Decision Trees

Marcelo Coca Perraillon

University of Colorado
Anschutz Medical Campus

Cost-Effectiveness Analysis
HSMP 6609
2020

Outline

- A silly but useful decision analysis example
- Decision tree components
- Review of basic probability
- A less silly example: doing CEA/CUA with decision trees
- Steps to construct models
- Limitations of decision models
- Recurrent decisions: Markov models

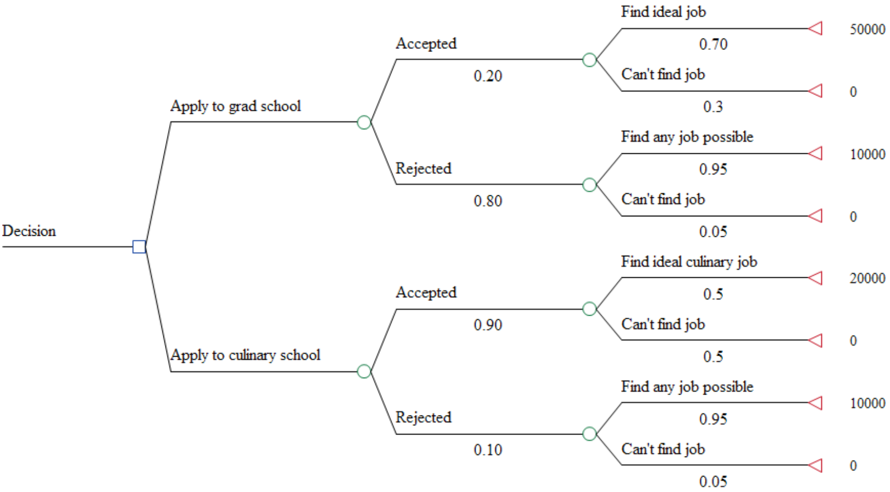
Big Picture

- We are about to learn how to calculate the ICER in a different way
- We have added up the costs over a time horizon and also added up the outcomes over the same time horizon: ICER $= \frac{C_A - C_B}{E_A - E_B}$. Most common in CEA: $\frac{C_A - C_B}{QALY_A - QALY_B} = \frac{\Delta Costs}{\Delta QALYs}$
- We used discounting to bring both to the present
- Now we will do the same but will incorporate **uncertainty**
- You could think of this class as a way of calculating a **expected ICER**
- Today, we will cover **decision trees**. Next week, **Markov models**. The following week, extensions of Markov models and other simulation techniques, including modeling of **infectious diseases**
- These tools have different origins so **language can be confusing**. Lots of new terms

Decision: Go to grad school or cooking school?

- Suppose that you are trying to decide between applying to graduate school (PhD) or cooking school
- There is considerable **uncertainty** in the decision
- We are going to isolate some **key elements** and **make assumptions** to simplify the problem
  1. What is the probability of being accepted into a grad school program or cooking school?
  2. What is the probability of finding an *ideal* job after graduation?
  3. What is the probability of finding *any* job?
  4. What is the yearly salary of the ideal job after graduation? Salary of any job?
- We will start by building a **decision tree** and by assigning some values to uncertain events and outcomes

The decision tree



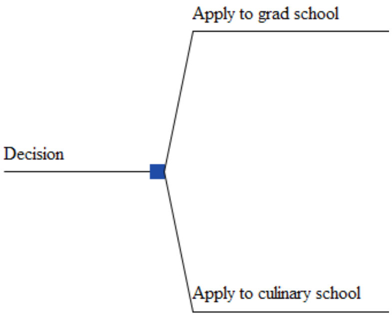
Trees

- A decision tree is a “convenience” tool
- Nothing magical about trees; they are useful for visualizing the problem and thinking about all the alternatives
- Trees are not needed to “solve” a decision problem, but drawing a tree helpful to organize options
- **The most difficult part of a decision problem is to isolate the most important considerations**
- The last point is not trivial: modeling is both science and art. We need to isolate key elements to simplify and ignore others but this is hard to do in practice
- Einstein: “Everything should be made as simple as possible but not simpler”

A digression about software

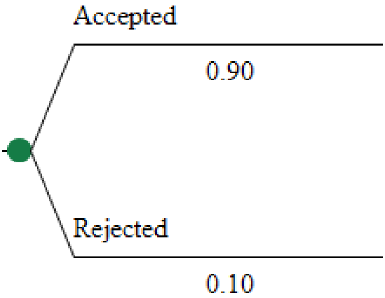
- We will use Excel to work with decision and Markov models (next class)
- A popular book on decision modeling (Briggs et al, 2006) has tons of examples of models using Excel (relevant chapters in Canvas)
- For **learning**, Excel is the best tool. You can see how *everything* is calculated
- **TreeAge** is the most popular alternative (I used TreeAge to draw the above tree)
- But TreeAge it's not easy to learn/use and is a "black box." Not great for learning. I would need several classes just to teach you the basics of TreeAge
- TreeAge does make life easier if you often work with decision/Markov models

Decision node



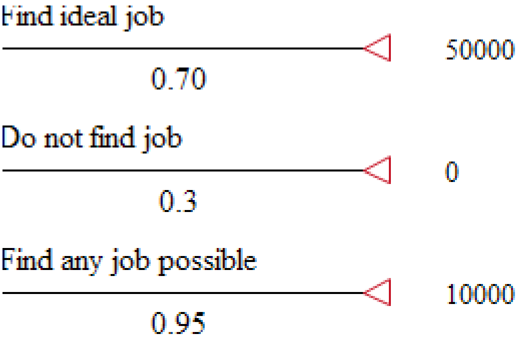
- A **decision node** is where a **decision** takes place; the **branches** are the **choices** (note the square)
- Usually at left of tree but it could be in the middle of the tree depending of the problem
- Our current example has two **choices**, but it could be more than two

Chance node



- A **chance node** marks the place where *chance* determines the outcome (no decision; note green circle)
- Outcomes (possibilities) must be **mutually exclusive** (only one event can happen) and **exhaustive** (probabilities must add up to 1)
- As with the chance node, there could be more than two possible outcomes

Terminal node

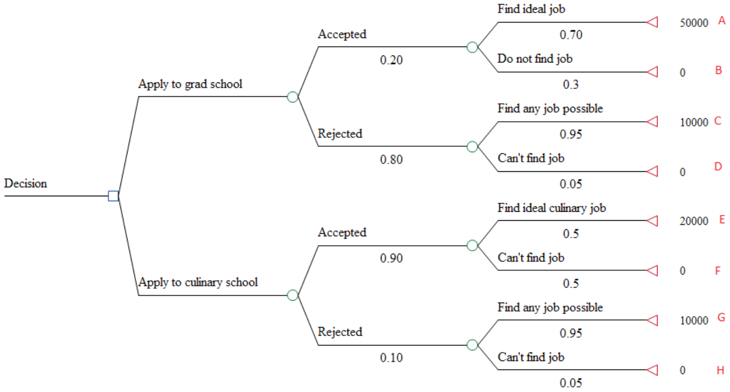


- The **terminal node** is the final outcome associated with each branch
- A final outcome has a **payoff**. In this example, the **payoff** is yearly income but it could be something more complicated and there could be **more than one** payoff (this will become important soon)

Department of Big Ideas

- Only three elements –**decision**, **chance**, and **terminal** nodes– provide all the elements needed in a decision tree (!)
- Yet, decision trees are surprisingly flexible
- Note the data needed:
  1. **probabilities** for chance nodes
  2. **payoffs** for terminal nodes (money so far but more to come)
- A decision tree defines many possibly **pathways**
- Pathways are **sequence of events** that lead to a payoff
- Think of pathways as a sequence of events
- That sequence of events has an end result or consequence: one or more payoffs

Pathways



- This simple example has 8 **pathways** (A to H)
- For example, pathway B: Accepted into grad school, attends, but can't find ideal job

The role of **time**

- Time is **implied**, but I did not explicitly stated the time horizon
- Grad school is about 5 years, cooking school is shorter (2-3 years)
- We could adjust the payoffs to more realistically describe the difference in time
- For example, we could use a 20-year time horizon considering that cooking school would provide two or three extra years of income
- We then bring the salary to the present using the discounting tools we learned
- Of course, the time horizon is important. **Remember**: perspective, time horizon, relevant of costs for the decision...

"Rolling back" the tree

- We have set up the tree, estimated probabilities, and payoffs. Now we can figure out what a **rational person should do**
- To figure out what we **should do**, we calculate the **expected payoff** of each decision
- In other words, **we want to know the expected payoff for each pathway** (that's all, really)
- We will do it in two ways. The first one, "rolling back" the tree (or the **rollback method**)
- For rolling back the tree, we start from the right and move towards the left in the tree (like reading Japanese or Hebrew) to calculate expected values
- First, a review of probability and expected values

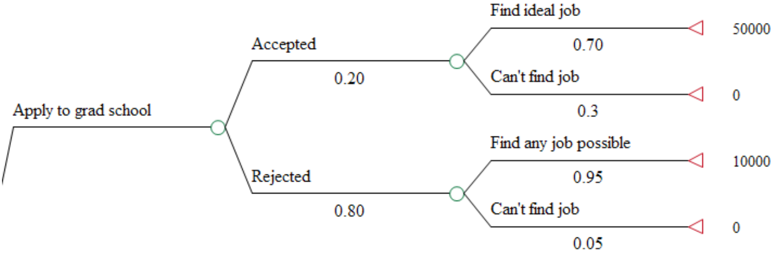
Expected value

- Think about the expected value as a weighted average. The simple average is also weighted
- If you have three numbers, x_1, x_2, x_3, their average is $\frac{x_1+x_2+x_3}{3}$, which can also be written as
  $\frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3$
- Each number is weighted by the same amount $(1/3)$ and the weights add up to 1
- The expected value are the chances outcomes multiplied by their probabilities and summed: $E[X] = \sum_{i=1}^{\infty} x_i p(x_i)$
- The expected value in the context of decision trees are the payoffs weighted by their probabilities
- Another way: the payoff times the probability of obtaining that payoff

Expected value

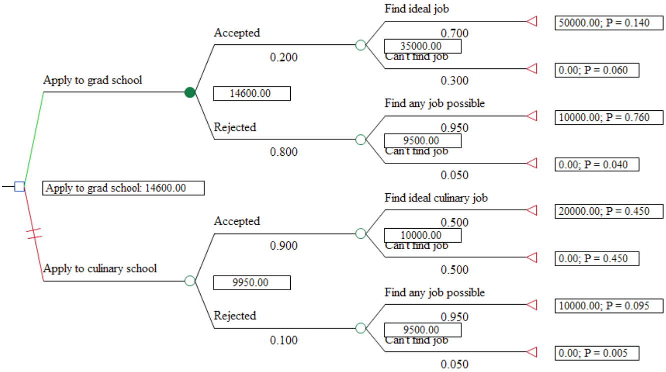
- Think about the expected value as a weighted average. The simple average is also weighted
- If you have three numbers, x_1, x_2, x_3, their average is $\frac{x_1+x_2+x_3}{3}$, which can also be written as
  $\frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3$
- Each number is weighted by the same amount $(1/3)$ and the weights add up to 1
- The expected value are the chances outcomes multiplied by their probabilities and summed: $E[X] = \sum_{i=1}^{\infty} x_i p(x_i)$
- The expected value in the context of decision trees are the payoffs weighted by their probabilities
- Another way: the payoff times the probability of obtaining that payoff

Expected value



- The expected value *after* being accepted into grad school is $0.7 \times 50,000 + 0.3 \times 0 = \$35,000$
- The expected value *after* being rejected is $0.95 \times 10,000 + 0.05 \times 0 = \$9,500$
- Using these results, we can then calculate the expected value of **applying to grad school**: $0.20 \times 35,000 + 0.80 \times 9,500 = \$14,600$

Rolling back with TreeAge



- The expected payoff from grad school is $14,600, which is larger than the expected payoff of culinary school, $9,950. Therefore, applying to grad school provides the highest expected payoff

Using Excel

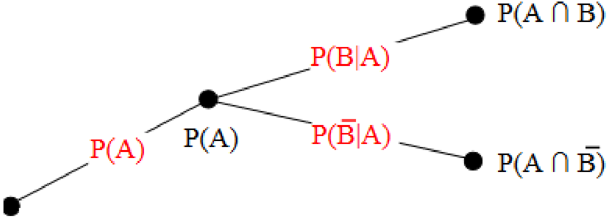
	Path	Branch probabilities		Probability Path	Payoff	Expected Payoff
	A	0.2	0.7	0.14	50000	7000
	B	0.2	0.3	0.06	0	0
Grad school	**C**	0.8	0.95	0.76	10000	7600
	D	0.8	0.05	0.04	0	0
				1		14600
	E	0.9	0.5	0.45	20000	9000
Culinary school	**F**	0.9	0.5	0.45	0	0
	G	0.1	0.95	0.095	10000	950
	H	0.1	0.05	0.005	0	0
				1		9950

- We can calculate the probability of **each pathway** and then multiply that probability by the corresponding payoff
- Note that we are using some rules of probability

More on probability

- **Conditional probability**: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- In words, the probability of event A happening given that event B has happened
- $P(A \cap B)$ is the probability of both events A and B happening (**joint probability**)
- The probability of **finding an ideal job if accepted to grad school** is a conditional probability because it depends on first being accepted intro grad school
- The probability of **pathway A** happening is the probability of two events happening: accepted intro grad school and finding an ideal job
- Solving for the joint probability: $P(A \cap B) = P(B) \times P(A|B)$. **That's why we multiply probabilities in decision trees**
- With more branches, events are **nested**, but the rule is the same. The probability of a pathway is a **joint probability**

Probability, graphically



- Graphically, from trusty Wikipedia
- If two events are independent, $P(A \cap B) = P(A) \times P(B)$ (but that's **not** why we multiply branches' probabilities)
- This implies that if A and B are independent $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times \cancel{P(B)}}{\cancel{P(B)}} = P(A)$
- In words, event B happening does not change the probability of event A happening

In case you got lost

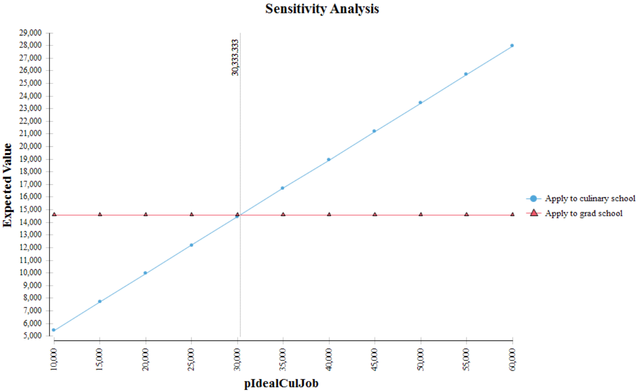
- We multiply the probabilities because their are **joint probabilities** (two or more events jointly occurring)
- We don't multiply because events are independent (they are not)
- Not sure why so many textbooks have it wrong
- You can calculate expected values **using the tree (rolling back) or using the probability of the path**. Both are equivalent
- You could also use Excel to draw a tree if that helps you
- Whichever way to choose to do it in the homework is fine

Using decision analysis

- Rolling the decision tree is never the end of the story. Actually, **it's just the beginning**
- Decision models are useful because they allow us to study the key elements of the decision
- One way to figure out which are the key elements is to see how changing some values affects the decision
- Changing some values also helps you figure out if the tree is not realistic or if you made assumptions that are not correct
- In other words, performing **sensitivity analyses**
- Some effects are obvious; others not so much

Changing the payoff

- What about if the salary after cooking school is too low? What happens if we change it?

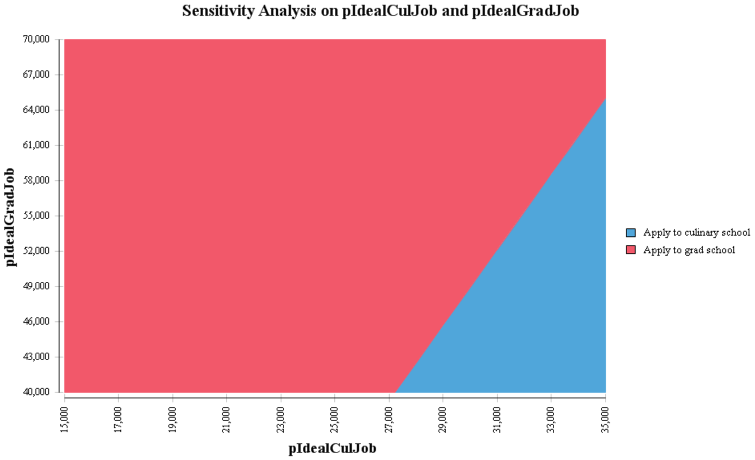


- Intial payoff ideal job after cooking school: 20K. EV grad school: 14,600; EV Cook 9,950

Changing payoff

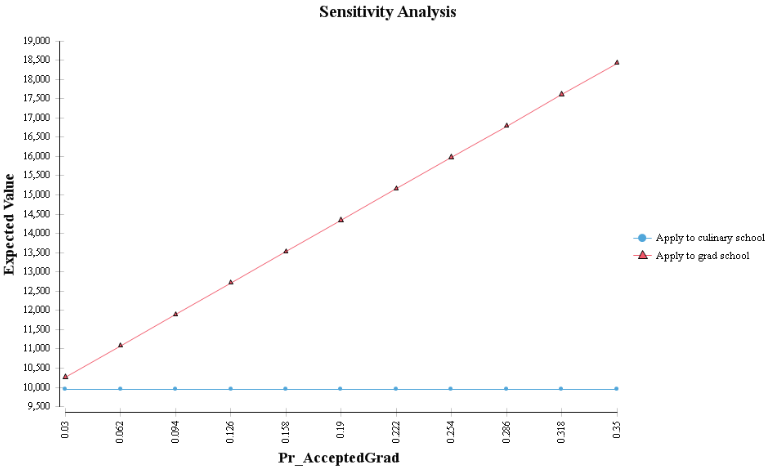
- The threshold is \$30,333
- If salary after cooking school is greater than \$30,333, then we should go to culinary school
- This is a **one-way sensitivity analysis** (we're changing only one parameter)
- Note that the line for applying to grad school is flat because we aren't changing the payoff after grad school
- Is 30K realistic? It could be. But not that likely perhaps. 50k after grad school is kind of low, too
- More interesting: let's change both at the same time time (**two-way sensitivity analysis**)

Changing both payoffs



- Range of 15K to 35K for culinary job and 40K to 70K after grad school

Changing the probability of acceptance to grad school

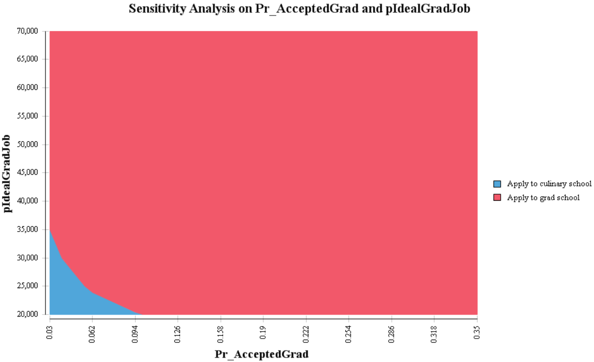


- This was **surprising**. Even at a very low probability of getting into grad school, the expected value of applying to grad school is higher

Why does this matter?

- Nothing surprising has happened other than relatively obvious things: bigger salaries for one option imply that that option is preferred
- On the other hand, we do have some quantitative measure about the importance of salary after each alternative
- Perhaps, though, our model is not that great. We could think of adding more salary options. What about a three-tier job prospect: high paying, average paying, no job?
- Going to cooking school is like aspiring to be a sports star: if you make it, it's great, but the chances of making it are thin
- On the other hand, a PhD (usually!!) gives you more choices that are often not badly paid, unless, of course, you decide that you want to teach...

Changing the Pr of acceptance to grad school and salary



- At low probability of getting into grad school, the salary after grad school would need to be very **low** to change the decision
- This one was surprising. I thought that a low probability of acceptance would change the decision but it doesn't

Sensitivity analysis, preview

- Sometimes **scenario analysis** are useful. For example: What about becoming a celebrity chef? Say, very low probability of ideal culinary job but very high salary
- Try some numbers using Excel. See if the decision would change
- We will see different ways of doing **sensitivity analyses**
- **But where do you get the numbers from**?

Real life

- In more serious modeling exercises, a lot of work goes into collecting the data
- Usually you present the scenario using the best parameter estimates (**baseline scenario**)
- Then you come up with a range of possible values for sensitivity analyses
- Sometimes you get weird results and make changes to the model

Adding another payoff

- We can of course consider other payoffs including a negative payoff
- One not-so-good thing about grad school is that it **requires a lot of effort** while cooking school may be more enjoyable (in general graduate school and mental health don't mix very well)
- We could come up with a number that reflects effort. Let's say a number between 0 and 1 (0 being no effort; 1 max effort)
- What is the expected effort level for each alternative?
- I chose some numbers

Salary and effort

	Path	Branch probabilities		Probability Path	Payoff	Expected Payoff	Effort	Expected Effort
	A	0.2	0.7	0.14	50000	7000	1	0.14
	B	0.2	0.3	0.06	0	0	1	0.06
Grad school	**C**	0.8	0.95	0.76	10000	7600	0.5	0.38
	D	0.8	0.05	0.04	0	0	0.5	0.02
				1		14600		0.6
	E	0.9	0.5	0.45	20000	9000	0.2	0.09
Culinary school	**F**	0.9	0.5	0.45	0	0	0.2	0.09
	G	0.1	0.95	0.095	10000	950	0.5	0.0475
	H	0.1	0.05	0.005	0	0	0.5	0.0025
				1		9950		0.23

- Each terminal node has an effort level now
- The expected effort for the grad school option is 0.6; the expected effort for the cooking school option is 0.23
- What about the salary per unit of effort? Or the incremental effort compared to the incremental salary?

Salary and effort

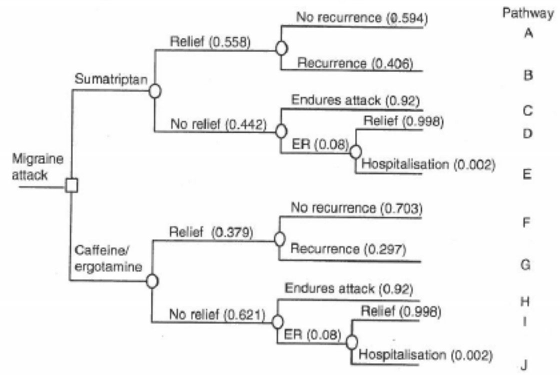
	Path	Branch probabilities		Probability Path	Payoff	Expected Payoff	Effort	Expected Effort
	A	0.2	0.7	0.14	50000	7000	1	0.14
	B	0.2	0.3	0.06	0	0	1	0.06
Grad school	**C**	0.8	0.95	0.76	10000	7600	0.5	0.38
	D	0.8	0.05	0.04	0	0	0.5	0.02
				1		14600		0.6
	E	0.9	0.5	0.45	20000	9000	0.2	0.09
Culinary school	**F**	0.9	0.5	0.45	0	0	0.2	0.09
	G	0.1	0.95	0.095	10000	950	0.5	0.0475
	H	0.1	0.05	0.005	0	0	0.5	0.0025
				1		9950		0.23

- Grad school versus cooking school: \$4650 extra (14,600-9,950) but 0.37 extra effort (0.6-0.23)
- We would a threshold or more comparisons to make sense of these numbers

What does this have to do with CEA?

- Hopefully you can see the connection by now
- In cost effectiveness, the decision node will have the alternatives we want to compare (e.g. new treatment versus usual care)
- The payoffs are 1) **costs** and 2) **outcomes** (life years, QALYs, natural units, etc)
- We can then calculated the (expected) ICER
- So what is different? We have incorporated **uncertainty**
- When using a decision tree to model CEA, **we do not really make a decision**; the final product will be an ICER, which, as usual, we need to compare to other ICERs or a threshold

CEA example



- Choosing between two medications (from Briggs et al, 2006); payoffs are costs and **utilities** (like assuming life years is 1 and doesn't change)
- Tree defines 10 pathways

Steps in decision modeling

1. **Identify the decision problem**: What are the alternatives? What is the research question? What are the key elements of the decision? Identify sub-populations
2. **Draw the tree**: It's very helpful to draw the tree to determine pathways
3. **Synthesize evidence**: Evidence usually comes from the literature. We need probabilities, costs, and benefits. Meta-analysis is a big part of CEA
4. **Analyze the tree and perform sensitivity analyses**: Always one-way to debug tree, set probabilities to 0 and 1 (things that should happen must happen), set payoffs to zero (alternatives should have same expected value)
5. **Go back to 1) when necessary**: Always an iterative process

The big picture

- Rolling back trees, defining probabilities, doing a sensitivity analysis are the **easy** parts
- The hardest part is to structure the tree. It requires **isolating the key elements** of the decision
- Without good knowledge of the problem, it is very hard to ascertain if a tree is modelling the situation correctly
- When reading a paper, always wonder what could be missing that is important
- And: **a paper without a sensitivity analysis is a bad paper**
- You will read this in every decision modeling book so I should just say it: *All models are wrong, some are useful*

Limitations of decision models for CEA

- Decision models do not easily handle recurrent events and disease progression
- We could add another tree to model a recurrent event, but this becomes very complicated very quickly
- For example, in the sumatriptan example, it is likely that another migrane attack will happen. The tree time horizon could be, say, a week. We would need to add another tree for next week. And another for the following week...
- For these reasons, decision models are commonly used to study short term CEA problems
- For long term CEA problems and recurrent events, Markov models are more common
- Markov models incorporate changes in disease states

Summary

- Decision models in CEA/CUA directly incorporate uncertainty
- The basics of measuring costs and benefits have not changed; but ICER now is an expected value
- The hardest part of decision models is to make sure the key elements of a decision are isolated and that the model represents the problem well
- Rolling back the tree, sensitivity analyses are the easy parts
- Decision models are limited when dealing with recurrent events and when we want to model disease progression
- Markov models to the rescue (next class)