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

Sommersemester 2011
Übungsblatt 9
4. Juli 2011

Diskrete Wahrscheinlichkeitstheorie

Abgabetermin: 12. Juli 2011, 12 Uhr in die **DWT Briefkästen**

Hausaufgabe 1 (5 Punkte)

Sei X eine kontinuierliche Zufallsvariable.

1. Zeigen Sie: Falls $X \sim \mathcal{N}(2, \frac{1}{2})$, dann gilt $2X + 1 \sim \mathcal{N}(5, 2)$.
2. Seien $d_1, d_2, c \in \mathbb{R}$ mit $d_1 < d_2$ und $c > 0$. Berechnen Sie $a, b \in \mathbb{R}$, so dass für $Y = aX + b$ gilt
$$\Pr[d_1 \le X \le d_2] = \Pr[-c \le Y \le c]\,.$$

Hausaufgabe 2 (5 Punkte)

Seien $X_1, X_2, \ldots$ kontinuierliche Zufallsvariable, die identisch verteilt und unabhängig sind mit $E[X_i] = 2$ und $\mathrm{Var}[X_i] = 4$. Wir betrachten $Y_n = \sum_{i=1}^{n} X_i$. Berechnen Sie die folgenden Grenzwerte und begründen Sie die Korrektheit Ihrer Herleitung.

1. $\lim_{n\to\infty} \Pr[\frac{Y_n}{n} = 2]$.
   Benutzen Sie für Ihre Herleitung <u>nicht</u> den Zentralen Grenzwertsatz.
2. $\lim_{n\to\infty} \Pr[1,9 < Y_n < 2,1]$.
3. $\lim_{n\to\infty} \Pr[1,99 < \frac{Y_n}{n} < 2,01]$.

Hausaufgabe 3 (5 Punkte)

Sei $X \sim \mathrm{Bin}(10^6, \frac{1}{10})$ eine binomialverteilte Zufallsvariable. Sei $k = 500$.

1. Zeigen Sie $\Pr[10^5 - k \le X \le 10^5 + k] \le 0,64$.
2. Approximieren Sie $\Pr[10^5 - k \le X \le 10^5 + k]$ unter der Annahme, dass der Satz von de Moivre anwendbar ist.

Hausaufgabe 4 (5 Punkte)

An der dänischen Grenze werden Grenzkontrollen durchgeführt. Im Schnitt treffen alle 30 Sekunden an der Grenzstation Personen ein, die zu kontrollieren sind. Die Zeit zwischen zwei Kontrollen sei exponentialverteilt mit Parameter $\frac{1}{30}$. Wenn 2 Minuten lang kein Kontrollfall eingetroffen ist, dann machen die Grenzbeamten Ruhepause.

Seien $T_1, T_2, \ldots$ die Zeitspannen zwischen dem Eintreffen von zu kontrollierenden Personen und W die Wartezeit bis zur nächsten Ruhepause.

1. Berechnen Sie $\mathbb{E}[T_1 \mid T_1 \ge 120]$.
2. 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

Seien $X_1, \ldots, X_n$ unabhängige Stichproben einer Zufallsvariablen X und sei $\bar{X}$ das arithmetische Mittel der X_i. Wir verwenden die Zufallsvariable

$$V = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

als Schätzer für die Varianz von X.

Berechnen Sie den Bias von V! Welche Aussage gilt für $n \to \infty$?

Vorbereitung 2

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

Die tatsächlich benötigte CPU-Zeit einer Benutzersitzung an einer Workstation werde als eine Zufallsvariable mit unbekanntem Erwartungswert μ und der Varianz $\sigma^2 = 6{,}25$ [sec^2] angenommen.

Seien X_i unabhängige Stichproben der CPU-Zeit und $\bar{X}$ das arithmetische Mittel der X_i. Wie viele unabhängige Stichproben sollten mindestens gemessen werden, damit

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

gilt. Man verwende zur Beantwortung der Frage

1. die Ungleichung von Chebyshev,
2. den Zentralen Grenzwertsatz.

Tutoraufgabe 2

Ziel dieser Aufgabe ist es, einen ML-Schätzer für die Erfolgswahrscheinlichkeit $p \in (0,1)$ einer negativ binomialverteilten Zufallsvariablen 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 3

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?