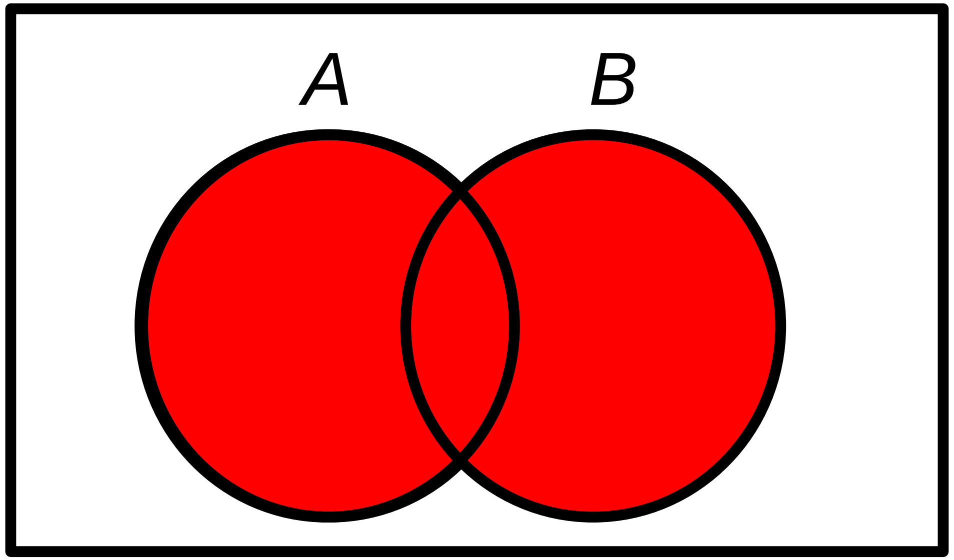
37161 Probability and Random Variables

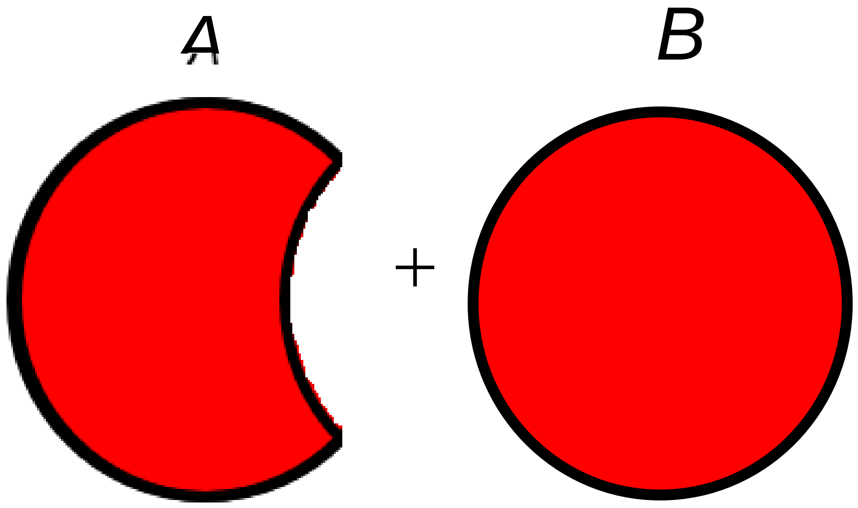
Lecture 2

Probabilities of Unions ("or" Events)



- We have previously seen how to calculate the probabilities of unions of events. That is, for two events, the probability that one event *or* the other (or both) occurs.

- It was easy to see from the Venn diagram for events A and B that $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

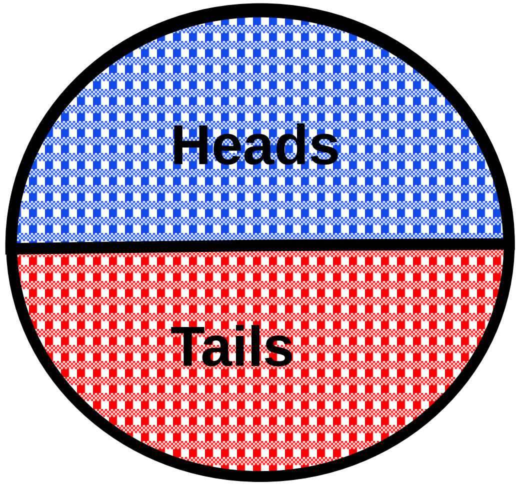


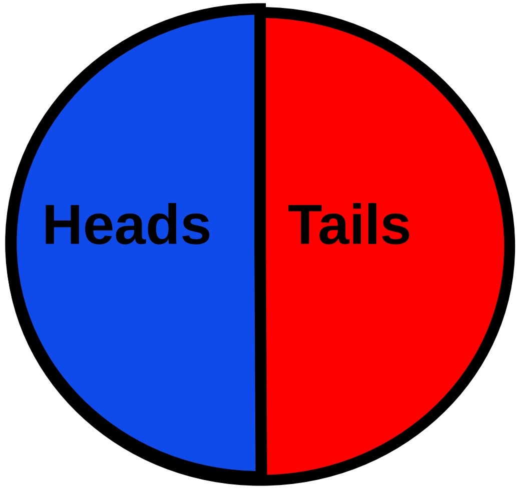
- Because any outcome which was contained in both A and B would otherwise have been counted twice (i.e. in both terms) we require the subtraction of the intersection.

Probabilities of Intersections (“and” Events)

- Consider now the probability of flipping two different fair coins at the same time. What is the probability that both coins land Heads?

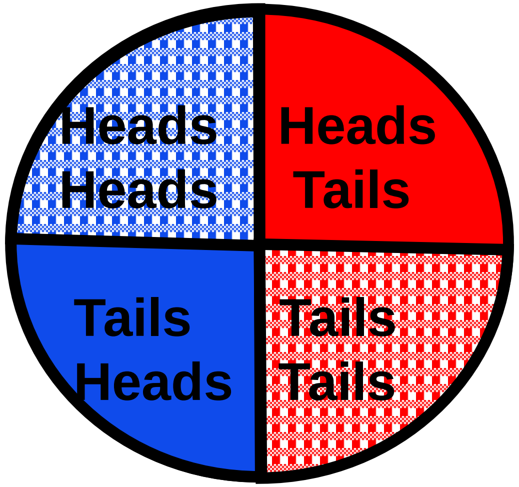
- $\frac{1}{2}$ of the time the first coin will lands Heads. $\frac{1}{2}$ of the time the second coin will land Heads.





- Overall then, the probability that both coins will land Heads is one half of one half of the time.

- That is, if the two coin experiment were to be repeated a large number of times, around $\frac{1}{4}$ of these would give Heads twice.



Probabilities of Intersections (“and” Events)

- Assume that a given population, 50% of people hold a valid driving licence.
- Assume that 50% of this population is female.
- One person is selected at random from the population whereby everyone is equally likely to be chosen.
- Why is it reasonable to conclude that we will select a female with a valid driving licence with probability 25%?
- $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

Probabilities of Intersections ("and" Events)

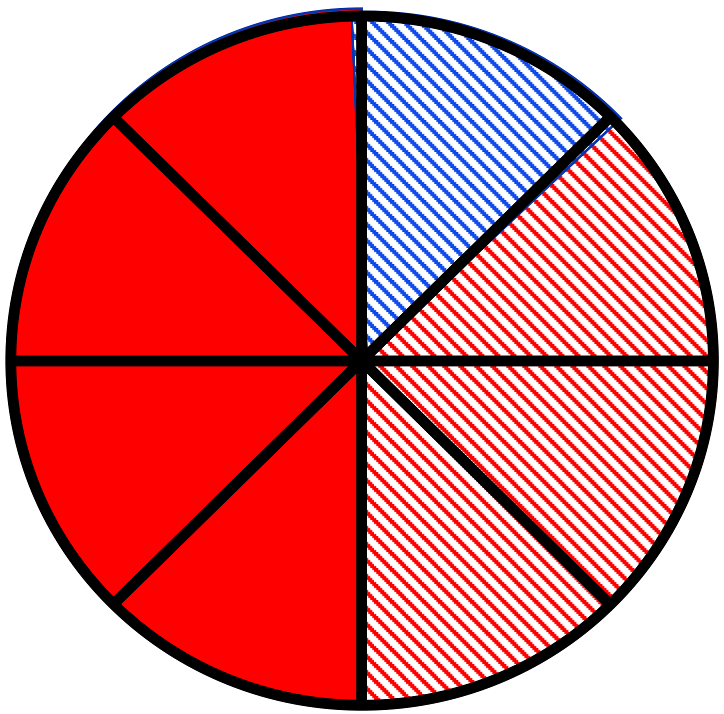
- Assume that a given population, 50% of people hold a valid driving licence.
- Assume that $\frac{1}{8}$ of this population is younger than 10 years old.
- One person is selected at random from the population whereby everyone is equally likely to be chosen.
- What is the probability that we will select a person younger than 10 years old who holds a valid drivers licence?
- It's zero, of course… What's the difference here?

Dependent vs Independent Events

- Considering the coin flipping example, the result of one coin does not affect what the other will give. These outcomes are **independent** of each other.

- The problem in the other case is that the proportion for one event (being an adult or child) does not apply equally to the outcomes relating to the other event (having a driving licence).

- In reality, $\frac{1}{2}$ of people having a driving licence means $\frac{4}{7}$ of adults and no children having a driving licence.



Solid: Licence

Shaded: No licence

Red: 10 years old or older

Blue: Under 10 years old

Conditional Probability

- In cases where knowledge of whether or not one event occurs influences our belief in whether or not another will occur, these events are **dependent**.

- For working out the probabilities of multiple events, we need conditional probabilities – the proportion of times the second event will happen given the first outcome.

- $$\text{Probability of second event given first event} = \frac{\text{Probability of first event and second event}}{\text{Probability of first event}}$$

- $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$ or, alternatively, $P(B) \times P(A|B) = P(A \cap B)$.

- If A and B are independent, then $P(A \cap B) = P(A) \times P(B)$.

Conditional Probability

- Consider selecting one card at random from a standard deck of 52.

- The probability that the card selected is a picture card (Jack, Queen, King) is $\frac{12}{52}$ or $\frac{3}{13}$.

- The probability that the card selected is a Queen is $\frac{4}{52}$ or $\frac{1}{13}$.

- So, the probability that a Queen is selected, given that a picture card is selected is $P(Queen|Picture) = \frac{P(Queen \cap Picture)}{P(Picture)} = \frac{\frac{4}{52}}{\frac{12}{52}} = \frac{1}{3}$

Dependent vs Independent Events

- Consider random experiments whose sample spaces are the sets containing the following observations. Are the possible outcomes of the two experiments dependent or independent?

- A student's height and the course he or she is studying.
- A student's Student ID number and his or her average marks.
- The temperature in Sydney CBD and the day of the week.
- A student's postcode and the colour of his or her shoes.
- The month of a student's birthday and the number of times his or her father has bought a newspaper this week.
- The sexes (male/female) of two students and how many times they obtain Tails when flipping a coin once each.
- A student's high school marks (ATAR) and his or her choice of degree course.

A Famous Mistake

- Consider random experiments whose sample spaces are the sets containing the followiIn 1996, the newborn son of a British woman, Sally Clark, died. His death was attributed to sudden infant death syndrome ("cot death").
- In 1998, she had a second son, who also died in near-identical circumstances.
- She was tried for the murder of both sons.
- The expert paediatrician, Sir Roy Meadows, who gave evidence in court argued that, in the UK, 1 out of every 8500 babies died in such circumstances.
- He argued that for this to happen twice would occur in 1/8500 of 1/8500 instances of siblings i.e. Once in 73 million.
- Given the birthrate, this would happen less than once a century.

Sally Clark

Friday, 26 November, 1999, 15:08 GMT

Mother given life for baby murders

Sally and Stephen Clark said they would fight the conviction

Sunday, 15 July, 2001, 02:58 GMT 03:58 UK

Doubt cast on baby killer case

Sally Clark has always protested her innocence

Evidence used at the trial of a solicitor found guilty of murdering her two baby sons has been called into question in a BBC investigation.

Friday, 16 February, 2001, 13:37 GMT

Cot death gene claim

Babies should be put to sleep on their backs

Genetic defects which may contribute to a baby's risk of cot death have been identified by scientists.

Wednesday, 29 January, 2003, 17:51 GMT

Solicitor accused of killing sons freed

Sally Clark and her husband Stephen outside court

A solicitor jailed for murdering her two baby sons has been cleared by the Court of Appeal.

- What was wrong with the probabilistic argument used for her prosecution?

Dependent Events in the Sally Clark case

- If 1 out of every 8500 seemingly healthy children will die in such circumstances, it is correct that two randomly selected, unrelated babies will both be affected is around 1 in 73 million.

- The prosecution argument rested on the fact that the death of one sibling was independent of the other i.e. There could be no “increased risk” of a later child having inherited any underlying condition which may have contributed to the first tragedy.

Bayes' Theorem

- In some contexts, we might know the conditional probability of one event given another but, in fact, we require the exact opposite.

- For example, with disease screening, we might know how likely a scan is to pick up a disease if the patient has it, but we really want to know the other way round – if the scan has "seen" the disease, how likely is the patient to be sick? (Or similarly, how likely is the "all clear" to be false hope?)

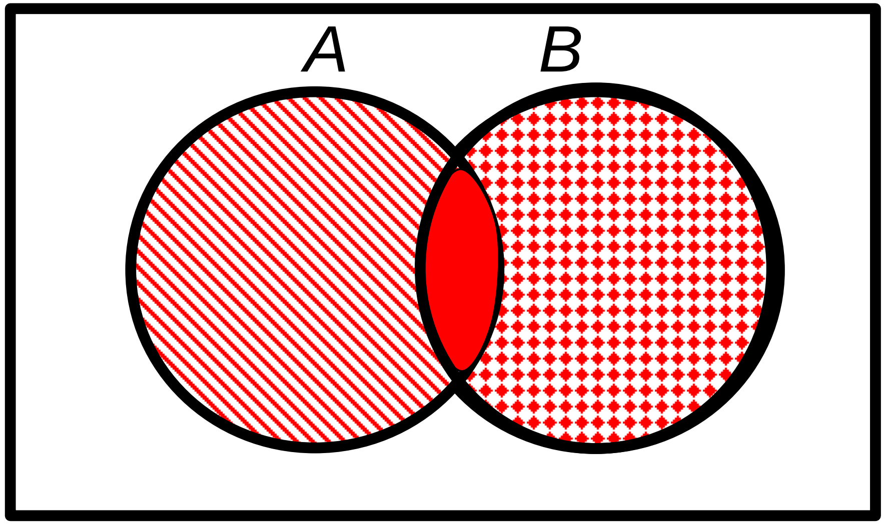
- $P(A \cap B) = P(A|B) \times P(B) = P(B|A) \times P(A)$ so $P(A|B) = \dfrac{P(B|A) \times P(A)}{P(B)}$.

- This is known as **Bayes' Theorem**.

Rev. Thomas Bayes

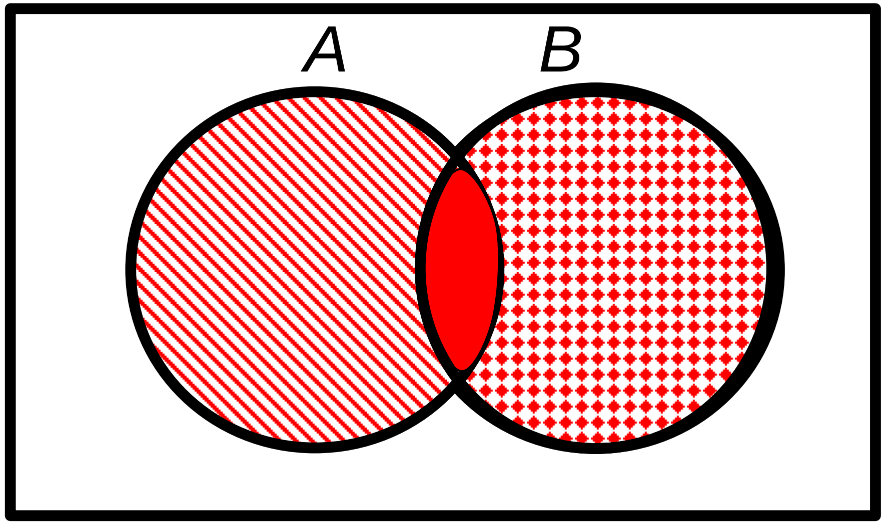
(1701-1761)

Bayes' Theorem: Venn Diagrams



- In the illustrated example, the total shaded area (checked, striped or solid) represents the event $A \cup B$.
- The conditional probability $P(A|B)$ represents the proportion of the event *B* that is also in *A*.
- On the diagram, the event *B* is represented by the union of the checked area and the solid area.
- Similarly, the event that both *A* and *B* occur is represented by the solid area.
- As such, the conditional probability $P(A|B)$ is denoted by the ratio of the solid area to the union of the solid and checked area.

Bayes' Theorem: Venn Diagrams



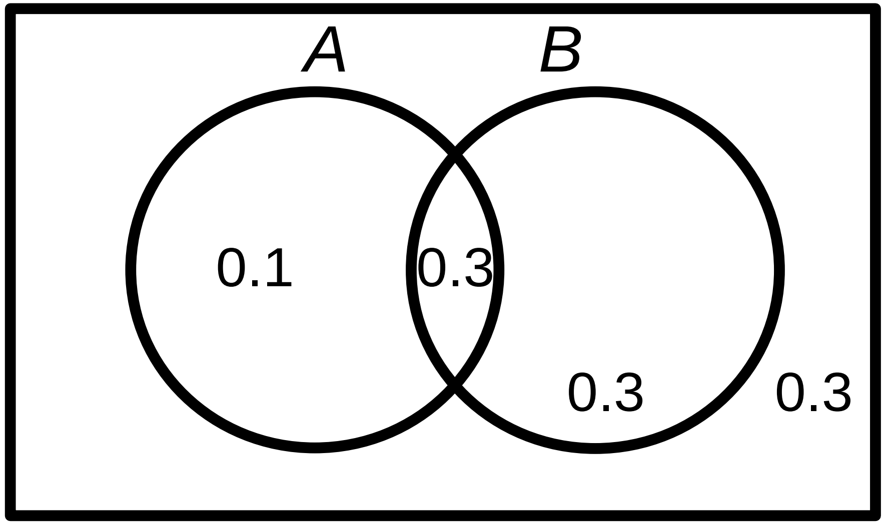
- Note that conditional probabilities are indeed probabilities, i.e. are always between 0 and 1 inclusively.

- The solid area cannot be larger than the total of the solid plus the checked.

$$A \cap B \subseteq B \text{ so } P(A \cap B) \leq P(B) \text{ which ensures } P(A|B) = \frac{P(A \cap B)}{P(B)} \leq 1$$

- Also $0 \leq P(A \cap B)$ which ensures $P(A|B) = \frac{P(A \cap B)}{P(B)} \geq 0$

Venn Diagrams: Example



- In the illustrated example, the values in each area represent the probability of an outcome corresponding to that event.

- This gives

$$P(A) = 0.4$$
$$P(B) = 0.6$$
$$P(A \cap B) = 0.3$$
$$P(A \cup B) = 0.7$$

- The conditional probabilities of A given B is therefore $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.3}{0.6} = \frac{1}{2}$

- Similarly, $P(B|A) = \frac{0.3}{0.4} = \frac{3}{4}$

Conditional Probability: Example

- Consider the following example:
- There is a rare disease which affects 1/40 (2.5%) of patients.

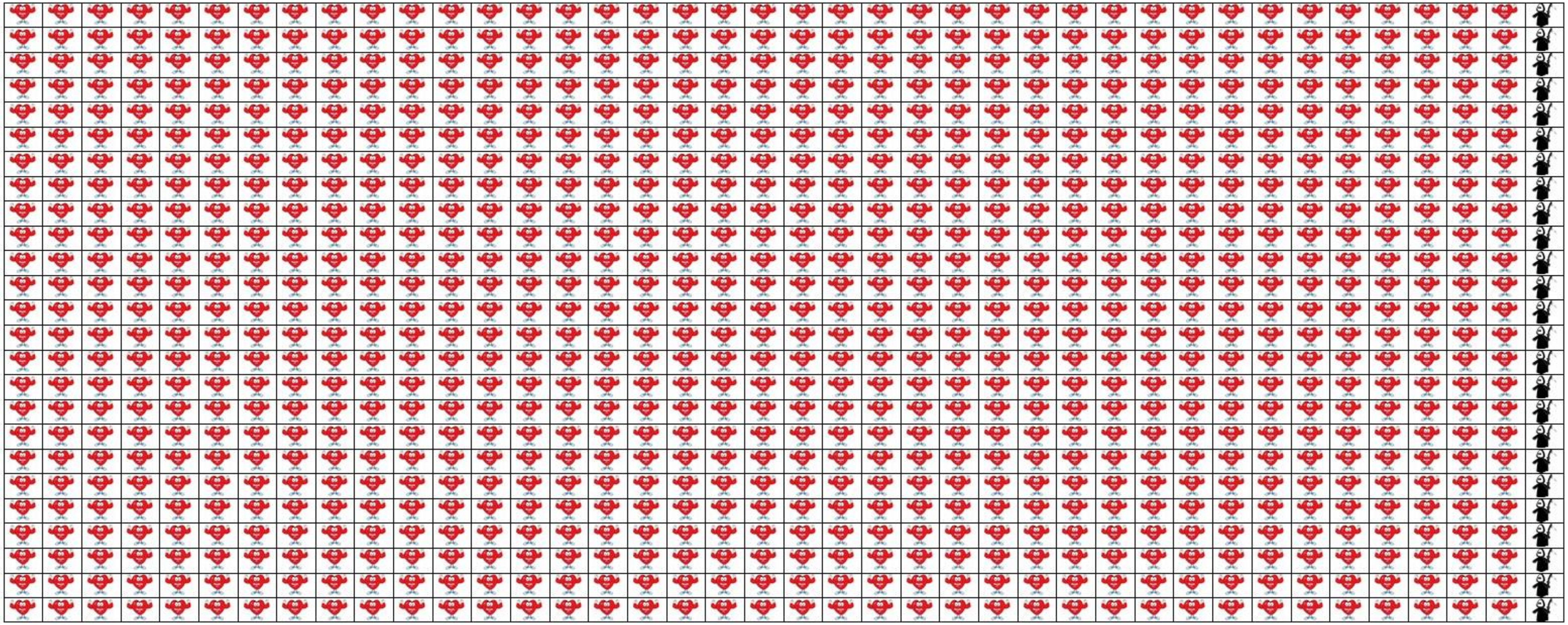
- Patients without the disease are classed as healthy ()
- Patients with the disease are classed as dying ()

- There is a scan to detect the disease.

- If the machine scans a healthy person, then 88% of the time it classifies the patient as "healthy"
- If the machine scans a dying person, then 96% it classifies the patient as "dying".

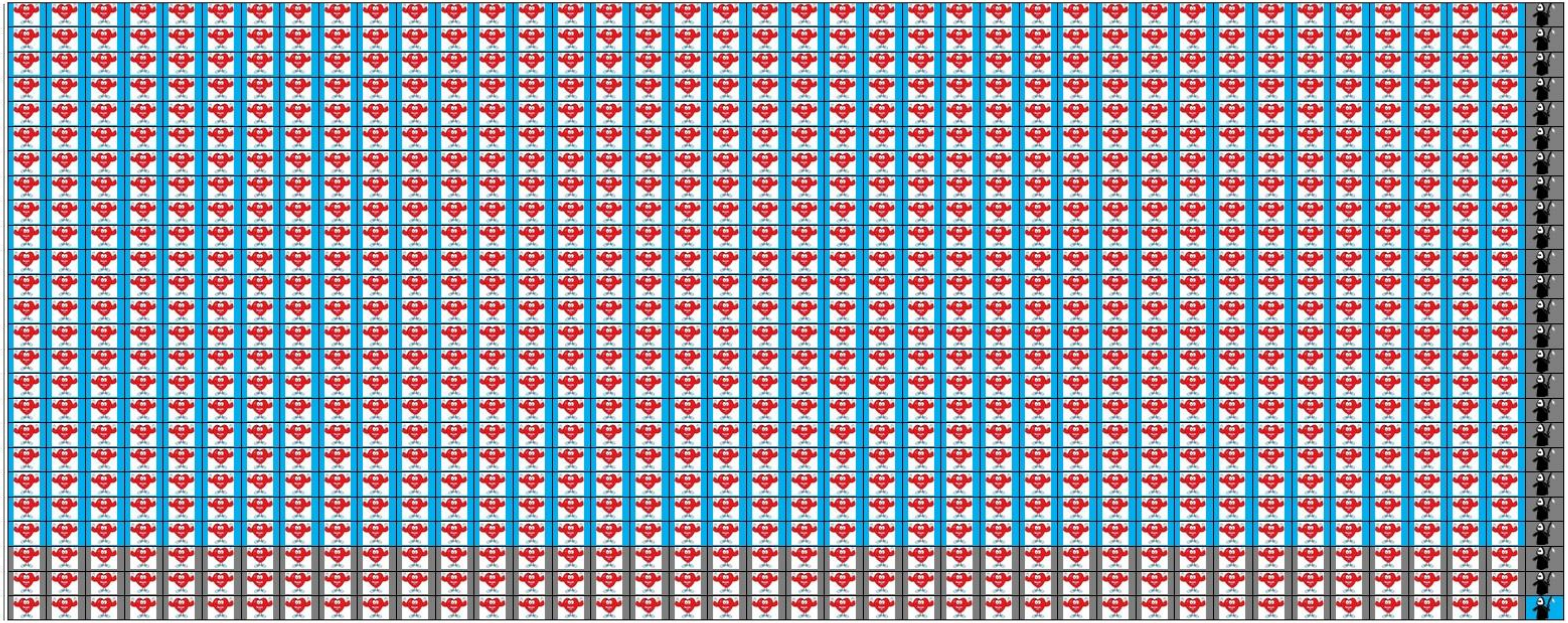
- How worried should a patient be with a "dying" diagnosis?

Example: A Typical 1000 Patients



- On average, out of every 1000 patients, 975 will be healthy and 25 will be dying.

Example: A Typical 1000 Patients



- What would the diagnoses for these be?

Example: A Typical 1000 Patients

- Out of a typical 1000:

- 25 are dying. Of these, 96% (i.e. 24 of the 25) are diagnosed as "dying".
- 975 are healthy. Of these, 12% (i.e. 117 of the 975) are wrongly diagnosed as "dying".

- So, the probability that a person classified as "dying" is really dying is only

$$P(Dying \mid "Dying") = \frac{P(Dying \cap "Dying")}{P("Dying")} = \frac{P(Dying \cap "Dying")}{P(Dying \cap "Dying") + P(Healthy \cap "Dying")}$$

$$= \frac{0.024}{0.024 + 0.117} \approx 17\%$$

- When diseases are relatively rare and false results are relatively common, many people who get bad news might in fact be fine.

A Famous (Worrying) Example

Psychological Review
1995, Vol. 102, No. 4, 684–704

0033-295X/95/$3.00

How to Improve Bayesian Reasoning Without Instruction: Frequency Formats

Gerd Gigerenzer
University of Chicago

Ulrich Hoffrage
Max Planck Institute for Psychological Research

- Given scan flagged up that a patient needed treatment, the chance that she needed treatment was only $\dfrac{0.01\times 0.80}{(0.01\times 0.80)+(0.99\times 0.096)} \approx 7.8\%.$

- Worryingly, just 15% of GPs got the calculation correct…

The Monty Hall Problem

- This famous problem is said to have originated on the American game show *Let's Make a Deal*, hosted by Monty Hall.
- As usually formulated, the game consists of three doors. Behind two of the doors there is a goat, behind the other is a sportscar.
- The game works as follows:

- The contestant chooses any one door
- The gameshow host then opens one other door, selected such that it always contains a goat
- The contestant is then offered two options: to stick with the original choice or to switch to the one other remaining closed door.

- Should the contestant stay with the original choice or change ? Does it make no difference?

The Monty Hall Problem: Solution

- Without loss of generality, assume the first door chosen is door 1.
- If the host then opens door 2 to reveal a goat what is the probability that the car is behind door 3?
- Let C_i be the event that the car is behind door *i*, H_i be the event that the host opens door *i* and F_i be the event that the first chosen door is *i*.
- What we want to know is $P(C_3 | F_1 \cap H_2) = \frac{P(C_3 \cap F_1 \cap H_2)}{P(F_1 \cap H_2)}$,

i.e. the conditional probability that if our first choice is door 1 and then we are shown a goat behind door 2, what is the probability that the car is behind door 3.

The Monty Hall Problem: Solution

- Bayes' Theorem gives $P(C_3|F_1 \cap H_2) = \dfrac{P(C_3 \cap F_1 \cap H_2)}{P(F_1 \cap H_2)} = \dfrac{P(C_3 \cap H_2|F_1)P(F_1)}{P(H_2|F_1)P(F_1)} = \dfrac{P(C_3 \cap H_2|F_1)}{P(H_2|F_1)}$

$$= \frac{P(C_3 \cap H_2|F_1)}{P(C_1 \cap H_2|F_1) + P(C_2 \cap H_2|F_1) + P(C_3 \cap H_2|F_1)}$$

- If door 1 is chosen and the car is behind door 3, the host opens door 2, thus

$$P(C_3 \cap H_2|F_1) = P(C_3) = \frac{1}{3}$$

- If door 1 is chosen and the car is behind door 2, the host CANNOT open door 2, thus

$$P(C_2 \cap H_2|F_1) = 0$$

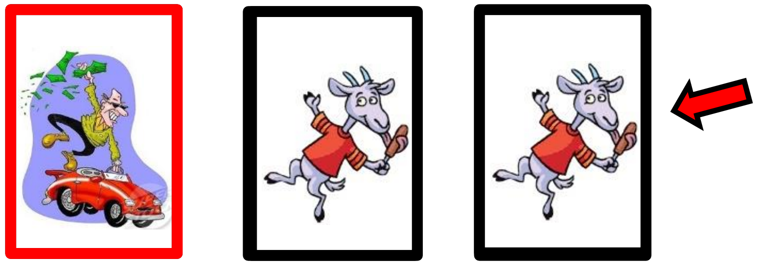
- If door 1 is chosen and the car is behind door 1, the host can open either door 2 or door 3, thus $P(C_1 \cap H_2|F_1) = \dfrac{1}{3} \times \dfrac{1}{2}$

The Monty Hall Problem: Solution

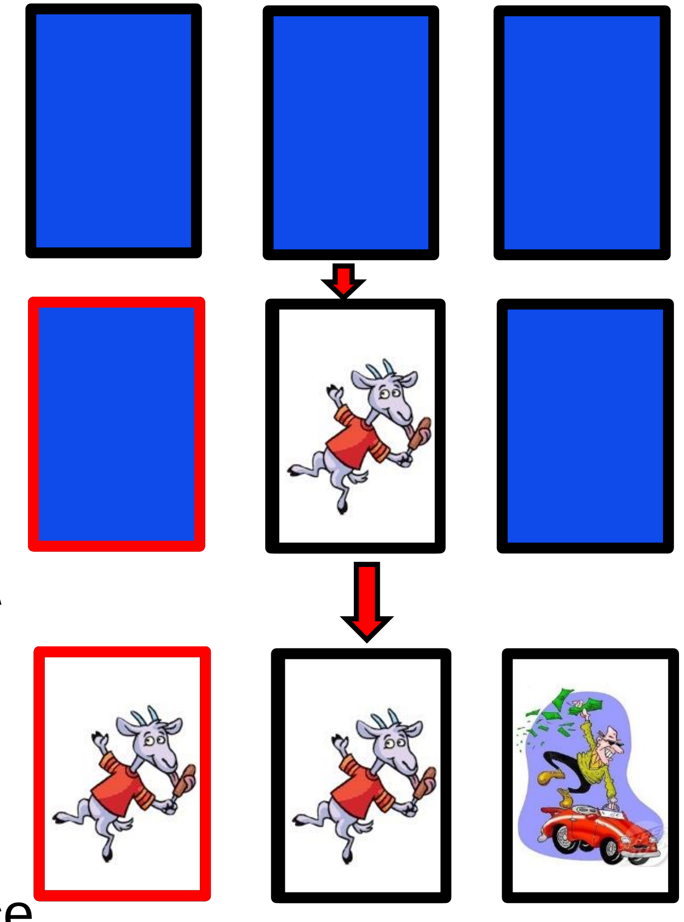
- Substituting all of these in gives:

$$P(C_3 | F_1 \cap H_2) = \frac{P(C_3 \cap H_2 | F_1) P(F_1)}{P(H_2 | F_1) P(F_1)} = \frac{\frac{1}{3}}{\frac{1}{6} + 0 + \frac{1}{3}} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

- In other words, the contestant is **twice** as likely to win the car by switching after the initial choice.



Probability 1/3



Probability 2/3

- The only way the contestant wins by sticking with the original choice is if that choice was correct. This occurs with probability 1/3.