Conditional probability and independence

Samy Tindel

Purdue University

Probability - MA 416

Mostly taken from *A first course in probability* by S. Ross

Outline

Outline

Global objective

Aim: Introduce conditional probability, whose interest is twofold

1. Quantify the effect of a prior information on probabilities
2. If no prior information is available, then independence
   $\hookrightarrow$ simplification in probability computations

Outline

Example of conditioning

Dice tossing: We consider the following situation

- We throw 2 dice
- We look for $\mathbf{P}(\text{sum of 2 faces is 9})$

Without prior information:

$$\mathbf{P}\left(\text{sum of 2 faces is 9}\right) = \frac{1}{9}$$

Changes with additional information:

If we know that first face is $= 4$, then

$\hookrightarrow$ how does it affect $\mathbf{P}\left(\text{sum of 2 faces is 9}\right)$?

Example of conditioning

Probability with additional information: If first face is $= 4$, then

- Only 6 possible results:

$$(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)$$

- Among them, only $(4,5)$ gives sum $= 9$
- Probability of having sum $= 9$ becomes

$$p = \frac{1}{6}$$

Conclusion:

We need to formalize this type of computation

General definition

Definition 1.

Let

- $\mathbf{P}$ a probability on a sample space S
- E, F two events, such that $\mathbf{P}(F) > 0$

Then

$$\mathbf{P}(E|\, F) = \frac{\mathbf{P}(E\, F)}{\mathbf{P}(F)}$$

Example: examination (1)

Situation:
Student taking a one hour exam

Hypothesis: For $x \in [0, 1]$ we have

$$\mathbf{P}(L_x) = \frac{x}{2}, \tag{1}$$

where the event L_x is defined by

$$L_x = \{\text{student finishes the exam in less than } x \text{ hour}\}$$

Question: Given that the student is still working after .75h
$\hookrightarrow$ Find probability that the full hour is used

Example: examination (2)

Model: We wish to find

$$\mathbf{P}\left(L_1^c \middle| L_{.75}^c\right)$$

Computation: We have

$$\begin{aligned}
\mathbf{P}\left(L_1^c \middle| L_{.75}^c\right) &= \frac{\mathbf{P}\left(L_1^c L_{.75}^c\right)}{\mathbf{P}\left(L_{.75}^c\right)} \\
&= \frac{\mathbf{P}\left(L_1^c\right)}{\mathbf{P}\left(L_{.75}^c\right)} \\
&= \frac{1-\mathbf{P}\left(L_1\right)}{1-\mathbf{P}\left(L_{.75}\right)}
\end{aligned}$$

Conclusion: Applying (1) we get

$$\mathbf{P}\left(L_1^c \middle| L_{.75}^c\right) = .8$$

Simplification for uniform probabilities

General situation: We assume

- $S = \{s_1, \ldots, s_N\}$ finite.
- $\mathbf{P}(\{s_i\}) = \frac{1}{N}$ for all $1 \leq i \leq N$

Alert:
This is an important but very particular case of probability space

Conditional probabilities in this case:
Reduced sample space, i.e

Conditional on F, all outcomes in F are equally likely

Example: family distribution (1)

Situation:

The Popescu family has 10 kids

Questions:

1. If we know that 9 kids are girls
   $\hookrightarrow$ find the probability that all 10 kids are girls

Example: family distribution (1)

Situation:

The Popescu family has 10 kids

Questions:

1. If we know that 9 kids are girls
   $\hookrightarrow$ find the probability that all 10 kids are girls
2. If we know that the first 9 kids are girls
   $\hookrightarrow$ find the probability that all 10 kids are girls

Example: family distribution (2)

Model:

- $S = \{G, B\}^{10}$
- Uniform probability: for all $s \in S$,

$$\mathbf{P}(\{s\}) = \frac{1}{2^{10}} = \frac{1}{1024}$$

Example: family distribution (3)

First conditioning: We take

$$F_1 = \{(G,\ldots,G);\,(G,\ldots,G,B);\,(G,\ldots,G,B,G);\cdots;\,(B,G,\ldots,G)\}$$

Reduced sample space:
Each outcome in F_1 has probability $\frac{1}{11}$

Conditional probability:

$$\mathbf{P}\left(\{(G,\ldots,G)\}\,|\,F_1\right) = \frac{1}{11}$$

Example: family distribution (4)

Second conditioning: We take

$$F_2 = \{(G, \ldots, G); (G, \ldots, G, B)\}$$

Reduced sample space:
Each outcome in F_2 has probability $\frac{1}{2}$

Conditional probability:

$$\mathbf{P}\left(\{(G, \ldots, G)\} \middle| F_2\right) = \frac{1}{2}$$

Example: bridge game (1)

Bridge game:

- 4 players, E, W, N, S
- 52 cards dealt out equally to players

Conditioning: We condition on the set

$$F = \{N + S \text{ have a total of 8 spades}\}$$

Question: Conditioned on F,
Probability that E has 3 of the remaining 5 spades

Example: bridge game (2)

Model: We take

$$S = \{\text{Divisions of 52 cards in 4 groups}\}$$

and we have

- Uniform probability on S
- $|S| = \binom{52}{13,13,13,13} \simeq 5.36\ 10^{28}$

Reduced sample space: Conditioned on F,

$$\tilde{S} = \{\text{Combinations of 13 cards among 26 cards with 5 spades}\}$$

and $|\tilde{S}| = 10,400,600$

Example: bridge game (3)

Conditional probability:

$$\mathbf{P}\left(\text{E has 3 of the remaining 5 spades} \mid F\right) = \frac{\binom{5}{3}\binom{21}{10}}{\binom{26}{13}} \simeq .339$$

Intersection and conditioning

Situation:

- Urn with 8 Red and 4 White balls
- Draw 2 balls without replacement

Question: Let

- R_1 = 1st ball drawn is red
- R_2 = 2nd ball drawn is red

Then find $\mathbf{P}(R_1 R_2)$

Intersection and conditioning (2)

Recall:

- Urn with 8 Red and 4 White balls
- Draw 2 balls without replacement

Computation: We have

$$\mathbf{P}(R_1 R_2) = \mathbf{P}(R_1)\mathbf{P}(R_2 |\, R_1)$$

Thus

$$\mathbf{P}(R_1 R_2) = \frac{8}{12}\,\frac{7}{11} = \frac{14}{33} \simeq .42$$

The multiplication rule

Proposition 2.

Let

- $\mathbf{P}$ a probability on a sample space S
- $E_1, \ldots, E_n$ n events

Then

$$\mathbf{P}(E_1 \cdots E_n) = \mathbf{P}(E_1) \prod_{k=1}^{n-1} \mathbf{P}(E_{k+1} | E_1 \cdots E_k) \tag{2}$$

Proof

Expression for the rhs of (2):

$$\mathbf{P}(E_1)\,\frac{\mathbf{P}(E_1E_2)}{\mathbf{P}(E_1)}\,\frac{\mathbf{P}(E_1E_2E_3)}{\mathbf{P}(E_1E_2)}\cdots\frac{\mathbf{P}(E_1\cdots E_{n-1}E_n)}{\mathbf{P}(E_1\cdots E_{n-1})}$$

Conclusion:
By telescopic simplification

Example: deck of cards (1)

Situation:

- Ordinary deck of 52 cards
- Division into 4 piles of 13 cards

Question: If

$$E = \{\text{each pile has one ace}\},$$

compute $\mathbf{P}(E)$

Example: deck of cards (2)

Model: Set

$$\begin{aligned} E_1 &= \{\text{the ace of S is in any one of the piles}\} \\ E_2 &= \{\text{the ace of S and the ace of H are in different piles}\} \\ E_3 &= \{\text{the aces of S, H \& D are all in different piles}\} \\ E_4 &= \{\text{all 4 aces are in different piles}\} \end{aligned}$$

We wish to compute

$$\mathbf{P}\left(E_1 E_2 E_3 E_4\right)$$

Example: deck of cards (3)

Applying the multiplication rule: write

$$\mathbf{P}(E_1 E_2 E_3 E_4) = \mathbf{P}(E_1)\ \mathbf{P}(E_2 | E_1)\ \mathbf{P}(E_3 | E_1 E_2)\ \mathbf{P}(E_4 | E_1 E_2 E_3)$$

Computation of $\mathbf{P}(E_1)$: Trivially

$$\mathbf{P}(E_1) = 1$$

Computation of $\mathbf{P}(E_2 | E_1)$: Given E_1,

- Reduced space is
  $\{$51 labels given to all cards except for ace S$\}$
- $\mathbf{P}(E_2 | E_1) = \frac{51-12}{51} = \frac{39}{51}$

Example: deck of cards (4)

Other conditioned probabilities:

$$\begin{aligned}\mathbf{P}(E_3 | E_1 E_2) &= \frac{50-24}{50} = \frac{26}{50}, \\ \mathbf{P}(E_4 | E_1 E_2 E_3) &= \frac{49-36}{49} = \frac{13}{49}\end{aligned}$$

Conclusion: We get

$$\begin{aligned}\mathbf{P}(E) &= \mathbf{P}(E_1)\,\mathbf{P}(E_2 | E_1)\,\mathbf{P}(E_3 | E_1 E_2)\,\mathbf{P}(E_4 | E_1 E_2 E_3) \\ &= \frac{39 \cdot 26 \cdot 13}{51 \cdot 50 \cdot 49} \simeq .105\end{aligned}$$

Outline

Thomas Bayes

Some facts about Bayes:

- England, 1701-1760
- Presbyterian minister
- Philosopher and statistician
- Wrote 2 books in entire life
- Bayes formula unpublished

Decomposition of $\mathbf{P}(E)$

Proposition 3.

Let

- $\mathbf{P}$ a probability on a sample space S
- E, F two events with $0 < \mathbf{P}(F) < 1$

Then

$$\mathbf{P}(E) = \mathbf{P}(E|\, F)\,\mathbf{P}(F) + \mathbf{P}(E|\, F^c)\,\mathbf{P}(F^c)$$

Bayes' formula

Proposition 4.

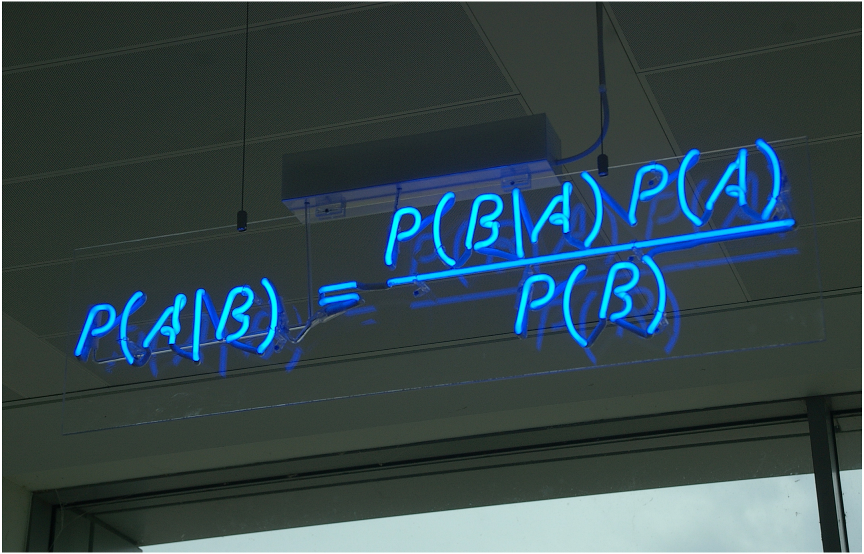
Let

- $\mathbf{P}$ a probability on a sample space S
- E, F two events with $0 < \mathbf{P}(F) < 1$

Then

$$\mathbf{P}(F|\,E) = \frac{\mathbf{P}(E|\,F)\,\mathbf{P}(F)}{\mathbf{P}(E|\,F)\,\mathbf{P}(F) + \mathbf{P}(E|\,F^c)\,\mathbf{P}(F^c)}$$

Iconic Bayes (offices of HP Autonomy)



Example: insurance company (1)

Situation:

- Two classes of people:
  those who are accident prone and those who are not.
- Accident prone: probability .4 of accident in a one-year period
- Not accident prone: probab .2 of accident in a one-year period
- 30% of population is accident prone

Question:

Probability that a new policyholder will have an accident within a year of purchasing a policy?

Example: insurance company (2)

Model: Define

- A_1 = Policy holder has an accident in 1 year
- A = Accident prone

Then

- $S = \{(A_1, A); (A_1^c, A); (A_1, A^c); (A_1^c, A^c)\}$
- Probability: given indirectly by conditioning

Aim:
Compute $\mathbf{P}(A_1)$

Example: insurance company (3)

Given data:

$$\mathbf{P}(A_1|\,A) = .4, \qquad \mathbf{P}(A_1|\,A^c) = .2, \qquad \mathbf{P}(A) = .3$$

Application of Proposition 3:

$$\mathbf{P}\left(A_1\right) = \mathbf{P}\left(A_1|\,A\right)\mathbf{P}(A) + \mathbf{P}\left(A_1|\,A^c\right)\mathbf{P}(A^c)$$

We get

$$\mathbf{P}\left(A_1\right) = 0.4 \times 0.3 + 0.2 \times 0.7 = 26\%$$

Example: swine flu (1)

Situation:

We assume that 20% of a pork population has swine flu.
A test made by a lab gives the following results:

- Among 50 tested porks with flu, 2 are not detected
- Among 30 tested porks without flu, 1 is declared sick

Question:

Probability that a pork is healthy while his test is positive?

Example: swine flu (2)

Model: We set F = "Flu", T = "Positive test"
We have

$$\mathbf{P}(F) = \frac{1}{5}, \quad \mathbf{P}(T^c \mid F) = \frac{1}{25}, \quad \mathbf{P}(T \mid F^c) = \frac{1}{30}$$

Aim:
Compute $\mathbf{P}(F^c \mid T)$

Example: swine flu (3)

Application of Proposition 4:

$$\begin{aligned}\mathbf{P}(F^c \mid T) &= \frac{\mathbf{P}(T \mid F^c)\,\mathbf{P}(F^c)}{\mathbf{P}(T \mid F^c)\,\mathbf{P}(F^c) + \mathbf{P}(T \mid F)\,\mathbf{P}(F)} \\ &= \frac{\mathbf{P}(T \mid F^c)\,\mathbf{P}(F^c)}{\mathbf{P}(T \mid F^c)\,\mathbf{P}(F^c) + [1 - \mathbf{P}(T^c \mid F)]\,\mathbf{P}(F)} \\ &= 0.12\end{aligned}$$

Conclusion:

12% chance of killing swines without proper justification

Henri Poincaré

Some facts about Poincaré:

- Born in **Nancy**, 1854-1912
- Cousin of Raymond Poincaré
  $\hookrightarrow$ French president during WW1
- Mathematician and engineer
- Numerous contributions in
  - Celestial mechanics
  - Relativity
  - Gravitational waves
  - Topology
  - Differential equations

An example by Poincaré (1)

Situation:

- We are on a train
- Someone gets on the train and proposes to play a card game
- The unknown person wins

Question:

Probability that this person has cheated?

An example by Poincaré (2)

Model: We set

- $p =$ probability to win without cheating
- $q =$ probability that the unknown person has cheated
- $W =$ "The unknown person wins"
- $C =$ "The unknown person has cheated"

Hypothesis on probabilities: We assume

$$\mathbf{P}(W \mid C^c) = p, \quad \mathbf{P}(W \mid C) = 1, \quad \mathbf{P}(C) = q$$

Aim:
Compute $\mathbf{P}(C \mid W)$

An example by Poincaré (3)

Application of Proposition 4:

$$\mathbf{P}(C \mid W) = \frac{\mathbf{P}(W \mid C)\,\mathbf{P}(C)}{\mathbf{P}(W \mid C)\,\mathbf{P}(C) + \mathbf{P}(W \mid C^c)\,\mathbf{P}(C^c)} = \frac{q}{q + p(1-q)}$$

Remarks:

(1) We have $\mathbf{P}(C \mid W) \geq q = \mathbf{P}(C)$.
$\hookrightarrow$ the unknown's win increases his probability to cheat

(2) We have

$$\lim_{p \to 0} \mathbf{P}(C \mid W) = 1$$

Odds

Definition 5.

Let

- $\mathbf{P}$ a probability on a sample space S
- A an event

We define the odds of A by

$$\frac{\mathbf{P}(A)}{\mathbf{P}(A^c)} = \frac{\mathbf{P}(A)}{1 - \mathbf{P}(A)}$$

Odds and conditioning

Proposition 6.

Situation: We have

- An hypothesis H, true with probability $\mathbf{P}(H)$
- A new evidence E

Formula: The odds of H after evidence E are given by

$$\frac{\mathbf{P}(H|\,E)}{\mathbf{P}(H^c|\,E)} = \frac{\mathbf{P}(H)}{\mathbf{P}(H^c)}\,\frac{\mathbf{P}(E|\,H)}{\mathbf{P}(E|\,H^c)}$$

Proof

Inversion of conditioning: We have

$$\mathbf{P}(H|\,E) = \frac{\mathbf{P}(E|\,H)\,\mathbf{P}(H)}{\mathbf{P}(E)}$$

$$\mathbf{P}(H^c|\,E) = \frac{\mathbf{P}(E|\,H^c)\,\mathbf{P}(H^c)}{\mathbf{P}(E)}$$

Conclusion:

$$\frac{\mathbf{P}(H|\,E)}{\mathbf{P}(H^c|\,E)} = \frac{\mathbf{P}(H)}{\mathbf{P}(H^c)}\,\frac{\mathbf{P}(E|\,H)}{\mathbf{P}(E|\,H^c)}$$

Example: coin tossing (1)

Situation:

- Urn contains two type A coins and one type B coin.
- When a type A coin is flipped,
  it comes up heads with probability $\frac{1}{4}$
- When a type B coin is flipped,
  it comes up heads with probability $\frac{3}{4}$
- A coin is randomly chosen from the urn and flipped

Question:

Given that the flip landed on heads
$\hookrightarrow$ What is the probability that it was a type A coin?

Example: coin tossing (2)

Model: We set

- A = type A coin flipped
- B = type B coin flipped
- H = Head obtained

Data:

$$\mathbf{P}(A) = \frac{2}{3}, \qquad \mathbf{P}(H|\,A) = \frac{1}{4}, \qquad \mathbf{P}(H|\,B) = \frac{3}{4}$$

Aim:

Compute $\mathbf{P}(A|\,H)$

Example: coin tossing (3)

Application of Proposition 6:

$$\frac{\mathbf{P}(A|\,H)}{\mathbf{P}(B|\,H)} = \frac{\mathbf{P}(A)}{\mathbf{P}(B)}\,\frac{\mathbf{P}(H|\,A)}{\mathbf{P}(H|\,B)}$$

Numerical result: We get

$$\frac{\mathbf{P}(A|\,H)}{\mathbf{P}(B|\,H)} = \frac{2/3}{1/3}\,\frac{1/4}{3/4} = \frac{2}{3}$$

Therefore

$$\mathbf{P}(A|\,H) = \frac{2}{5}$$

Generalization of Proposition 3

Proposition 7.

Let

- $\mathbf{P}$ a probability on a sample space S
- $F_1, \ldots, F_n$ partition of S, i.e
  - F_i mutually exclusive
  - $\cup_{i=1}^{n} F_i = S$
- E another event

Then we have

$$\mathbf{P}(E) = \sum_{i=1}^{n} \mathbf{P}(E|\, F_i)\, \mathbf{P}(F_i)$$

Generalization of Proposition 4

Proposition 8.

Let

- $\mathbf{P}$ a probability on a sample space S
- $F_1, \ldots, F_n$ partition of S, i.e
  - F_i mutually exclusive
  - $\cup_{i=1}^n F_i = S$
- E another event

Then we have

$$\mathbf{P}(F_j | E) = \frac{\mathbf{P}(E | F_j)\, \mathbf{P}(F_j)}{\sum_{i=1}^n \mathbf{P}(E | F_i)\, \mathbf{P}(F_i)}$$

Example: card game (1)

Situation:

- 3 cards identical in form (say Jack)
- Coloring of the cards on both faces:
  - 1 card RR
  - 1 card BB
  - 1 card RB
- 1 card is randomly selected, with upper side R

Question:

What is the probability that the other side is B?

Example: card game (2)

Model: We define the events

- RR: chosen card is all red
- BB: chosen card is all black
- RB: chosen card is red and black
- R: upturned side of chosen card is red

Aim:
Compute $\mathbf{P}(RB|\, R)$

Example: card game (3)

Application of Proposition 8:

$$\mathbf{P}(RB|\,R) = \frac{\mathbf{P}(R|\,RB)\,\mathbf{P}(RB)}{\mathbf{P}(R|\,RR)\,\mathbf{P}(RR) + \mathbf{P}(R|\,RB)\,\mathbf{P}(RB) + \mathbf{P}(R|\,BB)\,\mathbf{P}(BB)}$$

Numerical values:

$$\mathbf{P}(RB|\,R) = \frac{\frac{1}{2} \times \frac{1}{3}}{1 \times \frac{1}{3} + \frac{1}{2} \times \frac{1}{3} + 0 \times \frac{1}{3}} = \frac{1}{3}$$

Example: disposable flashlights

Situation:

- Bin containing 3 different types of disposable flashlights
- Proba that a type 1 flashlight will give over 100 hours of use is .7
- Corresponding probabilities for types 2 & 3: .4 and .3
- 20% of the flashlights are type 1, 30% are type 2, and 50% are type 3

Questions:

1. What is the probability that a randomly chosen flashlight will give more than 100 hours of use?
2. Given that a flashlight lasted over 100 hours, what is the conditional probability that it was a type j flashlight, for $j = 1, 2, 3$?

Example: disposable flashlights (2)

Model: We define the events

- A: flashlight chosen gives more than 100h of use
- F_j: type j is chosen

Aim 1:
Compute $\mathbf{P}(A)$

Example: disposable flashlights (3)

Application of Proposition 7:

$$\mathbf{P}(A) = \sum_{j=1}^{3} \mathbf{P}(A | F_j)\, \mathbf{P}(F_j)$$

Numerical values:

$$\mathbf{P}(A) = 0.7 \times 0.2 + 0.4 \times 0.3 + 0.3 \times 0.5 = .41$$

Example: disposable flashlights (4)

Aim 2:
Compute $\mathbf{P}(F_1 | A)$

Application of Proposition 8:

$$\mathbf{P}\left(F_1 | A\right) = \frac{\mathbf{P}\left(A | F_1\right)\mathbf{P}\left(F_1\right)}{\mathbf{P}(A)}$$

Numerical value:

$$\mathbf{P}\left(F_1 | A\right) = \frac{0.7 \times 0.2}{0.41} = \frac{14}{41} \simeq 34\%$$

Outline

Definition of independence

Definition 9.

Let

- $\mathbf{P}$ a probability on a sample space S
- E, F two events

Then E and F are independent if

$$\mathbf{P}(EF) = \mathbf{P}(E)\,\mathbf{P}(F)$$

Notation:

E and F independent denoted by $E \perp\!\!\!\perp F$

Some remarks

Interpretation: If $E \perp\!\!\!\perp F$, then

$$\mathbf{P}\left(E|\, F\right) = \mathbf{P}(E),$$

that is the knowledge of F does not affect $\mathbf{P}(E)$

Warning: Independent $\neq$ mutually exclusive!
Specifically

$$\begin{aligned} A, B \text{ mutually exclusive} \quad &\Rightarrow \quad \mathbf{P}(A\,B) = 0 \\ A, B \text{ independent} \quad &\Rightarrow \quad \mathbf{P}(A\,B) = \mathbf{P}(A)\,\mathbf{P}(B) \end{aligned}$$

Therefore A et B both independent and mutually exclusive
$\hookrightarrow$ we have either $\mathbf{P}(A) = 0$ or $\mathbf{P}(B) = 0$

Example: dice tossing (1)

Experiment: We throw two dice

Sample space:

- $S = \{1, \ldots, 6\}^2$
- $\mathbf{P}(\{(s_1, s_2)\}) = \frac{1}{36}$ for all $(s_1, s_2) \in S$

Events: We consider

$$A = \text{"1}^{\text{st}}\text{ outcome is 1"}, \qquad B = \text{"2}^{\text{nd}}\text{ outcome is 4"}$$

Question:
Do we have $A \perp\!\!\!\perp B$?

Example: dice tossing (2)

Description of A and B:

$$A = \{1\} \times \{1, \ldots, 6\}, \quad \text{and} \quad B = \{1, \ldots, 6\} \times \{4\}.$$

Probabilities for A and B: We have

$$\mathbf{P}(A) = \frac{|A|}{36} = \frac{1}{6}, \qquad \mathbf{P}(B) = \frac{|B|}{36} = \frac{1}{6}$$

Description of $A\,B$: We have $A\,B = \{(1,4)\}$. Thus

$$\mathbf{P}(A\,B) = \frac{1}{36} = \mathbf{P}(A)\,\mathbf{P}(B)$$

Conclusion: A and B are independent

Example: tossing n coins (1)

Experiment:
Tossing a coin n times

Events: We consider

$$A = \text{"At most one Head"}$$
$$B = \text{"At least one Head and one Tail"}$$

Question:
Are there values of n such that $A \perp\!\!\!\perp B$?

Example: tossing n coins (2)

Model: We take

- $S = \{h, t\}^n$
- $\mathbf{P}(\{s\}) = \frac{1}{2^n}$ for all $s \in S$

Description of A and B:

$$\begin{aligned} A &= \{(t, \ldots, t), (h, t, \ldots, t), (t, h, t, \ldots, t), (t, \ldots, t, h)\} \\ B &= \{(h, \ldots, h), (t, \ldots, t)\}^c \end{aligned}$$

Example: tossing n coins (3)

Computing probabilities for A and B: We have

$$\begin{aligned} \mathbf{P}(A) &= \frac{|A|}{2^n} = \frac{n+1}{2^n} \\ \mathbf{P}(B) &= 1 - \mathbf{P}(B^c) = 1 - \frac{1}{2^{n-1}} \end{aligned}$$

Description of $A\,B$ and

$$A\,B = A \backslash \{(t, \dots, t)\} \quad \Rightarrow \quad \mathbf{P}(A\,B) = \frac{n}{2^n}$$

Example: tossing n coins (4)

Checking independence: We have $A \perp\!\!\!\perp B$ iff

$$\frac{n+1}{2^n}\left(1-\frac{1}{2^{n-1}}\right)=\frac{n}{2^n} \quad\Longleftrightarrow\quad n-2^{n-1}+1=0$$

Conclusion: One can check that

$$x \mapsto x-2^{x-1}+1$$

vanishes for $x=3$ only on $\mathbb{R}_+$. Thus

We have $A \perp\!\!\!\perp B$ iff $n=3$

Independence and complements

Proposition 10.

Let

- $\mathbf{P}$ a probability on a sample space S
- E, F two events
- We assume that $E \perp\!\!\!\perp F$

Then

$$E \perp\!\!\!\perp F^c, \qquad E^c \perp\!\!\!\perp F, \qquad E^c \perp\!\!\!\perp F^c$$

Proof

Decomposition of $\mathbf{P}(E)$: Write

$$\begin{aligned}\mathbf{P}(E) &= \mathbf{P}(E\,F) + \mathbf{P}(E\,F^c) \\ &= \mathbf{P}(E)\,\mathbf{P}(F) + \mathbf{P}(E\,F^c)\end{aligned}$$

Expression for $\mathbf{P}(E\,F^c)$: From the previous expression we have

$$\begin{aligned}\mathbf{P}(E\,F^c) &= \mathbf{P}(E) - \mathbf{P}(E)\,\mathbf{P}(F) \\ &= \mathbf{P}(E)\,(1 - \mathbf{P}(F)) \\ &= \mathbf{P}(E)\mathbf{P}(F^c)\end{aligned}$$

Conclusion:
$E \perp\!\!\!\perp F^c$

Counterexample: independence of 3 events (1)

Warning:
In certain situations we have A, B, C pairwise independent, however

$$\mathbf{P}(A \cap B \cap C) \neq \mathbf{P}(A)\,\mathbf{P}(B)\,\mathbf{P}(C)$$

Example: tossing two dice

- $S = \{1, \ldots, 6\}^2$
- $\mathbf{P}(\{(s_1, s_2)\}) = \frac{1}{36}$ for all $(s_1, s_2) \in S$

Events: Define

$$\begin{aligned} A &= \text{"even number for the 1}^{\text{st}}\text{ outcome"} \\ B &= \text{"odd number for the 2}^{\text{nd}}\text{ outcome"} \\ C &= \text{"same parity for the two outcomes"} \end{aligned}$$

Counterexample: independence of 3 events (2)

Description of A, B, C:

$$\begin{aligned} A &= \{2,4,6\} \times \{1,\ldots,6\} \\ B &= \{1,\ldots,6\} \times \{1,3,5\} \\ C &= (\{2,4,6\} \times \{2,4,6\}) \cup (\{1,3,5\} \times \{1,3,5\}) \end{aligned}$$

Pairwise independence: we find

$$A \perp\!\!\!\perp B, \ A \perp\!\!\!\perp C \text{ and } B \perp\!\!\!\perp C$$

Independence of the 3 events: We have $A \cap B \cap C = \emptyset$. Thus

$$0 = \mathbf{P}(A \cap B \cap C) \neq \mathbf{P}(A)\,\mathbf{P}(B)\,\mathbf{P}(C) = \frac{1}{8}$$

Independence of 3 events

Definition 11.

Let

- $\mathbf{P}$ a probability on a sample space S
- 3 events A_1, A_2, A_3

We say that A_1, A_2, A_3 are independent if

$$\begin{aligned} \mathbf{P}(A_1 A_2) &= \mathbf{P}(A_1)\,\mathbf{P}(A_2), \quad \mathbf{P}(A_1 A_3) = \mathbf{P}(A_1)\,\mathbf{P}(A_3) \\ \mathbf{P}(A_2 A_3) &= \mathbf{P}(A_2)\,\mathbf{P}(A_3) \end{aligned}$$

and

$$\mathbf{P}(A_1 A_2 A_3) = \mathbf{P}(A_1)\,\mathbf{P}(A_2)\,\mathbf{P}(A_3)$$

Independence of n events

Definition 12.

Let

- $\mathbf{P}$ a probability on a sample space S
- n events $A_1, A_2, \ldots, A_n$

We say that $A_1, A_2, \ldots, A_n$ are independent if
for all $2 \leq r \leq n$ and $j_1 < \cdots < j_r$ we have

$$\mathbf{P}(A_{j_1} A_{j_2} \cdots A_{j_r}) = \mathbf{P}(A_{j_1}) \, \mathbf{P}(A_{j_2}) \cdots \mathbf{P}(A_{j_r})$$

Independence of an ∞ number of events

Definition 13.

Let

- $\mathbf{P}$ a probability on a sample space S
- A sequence of events $\{A_i;\, i \geq 1\}$

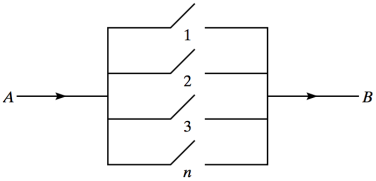
We say that the A_i's are independent if
for all $2 \leq r < \infty$ and $j_1 < \cdots < j_r$ we have

$$\mathbf{P}(A_{j_1} A_{j_2} \cdots A_{j_r}) = \mathbf{P}(A_{j_1})\, \mathbf{P}(A_{j_2}) \cdots \mathbf{P}(A_{j_r})$$

Example: parallel system (1)

Situation:

- Parallel system with n components
- All components are independent
- Probability that i-th component works: p_i



Question:

Probability that the system functions

Example: parallel system (2)

Model: We take

- $S = \{0,1\}^n$
- Probability **P** on S defined by

$$\mathbf{P}(\{(s_1, \dots, s_n)\}) = \prod_{i=1}^{n} p_i^{s_i} (1 - p_i)^{1-s_i}$$

Events:

$$A = \text{"System functions"} , \quad A_i = \text{"}i\text{-th component functions"}$$

Facts about A_i's:
The events A_i are independent and $\mathbf{P}(A_i) = p_i$

Example: parallel system (3)

Computations for $\mathbf{P}(A^c)$:

$$\begin{aligned}\mathbf{P}(A^c) &= \mathbf{P}\left(\cap_{i=1}^n A_i^c\right) \\ &= \prod_{i=1}^n \mathbf{P}\left(A_i^c\right) \\ &= \prod_{i=1}^n (1 - p_i)\end{aligned}$$

Conclusion:

$$\mathbf{P}(A) = 1 - \prod_{i=1}^n (1 - p_i)$$

Example: rolling dice (1)

Experiment:

- Roll a pair of dice
- Outcome: sum of faces

Event: We define

- $E =$ "5 appears before 7"

Question:
Compute $\mathbf{P}(E)$

Example: rolling dice (2)

Family of events: For $n \geq 1$ set

$$E_n = \text{no 5 or 7 on first } n-1 \text{ trials, then 5 on } n\text{-th trial}$$

Relation between E_n and E: We have

$$E = \text{5 appears before 7} = \cup_{n\geq 1} E_n$$

Example: rolling dice (3)

Computation for $\mathbf{P}(E_n)$: by independence

$$\mathbf{P}\left(E_n\right) = \left(1 - \frac{10}{36}\right)^{n-1} \frac{4}{36} = \left(\frac{13}{18}\right)^{n-1} \frac{1}{9}$$

Computation for $\mathbf{P}(E)$:

$$\mathbf{P}(E) = \sum_{n=1}^{\infty} \mathbf{P}\left(E_n\right) = \frac{1}{9} \frac{1}{1 - \frac{13}{18}}$$

Thus

$$\mathbf{P}(E) = \frac{2}{5}$$

Same example with conditioning (1)

New events: We set

- $E =$ "5 appears before 7"
- $F_5 =$ "1st trial gives 5"
- $F_7 =$ "1st trial gives 7"
- $H =$ "1st trial gives an outcome $\neq$ 5,7"

Same example with conditioning (2)

Conditional probabilities:

$$\mathbf{P}(E|\,F_5) = 1, \quad \mathbf{P}(E|\,F_7) = 0, \quad \mathbf{P}(E|\,H) = \mathbf{P}(E)$$

Justification: $E \perp\!\!\!\perp H$ since

$$E\,H = H \cap \{\text{Event which depends on } i\text{-th trials with } i \geq 2\}$$

Same example with conditioning (3)

Applying Proposition 7:

$$\mathbf{P}(E) = \mathbf{P}(E|\, F_5)\, \mathbf{P}(F_5) + \mathbf{P}(E|\, F_7)\, \mathbf{P}(F_7) + \mathbf{P}(E|\, H)\, \mathbf{P}(H) \quad (3)$$

Computation: We get

$$\mathbf{P}(E) = \frac{1}{9} + \frac{13}{18}\mathbf{P}(E),$$

and thus

$$\mathbf{P}(E) = \frac{2}{5}$$

Problem of the points

Experiment:

- Independent trials
- For each trial, success with probability p

Question:

What is the probability that n successes occur before m failures?

Pascal's solution

Notation: set

$$A_{n,m} = \text{"}n \text{ successes occur before } m \text{ failures"}, \qquad P_{n,m} = \mathbf{P}\left(A_{n,m}\right)$$

Conditioning on 1st trial: Like in (3) we get

$$P_{n,m} = pP_{n-1,m} + (1-p)P_{n,m-1} \tag{4}$$

Initial conditions:

$$P_{n,0} = p^n, \qquad P_{0,m} = (1-p)^m \tag{5}$$

Strategy:
Solve difference equation (4) with initial condition (5)

Fermat's solution

Expression for $A_{n,m}$: Write

$$A_{n,m} = \text{"at least } n \text{ successes in } m+n-1 \text{ trials"}$$

Thus $A_{n,m} = \cup_{k=n}^{m+n-1} E_{k,m,n}$ with

$$E_{k,m,n} = \text{"exactly } k \text{ successes in } m+n-1 \text{ trials"}$$

Expression for $P_{n,m}$: We get

$$P_{n,m} = \sum_{k=n}^{m+n-1} \binom{m+n-1}{k} p^k (1-p)^{m+n-1-k}$$