Schriftliche Übung Mathematik Stochastik I **Di 22.11.11**
SG29D **Lösungen**

1.) In einem Land der Dritten Welt leiden 1% der Menschen an einer bestimmten Infektionskrankheit. Ein Test zeigt die Krankheit bei den tatsächlich erkrankten zu 98% korrekt an. Leider zeigt der Test auch 3% der Gesunden als erkrankt an. Folgende Vierfeldtafel veranschaulicht die Zusammenhänge.

	T (positiv)	$\overline{T}$ (negativ)	
K (krank)	0,0098	0,0002	0,01
$\overline{K}$ (gesund)	0,0297	0,9603	0,99
	0,0395	0,9605	1

K : Die getestete Person ist krank
$\overline{K}$: Die getestete Person ist gesund
T : Das Testergebnis ist positiv
$\overline{T}$: Das Testergebnis ist negativ

Berechnen Sie: $P_T(K)$ und $P_K(\overline{T})$

Erklären Sie die Bedeutung der Ergebnisse und kommentieren Sie diese in Bezug auf die Güte des Testverfahrens.

$$P_T(K) = \frac{P(T \cap K)}{P(T)} = \frac{0{,}0098}{0{,}0395} \approx 0{,}248$$

Eine Person, von der man weiß, dass das Testergebnis positiv ist, ist mit einer Wahrscheinlichkeit von 0,248 auch tatsächlich krank.
Das Ergebnis von 0,248 ist nicht zufriedenstellend.
Nur etwa 25% aller positiv getesteten sind tatsächlich erkrankt.
Das bedeutet andersherum, dass ca. 75% der positiv getesteten Personen gesund sind. Es wäre wünschenswert, dass der Test verbessert wird.

$$P_K(\overline{T}) = \frac{P(K \cap \overline{T})}{P(K)} = \frac{0{,}0002}{0{,}01} \approx 0{,}02$$

Eine Person, von der man weiß, dass sie erkrankt ist, wird mit einer Wahrscheinlichkeit von 0,02 als gesund getestet.
Nur 2 % aller erkrankten Personen werden als gesund getestet.
Das Ergebnis ist zufriedenstellend.

2.) Viele Internetnutzer klagen über Spam- Mails.
Nehmen wir an, in 1% der guten und 40% der Spam- Mails komme das Wort " Viagra" vor.
Außerdem seien 10% der Mails gut und 90% Spam.
Wie groß ist die Wahrscheinlichkeit, dass eine Mail, von der man weiß, das in ihr das Wort "Viagra" vorkommt, eine Spam- Mail ist.?

Ereignisse :

A : Mail enthält das Wort Viagra — $\overline{A}$: Mail enthält nicht das Wort Viagra

B : Spam-Mail — $\overline{B}$: gute Mail

1. Aufstellen der Vierfeldtafel mit den vorgegebenen Daten.
   Die % Werte entsprechen relativen Häufigkeiten (Wahrscheinlichkeiten)
   90 % Spam bedutet Summe Spam = 0,9
   10% gute Mails bedeutet Summe gute Mails = 0,1
   40% der Spam-Mails mit Viagra bedeutet 0,9 x 0,4 = 0,36
   1% der guten Mails mit Viagra bedeutet =0,1 x 0,01 = 0,001
   Die restliche Werte kann man ausrechnen, da die Summen bekannt sind.

	B : Spam-Mail	$\overline{B}$: Gute Mail	Summe
A : mit Viagra	0,36	0,001	
$\overline{A}$: ohne Viagra			
Summe	0,9	0,1	1

Spam ohne Viagra: 0,9 – 0,36 = 0,54
Gute Mail ohne Viagra: 0,1 – 0,001 = 0,099
Summe aller Mails mit Viagra: 0,36 + 0,001 = 0,361
Summe aller Mails ohne Viagra: 0,54 + 0,099 = 0,639

Mit diesen Werten wird die Vierfeldtafel nun vervollständigt.

	B : Spam-Mail	$\overline{B}$: Gute Mail	Summe
A : mit Viagra	0,36	0,001	0,361
$\overline{A}$: ohne Viagra	0,54	0,099	0,639
Summe	0,9	0,1	1

Die Aufgabenstellung lautete:
Wie groß ist die Wahrscheinlichkeit, dass eine Mail, in der "Viagra" steht, Spam ist?
Gesucht ist also die Wahrscheinlichkeit von B unter der Bedingung, dass A eingetreten ist.

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

mit $P(A \cap B) = 0{,}36$ und $P(A) = 0{,}361$ ist $P_A(B) = \dfrac{0{,}36}{0{,}361} \approx \underline{\underline{0{,}997}}$

Das bedeutet, in 99,7% aller Fälle ist eine Mail, von der man weiß, das in ihr das Wort Viagra steht, eine Spam- Mail.

3.) Wann sind zwei Ereignisse A und B unabhängig voneinander?
Überprüfen Sie in Aufgabe **2.)** die Ereignisse A und B auf Unabhängigkeit.

Unabhängige Ereignisse	Das Ereignis B heißt unabhängig vom Ereignis A, wenn das Eintreten von A die Wahrscheinlichkeit für das Eintreten von B nicht beeinflusst. Es gilt: $P_A(B) = P(B)$ mit $P_A(B) = \frac{P(A \cap B)}{P(A)}$ Beispiel: Urnenziehung mit Zurücklegen.

$P_A(B) = P(B)$?

Aus obiger Rechnung:

$$P_A(B) = \frac{P(A \cap B)}{P(A)} = \frac{0,36}{0,361} \approx 0,997 \neq 0,9 = P(B)$$

Das Ereignis B ist statistisch abhängig vom Ereignis A da $P_A(B) \neq P(B)$