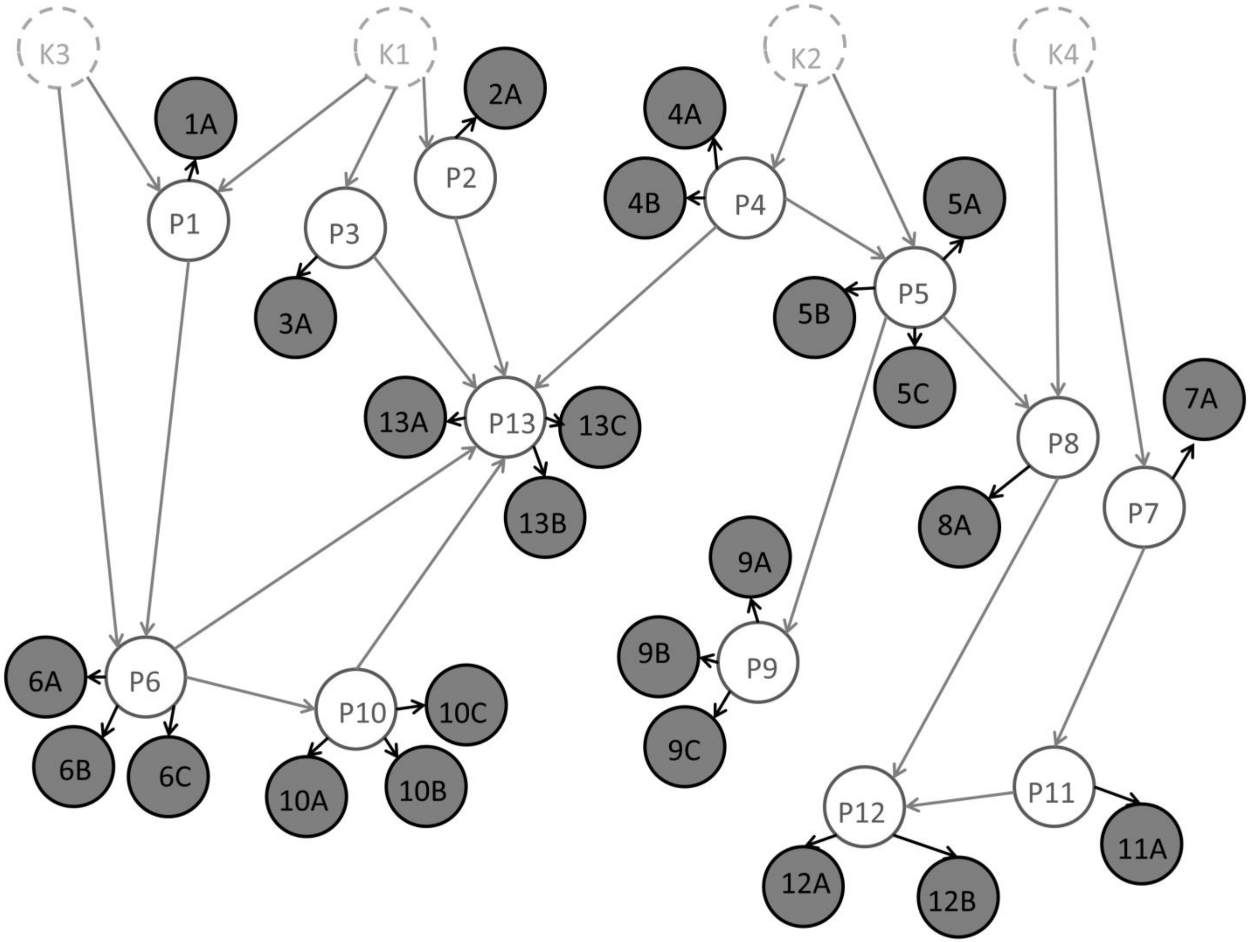
Decision Analysis and Value of Information

Jo Eidsvik, NTNU

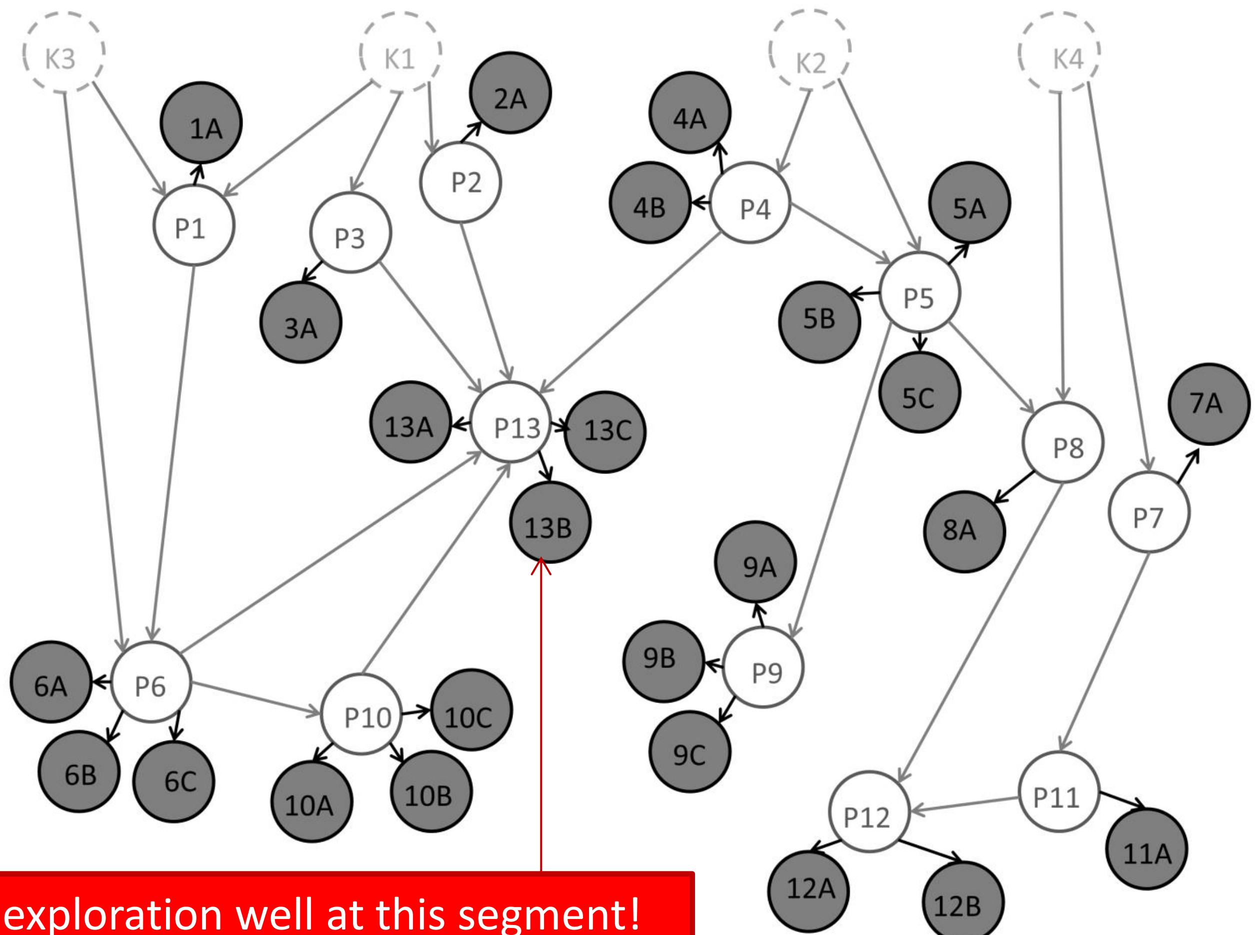
Motivation
(a petroleum exploration example)



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

Motivation
(a petroleum exploration example)

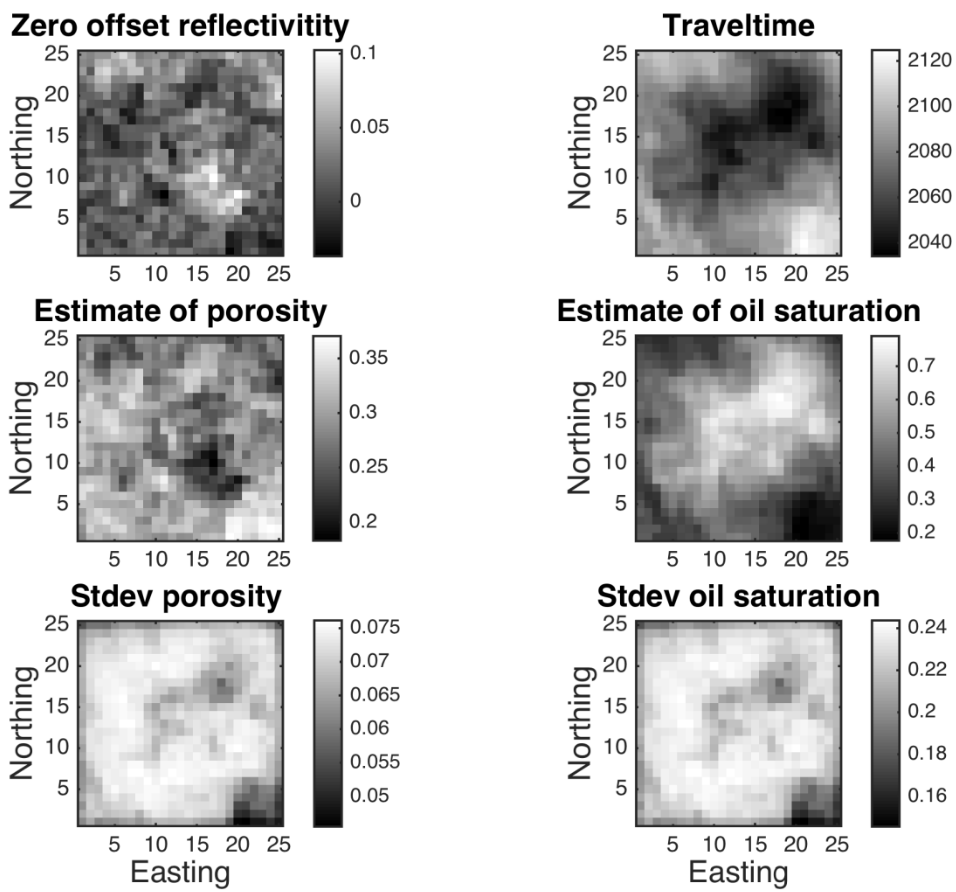
Gray nodes are petroleum reservoir segments where the company aims to develop profitable amounts of oil and gas.



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

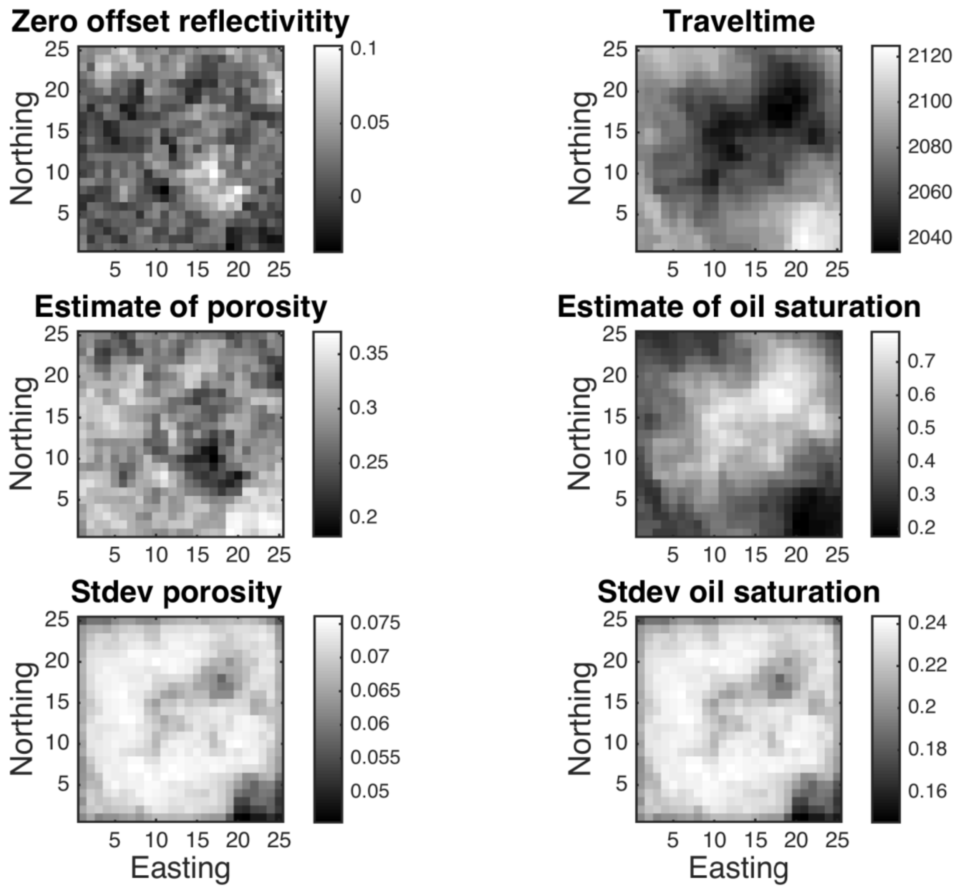
Motivation
(a petroleum development example)

Reservoir predictions from post-stack seismic data!



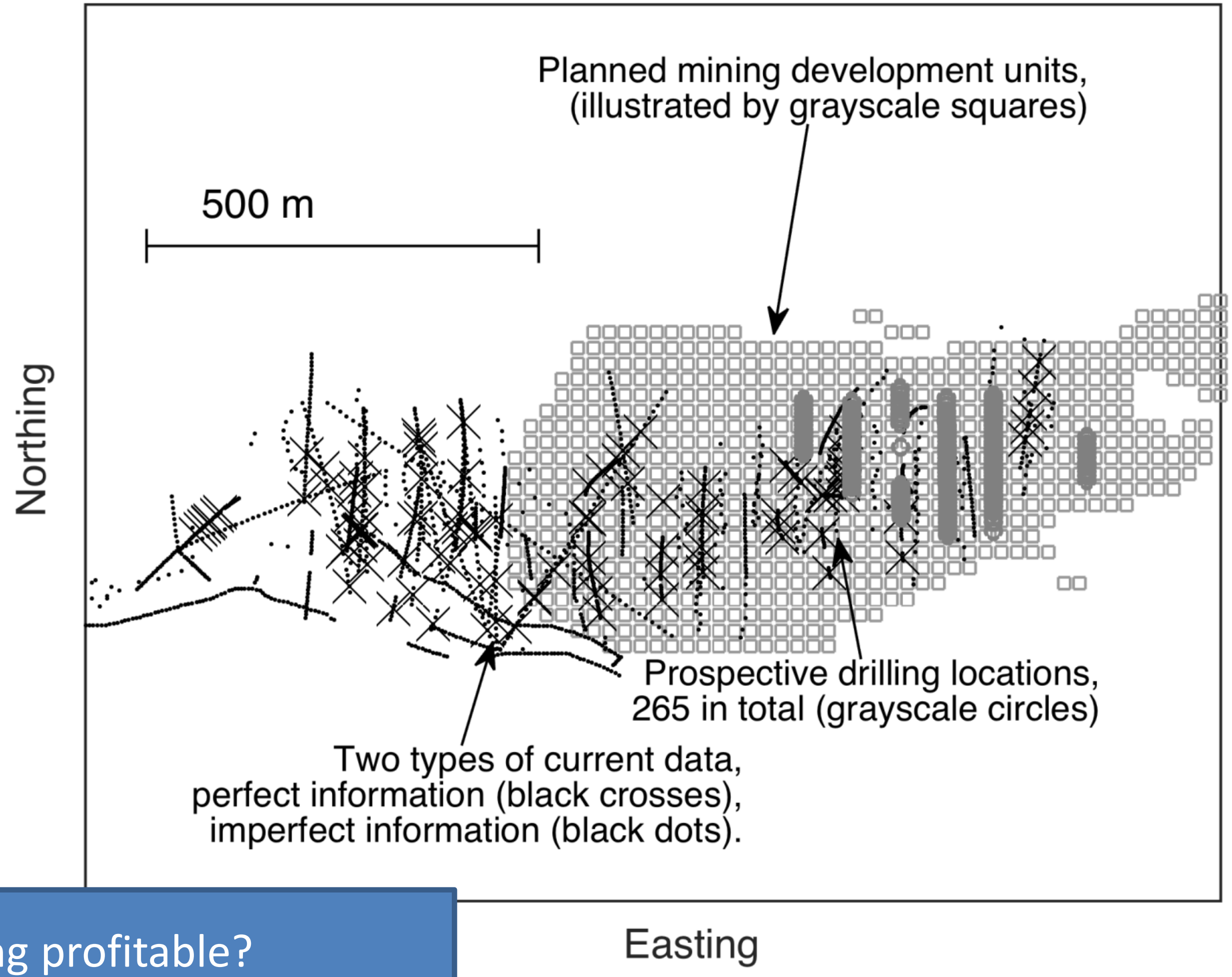
Motivation
(a petroleum development example)

Reservoir predictions from post-stack seismic data!



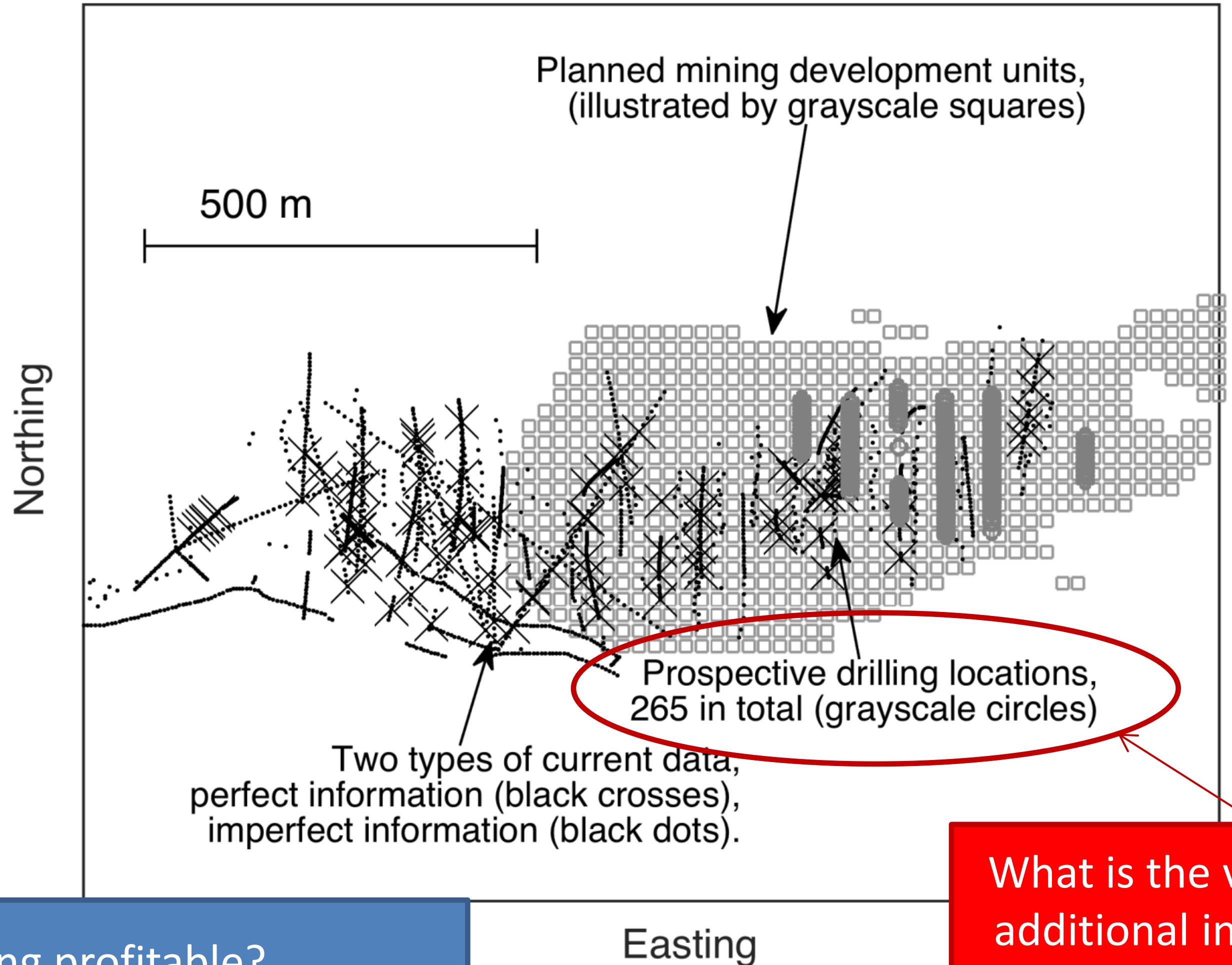
Process pre-stack seismic data, or electromagnetic data?

Motivation (an oxide mining example)



Is mining profitable?

Motivation (an oxide mining example)

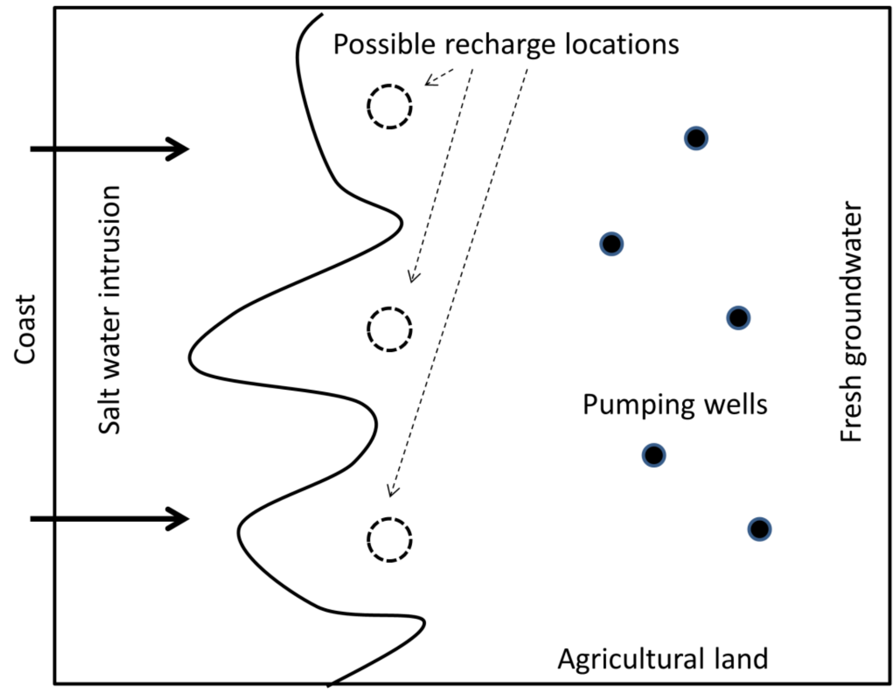


What is the value of this additional information?

Is mining profitable?

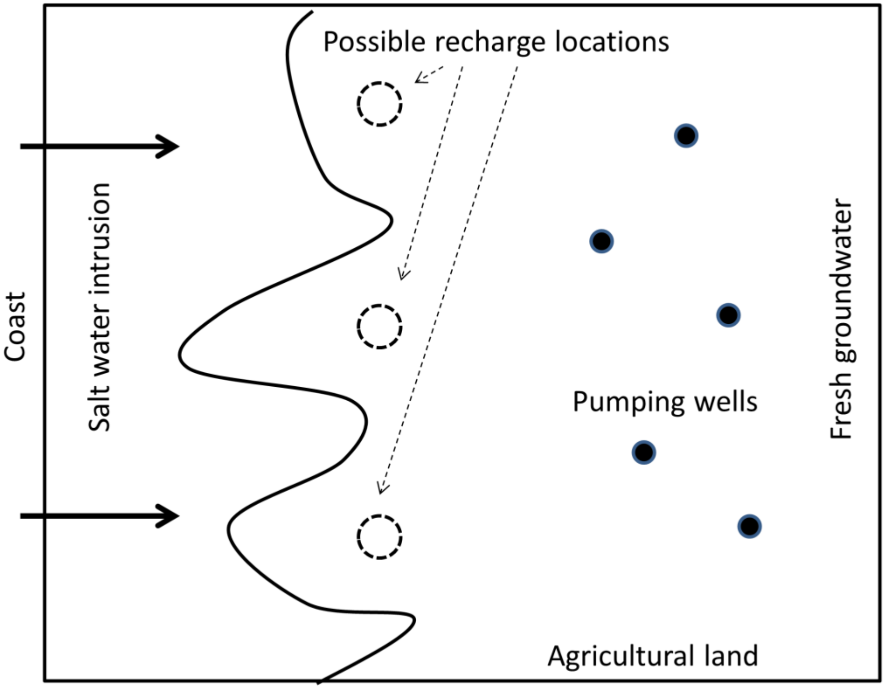
Motivation
(a groundwater example)

Which recharge location is better to prevent salt water intrusion?



Motivation
(a groundwater example)

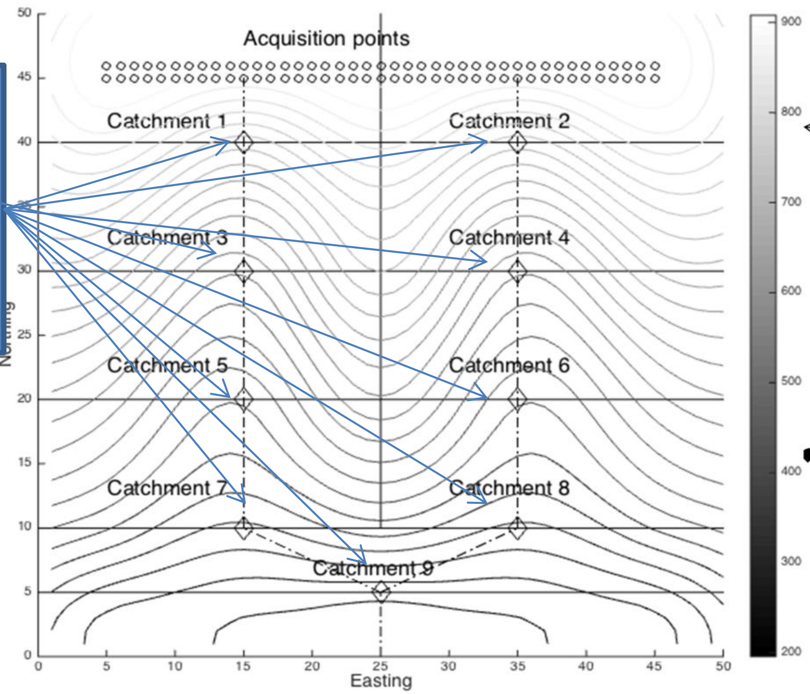
Which recharge location is better to prevent salt water intrusion?



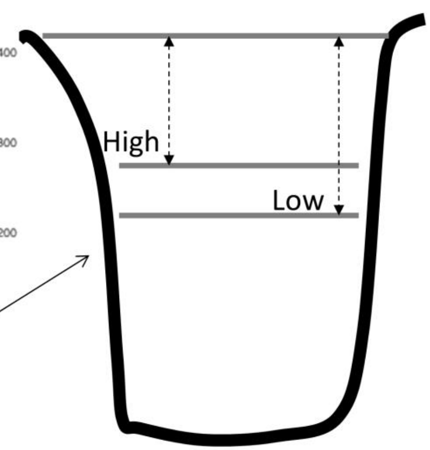
Is it worthwhile to acquire electromagnetic data before making the decision about recharge?

Motivation
(a hydropower example)

Adjusting water levels in 9 hydropower dams!



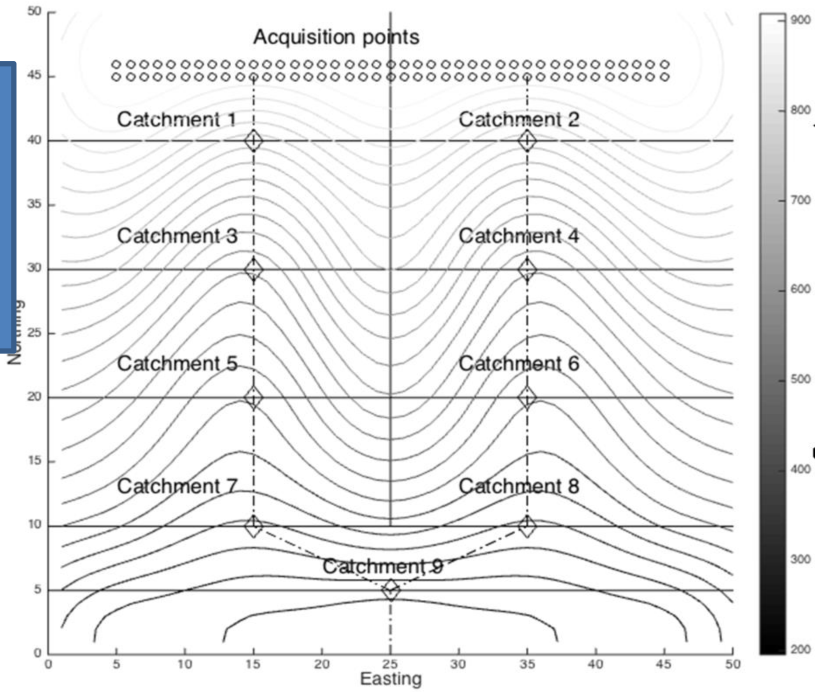
Map view illustration of a river basin. Dams are illustrated by diamonds.



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

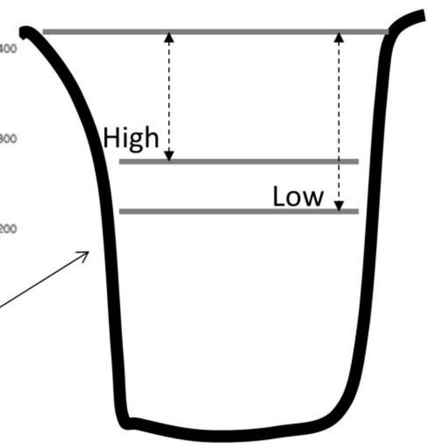
Motivation
(a hydropower example)

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?

Which data are valuable?

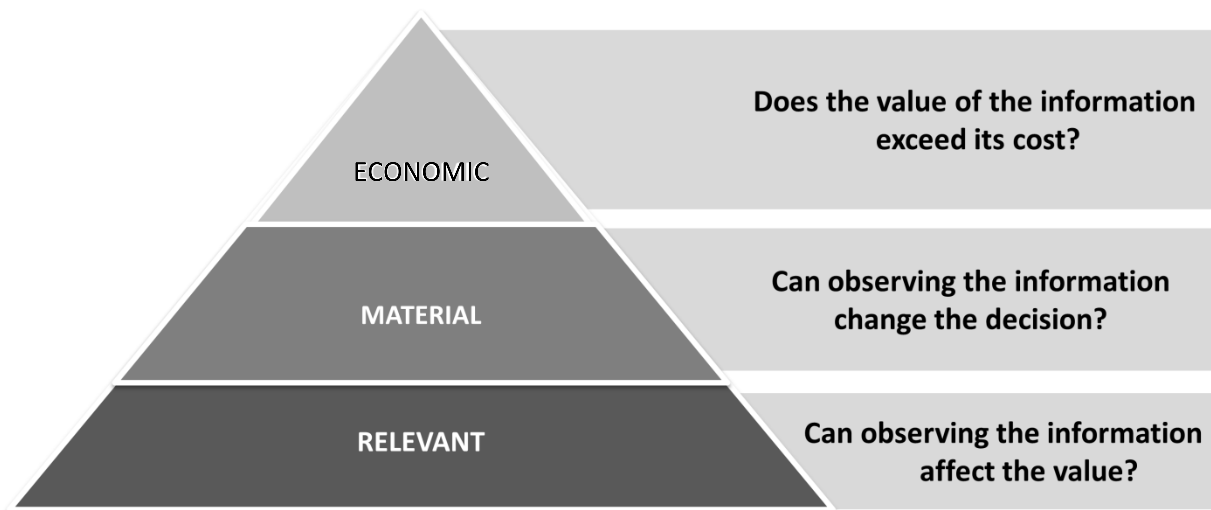
Five Vs of big data:

- Volume
- Variety
- Velocity
- Veracity
- Value

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

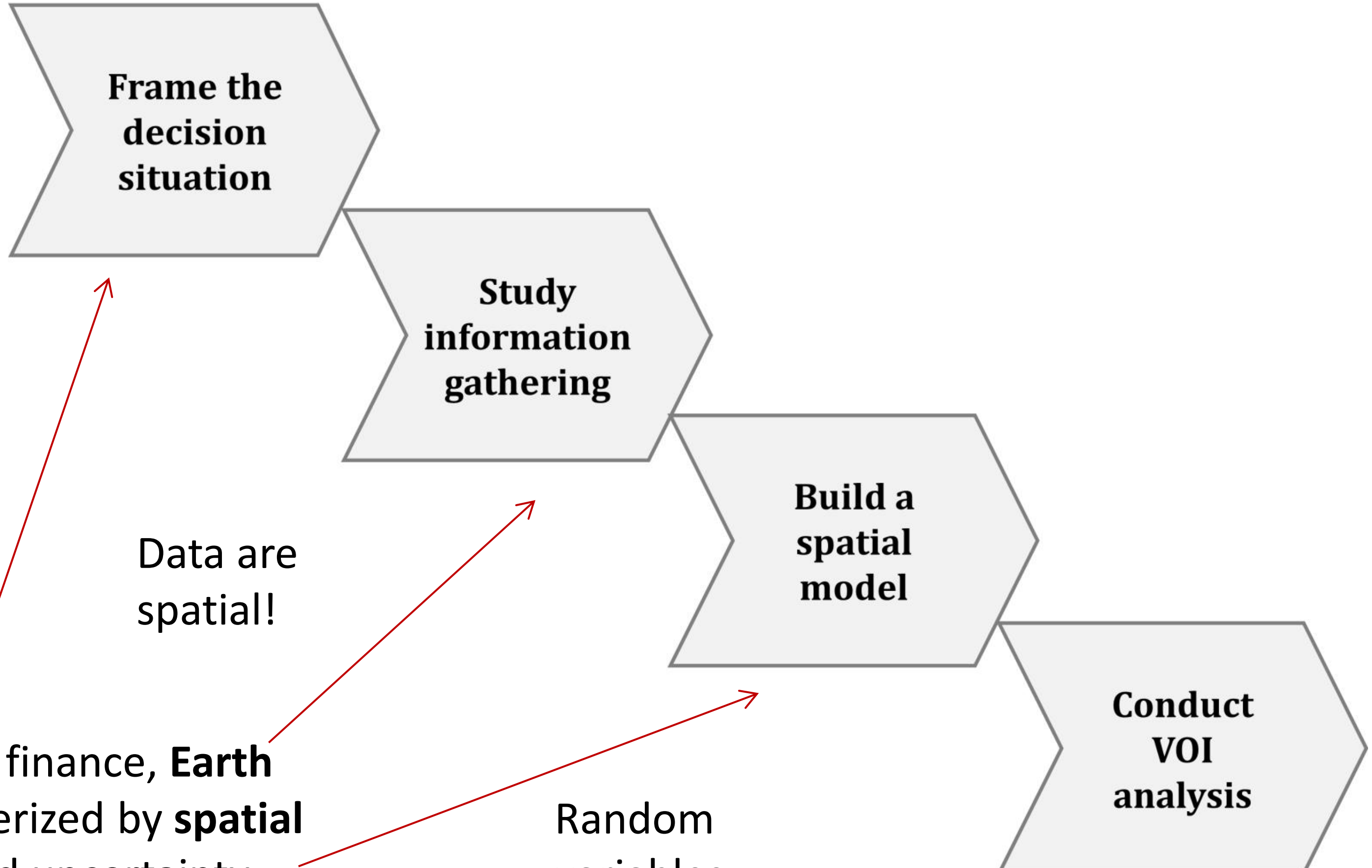
Value of information (VOI)

We often 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 reduction, variance reduction, prediction error, etc.)

VOI workflow



Alternatives are spatial!

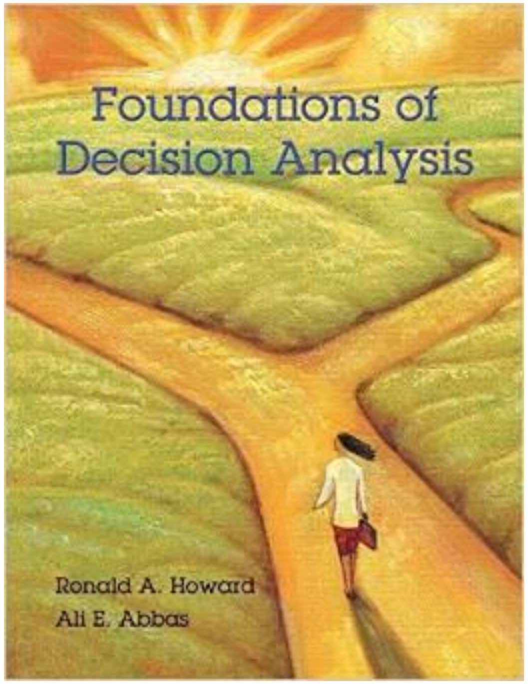
Data are spatial!

Random variables are spatial!

Unlike medicine and finance, **Earth sciences** are characterized by **spatial** alternatives, data and uncertainty. The VOI calculations must integrate these features realistically.

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

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 in any event.
- 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 in any event.
- Revenues if he finds the treasure .

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

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**

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

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

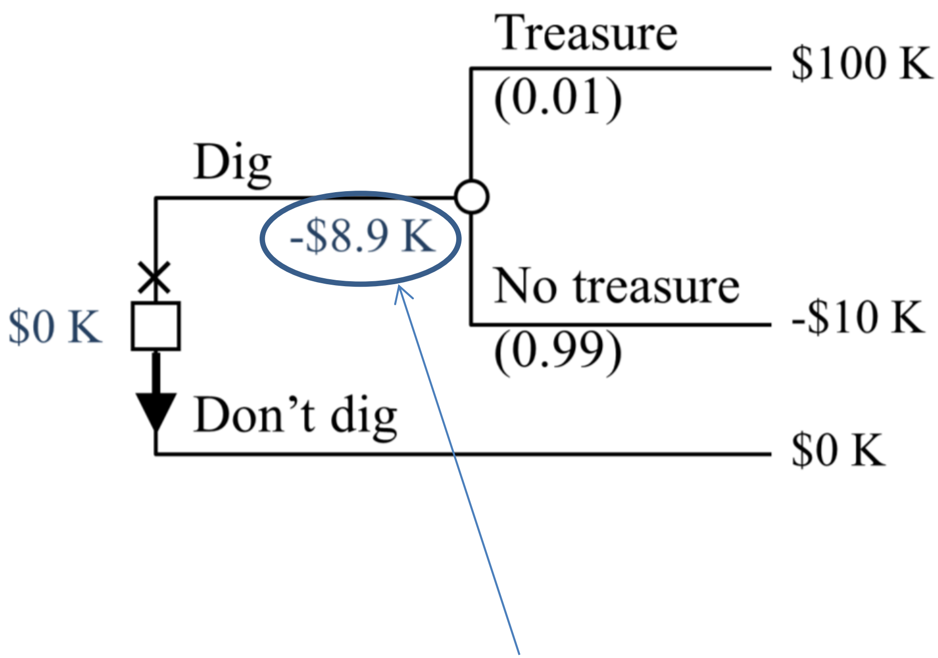
- **Risk profiles, utility function, certain equivalents**

$$u(v) = v \qquad u(v) = 1 - e^{-\gamma v}$$

- **Maximize expected utility**

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

Pirate's decision situation

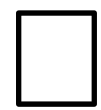


Risk neutral!

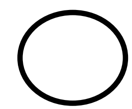
$$E\left(u\left(v_{dig}\right)\right) = E\left(v_{dig}\right) = 0.01(100000) + 0.99(-10000) = -8900$$

Decision trees

A way of structuring and illustrating a decision situation.



- Squares represent decisions
- Circles represent uncertainties

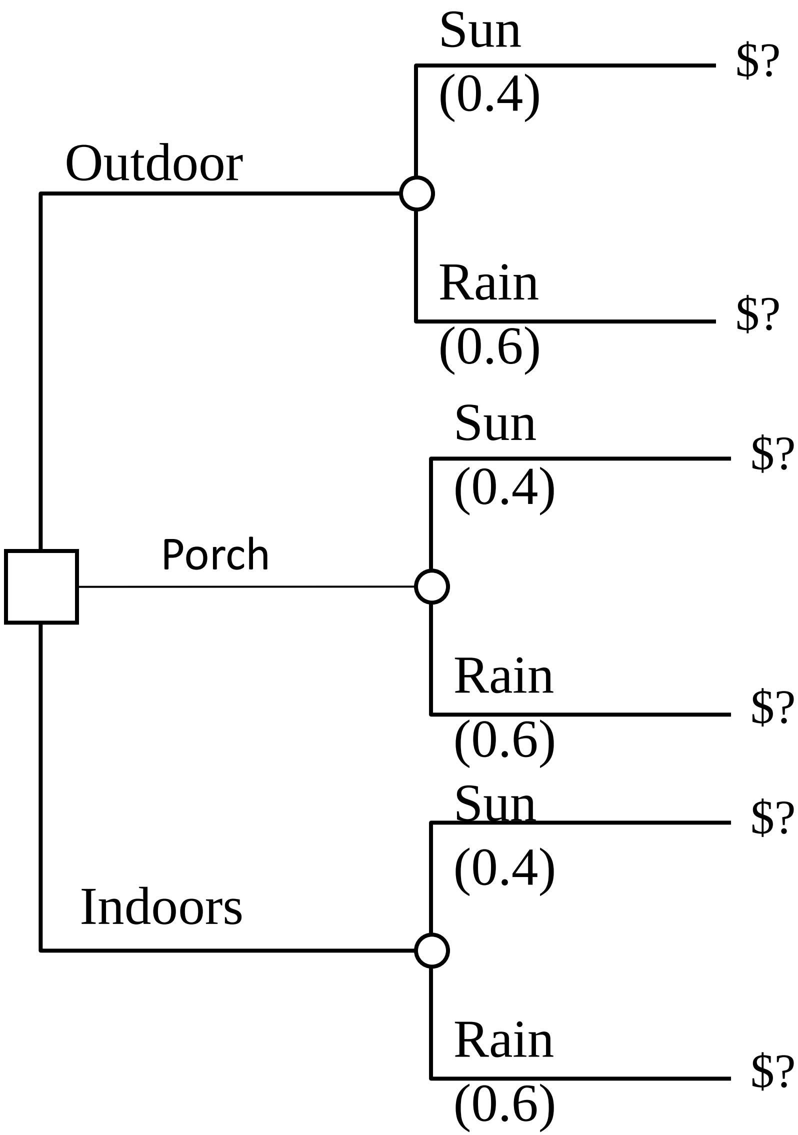


- Probabilities and values are shown by numbers.

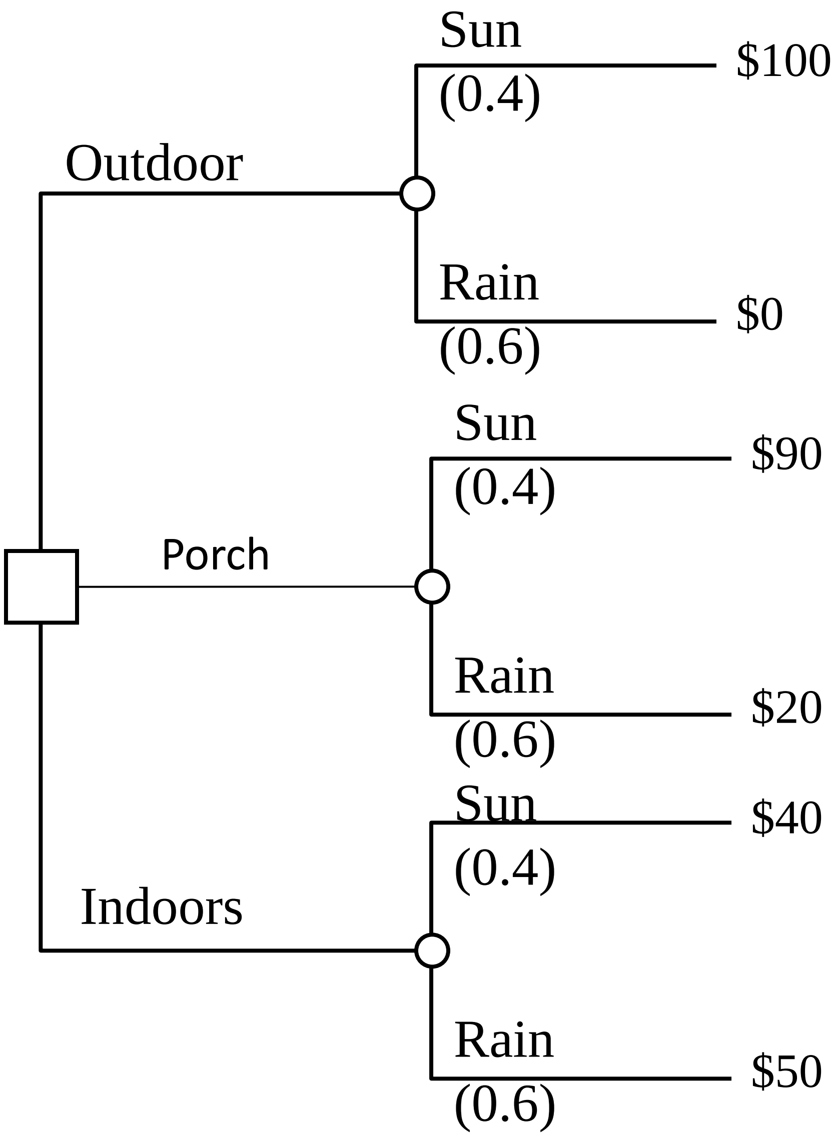
- Arrows indicate the optimal decision.

Kim's party problem

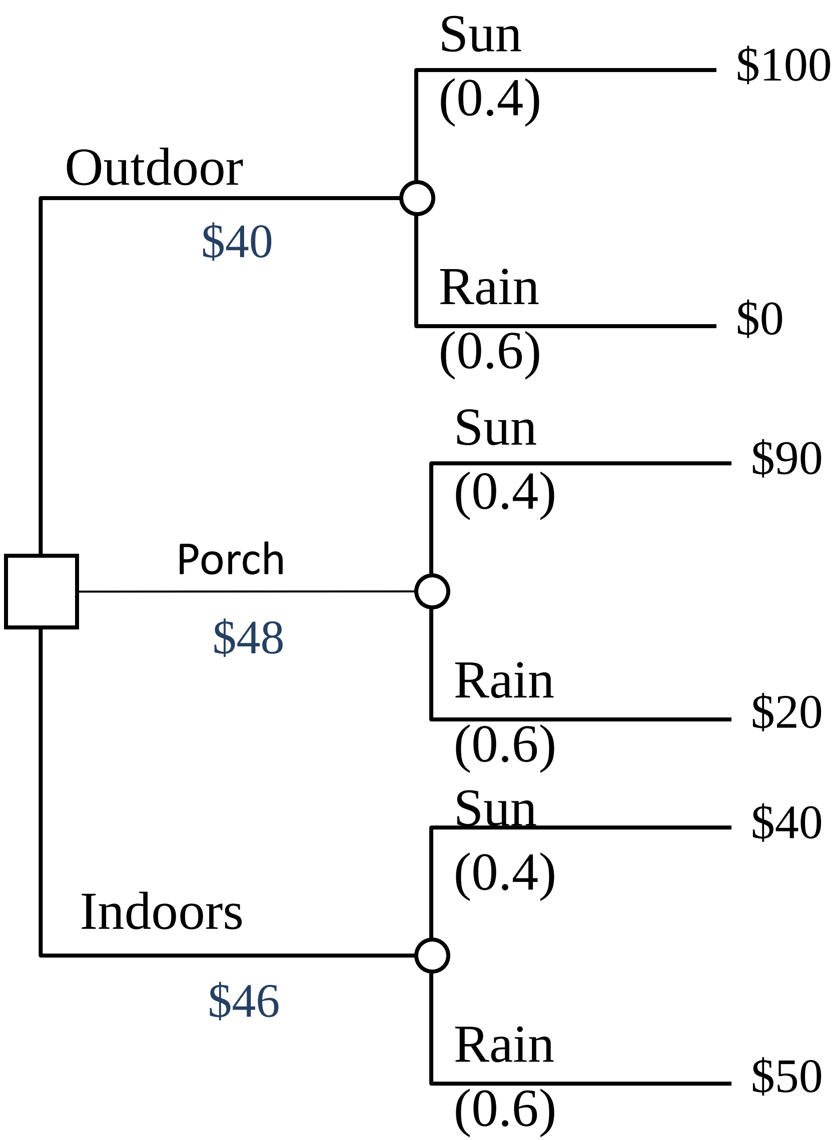
Howard, R.A. and Abbas, A., 2015, *Foundations of Decision Analysis*, Prentice Hall.



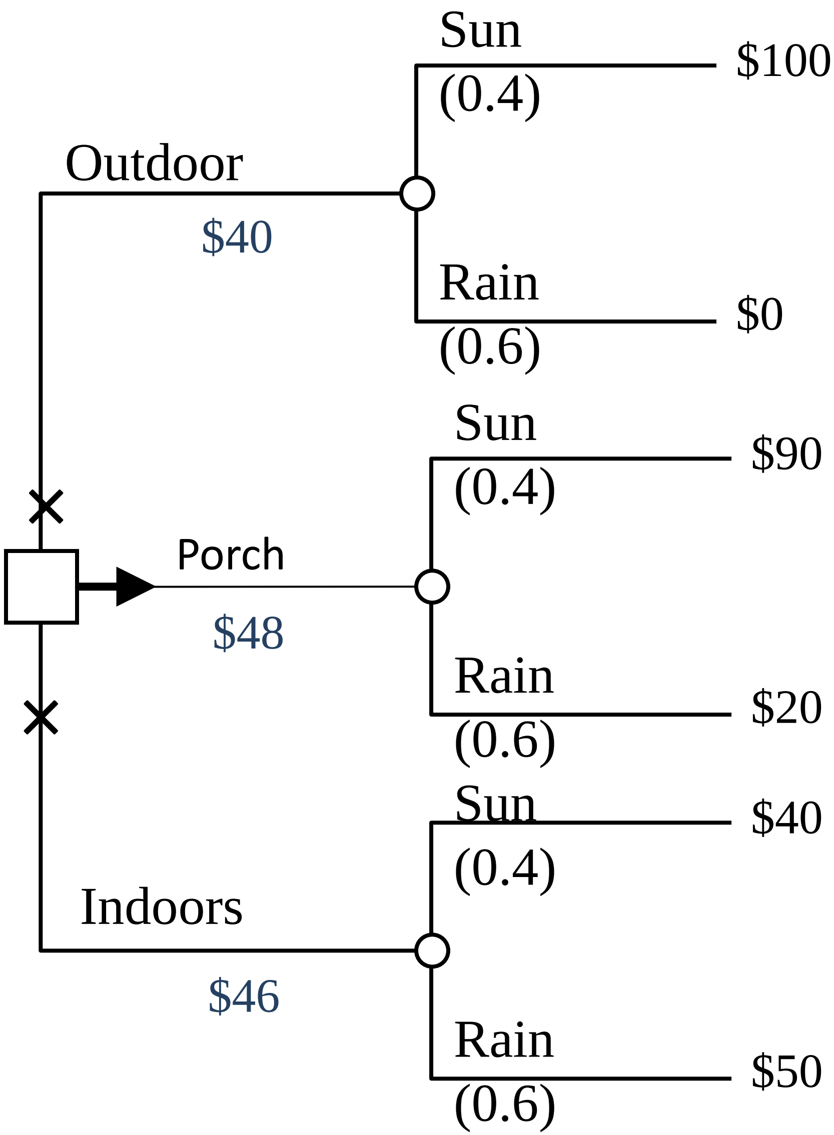
Kim’s party problem



Kim’s party problem



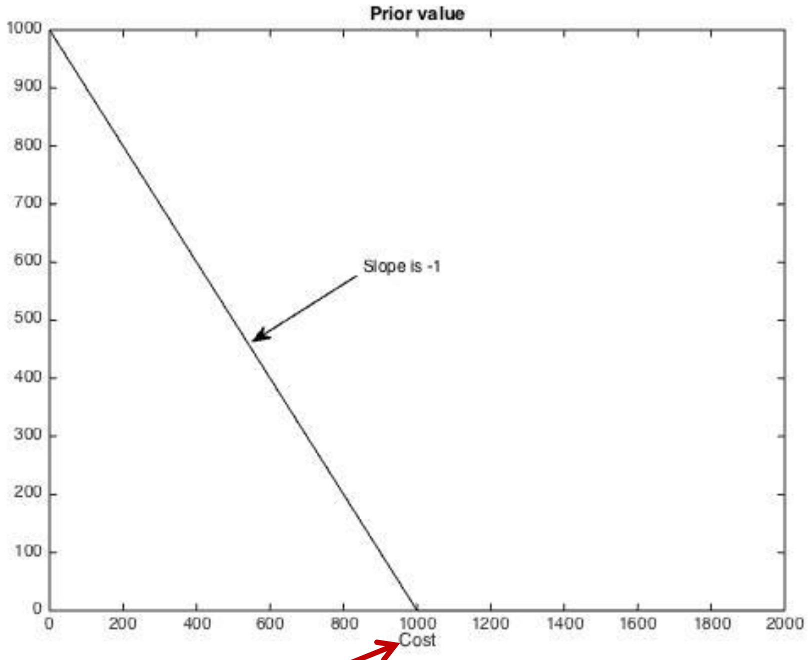
Kim’s party problem



Visualization - Prior value for pirate

$$p(x=1) = p = 1/100 = 0.01$$

$$PV = \max\{0, 100000 \cdot p - \text{Cost}\}$$



The pirate is indifferent for digging Cost=1000!
Digging for treasure (with uncertain profit, 100000 or -1000) has same expected value as not digging (with 0 profits for sure).

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

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

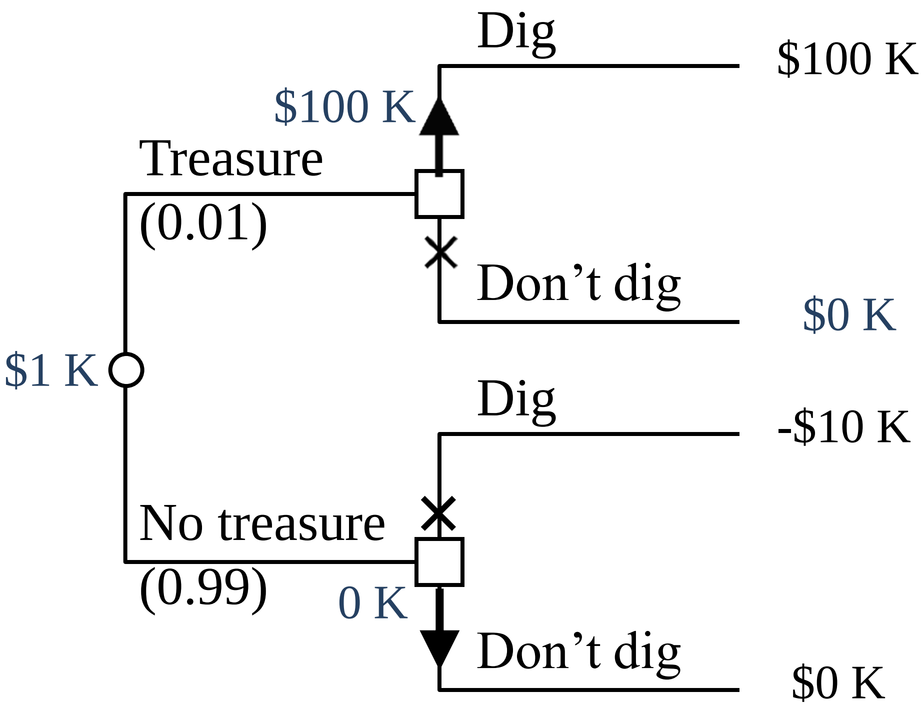
$$PoV(x) = \sum_{x} \max_{a \in A} \{v(x,a)\} p(x)$$

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

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

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 **utility** or **value**!

- Digging cost in any event.
- 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 **utility** or **value**!

- Digging cost in any event.
- Revenues if he finds the treasure .

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

Detector experiment

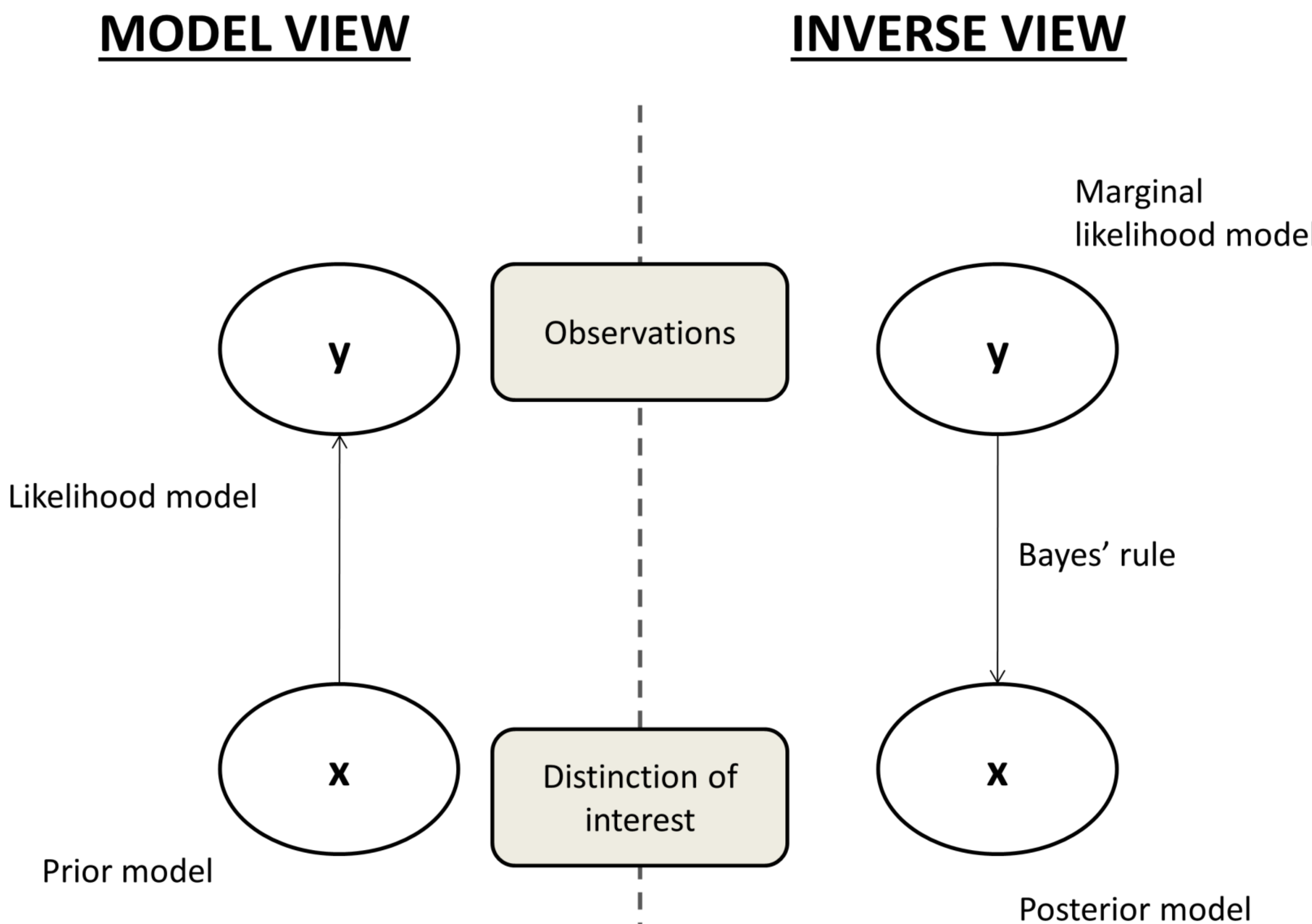
Accuracy of test:

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

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:

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

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

VOI – Pirate considers detector test

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

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

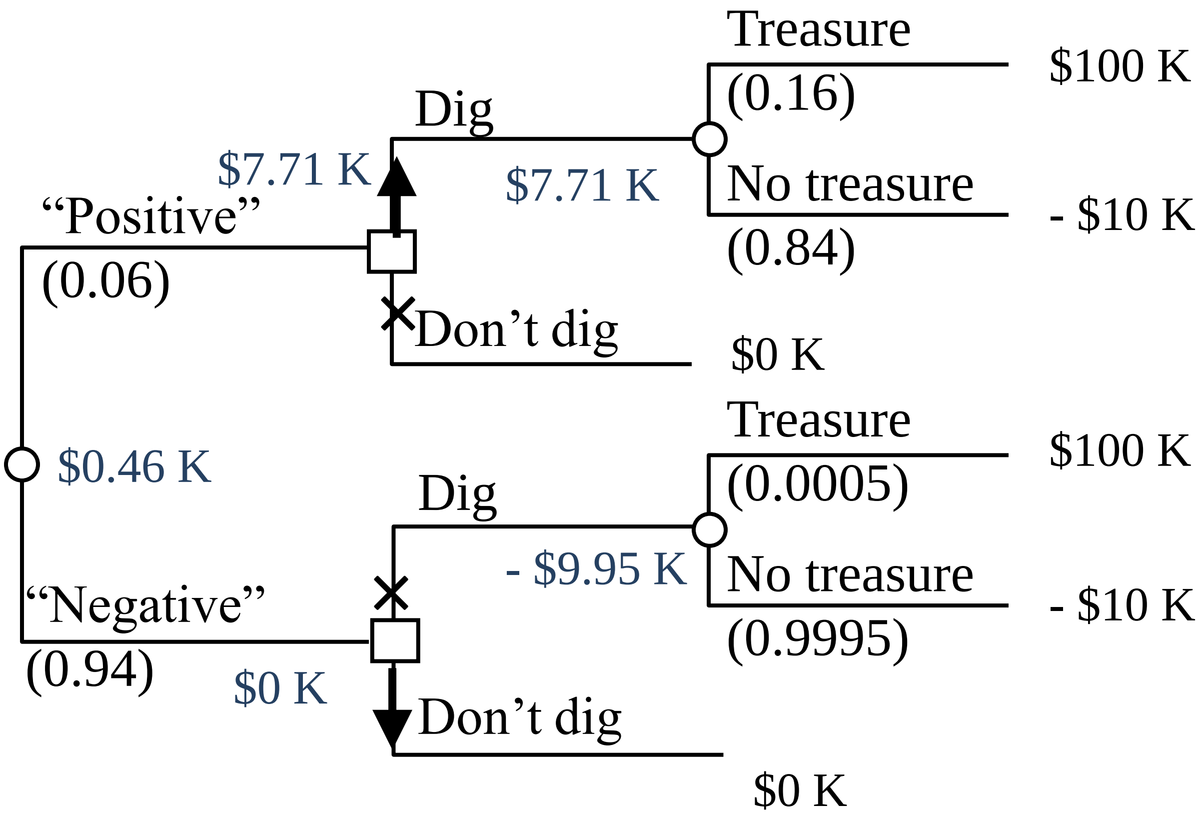
$$+ \left(0.94 \cdot \max\left\{0, \left(100 \cdot 0.0005\right) + \left(-10 \cdot 0.9995\right)\right\}\right)$$

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

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

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

PoV - imperfect information

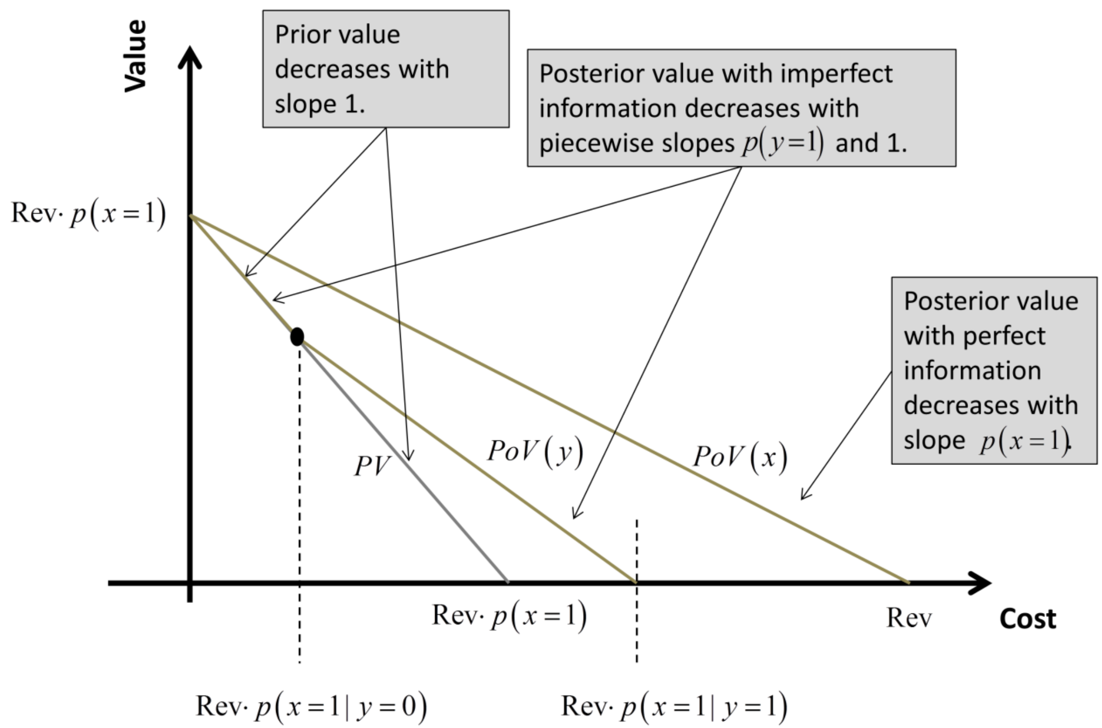


PV and PoV as a function of Digging Cost

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

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

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

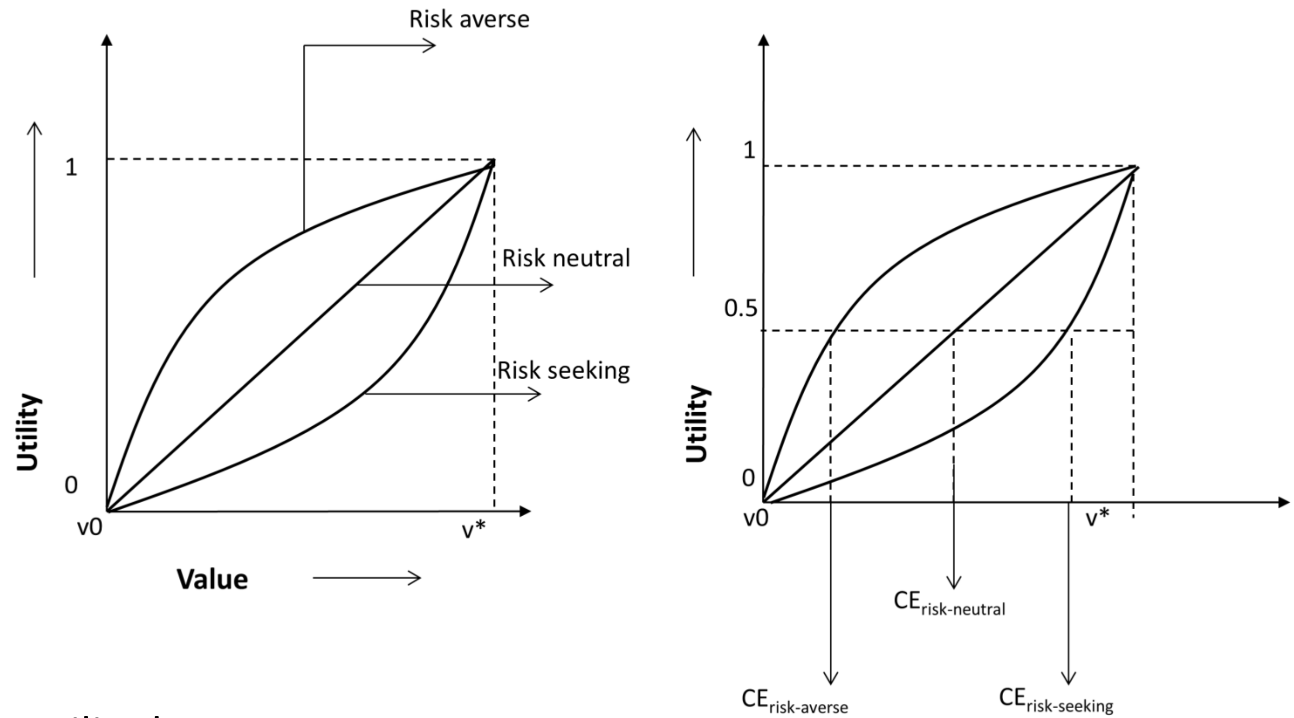


Value of information (VOI) - More general formulation

- 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

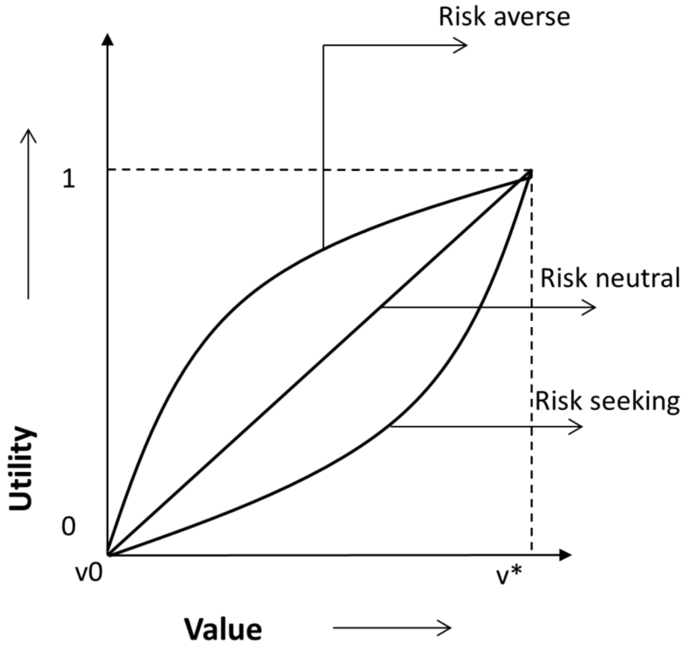


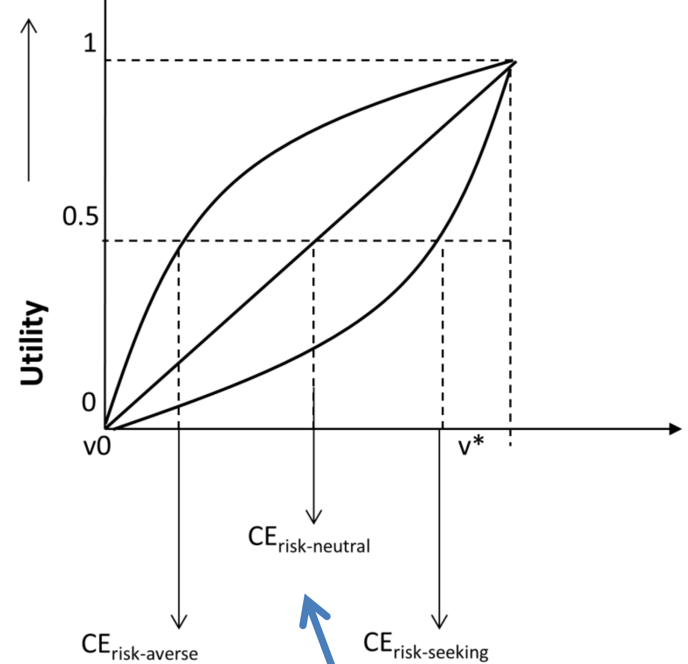
Exponential and linear utility have constant risk aversion coefficient:

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

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!

Value of information- Imperfect

Price of indifference.

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

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

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

VOI=Posterior value – Prior value

Assuming risk neutral decision maker!

Properties of VOI

a) VOI is always positive

- Data allow better, informed decisions.

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

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:

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

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).

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

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

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

VOI for CO_2 sequestration

The decision maker can proceed with CO_2 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 CO_2 may leak (x=1). If leakage occurs, there will be a fine of 60 monetary units (i.e. a cost of 90 in total).

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

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

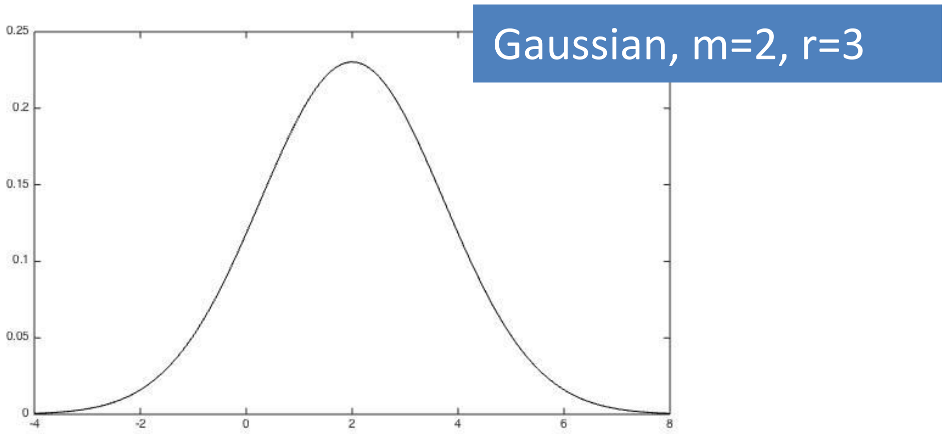
$$p(y=0|x=0)=0.95 \qquad p(y=1|x=1)=0.9$$

EXERCISE:

1. **Draw the decision tree without information.**
2. **Draw the decision tree with perfect information (clairvoyance).**
3. **Compute the VOI of perfect information.**
4. **Draw the decision tree with the geophysical experiment.**
5. **Compute conditional probabilities, expected values and the VOI of geophysical data.**

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 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

Gaussian cdf

Gaussian pdf

Result:

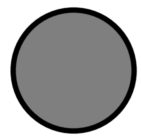
$$VOI = m\Phi\left(m/r\right) + r\phi\left(m/r\right) - \max\{0, m\}$$

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

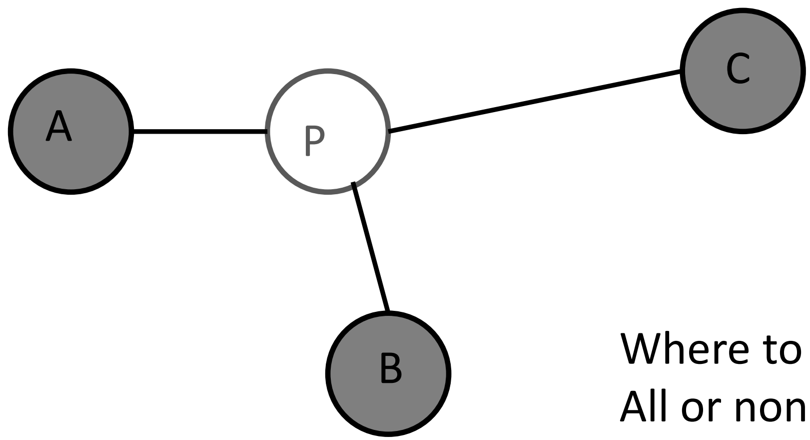
$$= m\int_{-m/r}^{\infty} \phi(z)\,dz + r\int_{-m/r}^{\infty} z\phi(z)\,dz = m\left(1-\Phi\left(-m/r\right)\right) + r\phi\left(-m/r\right)$$

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

What if several projects or treasures?



What if several projects or treasures?



Where to dig?
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, drilling sites, development, control rates, excavation opportunities, harvesting, possibilities, etc.

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

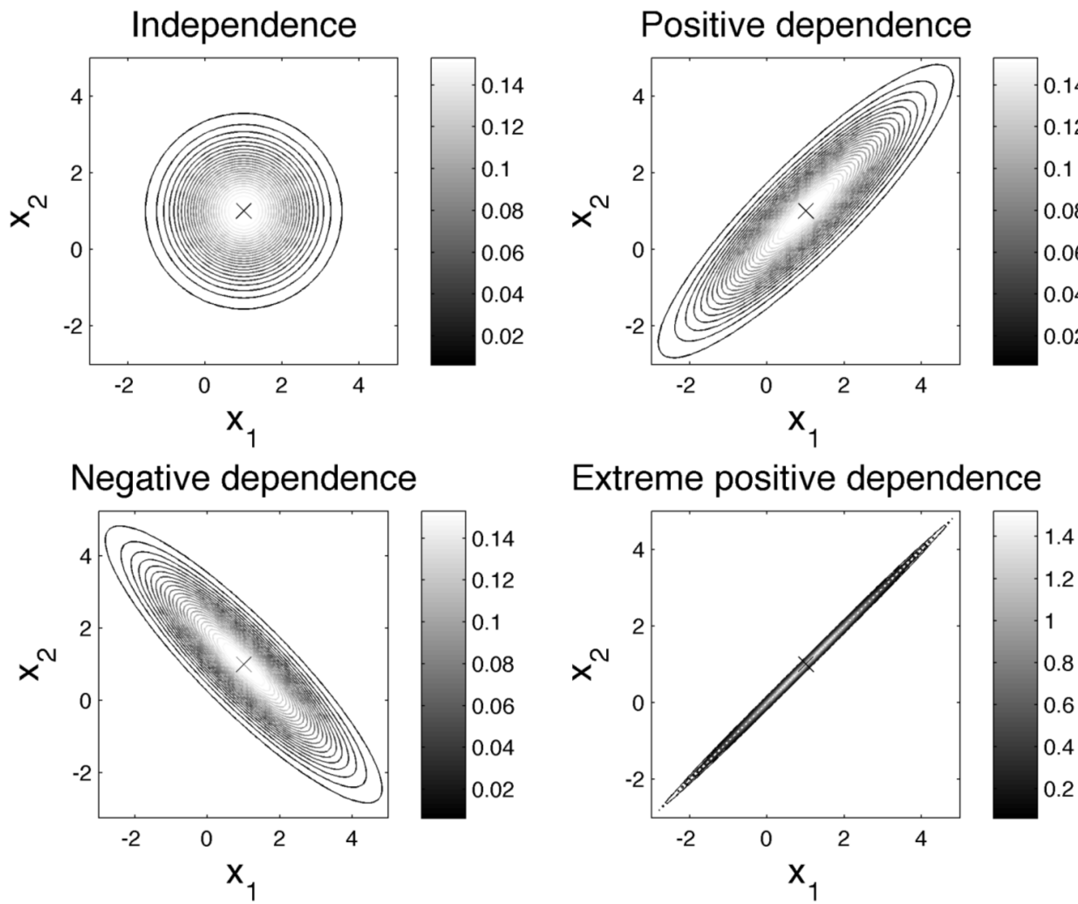
- **Value function is spatial**, typically involving coupled features, such as flow simulation. It can be defined by «physics» as well as economic attributes.

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

Two-project example

Two correlated projects.

Decision maker considers investing in project(s). There are uncertain profits.



Gaussian projects example

- **Alternatives**
  - Do not invest in project 1 (a1=0) - Invest in project 1 (a1=1)
  - Do not invest in project 2 (a2=0) - Invest in project 1 (a2=1)
  - Decision maker is free to select both, if profitable: Four sets of alternatives.
- **Uncertainty** (random variable)
  - Profits are *bivariate Gaussian*.
    Assume mean 0, variance 1 and fixed correlation.

  Value is sum of profits, if positive.

- **Information gathering**
  - Report can be written about one project (assume perfect).
  - Report can be written about both projects (assume imperfect).

Gaussian projects example

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

Prior model for profits: $p(\boldsymbol{x}) = N(\boldsymbol{0}, \boldsymbol{\Sigma}),\quad \boldsymbol{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$

$$PV = \sum_{i=1}^{2} \max\left\{0, E(x_i)\right\} = 0 + 0 = 0$$

$$PoV(\boldsymbol{y}) = \sum_{i=1}^{2} \int \max\left\{0, E(x_i \mid \boldsymbol{y})\right\} p(\boldsymbol{y})\, d\boldsymbol{y}$$

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

Gaussian projects example

$$PV = \sum_{i=1}^{2} \max\left\{0, E\left(x_i\right)\right\} = 0 + 0 = 0$$

$$PoV\left(\boldsymbol{y}\right) = \sum_{i=1}^{2} \int \max\left\{0, E\left(x_i \mid \boldsymbol{y}\right)\right\} p\left(\boldsymbol{y}\right) d\boldsymbol{y}$$

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

Must solve the integral expression!

Need conditonal expectation!

Need marginal for data!

Perfect information about 1 project

$$y = x_1$$
$$p(x_1) = N(0,1)$$
$$E(x_1) = x_1$$
$$E(x_2 \mid x_1) = \rho x_1$$

$$Var(x_1 \mid x_1) = 0, Var(x_2 \mid x_1) = 1 - \rho^2$$
$$PoV(x_1) = \int_0^\infty x_1 p(x_1)\,dx_1 + \int_0^\infty |\rho| x_1 p(x_1)\,dx_1$$
$$= \left(1 + |\rho|\right) \Big/ \sqrt{2\pi}$$

Get information about second project because of correlation!

Imperfect information, both projects

$$\boldsymbol{y} = \boldsymbol{x} + N\left(\boldsymbol{0}, \tau^2 \boldsymbol{I}\right)$$

$$p\left(\boldsymbol{y}\right) = N\left(\boldsymbol{0}, \tau^2 \boldsymbol{I} + \boldsymbol{\Sigma}\right) = N\left(\boldsymbol{0}, \boldsymbol{C}\right)$$

$$E\left(\boldsymbol{x} \mid \boldsymbol{y}\right) = \boldsymbol{\Sigma}\boldsymbol{C}^{-1}\boldsymbol{y}$$

Reduction in variances large, VOI is large.

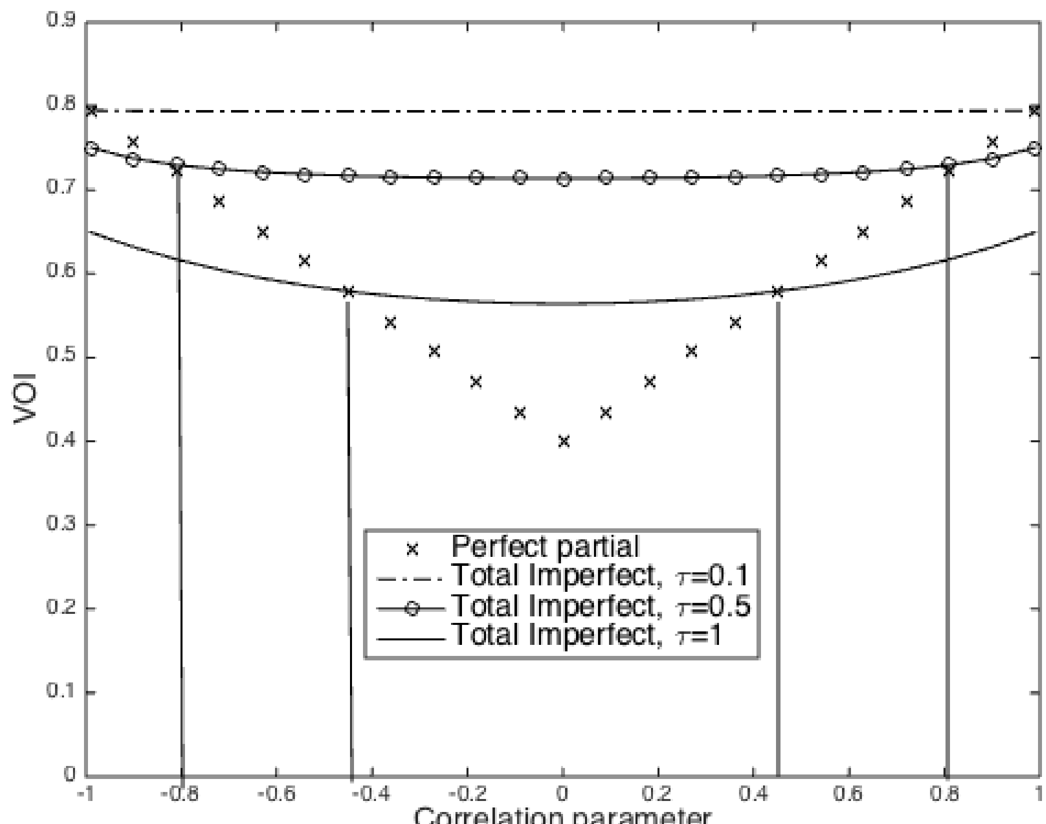
$$Var\left(\boldsymbol{x} \mid \boldsymbol{y}\right) = \boldsymbol{\Sigma} - \boldsymbol{S}, \qquad \boldsymbol{S} = \boldsymbol{\Sigma}\boldsymbol{C}^{-1}\boldsymbol{\Sigma}$$

$$PoV\left(\boldsymbol{y}\right) = \sum_{i=1}^{2} \int \max\left\{0, E\left(x_i \mid \boldsymbol{y}\right)\right\} p\left(\boldsymbol{y}\right) d\boldsymbol{y} = \left(\sqrt{S_{1,1}} + \sqrt{S_{2,2}}\right) \Big/ \sqrt{2\pi}$$

Gaussian projects results

$$PoV(\boldsymbol{y}) = \left(\sqrt{S_{1,1}} + \sqrt{S_{2,2}}\right) \Big/ \sqrt{2\pi}, \qquad \boldsymbol{S} = \boldsymbol{\Sigma C}^{-1}\boldsymbol{\Sigma}$$

$$PoV(x_1) = \left(1 + |\rho|\right) \Big/ \sqrt{2\pi}$$



EXERCISE

Gaussian univariate and bivariate distribution.

- Univariate Gaussian model.
- Bivariate Gaussian model.
- Study sensitivity to parameters.