Institut für industrielle Informationstechnik
Institute Industrial IT

Information Fusion – Basics on ET

Combination Techniques for Uncertain Information in Measurement and Signal Processing

3.1 Probability Theory

Information Fusion Prof. Dr.-Ing. Volker Lohweg

Institut für industrielle Informationstechnik
Institute Industrial IT

Lectures – Contents

Information Fusion Only valid with verbal explanations Prof. Dr.-Ing. Volker Lohweg 2

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Space S

- Sample Point
  - The outcome of a random experiment.
- Sample Space S
  - The set of all possible outcomes.
  - Discrete and Continuous.
- Events
  - A set of outcomes, thus a subset of S.
  - Certain, Impossible and Elementary.

Information Fusion Only valid with verbal explanations Prof. Dr.-Ing. Volker Lohweg 3

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Space S

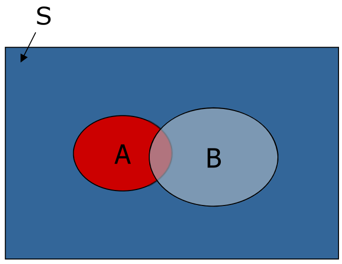
- Probability
  - Frequency (“fraction of times”) an event is seen over a series of observations (“trials”).
- A Priori Probability
  - Probability of an event before a trial.
- “Classical” Probability
  - All possible events uniformly distributed: $P = \frac{\#\text{of desired elements}}{\#\text{of all elements}}$
- Event
  - subset of sample space
  - Simple: 1 element (“sample point”)
  - Compound: >= 2 elements

Information Fusion Only valid with verbal explanations Prof. Dr.-Ing. Volker Lohweg 4

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Set Operations

- Union* $A \cup B$
- Intersection* $A \cap B$
- Complement* $\overline{A} = A^C = / A$
- Empty set $A \cap B = \{\ \} = \varnothing$
- Properties and terms
  - Commutation $A \cup B = B \cup A$
  - Associativity $A \cup (B \cup C) = (A \cup B) \cup C$
  - Distribution $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
  - De Morgan's Rule $(A \cup B)^C = A^C \cap B^C, \quad (A \cap B)^C = A^C \cup B^C$
  - Cardinal number (no. of elements in a set) $|A \cup B|, \quad |S|$



* In logic circuit design: or, and, not

Information Fusion　Only valid with verbal explanations　Prof. Dr.-Ing. Volker Lohweg　5

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Discrete Random Variables

- A is a Boolean-valued random variable, if A denotes an event, and there is some degree of uncertainty as to whether A occurs.
  - Examples
    - A = The german chancellor in 2023 will be female.
    - A = You wake up tomorrow with a headache.
    - A = You have an influenza (the flu).
- We write P[A] as "the fraction of possible worlds in which A is true"

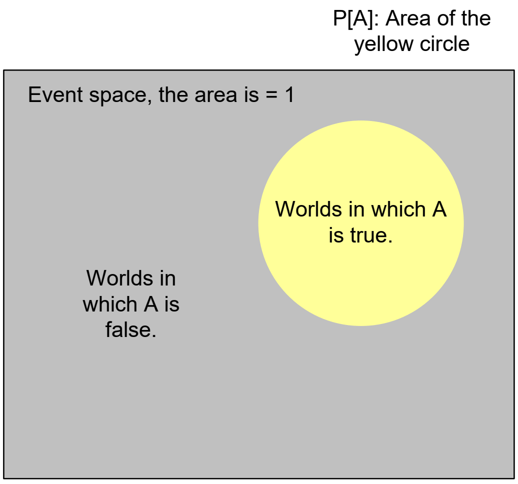
$P[A] = P(A)$

Information Fusion　Only valid with verbal explanations　Prof. Dr.-Ing. Volker Lohweg　6

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Discrete Random Variables

- Visualization of A

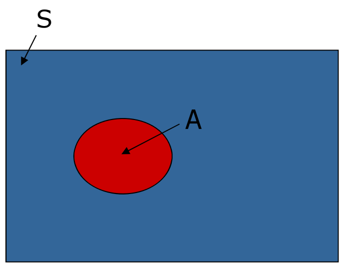


Institut für industrielle Informationstechnik
Institute Industrial IT

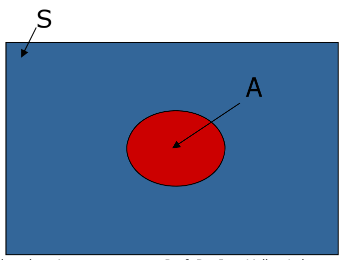
3.1 Probability Theory – Interpretations

- 1. Axiom

$$\underbrace{0 \leq P}\left[A\right] \leq 1$$



$$0 \leq \underbrace{P\left[A\right] \leq 1}$$



Institut für industrielle Informationstechnik
Institute Industrial IT

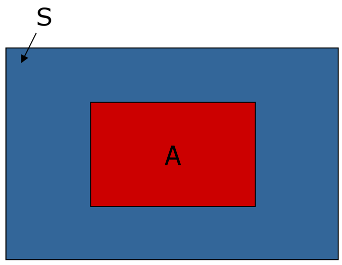
3.1 Probability Theory – Interpretations

- 2. Axiom

  $P[S]=1$

  - Corollary

    $P[A]\leq 1$



Information Fusion — Only valid with verbal explanations — Prof. Dr.-Ing. Volker Lohweg — 9

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Interpretations

- 3. Axiom
  - Condition

    $A\cap B=\emptyset$

    - Corollary

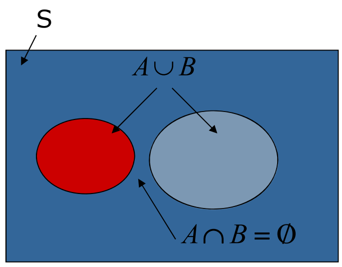
      $P[\varnothing]=0$

  - Result

    $P[A\cup B]=P[A]+P[B]$

  - Generalisation

    $$P\left[\bigcup_{k=1}^{\infty}A_k\right]=\sum_{k=1}^{\infty}P[A_k]$$



Information Fusion — Only valid with verbal explanations — Prof. Dr.-Ing. Volker Lohweg — 10

Institut für industrielle Informationstechnik
Institute Industrial IT

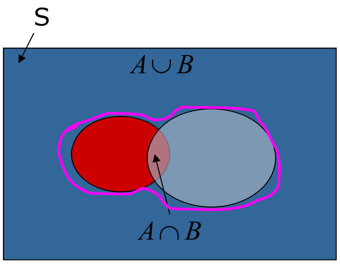
3.1 Probability Theory – Interpretations

- A fundamental corollary

$$P[A \cup B] =$$
$$P[A] + P[B] - P[A \cap B]$$

  - Condition

$$A \cap B \neq \emptyset$$



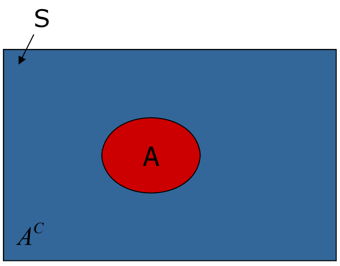
Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Interpretations

- A fundamental corollary

$$P[A^C] = 1 - P[A]$$

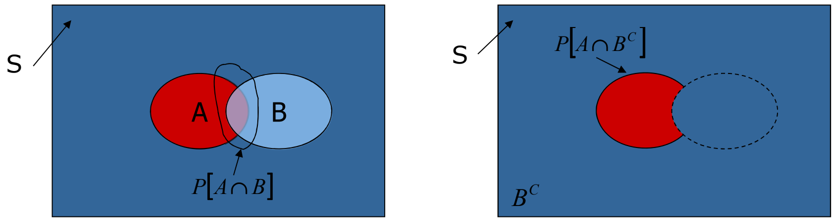
$$P[S] = 1 \rightarrow P[S] = P[A^C] + P[A]$$
$$P[S] - P[A] = P[A^C] \rightarrow$$
$$P[A^C] = P[S] - P[A] = 1 - P[A].$$



Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Interpretations

- An important theorem (Total Probability) P[A] = ?



$$P[A] = P[A \cap B] + P[A \cap B^C]$$

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Axioms

- **Axioms**
  - Uncertainty $0 \leq P[A] \leq 1$
  - Certain event $P[S] = 1$
  - Impossible event $P[\varnothing] = 0$
  - If A_1, A_2, … are pairwise exclusive $A \cap B = \emptyset$

$$P[A_1 \cup A_2] = P[A_1] + P[A_2]$$

$$P\left[\bigcup_{k=1}^{\infty} A_k\right] = \sum_{k=1}^{\infty} P[A_k]$$

- **Corollaries**
  - $P[A^C] = 1 - P[A]$
  - $P[A] \leq 1$
  - $P[\varnothing] = 0$

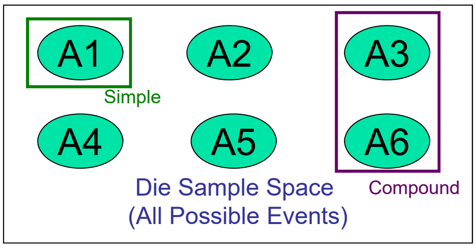
$$P[A \cup B] =$$
$$P[A] + P[B] - P[A \cap B]$$

  - $P[A] = P[A \cap B] + P[A \cap B^C]$

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – A priori Probability

- Example 3.1-1 "Toss a dice"



  - Simple Event: "toss a 1"
    - P[A1] = 1 Desired / 6 Possible = 1/6
  - Compound Event: "toss 3 or 6"
    - P[A3 U A6] = 2 Desired / 6 Possible = 1/3
  - Probability of *not* tossing a 1
    - $P[A1^c] = 1 - P[A1] = 1 - 1/6 = 5/6$
  - Probability of *not* tossing 3 or 6
    - $P[(A3 \cup A6)^c] = 1 - P[A3 \cup A6] = 1 - 1/3 = 2/3$

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – A priori Probability

- Example 3.1-1 – Union of Events, cont'd
  - Probability for *Union of Events* (One trial)
    For two events:

$$P[A \cup B] = P[A] + P[B] - P[A \cap B] = \frac{|A \cup B|}{|\text{Sample Space}|}$$

  - P [toss an even number ***or*** one no. divisible by 3]

$$P[(\text{toss an even no.}) \cup (\text{no. divisible by 3})] = P[A2 \cup A4 \cup A6] + P[A3 \cup A6] - P[A6] = \frac{3}{6} + \frac{2}{6} - \frac{1}{6} = \frac{4}{6} = \frac{2}{3}.$$

$$P[(\text{toss an even no.}) \cup (\text{no. divisible by 3})] = \frac{|\{A2, A4, A6\} \cup \{A3, A6\}|}{|\text{Sample Space}|} = \frac{|\{A2, A3, A4, A6\}|}{6} = \frac{4}{6} = \frac{2}{3}$$

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – A priori Probability

- Example 3.1-1 – Compound Events, cont'd
  - Probability of Two *Compound Events in a Single Trial*

$$P[A \cap B] = \frac{|A \cap B|}{|\text{Sample Space}|}$$

  - P [toss an even number ***and*** one no. divisible by 3]

$$P[(\text{toss an even no.}) \cap (\text{no. divisible by 3})] = \frac{|\{A2, A4, A6\} \cap \{A3, A6\}|}{|\text{Sample Space}|} = \frac{|\{A6\}|}{6} = \frac{1}{6}$$

Information Fusion Only valid with verbal explanations Prof. Dr.-Ing. Volker Lohweg 17

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – A priori Probability

- (Statistically) *Independent Events Across Trials*
  Do not affect one another (are "mutually independent (exclusive)").
  The events / measurements are de-correlated.

$$P[A, B, C, \ldots,] = P[A] \cdot P[B] \cdot P[C] \cdots$$

- Example 3.1-2: "toss a dice – revisited"
  - e.g.: P [toss an even number ***and then*** toss a no. divisible by 3]

$$P[\{A2, A4, A6\}, \{A3, A6\}] = P[\{A2, A4, A6\}] \cdot P[\{A3, A6\}] = \frac{3}{6} \cdot \frac{2}{6} = \frac{1}{6}$$

Information Fusion Only valid with verbal explanations Prof. Dr.-Ing. Volker Lohweg 18

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Conditional Probability

- What is conditional probability?

> Definition 3.1: *Conditional Probability*
>
> P[A|B] is the probability of an event A under the condition that an event B, P[B] >0, has occurred. *⁾

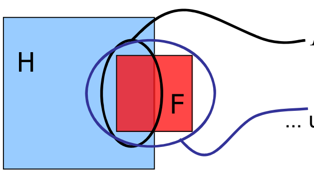
- E.g.:
  - P[Have an accident | drive fast]
  - P[headache | flu]
  - P[rain | clouds]
  - $P[y|x_1,x_2,\ldots,x_N]$

*⁾ It does not mean that P[A|B]=P[A] if P[B] does not occur, P[B] = 0!

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Conditional Probability

- Example 3.1-2 "Coming down with the flu"
  - "Headaches are rare and the flu is rarer, but if you're coming down with the flu, there's a 50-50 chance you'll have a headache."
    - Event H = "Have a headache" → P[H] = 1/10
    - Event F = "Coming down with the flu" → P[F] = 1/40



$H \cap F$

… under the condition of F.

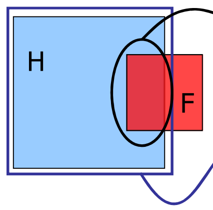
$$P[H|F] = \frac{P[H \cap F]}{P[F]}$$

$$P[H|F] = \frac{(1/2)(1/40)}{(1/40)} = \frac{1}{2}$$

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Conditional Probability

- Example 3.1-2 "Coming down with the flu", cont'd
  - How do we get P[F|H] ?



$H \cap F$

… under the condition of H.

$$P[F|H] = \frac{P[H \cap F]}{P[H]}$$

$$P[F|H] = \frac{(1/2)(1/40)}{(1/10)} = \frac{1}{8}$$

Information Fusion | Only valid with verbal explanations | Prof. Dr.-Ing. Volker Lohweg | 21

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Conditional Probability

- Conditional Probability – How to change the condition
  - Event Headache (H) under the condition of Event Flu (F).

$$P[H|F] = \frac{P[H \cap F]}{P[F]} \rightarrow P[H \cap F] = P[H|F] \cdot P[F].$$

  - Event Flu (F) under the condition of Event Headache (H).

$$P[F|H] = \frac{P[F \cap H]}{P[H]} \rightarrow P[F \cap H] = P[F|H] \cdot P[H].$$

$$P[H \cap F] = P[F \cap H] \;\rightarrow\; P[F|H] \cdot P[H] = P[H|F] \cdot P[F] \rightarrow P[F|H] = P[H|F] \cdot \frac{P[F]}{P[H]}$$

Information Fusion | Only valid with verbal explanations | Prof. Dr.-Ing. Volker Lohweg | 22

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Conditional Probability

- The Bayes Rule
  - Generalization $P[A|B] = \frac{P[A \cap B]}{P[B]} = P[B|A] \cdot \frac{P[A]}{P[B]}$
  - Joint probability $P[A \cap B] = P[A|B] \cdot P[B]$

Bayes, Thomas (1763) An essay towards solving a problem in the doctrine of chances. Philosophical Transactions of the Royal Society of London, **53:370-418**

Information Fusion   Only valid with verbal explanations   Prof. Dr.-Ing. Volker Lohweg   23

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Conditional Probability

- Variations
  - Total probability $P[A] = P[A \cap B] + P[A \cap B^C]$

$$P[A|B] = \frac{P[A \cap B]}{P[B]} = \frac{P[B|A] \cdot P[A]}{P[B \cap A] + P[B \cap A^c]} = \cdots$$

$$\cdots = \frac{P[B|A] \cdot P[A]}{\underbrace{P[B \cap A]}_{P[B|A] \cdot P[A]} + \underbrace{P[B \cap A^c]}_{P[B|A^c] \cdot P[A^c]}} = \frac{P[B|A] \cdot P[A]}{P[B|A] \cdot P[A] + P[B|A^c] \cdot P[A^c]}$$

Information Fusion   Only valid with verbal explanations   Prof. Dr.-Ing. Volker Lohweg   24

CH1

Institut für industrielle Informationstechnik
Institute Industrial IT

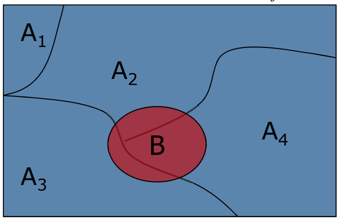
3.1 Probability Theory – Conditional Probability

- Variations (General Rule of Total Probability)
  - If $A_1, A_2, \ldots, A_n$ a ***partition*** of S, then

$$P[B] = P[B|A_1] \cdot P[A_1] + P[B|A_2] \cdot P[A_2] + \cdots + P[B|A_n] \cdot P[A_n]$$

$$A_i \cap A_j = \varnothing$$

Proof by you!



Example for n=4

Information Fusion | Only valid with verbal explanations | Prof. Dr.-Ing. Volker Lohweg | 25

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Conditional Probability

- Variations (General Rule of Total Probability)
  - If $A_1, A_2, \ldots, A_n$ a ***partition*** of S, then

$$P[A_j|B] = \frac{P[A_j \cap B]}{P[B]} = \frac{P[B|A_j] \cdot P[A_j]}{\sum_{j=1}^{n} P[B|A_j] \cdot P[A_j]}$$

Information Fusion | Only valid with verbal explanations | Prof. Dr.-Ing. Volker Lohweg | 26

Folie 25

CH1 Christoph-Alexander Holst; 02.04.2020 inIT Institut für industrielle Informationstechnik Institute Industrial IT

3.1 Probability Theory – Bayesian Inference

- Evaluating the posterior probability through Bayesian inference.

Likelihood Probability

A priori Probability

A posteriori Probability

$$P[A|B]=\frac{P[B|A]\cdot P[A]}{P[B]}$$

Evidence

$$\boxed{posterior=\frac{likelihood\ \cdot prior}{evidence}}$$

inIT Institut für industrielle Informationstechnik Institute Industrial IT

3.1 Probability Theory – Bernoulli Trials

- A trial with only two possible outcomes is used so frequently as a building block of a random experiment that it is called a Bernoulli trial.
- It is usually assumed that the trials that constitute the random experiment are independent. This implies that the outcome from one trial has no effect on the outcome to be obtained from any other trial.
- Furthermore, it is often reasonable to assume that the probability of a success on each trial is constant.

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Bernoulli Trials

- A random experiment consisting of *n* repeated trials which are
  - independent,
  - have only two outcomes (true, false),
  - and the probability *p* of the trials is always constant

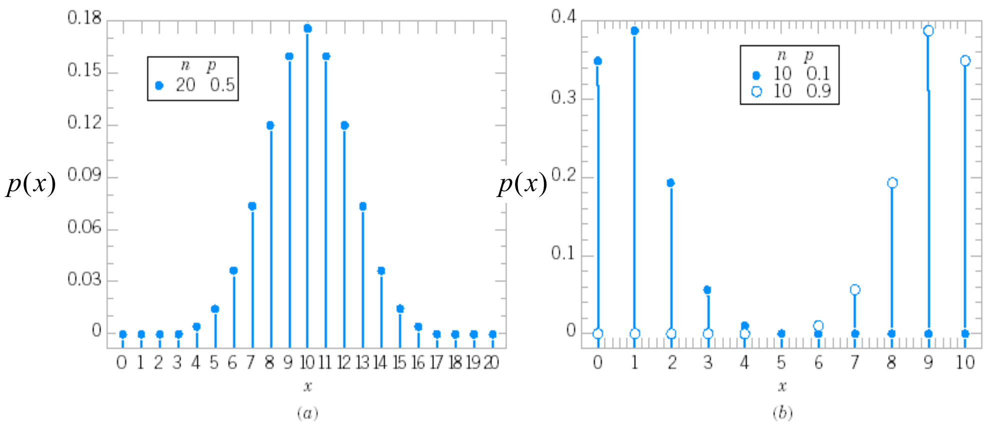
  can be described with a Bernoulli process – it is also called *binomial experiment*.
- The random variable *X* that equals the number of trials which are true has a **binomial distribution** with the parameters $0 < p < 1$ and $n = 1, 2, 3, \ldots$.

$$b(x) = \binom{n}{x} p^x \cdot (1-p)^{n-x}, x = 1,2,3,\ldots,n$$

$$\mu = E[X] = np, \quad \sigma^2 = V[X] = np(1-p)$$

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Bernoulli Trials



Binomial distribution for selected values of n and p.

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Poisson Process

- It is well known, that production processes can be described with a binominal probability distribution in the case of decorrelated occurrences of true and false decisions in a process.
- This fact is in general true for high volume production with low waste rates (e.g. banknotes, electronic devices, screws, etc.). It can also be assumed, that the process is
  - stationary – the occurrence is only dependent on the measurement and analysis time,
  - memoryless – the number of events is not dependent on the statistic of the occurrences itself and
  - ordinal – two events with different results can only occur consecutively.

Institut für industrielle Informationstechnik
Institute Industrial IT

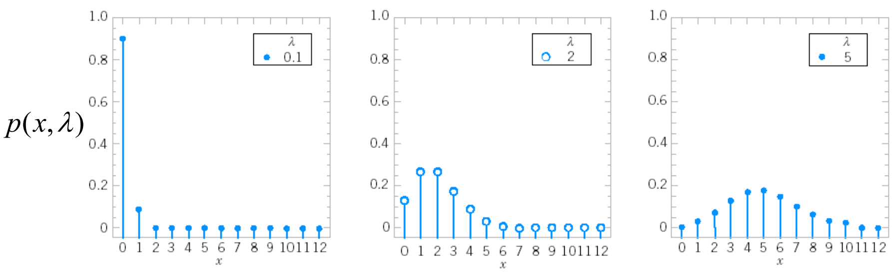
3.1 Probability Theory – Poisson Process

- Poisson density function as a continuous border process of a discrete binomial density function:

$$p(x,\lambda) = \frac{\lambda^x}{x!} e^{-\lambda}$$

- Mean and variance

$$E[X] = \lambda \qquad Var[X] = \lambda$$



Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Poisson Process

Example 3.1-5 "Disk error"

- Find the probability that zero particles occur on a disk under test. We can assume that
$$P(X=0)=e^{-10}.$$
- Find the probability that 12 or less particles are on the disc.
- Poisson distribution function (12 is the number of occurences)

$$P(X \le 12) = P(X=0)+P(X=1)+\cdots+P(X=12)=\cdots$$
$$\cdots=\sum_{x=0}^{12}\frac{10^x}{x!}e^{-10}=0.7916.$$

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Difficulties

- Unknown Probabilities
  - Often it is difficult or impossible to obtain all the required probabilities
  - As more types of evidence are added to a Bayesian framework, more probabilities are required
- Choosing Values "in Desperation"
  - If unsure, all simple events equally likely
  - Not always a reasonable assumption; can lead to inappropriate conclusions (e.g. cost-benefit of drilling for oil under your house)

Institut für industrielle Informationstechnik
Institute Industrial IT

3.1 Probability Theory – Summary

- Probability Theory is a classical concept in measurement analysis.
- Two strategies are in general possible, based on continuous or discrete variables.
- Inference is possible with Probability Theory.
- Inference is based on conditional probability and the prior.
- One main drawback is that the prior needs to be determined carefully.
- Incorrect priors give wrong or misleading results.