Technische Universität München
Institut für Informatik
Lehrstuhl für Effiziente Algorithmen
Prof. Dr. Ernst W. Mayr

SS 2001
Übungsblatt 9
12. Juli 2001

# Diskrete Strukturen II

Abgabe bis zum 19. Juli 2001, 10:00.

## Aufgabe 1

Das Einheitsintervall werde in $n$ gleich große Teilintervalle zerteilt. Jedes Teilintervall wird mit WS $\lambda/n$ markiert.
Wie ist die Anzahl $X$ der markierten Intervalle verteilt?
Betrachten Sie nun das $i$-te Teilintervall und das $i+m$-te (unter der Annahme, daß $n$ hinreichend groß ist, also $n \geq i+m$. Wie groß ist die WS $p_m$, daß die Intervalle $i+1, i+2, \ldots, i+m-1$ nicht markiert sind, das Intervall $i+m$ aber schon, unter der Voraussetzung, daß das $i$-te Intervall markiert ist.
Anschaulich beschreibt eine Zufallsvariable $Y$ mit der Dichte $f_Y(m/n) = p_m$ die Verteilung der Abstände zwischen zwei markierten Intervallen.
Gegen welche Verteilungsfunktionen konvergiert $F_X$ bzw. $F_Y$, wenn $n \to \infty$?

## Aufgabe 2

Berechnen Sie die Momente $\{\mathbb{E}\left[X^k\right]\}_{k=1}^{10}$, wenn $X$ standardnormalverteilt ist.
Vorschlag: Benutzen Sie die momenterzeugende Funktion $\int_{-\infty}^{\infty} e^{tx}\varphi(x;\ 0,1)\,dx = e^{t^2/2}$, wie sie in der Vorlesung berechnet wurde (und das Kommando `diff` in Maple).

## Aufgabe 3

Sei $X$ binomialverteilt. Berechnen Sie die Verteilungsfunktion von $X$ und der normierten Zufallsvariablen

$$\hat{X} := \frac{X - \mathbb{E}[X]}{\sqrt{\mathrm{Var}[X]}}.$$

Plotten Sie $F_{\hat{X}}$ zusammen mit der Verteilungsfunktion $\Phi$ einer standardnormalverteilten ZV, jeweils für $n = 10, p = 1/2$ und $n = 50, p = 1/2$.
Vergleichen Sie die Aussage des zentralen Grenzwertsatzes mit dem, was Sie auf den Plots sehen.

## Aufgabe 4

Sie besitzen eine verbogene Münze; '0' steht für Kopf und '1' für Zahl. Sie wollen die Wahrscheinlichkeit $\Pr[1]$ bestimmen, daß Zahl fällt. In $n$ unabhängigen Würfen beschaffen Sie sich eine Folge $(a_i)_{i=1}^n$ in $\{0,1\}^n$.

a) Wie berechnen Sie aus der gegebenen Folge einen brauchbaren Schätzwert für $\Pr[1]$?
b) Wie wahrscheinlich ist es, daß ihr Schätzwert für $\Pr[K]$ um mehr als 20% vom wahren Wert abweicht, als Funktion von $n$? Verwenden Sie Chernoff und Chebychev.
c) Berechnen Sie diese Wahrscheinlichkeit für feste Werte von $n$, nämlich $n = 50, 200, 1000$. Sie dürfen für die Rechnung die Verteilungsfunktion von normierten Summen von Zufallsvariablen durch die Verteilungsfunktion einer standardnormalverteilten ZV annähern, wie das durch den zentralen Grenzwertsatz zumindest asymptotisch gerechtfertigt ist.

## Aufgabe 5

**(Freiwillig!) Zusammenhang zwischen diskreter und kontinuierlicher Verteilung**
Sei $X$ eine diskrete ZV. Die diskrete Dichte $f_X(x)$ ist eine Abbildung von $W_X$ nach $\mathbb{R}_0^+$. Diese Dichte ist nur dann ungleich Null, wenn $x \in W_X$. Wir können uns vorstellen, daß in die reelle Achse abzählbar viele Bleikügelchen eingegossen sind, die zwar unendlich dicht sind, aber ein Gesamtgewicht von eins haben. Wir können diese Situation auch mit einer kontinuierlichen Dichte $f_{\hat{X}}$ annähern. In der physikalischen Welt gibt es keine Teilchen, die unendliche Dichte haben, sie können allenfalls sehr klein und sehr dicht sein.
Betrachten Sie zunächst die diskrete ZV $X$ mit Werten $1, 2, 4$, die mit WS $1/3, 1/2, 1/6$ angenommen werden. Plotten Sie die Verteilungsfunktion $F_X$.
Betrachten Sie nun die (kontinuierliche) ZV $\hat{X}^{(n)}$ mit der Dichte

$$f_{\hat{X}^{(n)}} := 1/3 \cdot f_{[1-n/2,1+n/2]} + 1/2 \cdot f_{[2-n/2,2+n/2]} + 1/6 \cdot f_{[4-n/2,4+n/2]}.$$

Dabei bezeichne $f_{[a,b]}$ die kontinuierliche Dichte einer auf dem Intervall $[a, b]$ gleichverteilten ZV. Plotten Sie $f_{\hat{X}}$ und $F_{\hat{X}}$, jeweils für $n = 2, 5, 20$.
Betrachten Sie auch die (kontinuierliche ZV) $\tilde{X}^{(n)}$

$$f_{\tilde{X}^{(n)}} := 1/3 \cdot \varphi(x;\ 1, 1/n) + 1/2 \cdot \varphi(x;\ 2, 1/n) + 1/6 \cdot \varphi(x;\ 2, 1/n).$$

Plotten Sie $f_{\tilde{X}}$ und $F_{\tilde{X}}$, jeweils für $n = 2, 5, 20$.
Was ist $\lim_{n\to\infty} F_{\hat{X}}$ und $\lim_{n\to\infty} F_{\tilde{X}}$?
Überzeugen Sie sich, daß die zugehörigen kontinuierlichen Dichten $f_{\hat{X}}$ und $f_{\tilde{X}}$ nicht konvergieren, und beschreiben Sie das Problem.

**Kommentar für besonders Interessierte:** Wir hatten gesehen, daß die Wahrscheinlichkeit $\Pr[\cdot]$ eine geeignete Abbildung von $\Sigma \subseteq 2^\Omega$ nach $[0, 1]$ ist. Man nennt solche Abbildungen auch *Wahrscheinlichkeitsmaße.*
Im Falle $\Omega = \mathbb{R}$ betrachtet man meist (die *Borelsche* $\sigma$-Algebra) $\Sigma_B$, welche sich in natürlicher Weise aus den Intervallen erzeugen läßt. Nun kann man jedes Wahrscheinlichkeitsmaß zerlegen in einen Anteil, der eine kontinuierliche Dichte hat, und einen Rest, der dem diskreten Anteil entspricht.
Dazu bedient man sich sogenannter Diracmaße $\Delta_{x_0}$, die intuitiv ein 'unendlich dichtes Kügelchen mit Masse eins an der Stelle $x_0$' beschreiben. Formal ist $\Delta_{x_0}[A] = 1$ genau dann, wenn $x_0 \in A$, und Null sonst. Jedes Wahrscheinlichkeitsmaß läßt sich durch eine Konvexkombination von (höchstens) abzählbar vielen Diracmaßen und einem kontinuierlichen Maß beschreiben.

$$\Pr[\cdot] = \sum_{i\in I} \lambda_i \cdot \Delta_{x_i}[\cdot] + \lambda_{\text{kont}} \cdot \Pr_{\text{kont}}[\cdot].$$

Die 'Gewichtungsfaktoren' $\lambda_i$ enstprechen i.wesentl. dem, was wir diskrete Dichte genannt haben. $\Pr_{\text{kont}}[\cdot]$ stammt von eine kontinuierlichen Dichte $\mu_{\text{kont}}$, wie in der Vorlesung besprochen. Es muß gelten $\sum_{i\in I} \lambda_i + \lambda_{\text{kont}} = 1$.
Manchmal 'tut man so', als hätten die Diracmaße eine Dichte, die man als $x \mapsto \delta_{x_0}(x)$ schreibt. Diese '*$\delta$-Funktionen*' sind aber keine 'echten Funktionen' von $\mathbb{R} \to \mathbb{R}_0^+$, da sie einerseits Integral eins haben sollten, aber andererseits überall gleich Null sein müßten, außer an der Stelle $x_0$. Wenn $x \mapsto \mu_{\text{kont}}(x)$ die kontinuierliche Dichte von $\Pr_{\text{kont}}[\cdot]$ bezeichnet, schreibt man symbolisch

$$x \mapsto \sum_{i\in I} \lambda_i \cdot \delta_{x_i}(x) + \lambda_{\text{kont}} \cdot \mu_{\text{kont}}(x)$$

für die Dichte einer Zufallsvariablen, die sowohl diskrete als auch kontinuierliche Anteile hat.
Wenn man die Verteilungsfunktion betrachtet, wird alles etwas einfacher. Die Diracmaße an den Stellen $x_i$ werden zu Sprungstellen, die Höhe dieser Sprungstellen ist jeweils $\lambda_i$. In einem geeigneten Sinne (*Distributionen*) stimmt die Aussage, daß die Ableitung der Verteilungsfunktion gleich der Dichte ist, auch in diesem Setting. So ist etwa die (sog. *schwache*) Ableitung der Sprungfunktion

$$\theta_{x_0}(x) := \begin{cases} 0 & x < x_0 \\ 1 & x \geq x_0 \end{cases}$$

gleich $\delta_{x_0}(x)$. Eine mathematisch saubere Diskussion würde unseren Rahmen völlig sprengen. Wir verweisen auf Vorlesungen zur Maßtheorie und, wegen der Distributionen, auch zur Funktionalanalysis.