Prof. Dr. Katharina Morik,
Prof. Dr. Claus Weihs
Dipl.-Inform. Marco Stolpe,
Dipl.-Stat. Julia Schiffner

Dortmund, 14.04.09
Abgabe: bis Di, 21.04., 10.00 Uhr an
`schiffner@statistik.tu-dortmund.de`

Übungen zur Vorlesung
Wissensentdeckung in Datenbanken
Sommersemester 2009
Blatt 1

Aufgabe 1.1 – Grundgesamtheit und Ereignisse (2 Punkte)
Ein handelsüblicher sechsseitiger Würfel mit den Zahlen 1 bis 6 wird zweimal geworfen.

a) Geben Sie die Grundgesamtheit Ω an.

b) Geben Sie die folgenden Ereignisse an:

   A: Die Summe der Augenzahlen der beiden Würfe ist eine Primzahl.

   B: Die Augenzahl des zweiten Wurfs ist um mindestens eins höher als die des ersten Wurfs.

   C: Die Augenzahlen beider Würfe betragen mindestens fünf.

Aufgabe 1.2 – bedingte Wahrscheinlichkeiten und Satz von Bayes (4 Punkte)
Die Eingänge eines Supermarkts sind mit einer Alarmanlage gegen Diebstahl gesichert. Wir betrachten die zwei Ereignisse

D: Der Kunde ist ein Dieb,

A: Der Alarm wird ausgelöst.

Wenn ein Dieb mit gestohlener Ware die Anlage passiert, wird mit einer Wahrscheinlichkeit von $P(A \mid D) = 0.995$ Alarm ausgelöst. Bei einem unbescholtenen Kunden beträgt diese Wahrscheinlichkeit $P(A \mid \bar{D}) = 0.006$. Nehmen Sie an, dass jeder 500. Kunde einen Diebstahl probiert (d. h. $P(D) = 0.002$).

a) Wie groß ist die Wahrscheinlichkeit $P(A)$, dass ein beliebiger Kunde mit dem Alarm konfrontiert wird?

b) Wie groß ist die Wahrscheinlichkeit $P(D \mid A)$, dass die Anlage zu Recht alarmiert?

Aufgabe 1.3 – Schätzung der Kreiszahl π (5 Punkte)
Bestimmen Sie eine Näherung der Kreiszahl π, indem Sie gleichverteilte Zufallsvektoren $X = (X_1, X_2)'$ in $[-1,1] \times [-1,1]$ erzeugen (in `R` ist dies mit der Funktion `runif()` möglich) und die Wahrscheinlichkeit für $P\big(||X|| \leq 1\big)$ schätzen.

Hinweis: Für die Dichtefunktion von zweidimensional gleichverteilten Zufallsvektoren in $[a,b] \times [c,d]$ gilt:

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} \frac{1}{(b-a)(d-c)} & \text{für } x_1 \in [a,b] \text{ und } x_2 \in [c,d] \\ 0 & \text{sonst} \end{cases}$$

sowie

$$P\big(a_1 \leq X_1 \leq b_1, a_2 \leq X_2 \leq b_2\big) = \int_{a_1}^{b_1} \int_{a_2}^{b_2} f_{X_1,X_2}(x_1,x_2)\, dx_2\, dx_1.$$

a) Zeigen Sie zunächst, dass $\hat{\pi} := 4 \cdot \hat{P}\big(\|X\| \leq 1\big)$ eine sinnvolle Schätzung für π darstellt.

b) Schätzen Sie $P\big(\|X\| \leq 1\big)$ durch die relative Häufigkeit

$$h_n\big(\|X\| \leq 1\big) = \frac{1}{n} \sum_{i=1}^{n} I_{\{\|x_i\| \leq 1\}}(x_i),$$

den Anteil an Realisierungen $x_i = (x_{i1}, x_{i2})'$ innerhalb des Einheitskreises, für unterschiedliche Anzahlen $n \in \{10, 20, 50, 100, 1000, 5000, 10000\}$ an erzeugten Zufallszahlen.

c) Stellen Sie $\hat{\pi}$ in Abhängigkeit von n grafisch dar. Wiederholen Sie nun für jedes n die Schätzung zehn mal und berechnen Sie die Standardabweichung in Abhängigkeit von n.