Aalto University
School of Business

Business Analytics 2
Lecture 1: Review of basic probability theory

- *Events*
- *Probability*
- *Independence*
- *Conditional probability*

What does probability have to do with Business Analytics?

- Business decisions are often made under uncertainty

- Probability is the dominant model for uncertainty in Management Science
  - Other models exist (e.g. fuzzy sets) - not mainstream, not covered here

- Probabilities can be subjective or objective
  - Computations are carried out using the same rules (=topic of this lecture)
  - Usually models contain both
  - Is there such a thing as objective probability? (Think about it!)

- Learning objective: Refresh/enhance skills in basic probability

The Sample Space

- Every probability model includes a set of all possible outcomes
  - This set is often called the sample space and denoted by S

- Examples:
  - Flipping a coin: S= {H,T}
  - Flipping two coins: S = {HH, TT, TH, HT}
  - Time for next customer arrival: $S = [0, \infty)$ minutes
  - Demand for a product: S={0,1,2,.....}
  - Number of customers arriving at the drive-in per day: S={0,1,2,.....}
  - Stock prices of four companies: $S = \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+ \times \mathbb{R}_+$

Simple Events (=outcomes) and Events

- Simple event: An individual outcome from sample space S
    - Flipping a coin: H
    - Flipping two coins: TH
    - Time for next customer arrival: 4.3 minutes
    - Demand for a product: 586 units
    - Stock prices of four companies: $ (3.7, 145.3, 45.1, 687.4)

- Event: a collection of one or more outcomes, i.e., a subset of sample space: $E \subseteq S$
    - Flipping two coins: First flip heads, E={HT, HH}
    - Time for next customer arrival: 5 minutes or more, $\mathrm{E} = [5, \infty)$
    - Demand for a product: Less than 200 units, E={0,1,2,...,199}
    - Stock prices of four companies:
      All priced above $100, $E = \{x \in \mathbb{R}^4 | x_i > 100 \text{ for all } i\}$

"E consists of vectors with 4 elements, such that all elements are greater than 100"

Sample Space and Events: Example on rolling two dice

- Simple events: (i,j), $i=1,\ldots,6$, $j=1,\ldots,6$
- Sample Space:

$$
\begin{array}{rl}
S = & \{(1,1),\ (1,2),\ (1,3),\ (1,4),\ (1,5),\ (1,6), \\
 & (2,1),\ (2,2),\ (2,3),\ (2,4),\ (2,5),\ (2,6), \\
 & (3,1),\ (3,2),\ (3,3),\ (3,4),\ (3,5),\ (3,6), \\
B & (4,1),\ (4,2),\ (4,3),\ (4,4),\ (4,5),\ (4,6), \\
 & (5,1),\ (5,2),\ (5,3),\ (5,4),\ (5,5),\ (5,6), \\
 & (6,1),\ (6,2),\ (6,3),\ (6,4),\ (6,5),\ (6,6)\} \quad A
\end{array}
$$

- Examples of events:
  - A="Both dice have the same number"
    - A= {(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)}
  - B="The sum of the numbers is 4"
    - B= {(1,3),(2,2),(3,1)}

Events define other events: Complement events

- Example
  - Sample Space S: "Q4 demand for Microsoft Surface and Apple Ipad"

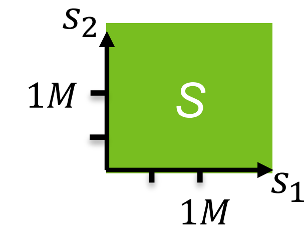
    $$S = \{(s_1, s_2) | s_1, s_2 \in [0, 2M]\}$$

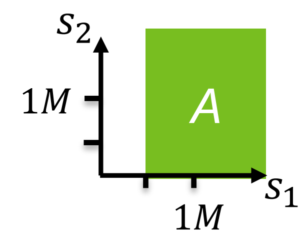
  - Event A: "Surface demand at least 500k"

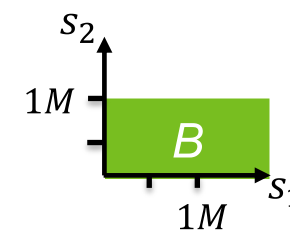
    $$A = \{(s_1, s_2) \in S | s_1 \geq 500{,}000\}$$

  - Event B: "Ipad demand at most 1M"

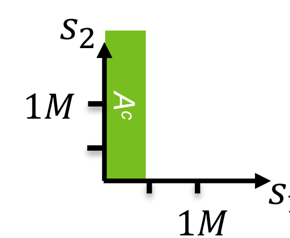
    $$B = \{(s_1, s_2) \in S | s_2 \leq 1{,}000{,}000\}$$





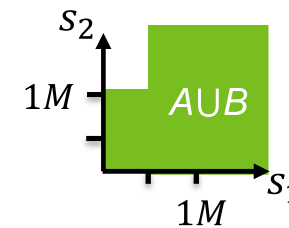


- **Complement** of event A, denoted A^C, consists of all outcomes in sample space (S) that are not in A
  - $A^C = \{(s_1, s_2) \in S | s_1 < 500{,}000\}$, i.e., "Surface demand below 500k"

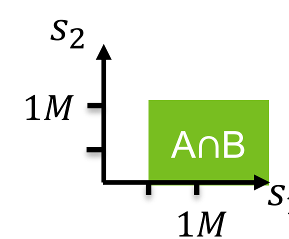


Events define other events: Union and intersection

- **Union** of two events A and B, denoted $A \cup B$, consists of all outcomes either in A or in B (or both)
  - $\mathrm{A} \cup B = \{(s_1, s_2) \in S | s_1 \geq 500{,}000 \text{ or } s_2 \leq 1{,}000{,}000\}$
  - i.e. “Surface demand at least 500k” **OR** “Ipad demand at most 1M”

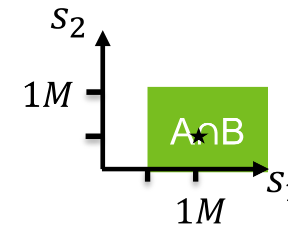


- **Intersection** of two events A and B , denoted $A \cap B$, consists of all outcomes that are in both events
  - $\mathrm{A} \cap B = \{(s_1, s_2) \in S | s_1 \geq 500{,}000 \text{ and } s_2 \leq 1{,}000{,}000\}$
  - i.e. “Surface demand at least 500k” **AND** “Ipad demand at most 1M”

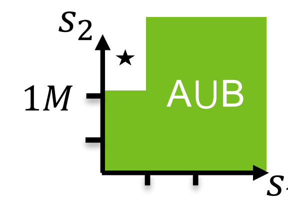


Mutually exclusive & collectively exhaustive events

- Two events A and B with no common outcomes are **mutually exclusive**, or disjoint events ($A \cap B = \emptyset$).
  - Events "Surface demand at least 500k" and "Ipad demand at most 1M" are not mutually exclusive
    - $A \cap B$ includes, for instance, the outcome ⋆ in which Surface demand is 1M and Ipad demand is 500k

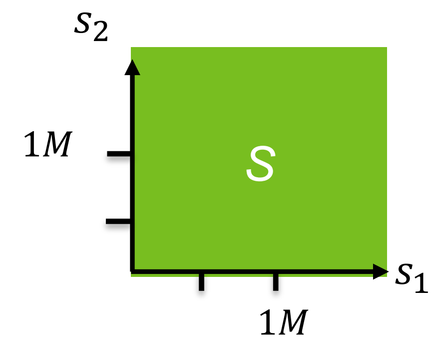


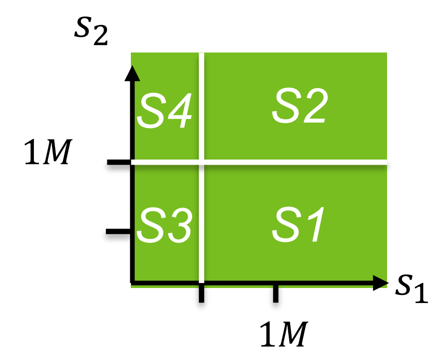
- Two events A and B are **collectively exhaustive** if $A \cup B = S$
  - Events "Surface demand at least 500k" and "Ipad demand at most 1M" are not collectively exhaustive
    - $A \cup B$ does not include, for instance, the outcome ⋆ in which Surface demand is 250k and Ipad demand is 1.5M



Mutually exclusive & collectively exhaustive events

- Events that are mutually exclusive and collectively exhaustive can sometimes be thought of as alternative (market) scenarios, e.g.,
  - Scenario 1 "High Surface demand, Low Ipad demand": $\mathrm{A} \cap B = \{(s_1, s_2) \in S | s_1 \geq 500{,}000 \text{ and } s_2 \leq 1{,}000{,}000\}$
  - Scenario 2 "High Surface demand, High Ipad demand": $\mathrm{A} \cap \mathrm{B}^{\mathrm{C}} = \{(s_1, s_2) \in S | s_1 \geq 500{,}000 \text{ and } s_2 > 1{,}000{,}000\}$
  - Scenario 3 "Low Surface demand, Low Ipad demand": $\mathrm{A}^{\mathrm{C}} \cap B = \{(s_1, s_2) \in S | s_1 < 500{,}000 \text{ and } s_2 \leq 1{,}000{,}000\}$
  - Scenario 4 "Low Surface demand, High Ipad demand": $\mathrm{A}^{\mathrm{C}} \cap \mathrm{B}^{\mathrm{C}} = \{(s_1, s_2) \in S | s_1 < 500{,}000 \text{ and } s_2 > 1{,}000{,}000\}$





Definition of Probability

- **Definition:** Probability measure P is a function that maps all events onto real numbers and satisfies the following three axioms:
  1. $P(S)=1$ *"Outcome is in the sample space with the probability of one"*
  2. $0 \leq P(A) \leq 1$: *"The probability of any event is between 0 and 1"*
  3. If $A \cap B = \emptyset$, then $P(A \cup B) = P(A) + P(B)$: *"Probabilities of mutually exclusive events are additive"*

- Questions:
  – What probability measure would you choose to model
  1. The outcome from tossing a die S={1,…,6}?
  2. Tomorrow's weather in Otaniemi: What probability would you give to events
     1. "It snows"
     2. "It does not snow"
     3. "It does not snow and the temperature is below –5 degrees Celsius"?

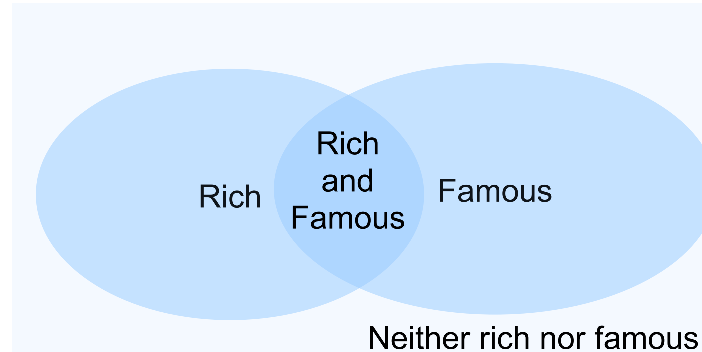
Properties of probability

- From the three axioms it follows that:
  I. $P(\emptyset) = 0$
  II. $P(A^c) = 1 - P(A)$
  III. If $A \subset B$, then $P(A) \leq P(B)$

- Questions:
  - What probabilities would you assign to the events
    1. “It snows”
    2. “It does not snow”
    3. “It does not snow and the temperature is below –5 degrees Celsius”?

Properties of probability (Cont'd)

- From the three axioms it follows that:
  I. $P(\emptyset) = 0$
  II. $P(A^c) = 1 - P(A)$
  III. If $A \subset B$, then $P(A) \leq P(B)$
  IV. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$



- Question:
  - In a certain population, 10% of the people are rich, 5% are famous, and 3% are both rich and famous. A person is randomly selected from this population. What is the probability that the person is
    - not rich?
    - rich or famous?
    - rich but not famous?

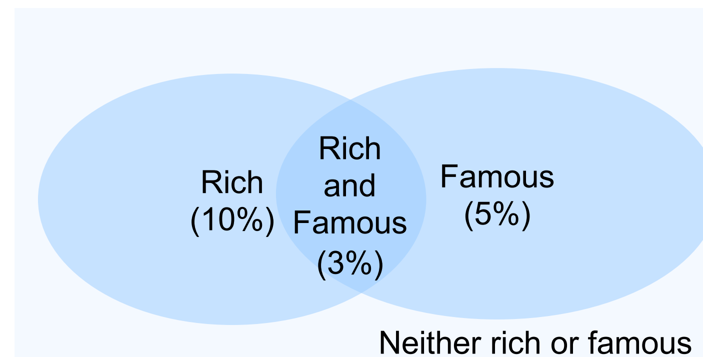
Independence

- **Definition:** Two events A and B are said to be independent if the probability of event 'A and B' is the prob. of 'A' times the prob. of 'B':

$$P(B \cap A) = P(A)P(B)$$

  - Uses:
    - Events are assumed to be independent to ease probability assessment: no need to assess P($A \cap B$) separately as it can computed from P(A) and P(B)
    - Statistically test from observations if events are independent

- Question:
  - A person is randomly selected from the population on the right
  - Are the events "the person is rich" and "the person is famous" independent?



Utilising independence: Example

- If a six-sided die is rolled six times, what is the probability of rolling no sixes?

  E_1 = 1,2,3,4 or 5 on the first roll

  E_2 = 1,2,3,4 or 5 on the second roll

  E_3 = 1,2,3,4 or 5 on the third roll

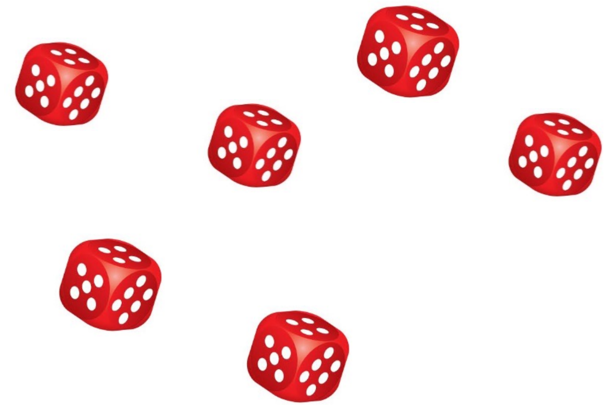
  E_4 = 1,2,3,4 or 5 on the fourth roll

  E_5 = 1,2,3,4 or 5 on the fifth roll

  E_6 = 1,2,3,4 or 5 on the sixth roll

  $P(E_i) = 5/6$ for all $i=1,\ldots,6$

  $\rightarrow$ Event "No sixes" is $E_1 \cap E_2 \cap E_3 \cap E_4 \cap E_5 \cap E_6$



  If the events E_i are independent, then the probability can be obtained as

  $P(E_1 \cap E_2 \cap E_3 \cap E_4 \cap E_5 \cap E_6) = P(E_1)P(E_2)P(E_3)P(E_4)P(E_5)P(E_6) = (5/6)^6 = 0.33$

Conditional probability

- **Definition:** Conditional probability for event A happening given that event B has happened, denoted by P(A | B), is

$$P(A|B) = \frac{P(A\cap B)}{P(B)}$$

  - $P(A \cap B)$: probability that both A and B happen (i.e, *joint probability* of A and B)
  - P(B): probability that B happens

– Example:

$$P(\text{"Stock up"}|\text{"DJIA up"}) = \frac{P(\text{"Stock up"}\cap\text{"DJIA up"})}{P(\text{"DJIA up"})}$$

– If A and B are independent, i.e., $P(A\cap B) = P(A)P(B)$, then information about A having happened does not change the probability of B happening:

$$P(B|A) = \frac{P(A\cap B)}{P(A)} = \frac{P(A)P(B)}{P(A)} = P(B)$$

$$P(\text{rich}|\text{famous}) = \frac{P(\text{rich\&famous})}{P(\text{famous})} = \frac{0.03}{0.05} = 0.6 > P(\text{rich}) = 0.1$$

Joint, marginal & conditional probabilities

Joint prob.
Marginal prob.

- Example:
  - A small vinery is deciding on the mix of grape varieties to grow. Experts have made the following forecast on white wine trends
  - Questions:
    - What is the probability of strong Chardonnay demand?
    - What is the probability of weak Riesling demand?
  - Assume Riesling demand is known before Chardonnay demand
    - What is the conditional probability of strong Chardonnay demand, if Riesling demand is weak?
    - Is Chardonnay demand independent of Riesling demand?

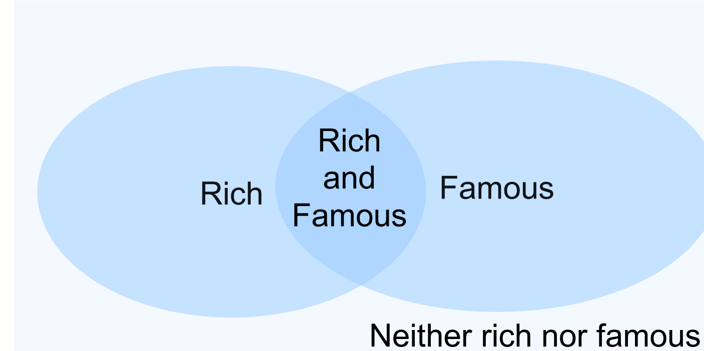
Chardonnay Demand	Riesling Demand: weak	Riesling Demand: strong	row sum
weak	0.05	0.5	0.55
strong	0.25	0.2	0.45
col. sum	0.3	0.7	

Chardonnay conditioned to Riesling

Chardonnay Demand	Riesling Demand: weak	Riesling Demand: strong	row sum
weak	0.17	0.71	0.88
strong	0.83	0.29	1.12
column sum	1	1	

$$P(\text{"C. Str."}|\text{"R. weak"}) = \frac{P(\text{"C.str"} \cap \text{"R.weak"})}{P(\text{"R.weak"})}$$

Summary



Concepts related to probability:

- Sample space, simple event, event
- Complement, union, intersection
- Mutually exclusive / collectively exhaustive events

Properties of probability:

I. $P(\emptyset) = 0$
II. $P(A^c) = 1 - P(A)$
III. If $A \subset B$, then $P(A) \leq P(B)$
IV. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Conditional probability of A given that B has happened: $P(A|B) = \frac{P(A \cap B)}{P(B)}$

If A and B are independent: $P(A \cap B) = P(A)P(B) \rightarrow P(A|B) = P(A)$