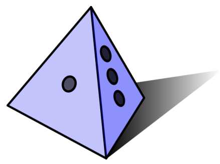
Dr. Jürg M. Stettbacher
Neugutstrasse 54
CH-8600 Dübendorf

Telefon: +41 43 299 57 23
Email: dsp@stettbacher.ch

Übung

Information und Entropie



Ein Würfel mit dreieckigen Seiten hat vier Flächen mit den Augenzahlen 1 bis 4. In einer Datenquelle, die zweistellige Zufallszahlen erzeugt, werden gleichzeitig zwei derartige Würfel geworfen, beispielsweise ein roter und ein blauer. Aus jedem Wurf entsteht eine Zufallszahl. Die Augenzahlen der beiden Würfel bezeichnen wir mit a und b, wobei $a \leq b$ sei. In jeder Zufallszahl S am Ausgang der Quelle[1] ist a die Zehnerziffer und b die Einerziffer.

1. Bestimmen Sie alle möglichen Ereignisse s_k der Datenquelle ($k = 1 \ldots K$).
2. Ermitteln Sie alle Wahrscheinlichkeiten $P(s_k)$.
3. Berechnen Sie den Erwartungswert $E\{S\}$ der Zufallszahl[2].
4. Berechnen Sie die beiden Informationsgehalte $I(s_i)$ und $I(s_j)$ für die Ereignisse $s_i = 22$ und $s_j = 34$.
5. Berechnen Sie die Entropie $H(S)$ der Quelle, resp. der Zufallsvariable S.
6. Vergleichen Sie die Entropie $H(S)$ mit dem theoretischen Maximum der Entropie H_{max} einer Quelle mit K Ereignissen.
7. Wir nehmen nun an, dass die Quelle die Zufallszahlen im BCD-Code ausgibt. Wie gross ist die Redundanz R dieses Codes?

[1] Wir bezeichnen S auch als *Zufallsvariable*, welche die Zufallswerte s_k annehmen kann. Die Zufallswerte nennen wir auch *Ereignisse*.

[2] Als Erwartungswert bezeichnen wir jenen Wert, den wir als Mittelwert aus einer sehr grossen Zahl von Ereignissen erhalten.

Antworten

1. Ereignisse s_k mit $k = 1 \ldots K$, wobie $K = 10$:

$$\begin{array}{ccccc} s_1 = 11 & s_2 = 12 & s_3 = 13 & s_4 = 14 & s_5 = 22 \\ s_6 = 23 & s_7 = 24 & s_8 = 33 & s_9 = 34 & s_{10} = 44 \end{array}$$

2. Wahrscheinlichkeiten $P(s_k)$:

$$\begin{array}{ccccc} P(s_1) = \frac{1}{16} & P(s_2) = \frac{2}{16} & P(s_3) = \frac{2}{16} & P(s_4) = \frac{2}{16} & P(s_5) = \frac{1}{16} \\ P(s_6) = \frac{2}{16} & P(s_7) = \frac{2}{16} & P(s_8) = \frac{1}{16} & P(s_9) = \frac{2}{16} & P(s_{10}) = \frac{1}{16} \end{array}$$

3. Erwartungswert $E\{S\}$:

$$E\{S\} = \sum_{k=1}^{K} P(s_k) \cdot s_k = 21.875$$

Zur Kontrolle betrachten wir $N = 1000$ Zufallszahlen. Im Mittel treten die Ereignisse s_k mit folgenden Häufigkeiten auf:

$$\begin{array}{rcl@{\qquad}rcl} N_1 = 63 & \text{x} & s_1 = 11 & N_6 = 125 & \text{x} & s_6 = 23 \\ N_2 = 125 & \text{x} & s_2 = 12 & N_7 = 125 & \text{x} & s_7 = 24 \\ N_3 = 125 & \text{x} & s_3 = 13 & N_8 = 63 & \text{x} & s_8 = 33 \\ N_4 = 125 & \text{x} & s_4 = 14 & N_9 = 125 & \text{x} & s_9 = 34 \\ N_5 = 62 & \text{x} & s_5 = 22 & N_{10} = 62 & \text{x} & s_{10} = 44 \end{array}$$

Dabei haben wir willkürlich N_1 und N_8 auf- und N_5 und N_{10} abgerunden, so dass die Summe aller Häufigkeiten zu stimmen kommt:

$$N = \sum_{k=1}^{10} N_k$$

Damit folgt die Schätzung $\widetilde{E}\{S\}$ des Erwartungswerts:

$$\widetilde{E}\{S\} = \frac{1}{N} \cdot \sum_{k=1}^{10} N_k \cdot s_k = 21.864$$

Die Schätzung stimmt nicht schlecht, so dass wir schreiben können $\widetilde{E}\{S\} \approx E\{S\}$.

Zu beachten ist aber, dass er lineare Mittelwert m, der die Häufigkeiten N_k nicht berücksichtigt, den folgenden Wert ergibt:

$$m = \frac{1}{10} \cdot \sum_{k=1}^{10} s_k = 23$$

Er ist verschieden vom Erwartungswert, der die Häufigkeiten beachtet.

4. Informationsgehalt von $s_i = s_5$:

$$I(s_5) = \log_2 \frac{1}{P(s_5)} = \log_2(16) = 4.00 \quad \text{Bit}$$

Informationsgehalt von $s_j = s_9$:

$$I(s_9) = \log_2 \frac{1}{P(s_9)} = \log_2(8) = 3.00 \quad \text{Bit}$$

5. Entropie $H(S)$:

$$\begin{aligned} H(S) &= \sum_{k=1}^{10} P(s_k) \cdot \log_2 \frac{1}{P(s_k)} \\ &= 4 \cdot \frac{1}{16} \cdot \log_2(16) + 6 \cdot \frac{2}{16} \cdot \log_2(8) \\ &= 3.25 \quad \text{Bit/Symbol} \end{aligned}$$

6. Bei einer Quelle, die $K = 10$ Symbole erzeugt, ist die maximal mögliche Entropie H_{max}:

$$H_{max} = \log_2 K = 3.32 \quad \text{Bit/Symbol}$$

   Die Entropie $H(S)$ liegt sehr nahe am theoretischen Maximum H_{max}.

7. Der BCD-Code hat 4 Bit/Ziffer. Eine 2-stellige Zahl hat demnach eine mittlere Codewortlänge von $L = 8$ Bit. Damit folgt die Redundanz R:

$$R = L - H(S) = 4.75 \quad \text{Bit/Symbol}$$

   Die Redundanz dieses Codes ist also grösser als die Entropie selbst. Oder anders ausgedrückt: Mehr als jedes zweite Bit der codierten Zufallszahlen trägt keine Information.