Technische Universität München
Fakultät für Informatik
Lehrstuhl für Effiziente Algorithmen
Prof. Dr. Ernst W. Mayr
Dr. Werner Meixner

Sommersemester 2010
Übungsblatt 11
7. Juli 2010

Diskrete Wahrscheinlichkeitstheorie

Abgabetermin: 13. Juli 2010, 14 Uhr in die **DWT Briefkästen**

Hausaufgabe 1 (5 Punkte)

Beweisen oder widerlegen Sie die folgenden Behauptungen.

1. Jede erwartungstreue Schätzvariable für einen Parameter ϑ schätzt den Erwartungswert von ϑ.
2. Bei echten Alternativtests ist die Summe der Wahrscheinlichkeit eines Fehlers 1.Art und der Wahrscheinlichkeit eines Fehlers 2. Art stets gleich 1.
3. Die Anwendung des Maximum-Likelihood-Prinzips setzt die Verfügbarkeit einer Stichprobe voraus.
4. Zu jeder stetigen Zufallsvariablen X mit Verteilung F_X gibt es ein 1-Quantil aus $\mathbb{R}$ von F_X.
5. Liegt beim Verfahren zum Test von Hypothesen H_0 der Wert der Testvariablen im Ablehnungsbereich, dann gibt die Fehlerwahrscheinlichkeit erster Art die Wahrscheinlichkeit an, mit der die Hypothese H_0 zutrifft.

Hausaufgabe 2 (5 Punkte)

1. Wir betrachten eine erwartungstreue, diskrete Schätzvariable X für einen Parameter α. Für die Verteilungsdichte f_X gelte $f_X(i) = \frac{\exp^{-3}(3^i)}{i!}$ für $i \in \mathbb{N}_0$.

   Mit welchem Wert wird α durch X geschätzt?

2. Für die Dichte f_X einer diskreten Schätzvariablen X für einen Parameter α gelte $f_X(i) = \binom{10}{i}(\frac{2}{9})^i(\frac{7}{9})^{10-i}$ und $W_X = \{0, 1, \ldots, 10\}$. Mit welchem Wert wird α durch X geschätzt, wenn der Bias $-\frac{2}{9}$ beträgt?

Hausaufgabe 3 (5 Punkte)

Sei $T = X_1 + X_2 + X_3$ eine Zufallsvariable, wobei X_1, X_2 und X_3 unabhängige Bernoulli-verteilte Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p seien. Wir betrachten T als Stichprobenvariable zum Test der Hypothese $H_0 : p \geq \frac{1}{3}$ mit Ablehnungsbereich $\tilde{K} = \{0, 1, 2\}$.

1. Berechnen Sie die Verteilungsfunktion F_T von T!
2. Berechnen Sie die Fehlerwahrscheinlichkeit 1. Art!
3. Berechnen Sie die Fehlerwahrscheinlichkeit 2. Art unter der Annahme, dass $H_1 : p \leq \frac{1}{4}$ eine echte Alternative zu H_0 sei!

Hausaufgabe 4 (5 Punkte)

Sei X eine Zufallsvariable mit unbekanntem Erwartungswert μ und einer Standardabweichung $\sigma_X = 2$. Wir verwenden $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ als Schätzvariable für μ, wobei die X_i unabhängige Wiederholungen von X seien. Wir setzen voraus, dass $\bar{X}$ für $n \geq 1500$ normalverteilt ist.

1. Berechnen Sie die Varianz von $\bar{X}$!

2. Leiten Sie eine möglichst kleine untere Schranke n_0 für n her, so dass für alle $n \geq n_0$ gilt

   $$\Pr[|\bar{X} - \mu| < 0{,}1] \geq 0{,}9\,.$$

   Für das 0,95-Quantil der Standardnormalverteilung ist dabei $z_{0,95} \approx 1{,}65$ zu verwenden.

3. Nun sei $n = 2500$. Geben Sie ein Konfidenzintervall für μ zum Konfidenzniveau 0,9 an!

***Hinweis:** Die als freiwillig gekennzeichneten Aufgaben werden nicht bewertet.*

Freiwillige Aufgabe 1

In einer großen Stadt gibt es N Taxis, die mit den Nummern $1, \ldots, N$ beschriftet sind. Wir stehen an der Straße und beobachten die Taxis, dabei notieren wir uns deren Nummer (Wiederholungen werden ignoriert). Sei $x_1, \ldots, x_n$ die aufsteigend sortierte Folge der beobachteten Nummern. Nun wollen wir die Anzahl N der Taxis schätzen.

1. Was ist der Maximum Likelihood Schätzer für N? Ist dieser erwartungstreu?

2. Geben Sie einen Schätzer für die Anzahl der Taxis an, der N dadurch abschätzt, dass er versucht, die Größe der (ev. nicht beobachteten) Lücke $x_n + 1, \ldots, N$ oberhalb von x_n abzuschätzen. Ist dieser Schätzer erwartungstreu?

Hinweis: *Auf den Übungsblättern in diesem Semester wird es grundsätzlich die drei Aufgabentypen Vorbereitungsaufgabe, Tutoraufgabe und Hausaufgabe geben. Die als Vorbereitung bezeichneten Aufgaben dienen der häuslichen Vorbereitung der Tutoraufgaben. Tutoraufgaben werden in den Übungsgruppen bearbeitet. Dabei wird die Lösung der Vorbereitungsaufgaben vorausgesetzt. Die Vorbereitungsaufgaben werden in der Zentralübung unterstützt.*

Vorbereitung 1

Wiederholung rekurrente Ereignisfolgen, Hausaufgabe 1, Blatt 9.

Tutoraufgabe 1

Auf zwei unabhängigen Servern stehe ein Web-Dienst zur Verfügung. Es soll festgestellt werden, welcher Server schnellere Antwortzeiten liefert.
Dazu werden $n = 1000$ Anfragen an die Server geschickt und es wird festgestellt, von welchem Server die Antwort zuerst eintrifft. Dabei gehen wir davon aus, dass Pakete nicht gleichzeitig empfangen werden können. In 540 Fällen antwortet Server A vor Server B.
Wir wählen als Nullhypothese H_0 die Aussage, dass Server B im Mittel schneller ist als Server A.
Kann man für einen entsprechenden statistischen Test auf einem Signifikanzniveau von $\alpha = 0,04$ die Nullhypothese annehmen? Formulieren Sie hierzu den Test und weisen Sie Ihre Behauptung nach.

Tutoraufgabe 2

Seien $(X_t)_{t\in\mathbb{N}_0}$ die Zufallsvariablen einer zeithomogenen Markov-Kette über den Zuständen $Q = \{0, 1, 2\}$ mit Übergangsmatrix

$$P = (p_{i,j}) = \begin{pmatrix} 0,25 & 0,25 & 0,5 \\ 0,25 & 0,25 & 0,5 \\ 0,25 & 0,25 & 0,5 \end{pmatrix}.$$

Die Dichtefunktion von X_0, d. h., die Startverteilung der Markov-Kette sei $q_0 = (s_0, s_1, s_2)$.

1. Berechnen Sie die Dichtefunktion q_1 von X_1.

2. Bestimmen Sie die Menge aller stationären Startverteilungen.

3. Beweisen Sie die Unabhängigkeit der beiden Variablen X_0 und X_1.

   Dabei sind X_0 und X_1 als Zufallsvariable über dem zugeordneten Wahrscheinlichkeitsraum $\langle\Omega, \Pr\rangle$ zu betrachten mit

$$\Omega = \{(x_0, x_1) : x_0, x_1 \in Q\}, \quad \Pr[(x_0, x_1)] = (q_0)_{x_0} \cdot \Pr[X_1 = x_1 | X_0 = x_0],$$

$$X_0((x_0, x_1)) = x_0 \qquad \text{und} \qquad X_1((x_0, x_1)) = x_1.$$

Tutoraufgabe 3

1. Wir betrachten Markov-Ketten M mit 6 Zuständen. Wie viele transiente Zustände kann M höchstens besitzen? Begründung!

2. Wie viele stationäre Verteilungen besitzt die Markovkette mit Übergangsmatrix

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}?$$

   Begründung!

3. Gegeben sei eine Markov-Kette M mit Zustandsmenge $\mathcal{S} = \{0, 1, 2, \dots\}$, $p_{n,(n+1)} = 2/3$ und $p_{n,0} = 1/3$ für alle $n \in \mathcal{S}$. Wie groß ist die Wahrscheinlichkeit, sich nach langer Zeit im Zustand i zu befinden?