Prüfung 29 Juli 2014, Fragen und Antworten - Abschlussklausur

Diskrete Wahrscheinlichkeitstheorie (IN0018) (Technische Universität München)

Sommersemester 2014
Abschlussklausur
29. Juli 2014

Diskrete Wahrscheinlichkeitstheorie

— Lösung —

Aufgabe 1 (3 Punkte)

Wahr oder falsch? Begründen Sie Ihre Antwort!

1. Die Menge $\mathcal{B}(\mathbb{R})$ der Borelschen Mengen über $\mathbb{R}$ enthält alle Intervalle $[a,\ b]$ mit $a,\ b \in \mathbb{R}$ und $a \leq b$.
2. Seien X_1 und X_2 unabhängige exponentialverteilte Zufallsvariablen. Dann ist die Summe $X_1 + X_2$ ebenfalls exponentialverteilt.
3. Der Maximum-Likelihood-Schätzwert für den Parameter δ einer Verteilungsdichte $f(x;\delta)$ ist definiert als der Erwartungswert für δ.

Lösung

1. Wahr! Begründung: $\mathcal{B}(\mathbb{R})$ enthät alle Intervalle $[a,\ b]$ mit $a,\ b \in \mathbb{R}$, also auch die Teilmenge dieser Intervalle, für die $a \leq b$ gilt.
2. Falsch! Begründung: Sei $X_1 \sim \text{Exp}(1),\ X_2 \sim \text{Exp}(1),\ Z \sim \text{Exp}(\beta)$. $X_1 + X_2$ und Z seien identisch verteilt, dann gilt:

   $$\mathbb{E}[X1 + X2] \overset{\text{unabh.}}{=} \mathbb{E}[X_1] + \mathbb{E}[X_2] \overset{!}{=} \mathbb{E}[Z]$$
   $$\Rightarrow 1 + 1 = \frac{1}{\beta} \Rightarrow \beta = \frac{1}{2}$$
   $$f_{X_1+X_2}(k) \overset{!}{=} f_Z(k)$$
   $$f_{X_1+X_2}(k) = 2 \cdot e^{-k} \neq \frac{1}{2} \cdot e^{-\frac{1}{2}\cdot k} = f_Z(k)$$

   Widerspruch! $X_1 + X_2$ und Z können nicht identisch verteilt sein, also kann $X_1 + X_2$ nicht exponentialverteilt sein.
3. Falsch! Begründung: δ ist keine Zufallsvariable.

Aufgabe 2 (6 Punkte)

Wir betrachten 5 Körbe, die jeweils weiße und schwarze Bälle enthalten. 3 von diesen Körben (Variante A) enthalten je 4 Bälle, nämlich 3 weiße Bälle und 1 schwarzen Ball. Die 2 übrigen Körbe (Variante B) enthalten je 3 Bälle, nämlich 2 schwarze Bälle und 1 weißen Ball.

Ein Gewinnspiel bestehe darin, dass ein Spieler zunächst mit Laplace-Wahrscheinlichkeit einen Korb wählt und anschließend aus dem Korb mit Laplace-Wahrscheinlichkeit einen Ball zieht.

Nun rät der Spieler, ob der Ball aus einem Korb mit mehr weißen Bällen gezogen wurde, d.h., ob der Korb zur Variante A gehört, oder andernfalls zu B gehört. Falls richtig geraten wurde, erhält der Spieler 1 Euro, andernfalls muss er 2 Euro zahlen.

Wir nehmen an, dass der Spieler stets den Korb der Variante A rät, falls er einen weißen Ball gezogen hat, und andernfalls die Variante B rät.

Sei X die Zufallsvariable des Spielergebnisses in Euro mit $W_X = \{1, -2\}$.

1. Berechnen Sie die Wahrscheinlichkeiten $\Pr[w]$ und $\Pr[s]$ für die Ereignisse, dass ein weißer bzw. schwarzer Ball gezogen wird.
2. Berechnen Sie die bedingten Wahrscheinlichkeiten $\Pr[A|w]$ und $\Pr[B|s]$, dass ein weißer Ball aus einem Korb der Variante A bzw. ein schwarzer Ball aus einem Korb der Variante B gezogen wurde.
3. Berechnen Sie den Erwartungswert $\mathbb{E}[X]$.

Lösung

1. $\Pr[w] = \frac{7}{12}$; $\Pr[s] = \frac{5}{12}$

   Rechenweg:

   $\Pr[w] = \Pr[A] \cdot \Pr[w|A] + \Pr[B] \cdot \Pr[w|B] = \frac{3}{5} \cdot \frac{3}{4} + \frac{2}{5} \cdot \frac{1}{3} = \frac{27+8}{60} = \frac{7}{12}$

   $\Rightarrow \Pr[s] = \frac{5}{12}$

2. $\Pr[A|w] = \frac{27}{35}$; $\Pr[B|s] = \frac{16}{25}$

   Rechenweg:

   $\Pr[A|w] = \frac{\Pr[w|A] \cdot \Pr[A]}{\Pr[w]} = \frac{\frac{3}{4} \cdot \frac{3}{5}}{\frac{7}{12}} = \frac{9 \cdot 12}{20 \cdot 7} = \frac{27}{35}$

   $\Pr[B|s] = \frac{\Pr[s|B] \cdot \Pr[B]}{\Pr[s]} = \frac{\frac{2}{3} \cdot \frac{2}{5}}{\frac{5}{12}} = \frac{4 \cdot 12}{15 \cdot 5} = \frac{16}{25}$

3. $\mathbb{E}[X] = \frac{3}{20}$

   Rechenweg:

   $\mathbb{E}[X] = 1 \cdot \Pr[\text{„Sieg"}] + (-2) \cdot \Pr[\text{„Verlust"}] \overset{\text{NR 1}}{=} 1 \cdot \frac{43}{60} - 2 \cdot \frac{17}{60} = \frac{43-34}{60} = \frac{3}{20}$

   NR 1: $\Pr[\text{„Sieg"}] = \Pr[w] \cdot \Pr[A|w] + \Pr[s] \cdot \Pr[B|s] = \frac{7}{12} \cdot \frac{27}{35} + \frac{5}{12} \cdot \frac{16}{25} = \frac{27+16}{60} = \frac{43}{60}$

   $\Rightarrow \Pr[\text{„Verlust"}] = \frac{17}{60}$

Aufgabe 3 (6 Punkte)

Wir nehmen an, dass sich in einem vorliegenden Kartenspiel mit 32 Karten 16 rote und 16 schwarze Karten befinden. Ein Geber verteilt an 2 Spieler A und B je 3 Karten. Die Zufallsvariablen X_A bzw. X_B zählen die roten Karten für A bzw. B.

1. Welchen Wert hat $\Pr[X_A \geq 2]$, d.h., mit welcher Wahrscheinlichkeit bekommt A mindestens 2 rote Karten?
2. Wie hoch ist die Wahrscheinlichkeit, dass sowohl Spieler A genau 2 rote Karten bekommt als auch Spieler B genau 2 rote Karten bekommt?

Lösung

1. $\Pr[X_A \geq 2] = 1 - \frac{\binom{16}{3}+16\cdot\binom{16}{2}}{\binom{32}{3}}$

   Rechenweg: Es gilt $X_A \sim$ Hyp(3,16,16). Daraus folgt:
   $\Pr[X_A \geq 2] = 1$ - $(\Pr[X_A = 0] + \Pr[X_A = 1]) =$
   $1 - \left(\frac{\binom{16}{0}\binom{16}{3-0}}{\binom{16+16}{3}} + \frac{\binom{16}{1}\binom{16}{3-1}}{\binom{16+16}{3}} \right) = 1 - \frac{\binom{16}{3}+16\cdot\binom{16}{2}}{\binom{32}{3}}$

2. $\Pr[X_A = 2 \wedge X_B = 2] = \frac{\binom{16}{2}\cdot 16\cdot\binom{14}{2}\cdot 15}{\binom{32}{3}\binom{29}{3}}$

   Rechenweg: Betrachte zweistufiges Experiment, bei dem o.B.d.A. zuerst A 3 Karten zieht, dann B 3 Karten zieht: $X'_A \sim \text{Hyp}(3, 16, 16)$, $X'_B \sim \text{Hyp}(3, 15, 14)$ Daraus folgt:
   $\Pr[{X_A}' = 2 \wedge {X_B}' = 2] \overset{\text{unabh.}}{=} \Pr[{X_A}' = 2] \cdot \Pr[{X_B}' = 2] =$
   $\frac{\binom{16}{2}\binom{16}{3-2}}{\binom{16+16}{3}} \cdot \frac{\binom{14}{2}\binom{15}{3-2}}{\binom{15+14}{3}} = \frac{\binom{16}{2}\cdot 16\cdot\binom{14}{2}\cdot 15}{\binom{32}{3}\binom{29}{3}}$

Aufgabe 4 (6 Punkte)

Seien X und Y kontinuierliche Zufallsvariablen mit gemeinsamer Dichtefunktion

$$f_{X,Y}(x,y) = \begin{cases} 12x^2y^3 & : \quad 0 \leq x \leq 1,\ 0 \leq y \leq 1 \\ 0 & : \quad \text{sonst} \end{cases}$$

1. Berechnen Sie die Randdichte $f_X(x)$.
2. Sind die Variablen X und Y unabhängig? Begründung!
3. Bestimmen Sie den Wert der Verteilungsfunktion $F_{X,Y}(a,\ b)$ für alle $a,\ b \in [0,\ 1]$.

Lösung

1. $f_X(x) = 3x^2$

   Rechenweg:

   $f_X(x) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)dy \overset{\text{s.u.}}{=} \int\limits_0^1 12x^2y^3dy = 12x^2 \cdot \left[\frac{1}{4}y^4\right]_0^1 = 12x^2 \cdot \frac{1}{4} = 3x^2$

   Als Begrenzung für das Integral kann 0 bis 1 hergenommen werden, da

   $$\forall x, y \notin\ [0,1]\ : f_{X,Y}(x,y) = 0 \Rightarrow \int_{-\infty}^{0} f_{X,Y}(x,y)dy = \int_{1}^{\infty} f_{X,Y}(x,y)dy = 0$$

   (analog für x) Das Integral wird durch diese Bereiche also nicht im Wert verändert.

2. Die Variablen sind unabhängig.

   Begründung: Wenn die Variablen unabhängig sind, muss gelten

   $$\begin{aligned} f_{X,Y}(x,y) &\overset{!}{=} f_X(x) \cdot f_Y(y) \\ 12x^2y^3 &\overset{!}{=} 3x^2 \cdot 4y^3 \\ 12x^2y^3 &= 12x^2y^3 \end{aligned}$$

   NR: $f_Y(y) = \int\limits_{-\infty}^{\infty} f_{X,Y}(x,y)dx = \int\limits_0^1 12x^2y^3dx = 12y^3 \cdot \left[\frac{1}{3}x^3\right]_0^1 = 12y^3 \cdot \frac{1}{3} = 4y^3$

3. $F_{X,Y}(a,b) = a^3b^4$

   Rechenweg:

   $F_{X,Y}(a,b) = \int\limits_{-\infty}^{b} \int\limits_{-\infty}^{a} f_{X,Y}(x,y)dxdy \overset{\text{s.o.}}{=} \int\limits_0^b \int\limits_0^a 12x^2y^3dxdy = \int\limits_0^b 12y^3 \left[\frac{1}{3}x^3\right]_{x=0}^a dy =$

   $\int\limits_0^b \left(12y^3 \cdot \frac{1}{3}a^3 - 12y^3 \cdot 0\right) dy = \int\limits_0^b 4a^3y^3 = 4a^3 \left[\frac{1}{4}y^4\right]_{y=0}^b = 4a^3 \cdot \frac{1}{4}b^4 - 4a^3 \cdot 0 = a^3b^4$

Aufgabe 5 (6 Punkte)

Seien X_1 und X_2 unabhängige Bernoulli-verteilte Indikatorvariablen mit Erfolgswahrscheinlichkeit p_1 bzw. p_2. Wir betrachten die Zufallsvariable $Z = X_1 + X_2$.

1. Berechnen Sie Erwartungswert und Varianz von Z in Abhängigkeit der Parameter p_1 und p_2.
2. Geben Sie für $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{3}$ die wahrscheinlichkeitserzeugende Funktion $G_Z(z)$ als quadratisches Polynom in z an.
3. Seien N eine geometrisch verteilte Zufallsvariable mit Dichte $f_N(i) = \frac{2}{3}(\frac{1}{3})^{i-1}$ und Z_i, $i = 1, 2, \ldots$ eine unabhängige Zufallsvariable, die gleichverteilt sind wie Z für $p_1 = \frac{1}{4}$, $p_2 = \frac{1}{3}$. Sei $S = \sum_{i=1}^{N} Z_i$ die von N abhängige Summe der Z_i.

   Berechnen Sie den Erwartungswert für $\mathbb{E}[S]$.

Lösung

1. $\mathbb{E}[Z] = p_1 + p_2$; $\mathrm{Var}[Z] = p_1 \cdot (1 - p_1) + p_2 \cdot (1 - p_2)$

   Rechenweg: Da beide Zufallsvariablen unabhängig sind, gilt für ihre Summe:

   $\mathbb{E}[Z] = \mathbb{E}[X_1] + \mathbb{E}[X_2] = p_1 + p_2$

   $\mathrm{Var}[Z] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] = p_1 \cdot (1 - p_1) + p_2 \cdot (1 - p_2)$

2. $G_Z(s) = \frac{1}{2} + \frac{5}{12}s + \frac{1}{12}s^2$

   Rechenweg: Z ist die Summe zweier Indikatorvariablen. Diese können jeweils nur die Werte 0 oder 1 annehmen. Das heißt Z kann die Werte 0, 1 und 2 annehmen.

   Das heißt, wir benötigen folgende Wahrscheinlichkeiten:

k	$f_Z(k)$	s^k	**Bedeutung**
0	$(1 - p_1) \cdot (1 - p_2) = \frac{3}{4} \cdot \frac{2}{3} = \frac{1}{2}$	1	Beide sind 0.
1	$(1 - p_1) \cdot p_2 + p_1 \cdot (1 - p_2) = \frac{5}{12}$	s	Genau einer ist 1.
2	$p_1 \cdot p_2 = \frac{1}{12}$	s^2	Beide sind 1.

   $G_Z(s) = \sum_{k=0}^{\infty} f_Z(k) \cdot s^k \Rightarrow G_Z(s) = \frac{1}{2} + \frac{5}{12}s + \frac{1}{12}s^2$

3. $\mathbb{E}[S] = \frac{7}{8}$

   Rechenweg: $N \sim \mathrm{Geo}(\frac{2}{3}) \Rightarrow \mathbb{E}[N] = \frac{1}{2/3} = \frac{3}{2}$. Nach dem Satz zufälliger Summen gilt:

   $\mathbb{E}[S] = \mathbb{E}[N] \cdot \mathbb{E}[Z] = \frac{3}{2} \cdot (p_1 + p_2) = \frac{3}{2} \cdot \left(\frac{1}{4} + \frac{1}{3}\right) = \frac{3 \cdot 7}{2 \cdot 12} = \frac{7}{8}$

Aufgabe 6 (6 Punkte)

Seien X_i, $i = 1, 2, 3$ unabhängige Bernoulli-verteilte Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p und $S = \sum_{i=1}^{3} X_i$ eine Stichprobenvariable zum Test der Hypothese $H_0 : p \leq \frac{1}{5}$ mit Ablehnungsbereich $\tilde{K} = \{2, 3\}$.

1. Berechnen Sie die Verteilungsfunktion F_S von S in Abhängigkeit von p!
2. Berechnen Sie die Fehlerwahrscheinlichkeit 1. Art α_1!
3. Welchen Wert hat die Fehlerwahrscheinlichkeit 2. Art α_2 unter der Annahme der trivialen Alternative $H_1 : p > \frac{1}{5}$, wenn man einen leeren Ablehnungsbereich $\tilde{K}$ wählt?

Lösung

1. $F_S(k) = \sum_{i=0}^{k} \binom{3}{i} \cdot p^i \cdot (1-p)^{3-i}$

   Rechenweg: Für S als Summe von drei Bernoulli-verteilten Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p gilt: $S \sim \mathrm{Bin}(3, p)$.

   $F_S(k) = \Pr[S \leq k] = \sum_{i=0}^{k} f_X(i) = \sum_{i=0}^{k} \binom{3}{i} \cdot p^i \cdot (1-p)^{3-i}$

2. $\alpha_1 = \frac{13}{125}$

   Rechenweg:

   $\alpha_1 = \sup_{\theta \in H_0} \Pr_\theta[S \in \tilde{K}] \overset{\text{s.u.}}{=} 1 - \Pr_{\frac{1}{5}}[S \leq 1] \overset{\text{NR 1}}{=} 1 - \frac{112}{125} = \frac{13}{125}$

   Wir benötigen das Supremum, also die obere Schranke, von $\Pr_\theta[S \in \tilde{K}]$. Da wir die Fehlerwahrscheinlichkeit 1. Art (H_0 gilt, wird aber abgelehnt) ermitteln, ist diese Wahrscheinlichkeit natürlich direkt an der Grenze zu H_1 maximal, also bei $\theta = \frac{1}{5}$.

   NR 1: $\Pr_{\frac{1}{5}}[S \leq 1] = \Pr_{\frac{1}{5}}[S = 0] + \Pr_{\frac{1}{5}}[S = 1] =$

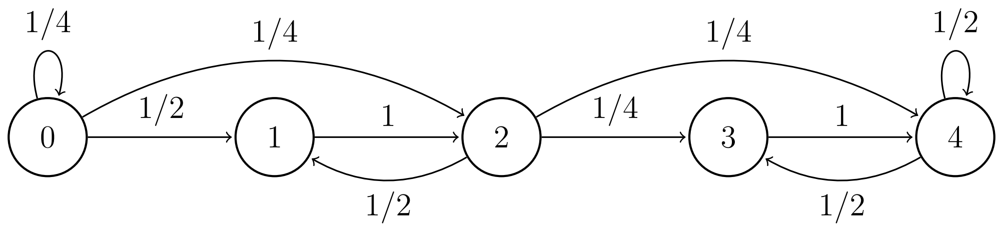
   $\binom{3}{0} \cdot (\frac{1}{5})^0 \cdot \left(1 - \frac{1}{5}\right)^3 + \binom{3}{1} \cdot (\frac{1}{5})^1 \cdot \left(1 - \frac{1}{5}\right)^2 = \left(\frac{4}{5}\right)^3 + 3 \cdot \frac{1}{5} \cdot \left(\frac{4}{5}\right)^2 = \frac{64}{125} + \frac{3 \cdot 16}{5 \cdot 25} = \frac{112}{125}$

3. $\alpha_2 = 1$

   Begründung: Die Fehlerwahrscheinlichkeit 2. Art ist die maximale Wahrscheinlichkeit dafür, dass H_0 angenommen wird, obwohl es *nicht* gilt. Da H_0 nur abgelehnt wird, wenn das Ergebnis ein Element des Ablehnungsbereiches ist, dieser aber in diesem Fall leer ist, nehmen wir H_0 immer an.

Aufgabe 7 (7 Punkte)

Sei $(X_t)_{t\in\mathbb{N}_0}$ eine endliche (zeit)homogene Markov-Kette mit diskreter Zeit über der Zustandsmenge $S = \{0, 1, 2, 3, 4\}$. Die positiven Übergangswahrscheinlichkeiten seien durch das folgende Übergangsdiagramm gegeben:



1. Geben Sie die Menge der transienten Zustände der Markov-Kette $(X_t)_{t\in\mathbb{N}_0}$ an. Begründung!
2. Berechnen Sie die Ankunftswahrscheinlichkeit f_{01}!
3. Berechnen Sie die erwartete Übergangszeit h_{24}!

Lösung

1. $\{0, 1, 2\}$

   Begründung: 3 und 4 bilden gemeinsam eine rekurrente Kommunikationsklasse (hat den gleichen Effekt wie ein absorbierender Zustand) und sind beide rekurrent.

2. $f_{01} = \frac{5}{6}$

   Rechenweg:
   $f_{01} = \sum_{k\in X_t} p_{0k} \cdot f_{k1} = p_{00} \cdot f_{01} + p_{01} \cdot f_{11} + p_{02} \cdot f_{21} + p_{03} \cdot f_{31} + p_{04} \cdot f_{41} =$

   $\frac{1}{4} \cdot f_{01} + \frac{1}{2} \cdot 1 + \frac{1}{4} \cdot f_{21} + 0 \cdot f_{31} + 0 \cdot f_{41} \overset{\text{NR 1}}{=} \frac{1}{4} \cdot f_{01} + \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} = \frac{1}{4} \cdot f_{01} + \frac{4+1}{8}$
   $\Rightarrow \frac{3}{4} \cdot f_{01} = \frac{5}{8} \Rightarrow f_{01} = \frac{5}{6}$

   NR 1: $f_{21} = p_{20} \cdot f_{01} + p_{21} \cdot f_{11} + p_{22} \cdot f_{21} + p_{23} \cdot f_{31} + p_{24} \cdot f_{41} \overset{\text{s.u.}}{=} \frac{1}{2} \cdot 1 = \frac{1}{2}$

   Aus der Tatsache, dass 3 und 4 eine rekurrente Kommunikationsklasse bilden, können wir folgern, dass $f_{31} = f_{41} = 0$ gilt.

3. $h_{24} = \frac{7}{2}$

   Rechenweg:
   $h_{24} = 1 + \sum_{k\in X_t} p_{2k} \cdot h_{k4} = 1 + p_{20} \cdot h_{04} + p_{21} \cdot h_{14} + p_{22} \cdot h_{24} + p_{23} \cdot h_{34} + p_{24} \cdot h_{44} =$

   $1 + 0 \cdot h_{04} + \frac{1}{2} \cdot h_{14} + 0 \cdot h_{24} + \frac{1}{4} \cdot h_{34} + \frac{1}{4} \cdot 0 \overset{\text{NR 2\&3}}{=} 1 + \frac{1}{2} \cdot (1 + h_{24}) + \frac{1}{4} \cdot 1 = 1 + \frac{1}{2} + \frac{1}{2} \cdot h_{24} + \frac{1}{4}$
   $\Rightarrow \frac{1}{2} \cdot h_{24} = \frac{7}{4} \Rightarrow h_{24} = \frac{7}{2}$

   NR 2: $h_{14} = 1 + p_{10} \cdot h_{04} + p_{11} \cdot h_{14} + p_{12} \cdot h_{24} + p_{13} \cdot h_{34} + p_{14} \cdot h_{44} = 1 + 1 \cdot h_{24} = 1 + h_{24}$

   NR 3: $h_{34} = 1 + p_{30} \cdot h_{04} + p_{31} \cdot h_{14} + p_{32} \cdot h_{24} + p_{33} \cdot h_{34} + p_{34} \cdot h_{44} = 1 + 1 \cdot 0 = 1$