Aalto University
School of Business

Business Analytics 2 – Lecture 5: Decision Making Under Uncertainty Cont'd

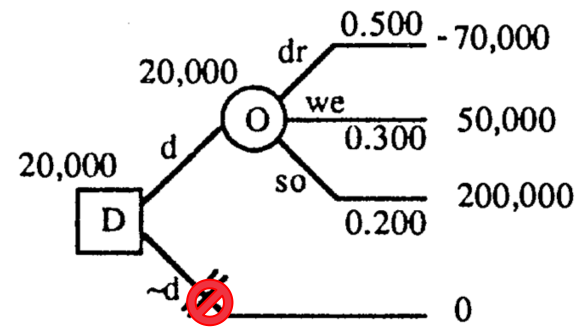
- *Probability distributions in decision trees*
- *Probability estimation: methods and biases*

Recap from week 3

We learned how to build and solve decision tree models to generate decision recommendations under uncertainty, when the objective is to maximize expected value / utility

We learned how to compute the expected value of

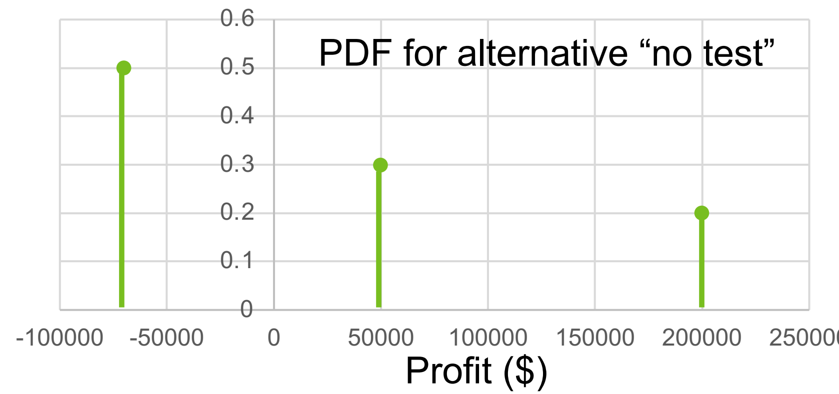
- Perfect information (i.e., the ability to make decisions after knowing how the uncertainties are resolved)
- Sample information (i.e., the ability to make decisions after obtaining imperfect information about how the uncertainties are resolved)

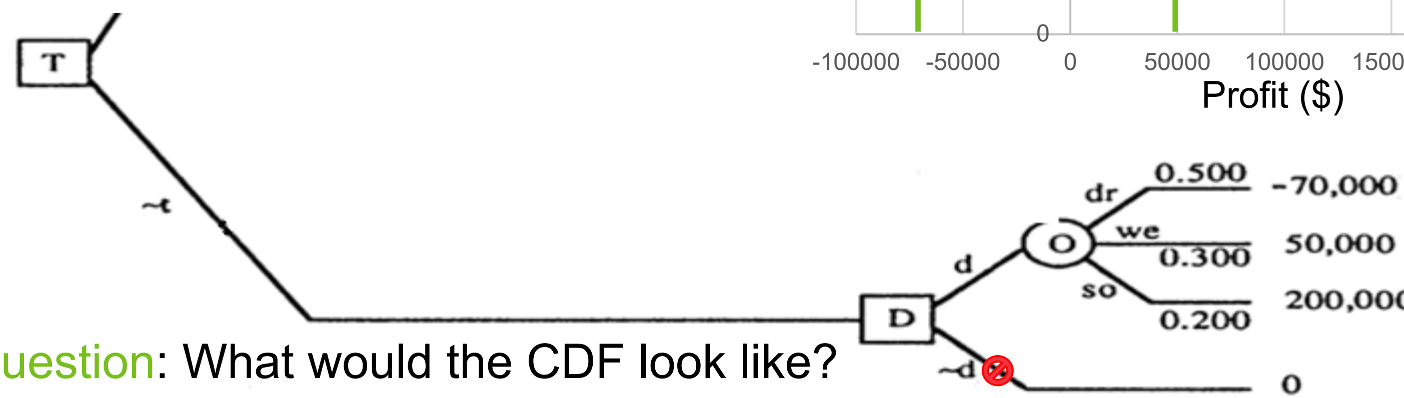


PDF of decision alternatives in decision trees

- Although we have focused on expected values, looking at the probability distribution over outcomes for different decision alternatives can be beneficial

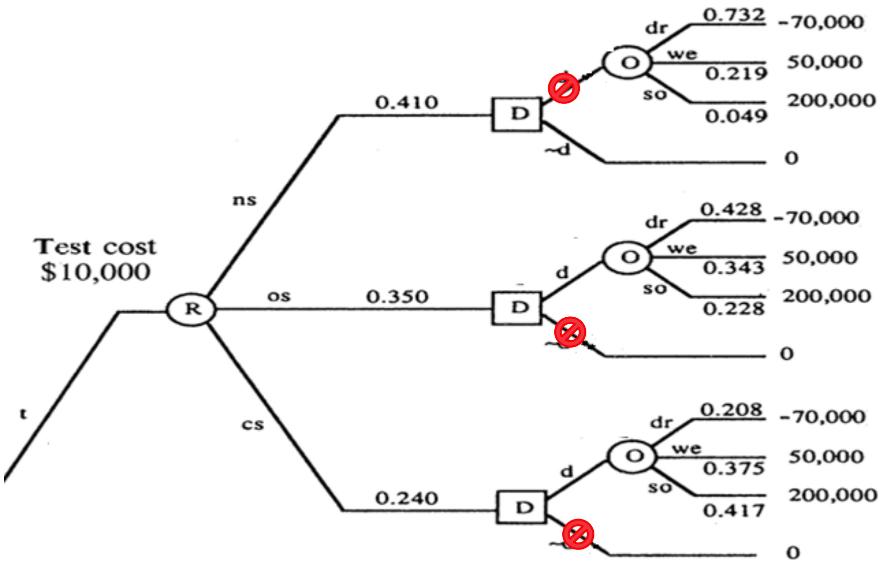
- Example: Oil wildcatting





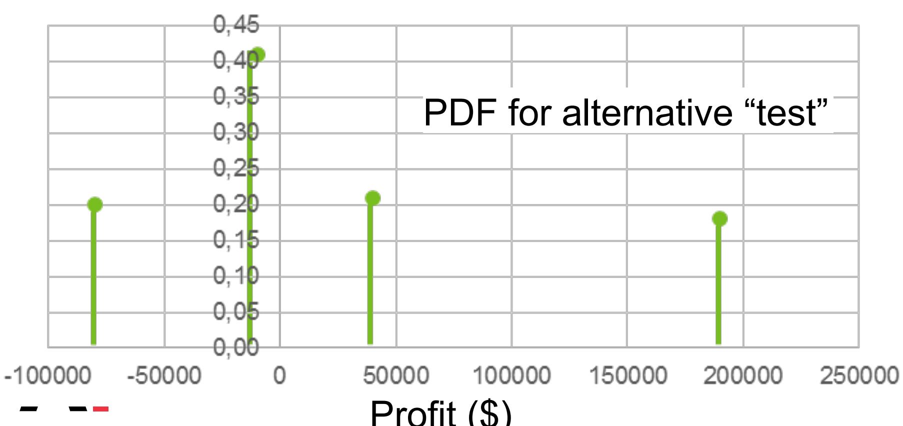
Question: What would the CDF look like?

PDF of decision alternatives in dec. trees (Cont’d)



R	P(R)	Profit	P(profit|R)	P(profit ∩ R)
ns	0.41	-10000	1	0.41
os	0.35	-80000	0.428	0.15
	0.35	40000	0.343	0.12
	0.35	190000	0.228	0.08
cs	0.24	-80000	0.208	0.05
	0.24	40000	0.375	0.09
	0.24	190000	0.417	0.10
			3.00	1.00

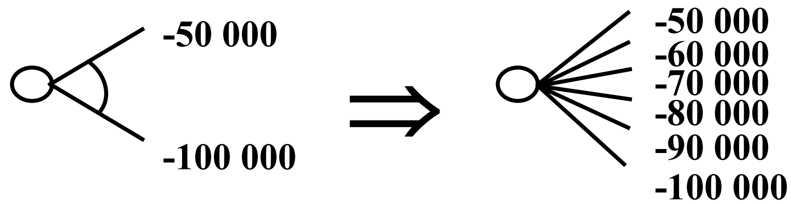
Profit	-80000	-10000	40000	190000	Sum
P(profit)	0.20	0.41	0.21	0.18	1.00
Prof*P(Prof.)	-15978	-4100	8402	34177	22502



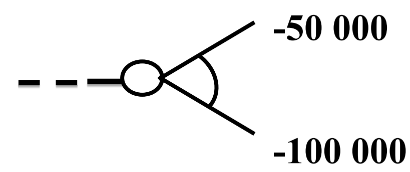
Question: What is the probability of negative profit if alternative “test” is chosen?

Continuous random variables in Decision Trees

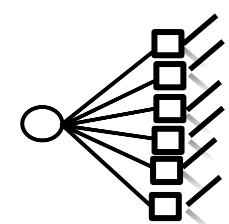
- Continuous r.v:s can be approximated with discrete events:



  - If no decisions are taken after the uncertainty is resolved, then the continuous distribution can be used, since only its expected value (later utility) is used in the tree

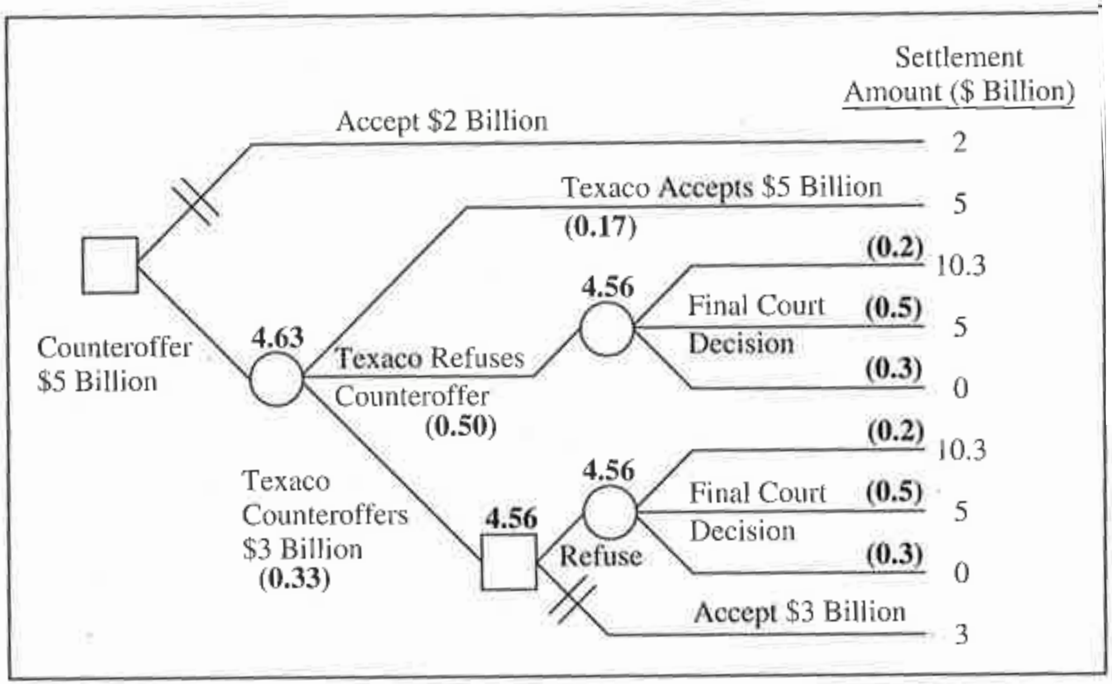


  - However, if decisions are made after the uncertainty is resolved, then continuous distributions may have to be discretized

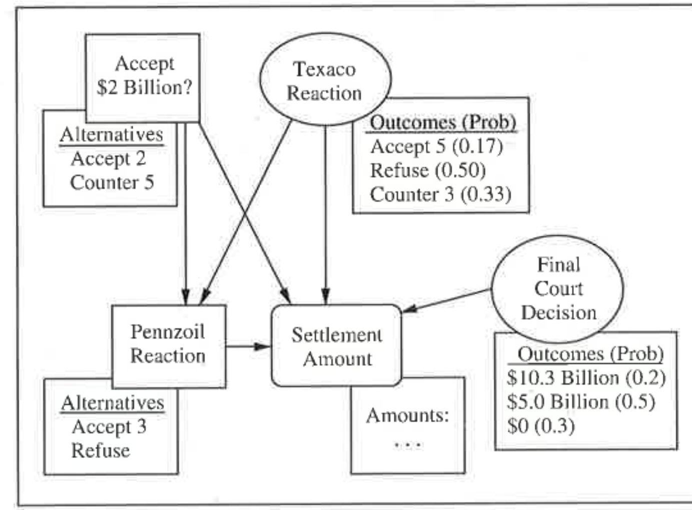


Decision tree example: Pennzoil v.s. Texaco*

Texaco has just offered 2 Billion to Pennzoil to settle the case.
Model for the Pennzoil decision:



Equivalent Influence diagram



*) Clemen (1996). Making Hard Decisions: An introduction to decision analysis

Estimation of probabilities

- How to obtain the probabilities needed in the models?

1. Use data if possible

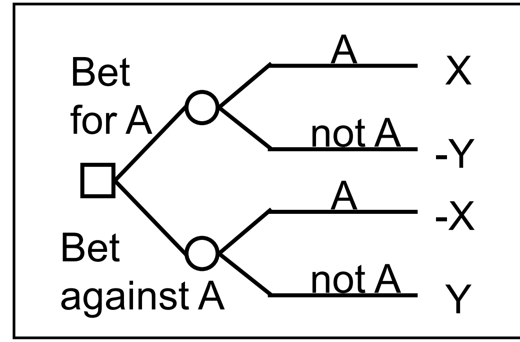
2. Often one has to rely on expert judgement methods
   - Betting approach
   - Reference lottery
   - Direct judgement

Example: ComputerWorld demand data for laptops selling for $4,300 over a period of 100 weeks.

PCs Demanded per Week	Frequency of Demand	Probability of Demand, $P(x)$
0	20	.20
1	40	.40
2	20	.20
3	10	.10
4	10	.10
	100	1.00

Estimation of probabilities: Betting approach

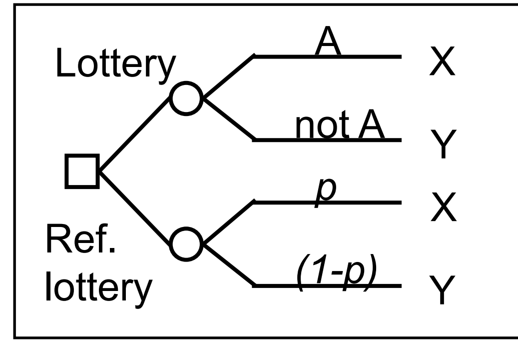
- Goal: Estimating the probability of event A
  - E.g. A="Demand over 145 units" or A="GDP growth below 5%"
  - Bet for A:
    - Win X € if A happens
    - Lose Y € if A does not happen.
  - Bet against A:
    - Lose X € if A happens
    - Win Y € if A does not happen.
  - Adjust X and Y until respondent is **indifferent** about which bet to take
  - Now the expected monetary values of the bets must be equal*:



$$XP(A) - Y[1 - P(A)] = -XP(A) + Y\big(1 - P(A)\big) \Rightarrow P(A) = \frac{Y}{X+Y}$$

* Strong assumption: The respondent maximizes expected monetary value (is risk neutral).

Estimation of probabilities: Reference lottery



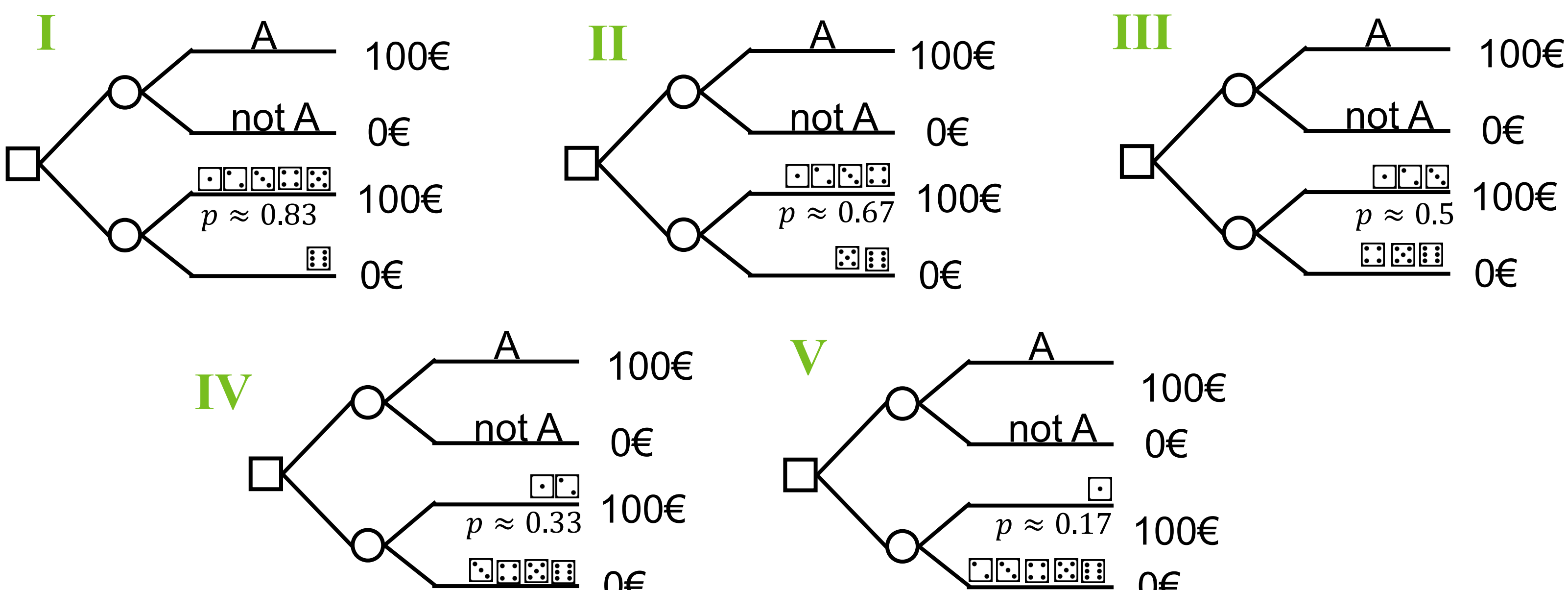
- Estimate the probability of event A
  - The Lottery:
    - Win X if A happens
    - Win Y if A does not happen.
  - The Reference Lottery
    - Win X with probability p
    - Win Y with probability $(1-p)$.
    - Probability p can be visualised e.g. with a wheel of fortune or a box with red and blue balls from which you want to draw a red one.
  - Adjust p until respondent is **indifferent** about which lottery to participate in

$$XP(A) + Y[1 - P(A)] = Xp + Y(1-p)$$
$$\Rightarrow P(A) = p$$

- The respondent's risk attitude does not affect the result (shown later)

Reference lottery: Exercise

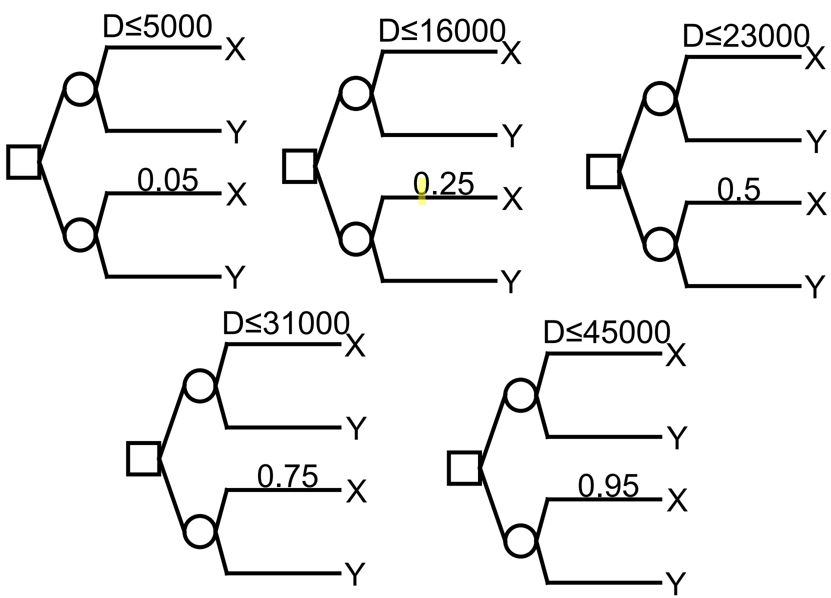
- Event A="Humans walk on the surface of Mars by the end of 2035"
  1. In each decision tree I-V choose the preferred lottery
  2. Derive an estimate for P(A)



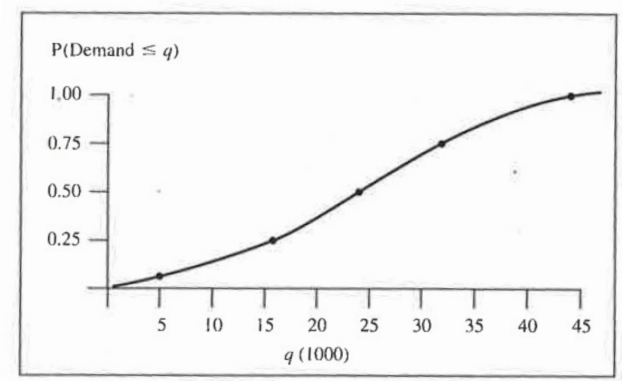
If you prefer the bottom lottery then $P(A) < p$
If you prefer the top lottery then $P(A) > p$

Estimation of continuous probability distributions

- Distribution of a continuous random variable can be approximated by estimating the several event probabilities
- Example: Assessing the distribution for uncertain demand
  - The expert is indifferent between the lotteries in each of the five cases:



CDF for Demand

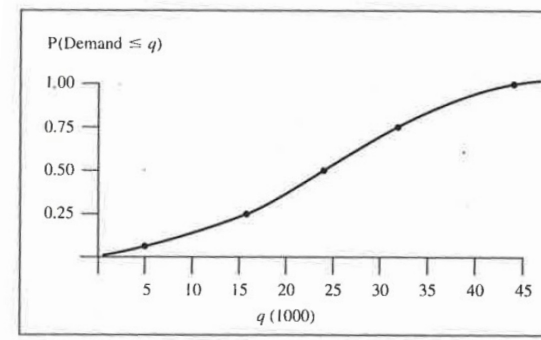


If DM is indifferent between the two lotteries then $P(A) = p$

Estimation of cont. probability distributions (Cont'd)

- Often experts directly assess some descriptive statistics of the distribution, e.g.,
  - the feasible range (min, max)
  - median (i.e. P(X<median) = 0.5)
  - other quantiles (e.g. 5%, 25%, 75%, 95%)
- Example: Draw the CDF based on expert judgements
  - "The 5% and 95% quantiles are 5000 and 45000"
  - "Demand is just as likely to be above 23000 as below it"
  - "There is a 25% chance that the demand will be below 16000"
  - "There is a 75% chance that the demand will be below 31000"

CDF for Demand



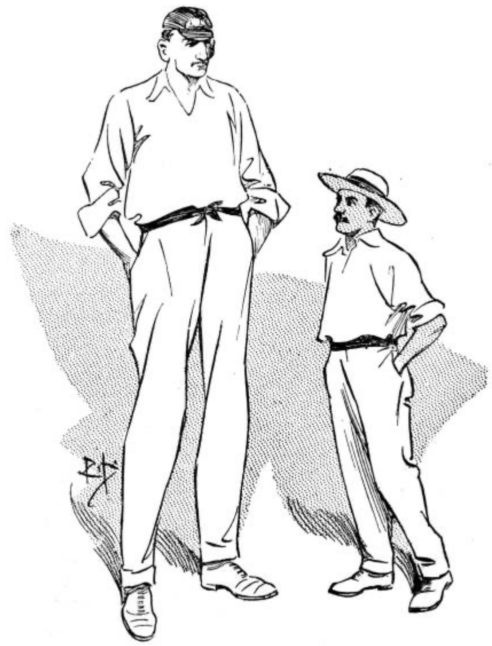
Probability estimation

Question: Consider two bags X and Y. Bag X contains 30 white balls and 10 black balls, whereas bag Y contains 30 black balls and 10 white balls. Suppose that you select one of these bags at random, and randomly draw five balls one-by-one by replacing them in the bag after each draw. Suppose you get four white balls and one black. What is the probability that you selected bag X with mainly white balls?

Probability estimation

Question: You see a very tall man in a bar. Is he more likely to be a professional basketball player or a teacher?

Question: You see a very photogenic woman in a bar. Is she more likely to be a model or a nurse?



Probability estimation: Tom W problem

Tom W. is of high intelligence, although lacking in true creativity. He has a need for order and clarity, and for neat and tidy systems in which every detail finds its appropriate place. His writing is rather dull and mechanical, occasionally enlivened by somewhat corny puns and by flashes of imagination of the sci-fi type. He has a strong drive for competence. He seems to have little feel and little sympathy for other people and does not enjoy interacting with others. Self-centered, he nonetheless has a deep moral sense.

How <u>similar</u> is Tom W. to the typical student in business, computer science, engineering, humanities, law, library science, medicine, life sciences, and social sciences?

Kahneman, Tversky (1973). On the psychology of prediction. *Psychological Review*. 80 (4): 237–251.

Probability estimation: The Linda problem

Linda is 31 years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in anti-war demonstrations.

Rank the following items in terms of probability:

A. Linda is a school teacher
B. Linda works in a bookstore and takes yoga classes
C. Linda is active in the feminist movement
D. Linda is a psychiatric social worker
E. Linda is a member of the League of Women Voters
F. Linda is a bank teller
G. Linda is an insurance salesperson
H. Linda is a bank teller and active in the feminist movement

Biases in probability elicitation

- People use heuristics, i.e., rules of thumb, in assessing probabilities
  - May lead to systematically incorrect estimates, i.e., biases
  - E.g. gambler's fallacy:
    - You toss a coin 7 times and get HHHHHHH (H=heads). Which (H or T) is likely to come next?
    - Is P(HHHHTTTT) < P(HTTHHTHT)?

- When estimating subjective probabilities it is important to mitigate the effects of these biases by
  - Making the experts aware of them
  - Using several modes of questioning

Biases in probability elicitation: Representativeness / similarity

- If x fits the description of a set A well, then prob. that x is a member of A is assumed to be high, even if similarity is not the same as likelihood
- Often this is due to base rate neglect
  - E.g. in the Tom W problem base rates for computer science students are low, so people thinking Tom is representative of comp sci neglect the base rate
- Example:
  - You see a very tall man in a bar. Is he more likely to be a professional basketball player or a teacher?
  - Many people are tempted to answer "a professional basketball player"
  - Yet, professional basketball players are rare, while there are far more teachers: thus, he is actually more likely to be a teacher (base rate of teachers is high)

Biases in probability elicitation: Availability

- People assess the probability of an event by the ease with which instances or occurrences of this event can be brought to mind

- Example:
  - In a typical sample of text in English, is it more likely that a word starts with the letter K or that K is the third letter?
  - Generally about 70 % think that words starting with K are more common. In reality, however, there are approximately twice as many words with K in the third position as there are words that begin with it.

Biases in probability elicitation: Anchoring

- In several situations people assess probabilities by adjusting a given number from a starting value
- Often the adjustment is not large enough and the final assessment is too close to the starting value
- Example:
  - Is the percentage of African countries in the UN greater or less than 65? What is the exact percentage?
    - Average answer: Less, 45%
  - Is the percentage of African countries in the UN greater or less than 10? What is the exact percentage?
    - Average answer: Greater, 25%
- How to reduce: Avoid giving starting values

Biases in probability elicitation: Conservatism

- People tend to change previous probability estimates more slowly than warranted by new data

Example:

- Consider two bags X and Y. Bag X contains 30 white balls and 10 black balls, whereas bag Y contains 30 black balls and 10 white balls. Suppose that you select one of these bags at random, and randomly draw five balls one-by-one by replacing them in the bag after each draw. Suppose you get four white balls and one black. What is the probability that you selected bag X with mainly white balls?

Biases in probability elicitation: Overprecision

- When asked to provide an estimate for a random variable and 95% confidence intervals, people tend to provide too narrow confidence intervals, being overprecise in their estimates
  - 95% confidence interval: the interval in which the value of a random variable should be with $p = 0.95$
- Fundamentally a problem of underestimating the variance of a random variable / placing too much weight on its expected value
- Experts such as weather forecasters are often better calibrated and less overprecise than lay people
- How to reduce the bias: widen your confidence intervals

Biases in probability elicitation: Overoptimism

- People tend to overestimate the probability of events that have a positive consequence, and underweight the probability of events that have a negative consequence
- Example:
  - What is the probability of BTC price being less than 10,000€ at the end of 2024?
  - Your probability judgment is likely to depend on whether you think this would be a good or a bad thing
- How to avoid:
  - Ask for credible storylines about how the event would come to occur
  - Present credible storylines for alternative outcomes
  - Elicit probability estimates from multiple sources with different preferences