Diskrete Wahrscheinlichkeitstheorie - Test 2

Sommersemester 2007

Name: ______________________________

Vorname: ______________________________

Matrikelnr.: ______________________________

Studiengang: ______________________________

Hinweise

- Sie sollten insgesamt 4 Blätter erhalten haben.
- Tragen Sie bitte Ihre Antworten in den dafür jeweils vorgesehenen Platz auf den Aufgabenblättern ein.
- Sie können maximal 10 Punkte erreichen.
- Füllen Sie die untenstehende Tabelle bitte nicht aus.
- Viel Erfolg.

A1	A2	A3	$\sum$

Aufgabe 1 $\quad$ 1.5P + 1.5P = 3 P

Miro und Ballack sind zwei Fußballspieler. In einem beliebigen Fußballspiel schießt Miro mit Wahrscheinlichkeit $1-p$ kein Tor, Ballack mit Wahrscheinlichkeit $1-p'$. Seien M bzw. B die Zufallsvariablen, die angeben, wie viele Spiele Miro bzw. Ballack brauchen, um zum ersten Mal ein Tor zu erzielen. Dann ist $M \sim \mathrm{Geo}(p)$ und $B \sim \mathrm{Geo}(p')$. Da die beiden in verschiedenen Ligen spielen, sind M und B unabhängig voneinander.

1.a) Sei $Z := \min\{M, B\}$ die Zahl der Spiele, die vergeht, bis einer von beiden ein Tor schießt. Zeigen Sie, dass auch Z geometrisch verteilt ist, und berechnen Sie den Parameter der Verteilung.

Antwort:

- 1. Möglichkeit:

  $M = k$ bedeutet, dass Miro erst im k.ten Spiel trifft, $B = l$, dass Ballack erst im l.ten Spiel trifft. Es ist dabei egal, ob die Spiele gleichzeitig, oder zeitlich versetzt stattfinden!

  $Z = \min\{B, M\}$ zählt damit einfach die Spieltage, bis schließlich einer der beiden das erste Mal trifft. Die Wahrscheinlichkeit, dass beide in ihrem jeweiligen i.ten Spiel nicht treffen, ist

  $$(1-p)(1-p').$$

  Die Wahrscheinlichkeit, dass mindestens einer der beiden trifft, ist also

  $$1-(1-p)(1-p') = p + p' - pp'.$$

  Da $Z = k$ nun gerade bedeutet, dass Miro und Ballack in ihren ersten $k-1$ Spielen jeweils nicht treffen, zumindest einer von beiden jedoch im k.ten Spiel, folgt

  $$\Pr[Z = k] = ((1-p)(1-p'))^{k-1}(p + p' - pp'),$$

  d.h. Z ist geometrisch verteilt mit Erfolgswahrscheinlichkeit $p + p' - pp'$.

- 2. Möglichkeit:

  Man zeigt zunächst, dass Z gedächtnislos ist für $k, l > 0$

  $$\begin{aligned}
\Pr[Z > k+l \mid Z > k+l] &= \tfrac{\Pr[M>k+l, B>k+l, M>l, B>l]}{\Pr[M>l, B>l]} \\
&= \tfrac{\Pr[M>k+l, B>k+l]}{\Pr[M>l, B>l]} \\
&= \tfrac{\Pr[M>k+l]}{\Pr[M>l]} \cdot \tfrac{\Pr[B>k+l]}{\Pr[B>l]} \\
&= \Pr[M > k+l \mid M > l] \cdot \Pr[B > k+l \mid B > l] \\
&= \Pr[M > k] \cdot \Pr[B > k] \\
&= \Pr[Z > k]
\end{aligned}$$

  Nach der Übung gilt dann, dass $Z \sim \mathrm{Geo}(\Pr[Z = 1])$ verteilt sind. Dabei gilt

  $$\Pr[Z = 1] = \Pr[M = 1, B > 1] + \Pr[M > 1, B = 1] + \Pr[M = 1, B = 1] = p + p' - pp'.$$

- 3. Möglichkeit:

  Man rechnet einfach nach entsprechend z.B. Aufgabe 1.1 oder der zweiten Aufgabe im ersten Test oder dem Beweis aus der Vorlesung bzgl. des Minimums zweier exp-verteilter, unabhängiger Zufallsvariablen:

  $$\Pr[Z \geq k] = \Pr[M \geq k, B \geq k] = \Pr[M \geq k] \cdot \Pr[B \geq k] = (1-p)^{k-1} \cdot (1-p')^{k-1},$$

  also

  $$\Pr[Z = k] = \Pr[Z \geq k] - \Pr[Z \geq k+1] = \left((1-p)(1-p')\right)^{k-1} \cdot (p + p' - pp').$$

  Es folgt wieder $Z \sim \mathrm{Geo}(p + p' - pp')$.

1.b) Berechnen Sie die Wahrscheinlichkeit, dass Miro vor Ballack ein Tor schießt, d.h. $\Pr[M < B]$.

Antwort:

Man verwendet hier am besten den Satz über die totale W'keit bzgl. der Ereignisse $[M = k]$:

$$\begin{aligned}
\Pr[M < B] &= \textstyle\sum_{k=1}^{\infty} \Pr[M = k, B \geq k+1] \\
&= \textstyle\sum_{k=1}^{\infty} p(1-p)^{k-1} \cdot (1-p')^k \\
&= p(1-p') \textstyle\sum_{k=1}^{\infty} \left((1-p)(1-p')\right)^{k-1} \\
&= \tfrac{p(1-p')}{1-(1-p)(1-p')} = \tfrac{p-pp'}{p+p'-pp'}.
\end{aligned}$$

Das Ereignis $[M = k, B \geq k+1]$ besagt dabei, dass Miro im k.ten Spiell trifft, Ballack jedoch frühestens im $k+1$.ten Spiel, d.h. Ballack darf in den ersten k Spielen nicht treffen - das geschieht gerade mit W'keit $(1-p')^k$, wobei man noch die Unabhängigkeit der Zufallsvariablen verwenden muss.

Aufgabe 2 $\quad$ 1P + 2P + 1P = 4 P

In einem zu Beginn schwarzen Bild der Auflösung $h \cdot b$ wird n-mal rein zufällig ein Pixel gleichverteilt ausgewählt und auf die Farbe Weiß gesetzt. Ein Pixel kann somit durchaus mehrmals ausgewählt werden.

Für $1 \leq i \leq h \cdot b$ sei X_i die Zufallsvariable, welche angibt, wie viele Pixel zufällig gewählt werden müssen, wenn bereits $i-1$ weiß eingefärbt sind, um noch einen schwarzen Pixel zu erhalten. Die Summe $S := \sum_{i=1}^{h \cdot b} X_i$ gibt dann gerade an, wie oft Pixel zufällig gewählt werden müssen, bis schließlich jeder Pixel auf Weiß gesetzt wurde.

2.a) Geben Sie die Dichte der Zufallsvariable X_i an. Begründen Sie Ihre Antwort.

Antwort:

Nach Aufgabenstellung zählt X_i, wie oft Pixel zufällig gezogen werden, bis schließlich ein schwarzer Pixel ausgewählt wird, wobei genau $i-1$ Pixel weiß sind.

D.h. $X_i = k$ bedeutet gerade, dass zunächst $k-1$ mal ein weißer Pixel gezogen wurde und schließlich der k.te Pixel schwarz war.

Da jeder Pixel mit derselben W'keit $\frac{1}{hb}$ gezogen wird (Gleichverteilung!), wird mit W'keit $\frac{i-1}{hb}$ ein weißer Pixel gezogen, mit W'keit $1 - \frac{i-1}{hb}$ ein schwarzer.

X_i ist daher geometrisch verteilt mit Erfolgswahrscheinlichkeit $p_i := 1 - \frac{i-1}{hb}$.

2.b) Schätzen Sie nun den Erwartungswert von S für $h = b = 10$ mittels

$$\sum_{i=1}^{k} \frac{1}{i} \approx 2.30 \cdot \log_{10} k + 0.58$$

ab. $\log_{10} x$ ist hierbei der Logarithmus von x zur Basis 10. Begründen Sie Ihre Rechenschritte.

Antwort:

Man nutzt zunächst die Linearität des Erwartungswerts aus und erhält:

$$\mathbb{E}[S] = \sum_{i=1}^{hb} \mathbb{E}[X_i]$$

Da X_i gemeometrisch verteilt ist mit Erfolgsw'keit $p_i = 1 - \frac{i-1}{hb} = \frac{hb-i+1}{hb}$ folgt

$$\mathbb{E}[X_i] = \frac{1}{p_i} = \frac{hb}{hb-i+1}.$$

Insgesamt ergibt sich also

$$\mathbb{E}[S] = \sum_{i=1}^{hb} \frac{hb}{hb-i+1} = hb \sum_{i=1}^{hb} \frac{1}{hb-i+1}.$$

Man muss entweder sehen oder sich aus der Vorlesung oder den Übungen daran erinnern, dass

$$\sum_{i=1}^{hb} \frac{1}{hb-i+1} = \sum_{i=1}^{hb} \frac{1}{i}$$

gilt.

Anwenden der Approximation liefert also

$$\mathbb{E}[S] = hb\,(2.3 \cdot \log_{10}(hb) + 0.58) = 100\,(4.6 + 0.58) = 518.$$

2.c) Ermitteln Sie mit Hilfe der Markov-Ungleichung

$$\Pr[S \geq t] \leq \frac{\mathbb{E}[S]}{t},$$

eine untere Schranke n^* für die Anzahl n von Pixeln, die zufällig auf Weiß gesetzt werden müssen, damit für $n \geq n^*$ das Bild mit Wahrscheinlichkeit ≥ 0.99 vollkommen weiß ist. Dabei gelte wieder $h = b = 10$.

Antwort:

Nach Aufgabenstellung werden zufällig insgesamt n Pixel gewählt und auf weiß gesetzt. Gesucht ist nun ein n^*, so dass für $n \geq n^*$ mit einer Wahrscheinlichkeit ≥ 0.99 das gesamte Bild weiß ist.

Wie viele Pixel wir insgesamt ziehen müssen, damit das Bild komplett weiß ist, gibt gerade S an. Gesucht ist daher ein möglichst kleines n^* mit

$$\Pr[S \leq n^*] \geq 0.99].$$

Wenn wir dann $n \geq n^*$ Pixel zufällig wählen, so ist mit W'keit ≥ 0.99 das gesamte Bild dann weiß.

Hierfür soll die Markov-Ungleichung verwendet werden:

$$\Pr[S \leq n] = 1 - \Pr[S > n] = 1 - \Pr[S \geq n+1] \geq 1 - \frac{\mathbb{E}[S]}{n+1} = 1 - \frac{518}{n+1} \overset{!}{\geq} 0.99.$$

Auflösen nach n führt auf

$$n \geq 51800 - 1 = 51799 =: n^*.$$

Aufgabe 3 **3 P**

Die Zufallsvariable Y gebe die Zeit an, die ein Professor braucht, um einen Beweis zu erklären. Hierbei nehmen wir an, dass Y die Dichte

$$f_Y(y) = \begin{cases} \lambda \cdot e^{-\lambda(y-1)} & \text{für } y \geq 1 \\ 0 & \text{sonst} \end{cases}$$

besitzt. Laut unbestätigten Untersuchungen ist der Prozentsatz der Studenten, welche dem Beweis bis zum Schluss folgen, dann durch die Zufallsvariable $Z := 1/Y$ gegeben.

Bestimmen Sie die Dichte der Zufallsvariable Z.

Antwort:

Die Aufgabe orientiert sich an A8.1: wieder ist nur die Verteilung bzw. die Dichte von Y und wie Z von Y abhängt bekannt.

In den Übungen wurde hierfür der Lösungsweg vorgestellt, dass zunächst die Verteilung von Z auf die Verteilung von Y zurückgeführt wird. Danach kann die Dichte durch Ableiten bestimmt werden.

Wegen $Z = 1/Y$ und $Y \in [1, \infty)$ folgt $Z \in (0, 1]$. Für $z \leq 0$ folgt $F_Z(z) = 0$, für $z > 1$ folgt $F_Z(z) = 1$. Es bleibt $z \in (0, 1]$:

$$F_Z(z) = \Pr[Z \leq z] = \Pr[1/Y \leq z] = \Pr[\frac{1}{z} \leq Y] = 1 - \Pr[Y < \frac{1}{z}] = 1 - F_Y(\frac{1}{z}).$$

Für $F_Y(y)$ mit $y \in [1, \infty)$ gilt

$$F_Y(y) = 1 - e^{-\lambda \cdot (y-1)},$$

also

$$F_Z(z) = e^{\lambda \cdot (1 - \frac{1}{z})}.$$

Die Dichte über $(0, 1]$ ergibt sich dann durch Ableiten zu:

$$e^{\lambda \cdot (1 - \frac{1}{z})} \cdot \frac{\lambda}{z^2}.$$

Insgesamt also

$$f_Z(z) = \begin{cases} e^{\lambda \cdot (1 - \frac{1}{z})} \cdot \frac{\lambda}{z^2} & \text{falls } z \in (0, 1] \\ 0 & \text{sonst} \end{cases}$$

Anmerkung:

Leider liegt das **Missverständnis** vor, dass auch bei stetig verteilten Zufallsvariablen $f_Y(y) = \Pr[Y = y]$ gelten würde - das ist aber **falsch**!

Man macht sich das am besten wie folgt klar: Sei X eine beliebige, stetig verteilt Zufallsvariable, dann gilt

$$\Pr\left[X \in [x - h, x + h]\right] = \Pr[x - h \leq X \leq x + h] = F_X(x + h) - F_X(x - h).$$

Lässt man nun $h \geq 0$ gegen 0 gehen, so folgt

$$\Pr[X = x] = \lim_{h \to 0} \Pr[x - h \leq X \leq x + h] = \lim_{h \to 0} (F_X(x + h) - F_X(x - h)) = F_X(x) - F_X(x) = 0.$$

Würde also $\Pr[X = x] = f_X(x)$ gelten, so müsste $f_X(x) = 0$ für alle $x \in \mathbb{R}$ folgen - was natürlich nicht sein kann, da $\int_{-\infty}^{\infty} f_X(x)dx = 1$ gelten muss.

Eine weiter Möglichkeit, sich diesen Sachverhalt zu merken ist, dass eine Dichte durchaus Werte größer 1 annehmen kann (siehe z.B. die Exponentialverteilung für $\lambda > 1$, dort gilt $\lambda e^{-\lambda \cdot 0} = \lambda > 1$).

Somit kann eine Dichte (im Fall stetiger Verteilungen) i.A. keine W'keit angeben! Der Ausdruck *Dichte* bezieht sich einzig darauf, dass sich die Wahrscheinlichkeit für ein gewisses Ereignis durch Summation (d.h. Integration) über die Dichte ergibt - analog zum diskreten Fall.