Technische Universität München
Fakultät für Informatik
Lehrstuhl Informatik XIV
Prof. Dr. Susanne Albers
Dr. Werner Meixner

Sommersemester 2014
Übungsblatt 9
16. Juni 2014

Diskrete Wahrscheinlichkeitstheorie

Abgabetermin: 25. Juni 2014, 10 Uhr in die **DWT Briefkästen**

Hausaufgabe 1 (5 Punkte)

Sie werfen eine kreisförmige Münze mit Radius 0.5 auf ein Quadrat mit Seitenlänge 4, wobei der Mittelpunkt der Münze stets innerhalb des Quadrats zu liegen kommt.

Wie hoch ist die Wahrscheinlichkeit, dass die Münze über den Rand des Quadrats hinausragt, wenn wir eine Gleichverteilung des Mittelpunkts der Münze auf der Quadratfläche annehmen?

Hausaufgabe 2 (5 Punkte)

Die Lebensdauer T eines Rechners habe die folgende mit $\lambda > 0$ und a parametrisierte Dichte:

$$f_T(t) = \begin{cases} a\lambda^2 t e^{-\lambda t} & : t \geq 0 \\ 0 & : t < 0 \end{cases}$$

1. Welchen Wert muss a besitzen, so dass für alle $\lambda > 0$ die Funktion f_T tatsächlich eine Dichte ist, d. h., dass $\int_{-\infty}^{\infty} f_T(t)\,dt = 1$ gilt. Beweis!
2. Berechnen Sie $\Pr[T \leq \frac{2}{\lambda}]$.
3. Berechnen Sie $\mathbb{E}[T]$ in Abhängigkeit von λ .

Hausaufgabe 3 (5 Punkte)

Wahr oder falsch? Begründen Sie Ihre Antwort!

1. Falls $X \sim \mathcal{N}(0, 1)$ und $Y \sim \mathcal{N}(0, 2)$ normalverteilt sind, dann folgt $\mathrm{Var}[X + Y] = 3$.
2. Seien X und Y standardnormalverteilt. Dann gilt $\Pr[X \leq 0] = \Pr[Y \geq 0]$.
3. Bei kontinuierlichen Zufallsvariablen existiert stets der Erwartungswert.
4. Seien $X \sim \mathrm{Po}(1)$ und $Y \sim \mathrm{Bin}(n, \frac{1}{n})$. Dann gilt $|\Pr[X = 2n] - \Pr[Y = 2n]| < 2^{-n}$.
5. Sei X exponentialverteilt. Dann gilt $\Pr[X > 2 \mid X > 1] + \Pr[X \leq 1] = 1$.

Hausaufgabe 4 (5 Punkte)

Sei $a > 0$, und seien X, Y kontinuierliche Zufallsvariablen mit gemeinsamer Dichtefunktion

$$f_{X,Y}(x, y) = \begin{cases} a \cdot (1 - x \cdot y) & : \quad 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & : \quad \text{sonst} \end{cases}$$

1. Berechnen Sie die Randdichten $f_X(x)$ und $f_Y(y)$.
2. Bestimmen Sie a.
3. Sind die Variablen X und Y unabhängig? Beweisen Sie Ihre Antwort.

Hinweis: *Die Vorbereitungsaufgaben bereiten die Tutoraufgaben vor und werden in der Zentralübung unterstützt. Tutoraufgaben werden in den Übungsgruppen bearbeitet. Hausaufgaben sollen selbstständig bearbeitet und zur Korrektur und Bewertung abgegeben werden.*

Vorbereitung 1

Seien $X_1, X_2, \ldots, X_{100}$ unabhängige Bernoulli-verteilte Zufallsvariable mit gleicher Erfolgswahrscheinlichkeit $\frac{1}{50}$. Sei $S_{100} = \sum_{i=1}^{100} X_i$.

1. Berechnen Sie Zahlen $a, b \in \mathbb{R}$ mit $a > 0$, so dass $\mathbb{E}[Y] = 0$ und $\mathrm{Var}[Y] = 1$ für $Y = a \cdot S_{100} + b$ gelten.
2. Wenden Sie den zentralen Grenzwertsatz an zur approximativen Berechnung eines Intervalls $[d_1, d_2]$, so dass
$$\Pr[d_1 \le S_{100} \le d_2] \approx 1 - \alpha = 1 - 0.05\,.$$
Benutzen Sie dabei das Quantil $c = x_{1-\frac{\alpha}{2}} \approx 1.96$ der Standardnormalverteilung.

Vorbereitung 2

Bei einem Einwahlserver für $n = 10^3$ Teilnehmer nehmen wir an, dass zu einem festen Zeitpunkt jeder Teilnehmer mit Wahrscheinlichkeit $p = 0,05$ Zugriff auf den Server wünscht.

Berechnen Sie eine Näherung der Wahrscheinlichkeit, mit der gleichzeitig mehr als 55 Verbindungswünsche auftreten? Approximieren Sie dabei die Binomialverteilung durch die entsprechende Normalverteilung und benutzen Sie ggf. geeignete Tabellen für die Werte der Standardnormalverteilung.

Vorbereitung 3

Sei X eine binomialverteilte Zufallsvariable mit $X \sim \mathrm{Bin}(2000, 0.05)$. Wir nehmen an, dass die Voraussetzungen sowohl für eine Approximation von entsprechenden Verteilungen mit Poisson-Verteilung als auch mit Normalverteilung vorliegen.
Berechnen Sie approximativ

1. $\Pr[X = 110]$,
2. $\Pr[X > 110]$.

Begründen Sie jeweils die Wahl einer der Approximationen.

Vorbereitung 4

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

Sei $X = \sum_{i=1}^{2000} X_i$ die Summe der Augenzahlen, wenn man 2000-mal mit einem idealen Würfel würfelt.

1. Berechnen Sie näherungsweise $\Pr[7000 \le X \le 7100]$!
2. Wie groß muss man Δ wählen, damit $\Pr[7000 - \Delta \le X \le 7000 + \Delta] \approx \frac{1}{2}$ gilt?

Approximieren Sie die entsprechenden Wahrscheinlichkeiten mit Hilfe des Zentralen Grenzwertsatzes.

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

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

Der Parameter m sei bekannt. Zu schätzen ist p.

1. Es sei $\vec{k} = (k_1, \ldots, k_n)$ ein Stichprobenvector (mit $k \ge m$). 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?