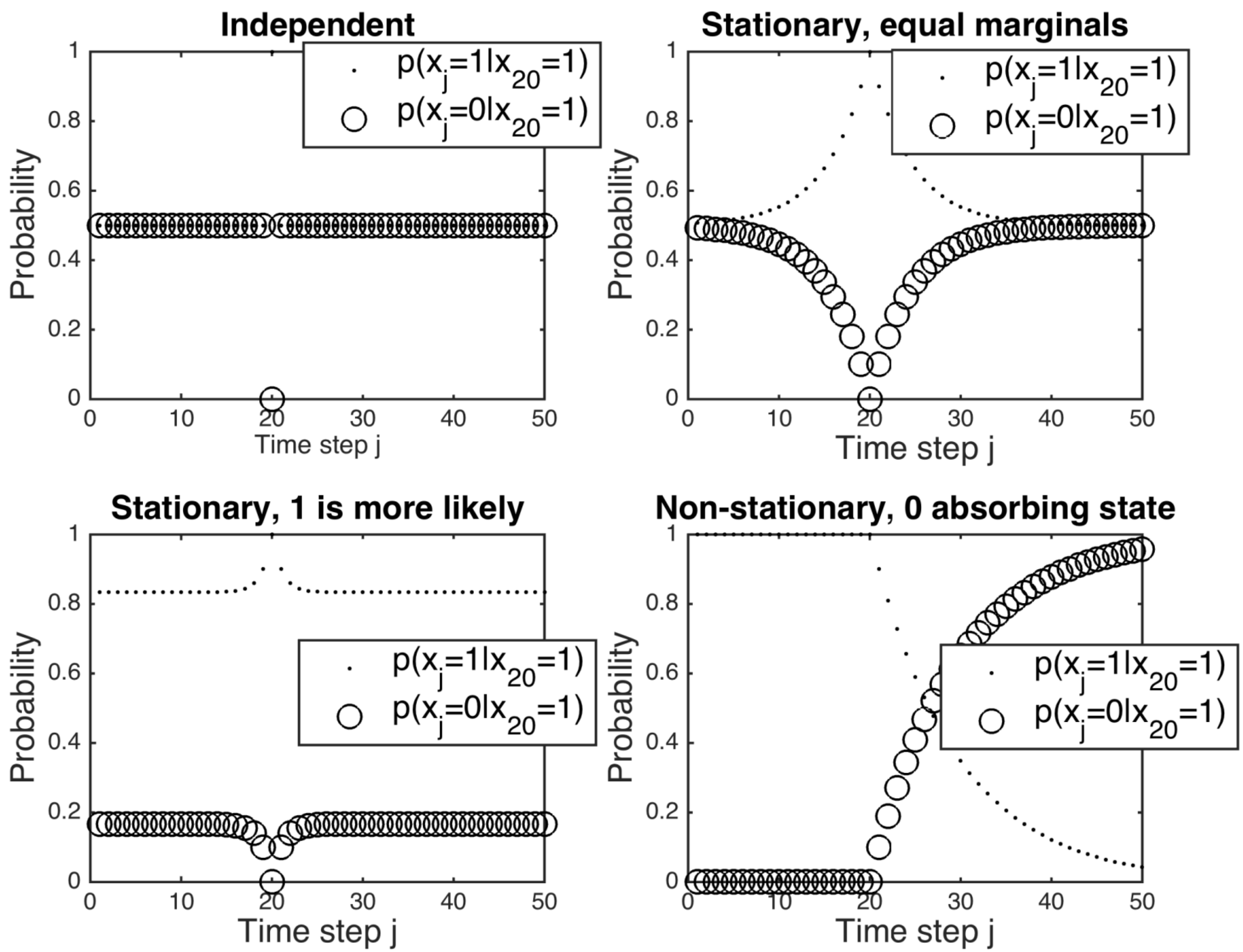
$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix} \quad P = \begin{bmatrix} 0.9 & 0.1 \\ 0.1 & 0.9 \end{bmatrix} \quad P = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix} \quad P = \begin{bmatrix} 1 & 0 \\ 0.1 & 0.9 \end{bmatrix}$$

Independence

Absorbing

EAGE EDUCATION

Markov chains (perfect information)



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:

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

$$C_1 = 0, C_2 = 5, C_3 = 20, C_4 = 40,$$

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

Prior value:

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

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

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

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.

Take home from this example:

.

- VOI varies with data gathering location
- Plan sensor locations wisely.

- VOI is largest when data are likely to help us make better decisions.

Plan for course

Time	Topic
Monday	Introduction and motivating examples
	Elementary decision analysis and the value of information
	Multivariate statistical modeling, dependence, graphs
	Value of information analysis for models with dependence
Tuesday	Examples of value of information analysis in Earth sciences
	Statistics for Earth sciences, spatial design of experiments
	Computational aspects
	Sequential decisions and sequential information gathering

Small problem sets along the way.