Probability judgement

Munich-Sydney-Tilburg Conference
Evidence, Inference, and Risk

Munich Center for Mathematical Philosophy
April 1, 2016
Glenn Shafer, Rutgers University

1. Introduction: Subjective and objective probability

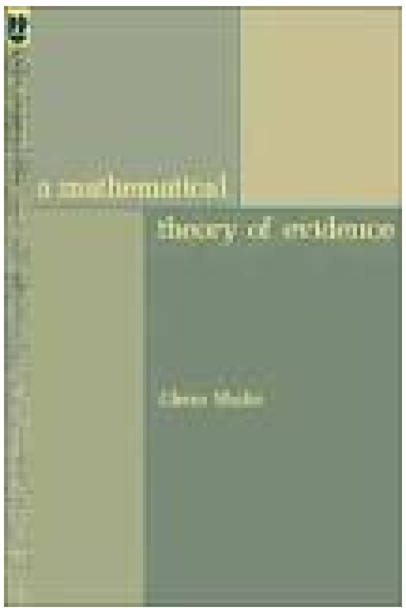
2. Probability judgment: From evidence to bets

3. Example: Glenn's house in 1978

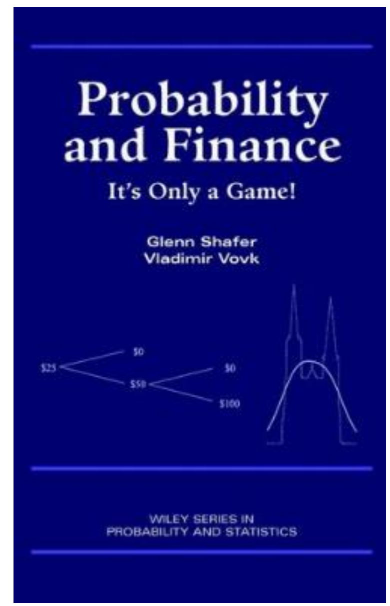
Two goals:

1. Illustrate Dempster-Shafer probability judgment.

2. Place Dempster-Shafer in the game-theoretic framework.



Dempster-Shafer theory
Glenn Shafer
1976



Game-theoretic probability
Glenn Shafer
Vladimir Vovk
2001

Part 1. Subjective & objective probability

- Price-maker vs price-taker
- Probability judgment = argument by analogy
- Causal conjecture = conjecture about objective probabilities

Part 2. From evidence to bets

- Using probability games from the outside
- Bayesian canonical examples
- Dempster-Shafer canonical examples

Part 3. Example: Glenn’s house in 1978

- The evidence
- A Dempster-Shafer assessment
- A Bayesian assessment

Part 1. Subjective & objective probability

- Two subjective interpretations
  1. price-maker
  2. price-taker

- Probability judgment
  = constructing subjective probabilities
  = assessing evidence
  = an argument by analogy

- Causal conjecture
  =conjecture about objective probabilities
  = conjecture about Nature's probabilities

Bruno de Finetti
1906-1985

- Right that subjective probability is about betting.
- Wrong to neglect strategic aspect of betting.

Betting involves two players.

- Price-maker expresses belief by offering bets.
  Maybe he believes he will break even in the long run.

- Price-taker expresses beliefs by deciding which bets to take.
  Maybe he believes he can multiply the capital he risks by a large factor.

The two points of view produce different subjective interpretations and different objective interpretations of probability.

Conceptually and historically, probability is about betting games

Ernest Nagel: different interpretations of probability are different ways of using Kolmogorov’s axioms.

Betting games provide a mathematical framework deeper and more powerful than axioms for measure.

Improving on Nagel: different interpretations of probability are different ways of using betting games.

Skeptic chooses bets
Reality decides outcomes

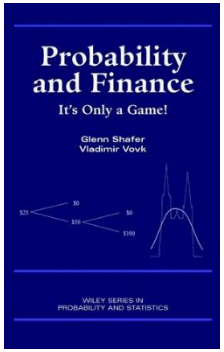
Suppose:

Outcomes are 1s and 0s.

Skeptic is allowed to bet at even odds on each round.

$$\boxed{\begin{array}{l} \text{Skeptic announces } \mathcal{K}_0 \in \mathbb{R}. \\ \text{FOR } n = 1, 2, \ldots: \\ \quad \text{Skeptic announces } L_n \in \mathbb{R}. \\ \quad \text{Reality announces } y_n \in \{0, 1\}. \\ \quad \mathcal{K}_n := \mathcal{K}_{n-1} + L_n(y_n - \frac{1}{2}). \end{array}}$$

- Perfect information game.
- $\mathcal{K}_n$ = Skeptic's capital after round n.
- $\mathcal{K}_0$ = Skeptic's initial capital = Skeptic's stake



Shafer/Vovk
2001, Wiley

Recall the game-theoretic definition of probability:

$$\mathbb{P}(A) = \{\text{stake Skeptic needs to get 1 if } A \text{ happens, 0 otherwise}\}$$

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
  Skeptic announces $L_n \in \mathbb{R}$.
  Reality announces $y_n \in \{0, 1\}$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + L_n(y_n - \frac{1}{2})$.

Ville showed that Skeptic has a strategy that multiplies his stake by ∞ if

$$\frac{\sum_{i=1}^n y_i}{n} \not\to \frac{1}{2}.$$

Therefore $\mathbb{P}\left(\frac{\sum_{i=1}^n y_i}{n} \not\to \frac{1}{2}\right) = 0.$

Parameters: $p_1, p_2, \ldots \in [0,1]$
Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Skeptic announces $L_n \in \mathbb{R}$.
Reality announces $y_n \in \{0,1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + L_n(y_n - p_n)$.

$$\mathbb{P}\left(\frac{\sum_{i=1}^{n}(y_i - p_i)}{n} \to 0\right) = 1$$

Generalize to three players

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
Forecaster announces $p_n \in [0,1]$.
Skeptic announces $L_n \in \mathbb{R}$.
Reality announces $y_n \in \{0,1\}$.
$\mathcal{K}_n := \mathcal{K}_{n-1} + L_n(y_n - p_n)$.

$$\mathbb{P}\left(\frac{\sum_{i=1}^{n}(y_i - p_i)}{n} \to 0\right) = 1$$

Using betting games from inside or outside

1. Testing (from inside the game)

- Play Skeptic to test Forecaster.
- Play Honest Skeptic to correct exaggerated p-values

2. Forecasting (from inside the game)

- Use theory (e.g., quantum mechanics) to give forecasts.
- Average forecasts (this is sometimes called Bayesian).
- Defensive forecasting (play against Skeptic's tests)

3. Causal modeling (from outside the game)

- Nature is the Forecaster, but we do not see her Forecasts.
- From statistical regularities, we conjecture relations in Nature's strategy.

4. Probability judgement (from outside the game)

- Use betting games as scale of canonical examples for probability judgement.
- Develop different designs for decomposing & recombining the evidence.

Skeptic announces $\mathcal{K}_0 \in \mathbb{R}$.
FOR $n = 1, 2, \ldots$:
  Forecaster announces $p_n \in [0, 1]$.
  Skeptic announces $L_n \in \mathbb{R}$.
  Reality announces $y_n \in \{0, 1\}$.
  $\mathcal{K}_n := \mathcal{K}_{n-1} + L_n(y_n - p_n)$.

Forecaster is the price-maker. May be…

- theory
- actual forecaster
- Nature
- market

Skeptic is the price-taker. May be…

- actual hypothesis-testing statistician
- Cournotian/Peircian limit of knowledge
- investor

Part 2. From evidence to bets

- Using probability games from the outside
- Bayesian canonical examples
- Dempster-Shafer canonical examples

George Hooper
1640-1727

Jakob Bernoulli
1655-1705

Johann Heinrich Lambert
1728-1777

How do we connect evidence with betting?

Probability theory is about betting.

Assessment of evidence is about

- analyzing the meaning of evidence,
- weighing competing arguments,
- combining arguments and conclusions.

After assessing the evidence…

- you might be inclined to bet (subjectivists), or
- you might believe certain bets cannot be beat (Cournotians).

But how do we get there?

How do we get from the evidence to the bets?

From evidence to probability

Assessment of evidence usually involves:

- Decomposition of the evidence.
- Judgement about the direction each item points.
- Assessment of the strength of individual items of evidence by comparison to canonical examples.
- Recombination of individual assessments.

Decomposition and recombination require a design.

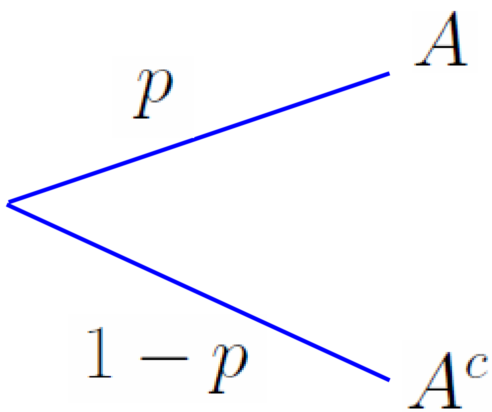
Different theories of probability judgement provide different canonical examples and different designs.

Bayes and Dempster-Shafer are two such theories.

Both use betting games as canonical examples and use Cournotian judgements of irrelevance in designs.

Languages and designs for probability judgment, by Glenn Shafer and Amos Tversky. *Cognitive Science* **9**:309-339. 1985.

The simplest canonical example



We consider this forecast unbeatable in the Cournotian sense.

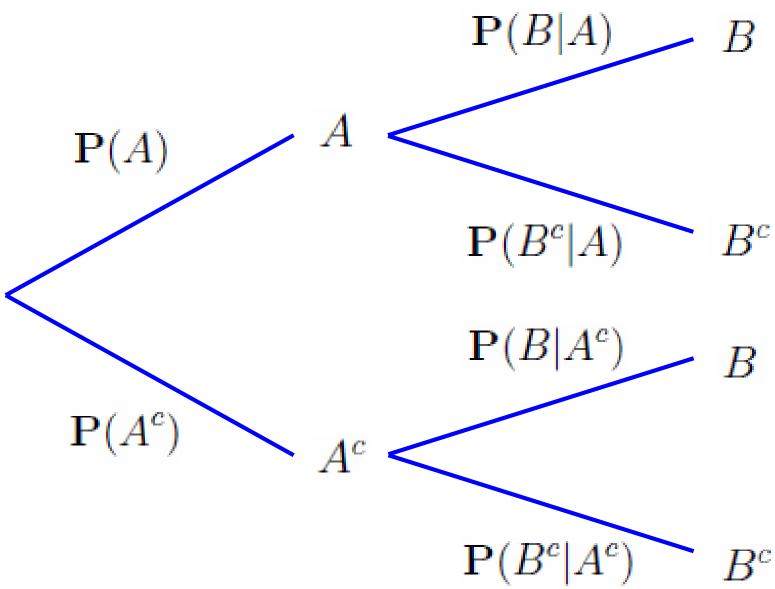
(That is to say, unbeatable when embedded in a sequence of similar forecasts.)

Scale of canonical examples:

Consider all $p \in [0, 1]$.

We can weigh given evidence for A by placing it on this scale.

A Bayesian canonical example



Initially we consider this forecasting strategy unbeatable in the Cournotian sense.

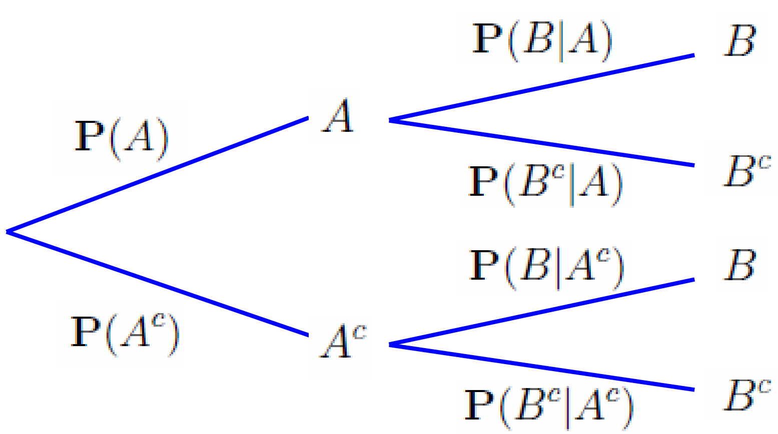
We learn that B is true and consider the way we learned it irrelevant in the Cournotian sense: it cannot help us beat the forecasting strategy.

This authorizes the betting rate

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A)\mathbf{P}(B|A)}{\mathbf{P}(A)\mathbf{P}(B|A) + \mathbf{P}(A^c)\mathbf{P}(B|A^c)}.$$

A Bayesian design

The scale of canonical examples:
This example with all different values of $\mathbf{P}(A)$, $\mathbf{P}(B|A)$, and $\mathbf{P}(B|A^c)$.



Initially we consider this forecasting strategy unbeatable in the Cournotian sense.

We learn that B is true and consider the way we learned it irrelevant in the Cournotian sense: it cannot help us beat the forecasting strategy.

This authorizes the betting rate

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A)\mathbf{P}(B|A)}{\mathbf{P}(A)\mathbf{P}(B|A) + \mathbf{P}(A^c)\mathbf{P}(B|A^c)}.$$

DESIGN TO ASSESS EVIDENCE FOR A

A. Decompose evidence into 4 parts.

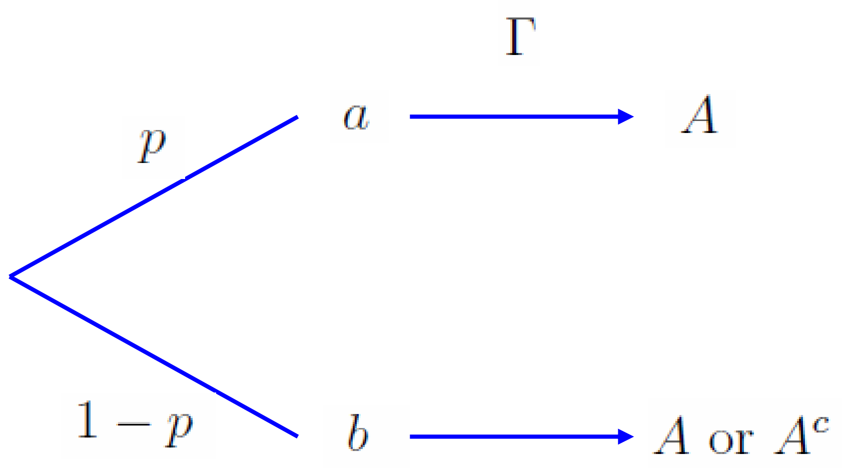
1. Knowledge of a salient fact B.
2. Evidence for A supposing we did not know B.
3. Evidence for B if we knew A.
4. Evidence for B if we knew A^c.

B. Use 2, 3, and 4 to assess $\mathbf{P}(A)$, $\mathbf{P}(B|A)$, and $\mathbf{P}(B|A^c)$, respectively.

C. Recombine by Bayes theorem.

In 1985, Tversky and I called this a "conditioning design". See our joint paper for ideas about other Bayesian designs.

The simplest Dempster-Shafer canonical example



We consider forecasting strategy given by p and 1-p unbeatable in the Cournotian sense.

We learn the multivalued mapping Γ, which tells us that A is true if a happened, and consider the way we learned it irrelevant in the Cournotian sense. This authorizes the betting rate p for A.

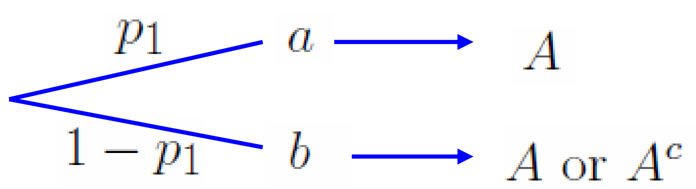
A Dempster-Shafer canonical example

To illustrate the idea of obtaining degrees of belief for one question from subjective probabilities for another, suppose I have subjective probabilities for the reliability of my friend Betty. My probability that she is reliable is 0.9, and my probability that she is unreliable is 0.1.

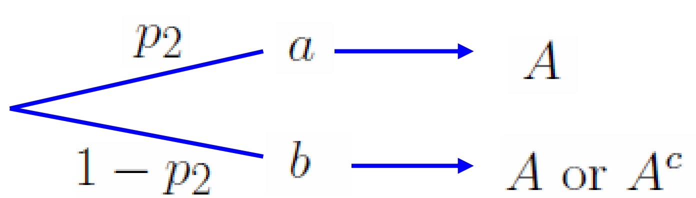
Suppose she tells me a limb fell on my car. This statement, which must true if she is reliable, is not necessarily false if she is unreliable. So her testimony alone justifies a 0.9 degree of belief that a limb fell on my car, but only a zero degree of belief (not a 0.1 degree of belief) that no limb fell on my car.

This zero does not mean that I am sure that no limb fell on my car, as a zero probability would; it merely means that Betty's testimony gives me no reason to believe that no limb fell on my car. The 0.9 and the zero together constitute a belief function.

The simplest Dempster-Shafer design



We judge $(p_1, 1 - p_1)$ unbeatable. We learn A is true if a happened, and we judge the way we learned it irrelevant to beating $(p_1, 1 - p_1)$. This authorizes p_1 for A.



We judge $(p_2, 1 - p_2)$ unbeatable. We learn A is true if c happened, and we judge the way we learned it irrelevant to beating $(p_2, 1 - p_2)$. This authorizes p_2 for A.

We judge each story irrelevant to beating the strategy in the other story. This authorizes Hooper's rule:

$$p = 1 - (1 - p_1)(1 - p_2).$$

THE DESIGN

A. Divide evidence into two intuitively independent arguments for A.

B. Assess each argument using the scale of simple Dempster-Shafer canonical examples.

C. Recombine by Hooper's rule.

The simplest Dempster-Shafer design

To illustrate Dempster's rule for combining degrees of belief, suppose I also have a 0.9 subjective probability for the reliability of Sally, and suppose she too testifies, independently of Betty, that a limb fell on my car.

The event that Betty is reliable is independent of the event that Sally is reliable, and we may multiply the probabilities of these events; the probability that both are reliable is 0.9x0.9 = 0.81, the probability that neither is reliable is 0.1x0.1 = 0.01, and the probability that at least one is reliable is 1 - 0.01 = 0.99.

Since they both said that a limb fell on my car, at least of them being reliable implies that a limb did fall on my car, and hence I may assign this event a degree of belief of 0.99.

Another Dempster-Shafer design

Suppose Betty and Sally contradict each other—Betty says a limb fell on my car, and Sally says no limb fell on my car.

In this case, they cannot both be right and hence cannot both be reliable—only one is reliable, or neither is.

The prior probabilities that only Betty is reliable, only Sally is reliable, and that neither is reliable are 0.09, 0.09, and 0.01, respectively.

The posterior probabilities (given that not both are reliable) are 9/19 , 9/19 , and 1/19 , respectively.

Hence we have a 9/19 degree of belief that a limb did fall on my car (because Betty is reliable) and a 9/19 degree of belief that no limb fell on my car (because Sally is reliable).

Dempster-Shafer design older than Bayesian design!

Used by

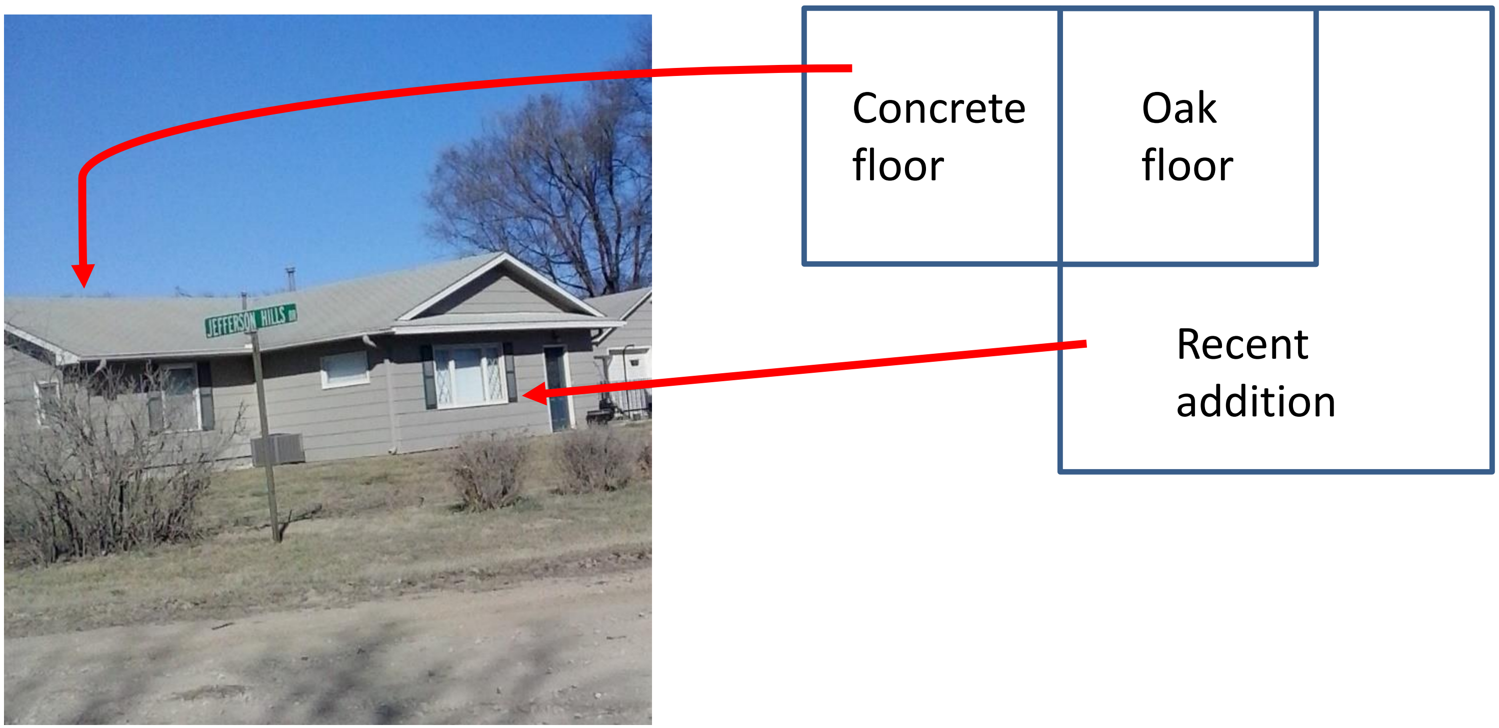
- Hooper in 1699
- Bernoulli in 1713
- Lambert in 1764

See

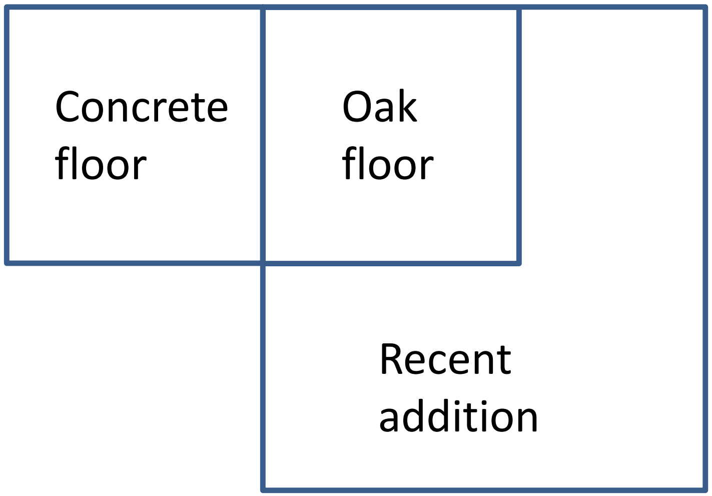
- Andrew Dale, On the authorship of “A Calculation of Human Testimony”, *Historia Mathematica* 19(4):414-417. 1992.
- Glenn Shafer, Non-additive probabilities in the work of Bernoulli and Lambert. *Archive for History of Exact Sciences* **19** 309-370. 1978.
- Glenn Shafer, The combination of evidence. *International Journal of Intelligent Systems* **1** 155-179. 1986.

Part 3. Example: Glenn's house in 1978

- The evidence
- A Dempster-Shafer assessment
- A Bayesian assessment



Which came first? Oak or concrete?



My house in 1978

Five items of evidence

1. Partition was probably original outside wall.
2. Concrete footing under partition appears to be part of concrete floor.
3. Footing under oak floor seems inadequate for outside wall.
4. Stubs in attic may have been rafters for concrete section.
5. Neighbor thinks the original building was quarry office.

I conclude that the concrete section probably came first.

This example from Two theories of probability, *PSA 1978*, Vol. 2, pp. 441-464. Peter D. Asquith and Ian Hacking, eds. Philosophy of Science Association, East Lansing, Michigan. 1981. At www.glennshafer.com.

Five items of evidence

1. Partition was probably original outside wall.
2. Concrete footing under partition appears to be part of concrete floor.
3. Footing under oak floor seems inadequate for outside wall.
4. Stubs in attic may have been rafters for concrete section.
5. Neighbor thinks the original building was quarry office.

This evidence does not consist of propositions.

The five statements are only summaries.

- I have evaluated visual impressions – the small quarry.
- I have evaluated nonverbal signals – my impression of neighbor's reliability.
- I have sampled from my memory of similar situations (rafters, footings,…).

Three issues

A = concrete section built first

A^c = oak section built first

B = partition on concrete floor

B^c = partition on separate concrete foundation

C = partition formerly outside wall

C^c = former outside wall was removed

I decide to rule out $A^c\&B\&C$.

Seven possibilities:

$a = A\&B\&C$

$b = A\&B^c\&C$

$c = A^c\&B^c\&C$

$d = A^c\&B\&C^c$

$e = A\&B\&C^c$

$f = A\&B^c\&C^c$

$g = A^c\&B^c\&C^c$

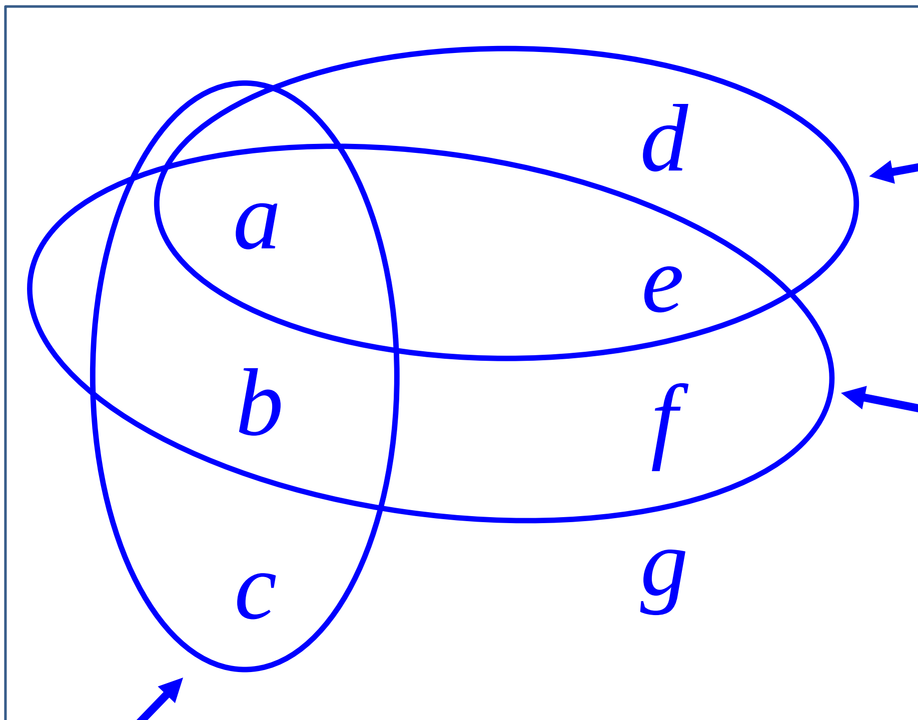
$A = \{a,b,e,f\}$

$B = \{a,d,e\}$

$C = \{a,b,c\}$

First step (already using the evidence)

Formulate *frame of discernment,* $\Theta = \{a,b,c,d,e,f,g\}$.



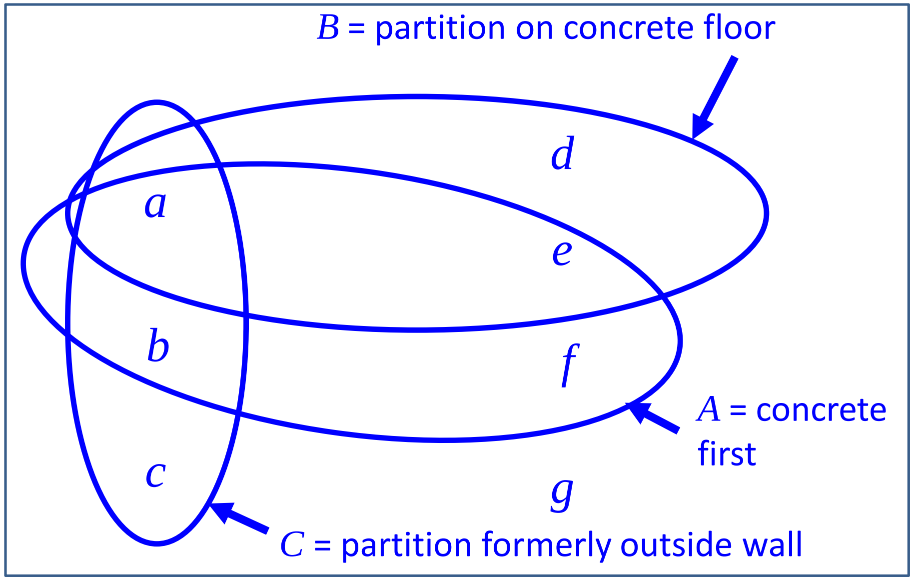
B = partition on concrete floor

A = concrete section built first

C = partition formerly outside wall

Want to make a probability judgement about A.

I have ruled out $A^c \& B \& C$.

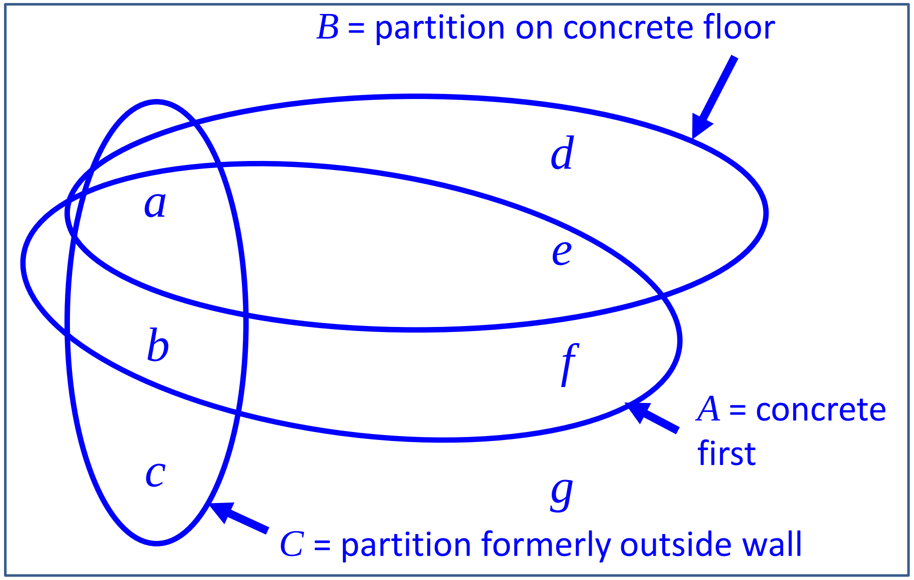


Making probability judgement about A.

I have already ruled out $A^c\&B\&C$.

Tendency of each item of evidence

1. Partition probably original outside wall. Supports C.
2. Concrete under partition apparently part of concrete floor. Supports B.
3. Footing under oak floor inadequate for an outside wall. Supports $\{a,b,c,e,f\}$.
4. Attic stubs suggest roof over concrete section. Supports $\{a,b,e\}$.
5. Neighbor says office for quarry. Supports A.

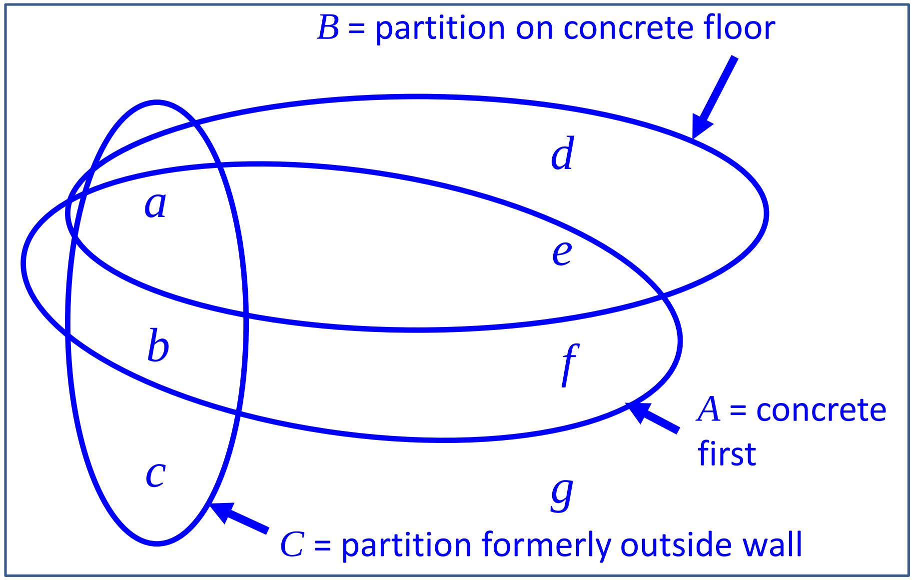


Dempster-Shafer assessment

1. Score each item using the scale of simple canonical examples.

Scoring each item of evidence

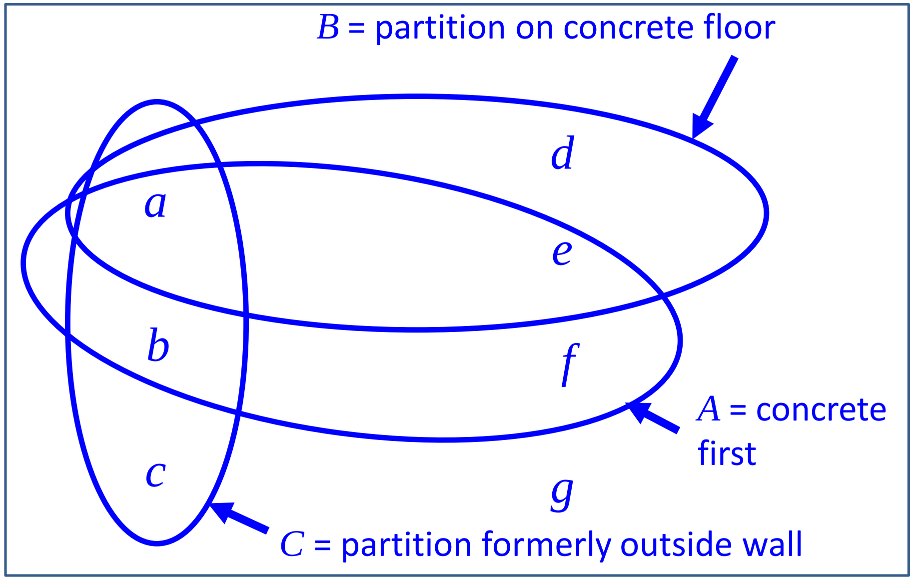
1.	Partition probably original outside wall. Supports C.	0.80
2.	Concrete under partition apparently part of concrete floor. Supports B.	0.98
3.	Footing under oak floor inadequate for an outside wall. Supports $\{a,b,c,e,f\}$.	0.50
4.	Attic stubs suggest roof over concrete section. Supports $\{a,b,e\}$.	0.80
5.	Neighbor says office for quarry. Supports A.	0.08



Dempster-Shafer assessment

Partition probably original outside wall.	
Supports C.	0.80
Concrete under partition part of concrete floor.	
Supports B.	0.98
Footing under oak floor inadequate.	
Supports $\{a,b,c,e,f\}$.	0.50
Attic stubs suggest roof over concrete section.	
Supports $\{a,b,e\}$.	0.80
Neighbor says office for quarry.	
Supports A.	0.08

2. Combine by Dempster’s rule, to obtain total belief in A of 0.98.



1. Partition probably original outside wall.
   0.80 for C 0.20 uncommitted
2. Concrete under partition part of concrete floor.
   0.98 for B 0.02 uncommitted
3. Footing under oak floor inadequate.
   0.50 for $\{a,b,c,e,f\}$ 0.50 uncommitted
4. Attic stubs suggest roof over concrete section.
   0.80 for $\{a,b,e\}$ 0.20 uncommitted
5. Neighbor says office for quarry.
   0.08 for A 0.92 uncommitted

Dempster's rule.

We make the Cournotian judgement that each item of evidence provides us with no information that can help us beat the probabilities underlying the other items of evidence.

Thus we can make one bet after another. As De Moivre showed, this means multiplying the probabilities.

Dempster's rule.

We make the Cournotian judgement that each item of evidence provides us with no information that can help us beat the probabilities underlying the other items of evidence.

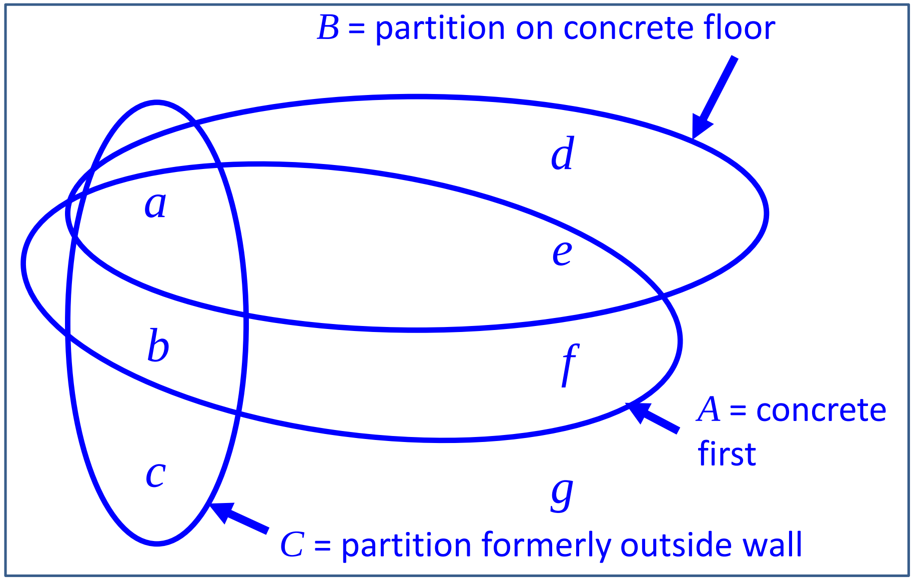
Thus we can make one bet after another. As De Moivre showed, this means multiplying the probabilities.

1. Partition probably original outside wall.
   0.80 for C 0.20 uncommitted
2. Concrete under partition part of concrete floor.
   0.98 for B 0.02 uncommitted
3. Footing under oak floor inadequate.
   0.50 for $D = \{a,b,c,e,f\}$ 0.50 uncommitted
4. Attic stubs suggest roof over concrete section.
   0.80 for $E = \{a,b,e\}$ 0.20 uncommitted
5. Neighbor says office for quarry.
   0.08 for A 0.92 uncommitted

$C\&B\&D\&E\&A=\{a\}$	.8 x .98 x .5 x .8 x .08
$C\&B\&D\&E\&\Theta=\{a\}$	.8 x .98 x .5 x .8 x .92
$C\&B\&D\&\Theta\&A=\{a\}$	.8 x .98 x .5 x .2 x .08
$C\&B\&\Theta\&E\&A=\{a\}$	.8 x .98 x .5 x .8 x .08
$C\&\Theta\&D\&E\&A=\{a\}$	.8 x .02 x .5 x .8 x .08
$\Theta\&B\&D\&E\&A=\{a\}$	.2 x .98 x .5 x .8 x .08
$C\&B\&D\&\Theta\&\Theta=\{a\}$	.8 x .98 x .5 x .2 x .92
$C\&B\&\Theta\&E\&\Theta=\{a\}$	.8 x .98 x .5 x .8 x .92
…	
$C\&B\&\Theta\&\Theta\&\Theta=\{b,c\}$	.8 x .98 x .5 x .2 x .08
…	
$C\&\Theta\&\Theta\&\Theta\&\Theta=C$	.8 x .02 x .5 x .2 x .92
…	
$\Theta\&\Theta\&\Theta\&\Theta\&\Theta=\Theta$	.2 x .02 x .5 x .2 x .92

32 rows all told

Dempster's rule. Collapsing the 32 rows, we obtain 9 focal elements.



1. Partition probably original outside wall.
 0.80 for C 0.20 uncommitted
2. Concrete under partition part of concrete floor.
 0.98 for B 0.02 uncommitted
3. Footing under oak floor inadequate.
 0.50 for $D = \{a,b,c,e,f\}$ 0.50 uncommitted
4. Attic stubs suggest roof over concrete section.
 0.80 for $E = \{a,b,e\}$ 0.20 uncommitted
5. Neighbor says office for quarry.
 0.08 for A 0.92 uncommitted

Focal element	Basic probability number
$\{a\}$*	0.784000
$\{a,e\}$*	0.177968
B	0.018032
$\{a,b\}$*	0.013056
$\{a,b,e\}$*	0.003200
C	0.002944
$\{a,b,c,e,f\}$	0.000368
Θ	0.000368
A*	0.000064

The degree of belief in a subset F is the sum for the focal elements contained in F.

For example, the degree of belief in A is the sum of the numbers with asterisks:

Bel(A) = 0.98.

Baysian assessment

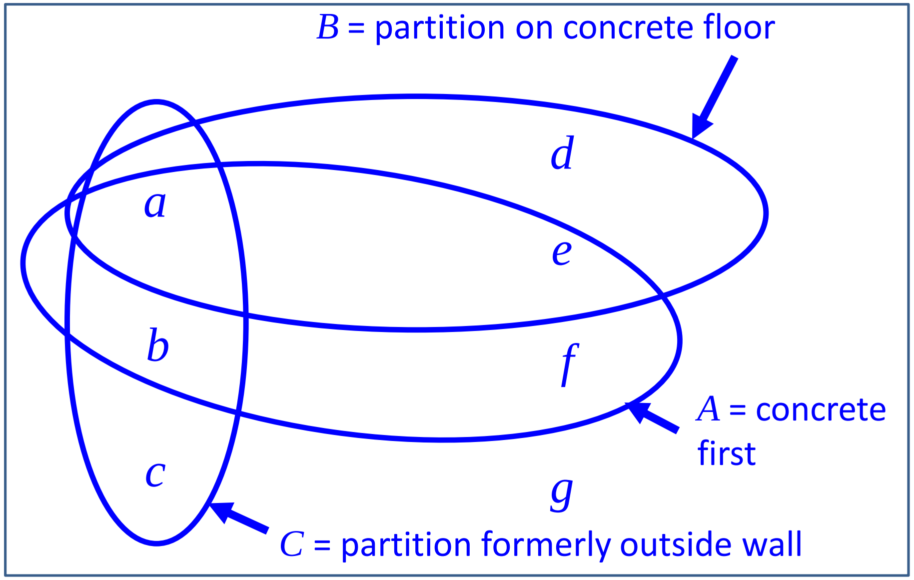
1. Choose a Bayesian design.

Let's use a conditioning design.

1. Divide evidence into "old" and "new".
2. Assess prior probabilities based on the old evidence,
3. Asssess likelihood of new evidence under different hypotheses.

We may assess the likelihood of different items of evidence separately and then multiply, appealing to a Cournotian independence judgement.

1. Partition probably original outside wall.
2. Concrete under partition apparently part of concrete floor.
3. Footing under oak floor inadequate for an outside wall.
4. Attic stubs suggest roof over concrete section.
5. Neighbor says office for quarry.

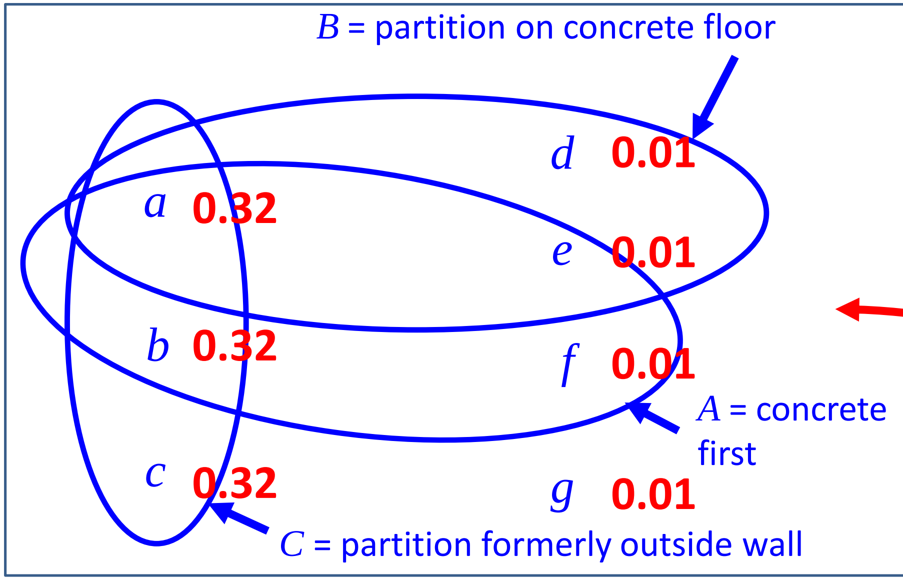


Baysian assessment

2. Divide the evidence into "prior" and "new".

Faute de mieux, I will call the first item my prior evidence.

1. Partition probably original outside wall. Supports C.
2. Concrete under partition apparently part of concrete floor. Supports B.
3. Footing under oak floor inadequate for an outside wall. Supports $\{a,b,c,e,f\}$.
4. Attic stubs suggest roof over concrete section. Supports $\{a,b,e\}$.
5. Neighbor says office for quarry. Supports A.

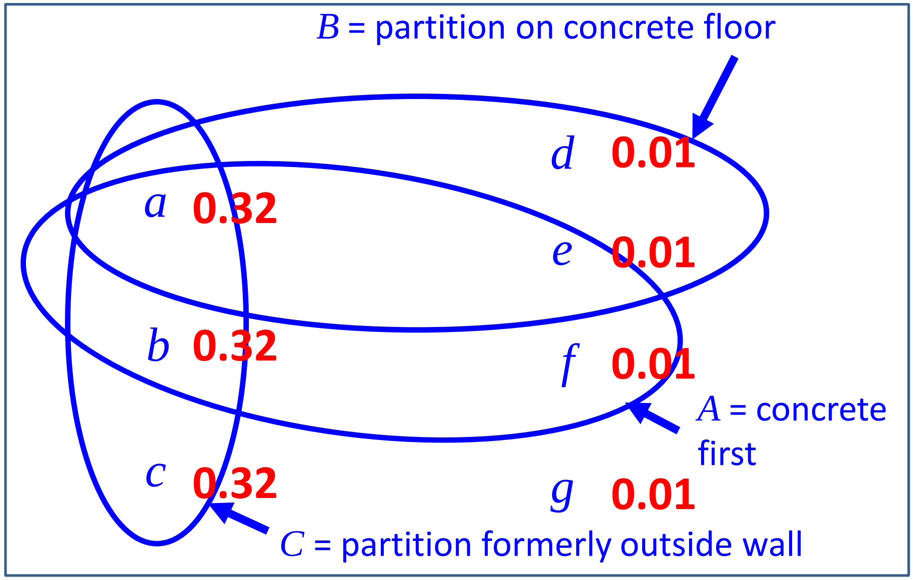


Baysian assessment

3. Assign prior probabilities based on the prior evidence.

Prior evidence.

1. Partition probably original outside wall. Supports C.
2. Concrete under partition apparently part of concrete floor. Supports B.
3. Footing under oak floor inadequate for an outside wall. Supports $\{a,b,c,e,f\}$.
4. Attic stubs suggest roof over concrete section. Supports $\{a,b,e\}$.
5. Neighbor says office for quarry. Supports A.

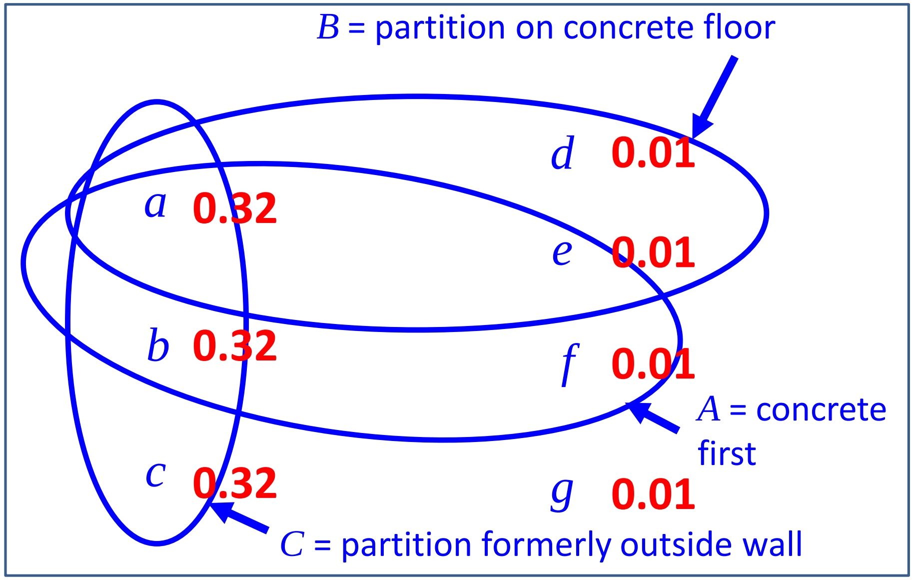


Baysian assessment

4. Make Cournotian judgement that the new items of evidence are conditionally independent.

New evidence.

1. Partition probably original outside wall. Supports C.
2. Concrete under partition apparently part of concrete floor. Supports B.
3. Footing under oak floor inadequate for an outside wall. Supports $\{a,b,c,e,f\}$.
4. Attic stubs suggest roof over concrete section. Supports $\{a,b,e\}$.
5. Neighbor says office for quarry. Supports A.



Baysian assessment

5. Assess relative likelihood of each new item.

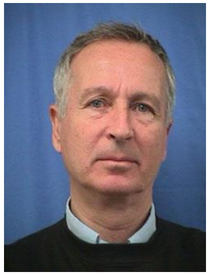
2. Concrete under partition apparently part of concrete floor. Supports B.
3. Footing under oak floor inadequate for an outside wall. Supports $\{a,b,c,e,f\}$.
4. Attic stubs suggest roof over concrete section. Supports $\{a,b,e\}$.
5. Neighbor says office for quarry. Supports A.

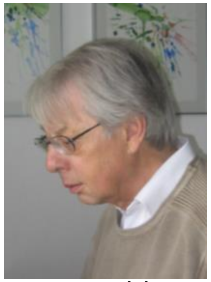
θ	P(θ)	P(2|θ)	P(3|θ)	P(4|θ)	P(5|θ)	P(θ|E)
a	0.32	$100K_2$	$2K_3$	$100K_4$	$1.5K_5$	0.960223
b	0.32	K_2	$2K_3$	$100K_4$	$1.5K_5$	0.009602
c	0.32	K_2	$2K_3$	K_4	K_5	0.000064
d	0.01	$100K_2$	K_3	K_4	K_5	0.000100
e	0.01	$100K_2$	$2K_3$	$100K_4$	$1.5K_5$	0.030007
f	0.01	K_2	$2K_3$	K_4	$1.5K_5$	0.000003
g	0.01	K_2	K_3	K_4	K_5	0.000001

Posterior probability for A = 0.98.

A few contributors to Dempster-Shafer theory…


Philippe Smets
1938-2005


Jean-Yves Jaffray
1939-2009


Jürg Kohlas
Born 1939


Radim Jiroušek
Born 1946


Chuanhai Liu
Born 1959


Thierry Denoeux
Born 1962

…with apologies to all the others

Fundamental idea: transferring belief

- Variable ω with set of possible values Ω.
- Random variable $\mathbf{X}$ with set of possible values $\mathcal{X}$.
- We learn a mapping $\Gamma : \mathcal{X} \to 2^{\Omega}$ with this meaning:

$$\text{If } \mathbf{X} = x, \text{ then } \omega \in \Gamma(x).$$

- For $A \subseteq \Omega$, our belief that $\omega \in A$ is now

$$\mathbb{B}(A) = \mathbb{P}\{x | \Gamma(x) \subseteq A\}.$$

Cournotian judgement of independence: Learning the relationship between $\mathbf{X}$ and ω does not affect our inability to beat the probabilities for $\mathbf{X}$.

Example: The sometimes reliable witness

- Joe is reliable with probability 30%. When he is reliable, what he says is true. Otherwise, it may or may not be true.

  $\mathcal{X} = \{\text{reliable}, \text{not reliable}\}$ $\quad\quad$ $\mathbb{P}(\text{reliable}) = 0.3$ $\quad\quad$ $\mathbb{P}(\text{not reliable}) = 0.7$

- Did Glenn pay his dues for coffee? $\quad\quad$ $\Omega = \{\text{paid}, \text{not paid}\}$

- Joe says "Glenn paid."

  $$\Gamma(\text{reliable}) = \{\text{paid}\} \quad\quad \Gamma(\text{not reliable}) = \{\text{paid}, \text{not paid}\}$$

- New beliefs:

  $$\mathbb{B}(\text{paid}) = 0.3 \quad\quad \mathbb{B}(\text{not paid}) = 0$$

Cournotian judgement of independence: Hearing what Joe said does not affect our inability to beat the probabilities concerning his reliability.

Example: The more or less precise witness

- Bill is absolutely precise with probability 70%, approximate with probability 20%, and unreliable with probability 10%.

$$\mathcal{X} = \{\text{precise}, \text{approximate}, \text{not reliable}\}$$

$$\mathbb{P}(\text{precise}) = 0.7 \qquad \mathbb{P}(\text{approximate}) = 0.2 \qquad \mathbb{P}(\text{not reliable}) = 0.1$$

- What did Glenn pay? $\Omega = \{0, \$1, \$5\}$

- Bill says "Glenn paid $ 5."

$$\Gamma(\text{precise}) = \{\$5\} \qquad \Gamma(\text{approximate}) = \{\$1, \$5\} \qquad \Gamma(\text{not reliable}) = \{0, \$1, \$5\}$$

- New beliefs:

$$\mathbb{B}\{0\} = 0 \qquad \mathbb{B}\{\$1\} = 0 \qquad \mathbb{B}\{\$5\} = 0.7 \qquad \mathbb{B}\{\$1, \$5\} = 0.9$$

Cournotian judgement of independence: Hearing what Bill said does not affect our inability to beat the probabilities concerning his precision.

Conditioning

- Variable ω with set of possible values Ω.

- Random variable $\mathbf{X}$ with set of possible values $\mathcal{X}$.

- We learn a mapping $\Gamma : \mathcal{X} \to 2^{\Omega}$ with this meaning:

$$\text{If } \mathbf{X} = x, \text{ then } \omega \in \Gamma(x).$$

- $\boxed{\Gamma(x) = \emptyset \text{ for some } x \in \mathcal{X}.}$

- For $A \subseteq \Omega$, our belief that $\omega \in A$ is now

$$\mathbb{B}(A) = \frac{\mathbb{P}\{x | \Gamma(x) \subseteq A \ \ \& \ \ \Gamma(x) \neq \emptyset\}}{\mathbb{P}\{x | \Gamma(x) \neq \emptyset\}}.$$

Cournotian judgement of independence: Aside from the impossibility of the x for which $\Gamma(x) = \emptyset$, learning Γ does not affect our inability to beat the probabilities for $\mathbf{X}$.

Example: The witness caught out

- Tom is absolutely precise with probability 70%, approximate with probability 20%, and unreliable with probability 10%.

$$\mathcal{X} = \{\text{precise}, \text{approximate}, \text{not reliable}\}$$

$$\mathbb{P}(\text{precise}) = 0.7 \qquad \mathbb{P}(\text{approximate}) = 0.2 \qquad \mathbb{P}(\text{not reliable}) = 0.1$$

- What did Glenn pay? $\Omega = \{0, \$1, \$5\}$

- Tom says "Glenn paid \$ 10."

$$\Gamma(\text{precise}) = \emptyset \qquad \Gamma(\text{approximate}) = \{\$5\} \qquad \Gamma(\text{not reliable}) = \{0, \$1, \$5\}$$

- New beliefs:

$$\mathbb{B}\{0\} = 0 \qquad \mathbb{B}\{\$1\} = 0 \qquad \mathbb{B}\{\$5\} = 2/3 \qquad \mathbb{B}\{\$1, \$5\} = 2/3$$

Cournotian judgement of independence: Aside ruling out his being absolutely precise, what Tom said does not help us beat the probabilities for his precision.

Independence

$\mathcal{X}_{\text{Bill}} = \{\text{Bill precise}, \text{Bill approximate}, \text{Bill not reliable}\}$

$\mathbb{P}(\text{precise}) = 0.7 \qquad \mathbb{P}(\text{approximate}) = 0.2 \qquad \mathbb{P}(\text{not reliable}) = 0.1$

$\mathcal{X}_{\text{Tom}} = \{\text{Tom precise}, \text{Tom approximate}, \text{Tom not reliable}\}$

$\mathbb{P}(\text{precise}) = 0.7 \qquad \mathbb{P}(\text{approximate}) = 0.2 \qquad \mathbb{P}(\text{not reliable}) = 0.1$

Product measure:

$$\begin{aligned} \mathcal{X}_{\text{Bill \& Tom}} &= \mathcal{X}_{\text{Bill}} \times \mathcal{X}_{\text{Tom}} \\ \mathbb{P}(\text{Bill precise}, \text{Tom precise}) &= 0.7 \times 0.7 = 0.49 \\ \mathbb{P}(\text{Bill precise}, \text{Tom approximate}) &= 0.7 \times 0.2 = 0.14 \\ \text{etc.} & \end{aligned}$$

Cournotian judgements of independence: Learning about the precision of one of the witnesses will not help us beat the probabilities for the other.

Nothing novel here. Dempsterian independence = Cournotian independence.

Example: Independent contradictory witnesses

- Joe and Bill are both reliable with probability 70%.
- Did Glenn pay his dues? $\Omega = \{\text{paid}, \text{not paid}\}$
- Joe says, "Glenn paid." Bill says, "Glenn did not pay."

$$\Gamma_1(\text{Joe reliable}) = \{\text{paid}\} \qquad \Gamma_1(\text{Joe not reliable}) = \{\text{paid}, \text{not paid}\}$$
$$\Gamma_2(\text{Bill reliable}) = \{\text{not paid}\} \qquad \Gamma_2(\text{Bill not reliable}) = \{\text{paid}, \text{not paid}\}$$

- The pair (Joe reliable, Bill reliable), which had probability 0.49, is ruled out.

$$\mathbb{B}(\text{paid}) = \frac{0.21}{0.51} = 0.41 \qquad \mathbb{B}(\text{not paid}) = \frac{0.21}{0.51} = 0.41$$

Cournotian judgement of independence: Aside from learning that they are not both reliable, what Joe and Bill said does not help us beat the probabilities concerning their reliability.

Dempster's rule (independence + conditioning)

- Variable ω with set of possible values Ω.
- Random variables $\mathbf{X}_1$ and $\mathbf{X}_2$ with sets of possible values $\mathcal{X}_1$ and $\mathcal{X}_2$.
- Form the product measure on $\mathcal{X}_1 \times \mathcal{X}_2$.
- We learn mappings $\Gamma_1 : \mathcal{X}_1 \to 2^\Omega$ and $\Gamma_2 : \mathcal{X}_2 \to 2^\Omega$:

  If $\mathbf{X}_1 = x_1$, then $\omega \in \Gamma_1(x_1)$. If $\mathbf{X}_2 = x_2$, then $\omega \in \Gamma_2(x_2)$.
- So if $(\mathbf{X}_1, \mathbf{X}_2) = (x_1, x_2)$, then $\omega \in \Gamma_1(x_1) \cap \Gamma_2(x_2)$.
- Conditioning on what is not ruled out,

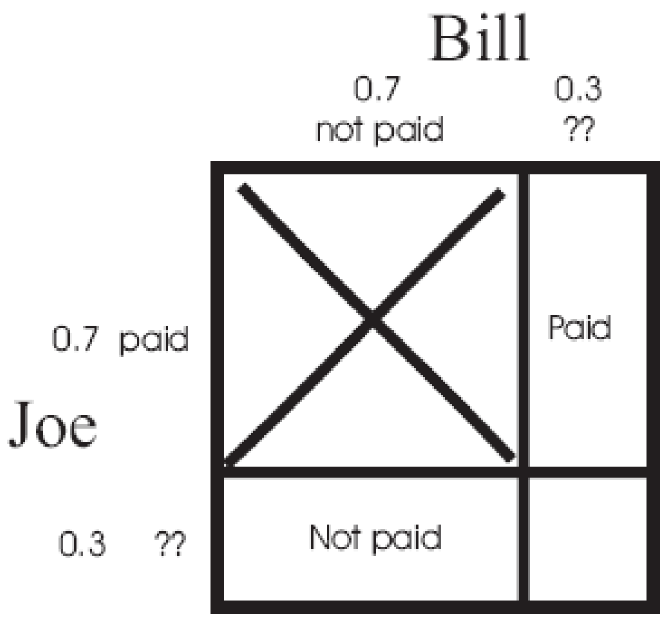
$$\mathbb{B}(A) = \frac{\mathbb{P}\{(x_1, x_2) | \emptyset \neq \Gamma_1(x_1) \cap \Gamma_2(x_2) \subseteq A\}}{\mathbb{P}\{(x_1, x_2) | \emptyset \neq \Gamma_1(x_1) \cap \Gamma_2(x_2)\}}$$

Cournotian judgement of independence: Aside from ruling out some (x_1, x_2), learning the Γ_i does not help us beat the probabilities for $\mathbf{X}_1$ and $\mathbf{X}_2$.

You can suppress the Γs and describe Dempster's rule in terms of the belief functions

Joe: $\mathbb{B}_1\{\text{paid}\} = 0.7$ $\mathbb{B}_1\{\text{not paid}\} = 0$

Bill: $\mathbb{B}_2\{\text{not paid}\} = 0.7$ $\mathbb{B}_2\{\text{paid}\} = 0$



$$\mathbb{B}(\text{paid}) = \frac{0.21}{0.51} = 0.41$$

$$\mathbb{B}(\text{not paid}) = \frac{0.21}{0.51} = 0.41$$

Dempster's rule is unnecessary. It is merely a composition of Cournot operations: formation of product measures, conditioning, transferring belief.

But Dempster's rule is a unifying idea. Each Cournot operation is an example of Dempster combination.

- Forming product measure is Dempster combination.
- Conditioning on A is Demspter combination with a belief function that gives belief one to A.
- Transferring belief is Dempster combination of (1) a belief function on $\mathcal{X} \times \Omega$ that gives probabilities to cylinder sets $\{x\} \times \Omega$ with (2) a belief function that gives probability one to $\{(x, \omega) | \omega \in \Gamma(x)\}$.

Implementing Dempster-Shafer in a specific problem generally involves two related tasks. First, we must sort the uncertainties in the problem into a priori independent items of evidence. Second, we must carry out Dempster's rule computationally.

These two tasks are closely related. Sorting the uncertainties into independent items leads to a structure involving items of evidence that bear on different but related questions, and this structure can be used to make computations feasible.

Suppose, for example, that Betty and Sally testify independently that they heard a burglar enter my house. They might both have mistaken the noise of a dog for that of a burglar. Because of this common uncertainty, I cannot combine their evidence directly by Dempster's rule.

But if I consider explicitly the possibility of a dog's presence, then I can identify three independent items of evidence: my other evidence for or against the presence of a dog, my evidence for Betty's reliability, and my evidence for Sally's reliability. I can combine these by Dempster's rule, and the computations are facilitated by the structure that relates the different questions involved.

Some publications on constructive probability judgment

Papers at www.glennshafer.com on constructive probability judgement:

1. Two theories of probability, *PSA 1978*, Vol. 2, pp. 441-464. Peter D. Asquith and Ian Hacking, eds. Philosophy of Science Association, East Lansing, Michigan. 1981.
2. Constructive probability. *Synthese* **48**:1-60. 1981.
3. Languages and designs for probability judgment (with Amos Tversky). *Cognitive Science* **9**:309-339. 1985. Reprinted in *Decision Making,* edited by David Bell, Howard Raiffa, and Amos Tversky, Cambridge University Press, 1988, pp. 237-265.
4. The combination of evidence. *International Journal of Intelligent Systems* **1**:155-179. 1986.
5. The construction of probability arguments (with discussion). *Boston University Law Review* **66**:799-823. 1986. Reprinted in *Probability and Inference in the Law of Evidence*, edited by Peter Tillers, Kluwer, 1988, pp. 185-204.

Papers 2 and 3 were reprinted as Chapters 9 and 13 of *Classic Works of the Dempster-Shafer Theory of Belief Functions*, edited by Ronald Yager and Liping Liu, Springer, 2008.

Papers at www.probabilityandfinance.com that interpret Bayes and Dempster-Shafer in terms of game-theoretic probability:

- A new understanding of subjective probability and its generalization to lower and upper prevision (with Peter R. Gillett and Richard B. Scherl). Instead of asking whether a person is willing to pay given prices for given risky payoffs, the article asks whether the person believes he can make a lot of money at those prices. *International Journal of Approximate Reasoning* **31**:1-49, 2003.
- A betting interpretation for probabilities and Dempster-Shafer degrees of belief. One way of interpreting numerical degrees of belief is to make the judgement that a strategy for taking advantage of such betting offers will not multiply the capital it risks by a large factor. Applied to ordinary additive probabilities, this can justify updating by conditioning. Applied to Dempster-Shafer degrees of belief, it can justify Dempster's rule of combination. *International Journal of Approximate Reasoning* **52**:127-136, 2011.