Analytics

Decision Analysis with Bayes Theorem

EMBA-J23

Hilary 2023

Decision Analysis Overview

- Decision Trees
- Expected Value Criterion, Risk Attitude and Other Criteria
- Risk Profiles
- Expected Value of Perfect Information
- Bayes Theorem (Today)
- Sensitivity Analysis (Today)

In baiting a mousetrap with cheese,
always leave room for the mouse.
Hector Hugh Munro

Disease Testing

- Accuracy of a test:
  - Of the people who have the disease, **98%** test positive.
  - Of the people who do not have the disease, **99%** test negative.

Sensitivity

$P(+|\text{Disease}) = \mathbf{0.98}$ $\quad$ $P(+|\text{NoDisease}) = \mathbf{0.01}$

$P(-|\text{Disease}) = \mathbf{0.02}$ $\quad$ $P(-|\text{NoDisease}) = \mathbf{0.99}$

Specificity

- You test positive. What is the probability you have the disease?

The optimist proclaims that we live in the best of all possible worlds; and the pessimist fears this is true.
The Silver Stallion, James Branch Cabell, 1926

P(Disease|+)

- Accuracy of the test:

$$P(+ \mid \text{Disease}) = \mathbf{0.98} \qquad P(+ \mid \text{NoDisease}) = \mathbf{0.01}$$

$$P(- \mid \text{Disease}) = \mathbf{0.02} \qquad P(- \mid \text{NoDisease}) = \mathbf{0.99}$$

- You test positive. What is the probability you have the disease?
  - To give the answer P(+|Disease) would be to fall victim to the *Prosecutor's Fallacy*.
  - Instead, we need P(Disease|+).
  - To compute this, use the accuracy AND the **prevalence**: P(Disease)

P(Disease) = **20%**

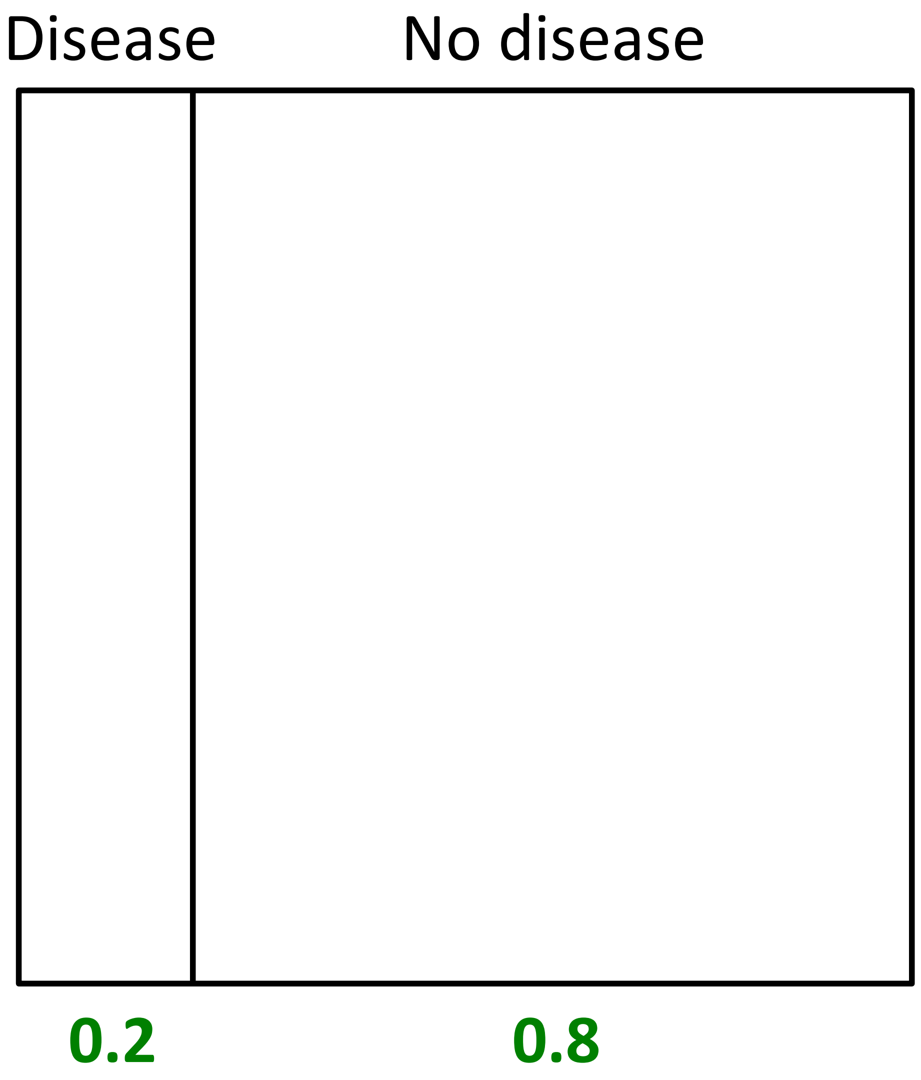
P(+ | Disease) = **0.98**

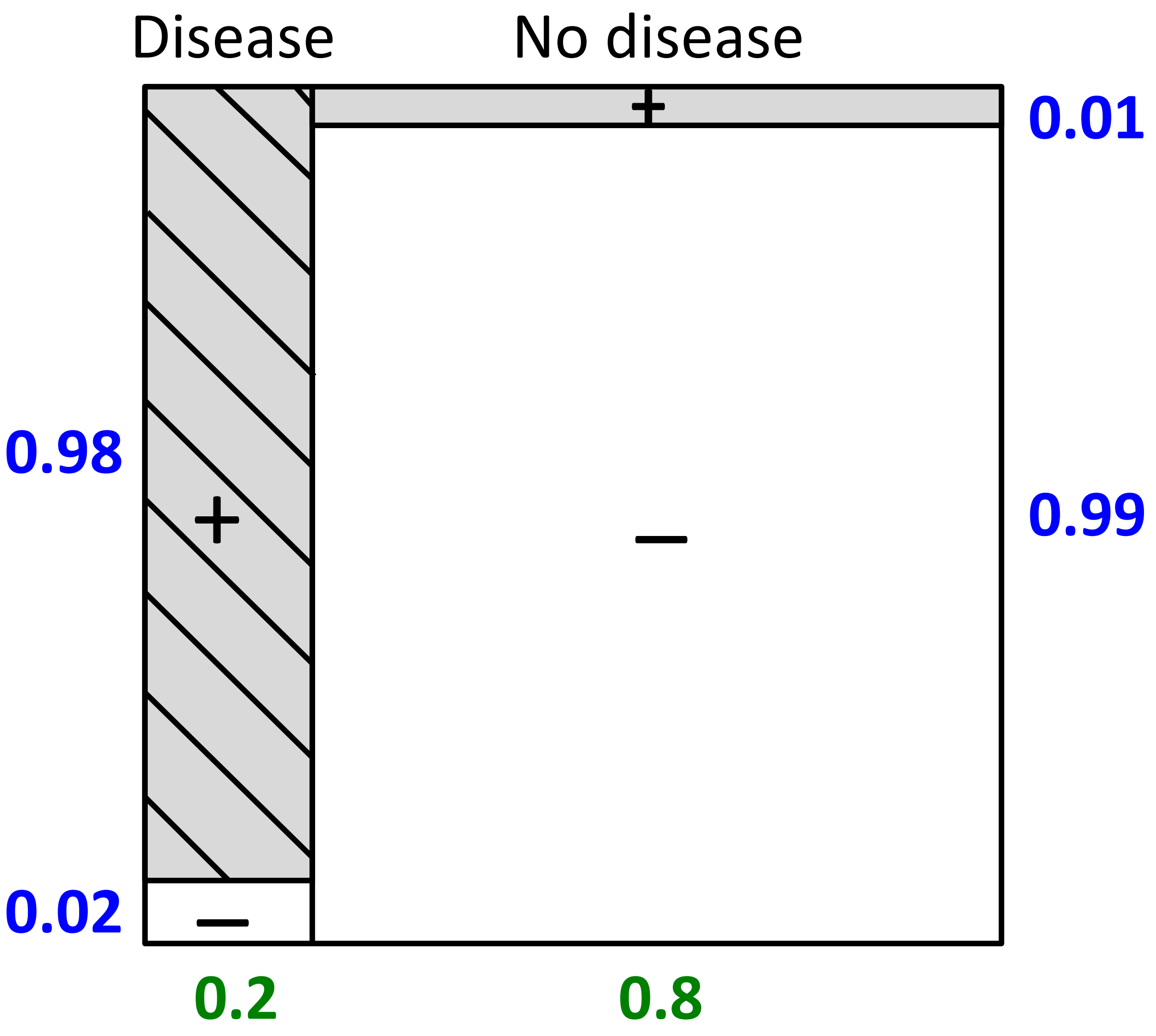
P(− | Disease) = **0.02**

P(+ | NoDisease) = **0.01**

P(− | NoDisease) = **0.99**

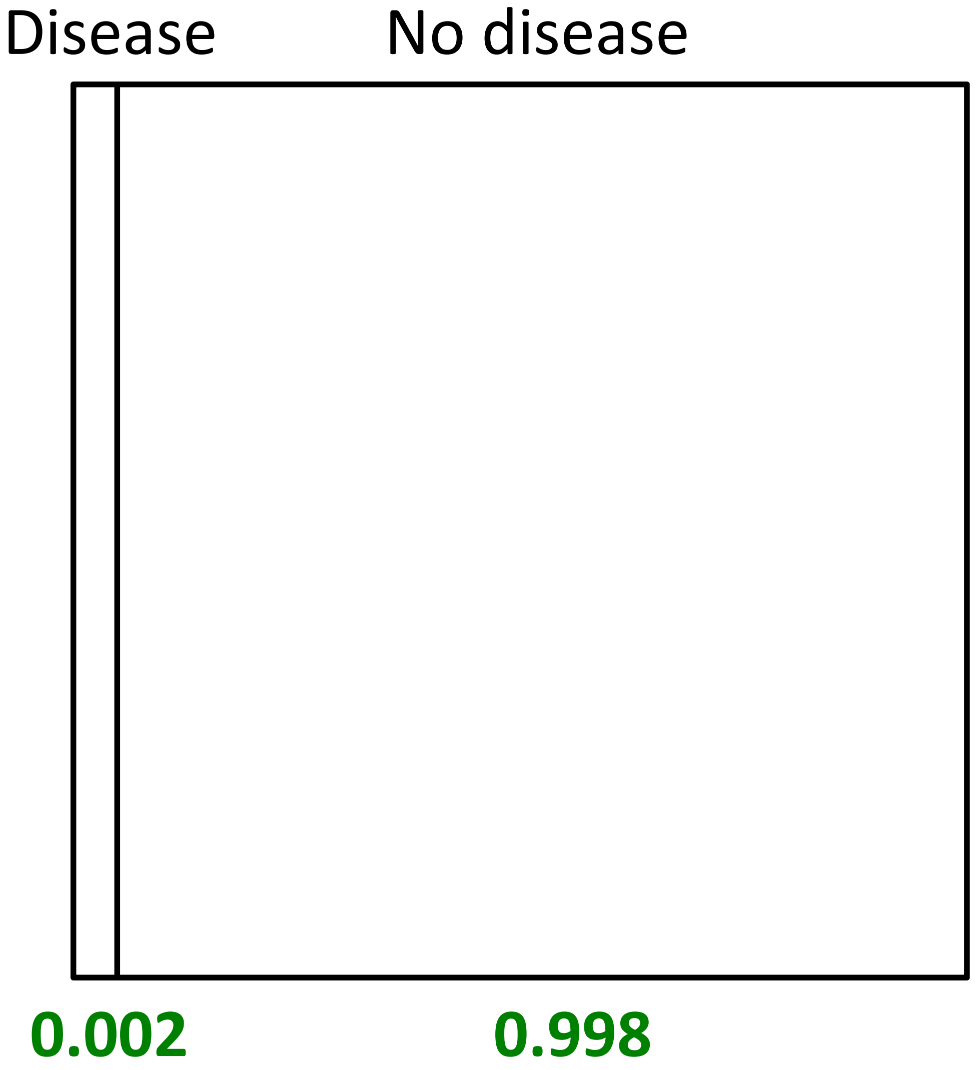
$P(\text{Disease}) = \mathbf{20\%}$

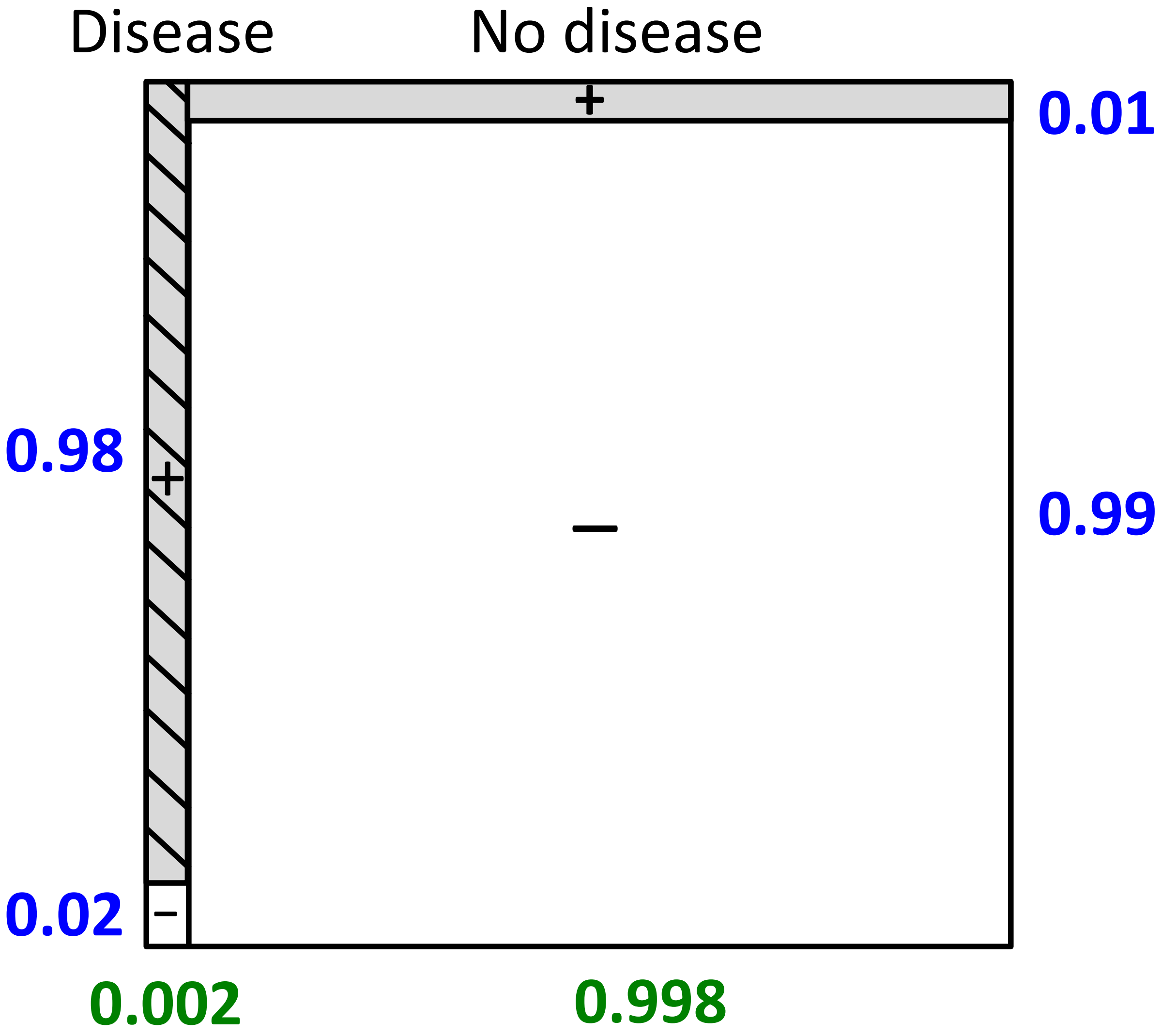




$$P(Disease|+) = \frac{0.98 \times 0.2}{0.98 \times 0.2 + 0.01 \times 0.8} = 0.961$$

$P(\text{Disease}) = 0.2\%$





$$P(Disease \mid +) = \frac{0.98 \times 0.002}{0.98 \times 0.002 + 0.01 \times 0.998} = 0.164$$

Disease Testing Terminology

- Accuracy of the test:

  Sensitivity

  $P(+\mid \text{Disease}) = \mathbf{0.98}$ $\quad$ $P(+\mid \text{NoDisease}) = \mathbf{0.01}$

  $P(-\mid \text{Disease}) = \mathbf{0.02}$ $\quad$ $P(-\mid \text{NoDisease}) = \mathbf{0.99}$

  Specificity

- Prior probability: P(Disease) Prevalence

- On previous 2 slides:
  - Posterior probability: $P(\text{Disease} \mid +)$ Positive Predictive Value
  - Areas within the Venn diagrams correspond to:

	Disease	No Disease
+	**True Positives**	**False Positives**
–	**False Negatives**	**True Negatives**

Bayes Updating of a Prior Probability

Prior probability

Test accuracy
(*likelihood*)

Bayes theorem → **Posterior** probability

e.g.

$P(Disease)$

$P(+|Disease)$

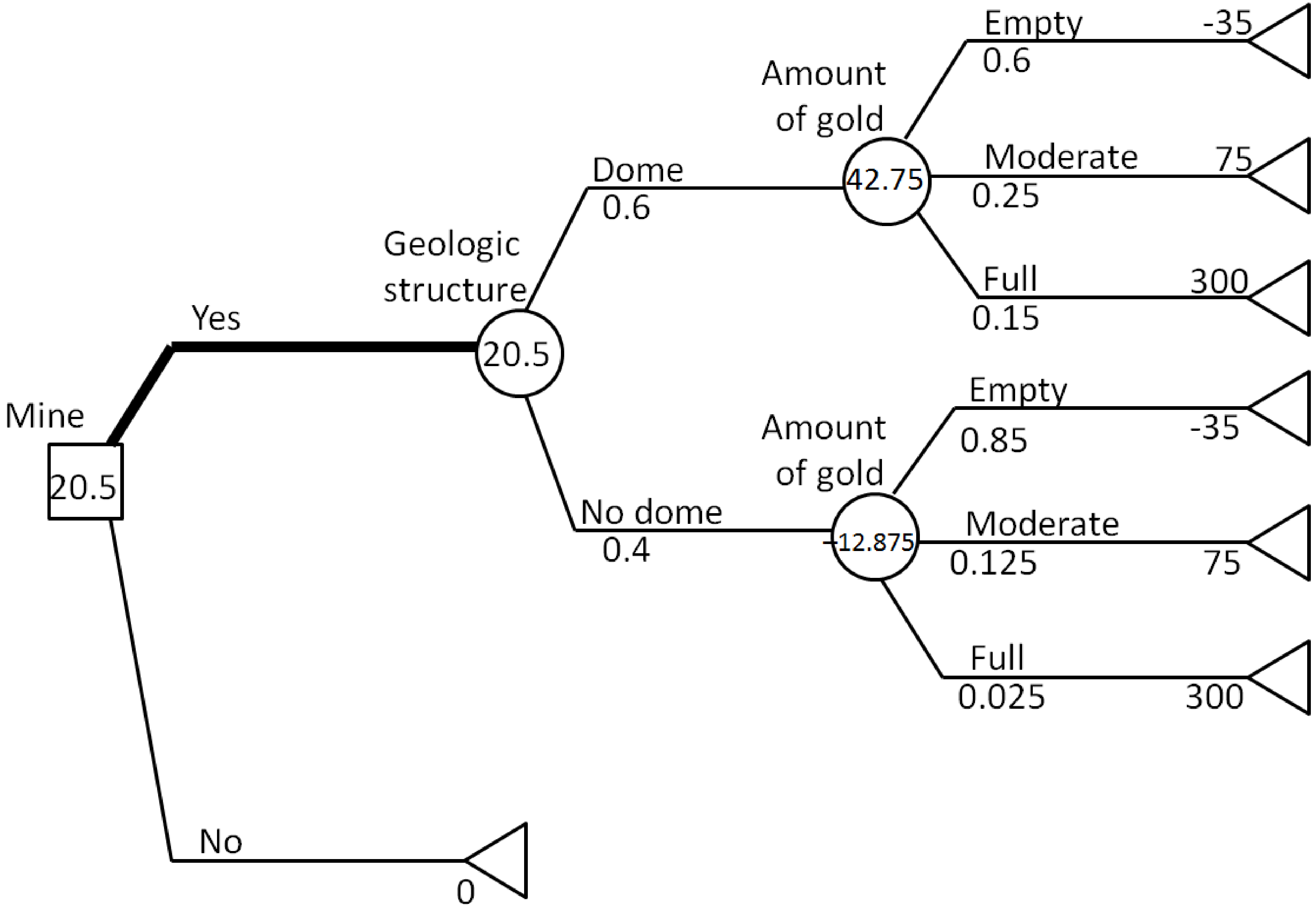
Bayes theorem → $P(Disease|+)$

Bayes theorem: $$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

See Appendix 2

Applications include medical testing, spam filtering, credit rating, forensic science.

Golden Retriever Decision Tree



In the face of uncertainty, the wise decision is the one we can live with serenely - no matter how unlucky it may turn out to be. *Anonymous*

Golden Retriever - Seismic Test Possibility

Charles, the company geologist, suggested that Golden Retriever might consider the possibility of taking a seismic test on the site before mining. This test would cost $5m. The seismic test would give an indication of the existence or non-existence of the dome structure. Charles emphasised that the seismic test was not foolproof. Sometimes intermediate layers of rock reflected the seismic soundings sufficiently to give the impression of a dome when none is there, and sometimes the soundings are misinterpreted to say that no dome exists when in fact it does. Charles gave the following estimates of probabilities describing the reliability of the seismic test:

	Reading from Seismic Test	
Actual State	Positive	Negative
Dome Exists	0.90	0.10
No Dome Exists	0.20	0.80

Each row gives the conditional probability of a seismic test reading for each actual state. For example, given there is actually a dome, the probability of a positive reading is 0.90.

What should Florence advise Golden Retriever to do?

Seismic Test for Dome

- Seismic test is available for $5m.

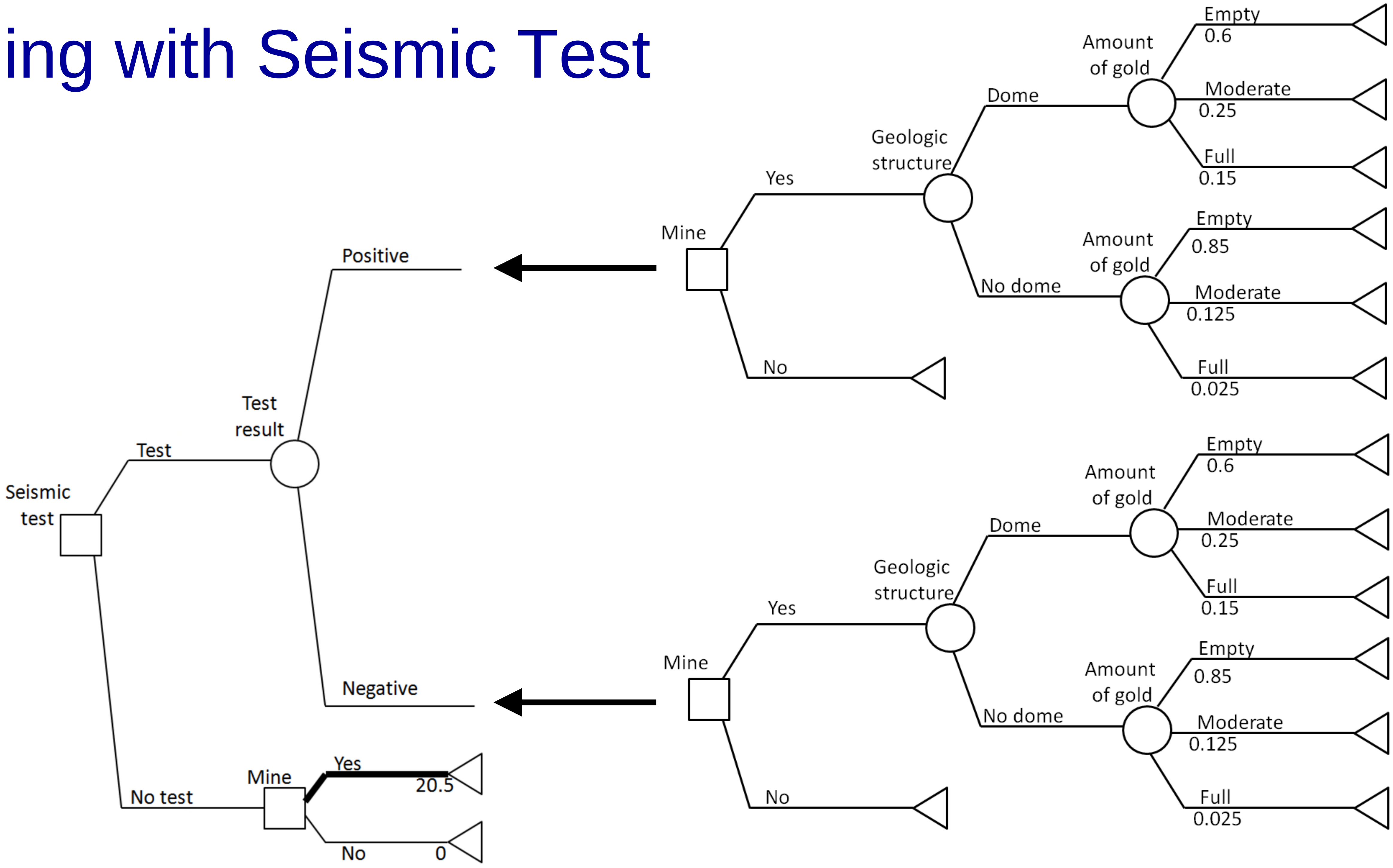
  It gives indication of existence of a dome.

- Test is not perfect.

  Estimates of test accuracy:

	Reading from Seismic Test	
Actual State	Positive	Negative
Dome Exists	0.90	0.10
No Dome Exists	0.20	0.80

Gold Mining with Seismic Test



After drawing the tree, we need probabilities and payoffs.

Venn Diagram

$P(\text{ Dome }) = \mathbf{0.6}$

$P(\text{ NoDome }) = \mathbf{0.4}$

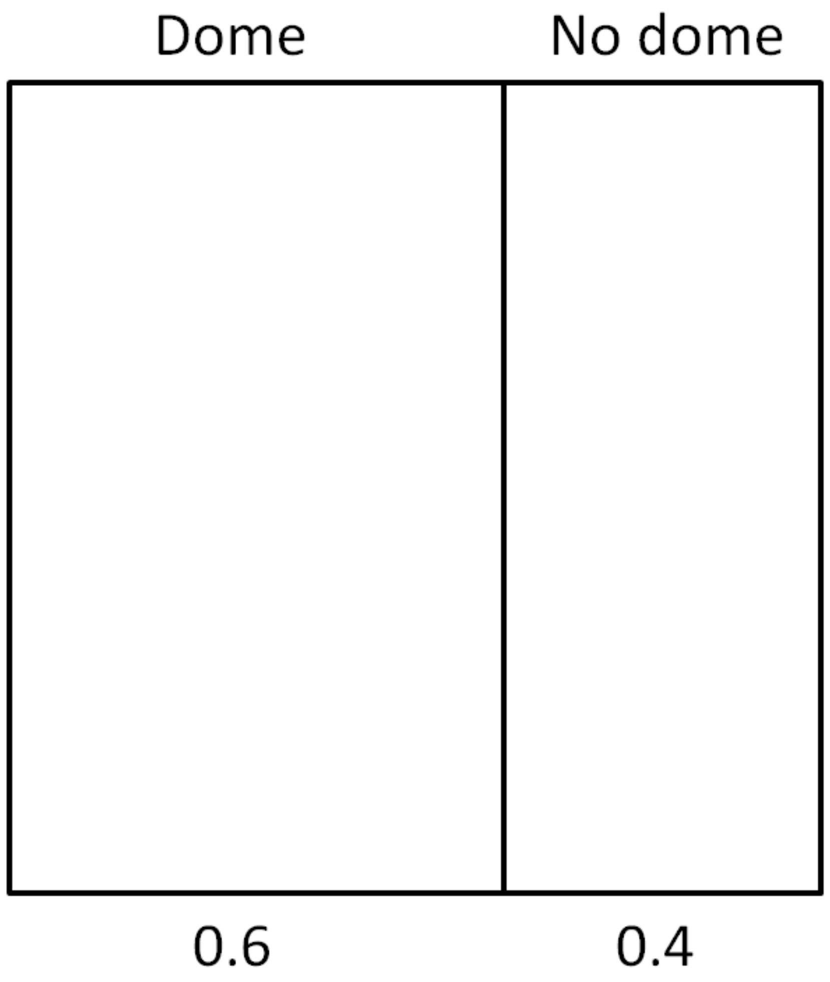
$P(+ \mid \text{Dome}) = \mathbf{0.9}$

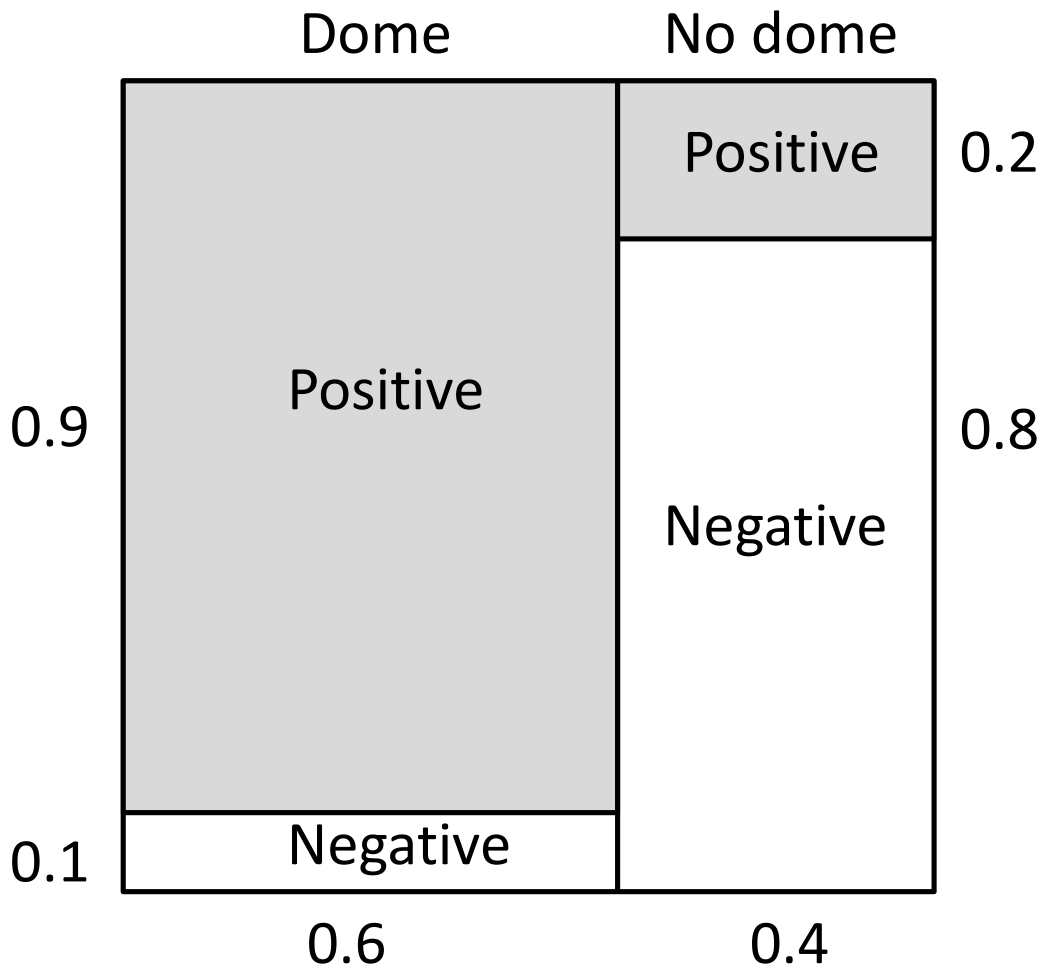
$P(- \mid \text{Dome}) = \mathbf{0.1}$

$P(+ \mid \text{NoDome}) = \mathbf{0.2}$

$P(- \mid \text{NoDome}) = \mathbf{0.8}$

Venn Diagram for Bayes Theorem





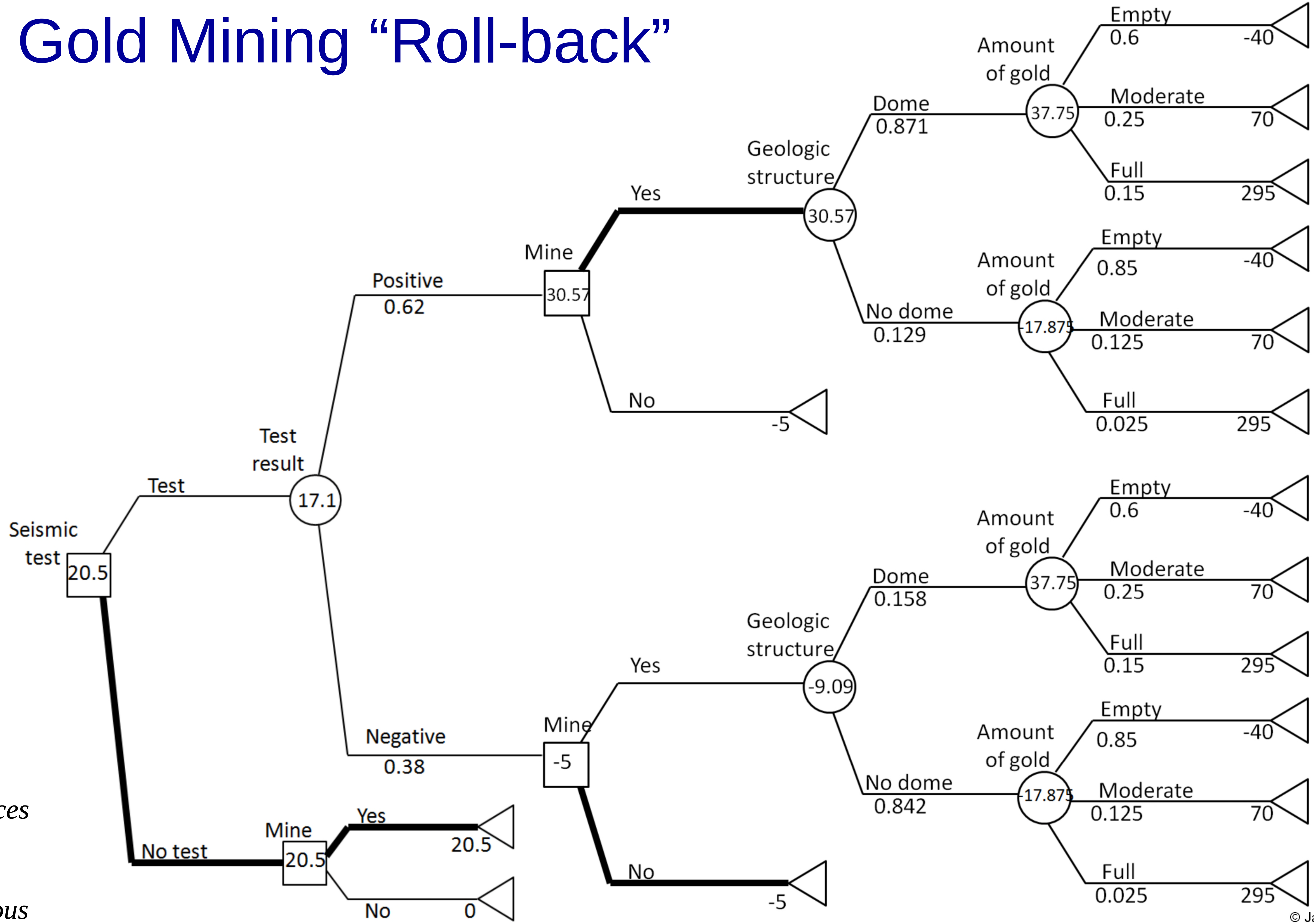
$$P\left(Dome \mid +\right) = \frac{0.9 \times 0.6}{0.9 \times 0.6 + 0.2 \times 0.4} = 0.871$$

$$P\left(Dome \mid +\right) = \frac{P\left(+ \mid Dome\right) P\left(Dome\right)}{P\left(+\right)}$$

If a man will begin with certainties, he shall end in doubts; but if he will be content to begin with doubts he shall end in certainties.

Sir Francis Bacon

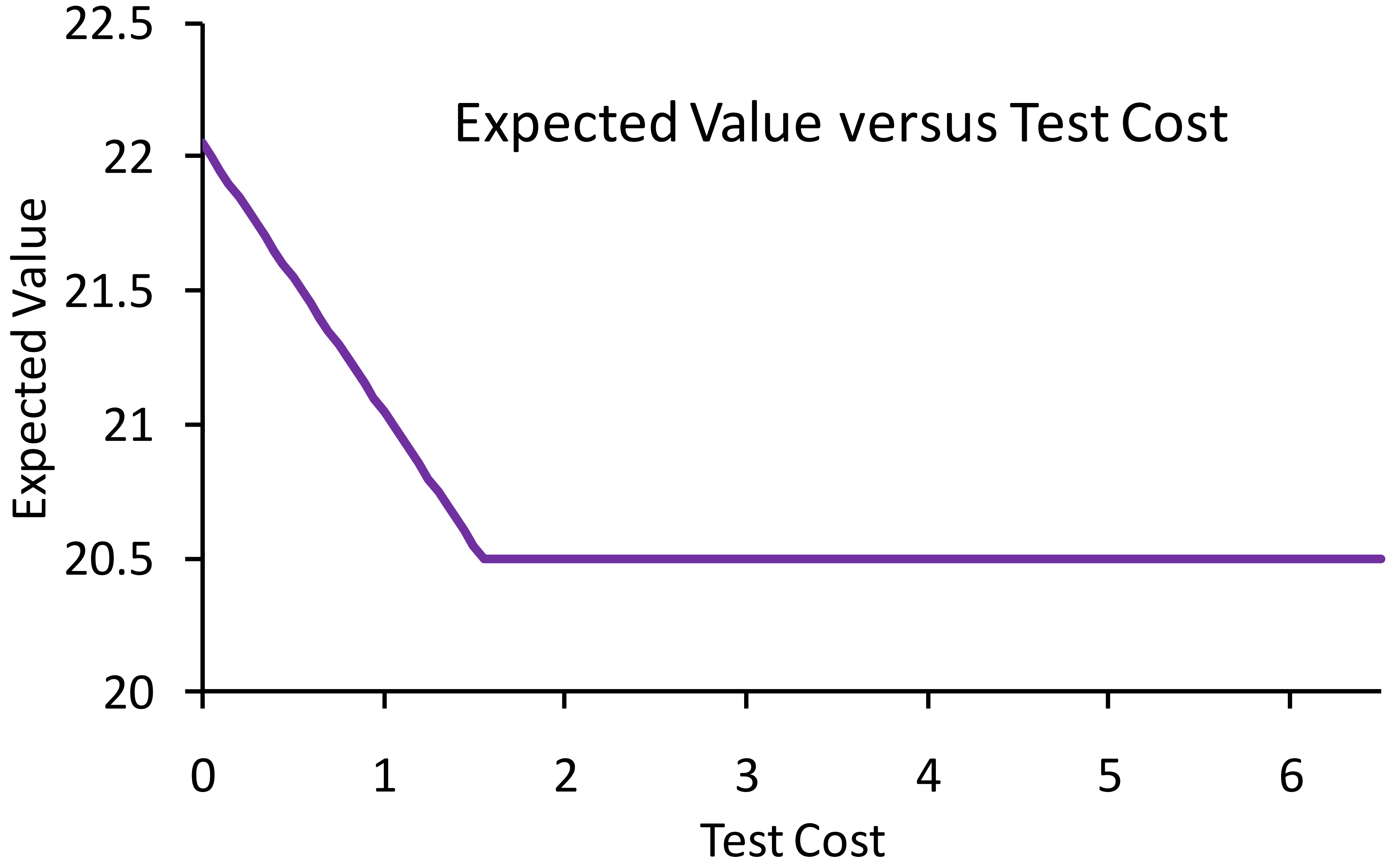
Gold Mining “Roll-back”



Opportunity dances with those on the dance floor.
Anonymous

One-Way Sensitivity Analysis for Test Cost

(see Excel Decision Analysis guides for one- and two-way sensitivity analysis)



Some men see things as they are and ask why;
others dream of things as they might be and ask why not. *Robert Kennedy*

Value of Information

- Tree shows:
  - Not testing has EV = \$20.5m
  - Paying \$5m for seismic test has EV = \$17.1m

- So, to use this test, its price needs to reduce by: 20.5 – 17.1 = \$3.4m
  i.e., reduce price to 5 – 3.4 = \$1.6m.
  This is value of information provided by this test.

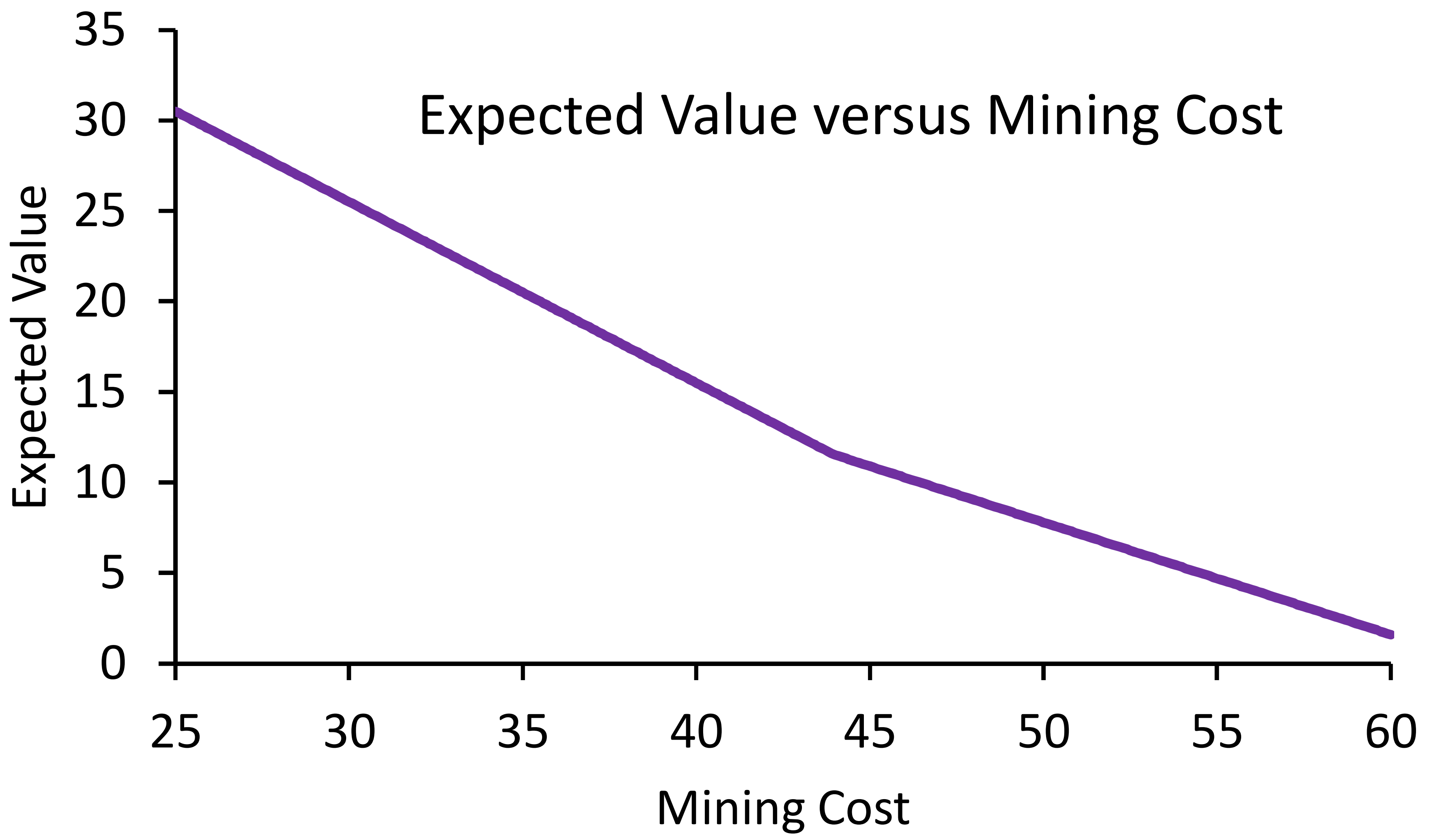
- If offered a ‘better’ test, what is the most worth paying?

See Appendix 3

A good decision is based on knowledge and not on numbers.

Plato

One-Way Sensitivity Analysis for Mining Cost



Decision Analysis Seeks to Improve

- **Insight** through the process of modelling and sensitivity analysis
- **Communication** through visual structure
- **Defensibility** with explicit modelling
- **Consensus** and commitment through shared involvement
- **Awareness** of important strategic drivers

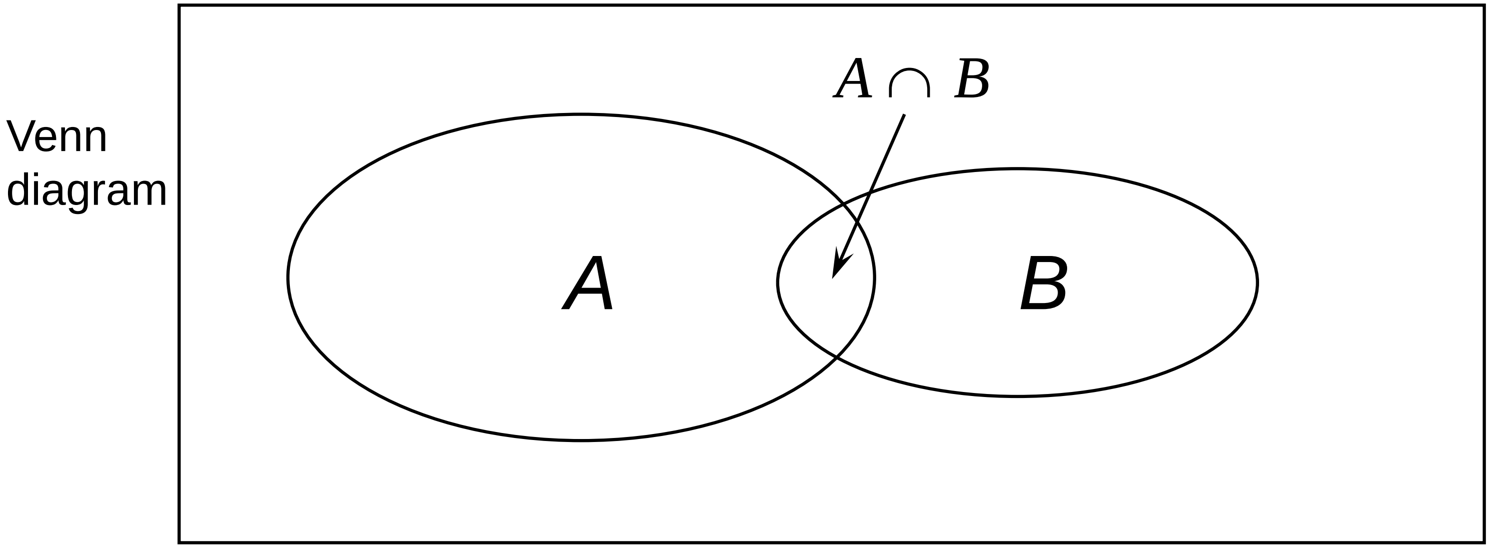
When a statistician passes the airport security check, they discover a bomb in his bag. He explains. "Statistics shows that the probability of a bomb being on an airplane is 1/1000. However, the chance that there are two bombs at one plane is 1/1000000. So, I am much safer..."

Summary

- Decision trees enable a convenient evaluation of decision problems.
- **Expected value criterion** is standard approach to evaluating decision trees.
- Should also consider **risk profiles** and **sensitivity analysis**.
- **Expected value of perfect information** provides an upper bound on value of new information.
- **Bayes theorem** updates prior probabilities.

Appendix 1 - Probability Formulae

- Useful formula: $P(A\,|\,B) = \dfrac{P(A \cap B)}{P(B)}$ i.e. $P(A \cap B) = P(A|B)\,P(B)$



- For example, we have

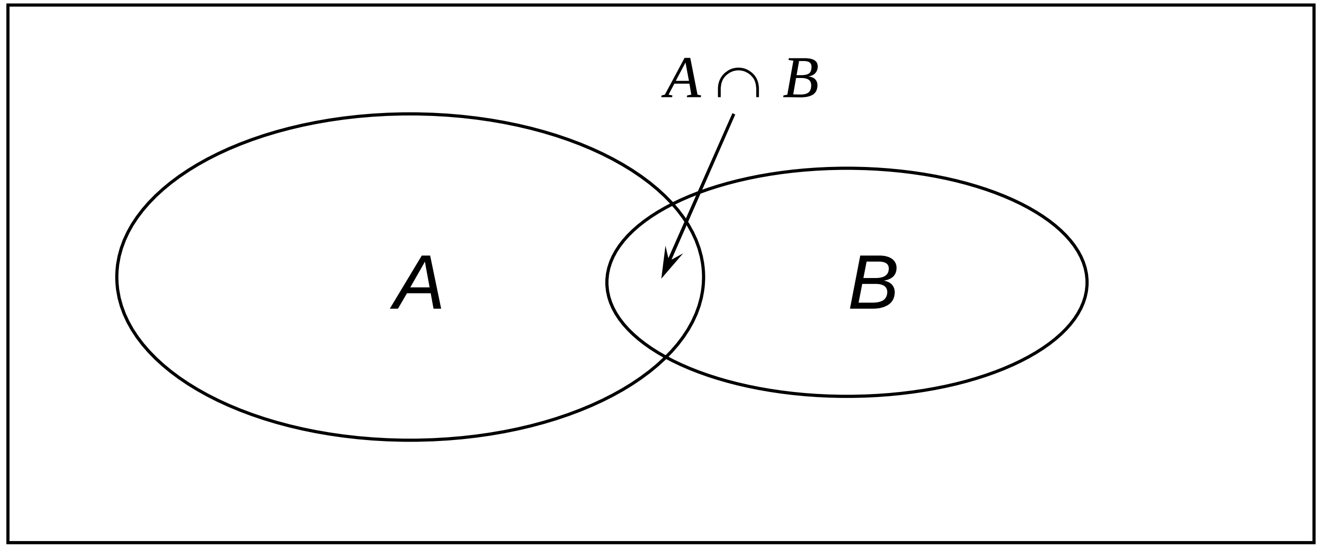
  $P(Empty \mid NoDome) = 0.85$

  $P(NoDome) = 0.4$

and using formula we get:

$$P(Empty \cap NoDome) = P(Empty|NoDome) \times P(NoDome) = 0.85 \times 0.4 = 0.34$$

Appendix 2 – Deriving Bayes Theorem



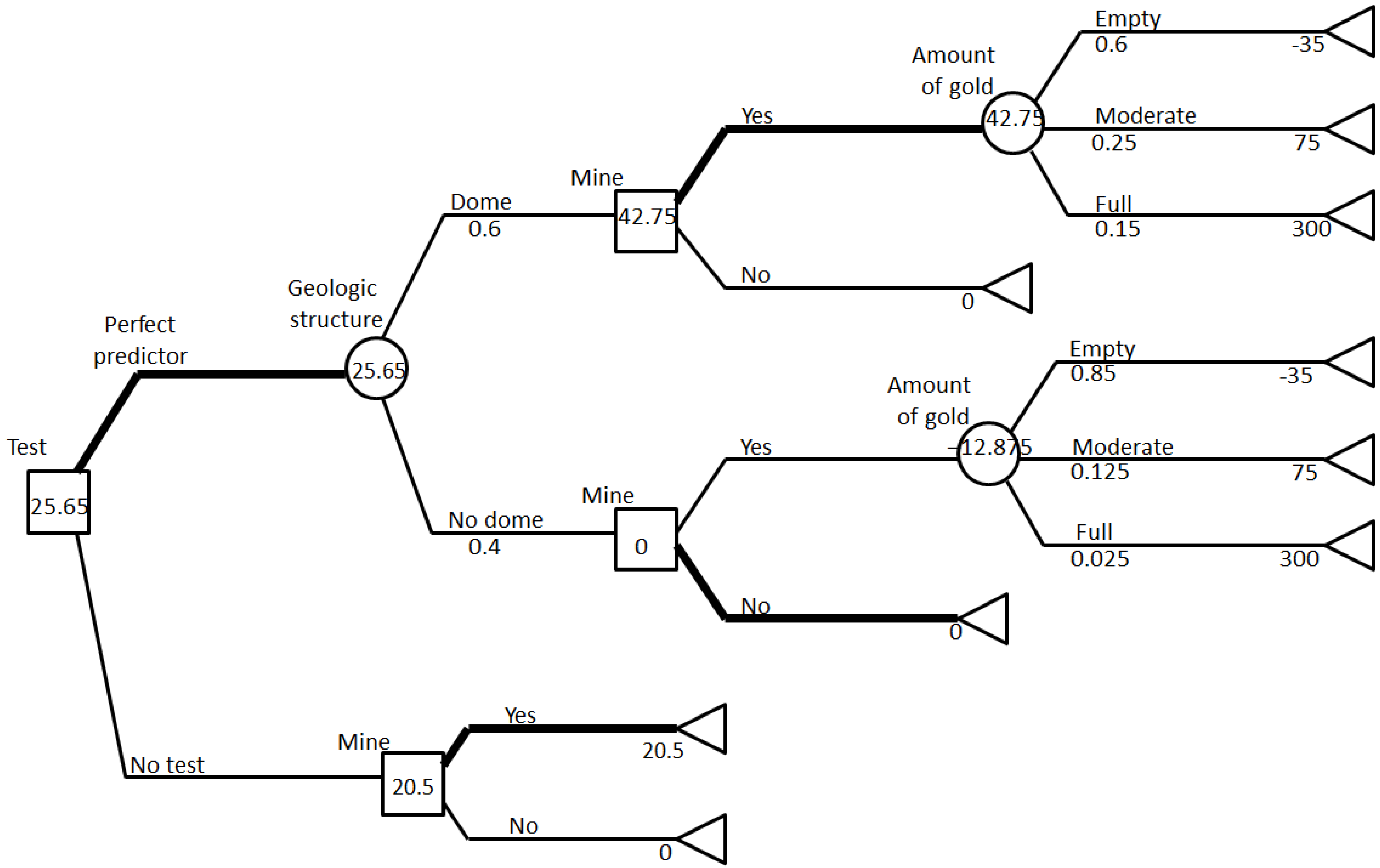
- From Appendix 1, we have: $P(A \cap B) = P(A|B)\, P(B)$

  Swapping A and B in above expression: $P(B \cap A) = P(B|A)\, P(A)$

  Left hand sides of above expressions are equal, so: $P(A|B)\, P(B) = P(B|A)\, P(A)$

  Rearranging gives: $P(A|B) = \dfrac{P(B|A)P(A)}{P(B)}$

Appendix 3 - Perfect Seismic Test for Dome



- With a perfect seismic test for presence of a dome:

  P(+) = P(Dome) and P(-) = P(NoDome)

- Expected value of perfect information is 25.65 – 20.5 = 5.15. Pay no more than this for information regarding existence of dome.