Technische Universität München (I7)
Prof. J. Esparza / Dr. M. Luttenberger / M. Schlund

Sommer 2012
25. Mai 2012

Lösung

Diskrete Wahrscheinlichkeitstheorie – Hausaufgabe 4

Abgabe bis zum 23.5. bis 8:30.

Alle Antworten sind unter Angabe des Rechenwegs zu begründen, soweit nicht anders gefordert! Fragen gerne im infler-Forum posten :).

Aufgabe 4.1 Abzugeben sind (a), (b) und (c) — 1P+2P+3P

Wir betrachten eine Multiple-Choice-Aufgabe, die aus genau 3 ja/nein Fragen besteht.

Die Fakultät zwingt uns eines der folgenden Bepunktungssysteme für die Aufgabe zu verwenden:

(1) Antwort richtig: +1 Punkt, Antwort falsch oder keine Antwort: 0 Punkte.

(2) Antwort richtig: +1 Punkt, Antwort falsch oder keine Antwort: -0.5 Punkt. Negative Punkte als Gesamtpunktzahl der Aufgabe sind möglich.

(3) Antwort richtig: +1 Punkt, Antwort falsch: -1 Punkt, keine Antwort: -0.5 Punkte. Erreicht ein Student eine negative Punktezahl, so wird die Aufgabe stattdessen mit 0 Punkten bewertet.

Wir sind natürlich daran interessiert, das System zu finden, bei dem ein ratender Student am schlechtesten abschneidet.

Wir nehmen an, dass ein ratender Student höchstens eine Antwort pro Frage ankreuzt. Sei 1/5 die W'keit, dass ein ratender Student keine Antwort ankreuzt, während er mit W'keit 2/5 "ja" bzw. "nein" ankreuzt.

Sei A_i die ZV, die die Gesamtpunktzahl eines ratenden Studenten unter dem Bepunktungssystem (i) angibt.

(a) Berechnen Sie $\mathbb{E}[A_1]$ und $\mathrm{Var}[A_1]$

(b) Bestimmen Sie Parameter $a, b \in \mathbb{R}$ so, dass $A_2 = a \cdot X + b$ mit $X \sim \mathrm{Bin}(3, 2/5)$. Bestimmen Sie dann $\mathbb{E}[A_2]$ sowie $\mathrm{Var}[A_2]$.

(c) Betrachten Sie nun das System (3). Bestimmen Sie die Dichte von A_3 sowie $\mathbb{E}[A_3]$ und $\mathrm{Var}[A_3]$.

Wie hängt die Dichtefunktion von A_3 mit den Koeffizienten c_j in $(\frac{2}{5}z^2 + \frac{2}{5}z^{-2} + \frac{1}{5}z^{-1})^3 = \sum_{j=-6}^{6} c_j z^j$ zusammen?

Lösung:

(a) System (1): $A_1 \sim \mathrm{Bin}(3, 2/5)$. Daher $\mathbb{E}[A_1] = 6/5 (= 1.2)$ und $\mathrm{Var}[A_1] = 6/5 \cdot 3/5 = 18/25 (= 0.72)$.

(b) System (2): $A_2 = F_1 + F_2 + F_3$. Die F_i sind folgendermaßen verteilt: $\Pr[F_i = 1] = 2/5$ und $\Pr[F_i = -0.5] = 3/5$. Sie lassen sich aber auf Bernoulli-Variablen transformieren. Sei dazu $X_i := 2/3 \cdot (1/2 + F_i)$ dann ist $X_i \sim \mathrm{Bin}(1, 2/5)$. Damit ist $A_2 = 3/2(X_1 + X_2 + X_3) - 3/2$. Also ist $2/3 \cdot (A_2 + 3/2) \sim \mathrm{Bin}(3, 2/5)$ verteilt. Es ergibt sich $\mathbb{E}[A_2] = 3/2(6/5) - 3/2 = 3/10 (= 0.3)$ sowie $\mathrm{Var}[A_2] = (3/2)^2 \cdot (3 \cdot 2/5 \cdot 3/5) = 81/50 (= 1.62)$.

(c) System (3):

Sei R_3 die Anzahl der richtigen Antworten in der Aufgabe, F_3 die Anzahl der falschen Antworten, und somit $3 - R_3 - F_3$ die Anzahl der Fragen, in denen der Student keine Antwort angekreuzt. R_3 und F_3 sind nicht unabhängig!

Dann ist

$$A_3 = \max\{0, 3/2R_3 - 1/2F_3 - 3/2\}$$

die Endpunktzahl in der Aufgabe.

Um die Dichte von A_3 zu bestimmen bestimmen wir alle Möglichkeiten $(R_3, F_3, 3 - R_3 - F_3)$ um $1/2, 1, 3/2, 2, 5/2$ oder 3 Punkte zu erhalten (manche Punktezahlen sind nicht möglich) und die W'keit mit der sie auftreten (die Wahrscheinlichkeit 0 Punkte zu erhalten ergibt sich einfach aus der Summe der Komplemente):

Punkte	(Richtig,Falsch,Nicht)
1/2	—
1	(2,1,0)
3/2	(2,0,1)
2	—
5/2	—
3	(3,0,0)

Die Dichte von A_3 ergibt sich also zu:

$$f_{A_3}\begin{cases} (1) & = & \binom{3}{2}(\frac{2}{5})^2(\frac{2}{5}) = \frac{24}{125} \\ (3/2) & = & \binom{3}{2}(\frac{2}{5})^2(\frac{1}{5}) = \frac{12}{125} \\ (3) & = & (\frac{2}{5})^3 = \frac{8}{125} \\ (0) & = & 1 - f_{A_3}(1) - f_{A_3}(3/2) - f_{A_3}(3) = \frac{81}{125} \end{cases}$$

Und dann ist $\mathbb{E}[A_3] = \frac{1}{125}(3 \cdot 8 + 3/2 \cdot 12 + 24) = \frac{66}{125} = 0.528$ und $\mathrm{Var}[A_3] = \mathbb{E}[A_3^2] - \mathbb{E}[A_3]^2 = \frac{72+9/4\cdot 12+24}{125} - (\frac{66}{125})^2 = \frac{123\cdot 125 - 66^2}{125^2} = 0.705\ldots$.

Erneut sehen wir einen schönen Zusammenhang zu erzeugenden Funktionen: $f_{A_3}(i) = c_{2i}$ für $i > 0$ und $f_{A_3}(0) = \sum_{i=-6}^{0} c_i$.

System (1) ist also Studentenfreundlich—ob (2) oder (3) als "schlechter" einzustufen sind, hängt von der Risikoaffinität von Prof. E. ab! (Erwartungswert bei (2) ist kleiner, aber "Streuung" ist größer, bei (3) genau umgekehrt).

Ausflug in die GF-Welt: Allgemein können wir erstmal eine Variante von System (3) mit negativen Punkten betrachten. Sei $\hat{A}_k$ die ZV, die die Anzahl der **halben** Punkte nach k ja/nein-Fragen zählt. Sei $G_k(z)$ die "Erzeugendenfunktion" der Folge $\{\Pr\left[\hat{A}_k = i\right]\}_{i=-2k}^{2k}$, sprich $\Pr\left[\hat{A}_k = i\right] = [z^i]G_k(z)$ (in Anführungszeichen, da EF typischerweise Potenzreihen sind, d.h. sie haben nur positive Exponenten da sie "Dinge" zählen—hier zählen sie eben auch "negative Dinge"). Für $G_k(z)$ kann man sofort eine schöne Rekursion aufstellen:

$$G_0(z) = 1z^0$$

und

$$G_{k+1}(z) = (2/5z^2 + 2/5z^{-2} + 1/5z^{-1})G_k(z).$$

Und daher ist

$$G_k(z) = (2/5z^2 + 2/5z^{-2} + 1/5z^{-1})^k.$$

$A_k := \max\{0, \hat{A}_k\}$ zählt also die Anzahl der halben Punkte in System (3). Die W'keiten für positive Werte sind die selben wie für $\hat{A}_k$, lediglich die W'keit für ≤ 0 Punkte ändert sich (deren Masse wird auf die 0 geworfen):

$$\Pr[A_k = i] = \begin{cases} [z^i]G_k(z), \text{für } i > 0 \\ \sum_{i=-2k}^{0}[z^i]G_k(z), \text{für } i = 0 \end{cases}$$

Aufgabe 4.2 **Abzugeben** **4P**

Seien W, X, Y, Z (nicht nur paarweise!) unabhängige (diskrete) ZVen mit folgenden Verteilungen:

- $W \sim \mathrm{Bin}(5, 1/3)$.
- $X \sim \mathrm{Geo}(1/2)$.
- $Y \sim \mathrm{Bin}(20, 3/4)$.
- $Z \sim \mathrm{Geo}(2/5)$.

Bestimmen Sie $\mathbb{E}[\frac{(X+Z)(Y+Z)}{W+1}]$. Begründen Sie jeden Ihrer Rechenschritte!

Lösung: Da $W + 1$ nur Werte in [11] mit positiver W'keit annimmt, ist $(W + 1)^{-1}$ wohldefiniert.

Da W, X, Y, Z unabhängig, sind auch $W + 1, X, Y, Z$ unabhängig, und ebenso $(W + 1)^{-1}, X, Y, Z$. Somit gilt:

$$\mathbb{E}[\frac{(X+Z)(Y+Z)}{W+1}] = \mathbb{E}[(X+Z)(Y+Z)] \cdot \mathbb{E}[(W+1)^{-1}] = \mathbb{E}[XY + XZ + ZY + Z^2] \cdot \mathbb{E}[(W+1)^{-1}]$$

Mit der Linearität des EW folgt:

$$\mathbb{E}[\frac{(X+Z)(Y+Z)}{(W+1)}] = (\mathbb{E}[XY] + \mathbb{E}[XZ] + \mathbb{E}[YZ] + \mathbb{E}[Z^2]) \cdot \mathbb{E}[(W+1)^{-1}].$$

Wegen der Unabhängigkeit von jeweils X, Y, XZ bzw. Y, Z:

$$\mathbb{E}[\frac{(X+Z)(Y+Z)}{W+1}] = (\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Z] + \mathbb{E}[Y]\mathbb{E}[Z] + \mathbb{E}[Z^2]) \cdot \mathbb{E}[(W+1)^{-1}].$$

Mit $\mathrm{Var}[Z] = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2$ folgt:

$$\mathbb{E}[\frac{(X+Z)(Y+Z)}{W+1}] = (\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Z] + \mathbb{E}[Y]\mathbb{E}[Z] + \mathrm{Var}[Z] + \mathbb{E}[Z]^2) \cdot \mathbb{E}[(W+1)^{-1}].$$

Nach Vorlesung:

- $\mathbb{E}[X] = 2$.
- $\mathbb{E}[Y] = 15$.
- $\mathbb{E}[Z] = 5/2$ und $\mathrm{Var}[Z] = 5/2 \cdot 5/2 \cdot 3/5 = 15/4$.

Bleibt:

$$\mathbb{E}[(W+1)^{-1}] = \sum_{k=0}^{5} \frac{1}{k+1} \Pr[W = k] = \sum_{k=0}^{5} \frac{1}{k+1} \binom{5}{k} (1/3)^k \cdot (2/3)^{5-k} = \ldots = \frac{665}{1458} \approx 0.456.$$

Alles zusammen: $\mathbb{E}[\frac{(X+Z)(Y+Z)}{W+1}] = \frac{36575}{972} \approx 37.6$

Aufgabe 4.3 **Abzugeben sind (a), (b) und (c).** **2P+1P+2P**

Eine (idealisierte) Hashfunktion ist eine Abbildung $h\colon \{0,1\}^* \to \{0,1\}^l$ für ein festes $l \in \mathbb{N}$. So bildet SHA-512 Strings (quasi) beliebiger Länge auf Hashwerte bestehend aus $l = 512$ Bit ab.

Für kryptographische Anwendungen fordert man, dass es "schwierig" sein muss, eine Kollision zu finden, d.h. zwei Urbilder $u_1, u_2 \in \{0,1\}^*$ mit $u_1 \neq u_2 \wedge h(u_1) = h(u_2)$.

Um eine untere Schranke für die Ausgabelänge l einer Hashfunktion herzuleiten, betrachtet man folgendes Experiment:

$X_1, X_2, \ldots, X_m$ seien unabhängige ZVen, jede gleichverteilt über $\{0,1\}^{2l}$. Also $\Pr[X_i = u] = 2^{-2l}$ für alle $i \in [r]$ und $u \in \{0,1\}^{2l}$.

Weiterhin sei idealisiert angenommen, dass die Hashfunktion die Gleichverteilung erhält, das heißt, dass auch $\Pr[h(X_i) = w] = 2^{-l}$ für alle $i \in [r]$ und $w \in \{0,1\}^l$.

Es sei $K_u := \{\omega \in \Omega \mid |\{X_1(\omega), \ldots, X_m(\omega)\}| < m\}$ das Ereignis, dass eine Kollision bereits in den Urbildern vorliegt.

Entsprechend sei $K_h := \{\omega \in \Omega \mid |\{h(X_1(\omega)), \ldots, h(X_m(\omega))\}| < m\}$.

(a) Zeigen Sie, dass $\Pr[K_h] \geq 1/2 \cdot \binom{m}{2} 2^{-l}$ für $\binom{m}{2} \leq 2^l$.

Hinweise: Argumentieren Sie zunächst, warum Sie die Formel aus Beispiel 20 auf $\Pr\left[\overline{K}_h\right]$ anwenden dürfen. Skizzieren Sie sich dann die Abbildung $1 - e^{-x}$ über dem Intervall $[0,1]$.

(b) Zeigen Sie, dass auch $\Pr[K_h] \leq \binom{m}{2} 2^{-l}$.

Hinweis: Benutzen Sie die boolsche Ungleichung aus den Folien.

(c) Zeigen Sie, dass $\Pr\left[\overline{K}_u \mid K_h\right] \geq 1 - 2^{-l+1}$.

Lösung:

(a) Da die X_i unabhängig sind, sind auch die $h(X_i)$ von einander unabhängig. Damit kann die Formel aus Beispiel 20 direkt angewendet werden:

$$\Pr\left[\overline{K}_h\right] = \prod_{j=1}^{m} \left(1 - \frac{j-1}{2^l}\right) \leq e^{-\binom{m}{2} 2^{-l}}.$$

Wegen der Konvexität von e^{-x} folgt, dass $e^{-x} \leq e^{-0}(1-x) + e^{-1}x = 1 - (1 - e^{-1})x \leq 1 - 1/2x$ für $x \in [0,1]$, woraus sich die Abschätzung sofort ergibt.

(b) Wegen der Unabhängigkeit der $h(X_i)$ folgt:

$\Pr[K_h] = \Pr\left[\bigvee_{i \neq j} h(X_i) = h(X_j)\right] \leq \sum_{i \neq j} \Pr[h(X_i) = h(X_j)] = \binom{m}{2} \sum_{w \in \{0,1\}^l} \Pr[h(X_2) = w \wedge h(X_1) = w] = \binom{m}{2} 2^l \cdot 2^{-l} \cdot 2^{-l}$.

(c) Es gilt

$$\Pr[K_h] = \Pr[K_h \wedge K_u] + \Pr[K_h \wedge \neg K_u]\,,$$

also

$$\Pr[\neg K_u \mid K_h] = 1 - \frac{\Pr[K_h \mid K_u] \cdot \Pr[K_u]}{\Pr[K_h]} = 1 - \frac{\Pr[K_u]}{\Pr[K_h]}.$$

Offensichtlich gilt $\Pr[K_h \mid K_u] = 1$.

Für $\binom{m}{2} \leq 2^l$ folgt somit:

$$\Pr[\neg K_u \mid K_h] \geq 1 - \frac{\binom{m}{2} \cdot 2^{-2l}}{1/2\binom{m}{2} \cdot 2^{-l}} = 1 - 2^{-l+1}.$$

Aufgabe 4.4 **Abzugeben sind (a-i) und (a-ii).** **2P+3P**

(a) Die kleine Maxi hat vor Kurzem von ihrem Vater Xaver eine CD von Justus Nagetier geschenkt bekommen, obwohl sie sich viel lieber die CD von den Flambierten gewünscht hätte. Dummerweise zwingt Xaver sie auch noch die ganze Zeit, sich diese CD anzuhören.

Um sich davon abzulenken, hat sie sich das folgende spannende Spiel ausgedacht:

Sie wirft eine faire Münze, bis zum ersten Mal das Muster 110 auftritt (1 sei Zahl, 0 sei Kopf).

(i) Wie oft muss Maxi die Münze im Mittel werfen, bis das Spiel endet?

Hinweis: Orientieren Sie sich an den Rechenwegen aus den Beispielen 14 und 15.

(ii) Bestimmen Sie auch die Varianz der entsprechenden ZV aus (i).

Hinweis: Die Varianz lässt sich mit Hilfe des Erwartungswerts darstellen. Gehen Sie dann wie in (i) vor.

(b) Nachdem der kleine Michel zu häufig auf 4chan gelandet ist, lässt ihn sein Vater Xaver nicht mehr an den Computer. Notgedrungen muss er mit seiner kleinen Schwester Maxi spielen:

Michel ist natürlich auch sofort von Maxis Spiel begeistert. Sie wandeln das Spiel daher so ab, dass sie die faire Münze werfen, bis das erste Mal entweder das Muster 110 oder 100 auftritt. Im Fall von 110 gewinnt Maxi, im Fall von 100 gewinnt Michel.

(i) Wie viele Würfe dauert ein Spiel im Mittel?

(ii) Nach ein paar Spielen fängt Michel an, seine kleine Schwester zu beschimpfen, dass sie bei ihrem dummen Spiel viel häufiger gewinnen würde.

Stimmt das?

Lösung:

(a) (i) Ansatz wie in der Vorlesung: Man partitioniert die Elementarereignisse geeignet nach ihrem Präfix. Der Präfix entspricht dabei dem Zustand eines minimalen DFAs, der genau die Wörter über $\{0,1\}$ erkennt, welche genau einmal, nämlich am Ende, 110 enthalten. Man entrollt zuerst den Erwartungswert anhand des Präfix (ersten beiden Spalten), dann löst man das so erhaltene LGS (drittte Spalte). Für die Übungen bietet es sich an, auch das entsprechende Markov-Diagramm/DFA anzuzeichnen mit den Zuständen $\varepsilon, 1, 11, 110$. Für die formale Begründung, warum man so rechnen darf, siehe Beispiel 14 aus den Folien.

$$\begin{array}{lll}
\mathbb{E}[W] & = \mathbb{E}[W|0]1/2 + \mathbb{E}[W|1]1/2 & = 1/2 + \mathbb{E}[W]1/2 + 3/2 + \mathbb{E}[W]1/4 = 8 \\
\mathbb{E}[W|0] & = \mathbb{E}[W+1] = \mathbb{E}[W] + 1 & = 9 \\
\mathbb{E}[W|1] & = \mathbb{E}[W|10]1/2 + \mathbb{E}[W|11]1/2 & = 1 + \mathbb{E}[W]1/2 + 2 = 7 \\
\mathbb{E}[W|10] & = \mathbb{E}[W+2] = \mathbb{E}[W] + 2 & = 10 \\
\mathbb{E}[W|11] & = \mathbb{E}[W|110]1/2 + \mathbb{E}[W|111]1/2 & = 3/2 + 1/2 + \mathbb{E}[W|11]1/2 = 4 \\
\mathbb{E}[W|110] & = 3 & \\
\mathbb{E}[W|111] & = \mathbb{E}[W+1|11] = \mathbb{E}[W|11] + 1 & = 5
\end{array}$$

(ii) Wegen $\mathrm{Var}[W] = \mathbb{E}[W^2] - \mathbb{E}[W]^2$ verfährt man mit $\mathbb{E}[W^2]$ wie in (i). Man muss nur beachten, dass z.B.

$$\mathbb{E}[W^2|111] = \mathbb{E}[(W+1)^2|11] = \mathbb{E}[W^2|11] + 2\mathbb{E}[W|11] + 1$$

gilt:

$$
\begin{array}{lll}
\mathbb{E}[W^2] & = \mathbb{E}[W^2|0]1/2 + \mathbb{E}[W^2|1]1/2 & = \mathbb{E}[W^2]1/2 + 17/2 + \mathbb{E}[W^2]1/4 + 27/2 = 88 \\
\mathbb{E}[W^2|0] & = \mathbb{E}[(W+1)^2] = \mathbb{E}[W^2] + 2\mathbb{E}[W] + 1 = \mathbb{E}[W^2] + 17 & \\
\mathbb{E}[W^2|1] & = \mathbb{E}[W^2|10]1/2 + \mathbb{E}[W^2|11]1/2 & = \mathbb{E}[W^2]1/2 + 18 + 9 \\
\mathbb{E}[W^2|10] & = \mathbb{E}[(W+2)^2] = \mathbb{E}[W^2] + 4\mathbb{E}[W] + 4 = \mathbb{E}[W^2] + 36 & \\
\mathbb{E}[W^2|11] & = \mathbb{E}[W^2|110]1/2 + \mathbb{E}[W^2|111]1/2 & = 9/2 + \mathbb{E}[W^2|11]1/2 + 9/2 = 18 \\
\mathbb{E}[W^2|110] & = 9 & \\
\mathbb{E}[W^2|111] & = \mathbb{E}[(W+1)^2|11] = \mathbb{E}[W^2|11] + 9 &
\end{array}
$$

Somit: $\mathrm{Var}[W] = 88 - 64 = 24$.

(b) (i)

$$
\begin{array}{lll}
\mathbb{E}[E] & = \mathbb{E}[E|0]1/2 + \mathbb{E}[E|1]1/2 & = \mathbb{E}[E]1/2 + 1/2 + 13/6 = 16/3 \\
\mathbb{E}[E|0] & = \mathbb{E}[E+1] = \mathbb{E}[E] + 1 & \\
\mathbb{E}[E|1] & = \mathbb{E}[E|10]1/2 + \mathbb{E}[E|11]1/2 & = 3/4 + \mathbb{E}[E|1]1/4 + 1/2 + 2 = 13/3 \\
\mathbb{E}[E|10] & = \mathbb{E}[E|100]1/2 + \mathbb{E}[E|101]1/2 & = 3/2 + \mathbb{E}[E|1]1/2 + 1 \\
\mathbb{E}[E|100] & = 3 & \\
\mathbb{E}[E|101] & = \mathbb{E}[E+2|1] = \mathbb{E}[E|1] + 2 & \\
\mathbb{E}[E|11] & = \mathbb{E}[E|110]1/2 + \mathbb{E}[E|111]1/2 & = 3/2 + \mathbb{E}[E|11]1/2 + 1/2 = 4 \\
\mathbb{E}[E|110] & = 3 & \\
\mathbb{E}[E|111] & = \mathbb{E}[E+1|11] = \mathbb{E}[E|11] + 1 &
\end{array}
$$

(ii)

$$
\begin{array}{lll}
\Pr[G] & = \Pr[G|0]\,1/2 + \Pr[G|1]\,1/2 & = \Pr[G]\,1/2 + 1/3 = 2/3 \\
\Pr[G|0] & = \Pr[G] & = 1 \\
\Pr[G|1] & = \Pr[G|11]\,1/2 + \Pr[G|10]\,1/2 & = 1/2 + \Pr[G|1]\,1/4 = 2/3 \\
\Pr[G|11] & = \Pr[G|110]\,1/2 + \Pr[G|111]\,1/2 & = 1/2 + \Pr[G|11]\,1/2 = 1 \\
\Pr[G|110] & = 1 & \\
\Pr[G|111] & = \Pr[G|11] & = 1 \\
\Pr[G|10] & = \Pr[G|101]\,1/2 + \Pr[G|100]\,1/2 & = \Pr[G|1]\,1/2 \\
\Pr[G|101] & = \Pr[G|1] & = 2/3 \\
\Pr[G|100] & = 0 &
\end{array}
$$

Es sollte auch intuitiv klar sein, dass $\Pr[G|11] = 1$, da ab hier Maxi nur noch warten muss, bis die 0 kommt.