SS 2011

Zentralübung
Diskrete Wahrscheinlichkeitstheorie (zur Vorlesung Prof. Mayr)

Dr. Werner Meixner

Fakultät für Informatik
TU München

http://www14.in.tum.de/lehre/2011SS/dwt/uebung/

30. Juni 2011

ZÜ V

Übersicht:

1. Termine
2. Thema: Ausgewählte Aufgaben der Midterm
3. Vorbereitung auf Tutoraufgaben: VA 2 von Blatt 8

1. Termine

Die nächste und letzte Zentralübung findet statt am

Donnerstag, den 14. Juli 2011

2. Thema: Ausgewählte Aufgaben der Midterm

2.1 Vorbemerkungen

Die Midtermklausur hatte mittleren Schwierigkeitsgrad:

Aufgaben 1 und 2 waren meist sehr leicht bis leicht,

Aufgabe 1.6 war mittelschwer.

Aufgaben 4 und 5 waren leicht bis mittelschwer.

Aufgabe 3.3 war eher schwierig.

Ergebnisse: Insgesamt erfreulich!

2.2 Aufgabe 3

Wir werfen gleichzeitig und unabhängig mit einem blauen und einem roten fairen Würfel und definieren eine diskrete Zufallsvariable X wie folgt:

Falls der rote Würfel eine höhere Augenzahl zeigt als der blaue Würfel, dann sei der Wert von X gleich 0.

Andernfalls sei X durch die Augenzahl des blauen Würfels gegeben.

Wenn beispielsweise der blaue Würfel die Augenzahl 6 zeigt, dann hat X stets den Wert 6.

Es gilt $W_X = \{0, 1, 2, \ldots, 6\}$.

1. Geben Sie die Dichtefunktion f_X für X an.

Lösung:

$f_X(0) = \frac{15}{36}$,

$f_X(1), \ldots, f_X(6) = \frac{1}{36}, \frac{2}{36}, \frac{3}{36}, \frac{4}{36}, \frac{5}{36}, \frac{6}{36}$.

(3 P.)

2. Berechnen Sie den Erwartungswert $\mathbb{E}[X|X \neq 0]$ der bedingten Variablen $X|X \neq 0$.

Lösung:

Sei X' die bedingte Variable $X|X \neq 0$. Dann gilt für ihre Dichte

$f_{X'}(0) = 0$ und
$f_{X'}(i) = \frac{f_X(i)}{\Pr[X \neq 0]} = \frac{36}{21} f_X(i)$ für $i \neq 0$. (2 P.)

Es folgt

$\mathbb{E}[X|X \neq 0] = \frac{1}{21}(1 + 2^2 + 3^2 + 4^2 + 5^2 + 6^2) = \frac{91}{21}$. (1 P.)

3. Wir wiederholen das Werfen der Würfel so lange, bis im n-ten Wurf zum ersten Mal der rote Würfel eine höhere Augenzahl zeigt als der blaue Würfel.

   Sei X_i für $i \in \mathbb{N}$ die Zufallsvariable für den i-ten Wurf. Die Verteilung der X_i ist also identisch mit der Verteilung von X.

   Seien N und Y Zufallsvariable, wobei N den Wert n liefere und $Y = \sum_{i=1}^{N} X_i$ gelte.

   Berechnen Sie den Erwartungswert $\mathbb{E}[N]$ von N.

   Berechnen Sie nun den Erwartungswert $\mathbb{E}[Y]$ von Y.

Lösung:

N ist geometrisch verteilt mit Erfolgswahrscheinlichkeit $p = \frac{15}{36}$.

Es folgt $\mathbb{E}[N] = \frac{36}{15}$.

(2 P.)

Berechnung von $\mathbb{E}[Y]$:

$$\begin{aligned}\mathbb{E}[Y] &= \sum_{n=1}^{\infty} \mathbb{E}[Y|N=n] \cdot p(1-p)^{n-1} \\ &= \sum_{n=1}^{\infty} \mathbb{E}[X|X \neq 0] \cdot (n-1) \cdot p(1-p)^{n-1} \\ &= \mathbb{E}[X|X \neq 0] \cdot \mathbb{E}[N-1] \\ &= \frac{91}{21} \cdot \frac{21}{15} = \frac{91}{15}.\end{aligned}$$

(2 P.)

Es gilt auch

$$\mathbb{E}[Y] = \mathbb{E}[N] \cdot \mathbb{E}[X]\,.$$

Die Berechnungsformel von Tutoraufgabe 3 von Blatt 6 ist anwendbar, obwohl die Begründung der Formel hier anders lautet.

Da die Begründung der Formel in der Prüfung nicht verlangt war, gibt es auch bei dieser Anwendung volle Punktzahl.

3. Vorbereitung auf Tutoraufgaben: VA 2 von Blatt 8

In einem Unfallkrankenhaus treffen im Schnitt alle 20 Minuten Patienten zur Behandlung ein.

Die Zeit zwischen zwei Behandlungsfällen sei exponentialverteilt mit Parameter $\frac{1}{20}$.

Wenn 1 Stunde lang kein Patient eingetroffen ist, macht das Personal Ruhepause.

Wir wollen wissen,
welcher Zeitabstand zwischen zwei Ruhepausen zu erwarten ist.

Seien $T_1, T_2, \ldots$ die Zeitspannen zwischen dem Eintreffen zweier Behandlungsfälle und W die Wartezeit bis zur nächsten Ruhepause.

1. Geben Sie $\mathbb{E}[T_1 \mid T_1 \geq 60]$ an.
2. Geben Sie $\mathbb{E}[W \mid T_1 \geq 60]$ an.
3. Berechnen Sie $\mathbb{E}[W]$.

Vorsicht bei Aufgaben, die als Anwendungen verkleidet sind!

Der schwierige erste Schritt besteht bei diesen Aufgaben zunächst darin, eine adäquate mathematische Abstraktion anzugeben.

Vorüberlegungen:

1. Es wird offenbar eine Folge von Zeitpunkten z_i betrachtet, zu denen Patienten im Krankenhaus eintreffen.

2. Die Zeitdifferenzen $T_i = z_i - z_{i-1}$ werden als exponentialverteilt angenommen. Dies bedeutet, dass diese T_i nicht davon abhängen, wie lange noch kein Patient eingetroffen ist (Gedächtnislosigkeit).

Der Parameter λ bedeutet „Anzahl der Patienten pro Zeiteinheit" im Durchschnitt, hier also $\frac{1}{20}$ Patient pro Zeiteinheit als Erwartungswert.

Alle T_i sind unabhängig, d. h. die Menge der T_i ist unabhängig.

3. Falls $T_i > 60$, dann gibt es eine Ruhepause.

4. Die Frage ist, wie lange man warten muss,
bis erstmalig $T_N > 60$ festgestellt wird?

Die Wartezeit ist dann

$$W = 60 + \sum_{j=1}^{N-1} T_j$$

oder

$$W = 60 + W'$$

mit

$$W' = \sum_{j=1}^{N-1} T_j \, .$$

1. Geben Sie $\mathbb{E}[T_1 \mid T_1 \geq 60]$ an.

Lösung:

Sei $\lambda = \frac{1}{20}$.

Dann gilt $\mathbb{E}[T_1] = \frac{1}{\lambda} = 20$.

Da die Exponentialverteilung gedächtnislos ist, gilt

$$\mathbb{E}[T_1 \mid T_1 \geq 60] = 60 + \mathbb{E}[T_1] = 80\,.$$

Wir zeigen die Gleichung direkt durch Berechnung wie folgt:

T_1 ist exponentialverteilt:

$$F_{T_1}(t) = \begin{cases} 1 - e^{-\lambda t} & : t \geq 0 \\ 0 & : \text{sonst}\,. \end{cases}$$

Verteilung der bedingten Variablen $T_1 \mid T_1 \geq 0$:

$$\begin{aligned} F_{(T_1 | T_1 \geq 60)}(t) &= \Pr[(T_1 \mid T_1 \geq 0) \leq t] \\ &= \frac{\Pr[T_1 \leq t, T_1 \geq 60]}{\Pr[T_1 \geq 60]} \end{aligned}$$

Für $t \geq 60$ folgt

$$
\begin{aligned}
F_{(T_1 | T_1 \geq 60)}(t) &= \frac{(1 - e^{-\lambda \cdot t}) - (1 - e^{-\lambda \cdot 60})}{e^{-\lambda \cdot 60}} \\
&= 1 - e^{-\lambda \cdot (t-60)} \,.
\end{aligned}
$$

Damit ist die Variable $T' = (T_1 \mid T_1 \geq 60) - 60$ gleichverteilt wie T_1, d. h. entsprechend exponentialverteilt.

Es folgt

$$
\begin{aligned}
\mathbb{E}[(T_1 \mid T_1 \geq 60)] &= \mathbb{E}[T' + 60] \\
&= \mathbb{E}[T_1] + 60 \\
&= 80 \,.
\end{aligned}
$$

2. Geben Sie $\mathbb{E}[W \mid T_1 \geq 60]$ an.

Lösung:

Offenbar gilt $\mathbb{E}[W \mid T_1 \geq 60] = 60\,.$

3. Berechnen Sie $\mathbb{E}[W]$.

Lösung:

Es gilt $$W = 60 + \sum_{j=1}^{N-1} T_j$$

oder $$W = 60 + W'$$

mit $$W' = \sum_{j=1}^{N-1} T_j\,.$$

N ist eine geometrisch verteilte Zufallsvariable mit Erfolgswahrscheinlichkeit $p = \Pr[T \geq 60]$. Es gilt

$$p = \Pr[T \geq 60] = e^{-3}\,.$$

Berechnung von $\mathbb{E}[W']$.

Wir setzen $T = T_1$.

$$
\begin{aligned}
\mathbb{E}[W'] &= \sum_{n=1}^{\infty} \mathbb{E}[W' | N = n] \cdot p(1-p)^{n-1} \\
&= \sum_{n=1}^{\infty} \mathbb{E}[T | T \leq 60] \cdot (n-1) \cdot p(1-p)^{n-1} \\
&= \mathbb{E}[T | T \leq 60] \cdot \mathbb{E}[N-1].
\end{aligned}
$$

Es gilt

$$\mathbb{E}[N-1] = e^3 - 1\,.$$

$\mathbb{E}[T|T \le 60]$ erhalten wir aus den Gleichungen

$$\begin{aligned}
\mathbb{E}[T] &= \mathbb{E}[T|T \le 60] \cdot \Pr[T \le 60] \;+\; \mathbb{E}[T|T \ge 60] \cdot \Pr[T \ge 60] \\
&= \mathbb{E}[T|T \le 60] \cdot (1 - e^{-3}) + 80 \cdot e^{-3} \\
&= 20\,.
\end{aligned}$$

Es folgt

$$\mathbb{E}[T|T \le 60] = \frac{20 - 80 \cdot e^{-3}}{1 - e^{-3}} = \frac{20 \cdot e^3 - 80}{e^3 - 1}\,.$$

Ergebnis:

$$\begin{aligned}\mathbb{E}[W'] &= 20 \cdot e^3 - 80\,,\\ \mathbb{E}[W] &= 20 \cdot e^3 - 80 + 60\\ &\approx 382 \text{ (Minuten)}\,.\end{aligned}$$