Vorlesung: Statistik II für Wirtschaftswissenschaft

Prof. Dr. Helmut Küchenhoff

Institut für Statistik, LMU München

Sommersemester 2017

Axiome nach Kolmogorov

Gegeben sei ein Zufallsexperiment mit Ergebnisraum Ω (Menge der möglichen Ergebnisse)

Axiom 1

Jedem Ereignis A , $A \subset \Omega$ ist eine Wahrscheinlichkeit $P(A)$ zugeordnet, die Werte zwischen 0 und 1 annehmen kann:

$$0 \leq P(A) \leq 1.$$

Axiom 2

Das sichere Ereignis hat die Wahrscheinlichkeit 1:

$$P(\Omega) = 1.$$

Axiom 3

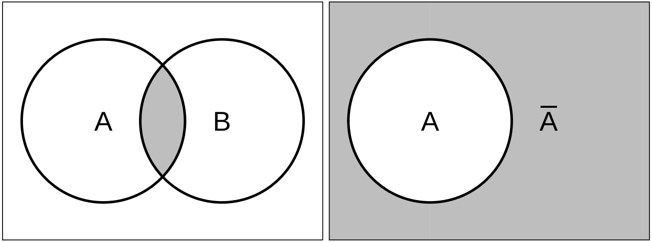
Sind A_1 und A_2 disjunkte Ereignisse, so ist

$$P(A_1 \cup A_2) = P(A_1) + P(A_2).$$

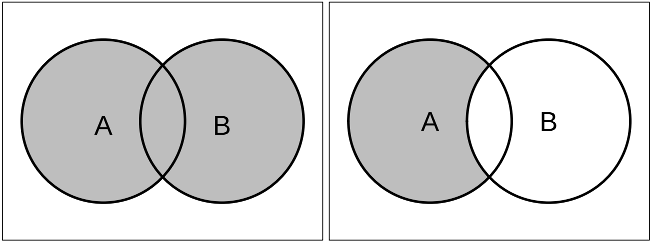
Venn Diagramme

Veranschaulichung von Wahrscheinlichkeiten durch Flächen :

Schnittmenge und Komplement:



Vereinigung und Differenz



Folgerungen

Folgerung 1

Die Wahrscheinlichkeit für das zu A komplementäre Ereignis $\bar{A}$ ist

$$P(\bar{A}) = 1 - P(A)$$

Beweis

$$\begin{aligned}
\text{Axiom 2:} \quad & P(\Omega) = 1 \\
\Leftrightarrow \quad & P(A \cup \bar{A}) = 1 \\
\stackrel{\text{Axiom 3}}{\Leftrightarrow} \quad & P(A) + P(\bar{A}) = 1 \\
\Leftrightarrow \quad & P(\bar{A}) = 1 - P(A)
\end{aligned}$$

Folgerungen

Folgerung 2

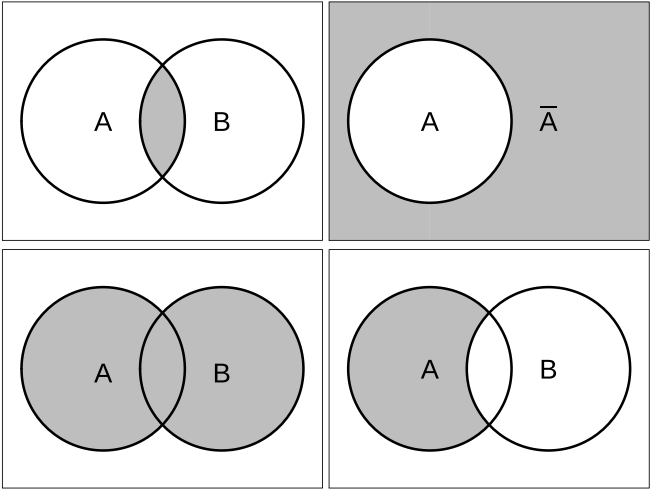
Die Wahrscheinlichkeit des unmöglichen Ereignisses $\emptyset$ ist

$$P(\emptyset) = 0$$

Beweis

$$P(\emptyset) = P(\bar{\Omega}) \overset{\text{Folgerung 1}}{=} 1 - P(\Omega) \overset{\text{Axiom 2}}{=} 0$$

Venn Diagramme



Folgerungen

Folgerung 3

Die Wahrscheinlichkeit, dass von zwei Ereignissen A_1 und A_2, die sich nicht notwendig gegenseitig ausschließen, mindestens eins eintritt, ist

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Beweis

$$\begin{aligned}
P(A_1 \cup A_2) &\overset{\substack{\text{disjunkte}\\ \text{Zerlegung}}}{=} P\left(A_1 \backslash A_2 \cup A_2 \backslash A_1 \cup (A_1 \cap A_2)\right) \\
&\overset{\text{Axiom 3}}{=} P(A_1 \backslash A_2) + P(A_2 \backslash A_1) + P(A_1 \cap A_2) \\
&\overset{\text{kreative 0}}{=} \underbrace{P(A_1 \backslash A_2) + P(A_1 \cap A_2)}_{P(A_1)} \\
&\quad + \underbrace{P(A_2 \backslash A_1) + P(A_1 \cap A_2)}_{P(A_2)} \\
&\quad - P(A_1 \cap A_2) \\
&= P(A_1) + P(A_2) - P(A_1 \cap A_2)
\end{aligned}$$

Folgerungen

Folgerung 4

Für $A \subseteq B$ gilt stets

$$P(A) \leq P(B)$$

Beweis

$$\begin{aligned}
& B & \overset{\text{disjunkte Zerlegung}}{=} & \; A \cup (\bar{A} \cap B) \\
\Rightarrow \; & P(B) & = & \; P(A \cup (\bar{A} \cap B)) \\
\overset{\text{Axiom 3}}{\Leftrightarrow} \; & P(B) & = & \; P(A) + \underbrace{P(\bar{A} \cap B)}_{\geq 0 \text{ (Axiom 1)}} \\
\Rightarrow \; & P(B) & \geq & \; P(A)
\end{aligned}$$

Folgerungen

Folgerung 5

Sei $A_1, ..., A_n$ eine vollständige Zerlegung des Ereignisraums Ω in paarweise disjunkte Ereignisse. Für ein beliebiges Ereignis B gilt dann

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i)$$

Zusammenfassung

Rechenregeln für Wahrscheinlichkeiten

- $0 \leq P(A) \leq 1$
- $P(\Omega) = 1$
- $P(\emptyset) = 0$
- $P(\bar{A}) = 1 - P(A)$
- $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$
- $P(A_1 \cup A_2) = P(A_1) + P(A_2)$, falls A_1 und A_2 disjunkt sind
- $P(B) = \sum_{i=1}^{n} P(B \cap A_i)$, falls A_i eine vollständige Zerlegung von Ω bilden

Laplacesche Wahrscheinlichkeit

Definition Laplacesche Wahrscheinlichkeit

Liegt ein Zufallsexperiment zugrunde, bei dem

- die Ergebnismenge *endlich* ist und
- alle Ergebnisse *gleichwahrscheinlich* sind,

dann bildet der Quotient aus

$$\frac{\text{Anzahl der für } A \text{ günstigen Fälle}}{\text{Anzahl aller möglichen Fälle}} = \frac{|A|}{|\Omega|} = P(A)$$

die Laplace-Wahrscheinlichkeit.

Die Mächtigkeiten $|A|$ und $|\Omega|$ können z.B. mit Hilfe von kombinatorischen Regeln bestimmt werden.

Ziehen aus einer Grundgesamtheit

Beispiel: Es wird ein Studierender der Vorlesung gezogen und nach seiner Wahlabsicht gefragt.

Dazu nehmen wir an, dass es N Studierende in der Vorlesung gibt und dass sie durchnummeriert sind $n = 1, ..., N$

$$P(\text{Student Nr n wird gezogen}) = 1/N$$

Alle haben die gleiche Ziehungswahrscheinlichkeit.

- Wie groß ist die Wahrscheinlichkeit, dass er/sie ein SPD Wähler ist?
- Wie groß ist die Wahrscheinlichkeit, dass eine Frau gezogen wird?

Lösung nach Laplace

- Wahrscheinlichkeit für „SPD-Wähler“

$$
\begin{aligned}
P(SPD) &= \frac{\text{Anzahl der für } SPD \text{ günstigen Ergebnisse}}{\text{Anzahl aller möglichen Ergebnisse}} \\
&= \frac{\text{Anzahl der SPD Wähler}}{\text{Anzahl aller Studierenden der Vorlesung}}
\end{aligned}
$$

  Die Wahrscheinlichkeit ist also die relative Häufigkeit f_{SPD} der SPD Wähler in der Grundgesamtheit.

- Wahrscheinlichkeit für Frau ?

Relative Häufigkeiten und Wahrscheinlichkeiten

Die Argumentation des Beispiels gilt ganz allgemein.

$$P(\text{Eine Person mit der Eigenschaft E wird gezogen }) = f_E$$

- Die relativen Häufigkeiten/Anteile aus der Grundgesamtheit pflanzen sich also in der entsprechenden Wahrscheinlichkeitsverteilung in der Stichprobe fort.
- Dies ist ganz entscheidend, denn dadurch kann man also durch eine Stichprobe etwas über die Häufigkeitsverhältnisse in der Grundgesamtheit lernen.

Zufallsstichproben

- Ziehung von mehreren n Einheiten aus der Grundgesamtheit
- Ziehung mit und ohne Zurücklegen
- Typischerweise sind Stichproben ohne Zurücklegen praktisch einfacher zu realisieren und zu rechtfertigen.
- Für sehr große Grundgesamtheiten sind die Unterschiede zwischen mit und ohne Zurücklegen verschwindend gering.

Die praktische Umsetzung:

- Mit Hilfe einer nummerierten Liste der Grundgesamtheit Hilfe von Computerprogrammen
- Ersatzmechanismen : Random dialing (Telefon), Random Walks etc.
- **Nicht aufs gerate Wohl**. (Ich spreche Leute an)

Ziehen mit Zurücklegen

- Grundgesamtheit mit N Zahlen $G = \{1, \dots, N\}$.
- Ziehe Stichprobe vom Umfang n **mit** Zurücklegen.
- Zur Beschreibung des Zufallsvorgangs müssen wir die Anzahl der potentiell möglichen Stichprobenergebnisse bestimmen (jede Stichprobe ist gleichwahrscheinlich).
- $\Omega = \{(\omega_1, \dots, \omega_n) | \omega_j \in \{1, \dots, N\}\}$, das selbe Element kann mehrfach vorkommen.
- $|\Omega| = \underbrace{N \cdot N \cdot \ldots \cdot N}_{n\text{-mal}} = N^n$, d.h. N^n potentiell mögliche Stichproben vom Umfang n.

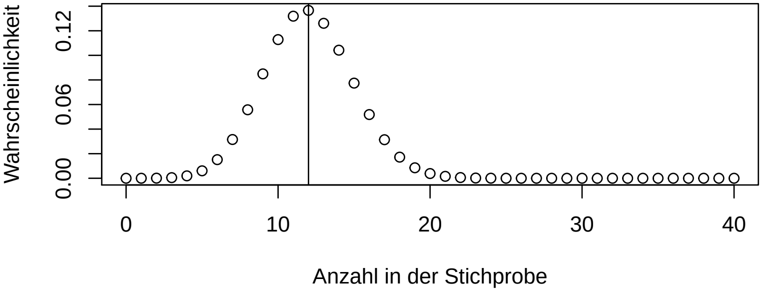
Beispiel: Stichprobentheorie

Ziehe Stichprobe vom Umfang n aus Grundgesamtheit von N=1000 mit Zurücklegen. Annahme: In Grundgesamtheit sind 300 SPD Wähler

$$
\begin{aligned}
n = 1 \qquad & P(1SPD) = 0.3 \\
n = 2 \qquad & P(0SPD) = \tfrac{700 \cdot 700}{1000 \cdot 1000} = 0.49 \\
& P(1SPD) = \tfrac{300 \cdot 700}{1000 \cdot 1000} \cdot 2 = 0.42 \\
& P(2SPD) = \tfrac{300 \cdot 300}{1000 \cdot 1000} = 0.09
\end{aligned}
$$

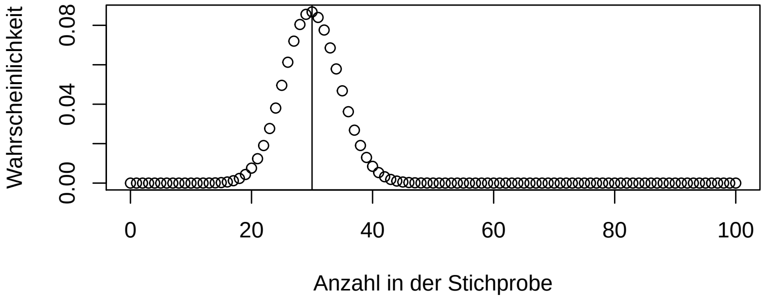
Beispiel: n=40

Ziehe Stichprobe vom Umfang n aus Grundgesamtheit von N=1000 mit Zurücklegen. Annahme: In Grundgesamtheit sind 300 SPD Wähler
Berechnung für große n mit Hilfe der Binomialverteilung



Beispiel: n=100

Ziehe Stichprobe vom Umfang n aus Grundgesamtheit von N=1000 mit Zurücklegen. Annahme: In Grundgesamtheit sind 300 SPD Wähler
Berechnung für große n mit Hilfe der Binomialverteilung



Einfache Zufallsstichprobe

Ziehen **ohne** Zurücklegen **ohne** Berücksichtigung der Reihenfolge

- Ziehe n Kugeln aus einer Urne mit N nummerierten Kugeln. Die Reihenfolge der Ziehungen spielt keine Rolle, d.h. die Stichprobe „4,1,7“ wird nicht unterschieden von „7,1,4“.
- $\Omega = \{\{\omega_1, \ldots, \omega_n\} : \omega_j \in \{1, \ldots, N\}, \omega_j \neq \omega_i \text{ für } j \neq i\}$
- Anzahl der Stichproben:

$$|\Omega| = \frac{N!}{(N-n)!n!} = \binom{N}{n}$$

Bedingte Wahrscheinlichkeit I

„Herzoperation in Krankenhaus“

Überleben der Operation

Alle Fälle	Operation überlebt	Operation nicht überlebt	P(nicht ü) „Risiko“
Krankenhaus U	500	500	0.5
Krankenhaus K	900	100	0.1

Frage: „In welchem Krankenhaus würden Sie sich behandeln lassen?“

Bedingte Wahrscheinlichkeit II

Schwere der behandelten Fälle

	schwere Fälle	leichte Fälle
Krankenhaus U	900	100
Krankenhaus K	100	900

Frage: „Bleiben Sie bei Ihrer Entscheidung?“

Bedingte Wahrscheinlichkeit III

Überleben der Operation aufgeteilt nach der Schwere der behandelten Fälle

Schwere Fälle	Operation überlebt	Operation nicht überlebt	P(nicht ü) „Risiko“
Krankenhaus U	400	500	0.56
Krankenhaus K	30	70	0.7

Leichte Fälle	Operation überlebt	Operation nicht überlebt	P(nicht ü) „Risiko“
Krankenhaus U	100	0	0
Krankenhaus K	870	30	0.033

Definition der bedingten Wahrscheinlichkeit

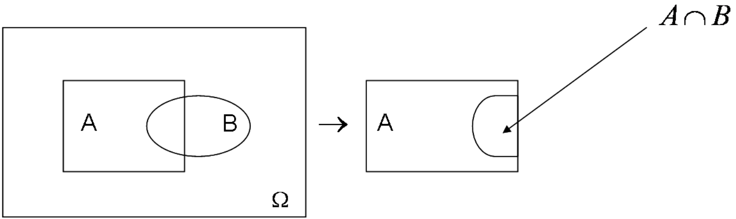
In dem Beispiel betrachten wir das Risiko gegeben „schwerer Fall“.
Das Risiko wird berechnet durch

$$\frac{\text{Anzahl (schwere Fälle und nicht überlebt)}}{\text{Anzahl(schwere Fälle)}}$$

Allgemein definieren wir die Wahrscheinlichkeit von
„Ereignis B gegeben A“

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

Einschränkung des Ergebnisraumes und bedingte Wahrscheinlichkeit



Bedingte Wahrscheinlichkeit: Beispiel

B: Nicht überleben
A: Schwerer Fall

Krankenhaus U

$$\begin{aligned} P(B) &= 500/1000 = 0.5 \\ P(A) &= 900/1000 = 0.9 \\ P(A \cap B) &= 500/1000 = 0.5 \\ P(B|A) &= 0.5/0.9 = 0.56 \end{aligned}$$

Krankenhaus K

$$\begin{aligned} P(B) &= 100/1000 = 0.1 \\ P(A) &= 100/1000 = 0.1 \\ P(A \cap B) &= 70/1000 = 0.07 \\ P(B|A) &= 0.07/0.1 = 0.7 = 70\% \end{aligned}$$

Schwere Fälle	OP überlebt	OP nicht überl.	P(nicht ü) „Risiko“
Krankenh U	400	500	0.56
Krankenh K	30	70	0.7

Leichte Fälle	OP überlebt	OP nicht überl.	P(nicht ü) „Risiko“
Krankenh U	100	0	0
Krankenh K	870	30	0.033

Beispiel: Würfeln

$$
\begin{array}{ll}
\Omega & = \{1,2,3,4,5,6\} \\
A & = \{2,4,6\} \quad \text{„gerade“} \\
B & = \{4,5,6\} \quad \text{„groß“} \\
A \cap B & = \{4,6\} \\
 & \\
P(A) & = 3/6 \\
P(A \cap B) & = 2/6 \\
P(B|A) & = P(A \cap B)/P(A) = (2/6)/(3/6) = 2/3
\end{array}
$$

Interpretation:
Wenn bekannt ist, dass die gewürfelte Zahl gerade ist, steigt die Wahrscheinlichkeit für „groß“ auf $2/3$.

Multiplikationssatz

Satz

Für zwei beliebige Ereignisse A und B gilt:

$$P(A \cap B) = P(A|B) \cdot P(B) = P(B|A) \cdot P(A).$$

Beweis

Nach Definition gilt:

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

$$\text{und} \quad P(A|B) = \frac{P(A \cap B)}{P(B)} \Leftrightarrow P(A|B) \cdot P(B) = P(A \cap B)$$

zusammen ergibt sich

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

Fußball Beispiel

Wie groß ist die Wahrscheinlichkeit, das Halbfinale zu gewinnen ?
Gesucht: $P(B)$ mit $B =$ „Sieg im Halbfinale“
Siegchancen sind abhängig vom jeweiligen Gegner!
$\Longrightarrow$ bedingte Wahrscheinlichkeiten.

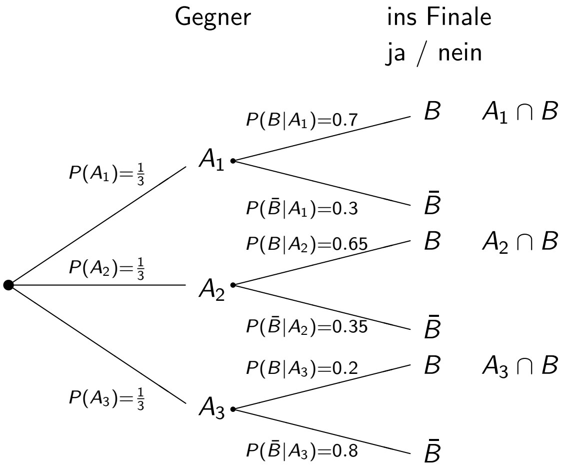
A_1	Gegner ist Mannschaft	1
A_2	"	2
A_3	"	3

Bedingte Wahrscheinlichkeiten leicht(er) anzugeben:

$$P(B|A_1) = 0.7$$
$$P(B|A_2) = 0.65$$
$$P(B|A_3) = 0.2$$

Gegner wird ausgelost $\Longrightarrow$ Annahme: $P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}$

Wahrscheinlichkeitsbaum (Fußball Beispiel)



Fußball Beispiel(2)

Welche „Wege“ im Wahrscheinlichkeitsbaum führen zu B?
Nutze Multiplikationssatz

$$\left.\begin{array}{rcl} P(A_1 \cap B) & = & P(A_1) \cdot P(B|A_1) = \frac{1}{3} \cdot 0.7 \\ P(A_2 \cap B) & = & P(A_2) \cdot P(B|A_2) = \frac{1}{3} \cdot 0.65 \\ P(A_3 \cap B) & = & P(A_3) \cdot P(B|A_3) = \frac{1}{3} \cdot 0.2 \end{array}\right\} \text{insgesamt: } 0.52$$

Verallgemeinerung: Vollständige Zerlegung

- A_1, A_2, A_3 bilden eine vollständige Zerlegung.
- $(A_1 \cap B)$, $(A_2 \cap B)$ und $(A_3 \cap B)$ sind disjunkt und ergeben in der Vereinigung B

Damit ergibt sich

$$
\begin{aligned}
P(B) &= P\big((A_1 \cap B) \cup (A_2 \cap B) \cup (A_3 \cap B)\big) \\
&= P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) \\
&= P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3) \; = \; 0.52
\end{aligned}
$$

Entlang der Äste multiplizieren, dann summieren

Satz von der totalen Wahrscheinlichkeit

Satz

Bilden die Ereignisse $A_1, ..., A_n$ eine *vollständige* Zerlegung von $\Omega = \bigcup_{i=1}^{n} A_i$ in paarweise disjunkte Ereignisse, so gilt für ein beliebiges Ereignis B:

$$P(B) = \sum_{i=1}^{n} P(B|A_i) \cdot P(A_i).$$

Satz von Bayes

Satz

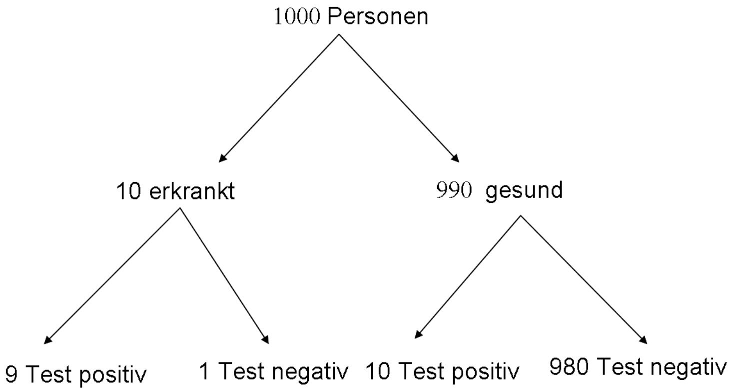
Für beliebige Ereignisse A und B mit $P(A), P(B) > 0$ gilt:

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

Bilden die A_i eine vollständige Zerlegung von Ω und ist B irgendein Ereignis, so gilt unter Zuhilfenahme des *Satzes von der totalen Wahrscheinlichkeit*:

$$P(A_j|B) = \frac{P(B|A_j) \cdot P(A_j)}{\sum_{i=1}^{n} P(B|A_i) \cdot P(A_i)}.$$

Medizinische Tests



Medizinische Tests 2

K:	Krank
G:	Gesund
TP:	test positiv
TN:	Test negativ

Gegeben:

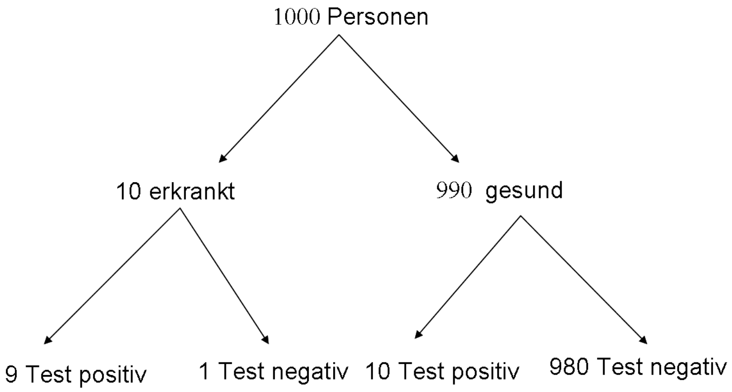
$$P(K) = 10/1000 = 0.01$$
$$P(TP|K) = 9/10 = 0.9$$
$$P(TP|G) = 10/990 = 0.0101$$

$$P(K|TP) = ???$$

Lösung mit Satz von Bayes

$$
\begin{aligned}
P(K|TP) &= \frac{P(K \cap TP)}{P(TP)} \\
&= \frac{P(TP|K) \cdot P(K)}{P(TP|K) \cdot P(K) + P(TP|G) \cdot P(G)} \\
&= \frac{0.9 \cdot 0.01}{0.9 \cdot 0.01 + 0.0101 \cdot 0.99} = 0.474
\end{aligned}
$$

Lösung mit Population



Beachte: Die Bedingung entspricht der Bezugspopulation 9 von 19 Patienten mit positivem Test sind tatsächlich krank:

$$P(K|TP) = 9/19 = 0.474$$

Unabhängigkeit

Definition stochastisch unabhängig

Zwei zufällige Ereignisse A und B heißen genau dann voneinander stochastisch unabhängig, wenn

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

gilt, d.h., wenn die Wahrscheinlichkeit für das gleichzeitige Eintreten von A und B gleich dem Produkt der beiden Einzelwahrscheinlichkeiten ist.

Folgerungen

- Sind zwei Ereignisse A und B unabhängig so folgt, dass das Eintreten des Ereignisses B keinen Einfluss auf das Eintreten von A hat, d.h. es gilt:

$$P(A|B) = P(A|\bar{B}) = P(A)$$

- Man kann unter der der Annahme der Unabhängigkeit Wahrscheinlichkeiten berechnen:
  A: Beim ersten Wurf 6
  B: Beim zweiten Wurf 6

$$P(A \cap B) = P(A)P(B) = 1/36$$