Übungen zu Diskrete Wahrscheinlichkeitstheorie

Dieses Blatt wird am besprochen.

Aufgabe 5.1

Ein Hund bewacht das Grundstück seines Besitzers und bewegt sich dazu am Gartenzaun auf und ab. Nach jeweils einer Minute bewegt er sich entweder einen Meter nach links oder nach rechts oder behält seine Position bei.

Wenn wir die anfängliche Position des Hundes als $P_0 := 0$ bezeichnen, so befindet er sich nach n Minuten an einer Position im Intervall $[-n, n]$. Die Positionsveränderungen modellieren wir als Zufallsvariablen $X_1, \ldots, X_n$ (mit $W_{X_i} = \{-1, 0, 1\}$ für $i = 1, \ldots, n$), wobei wir annehmen, dass die X_i identisch verteilt und unabhängig voneinander sind. Die Position des Hundes nach n Minuten ist dann $P_n := \sum_{i=1}^{n} X_i$.

(a) Nehmen Sie an, dass

$$\Pr[X_i = 1] =: p \geq 0,\ \Pr[X_i = -1] =: q \geq 0 \text{ und } \Pr[X_i = 0] = 1 - p - q \geq 0$$

gilt für alle $i = 1 \ldots n$.

Berechnen Sie dann $\mathbb{E}[P_n]$ und $\mathrm{Var}[P_n]$.

(b) Es sei nun $p + q = 1$, d.h., der Hund bewegt sich in jedem Schritt nach links oder rechts und bleibt nicht stehen. Ferner seien $\mathbb{E}[P_n] = 2.8$ und $\mathrm{Var}[P_n] = 5.88$. Bestimmen Sie n und p.

(c) Berechnen Sie für den Fall $p + q = 1$ die Dichte von P_n.

Hinweis: Überlegen Sie sich, dass $Y_i := \frac{X_i+1}{2}$ $\mathrm{Bin}(1, p)$-verteilt ist.

(d) Ein einfaches Modell für die Entwicklung einer Aktie sieht wie folgt aus:

Zwischen zwei Zeitpunkten t_i und t_{i+1} wächst der Wert der Aktie entweder mit W'keit $0 < p < 1$ um einen Faktor $u > 1$ oder fällt um Faktor $0 < d < 1$ mit W'keit $1 - p$.

Sind $R_1, \ldots, R_n$ unabhängige, identisch verteilte Zufallsvariablen mit

$$\Pr[R_1 = u] = p \text{ und } \Pr[R_1 = d] = 1 - p,$$

so beschreibt also die Zufallsvariable

$$S_n = S_0 \prod_{i=1}^{n} R_i$$

den Wert der Aktie zum Zeitpunkt t_n - wobei $S_0 \in (0, \infty)$ der anfängliche Wert der Aktie sei.

S_n hat damit die Wertemenge

$$W_{S_n} = \{S_0 \cdot u^i \cdot d^{n-i} \mid 0 \leq i \leq n\}.$$

Bestimmen Sie die Dichte und den Erwartungswert von S_n.

Hinweis:

$$(a + b)^n = \sum_{i=0}^{n} \binom{n}{i} a^i b^{n-i}.$$

Lösungsvorschlag

(a) Es ergibt sich sofort aus der Linearität des Erwartungswerts und der identischen Verteilung der ZVariablen, dass

$$\mathbb{E}[P_n] = \mathbb{E}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n \cdot \mathbb{E}[X_1] = n(p-q).$$

Wegen der Unabhängigkeit folgt für die Varianz

$$\mathrm{Var}[P_n] = \mathrm{Var}[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} \mathrm{Var}[X_i] = n \cdot \mathrm{Var}[X_1] = n(p+q-(p-q)^2).$$

(b) Nach (a) gilt

$$\mathbb{E}[P_n] = n \cdot (p-q) = n \cdot (p-(1-p)) = n \cdot (2p-1)$$

und

$$\mathrm{Var}[P_n] = n \cdot (p+q-(p-q)^2) = n \cdot (p+1-p-(2p-1)^2) = n \cdot (1-4p^2+4p-1) = n \cdot 4p(1-p).$$

Damit folgt

$$\begin{aligned} & \frac{\mathrm{Var}[X]}{\mathbb{E}[X]} = \frac{4p(1-p)}{2p-1} = \frac{5.88}{2.8} = 2.1 \\ \Leftrightarrow\quad & 4p - 4p^2 = 4.2p - 2.1 \\ \Leftrightarrow\quad & 4p^2 + 0.2p - 2.1 = 0 \\ \Leftrightarrow\quad & p = \frac{-0.2 \pm \sqrt{00.4+33.6}}{8} = \frac{-0.2 \pm 5.8}{8} \in \{0.7, -0.75\}. \end{aligned}$$

Also gilt $p = 0.7$ und damit dann

$$n = \frac{\mathbb{E}[P_n]}{2p-1} = \frac{2.8}{0.4} = 7.$$

(c) Es gilt

$$\Pr[Y_i = 1] = \Pr[\frac{X_i+1}{2} = 1] = \Pr[X_i = 1] = p \text{ und } \Pr[Y_i = 0] = \Pr[\frac{X_i+1}{2} = 0] = \Pr[X_i = -1] = 1-p.$$

Also ist Y_i $\mathrm{Bin}(1,p)$-verteilt.

Natürlich sind $Y_1, \ldots, Y_n$ wieder unabhängig.

Es folgt dann, dass

$$P_k = \sum_{i=1}^{k} (2Y_i - 1) = 2\sum_{i=1}^{k} Y_i - k,$$

wobei $\sum_{i=1}^{k} Y_i$ nach der Vorlesung $\mathrm{Bin}(k,p)$-verteilt ist.

Für die Dichte erhält man dann sofort

$$\Pr[P_k = j] = \Pr[-k + 2\sum_{i=1}^{k} Y_i = j] = \Pr[\sum_{i=1}^{k} Y_i = \frac{j+k}{2}] = \binom{k}{\frac{j+k}{2}} p^{\frac{j+k}{2}} (1-p)^{k-\frac{k+j}{2}}$$

für $\frac{j+k}{2} \in \{0, \ldots, k\}$ und ansonsten $\Pr[P_k = j] = 0$.

Man kann das Resultat natürlich auch direkt herleiten:

Für die Dichte von P_k überlegt man sich zunächst, dass wegen $p+q=1$ für ungerades k P_k nur ungerade Werte mit W'keit > 0 annehmen kann, entsprechend kann P_k für gerades k nur gerade Werte mit positiver W'keit annehmen - ggf. kann dies auch induktiv bewiesen werden.

Weiter gilt sicherlich $\Pr[P_k = j] = 0$ für $j > k$, da $P_k \in [-k, k]$.

Wir betrachten zunächst den Fall, dass $k = 2i$ gerade ist, dann folgt für $0 \leq 2j \leq 2i$, dass wir $2j + (i-j)$ mal nach rechts gehen müssen und $i - j$ mal nach links, wobei die Reihenfolge der Bewegungen nach Links und Rechts unerheblich ist, d.h.:

$$\Pr[P_{2i} = 2j \geq 0] = \binom{2i}{i-j} (1-p)^{i-j} p^{3i-j}.$$

Entsprechend folgt für $-2i \leq 2j < 0$

$$\Pr[P_{2i} = 2j < 0] = \binom{2i}{i-j}(1-p)^{3i-j}p^{i-j}.$$

Für den ungeraden Fall folgt entsprechend

$$\Pr[P_{2i+1} = 2j+1 \geq 0] = \binom{2i+1}{i-j}(1-p)^{i-j}p^{3i-j+1} \text{ und } \Pr[P_{2i+1} = 2j-1 < 0] = \binom{2i+1}{i-j}(1-p)^{3i-j+1}p^{i-j}$$

(d) Da die relativen Änderungen R_i unabhängig sind, folgt

$$\Pr[S_k = S_0 \cdot u^i \cdot d^{k-i}] = \binom{k}{i}p^i(1-p)^{k-i}.$$

$$\mathbb{E}[S_k] = S_0 \cdot \mathbb{E}[R_1 \cdots R_k] = S_0 \cdot \mathbb{E}[R_1] \cdots \mathbb{E}[R_k] = S_0 \cdot \mathbb{E}[R_1]^k.$$

Aufgabe 5.2

In einer Klausur kommt ein Multiple-Choice-Test mit $n = 8$ Fragen vor, für die es jeweils eine richtige und eine falsche Antwort gibt.

Wir nehmen an, dass es pro korrekter Antwort $K > 0$ Punkte gibt, pro falscher Antwort jedoch $F \leq 0$ Punkte. Die Endpunktzahl ist dann einfach die Summe der Punkte, welche in den einzelnen Aussagen erzielt wurde, mindestens aber 0 Punkte.

Unter der Annahme, dass jede Aussage mit Wahrscheinlichkeit $p = 0.5$ korrekt beantwortet und die Antworten dabei unabhängig voneinander gewählt werden, bestimmen Sie den Erwartungswert und die Varianz der Endpunktzahl für die folgenden Fälle:

(a) $K = 1$, $F = 0$.

(b) $K = 1$, $F = -1$.

Lösungsvorschlag Wir können jede Aussage mittels einer $\text{Bin}(1, p)$-verteilten Zufallsvariable X_i beschreiben, so dass X_i genau dann den Wert 1 annimmt, falls die i.te Aussage korrekt beantwortet wurde, und die Zufallsvariablen $X_1, \ldots, X_n$ unabhängig sind.

Die Anzahl der korrekten Anworten ist dann durch

$$S := \sum_{i=1}^{n} X_i$$

gegeben, die Summe der Punkte durch

$$P := K \cdot S + (n - S) \cdot F = (K - F) \cdot S + n \cdot F$$

und die Endpunktzahl durch

$$E := \max\{0, P\}.$$

Nach der Vorlesung ist S $\text{Bin}(n, p)$ verteilt.

(a) Im Fall $K = 1$ und $F = 0$ gilt

$$E = \max\{0, P\} = \max\{0, (1-0) \cdot S + n \cdot 0\} = \max\{0, S\} = S.$$

Damit folgt

$$\mathbb{E}[E] = \mathbb{E}[S] = n \cdot p = 4.$$

Entsprechend gilt auf Grund der Unabhängigkeit

$$\text{Var}[E] = \text{Var}[S] = n \cdot \text{Var}[X_1] = n \cdot p \cdot (1-p) = 2.$$

(b) Für $K=1$ und $F=-1$ gilt für die Dichte von P für $k \in \{1,\ldots,8\}$:

$$\Pr[P=k] = \Pr[2\cdot S - n = k] = \Pr[S = \frac{n+k}{2}] = \binom{n}{\frac{n+k}{2}} p^{\frac{n+k}{2}} (1-p)^{n-\frac{n+k}{2}} = \binom{8}{\frac{8+k}{2}} 2^{-8},$$

wobei der Binomialkoeffizient hier gleich 0 sein soll, falls $\frac{8+k}{2}$ keine ganze Zahl ist.

Damit folgt für den Erwartungswert:

$$\mathbb{E}[E] = \sum_{k\in W_E} k\cdot\Pr[\max 0, P=k] = \sum_{k=1}^{8} k\cdot\Pr[P=k] = \sum_{l=1}^{4} 2l\cdot\binom{8}{4+l} 2^{-8} = 2^{-7}(1\cdot56+2\cdot28+3\cdot8+4\cdot1) = \frac{140}{128} \approx 1.094.$$

Entsprechend gilt

$$\mathbb{E}[E^2] = \sum_{k=1}^{8} k^2\cdot\Pr[P=k] = \sum_{l=1}^{4}(2l)^2\cdot\binom{8}{4+l} 2^{-8} = 2^{-6}(1\cdot 56 + 4\cdot 28 + 9\cdot 8 + 16\cdot 1) = 4.$$

Schließlich:

$$\mathrm{Var}[E] = \mathbb{E}[E^2] - \mathbb{E}[E]^2 = 4 - \frac{140^2}{2^{14}} \approx 2.804.$$

Aufgabe 5.3

Sei X eine diskrete Zufallsvariable mit Wertemenge $W_X = \{1,2,\ldots\}$. Wir schreiben kurz p_i für $\Pr[X=i]$ und nehmen an, dass $0 < p_1 < 1$ gilt.

Weiterhin sei X gedächtnislos, d.h., es gilt

$$\Pr[X > k+l \mid X > k] = \Pr[X > l]$$

für all k, l.

Zeigen Sie, dass $p_i = p_1 \cdot (1-p_1)^{i-1}$ gilt!

Lösungsvorschlag

- Zunächst schreibt man die Gedächtnislosigkeit

$$\Pr[X > k+l \mid X > k] = \Pr[X > l]$$

um in

$$\frac{1-\Pr[X\le k+l]}{1-\Pr[X\le k]} = 1-\Pr[X\le l] \text{ bzw. } \frac{1-p_1-p_2-\ldots-p_{k+l}}{1-p_1-p_2-\ldots-p_k} = 1-p_1-p_2-\ldots-p_l,$$

da $\Pr[X > k+l, X > k] = \Pr[X > k+l]$.

Für $l=1$ folgt dann speziell:

$$1-p_1-\ldots-p_k-p_{k+1} = (1-p_1)\cdot(1-p_1-p_2-\ldots-p_k) \text{ bzw } p_{k+1} = p_1\cdot(1-\sum_{i=1}^{k} p_i).$$

Für $k=1$ folgt also direkt $p_2 = p_1\cdot(1-p_1)$. Induktiv sei die Aussage dann bis $k \ge 0$ bewiesen. Dann gilt nach Induktion zunächst

$$\sum_{i=1}^{k} p_i = p_1 \cdot \sum_{i=1}^{k}(1-p_1)^{i-1} = p_1\cdot\frac{1-(1-p_1)^k}{p_1} = 1-(1-p_1)^k.$$

Und damit

$$p_{k+1} = p_1\cdot(1-\sum_{i=1}^{k} p_i) = p_1\cdot(1-(1-(1-p_1)^k)) = p_1\cdot(1-p_1)^k.$$

Alternativ unter Verwendung des Multiplilationssatzes (Satz 17):

Man setze $A_n = \{\omega \in \Omega \mid X(\omega) > n\}$ für $n \geq 0$.

- Mittels des Multiplikationssatzes folgt:

  $$\Pr[A_n] = \Pr[A_0] \cdot \Pr[A_1 \mid A_0] \cdot \ldots \cdot \Pr[A_n \mid A_0 \cap A_1 \cap \ldots \cap A_{n-1}].$$

  Beachtet man nun, dass $A_0 \cap \ldots \cap A_k = A_k$ gilt, so folgt:

  $$\Pr[A_n] = \Pr[A_0] \cdot \prod_{k=1}^{n} \Pr[A_k \mid A_{k-1}].$$

  Auf Grund der Gedächtnislosigkeit gilt weiter $\Pr[A_k \mid A_{k-1}] = \Pr[A_1] = 1 - p_1$, daher sind die betrachteten bedingten W'keiten definiert, und der Multiplikationssatz darf tatsächlich verwendet werden.

  Mit $\Pr[A_0] = 1$ folgt schließlich

  $$\Pr[A_n] = (1 - p_1)^n.$$

- oder unter indirekter Anwendung des Multiplikationssatzes:

  $$\begin{aligned} \Pr[A_n] &= \Pr[A_n \mid A_{n-1}] \cdot \Pr[A_{n-1}] \\ &\overset{\text{Gedächtnislosigkeit}}{=} \Pr[A_1] \cdot \Pr[A_{n-1}] \\ &= (1 - p_1) \cdot \Pr[A_{n-1}] \\ &\overset{\text{Induktion über } n}{=} (1 - p_1)^n. \end{aligned}$$

Schließlich berechnet man

$$\Pr[X = n] = \Pr[A_{n-1} \setminus A_n] = \Pr[A_{n-1}] - \Pr[A_n] = (1 - p_1)^{n-1} \cdot (1 - (1 - p_1)) = p_1 \cdot (1 - p_1)^{n-1}.$$