ETH Zürich
Dept. Informatik

Prof. Markus Gross
WS 05/06

Informationstheorie

Übung 1

Ausgabe: 31. Oktober 2005
Abgabe: 7. November 2005

1.1 Erwartungswerte

Gegeben sei ein unfairer Würfel mit der folgenden Wahrscheinlichkeitsverteilung:

X	1	2	3	4	5	6
P_X	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

Berechnen Sie für $Y_1 := X$, $Y_2 := X^2$ und $Y_3 := P_X(X)$ die Erwartungswerte.

1.2 Bedingte Wahrscheinlichkeiten

Wir betrachten ein Testverfahren für eine Krankheit, an der eine Person mit Wahrscheinlichkeit $p_K = 0.00001$ erkrankt, d.h. es tritt durchschnittlich ein Erkrankter auf 100.000 Einwohner auf. Das Verfahren hat eine Fehlerwahrscheinlichkeit von $p_F = 0.01$, d.h. das Testergebnis ist mit Wahrscheinlichkeit 0.99 korrekt (sowohl im positiven wie im negativen Fall).

a) Alice lässt sich testen; der Test fällt positiv aus. Mit welcher Wahrscheinlichkeit leidet Alice trotz des Testergebnisses nicht an der Krankheit? Vergleichen Sie mit der A-priori-Wahrscheinlichkeit.

b) Bobs Test fällt negativ aus. Mit welcher Wahrscheinlichkeit ist Bob dennoch von der Krankheit betroffen? Vergleichen Sie mit der A-priori-Wahrscheinlichkeit.

1.3 Gemeinsame Wahrscheinlichkeitsverteilungen

Die folgende Tabelle zeigt die gemeinsame Wahrscheinlichkeitsverteilung der Zufallsvariablen K = "Wetter in Kanada" und S = "Wetter in der Schweiz". Beide Zufallsvariablen können folgende Werte annehmen: h= "heiss", w= "warm", k= "kühl" und e= "eiskalt".

Die folgende Tabelle gibt die gemeinsame Verteilung an:

		S			
	P_{KS}	h	w	k	e
	h	0.17	0.12	0.03	0
K	w	0.04	0.09	0.06	0.01
	k	0.01	0.06	0.10	0.03
	e	0	0.03	0.11	0.14

a) Ist P_{KS} eine Wahrscheinlichkeitsverteilung?

b) Berechne die Verteilungen P_S und P_K.

c) Sind die *Ereignisse* $K = k$ und $S = w$ und die *Zufallsvariablen* K und S unabhängig?

d) Wie gross ist die Wahrscheinlichkeit, dass es in Kanada warm ist, gegeben dass es in der Schweiz heiss ist?

e) Angenommen Sie mögen warmes Wetter und wissen, dass es nicht kalt sein wird in Kanada. Wo sollten Sie ihre Ferien verbringen (Kanada oder Schweiz)?

1.4 Kenngrössen bei Summen von Zufallsvariablen

a) Zeigen Sie, dass für den Erwartungswert $E[X_1 + X_2]$ der Summe von 2 Zufallsvariablen X_1 und X_2 gilt:

$$E[X_1 + X_2] = E[X_1] + E[X_2].$$

b) Zeigen Sie, dass für die Varianz $\mathrm{Var}[X_1 + X_2]$ der Summe von 2 unabhängigen Zufallsvariablen X_1 und X_2 gilt:

$$\mathrm{Var}[X_1 + X_2] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2].$$

1.5 Ausblick Codierung von Information

Bekanntlich können Buchstaben durch Bitfolgen repräsentiert werden. Zum Beispiel ist es möglich, die acht Buchstaben a bis h mit jeweils drei Bits darzustellen. Dies ist aber nicht die einzige Möglichkeit, insbesonders, wenn wir auch verschiedene Bit-String-Längen zulassen. Eine solche Repräsentation macht zum Beispiel dann Sinn, wenn der Buchstabe a im Text viel häufiger als h vorkommt.

$$\begin{array}{ll}
a \leftarrow 1 & b \leftarrow 01 \\
c \leftarrow 001 & d \leftarrow 0001 \\
e \leftarrow 00001 & f \leftarrow 000001 \\
g \leftarrow 0000001 & h \leftarrow 00000001
\end{array}$$

Gegeben seien nun drei Texte (die nur Buchstaben a bis h enthalten) und die alle gleich lang sind. Die relativen Häufigkeiten der Buchstaben in den drei Texten ist durch folgende Tabellen gegeben:

P_{T_1}	
a	0.125
b	0.125
c	0.125
d	0.125
e	0.125
f	0.125
g	0.125
h	0.125

P_{T_2}	
a	0.65
b	0.05
c	0.05
d	0.05
e	0.05
f	0.05
g	0.05
h	0.05

P_{T_3}	
a	0.5
b	0.25
c	0.15
d	0.05
e	0.03
f	0.01
g	0.01
h	0

Versuchen Sie ein Gefühl dafür zu entwickeln, welcher Text sich am stärksten komprimieren lässt.