Tutorium
Statistische Methoden in der Sprachverarbeitung

Tutorin: Suteera Seeha

Kontaktmöglichkeiten
- s.seeha@campus.lmu.de
- discord server (Statistische Methode Kanal, DM)

Tutorium-Website: https://tutorium-statistische-methoden.github.io/winter_2021_2022/

Ziele des Tutoriums

- helfen den Studierenden, die Inhalte aus der Vorlesung und Übungen zu verstehen
- helfen den Studierenden bei der Vorbereitung auf die Prüfungen
- Fragen beantworten

Was werden wir in diesem Tutorium tun?

- Besprechung der Vorlesungsunterlagen + kleine Übungen zum besseren Verständnis
- Besprechung der Übungsblätter, falls erforderlich

Überblick

1. Grundlagen
   - Textkorpora
   - Mathematische Grundlagen

ok

2. Generative Modelle und Anwendungen
   - Statistische Tests: Kollokationsextraktion
   - Markowmodelle: Sprachidentifizierung
   - Parameterglättung
   - Naïve-Bayes-Modelle: Wortbedeutungs-Desambiguierung
   - Hidden-Markov-Modelle: Wortart-Tagging
   - PCFGs: Statistisches Parsen
3. Diskriminative Modelle
   - Perzeptron-Algorithmus
   - Conditional Random Fields

difficult

4. Zusammenfassung

Tips to prepare for the exams

- Vorlesungklausur:
  - make sure you understand every single slide
  - if you don‘t understand any concept, formula, symbol, ask Herr Schmid in Repetitorium / or ask me
  - solve old exams (if you don‘t know the answer, ask Herr Schmid)
  - go to the lecture...
  - pay attention when someone asks about something in class because it could become questions in the exam

- Übungklausur:
  - try to understand the provided solution(Musterlösung)
  - if you are not good at Python, try to improve your skill as soon as possible

Online- oder Präsenzveranstaltung ?

Die meisten Studenten bevorzugen das Online-Format, daher werden wir den Kurs mit Zoom weiterführen.

Textkorpus

Korpus: Sammlung von Texten für linguistische Zwecke

Einfache Korpusstatistiken für die Erzählung "Tom Sawyer":

- Gesamtzahl der Wörter: 71370 (Tokens)
- Zahl der unterschiedlichen Wörter: 8018 (Types)
- Zahl der Wörter, die einmal auftraten: 3993 (Hapax Legomena)

- sentence length analysis,
- average word length analysis,

Why are we interested in computing these statistics ?

-Before building an NLP model, it is important to know the data that we will be working with.
-This step is also called **Exploratory Data Analysis** (explore and visualize data)

Wörter sortiert nach Häufigkeit

-This statistic tells us that most words are function words
-for some NLP tasks, we might consider removing them
-content words are likely to have a lower frequency

Wort	Häufigkeit
the	3332
and	2972
a	1775
to	1725
of	1440
was	1161
it	1027
in	906
that	877
he	877
I	783
his	772
you	686
Tom	679
with	642

Funktionswörter: the, and, a, ...

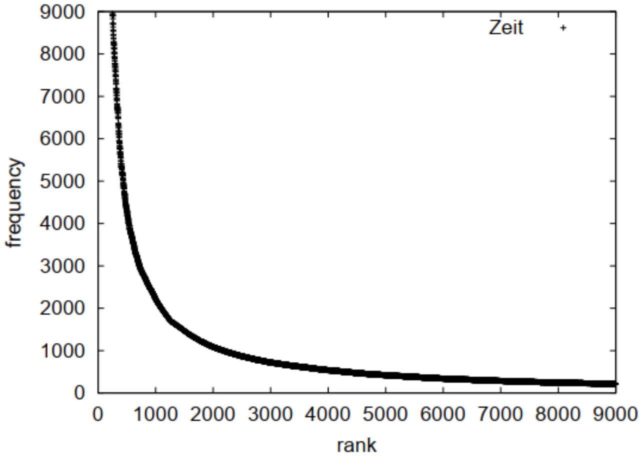
Inhaltswörter: Tom

rank

1
2
3
4
5
…

Wort	Häufigkeit
the	3332
and	2972
a	1775
to	1725
of	1440
was	1161
it	1027
in	906
that	877
he	877
I	783
his	772
you	686
Tom	679
with	642

This graph represents a relationship between the frequency and the rank of words
- low rank -> high freq
- high rank -> low freq



lineare Skala

$\Rightarrow$ Zipf'sche Verteilung

The graph also shows us the distribution of words in this corpus
- a small number of words have high freq
- most of the words have quite low freq

Zipf's Gesetz

Wort	f	r	$f \cdot r$
the	3332	1	3332
and	2972	2	5944
a	1775	3	5235
he	877	10	8770
but	410	20	8400
be	294	30	8820
two	104	100	10400
turned	51	200	10200
you'll	30	300	9000
family	8	1000	8000
brushed	4	2000	8000
sins	2	3000	6000
Could	2	4000	8000
Applausive	1	8000	8000

Zipf's Gesetz:

$f \sim \frac{1}{r} \quad (f \cdot r \approx K)$

$\Rightarrow$ Wenige Wörter sind sehr häufig

$\Rightarrow$ Die meisten Wörter sind sehr selten.

- the freq of a word is inverse proportional to the rank
- if f becomes larger, r becomes smaller
- if we compute f* r we will always get the same value (constant K) (this statement is not very strict)

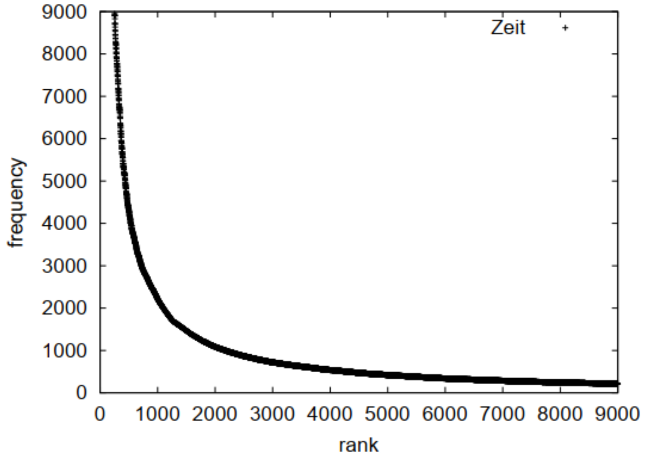
Aufgabe 1) Wie lautet das Zipfsche Gesetz und was sagt es aus? (2 Punkte)

Aufgabe 1) Wie lautet das Zipfsche Gesetz und was sagt es aus? (2 Punkte)

- Zipf's law states that given some corpus of natural language, the frequency of any word is inversely proportional to its rank in the frequency table.
- This tells us that most of the words in a corpus are normally rare (have low frequency)

- Formula? $f \sim \frac{1}{r}$

- How does the graph look like?



Zipf's law (/zɪf/, not /tsɪpf/ as in German) is an empirical law formulated using mathematical statistics that refers to the fact that for many types of data studied in the physical and social sciences, the rank-frequency distribution is an inverse relation. The **Zipfian distribution** is one of a family of related discrete power law probability distributions. It is related to the zeta distribution, but is not identical.

Zipf's law was originally formulated in terms of quantitative linguistics, stating that given some corpus of natural language utterances, the frequency of any word is inversely proportional to its rank in the frequency table. Thus the most frequent word will occur approximately twice as often as the second most frequent word, three times as often as the third most frequent word, etc. For example, in the Brown Corpus of American English text, the word "*the*" is the most frequently occurring word, and by itself accounts for nearly 7% of all word occurrences (69,971 out of slightly over 1 million). True to Zipf's Law, the second-place word "*of*" accounts for slightly over 3.5% of words (36,411 occurrences), followed by "*and*" (28,852). Only 135 vocabulary items are needed to account for half the Brown Corpus.[1]

Wichtige Begriffe

Zufallsexperiment: Experiment (Versuch) mit mehreren möglichen Ausgängen

Ergebnis: Resultat eines Experimentes

Ergebnisraum: Menge aller möglichen Ergebnisse

Ereignis: Teilmenge des Ergebnisraumes

Elementarereignis: anderes Wort für Ergebnis

Stichprobe: Folge von Ergebnissen bei einem wiederholten Experiment

Beispiel

Experiment (Prozess/Situation): wir werfen 2 Würfel
Observation: Wir beobachten die Zahl auf jedem Würfel

Ergebenis/Elementarereignis(outcome):
z.B. (1,3) oder (4,3)
wobei die erste Zahl die Nummer des ersten Würfels und die zweite Zahl die Nummer des zweiten Würfels ist.

Ergebnisraum (sample space):
{ (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)
(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) }

Ereignis(event):
A = {(1,2)(2,1)} -- Das Ereignis, bei dem die Summe der beiden Würfel gleich 3 ist.
B = { (1,1)(1,2)(1,3)(1,4)(1,5) (1,6) } -- Das Ereignis, bei dem der erste Würfel eine 1 ergibt.
C = { (5,6) } -- Das Ereignis, bei dem das Ergebnis (5,6) auftritt.

Stichprobe (sample):
- Ergebnisse, die wir nach Durchführung des Experiments tatsächlich erhalten.
- Wenn wir das Experiment zum Beispiel 3 Mal durchführen, erhalten wir (1,2),(3,2),(4,3)

Eine Stichprobe ist zunächst einmal eine Teilmenge einer Grundgesamtheit. Für eine Zufallsstichprobe werden zusätzliche Bedingungen gestellt:

- Die Elemente werden zufällig aus der Grundgesamtheit gezogen und
- die Wahrscheinlichkeit, mit der ein Element aus der Grundgesamtheit gezogen wird, ist angebbar.

https://de.wikipedia.org/wiki/Zufallsstichprobe

Ergebnisraum (sample space):

{ (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)
(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) }

Question

Are (3,4) and (4,3) the same?

Answer:

No, (3,4) means for the first die we get value 3 and for the second die we get 4

(4,3) means for the first die we get value 4 and for the second die we get 3

They are not the same outcome

An experiment is a process by which we observe something uncertain.

Suppose, we are tossing a coin.

Here we get only two results that are head or tail.

But before the toss we can’t predict the result, so it is said to be a random experiment.

Zufallsexperiment

In der Statistik geht es um die Wahrscheinlichkeit von Ereignissen:

Beispiel: Wie wahrscheinlich ist es, sechs Richtige im Lotto zu haben?

Zufallsexperiment: Experiment (Versuch) mit mehreren möglichen Ausgängen (*Wurf von zwei Würfeln*)

Ergebnis: Resultat eines Experimentes (*3 Augen auf Würfel 1 und 4 Augen auf Würfel 2*)

Ergebnisraum: Menge aller möglichen Ergebnisse

Ereignis: Teilmenge des Ergebnisraumes (*7 Augen auf zwei Würfeln*)

Elementarereignis: anderes Wort für Ergebnis

Stichprobe: Folge von Ergebnissen bei einem wiederholten Experiment

Wenn wir ein Zufallsexperiment entwerfen, müssen wir genau definieren, was wir beobachten

- Wir nehmen 2 Kugeln aus einem Schachtel (draw with replacement) ---- > Das ist der Prozess/die Situation
- Nachdem dieser Prozess abgeschlossen ist, können wir z.B. Folgendes beobachten:
  - die Nummer auf der Kugeln
    - Ergebnisraum: {7,4,1,9,5,8 }
  - die Farbe der Kugeln
    - Ergebnisraum : {rot, blau, gelb, hellblau}
  - die Nummer und die Farbe
    - Ergebnisraum : {(rot, 7), (blau, 4), (gelb, 9), ...}

- Beispiel für ein **Ereignis** aus dem Ergebnisraum {(rot, 7), (blau, 4), (gelb, 9), ...}
  - **A** ist ein Ereignis, bei dem blaue Kugeln ausgewählt werden.
  - A = { (blau, 4), (blau, 8)}

Example 1

Tossing three distinct coins once.

Then, with H and T standing for "heads" and "tails," respectively, a sample space is:

The event A = "no more than 1 H occurs" is given by:

$A = \{$.

Example 2

Rolling once two distinct dice.

Then a sample space is:

and the event B = "the sum of numbers on the upper faces is ≤ 5" is:

$B =$

Example 1

Tossing three distinct coins once.

Then, with H and T standing for "heads" and "tails," respectively, a sample space is:

$$\mathscr{S} = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}.$$

The event A = "no more than 1 H occurs" is given by:

$$A = \{TTT, HTT, THT, TTH\}.$$

Example 2

Rolling once two distinct dice.

Then a sample space is:

$$\mathscr{S} = \{(1,1),(1,2),\ldots,(1,6),\ldots,(6,1),(6,2),\ldots,(6,6)\},$$

and the event B = "the sum of numbers on the upper faces is ≤ 5" is:

$$B = \{(1,1),(1,2),(1,3),(1,4),(2,1),(2,2),(2,3),(3,1),(3,2),(4,1)\}.$$

Übung

Wir werfen 2 Mal eine Münze und beobachten das Ergebnis (Kopf oder Zahl)

- Geben Sie ein Beispiel für ein Ergebnis:
- Geben Sie den Ergebnisraum an:
- Geben Sie ein Beispiel für ein Ereignis:

Übung

Wir werfen 2 Mal eine Münze und beobachten das Ergebnis (Head order Tail)

- Geben Sie ein Beispiel für ein Ergebnis: (tail,tail)
- Geben Sie den Ergebnisraum an: {(tail,tail), (head,head),(head,tail),(tail,head)}
- Geben Sie ein Beispiel für ein Ereignis: {(tail,tail),(head,head)}

Wahrscheinlichkeitsverteilung: Funktion, die jedem Ergebnis o einen Wert zwischen 0 und 1 zuweist

Z.B. wir haben ein Ergebnis „head“ und definieren p(„head“) = 0.5

Außerdem muss auch gelten, dass die Summe aller Ergebnisse (aus dem Ergebnisraum) gleich 1 ist.

$$\sum_{o} p(o) = 1$$

o ist ein Element des Ergebnisraums

Beispiel

- Gegeben ein Ergebnisraum: {sehr gut, gut, schlecht}
- Wir wollen eine Wk-Verteilung definieren. Wir geben also jedem Ergebnis eine Nummer (zwischen 0 und 1)
  p(sehr gut) - > 0.25
  p (gut) - > 0.5
  p(schlecht) - > 0.25

- Dazu müssen wir auch schauen, ob die Summe gleich 1 ist.
  - p(sehr gut) + p(gut) + p(schlecht) = 0.25+0.5 +0.25 = 1

Wahrscheinlichkeitsverteilung: Funktion, die jedem Ergebnis o einen Wert zwischen 0 und 1 zuweist, so dass

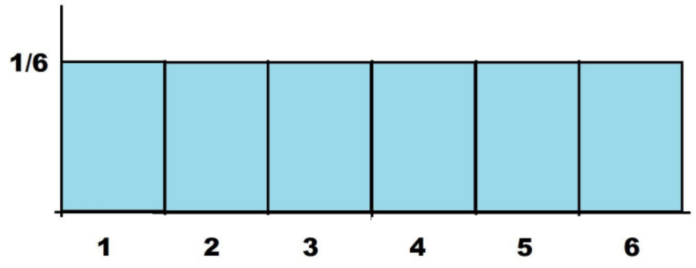
$$\sum_{o} p(o) = 1$$

Die Definition sagt nicht, wie man p(Ergebnis) berechnet soll

- In manchen Experimenten ist es leicht zu erkennen, was p ist. Z.B Wurf von fairen Münzen oder Würfel.
- In manchen schwieriger. Es gibt verschiedene Methoden, um p für bestimmte Anwendungsfälle zu definieren.

Wie ist die Wahrscheinlichkeitsverteilung beim Werfen einer Münze und beim Würfeln?

- Normalerweise können wir davon ausgehen, dass Münzen und Würfel fair sind (wenn nicht anders angegeben)
- D.h. alle Ergebnisse haben die gleiche Wahrscheinlichkeit (uniforme Verteilung)
  - Beim Münzwurf gibt es 2 Ergebnisse, der WK wird wie folgt berechnet.
    - p(outcome) = 1 / Anzahl der Ergebnisse
      - p(head) = p(tail) = 1 / 2
  - für das Würfel-Experiment haben wir 6 Ergebnisse, also
    - p(outcome) = 1/6 (für alle Ergebnisse)
- Wenn eine Münze p(head) = 0.8, p(tail) = 0.2 hat, ist diese Münze verzerrt (biased).



(uniforme Verteilung)

NLP-Beispiel

- Angenommen, wir haben einen Korpus mit 3-Grammen, wir wählen eine 3-Gramm-Sequenz aus und beobachten die Wörter in der Folge.

- Ein Ergebnis dieser Beobachtung könnte wie folgt aussehen: (the, dog, barks)
- Der Ergebnisraum enthält also alle möglichen 3-Gramme dieses Textes
    - S = { (the, dog, barks), (dog, barks, at), ...}
    - note: S steht für sample space

- Wir wollen die Wahrscheinlichkeit dafür berechnen, dass wir die Reihenfolge (the, man, eat) wählen/sehen. Wie können wir das machen?
    - Es gibt mehrere Methoden diese Wahrscheinlichkeit zu schätzen
    - Damit werden wir uns in späteren Kapiteln beschäftigen.

- **Anwendungsfall:** Wir berechnen die WK der N-Gramme, um ein Sprachmodell zu schätzen. Das Sprachmodell kann uns sagen, wie wahrscheinlich es ist, dass eine bestimmte Folge von Wörtern in einem Text vorkommt.

Die Wahrscheinlichkeit eines Ereignisses ist die Summe der Wahrscheinlichkeiten der entsprechenden Ergebnisse.

Wiederholung: Was sind noch mal **Ereignisse** und **Ergebnisse?**

Wichtige Begriffe

Zufallsexperiment: Experiment (Versuch) mit mehreren möglichen Ausgängen

Ergebnis: Resultat eines Experimentes

Ergebnisraum: Menge aller möglichen Ergebnisse

Ereignis: Teilmenge des Ergebnisraumes

Elementarereignis: anderes Wort für Ergebnis

Stichprobe: Folge von Ergebnissen bei einem wiederholten Experiment

Beispiel

Experiment (Prozess/Situation): wir werfen 2 Würfel
Observation: Wir beobachten die Zahl auf jedem Würfel

Ergebenis/Elementarereignis(outcome):
z.B. (1,3) oder (4,3)
wobei die erste Zahl die Nummer des ersten Würfels und die zweite Zahl die Nummer des zweiten Würfels ist.

Ergebnisraum (sample space):
{ (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)
(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) }

Ereignis(event):
A = {(1,2)(2,1)} -- Das Ereignis, bei dem die Summe der beiden Würfel gleich 3 ist.
B = { (1,1)(1,2)(1,3)(1,4)(1,5) (1,6) } -- Das Ereignis, bei dem der erste Würfel eine 1 ergibt.
C = { (5,6) } -- Das Ereignis, bei dem das Ergebnis (5,6) auftritt.

Stichprobe (sample):
- Ergebnisse, die wir nach Durchführung des Experiments tatsächlich erhalten.
- Wenn wir das Experiment zum Beispiel 3 Mal durchführen, erhalten wir (1,2),(3,2),(4,3)

Die Wahrscheinlichkeit eines Ereignisses ist die Summe der Wahrscheinlichkeiten der entsprechenden Ergebnisse.

- Was brauchen wir, um die WK eines Ereignisses zu berechen?
  - Wir müssen wissen, welche Ergebnisse in diesem Ereignis enthalten sind. Z.b. A = {2, 4, 6 }
  - Wir müssen die WK für jedes Ergebnis kennen, also p(2), p(4), p(6)

- Wir können dann P(A) berechnen
  - P(A) = p(2) + p(4) + p(6)

o steht für ein Ergebnis aus A

$$P(A) = \sum_{o \in A} p(o)$$

Recap:

- **Was is die „Wahrscheinlichkeitverteilung"?**
  - Funktion, die jedem Ergebnis eine Zahl zwischen 0 und 1 zuweißt
  - + die Summe der Wahrscheinlichkeit aller Ergebnisse p(o) muss gleich 1 sein

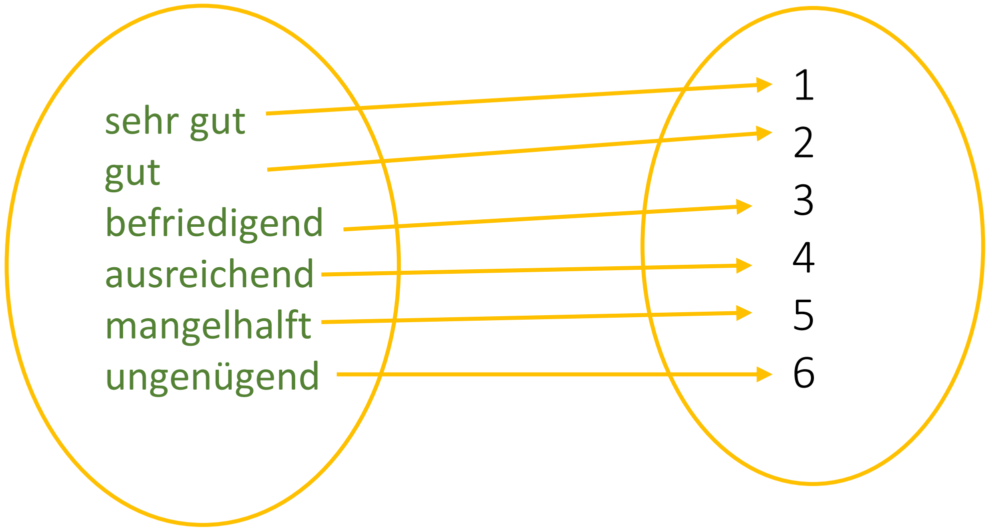
- **Wie berechnet man die Wahrscheinlichkeit eines Ereignisses?**
  - $A = \{1,2\}$
  - $P(A) = p(1) + p(2)$
  - alle Wahrscheinlichkeiten der entspechenden Ergebnisse summieren

- **Kennen wir immer die Wahrscheinlichkeit des Ergebnisses eines Zufallsexperiments?**
  - nein, aber mit bestimmten Angaben/Informationen kann die Wahrscheinlichkeit berechnet werden.
    - Wenn angegeben ist, dass die Münze fair ist, oder dass die Ergebnisse eines Experiments eine Uniformverteilung (oder eine bestimmte Art der Verteilung) haben, dann können wir die WK berechnen

Zufallsvariable: Funktion, welche jedem Ergebnis eine reelle Zahl zuweist.

Beispiel: Abbildung der Noten *sehr gut*, *gut*, *befriedigend*, *ausreichend*, *mangelhaft*, *ungenügend* auf die Zahlen 1, 2, 3, 4, 5, 6.

Y ist eine Zufallsvariable (ZV) mit folgender Zuordnung



Notation

X is a random variable that can take 6 values (1,2,3,...,6) .

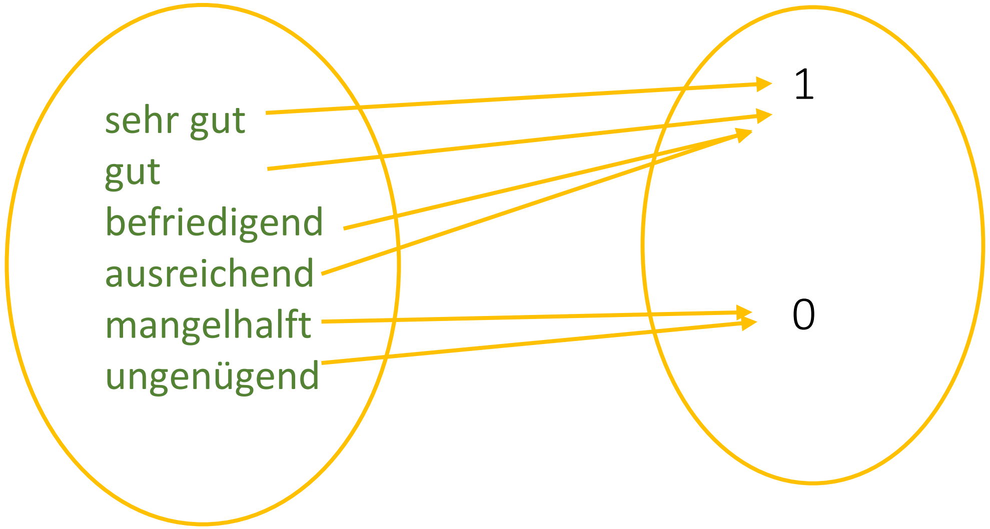
We refer to the value of a random variable with the notation ***X=value***

For example,
X=1 is the set { sehr gut}
X=2 is the set {gut}

Man kann eine andere ZV auch definieren, die eine andere Zuordnung hat.

Wir definieren hier B als eine Zufallsvariable, die den Noten, die als bestanden gelten, die Zahl 1 zuordnet und den Noten, die als nicht bestanden gelten, die Zahl 0 zuordnet.

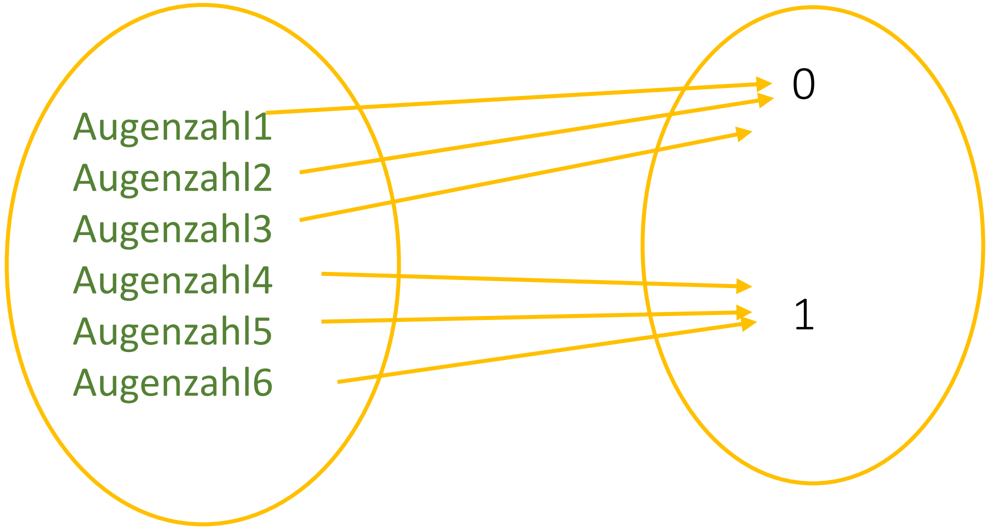
B ist eine ZV mit folgender Zuordnung



Der Wert 1 der Zufallsvariable B gruppiert die Ergebnisse, die als bestanden gelten.
X=1 : {sehr gut, gut, befriedigend, ausreichend }

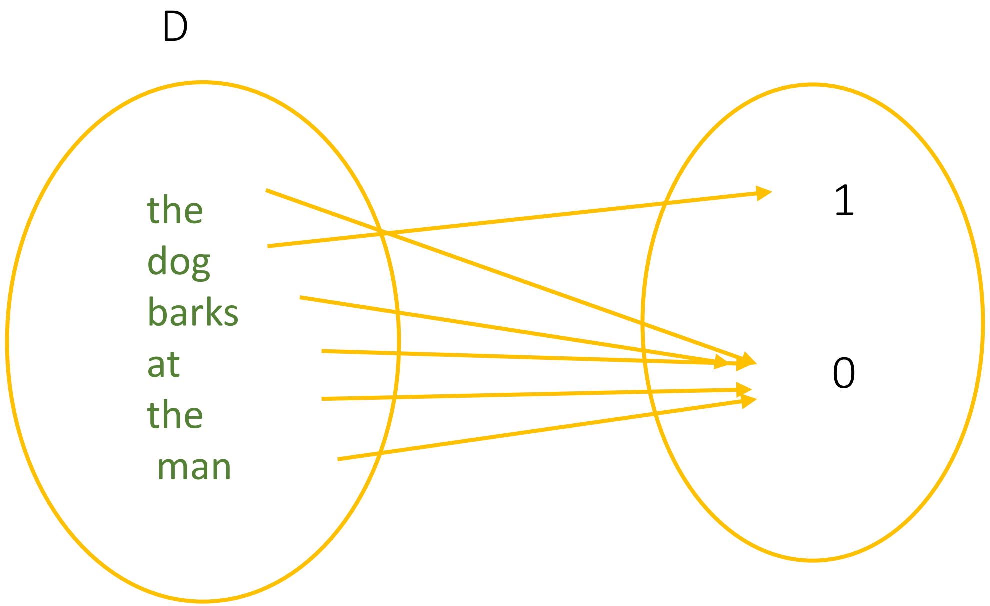
und der Wert 0 gruppiert die Ergebnisse, die als nicht bestanden gelten.
X=0: {mangelhaft, ungenügend}

C ist eine ZV mit folgender Zuordnung



wobei der Wert 1 der Zufallsvariable C die Ergebnisse gruppiert, die **größer als 3 sind**.
X=1 repräsentiert { AZ4, AZ5, AZ6}
(AZ: Augenzahl)

und der Wert 0 gruppiert die Ergebnisse, die nicht größer als 3 sind.
X=0 repräsentiert { AZ1, AZ2, AZ3}



D=1 steht für die Menge der Ergebnisse, in der " dog " vorkommt.

D=0 steht für die Menge der Ergebnisse, in denen andere Wörter vorkommen (die nicht „dog" sind).

Random Variables

A ***random variable***, usually written *X*, is a variable whose possible values are numerical outcomes of a random phenomenon. There are two types of random variables, ***discrete*** and ***continuous***.

Discrete Random Variables

A ***discrete random variable*** is one which may take on only a countable number of distinct values such as 0,1,2,3,4,........ Discrete random variables are usually (but not necessarily) counts. If a random variable can take only a finite number of distinct values, then it must be discrete. Examples of discrete random variables include the number of children in a family, the Friday night attendance at a cinema, the number of patients in a doctor's surgery, the number of defective light bulbs in a box of ten.

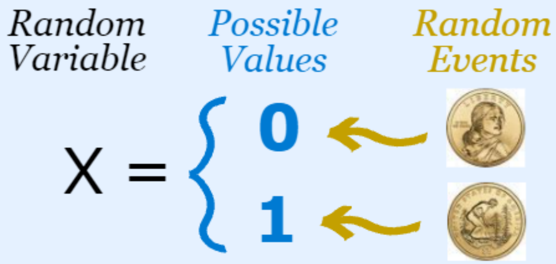
The ***probability distribution*** of a discrete random variable is a list of probabilities associated with each of its possible values. It is also sometimes called the probability function or the probability mass function.

Random Variables

A Random Variable is a set of **possible values** from a random experiment.

Example: Tossing a coin: we could get Heads or Tails.

Let's give them the values **Heads=0** and **Tails=1** and we have a Random Variable "X":



In short:

$$X = \{0, 1\}$$

Note: We could choose Heads=100 and Tails=150 or other values if we want! It is our choice.

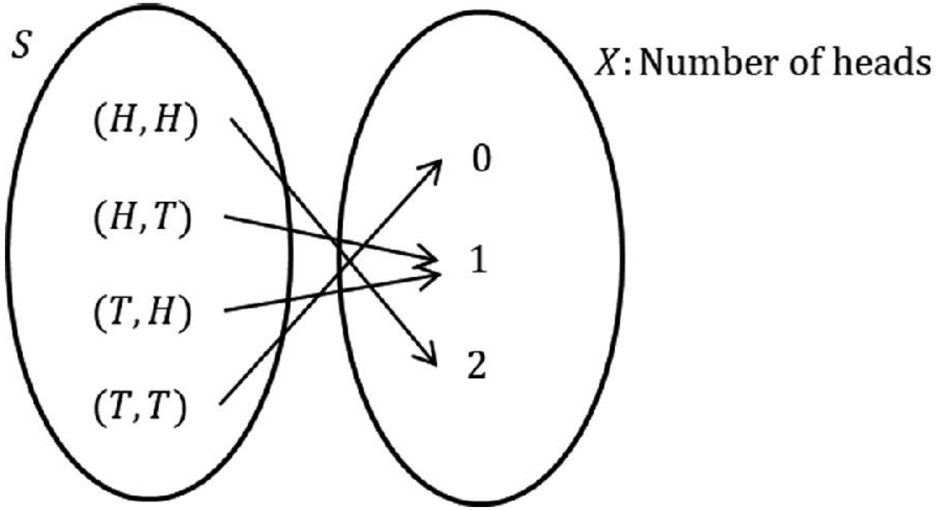
X ist eine Zufallsvariable, die zwei Werte annehmen kann, 0 und 1. Hier wird Head auf 0 und Tail auf 1 abgebildet.

Random Event Examples

Rolling a die is a random event and you can quantify (i.e. give a number to) the outcome. Let's say you wanted to know how many sixes you get if you roll the die a certain number of times. Your random variable, X could be equal to 1 if you get a six and 0 if you get any other number.
This is just an example; You can define X and Y however you like (i.e. 2 if you roll a six and 9 if you don't).

A few more example of random variables:

- X = total of lotto numbers.
- Y = number of open parking spaces in a parking lot.
- Z = number of aces in a card hand.

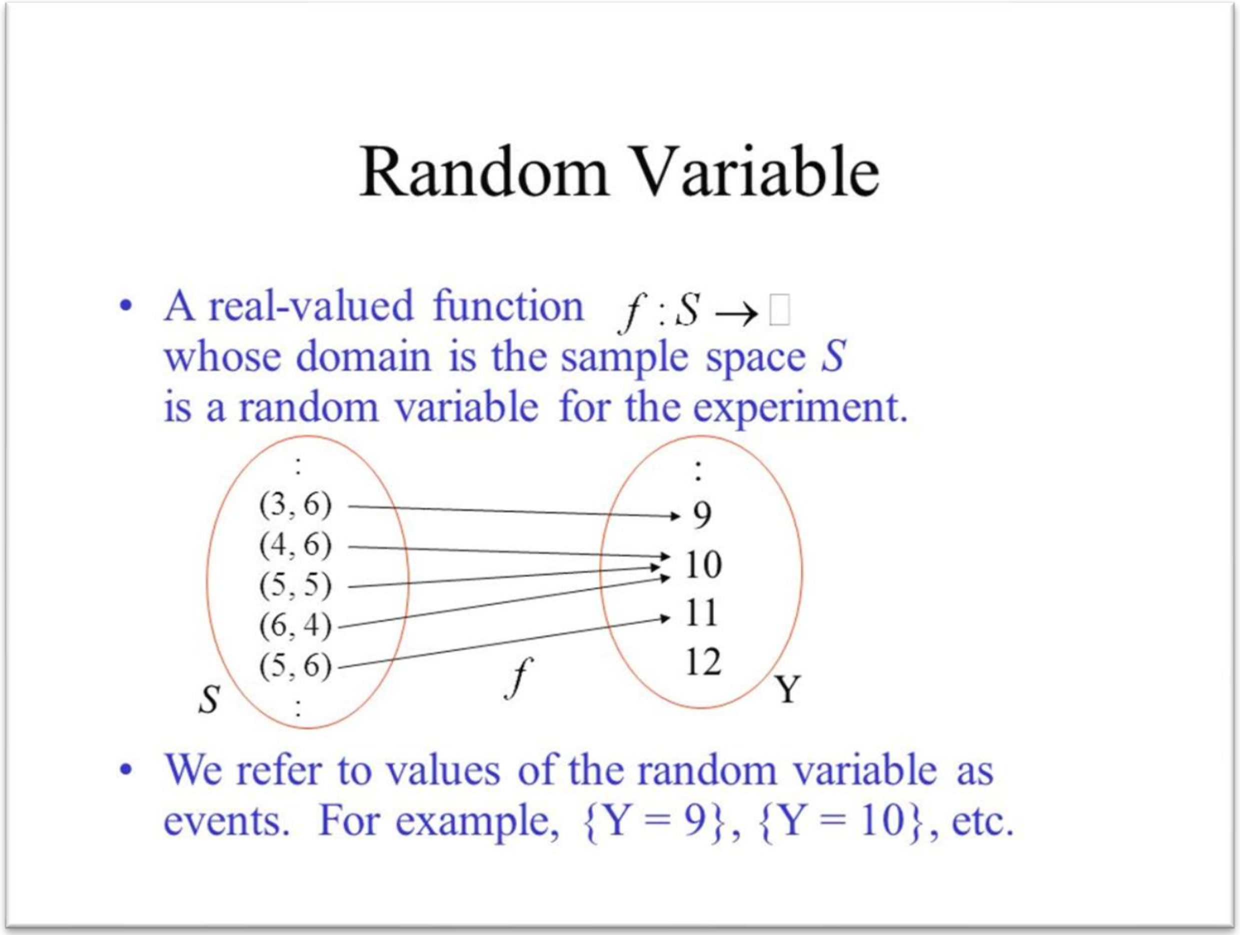


The concept of random variable is similar to the concept of event.

X=1 is a subset of the sample space, so it is an event.

We use X=1 as a way to refer to a group of particular outcomes.

We define **an event** also as a way to group particular outcomes.



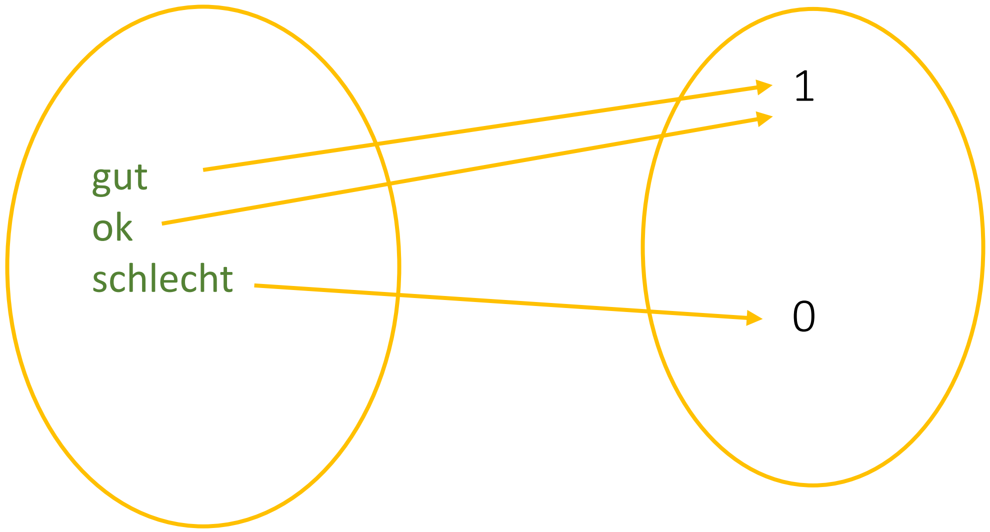
Wahrscheinlichkeit eines Wertes x der Zufallsvariablen X:

- Previously we have learned how to compute the probability of an event A.
- Now we want to compute the probability of a value x of a random variable X.
- The calculation will be very similar because the value of a random variable is also an event (it is a subset of the sample space).

we want to compute $P(X = x)$

$P(X = x)$

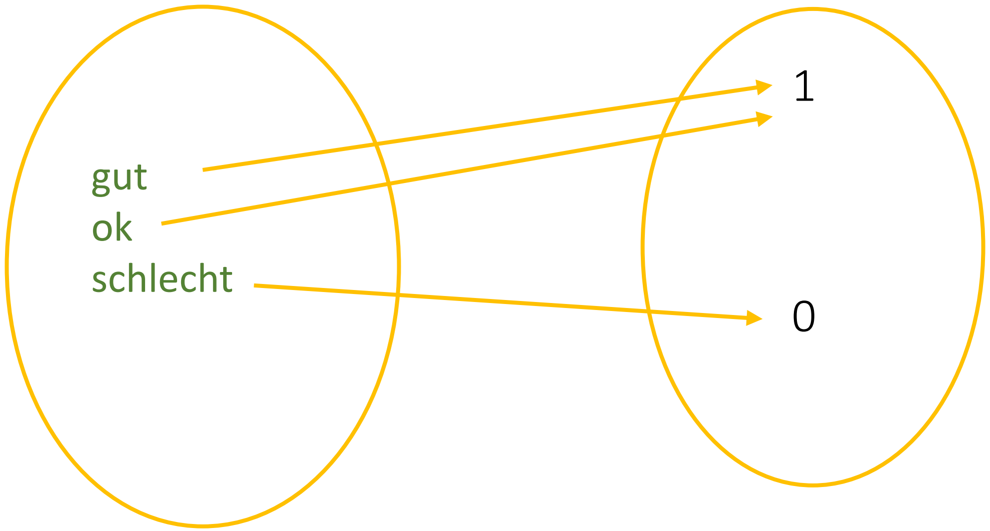
For example:
Given a random variable X with the following mapping, we want to compute P(X=1) and P(X=0)



Steps

- what is X=1 ?
  - it is the set { gut, ok}
  - this is an event and to compute the probability of an event we have to compute *the sum of each outcome in that event*
    - P(X=1) = p(gut) + p(ok)

- compute P(X= 0)
  - ......



Steps

- what is X=1 ?
  - it is the set { gut, ok}
  - this is an event and to compute the probability of an event we have to compute
    - the sum of each outcome in that event
    - P(X=1) = p(gut) + p(ok)
- compute P(X= 0)
  - P(X=0) = p(schlecht)

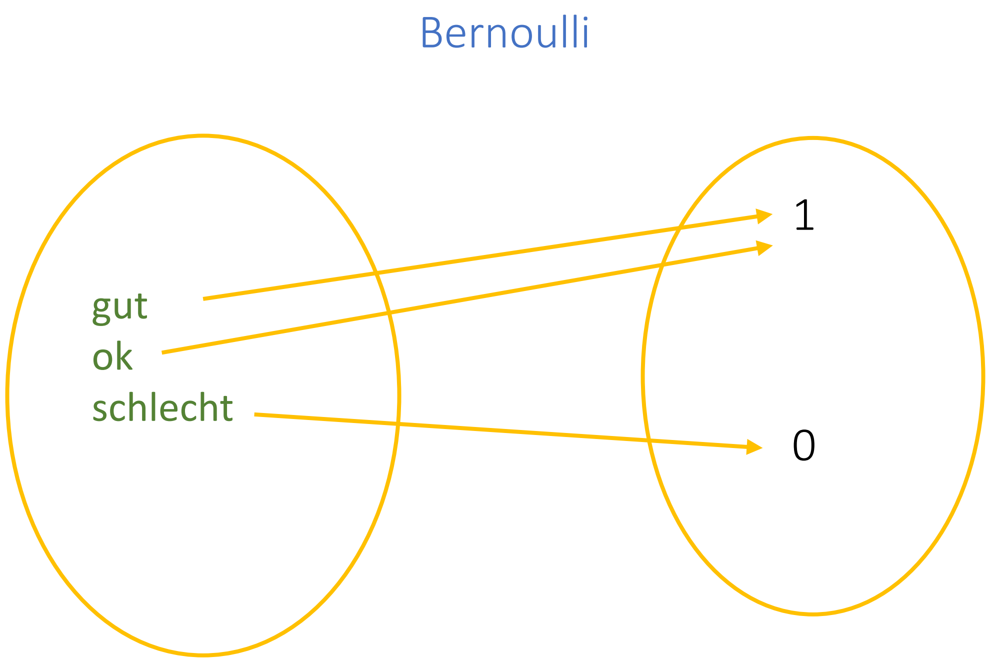
Notations

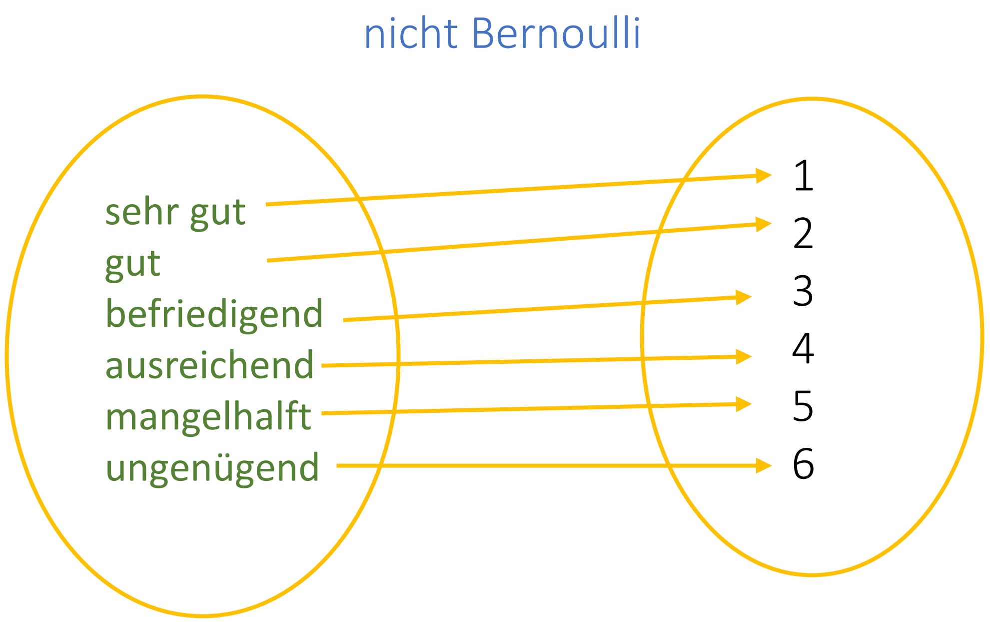
Wahrscheinlichkeit eines Wertes x der Zufallsvariablen X:

$P(X = x)$	Wertes x der Zufallsvariablen X
$p(x)$	$p(x)$ ist eine abkürzende Schreibweise, die oft verwendet wird, wenn klar ist, zu welcher Zufallsvariablen der Wert x gehört.
$P(A_x)$	A_x: Menge der Ergebnisse, für welche die Zufallsvariable X den Wert x liefert.

Die Notationen beziehen sich auf die gleiche Sache -> die Wahrscheinlichekeit einer Menge von Ergebnissen aus dem Ergebnisraum

Eine Zufallsvariable, die nur die Werte 0 und 1 liefert, nennt man **Bernoulli-Experiment**.





Recap:

- Wie ist die Wahrscheinlichkeitsverteilung(WK) definiert?
- Wie kann man die WK eines Ergebnisses p(Ergebnis) berechnen?
- Wie kann man die WK eines Ereignisses P(A) berechnen?
- Wie kann man die abstrakte Beschreibung von Ergebnissen in eine numerische Zahl umwandeln?
- Wie kann man die WK eines Wertes einer Zufallsvariablen berechnen?

Recap:

- Wie ist die Wahrscheinlichkeitsverteilung(WK) definiert?

  Wahrscheinlichkeitsverteilung: Funktion, die jedem Ergebnis o einen Wert zwischen 0 und 1 zuweist, so dass

  $$\sum_{o} p(o) = 1$$

- Wie kann man p(Ergebnis) berechnen?
  - Das hängt von dem Experiment ab...
    - Bei Münzen / Würfen kann man die WK schon abschätzen
    - bei andere Experimente nicht so einfach (das haben wir auch noch nicht gelernt)
- Wie kann man die WK eines Ereignisses berechnen?
    - Summe aller p(o) wobei o ein Ergebnis des Ereignisses ist
- Wie kann man die abstrakte Beschreibung von Ergebnissen in eine numerische Zahl umwandeln?
    - Man definiert eine Zufallsvariable
- Wie kann man die WK eines Wertes einer Zufallsvariablen berechnen?
    - Man muss zuerst wissen, wo für der Wert der ZV steht
    - Die Berechnung ist wie bei der Berechung der WK eines Ereignisses