

LÖSUNG

Wiederholungsklausur Diskrete Wahrscheinlichkeitstheorie

Sommersemester 2008

Hinweis: Alle Antworten sind zu begründen. Insbesondere sollte bei nicht-trivialen Umformungen kurz angegeben werden, weshalb diese Umformungen erlaubt sind (z.B.: Unabhängigkeit von ZV/Ereignissen, Disjunktheit von Ereignissen, Approximation mittels ZGWS, etc.)

Aufgabe 1

3P+2P+2P=7P

Es soll ein Codewort $X_1 X_2 \dots X_5$ bestehend aus fünf Zeichen X_i erzeugt werden, wobei jedes Zeichen X_i ein Großbuchstabe ist, und in welchem höchstens zwei Konsonanten bzw. Vokale aufeinanderfolgen.

Hierfür bezeichne Σ die 26 Großbuchstaben des deutschen Alphabets *ohne* Umlaute, also $\Sigma = \{A, B, C, \dots, Z\}$. Weiterhin bezeichne $\Sigma_V = \{A, E, I, O, U\}$ die Menge der Vokale und $\Sigma_K = \Sigma \setminus \Sigma_V$ die Menge der Konsonanten.

Das Codewort wird nun wie folgt erzeugt:

- Die ersten beiden Zeichen X_1 und X_2 werden zufällig gleichverteilt aus Σ gewählt.
 - Die verbleibenden drei Zeichen werden nacheinander erzeugt. Für $i = 3$ bis 5 gilt dabei:
 - Sind X_{i-2} und X_{i-1} Vokale ($X_{i-1}, X_{i-2} \in \Sigma_V$), so wird X_i zufällig gleichverteilt aus Σ_K gewählt.
 - Sind X_{i-2} und X_{i-1} Konsonanten ($X_{i-1}, X_{i-2} \in \Sigma_K$), so wird X_i zufällig gleichverteilt aus Σ_V gewählt.
 - Ansonsten wird X_i zufällig gleichverteilt aus Σ gewählt.
- a) Berechnen Sie die Wahrscheinlichkeit, dass $X_3 \in \Sigma_K$, d.h. dass das dritte Zeichen ein Konsonant ist.
- b) Berechnen Sie die bedingte Wahrscheinlichkeit, dass ein Codewort, das mit zwei Konsonanten beginnt, auch auf einen Konsonant endet.
- c) Berechnen Sie die erwartete Anzahl von Konsonanten unter der Bedingung, dass die ersten beiden Zeichen Vokale sind.

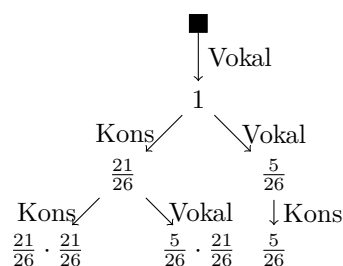
Hinweis: Es empfiehlt sich, die Bäume zu den mehrstufigen Experimenten aufzuzeichnen.

Lösungsvorschlag:

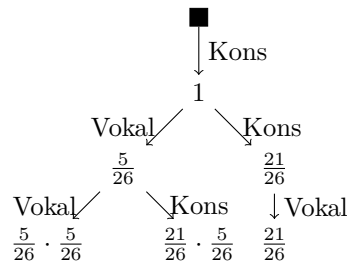
- a) Fallunterscheidung, ob man zu Beginn zwei Vokale oder einen Vokal und einen Konsonanten erzeugt.

$$\left(\frac{5}{26}\right)^2 \cdot 1 + 2 \cdot \frac{5}{26} \cdot \frac{21}{26} \cdot \frac{21}{26} \approx 0,288.$$

- b) Die gesuchte W'keit beträgt $\frac{21}{26} \cdot \frac{21}{26} + \frac{5}{26} = 1 - \frac{5}{26} \cdot \frac{21}{26}$, wie folgendem Entscheidungsbaum zu entnehmen ist:



- c) Den Entscheidungsbaum zu c) erhält man aus dem Baum zu b) durch Vertauschen der Label und der W'keiten:



Damit folgt für die erwartete Anzahl von Konsonanten unter der Bedingung, dass X_1 und X_2 Vokale sind:

$$1 \cdot \left(\frac{5}{26}\right)^2 + 2 \cdot \left(1 - \left(\frac{5}{26}\right)^2\right) = 2 - \left(\frac{5}{26}\right)^2.$$

Aufgabe 2

3P+3P+3P=9P

Der (relative) Anteil der Frauen unter den Informatik-Studenten der TU München sei $p = 0.1$ (mit $q := 1 - p$). Nehmen Sie an, dass zufällige Informatik-Studenten den Tag über das Service-Büro betreten.

Die Zufallsvariable Y gebe an, wie viele Studenten der Service-Büro-Mitarbeiter abwarten muss, bis zwei Frauen *hintereinander* das Büro betreten haben. Mit G sei die erzeugende Funktion von Y bezeichnet.

- a) Es sei M das Ereignis, dass ein männlicher Student als erstes das Büro betritt. Entsprechend bezeichne FM das Ereignis, dass zu Beginn des Experiments eine Studentin gefolgt von einem Studenten das Büro betreten. Das Ereignis FF sei entsprechend definiert.

Zeigen Sie unter Verwendung von

$$\Pr[Y = k + 3|M] = \Pr[Y = k + 2] \text{ und } \Pr[Y = k + 3|FM] = \Pr[Y = k + 1],$$

dass für alle $k \in \mathbb{N}$ gilt:

$$\Pr[Y = k + 3] = \Pr[Y = k + 2] \cdot q + \Pr[Y = k + 1] \cdot pq.$$

- b) Zeigen Sie, dass G für alle $z \in [0, 1]$ folgende Gleichung erfüllen muss:

$$G(z) = p^2 z^2 + qz G(z) + pqz^2 G(z).$$

- c) Geben Sie eine geschlossene Form für $G(z)$ (über $[0, 1]$) an und bestimmen Sie damit die Varianz von Y .

Hinweis: Verwenden Sie $G''(1) = 23780$ direkt. Sie müssen also nicht die zweite Ableitung berechnen.

Lösungsvorschlag:

- a) Die Ereignisse M, FM, FF bilden eine Partition von Ω . Damit folgt:

$$\Pr[Y = k + 3] = \Pr[Y = k + 3|M] \cdot \Pr[M] + \Pr[Y = k + 3|FM] \cdot \Pr[FM] + \Pr[Y = k + 3|FF] \cdot \Pr[FF].$$

Da $k \geq 0$ und $FF = \Pr[Y = 2]$ gilt:

$$\Pr[Y = k + 3|FF] = \frac{\Pr[Y = k + 3, Y = 2]}{\Pr[Y = 2]} = 0.$$

Schließlich gilt:

$$\Pr[M] = 1 - p =: q \text{ und } \Pr[F] = p.$$

Mit den angegebenen Beziehungen folgt die angegebene Rekursionsgleichung.

- b) Unter Beachtung von $\Pr[Y \leq 1] = 0$ und $\Pr[Y = 2] = p^2$ folgt (mit $q := 1 - p$):

$$\begin{aligned} G(z) &= \sum_{k \geq 0} z^k \cdot \Pr[Y = k] \\ &= z^2 \cdot \Pr[Y = 2] + \sum_{k \geq 0} z^{k+3} \cdot \Pr[Y = k + 3] \\ &= p^2 z^2 + \sum_{k \geq 0} z^{k+3} (q \cdot \Pr[Y = k + 2] + pq \cdot \Pr[Y = k + 1]) \\ &= p^2 z^2 + qz \sum_{k \geq 0} z^{k+2} \cdot \Pr[Y = k + 2] + pqz^2 \sum_{k \geq 0} z^{k+1} \cdot \Pr[Y = k + 1] \\ &= p^2 z^2 + qz G(z) + pqz^2 G(z). \end{aligned}$$

c) Auflösen der Gleichung aus b) liefert:

$$G(z) = \frac{p^2 z^2}{1 - qz - pqz^2} = \frac{z^2}{100 - 90z - 9z^2}.$$

Es folgt:

$$G'(z) = \frac{2z}{100 - 90z - 9z^2} - \frac{z^2}{(100 - 90z - 9z^2)^2}(-90 - 18z) = \frac{200z - 180z^2 - 18z^3 + 90z^2 + 18z^3}{(100 - 90z - 9z^2)^2} = \frac{200z - 90z^2}{(100 - 90z - 9z^2)^2}$$

Damit:

$$\text{Var}[Y] = G''(1) + G'(1) - (G'(1))^2 = 23780 + 110 - 12100 = 11790.$$

Aufgabe 3

2P+3P+3P=8P

Sie haben einen fairen 3-seitigen Würfel, der die Augenzahlen 1, 2 und 3 aufweist. Sie würfeln, bis alle drei Augenzahlen einmal gefallen sind. Sei X die Zahl der Würfe, die Sie dafür benötigen. Erinnern Sie sich aus der Vorlesung, dass $X = X_1 + X_2 + X_3$ gilt, wobei die X_i unabhängig und geometrisch verteilt sind mit Erfolgsw'keiten $(p_1, p_2, p_3) = (1, \frac{2}{3}, \frac{1}{3})$.

- Berechnen Sie $\mathbb{E}[X]$ und $\text{Var}[X]$.
- Bestimmen Sie mithilfe der Chebyshev-Ungleichung eine möglichst kleine natürliche Zahl n , sodass $\Pr[X \geq n] \leq 0.5$ gilt.
- Bestimmen Sie die Wahrscheinlichkeit dafür, dass Sie als erstes genau eine 3 würfeln, dann eine oder mehrere 2, und dann eine 1.

Lösungsvorschlag:

- a) Es gilt $\mathbb{E}[X_i] = \frac{1}{p_i}$ und $\text{Var}[X_i] = \frac{1 - p_i}{p_i^2}$.

i	1	2	3
$\mathbb{E}[X_i]$	1	$\frac{3}{2}$	3
$\text{Var}[X_i]$	0	$\frac{3}{4}$	6

Also $\mathbb{E}[X] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \mathbb{E}[X_3] = 1 + \frac{3}{2} + 3 = 5.5$ und wegen Unabhängigkeit $\text{Var}[X] = \text{Var}[X_1] + \text{Var}[X_2] + \text{Var}[X_3] = 0 + \frac{3}{4} + 6 = \frac{27}{4}$.

b)

$$\begin{aligned} \Pr[X \geq n] &= \Pr[X - \mathbb{E}[X] \geq n - \mathbb{E}[X]] \\ &\leq \Pr[|X - \mathbb{E}[X]| \geq n - \mathbb{E}[X]] \\ &\leq \frac{\text{Var}[X]}{(n - \mathbb{E}[X])^2} && \text{(Chebyshev-Ungleichung)} \\ &= \frac{27/4}{(n - 5.5)^2} && \text{(mit a)} \\ &\stackrel{!}{\leq} 0.5 \end{aligned}$$

führt zu $(n - 5.5)^2 \geq \frac{27}{2}$ oder $n \geq \sqrt{\frac{27}{2}} + 5.5 \approx 9.2$. Also $n = 10$.

- c) $\Omega = \{aa^k b(b|a)^l c \mid \{a, b, c\} = \{1, 2, 3\}, k, l \in \mathbb{N}\}$.

$$\Pr[\omega] = 3^{-|\omega|}.$$

Das beschriebene Ereignis ist $E = \{322^l 1\}$. Es gilt $\Pr[E] = \sum_{l \in \mathbb{N}} \left(\frac{1}{3}\right)^{3+l} = \left(\frac{1}{3}\right)^3 \cdot \frac{1}{1-\frac{1}{3}} = \left(\frac{1}{3}\right)^3 \cdot \frac{3}{2} = \frac{1}{18}$.

Die Fluggesellschaft EsparzAir bietet regelmäßig einen Flug von Stuttgart nach München an, wobei 96 Sitzplätze zur Verfügung stehen. Auf Grund (älterer) Statistiken vergangener Jahre geht EsparzAir davon aus, dass jeder zehnte Fluggast nicht zu diesem Flug erscheint.

Hinweis: Gehen Sie davon aus, dass sich die Fluggäste unabhängig von einander, aber ansonsten gleich verhalten.

- a) Wie viele Tickets kann die Fluggesellschaft für einen Flug von Stuttgart nach München mit 96 Plätzen maximal verkaufen, damit mit mindestens 95%-iger Wahrscheinlichkeit genug freie Plätze für alle erschienenen Passagiere zur Verfügung stehen?

Verwenden Sie den Zentralen Grenzwertsatz.

Hinweis: $z_{0.95} \approx 1.65$ und $z_{0.05} \approx -1.65$.

- b) Die Fluggesellschaft will die W'keit neu schätzen, dass ein Passagier *nicht* zum Flug erscheint. Dafür betrachtet sie die letzten vier Flüge Stuttgart-München. Bei den Flügen wurden stets alle 100 Tickets verkauft, wobei jeweils k_i viele Passagiere nicht zum Flug erschienen:

$$k_1 = 2 \quad k_2 = 4 \quad k_3 = 2 \quad k_4 = 5$$

Die Fluggesellschaft geht davon aus, dass die Anzahl der Personen, die nicht zum Flug erscheinen, $\text{Bin}(100, p)$ -verteilt ist, also binomial-verteilt mit Erfolgsw'keit p . Bestimmen Sie $p \in (0, 1)$ aus diesen Daten mit Hilfe des Maximum-Likelihood-Prinzips.

Hinweis: Das Maximum-Likelihood-Prinzip sollte klar ersichtlich sein. Sie müssen jedoch nicht die zweite Ableitung berechnen.

Lösungsvorschlag: Notation: $X_i \sim \text{Bin}(1; 0.9)$ unabhängig mit $X_i = 1$, falls i -ter Passagier rechtzeitig zum Flug kommt. $S_n := \sum_{i=1}^n X_i$ mit $\mathbb{E}[S_n] = n \cdot 0.9$ und $\text{Var}[S_n] = n \cdot 0.09$.

- a) Approximation mit Φ :

$$\Pr[S_n > 96] = \Pr\left[\frac{S_n - \mathbb{E}[S_n]}{\sqrt{\text{Var}[S_n]}} > \frac{96 - n \cdot 0.9}{0.3\sqrt{n}}\right] \approx 1 - \Phi\left(\frac{960 - 9n}{3\sqrt{n}}\right) \stackrel{!}{\leq} 0.05.$$

Mit $z_{0.95} \approx 1.65$ folgt

$$\begin{aligned} \Phi\left(\frac{960-9n}{3\sqrt{n}}\right) &\approx 0.95 \approx \Phi(1.65) \\ \rightsquigarrow \frac{960-9n}{3\sqrt{n}} &\approx 1.65 \\ \rightsquigarrow 9n + 4.95\sqrt{n} - 960 &\approx 0 \\ \rightsquigarrow \sqrt{n} &\approx \frac{-4.95 \pm 185.97}{18} \\ \rightsquigarrow n &\approx 101.14 \end{aligned}$$

- b) Es soll p so gewählt werden, dass das Produkt

$$\prod_{i=1}^4 \binom{100}{k_i} p^{k_i} (1-p)^{100-k_i}$$

maximiert wird. Die Binomialkoeffizienten sind dabei offensichtlich irrelevant.

Mit $k_1 + k_2 + k_3 + k_4 = 13$ ergibt sich

$$p^{13}(1-p)^{387}.$$

Ableiten nach p führt auf

$$13p^{12}(1-p)^{387} - 387p^{13}(1-p)^{386} = p^{12}(1-p)^{386}(13(1-p) - 387p) = p^{12}(1-p)^{371}(13 - 400p) \stackrel{!}{=} 0$$

bzw.

$$p = \frac{13}{400}.$$

Aufgabe 5

1P+2P+1P+2P+3P=9P

Die Adler spielen gegen die Bären Eishockey. Die Adler benötigen für ein Tor im Schnitt 12 Minuten, die Bären nur 10 Minuten. Nehmen Sie an, dass die Adler und die Bären unabhängig voneinander Tore schießen und dass die Wartezeit bis zum nächsten Tor der Adler (bzw. der Bären) exponential-verteilt ist mit Parameter λ_A (bzw. λ_B).

- Geben Sie λ_A und λ_B an.
- Sei T die Wartezeit bis zum nächsten Tor (egal wer es schießt). Geben Sie die Verteilung und den Erwartungswert der Zufallsvariablen T an.
- Berechnen Sie den Erwartungswert für die Zahl B_{40} der **Bären**-Tore in den ersten 40 Spielminuten.
- Berechnen Sie $\Pr[B_{40} \leq 1]$.
- Nehmen Sie jetzt an, dass es eine Verlängerung gibt. Sie dauert so lange, bis eine der Mannschaften ein Tor schießt. Berechnen Sie die Wahrscheinlichkeit, dass dieses Tor von den Adlern geschossen wird.
Hinweis: Sie benötigen dafür nur Teilaufgabe a).

Lösungsvorschlag:

- $\lambda_A = \frac{1}{12}$, $\lambda_B = \frac{1}{10}$.
- Seien T_A und T_B die Wartezeiten bis zum nächsten Adler-Tor und zum nächsten Bären-Tor. Laut Aufgabe sind T_A und T_B exponential-verteilt mit Parameter λ_A bzw. λ_B . Es gilt $T = \min\{T_A, T_B\}$. Laut Vorlesung ist T exponential-verteilt mit Parameter $\lambda = \lambda_A + \lambda_B = \frac{1}{12} + \frac{1}{10} = \frac{11}{60}$. Es gilt $\mathbb{E}[T] = 1/\lambda = \frac{60}{11}$.
- Laut Vorlesung ist B_{40} Poisson-verteilt mit Parameter $40 \cdot \lambda_B = 40 \cdot \frac{1}{10} = 4$. Also gilt $\mathbb{E}[B_{40}] = 4$.
-

$$\Pr[B_{40} \leq 1] = \Pr[B_{40} = 0] + \Pr[B_{40} = 1] = e^{-4} \left(\frac{4^0}{0!} + \frac{4^1}{1!} \right) = 5e^{-4}$$

- Seien f_A, f_B, F_A, F_B die Dichte- und Verteilungsfunktionen für die Wartezeiten T_A und T_B . Laut Aufgabenstellung und Teil a) handelt es sich um Exponentialverteilungen mit Parameter λ_A bzw. λ_B :

$$f_A(x) = \begin{cases} \lambda_A e^{-\lambda_A x} & x \geq 0 \\ 0 & x < 0 \end{cases} \quad F_A(x) = \begin{cases} 1 - e^{-\lambda_A x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

Es gilt:

$$\begin{aligned} \Pr[T_A \leq T_B] &= \int_{y=-\infty}^{\infty} \int_{x=-\infty}^y f_A(x) f_B(y) dx dy \\ &= \int_{y=-\infty}^{\infty} f_B(y) \left(\int_{x=-\infty}^y f_A(x) dx \right) dy \\ &= \int_{y=0}^{\infty} f_B(y) F_A(y) dy \\ &= \int_{y=0}^{\infty} (\lambda_B e^{-\lambda_B y} (1 - e^{-\lambda_A y})) dy \\ &= [-e^{-\lambda_B y}]_0^{\infty} + \frac{\lambda_B}{\lambda_A + \lambda_B} [e^{-(\lambda_A + \lambda_B)y}]_0^{\infty} \\ &= 1 - \frac{\lambda_B}{\lambda_A + \lambda_B} \\ &= \frac{\lambda_A}{\lambda_A + \lambda_B} \\ &= \frac{1}{12} / \left(\frac{1}{12} + \frac{1}{10} \right) = \frac{1}{12} \cdot \frac{60}{11} = \frac{5}{11} \end{aligned}$$