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

Sommersemester 2010
Übungsblatt 10
30. Juni 2010

Diskrete Wahrscheinlichkeitstheorie

Abgabetermin: 6. Juli 2010, 14 Uhr in die **DWT Briefkästen**

Hausaufgabe 1 (5 Punkte)

Wir betrachten unabhängige stetige Zufallsvariablen X und Y, die beide auf dem Intervall $[0,1] \subseteq \mathbb{R}$ gleichverteilt sind. Sei $Z = \max\{X, Y\}$.

1. Berechnen Sie die Verteilungsfunktion F_Z.
2. Bestimmen Sie eine Funktion $u : [0,1] \to [0,1]$, so dass $u(X)$ die gleiche Verteilung wie Z besitzt.

Hausaufgabe 2 (5 Punkte)

Sei $X \sim \mathrm{Bin}(10^5, \frac{1}{2})$ eine binomialverteilte Zufallsvariable. Sei $k \in \{100, 1000\}$.

1. Berechnen Sie mit Hilfe der Ungleichung von Chebyshev eine untere Schranke für $\Pr[5\cdot 10^4 - k \leq X \leq 5\cdot 10^4 + k]$.
2. Approximieren Sie $\Pr[5\cdot 10^4 - k \leq X \leq 5\cdot 10^4 + k]$ mit Hilfe des Satzes von de Moivre.

Hausaufgabe 3 (5 Punkte)

Seien $X_1, X_2, \ldots$ identisch verteilt und unabhängig mit $E[X_i] = 0,6$ und $\sigma = 0,3$. Wir betrachten $Y_n = \sum_{i=1}^{n} X_i$. Berechnen Sie die folgenden Grenzwerte.

1. $\lim_{n\to\infty} \Pr[\frac{Y_n}{n} = 0,6]$.
2. $\lim_{n\to\infty} \Pr[0,5 < Y_n < 0,7]$.
3. $\lim_{n\to\infty} \Pr[0,59 < \frac{Y_n}{n} < 0,61]$.

Hausaufgabe 4 (5 Punkte)

In einem Unfallkrankenhaus treffen im Schnitt alle 20 Minuten Patienten zur Behandlung ein. Die Zeit zwischen zwei Behandlungsfällen sei exponentialverteilt mit Parameter $\frac{1}{20}$. Wenn 1 Stunde lang kein Patient eingetroffen ist, macht das Personal Ruhepause. Wir wollen wissen, welcher Zeitabstand zwischen zwei Ruhepausen zu erwarten ist.
Seien $T_1, T_2, \ldots$ die Zeitspannen zwischen dem Eintreffen zweier Behandlungsfälle und W die Wartezeit bis zur nächsten Ruhepause.

1. Geben Sie $\mathbb{E}[T_1 \mid T_1 \geq 60]$ an.
2. Geben Sie $\mathbb{E}[W \mid T_1 \geq 60]$ an.
3. Berechnen Sie $\mathbb{E}[W]$.

***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

Eine Werbeagentur möchte am letzten Tag der Fußballweltmeisterschaft mit einer Blitzumfrage schätzen, welcher Anteil ϑ der per Bahn anreisenden Fußballfans einen Platz im Stadion hat. Jeder der 12 Mitarbeiter befragt so lange zufällig ausgewählte Fans, bis er einen Fan gefunden hat, der eine Karte für das Stadion besitzt. Die Anzahl der vom Mitarbeiter i befragten Fans sei X_i.

Wir nehmen an, dass alle X_i die gleiche geometrische Verteilung besitzen mit

$$\Pr_\vartheta[X_i = k] \;=\; (1-\vartheta)^{k-1} \cdot \vartheta\,,\ \forall k \in \mathbb{N}\,.$$

1. Man bestimme auf der Basis der ermittelten Stichprobenwerte

   $$3,\ 1,\ 2,\ 3,\ 2,\ 1,\ 2,\ 1,\ 5,\ 3,\ 2,\ 3$$

   einen Maximum-Likelihood-Schätzwert für ϑ.

2. Man gebe mit Hilfe des Zentralen Grenzwertsatzes ein konkretes 95%-Konfidenzintervall für ϑ an.

Tutoraufgabe 1

Ziel dieser Aufgabe ist es, einen ML-Schätzer für die Erfolgswahrscheinlichkeit $p \in (0,1)$ einer negativ binomialverteilten ZV zu bestimmen. Hierfür seien $X_1, \ldots, X_n$ unabhängige Zufallsvariable, wobei jedes X_i negativ binomialverteilt mit Erfolgswahrscheinlichkeit p bei m zu erzielenden Erfolgen sei, d. h., jedes X_i hat die Dichte

$$f_{X_i}(k) = \binom{k-1}{m-1} p^m (1-p)^{k-m} \quad (\text{mit } k \geq m)\,.$$

Der Parameter m sei bekannt. Zu schätzen ist p.

1. Es sei $\vec{k} = (k_1, \ldots, k_n)$ ein Stichprobenvector (mit $k \geq m$). Stellen Sie die Likelihood-Funktion $L(\vec{k}; p)$ auf.

2. Maximieren Sie $L(\vec{k}; p)$ und bestimmen Sie den entsprechenden ML-Schätzer für p.

3. Zeigen Sie, dass der hergeleitete ML-Schätzer i. A. nicht erwartungstreu ist.

   Hinweis: Verwenden Sie $\ln(1+x) = \sum_{i=1}^{\infty} (-1)^{i+1} \frac{x^i}{i}$ für $x \in (-1, 1]$.

Tutoraufgabe 2

Beim Testen von Hypothesen bezeichnet man die zu überprüfende Hypothese (Nullhypothese) generell mit H_0 und die Alternative mit H_1. Ein Tierhändler erhält ein Paket mit 100 Frettchen. Er will testen, ob weniger als zehn (< 10) dieser Frettchen aggressiv und bissig sind. Dazu hält er zehn Frettchen seinen Finger hin und nimmt das Paket nur an, wenn ihn keines davon beißt (wir nehmen an, dass ein aggressives Frettchen sofort zubeißen würde). Wie lauten die Hypothesen des Händlers? Was ist das Signifikanzniveau des Tests?

Tutoraufgabe 3

Sei X eine Bernoulli-verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit p. Mit einer Stichprobe von nur 2 Elementen soll die Hypothese $H_0 : p \geq \frac{2}{3}$ getestet werden. Die Testvariable sei $T = X_1 + X_2$, wobei X_1 und X_2 unabhängige Kopien von X sein sollen, d. h., X_1, X_2 sind ebenfalls Bernoulli-verteilt mit gleichem Parameter p. Der Ablehnungsbereich des Tests sei $K = \{0\}$.

1. Geben Sie die Verteilungsfunktion von T an. (Es ist nicht nach der Dichtefunktion von T gefragt!)

2. Berechnen Sie die Fehlerwahrscheinlichkeit 1. Art (Signifikanzniveau) α_1 .

3. Wir nehmen an, dass die Alternative $H_1 : p \leq \frac{1}{3}$ echt ist. Dies bedeutet, dass H_1 gelte, wenn H_0 nicht gilt.

   Berechnen Sie damit die Fehlerwahrscheinlichkeit 2. Art α_2 .