Session 6:
Decision Theory and Decision Analysis

Marek J. Druzdzel

University of Pittsburgh

School of Information Sciences and Intelligent Systems Program

marek@sis.pitt.edu
http://www.pitt.edu/~druzdzel

Summer School on *"Modeling and Decision Making Using Bayesian Statistics"*
Aalto University, Department of Applied Mechanics, Marine Technology in Espoo, Finland
4th – 8th June, 2012

Course schedule

Day 1 Monday

Session 1: Introduction to Bayesian inference
Session 2: Bayesian networks
Session 3: Building Bayesian networks
Session 4: Hands-on exercises (Bayesian networks)

Day 2 Tuesday

Session 5: Learning Bayesian networks and causal discovery
Session 6: Decision theory and decision analysis
Session 7: Hands-on exercises (learning)
Session 8: Hands-on exercises (decision modeling)

Session overview

- Rationality, rational behavior, elements of behavioral decision theory
- Decision theory, and decision analysis
- Multi-attribute utility models
- Representation and solving of decision problems: Decision trees and influence diagrams
- Sensitivity analysis
- Value of information

What I want you to know after this session?

- Understand the foundations of decision theory and decision analysis
- Understand the concept of utility and multi-attribute utility
- Understand the extension of Bayesian networks to influence diagrams
- Know how to create influence diagrams, perform (true) sensitivity analysis and value of information computation

The Normative Foundations of Decision Theory and Decision Analysis?

What is a good decision?

As many good questions, this question does not have a crisp-cut answer ☺.

Can you judge decisions by their outcomes?

The story of Bill and Bob

What is a good decision then?

One possible answer:

- One that results from a good decision making process
- Improving decisions means mostly improving the decision-making process.

Elements of decisions

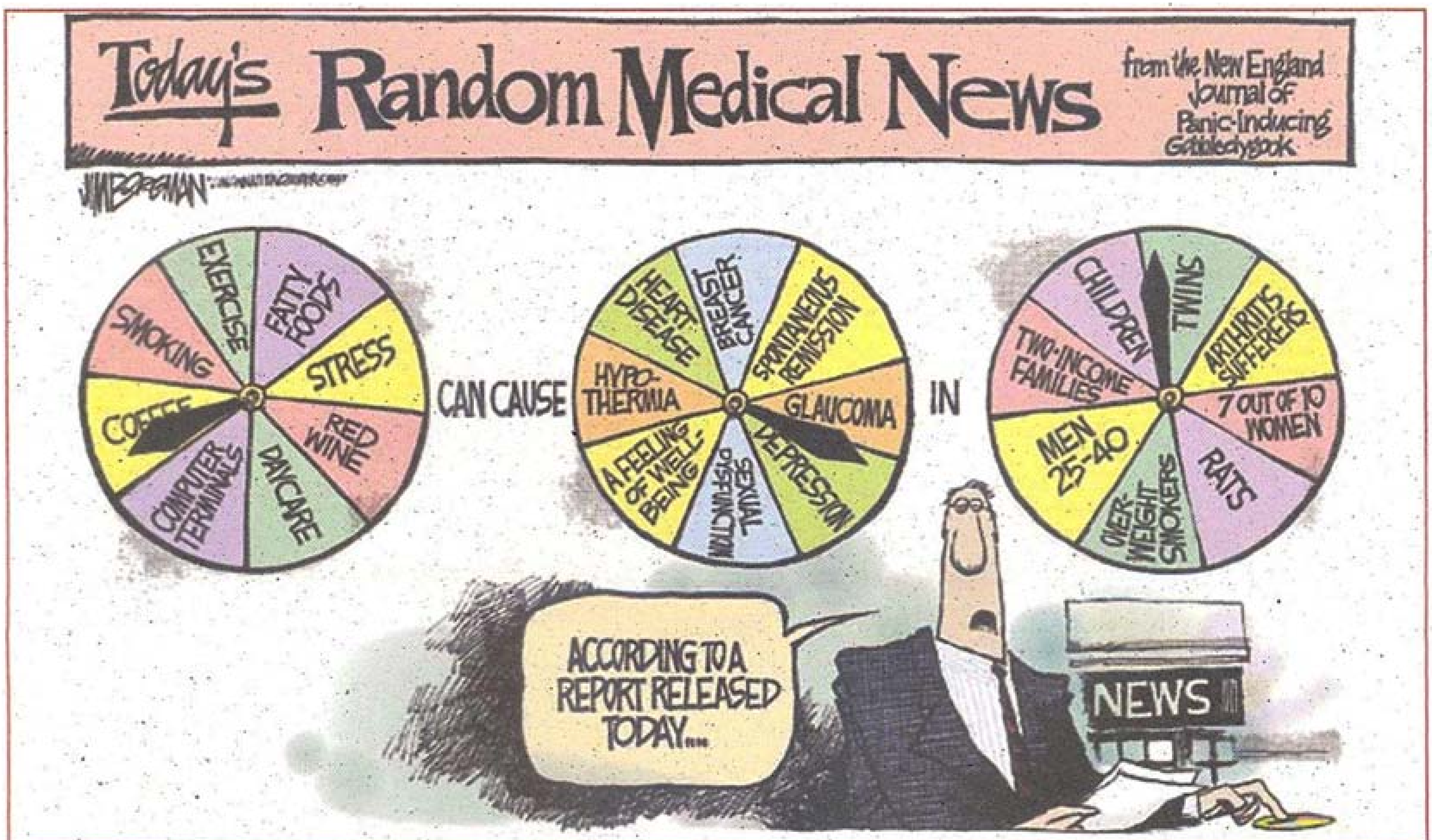
Decisions are made everywhere, including science. What are their elements?

- Preferences (a.k.a. objectives)
- Actions (a.k.a. decision options)
- Uncertainty (nuisance but, unfortunately a fact of life ☹)

Other relevant concepts:

- Context of a decision (situation)
- Consequences (outcomes)
- Dynamic character of decision problems (often leads to sequential decisions)

Why are decisions hard?



©4/27/97 Jim Borgman, Cincinnati Enquirer.
Reprinted with special permission of King Features Syndicate.

- **Uncertainty**

Why are decisions hard?

- Complexity
- Conflicting objectives
- Many decision alternatives

Why are decisions hard?

- Multiple decision makers
- Our cognitive limitations

Normative vs. descriptive decision support

- Traffic laws vs. actual behavior of drivers
- Bible vs. actual behavior of people

Decision theory and decision analysis

Decision theory:

A mathematical theory of how decisions should be made

(based on the idea that uncertainty and preferences should combine like mathematical expectation)

Decision analysis:

The art and craft of applying decision theory in practice

Foundations of decision analysis

The foundation of decision analysis (assumption but confirmed by observations):

Humans can provide reliably the structure of a problem and reliable numbers (judgments) but are weak in combining these

The goal of decision analysis

Insight not numbers!

- Decision analysis provides structure and guidance to thinking systematically about hard decisions
- A DA exercise will be successful if the decision maker has learned something about the problem
- Sometimes it offers justification of previously made choices, but even then it is useful by offering insight

Utility

“Immeasurables”

Some things are difficult to express in numerical terms.

Imagine that you are a juror. How much is it worth to condemn an innocent man or to release a guilty one?

How do you judge money vs. health or happiness?

The need for utility

Even if you can express “immeasurables” in numbers, there are problems with expected value, found quite a while ago (even though probability is quite young).

Bernoulli (17th century) pointed out these problems and the need to have some measure of preferences.

Then there was long nothing, just a qualitative, ordinal notion (note the gymnastics around qualitative notion of utility in economics) and finally a quantitative, cardinal utility in 1940s due to von Neuman & Morgenstern.

Problems with expected value

Who would call a pauper foolish for selling a lottery ticket paying 20K ducats tomorrow with p=0.5 for 9K today?

Problems with expected value

(a.k.a., St. Petersburg's paradox)

Imagine a game that involves flipping a coin infinitely many times and that pays progressively more for reaching each step.

If you get just one heads (p=0.5), you get \$2, if you get two heads in a row (p=0.25), you get \$4, if you get three heads in a row (p=0.125), you get \$8, etc. The expected value of this game is:

$$EV = \sum_{i=1}^{\infty} 2^i \left(\tfrac{1}{2}\right)^i = \sum_{i=1}^{\infty} 1 = \infty$$

What would you pay to participate in this game? Will it be wise to pay a lot for it? Note that the expected value of this game is infinity!

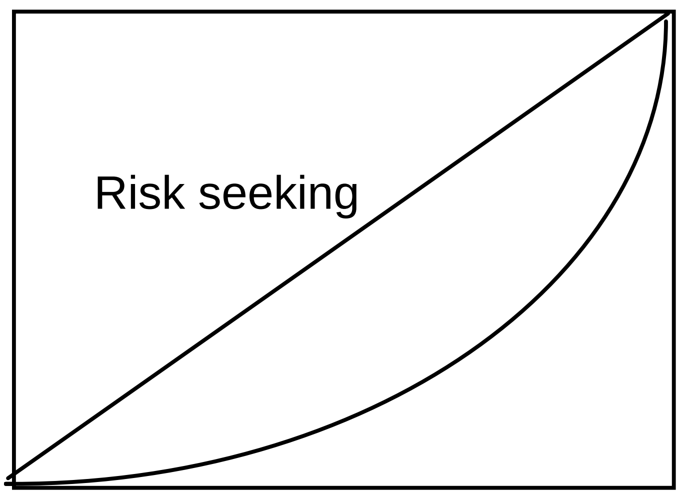
Bernoulli's solution

The solution proposed by Bernoulli is that, what he called, "moral worth" of a quantity is different from that quantity.

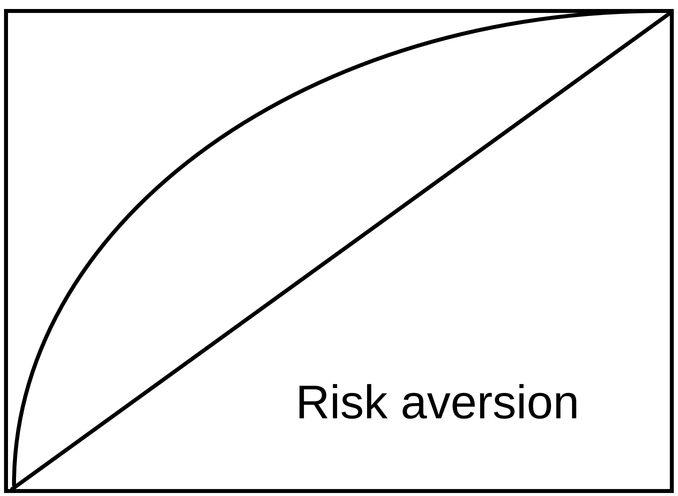
He introduced the law of diminishing marginal utility and proposed the logarithm function as one that satisfies this law. (Just take the logarithm of the value to get the utility.)

Risk attitudes

- Three theoretical attitudes: risk neutrality, risk seeking, and risk aversion.
- Easy to understand in terms of the second derivative of the utility function:
  - If each additional dollar is worth more to you than the one before, you are out to win big and you are willing to take risks
  - If the value of each additional dollar is worth less than the last dollar, then you are risk averse.



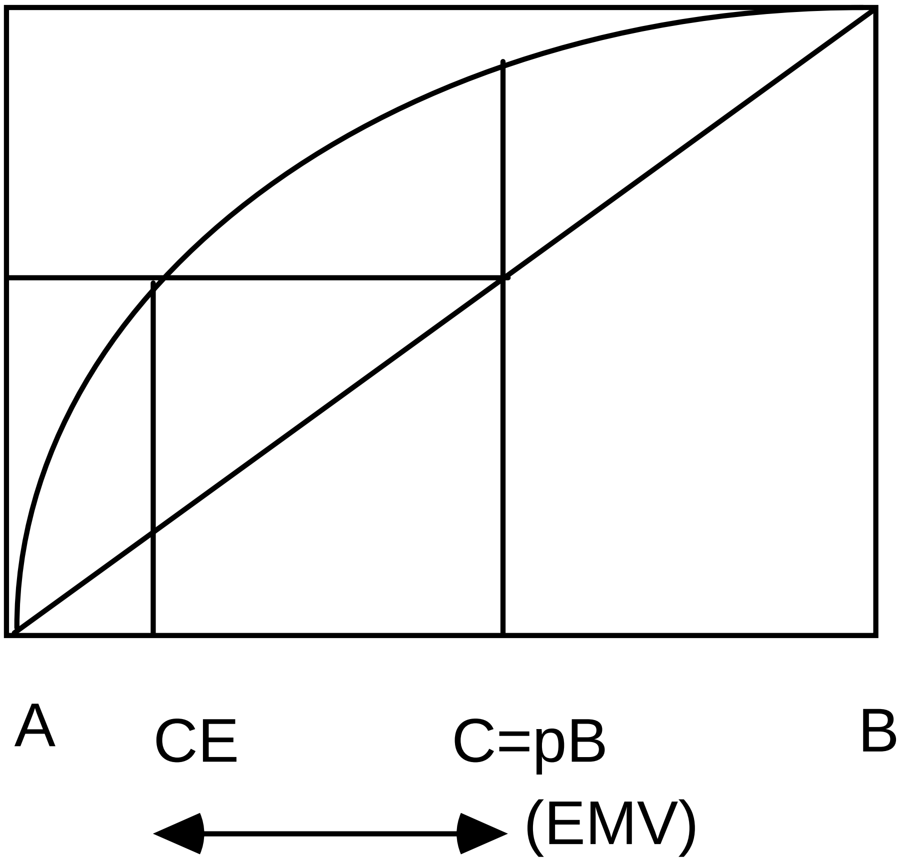
e.g., lottery players



e.g., people buying flood or health insurance

Certainty equivalent

Certainty equivalent of a gamble: How much would you pay for an opportunity to participate in this gamble?

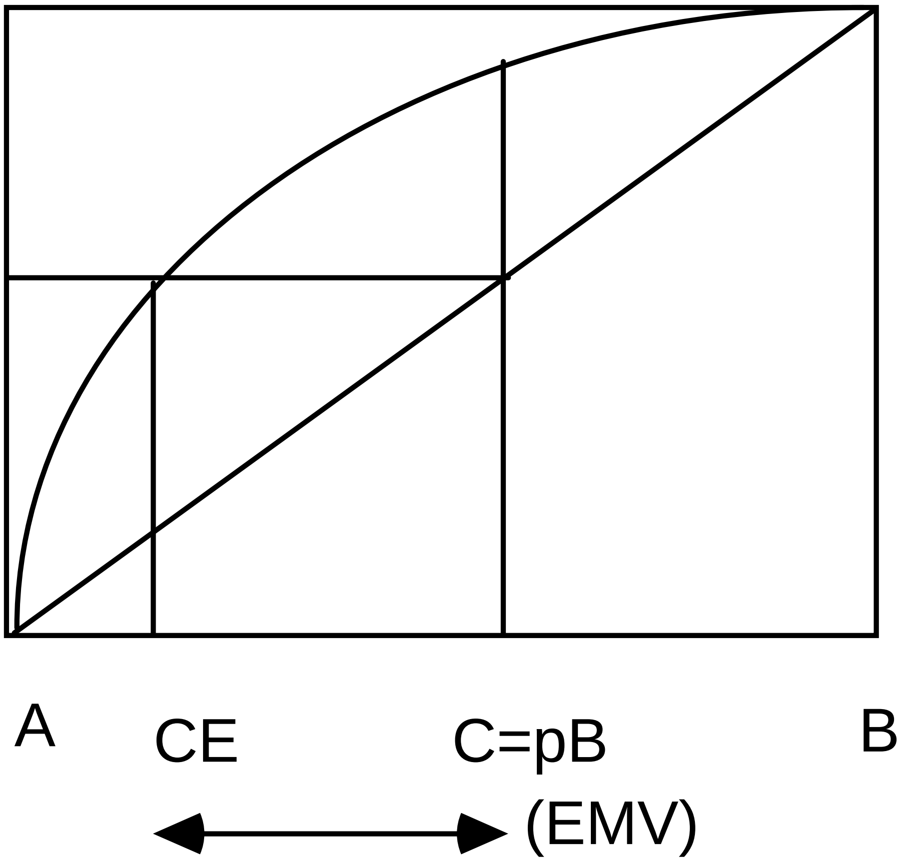


Risk premium can be positive or negative (How will the picture look for somebody who is risk prone ☺?).

Note that each of CE, EMV and RP are not in terms of utility but rather in terms of the quantity that we are measuring.

Risk Premium = EMV – CE

Certainty equivalent



Restatement: If the gamble is worth to you less than the expected value, then CE is to the left of EV.

Restatement: If the utility of a value is higher than the utility of that value when it is only expected, then we are dealing with risk aversion.

Utility

Axiomatization proposed by von Neumann and Morgenstern in 1940s

A peculiar measure with no scale and no zero point

If U(x) is a utility function, then U'(x)=aU(x)+b is also a utility function, i.e., utility is determined up to a linear transformation

Any other measures that behave like this ☺?

Measurement of utility

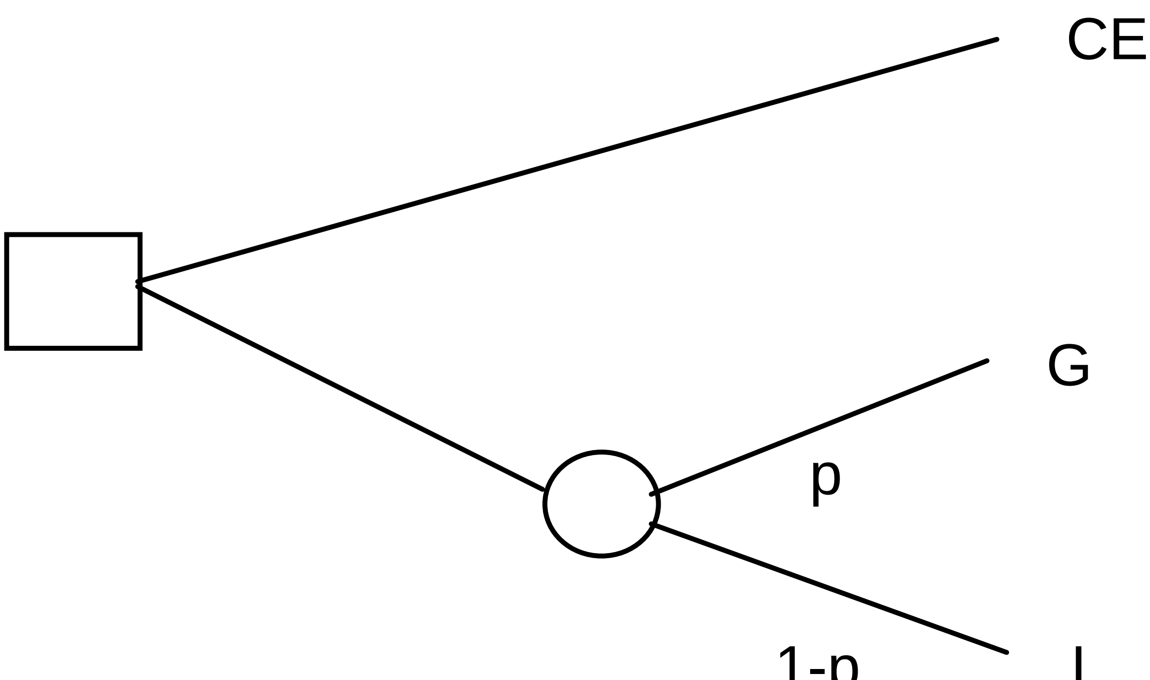
We have four variables: p, CE, G, L.

Two are for free and determined by the axioms (these are the lower and the upper bounds of the utility range).

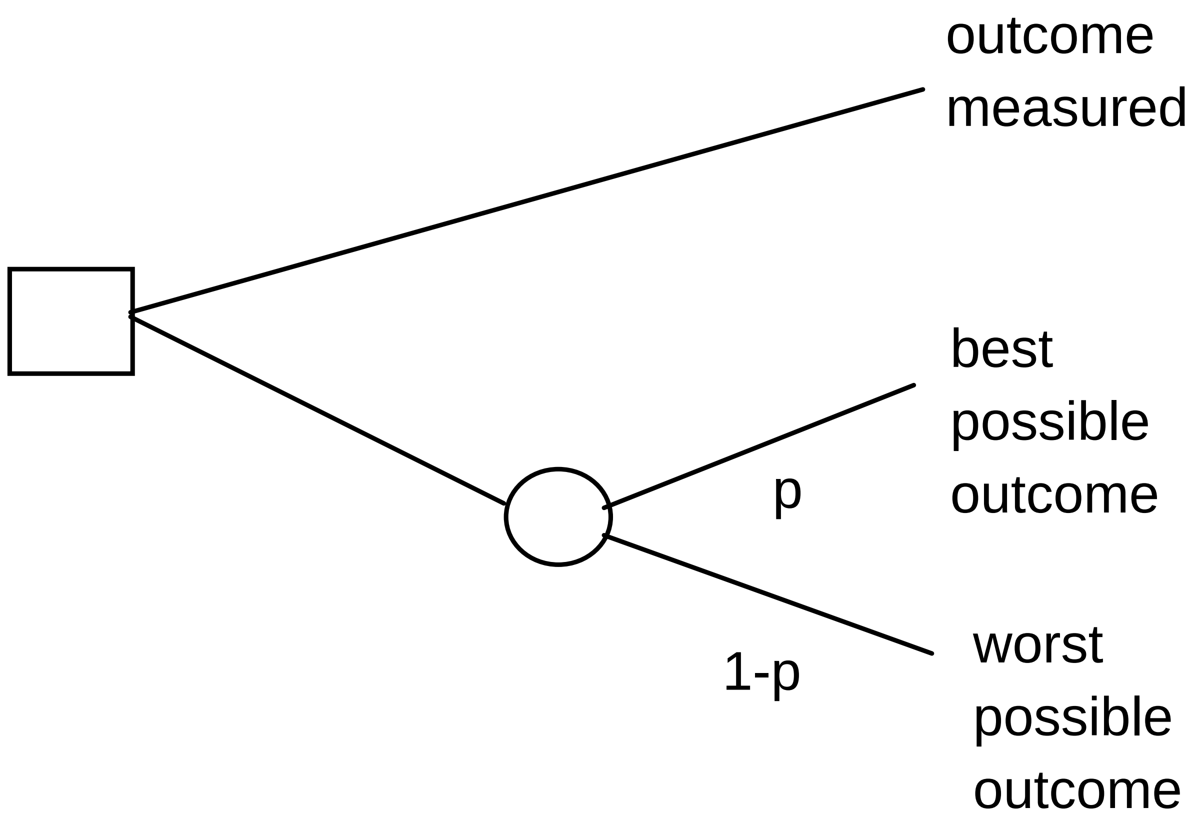
We need to fix (preset) the third and then obtain the fourth.

CE method: fix G, L, and p, assess CE

PE method: fix G, L, and CE, assess p



Measurement of utility: Probability equivalent

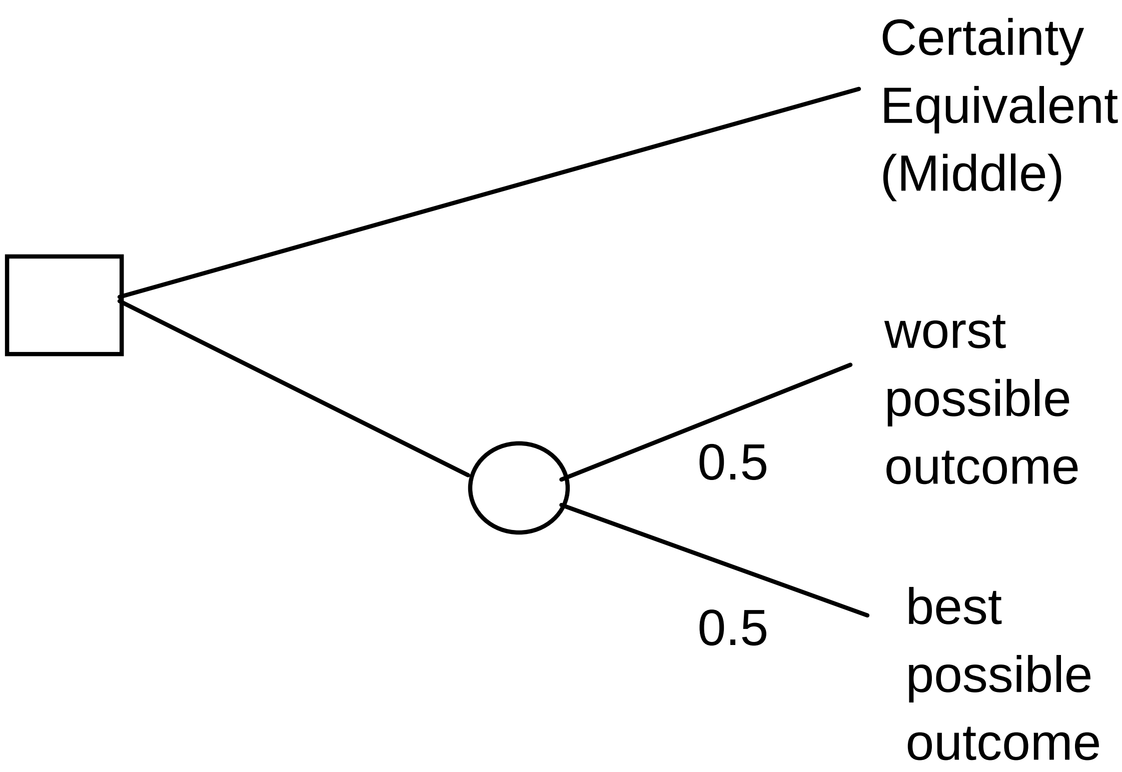


Manipulate p until the decision maker is indifferent between the two choices. Then,

U(Measured) = p U(Best) + (1-p) U(Worst)

U(Measured) = p 100 + (1-p) 0 = p 100

Measurement of utility: Certainty Equivalent



Manipulate CE until the decision maker is indifferent between the two choices. Then,

U(CE) = 0.5 U(Best) + 0.5 U(Worst)

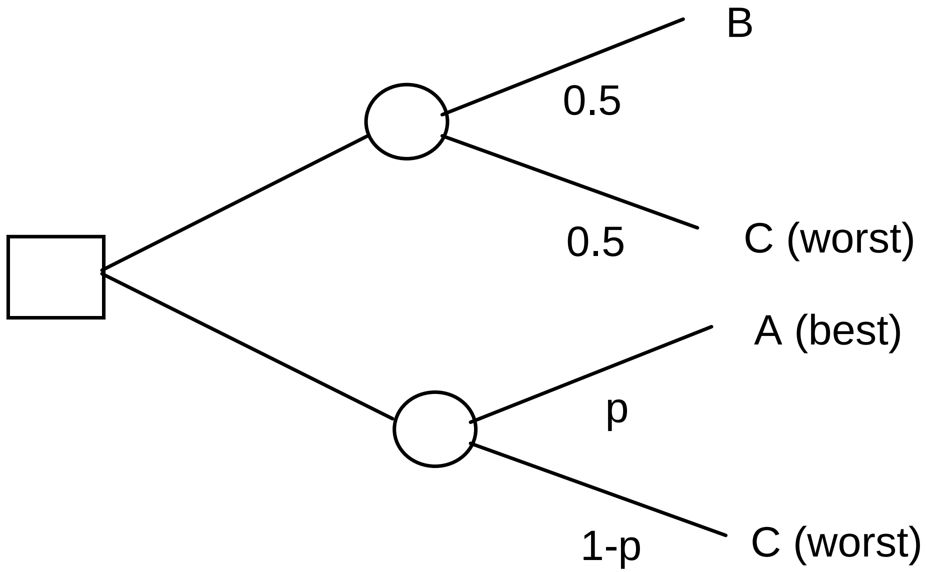
U(CE) = 0.5 100 + 0.5 0 = 50

Measurement of utility: Comparison of PE and CE

CE leads to more risk-averse responses in gains and risk seeking in losses.

In PE, p=0.5 is the best, as people exhibit probability distortions at more extreme probabilities ("certainty effect").

One possible answer to this problem is to use the following lotteries:



This is known as The McCord-De Neufville utility assessment procedure.

Risk tolerance-based utility functions

Exponential utility function: $U(x) = 1 - e^{-x/R}$

This is a "poor-man's" utility function and one can argue whether it models well a decision maker's preferences.

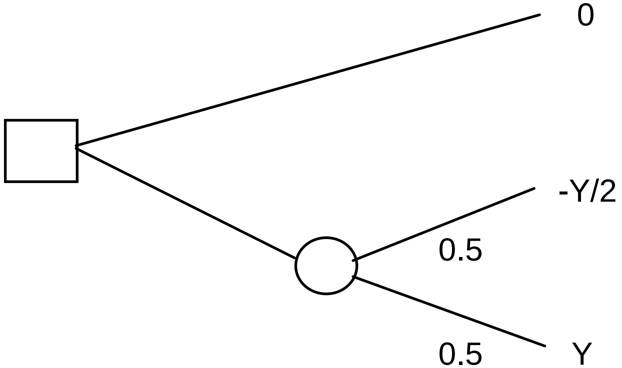
Useful as a first-cut approximation in cases when we want to model risk aversion.

A quick sensitivity analysis can determine a critical risk tolerance, and the decision maker can be asked, via a simple assessment question whether his/her risk tolerance exceeds the critical value.

If the choice is clear, then there is no need for further preference modeling.

If the choice is not clear, it may be a good idea to assess a utility function more carefully.

Risk Tolerance R is the maximal value of Y for which the DM is indifferent between the choices in the following gamble:

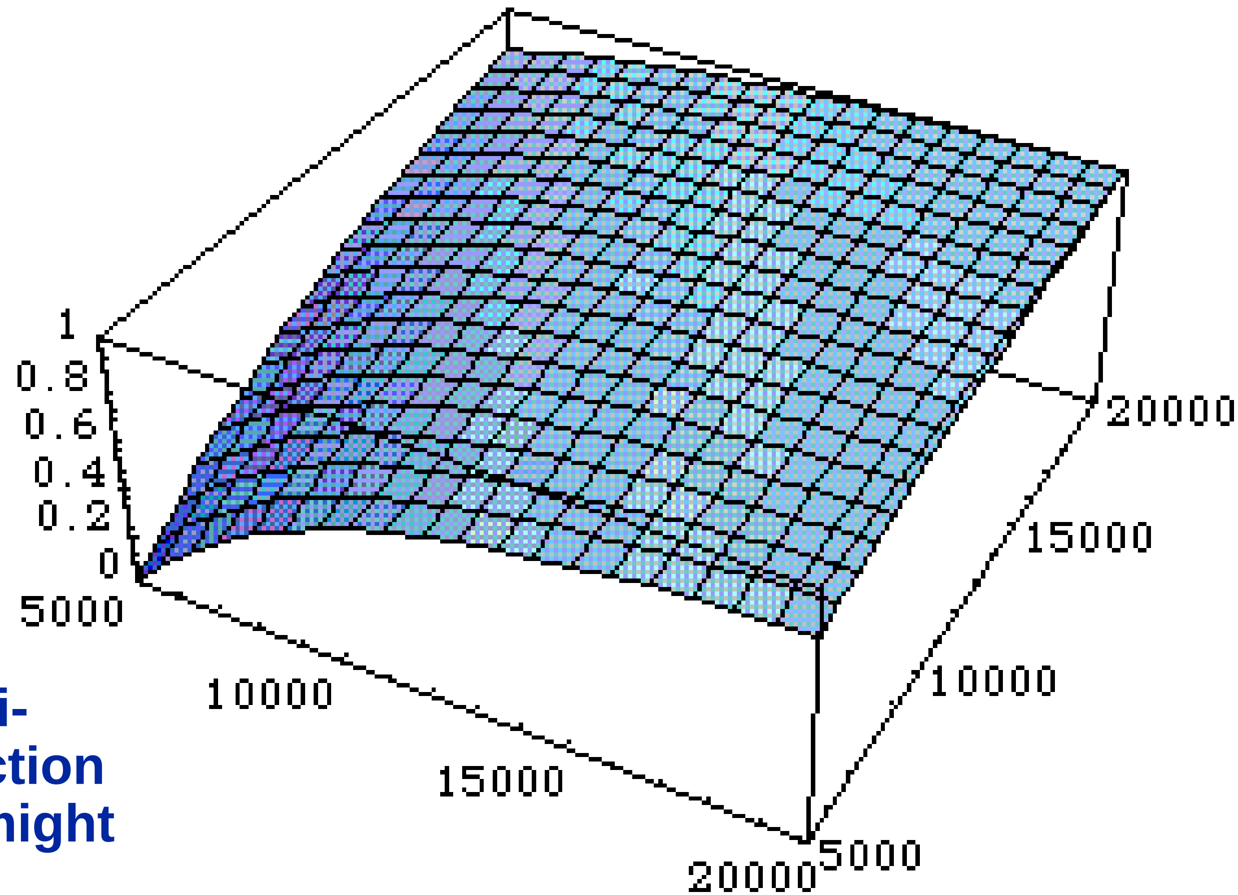


Risk tolerance-based utility functions

Multi-Attribute Utility

Multi-attribute utility

When there are multiple attributes of a decision (quite typical ☺), we are facing a hard problem, a function of multiple arguments

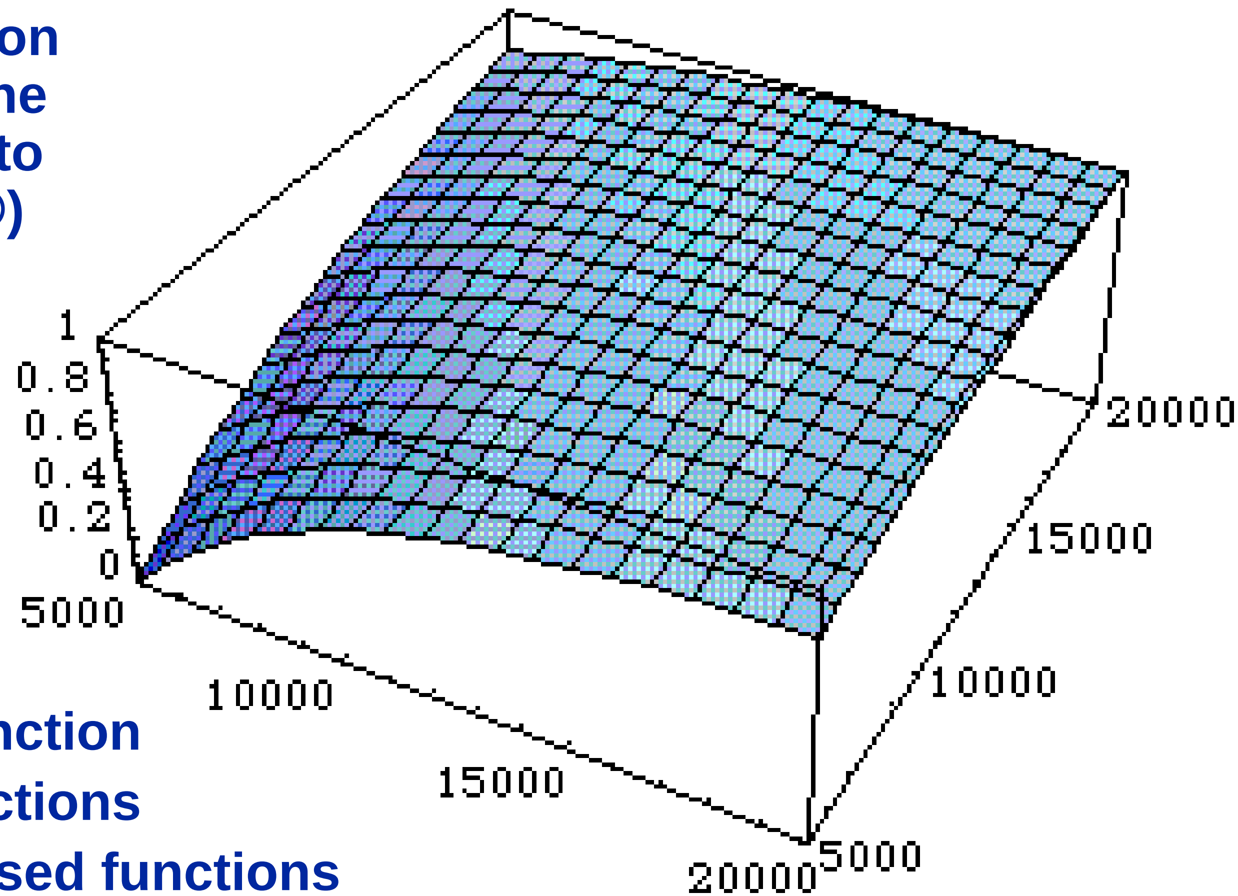


Here is what a multi-attribute utility function of two arguments might look like.

Multi-attribute utility

Elicitation of a MAU function is hard (number of points exponential in the number of attributes)

An obvious solution is standardizing the shapes (similarly to canonical gates ☺)



Several solutions:

- **Additive linear function**
- **Multiplicative functions**
- **Risk tolerance-based functions**

Multi-attribute utility: Simplification of the problem

Simplifications of the problem starts with a series of attribute independence tests:

- **preferential independence**
- **utility independence**
- **additive independence**

Preferential independence

An attribute Y is said to be preferentially independent of X if preferences for specific outcomes of Y do not depend on the level of attribute X. In other words, the value of X does not influence our ordinal preferences for Y.

This condition is pretty intuitive and it holds most of the time.

Examples of violations?

1. **The amount of homework and the course topic.**
2. **Car type and location.**

Utility independence

An attribute Y is considered utility independent of attribute X if preferences for uncertain choices involving different levels of Y are independent of the value of X. In other words, the value of X does not influence the certainty equivalent of a lottery involving Y.

Mutual utility independence: When the relation holds both ways.

Example when this is violated (from Keeney and Raiffa): Serious crime rates in two police precincts. The region's police chief does not want to appear as though he neglects one of the two precincts. An easy fix in that case is adding bonus to some values or transforming the function.

Implication of utility independence

When mutual utility independence holds, we can write a two-attribute utility function as follows:

$$U(x,y) = w_x\, U_x(x) + w_y\, U_y(y) + (1 - w_x - w_y)\, U_x(x)\, U_y(y)$$

$U_x(x)$ and $U_y(y)$ are utility functions scaled to the interval [0,1], $w_x=U(x_1,y_0)$, $w_y=U(x_0,y_1)$.

Multiplicative form of multi-attribute utility

This is known as the multiplicative form of a MAU function. It is a special functional form that gives a curvature in the utility function of multiple attributes and is capable of modeling such non-linearities as complements and substitutes.

$$U(x,y) = w_x\, U_x(x) + w_y\, U_y(y) + (1 - w_x - w_y)\, U_x(x)\, U_y(y)$$

The product term is what allows for modeling the interaction between the two attributes.

Complements and substitutes

$$U(x,y) = w_x U_x(x) + w_y U_y(y) + (1 - w_x - w_y) U_x(x) U_y(y)$$

The coefficient $(1 - w_x - w_y)$ can be interpreted quite nicely.

If positive, then higher values of both attributes at the same time will drive up the value of the utility function even higher (the attributes complement each other, e.g., two battles on one front, you need to win both, defeat on one is almost just as bad as defeat on both).

If negative, we are quite happy with having one or the other and don't necessarily need to have both (they substitute each other, e.g., two branches of a company, two investments).

Utility independence

How do we demonstrate that this functional form implies mutual utility independence?

Take one value of y: The function will transform to the utility U_x, although it will be its linear transformation.

For another value of y, it will be another linear transformation.

The utility function for x will be exactly the same, because it is determined up to a linear transformation anyway.

How to go the other way, i.e., demonstrate that you need this functional form to have mutual utility independence?

Left as a homework exercise ☺.

Additive independence

When $w_x+w_y=1$, the multiplicative function simplifies to

$$U(x,y) = w_x\, U_x(x) + w_y\, U_y(y)$$

This is precisely when additive independence holds.

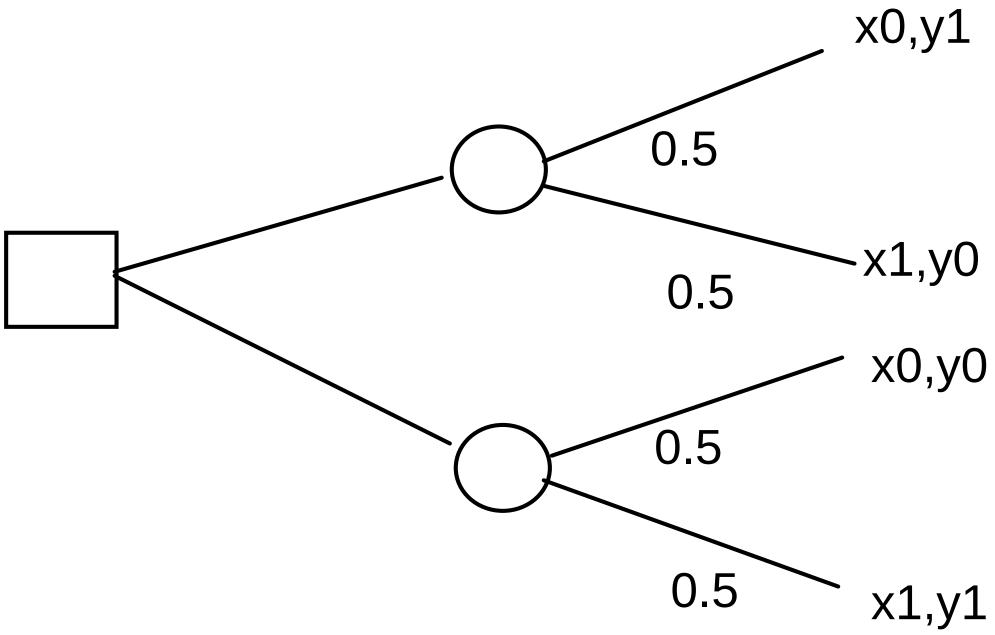
In general

- $U(x_1, x_2, \ldots, x_m) = k_1\, U(x_1) + k_2\, U(x_2) + \ldots + k_m\, U(x_m)$
- Condition on weights: $k_1 + k_2 + \ldots + k_m = 1$)

Additive linear utility function is quite often used and abused (used without checking whether it is a good approximation).

Multi-attribute utility assessment

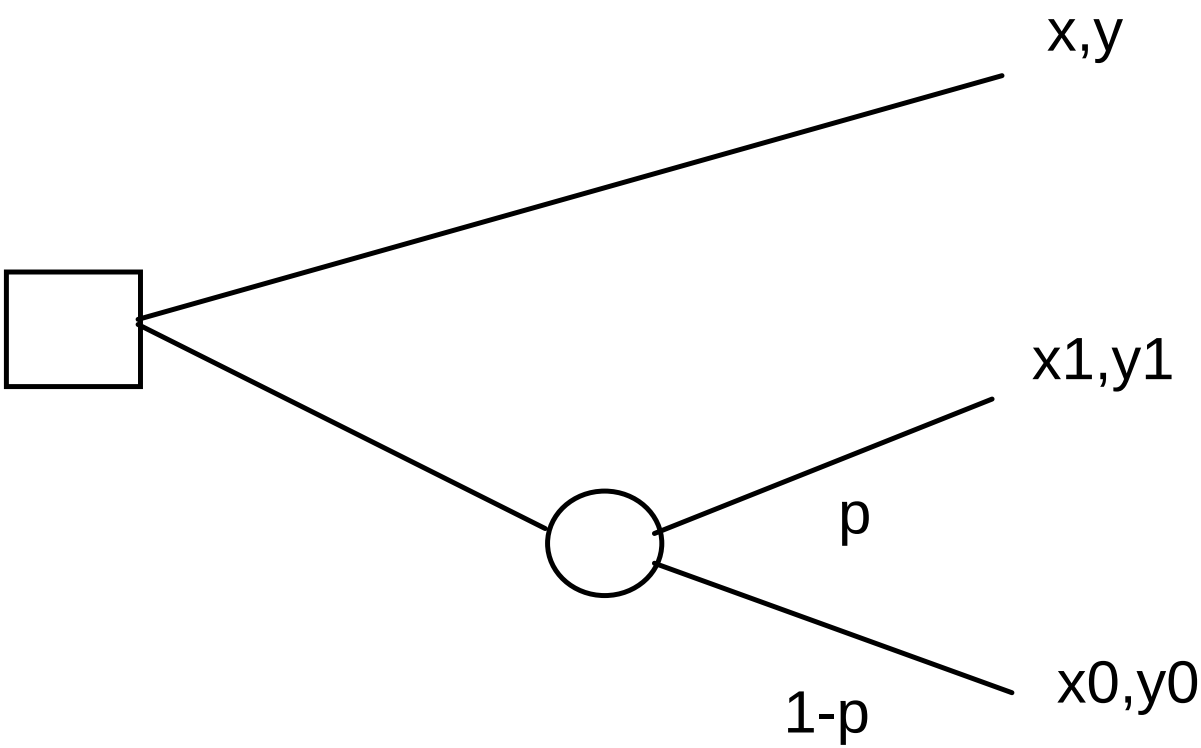
Are you indifferent between the two choices? If so, then they are additively independent, but if you prefer one over the other, then they are not. A good example: service and reliability — most of us prefer when at least one of them is good to the situation when you can be screwed up on both or have both good.



MAU assessment: When everything fails

What is mutual utility independence fails? You can always use **direct assessment**.

Sometimes transformations of the individual utility functions will work (e.g., instead of individual crime rates, take the average and difference between the two crime rates).

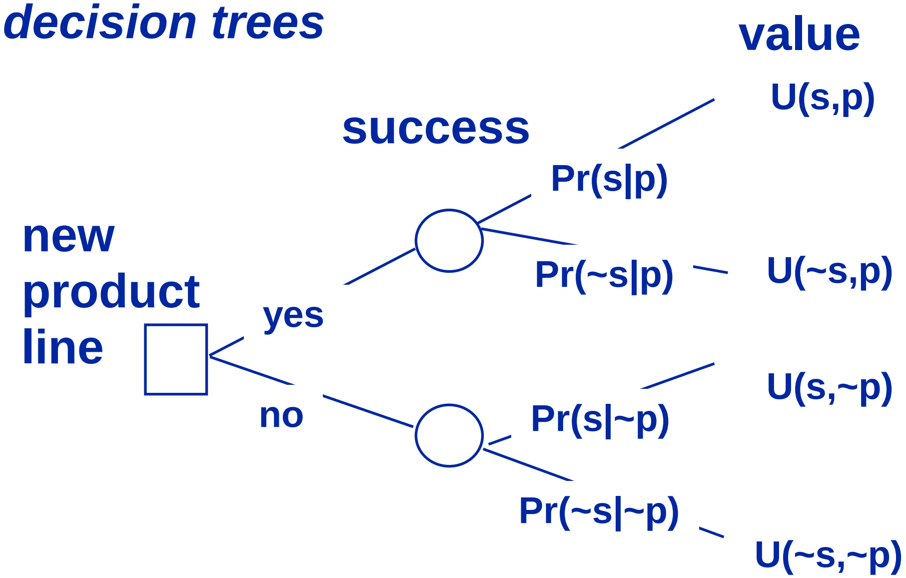


Decision Modeling Tools

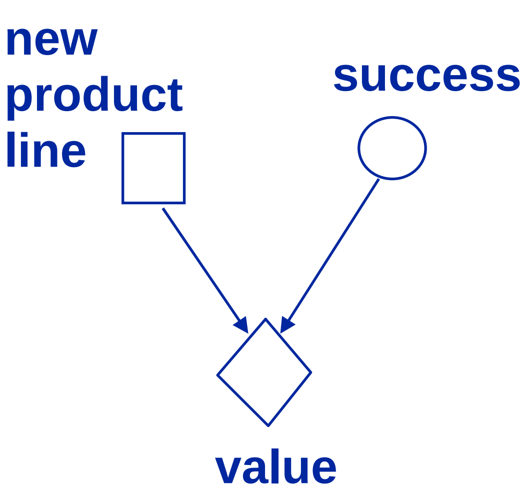
Decision-analytic modeling tools

The essence of a decision-analytic model is a representation of uncertainties (joint probability distribution over the model's variables), decision options, and values (utility function over the outcomes).

decision trees



influence diagrams



Most of the time it comes down to significant simplifications, such as discretization of the model variables to allow this specification.

Sensitivity Analysis

Sensitivity analysis

Every model rests on a variety of assumptions regarding:

- **the decision options available**
- **possible states of nature and the probabilities associated with these states**
- **the values of different outcomes**

- **In many cases we will be uncertain about the validity of some of our assumptions. Or, in social settings, different people may differ in what they find reasonable to assume.**
- **The purpose of sensitivity analysis is to determine which assumptions really do have a substantial impact on the decision.**
- **Once we know which assumptions matter most, we can focus our attention there.**

Sensitivity analysis

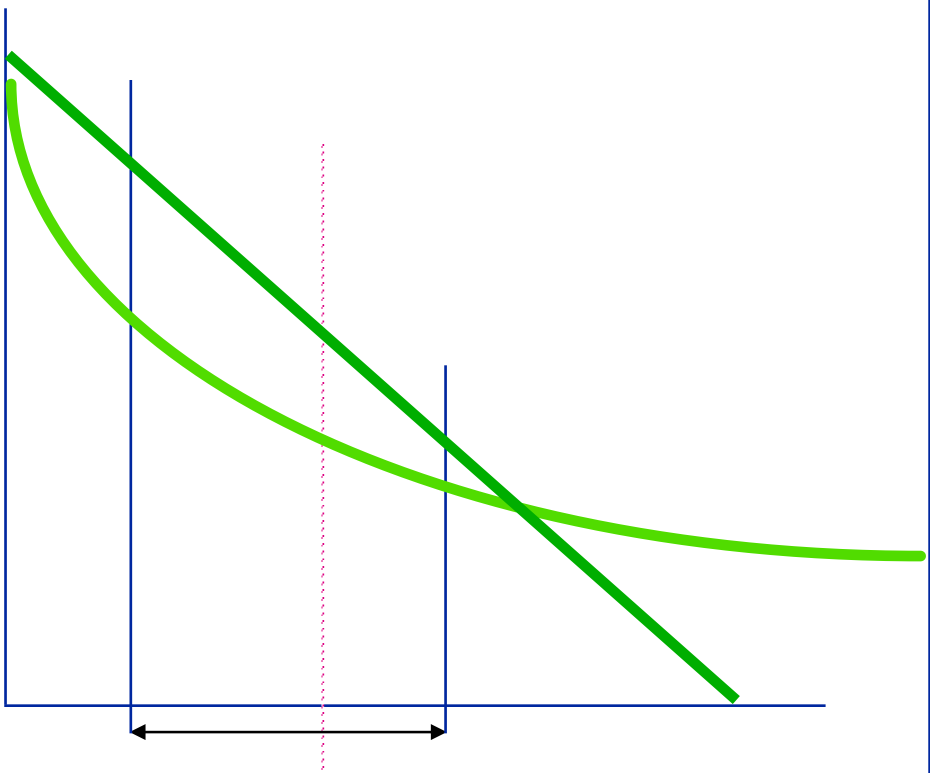
- Sensitivity analysis answers the question: "What matters in this decision?"
- Determining what matters requires incorporating sensitivity analysis throughout the modeling process.
- It is important to keep in mind that the purpose of sensitivity analysis is to refine the decision model, with the ultimate objective of obtaining a requisite model.
- Sensitivity analysis can lead to reconsidering the very nature of the problem ("Are we solving the right problem?")
- There may be synergy effects among various model parameters, so the problem is very complex in general. No optimal procedure exists for performing sensitivity analysis. It is essentially an art with a few basic heuristics, most of which are covered in the textbook.

One-way sensitivity analysis

Indicates the sensitivity of a proposed decision to changes in the value of a single parameter.

Plots the graph of variations in total value with respect to the possible range of values of a parameter.

Very often we have the lower and the upper bound for a value. We can check the values of the outcome variable for the range of possible values of decision variables and uncertain quantities between their upper and lower bounds.



Multi-way sensitivity analysis

- Sometimes we may want to consider sensitivity to several variables at the same time.
- This gives us an idea about how combinations of parameters impact the decision.
- Quite complex in general (exponential in the number of parameters). Two-way sensitivity analysis is quite complex already.
- I have not seen more than two-way done in practice. On the other hand, I have read opinions that going higher rarely gives insight that we have not gained from one and two-way sensitivity analysis.

Sensitivity analysis: Basic question

The basic question in sensitivity analysis is whether a parameter changes our decision

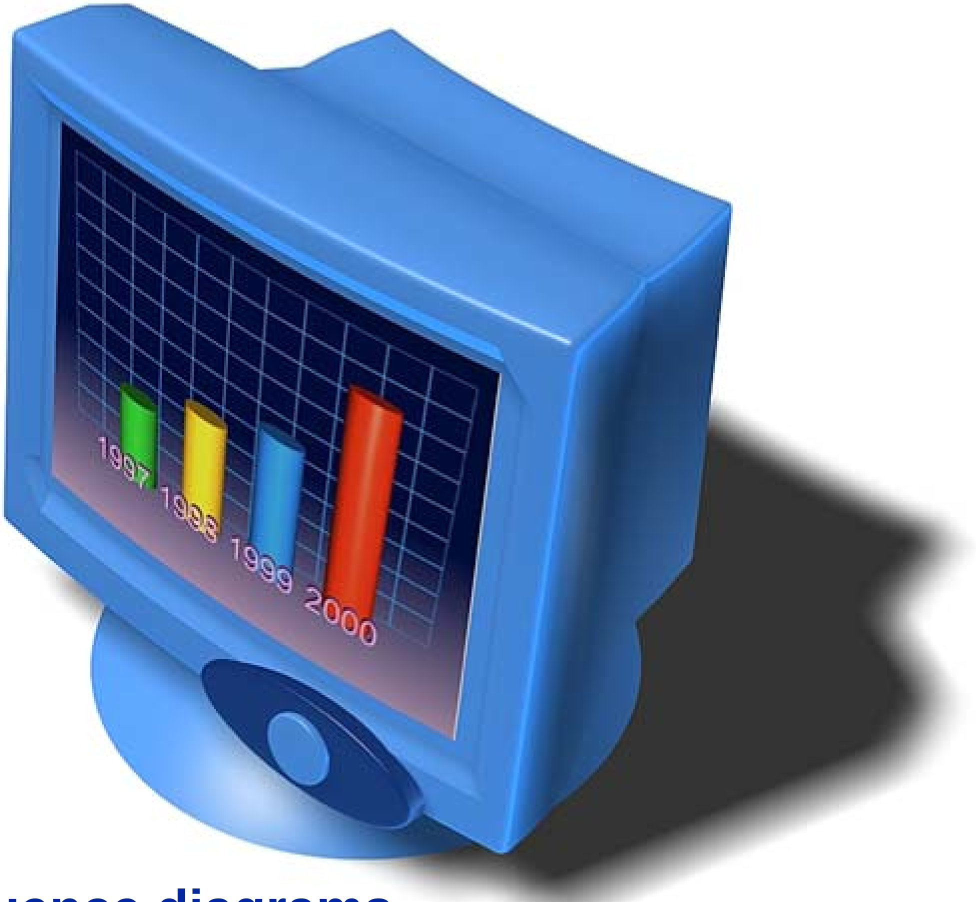
That's why sensitivity analysis in Bayesian networks is standing on a somewehat uncertain ground ☺

Value of Information

Value of information (VOI)

- It is good to know things ☺
- Since knowledge does not typically come for free, one might ask the question how much is a piece of information worth.
- This is typically translated into additional benefit beyond what we have already, i.e., VOI is defined as the difference between the expected utility when we have the additional information and the expected utility without it.
- What does it mean that VOI is zero?

The remainder



- Building and solving influence diagrams
- Performing value of information analysis
- Performing sensitivity analysis