WAHRSCHEINLICHKEIT / EINIGE BEGRIFFE

***Ereignisraum Ω*.** Ω. ist die Menge aller möglichen Elementarereignisse

***Elementarereignis A*:** $A \in \Omega$, Element des Ereignisraumes

Beispiele

1. Man werfe einen Würfel: $\Omega = \{1,2,3,4,5,6\}$, 6 Elementarereignisse. Jedes Elementarereignis kommt gleich oft vor.
2. Werden 2 Würfel geworfen, dann ist der Ereignisraum ..$\Omega = \{(1,1), (1,2),.....(6,6)\}$ die Anzahl aller möglichen Kombinationen der Augenzahlen der beiden Würfel; Den gleichen Ergebnisraum erhält man, wenn 1 Würfel 2-mal hintereinander geworfen wird. In beiden Fällen gibt es 36 Elementarereignisse.

Für die Definition der Wahrscheinlichkeit ist nun wesentlich, was als Elementarereignis definiert wird.

Im gegebenen Beispiel mit X = 6, ist X die Ziffernsumme aus 2 Würfen oder 1 Wurf mit 2 Würfeln. Die Wahrscheinlichkeit p(X=6).ist dann,
..

$A = \{(1,5);(2,4);(3,3);(4,2);(5,1)\}$; 5 Paarungen erfüllen die Bedingung

$$p(X=6) = \frac{\text{Anzahl der Elementarereignisse mit } X = 6}{\text{Anzahl der Elemente von } \Omega}$$

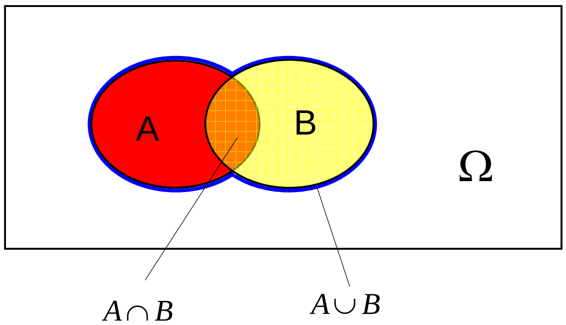
$$p(X=6) = \frac{5}{36}; \quad \text{allgemein } p(X=A) = \frac{|A|}{|\Omega|}$$

Bem: Die Anzahl der Elemente in einer Menge A wird mit dem Mächtigkeitssymbol | A | gekennzeichnet.

2. Zur persönlichen Kontemplation ein ähnlich gelagertes Problem von Galilei:
Wie groß ist die Wahrscheinlichkeit bei 3maligen Münzwurf die Ziffernsumme X = 10 zu erzielen?
$P(X=10) = ???$
Historisches: Spieler hatten entdeckt, dass beim Wurf mit 3 Würfeln die Augensumme 10 leichter zu erreichen ist als die Augensumme 9. Galilei (1564-1642) fand dafür die richtige Erklärung. Zeigen Sie durch Berechnung der Wahrscheinlichkeiten der beiden Augensummen, welch kleiner Unterschied durch die Spieler damals bemerkt worden war.

Ein paar Regeln für das Rechnen mit Wahrscheinlichkeiten

Venn Diagramm: Dient der graphischen Darstellung von Wahrscheinlichkeiten. In der Darstellung schließen sich A und B nicht gegenseitig aus !!



Regel 1: Für sich gegenseitig ausschließende Ereignisse gilt: Σ P(i) =1.

$P_{(1)} + P_{(2)} + ...P_{(n)} = 1$

Regel 2: Für beliebige Ereignisse A und B gilt die allgemeine Additionsregel,

$$P(A \text{ oder } B) = P(A) + P(B) - P(A \text{ und } B).$$

Bsp: Die Wahrscheinlichkeit, dass bei einer zufälligen Stichprobe aus der Bevölkerung eine Person männlich (*m*) oder älter als 30 Jahre (*A*>*30*) ist, wird demnach wie folgt berechnet:

$$P(m \vee (A > 30)) = P(m) + P(A>30) - P(m \wedge (A>30))$$

Regel 3: Bei sich gegenseitig ausschließenden Ereignissen A und B gilt die spezielle Additionsregel

$$P(A \, oder \, B) = P(A) + P(B)$$

Bsp: Wie groß ist die Wahrscheinlichkeit, dass in einer Stichprobe aus einer Rinderherde ein Exemplar weiblich (*w*) oder ein Jungstier mit weniger als 2 Jahren ist (*J* < *2*). In der Herde mit 120 Stück gibt es 85 weibliche Rinder und 21 Jungstiere, die maximal 2 Jahre alt sind.

$$P(w \vee (J<2)) = P(w) + P(J<2).$$

Regel 4: Häufig ist die Bestimmung der Gegenwahrscheinlichkeit $P(\overline{A})$ des Ereignisses A leichter als die Bestimmung von $P(A)$. Dann kann die Subtraktionsregel angewendet werden: Z. B. $P(A) = 1 - P(\overline{A})$

Regel 5: Für unabhängige Ereignisse gilt die spezielle Multiplikationsregel. Unabhängig bedeutet, dass A keinen Einfluß auf B hat und umgekehrt. Z. B. ist die Wahrscheinlichkeit bei 3 Würfen 3mal hintereinander eine "6" zu werfen gleich 1/6*1/6 *1/6 = 1/216

$$P(A\,und\;B) = P(A)\cdot P(B)$$

Bedingte Wahrscheinlichkeit (Bayes Statistik)

Die grundlegende Fragestellung lautet: Wie hoch ist die Wahrscheinlichkeit für das Eintreten des Ereignisses A, unter der Voraussetzung, dass B bereits eingetreten ist. So ist auch die Schreibweise zu lesen: $P(A \mid B)$.
P(A) wird als a priori Wahrscheinlichkeit und P(A|B), die bedingte Wahrscheinlichkeit, als a-posteriori-Wahrscheinlichkeit bezeichnet.

$$P(A \mid B) = \frac{P(A\;und\;\;B))}{P(B)}$$

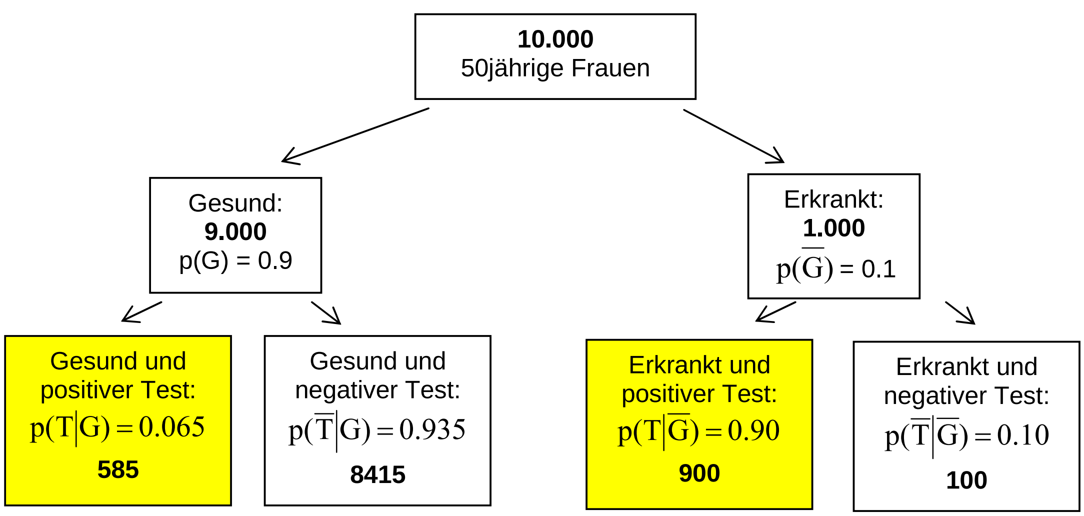
oder, zwar anders geschrieben, aber mathematisch ident:

$$P(A|B) = \frac{P(A)\cdot P(B \mid A)}{P(A)\cdot P(B \mid A) + P(\overline{A})\cdot P(B \mid \overline{A})}$$

Diese Formel ist vielleicht etwas unanschaulich und verwirrend. Sie wird sicher klarer, wenn man sich den Sachverhalt einmal anhand eines numerischen Beispieles vor Augen führt. (Im Buch "Das Einmaleins der Skepsis von Gerd Gigerenzer gibt es dazu sehr gute Erläuterungen). Nehmen wir zunächst ein aktuelles Beispiel:

> *In der Ausgabe 4/2004 der Zeitschrift Forum Gesundheit der Salzburger Gebietskrankenkasse werden in einem Artikel über Brustkrebs-Screening dessen Vorteile in Frage gestellt. Die Kritik beruft sich dabei auf Fakten, nach denen die Spezifität (Richtig-negativ-Rate) der Mammographie bei 93,5% und die Sensitivität (Richtig-positiv-Rate) bei 90% liegt. Weiters ist bekannt, dass bei 50 jährigen Frauen statistisch bei 1 von 10 ein Brustkrebs erwartet wird. Wie hoch ist nun der Anteil der positiv getesteten Frauen bei einem Massen-Screening, und die Wahrscheinlichkeit, dass bei einem positiven Test Brustkrebs tatsächlich vorliegt?*

Nehmen wir an, in einem Massen-Screening werden 10.000 50-jährige Frauen getestet, dann sieht der darauf aufbauende Entscheidungsbaum folgendermassen aus:



Es werden demnach insgesamt 900 + 585 = 1485 Frauen positiv getestet. Von diesen sind aber nur 900 tatsächlich an Brustkrebs erkrankt.

In der oben verwendeten Notation hieße dies:

$p(G) = 0.9$	Wahrscheinlichkeit, gesund zu sein
$p(\overline{G}) = 0.1$	W. erkrankt zu sein, bzw. nicht gesund zu sein
$p(\overline{T}\mid\overline{G}) = 0.10$	W. für ein negatives Testergebnis bei Kranken (Nicht Gesunden)
$p(T\mid G) = 0.065$	W. für ein positives Testergebnis bei Gesunden (Falsch positiv)
$p(T\mid\overline{G}) = 0.90$	W. für ein positives Testergebnis bei Kranken (Sensitivität = richtig positiv)
$p(\overline{T}\mid G) = 0.935$	Wahrscheinlichkeit für negatives Testergebnis bei Kranken (Spezifität = richtig negativ)

Die Wahrscheinlichkeit, dass es sich bei den positiven getesteten Frauen um eine tatsächlich an Brustkrebs erkrankte Frau handelt ist demnach nicht 100% sondern nur 900/(585 +900) = 60,6% !! Massen-Screenings sind daher mit einer gebotenen Skepsis zu bewerten, die sich, wie in dem gerade gezeigten Beispiel, quantitativ auf die Angaben der Spezifität, der Sensitivität und Angaben über die tatsächliche Erkrankungshäufigkeit bzw. deren Schätzung stützen muß.

Die Wahrscheinlichkeit, dass ein positives Testergebnis von einer gesunden Person stammt ist daher:

$$P(G/T) = \frac{P(T/G)P(G)}{P(T/G)P(G) + P(T/\overline{G})P(\overline{G})} = \frac{0.065 \cdot 0.9}{0.065 \cdot 0.9 + 0.90 \cdot 0.10} = 39.4\%$$

Ein weiteres Beispiel: Wie hoch ist die Wahrscheinlichkeit, bei einem zweiten Wurf mit einem Würfel eine "6" zu würfeln, wenn der erste Wurf auch eine "6" war.

$$P(x_2 = 6 \mid x_1 = 6) = \frac{P((x_1 \wedge x_2) = 6)}{P(x_1 = 6)} = \frac{1/6 \cdot 1/6}{1/6} = \frac{1}{6}$$

Das zweite Ereignis ist vom ersten Ereignis unabhängig !!

KOMBINATORIK

a) Permutationen

ohne Wiederholung:

$$P_n = n! \text{ (Faktorielle)}$$

Bsp: wie viele Anordungsmöglichkeiten der Elemente a,b,c, gibt es ?
Antwort: 3! = 6 verschiedene Möglichkeiten (abc,acb,bac,bca,cab,cba)

mit Wiederholung:
wenn von n Elementen r verschieden sind , $r \leq n$. Einzelne oder alle Elemente kommen mehr als einmal vor.

$${}_wP_n = \frac{n!}{r_1!.r_2!....r_n!}$$

Bsp.: Wie viele sechsstellige Zahlen sind mit den Ziffern {1,1,1,2,2,3} möglich?
A: ${}_wP_n = 6!/(2!.3!.1!) = 60$

b) Kombinationen

Ohne Wiederholung

$$C_k^n = \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

..entspricht der Auswahl von Klassen mit k Elementen aus n Elementen. (Vgl. Binomialkoeffizienten, Pascalsches Dreieck)

Bsp: Wie viele Klassen mit 2 Elementen können aus den 3 Elementen a,b,c, gebildet werden?
A: Es gibt 3 Möglichkeiten $C_2^3 = \binom{3}{2} = \frac{3!}{2!.1!}$, diese sind: ab, ac, bc

Mit Wiederholung
...wenn sich einzelne Elemente bis zu k-mal in einer Kombination wiederholen können.
Bsp: Aus den Elementen a,b, (n = 2, Anzahl der für die Auswahl zur Verfügung stehenden Elemente) können folgende Kombination zu Klassen mit 3 Elementen gebildet werden: aaa,aab,abb,bbb.

$${}_wC_k^n = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}$$

c) Variationen

Bei Variationen ist auch die Reihenfolge relevant, im Unterschied zu den Kombinationen, bei denen die Reihenfolge keine Rolle spielt.
Fragestellung:
Wie viele Möglichkeiten gibt es.
k Elemente aus der Ausgangsmenge mit n Elementen auszuwählen und in einer Reihenfolge anzuordnen

a) Variationen ohne Wiederholung:

$$V_k^n = \binom{n}{k} k! = \frac{(n)!}{(n-k)!} \wedge$$

1. Alle n Elemente der Ausgangsmenge unterscheiden sich voneinander.
2. Es werden einige k Elemente ausgewählt.
3. Ein Element **kann nicht mehrmals** ausgewählt werden.

Beispiel: Wie viele Möglichkeiten (Variationen) gibt es aus einer Pesonengruppe von 10 Personen 3 auszuwählen und diese in unterschiedliche Reihenfolge zu bringen ?

Es gibt insgesamt 10!(3! (10-3)!) 3er Kombinationen aus 10 Personen. Diese können in sich jeweils mit 3! permutiert werden, es gibt daher 10!/(10-3)! = 720 Variationen ohne Wiederholung, denn jede Person kann nur einmal vorkommen.

b) Variationen mit Wiederholung

$${}_wV_k^n = n^k$$

1. Alle n Elemente der Ausgangsmenge unterscheiden sich voneinander.
2. Es werden einige k Elemente ausgewählt.
3. Ein Element **kann mehrmals** ausgewählt werden.

Beispiel:
Toto spielen; Es gibt insgesamt 12 Tipps, Für jeden Tipp gibt es 3 Möglichkeiten. Es können daher 3^{12} unterschiedliche (Tipps) Variationen gebildet werden.

Im Dezimalsystem gibt es für eine 4-stellige Zahl 10^4 Variationen. k entspricht der Anzahl der Stellen, der Dezimalzahl, n – die Anzahl der Elemente des Ziffernvorrats von 0 bis 9 - ist die Basis des Zahlensystems.

Ein Zahlenschloss mit 3 Stellen hat demnach wie viele unterschiedliche Einstellungsmöglichkeiten? (Im täglichen Sprachgebrauch als Ziffernkombinationen bezeichnet, eigentlich handelt es sich dabei aber um Variationen mit Wiederholung – wie sie nun wissen)