Bemerkung 2 (Gegenbeispiel für die allgemeine Gültigkeit der Aussage)

Man kann ein Modell angeben, bei dem $X_1, \ldots, X_n$ stochastisch unabhängig sind, jedes X_i Bernoulli(p)-verteilt ist und

$$\begin{aligned} P(A_i = 4|X_i = 1) &= 0,5 \\ P(A_i = 12|X_i = 1) &= 0,3 \\ P(A_i = 0|X_i = 1) &= 0,2 \\ P(A_i = 0|X_i = 0) &= 1 \end{aligned}$$

für $i = 1, \ldots, n$ gilt, <u>aber nicht</u>

$$E(\sum_{i=1}^{n} A_i | X = x) = 5,6 \cdot x$$

wobei zur Abkürzung $X := \sum_{i=1}^{n} X_i$ sei.

Begründung:

Zur Abkürzung sei $S_i := X - X_i = \sum\limits_{j=1, j\neq i}^{n} X_j$ für $i = 1, \ldots, n$.

Die Verteilung von A_i sei folgendermaßen festgelegt:

$$\begin{aligned} P(A_i = 0|X_i = 0) &= 1 \\ P(A_i = 12|X_i = 1 \text{ und } S_i = s) &= 0,3 \text{ für alle } s \in \{0, \ldots, n-1\} \\ P(A_i = 4|X_i = 1 \text{ und } S_i = s) &= 0,5 - \delta \text{ für } s = 0, 1, \ldots, c \\ P(A_i = 4|X_i = 1 \text{ und } S_i = s) &= 0,5 + \delta \cdot \frac{P(S_i \leq c)}{1 - P(S_i \leq c)} \text{ für } s = c+1, \ldots, n-1 \\ P(A_i = 0|X_i = 1 \text{ und } S_i = s) &= 1 - 0,3 - (0,5 - \delta) \\ &= 0,2 + \delta \text{ für } s = 0, 1, \ldots, c \\ P(A_i = 4|X_i = 1 \text{ und } S_i = s) &= 1 - 0,3 - (0,5 + \delta \cdot \frac{P(S_i \leq c)}{1 - P(S_i \leq c)}) \\ &= 0,2 - \delta \cdot \frac{P(S_i \leq c)}{1 - P(S_i \leq c)} \text{ für } s = c+1, \ldots, n-1 \end{aligned}$$

wobei δ und c so gewählt werden, dass $P(S_i \leq c) \leq \frac{1}{2}$ und $0 < \delta < 0,2$ gilt (dadurch ist sichergestellt, dass die gewählten Wahrscheinlichkeiten tatsächlich zwischen 0 und 1 liegen).

A_i kann nur dann Werte $b \neq 0$ anehmen, wenn $X_i = 1$, deshalb gilt:

1. Für alle $b \neq 0$:

$$P(A_i = b|X_i = 1) = \frac{P(A_i = b \text{ und } X_i = 1)}{P(X_i = 1)} = \frac{P(A_i = b)}{P(X_i = 1)} = \frac{1}{p} \cdot P(A_i = b)$$

2. Für alle Teilmengen B von $\{0, 1, \ldots, n-1\}$ und alle $b \neq 0$:

$$\begin{aligned} P(A_i = b|X_i = 1 \text{ und } S_i \in B) &= \frac{P(A_i = b \text{ und } X_i = 1 \text{ und } S_i \in B)}{P(X_i = 1 \text{ und } S_i \in B)} \\ &\overset{\text{weil } b \neq 0}{=} \frac{P(A_i = b \text{ und } S_i \in B)}{P(X_i = 1 \text{ und } S_i \in B)} \\ &= \frac{P(A_i = b \text{ und } S_i \in B)}{P(X_i = 1)P(S_i \in B)} \end{aligned}$$

Der letzte Schritt gilt, weil $X_1, \ldots, X_n$ stochastisch unabhängig und somit auch X_i und S_i stochastisch unabhängig sind. Damit ergibt sich:

$$P(A_i = b | X_i = 1 \text{ und } S_i \in B) = \frac{P(A_i = b \text{ und } S_i \in B)}{P(X_i = 1) P(S_i \in B)} = \frac{1}{p} \cdot P(A_i = b | S_i \in B)$$

3. Es gilt für $s = 0, \ldots, c$:

$$P(A_i = 4 | S_i = s) \overset{\text{wegen 2.}}{=} p \cdot P(A_i = 4 | X_i = 1 \text{ und } S_i = s) = p \cdot (0,5 - \delta)$$

4. Es gilt

$$\begin{aligned}
P(A_i = 4 | S_i \leq c) &= \frac{P(A_i = 4 \text{ und } S_i \leq c)}{P(S_i \leq c)} \\
&\overset{\text{Satz v.d. totalen WS}}{=} \frac{1}{P(S_i \leq c)} \sum_{s=0}^{c} P(A_i = 4 | S_i = s) P(S_i = s) \\
&\overset{\text{wegen 3.}}{=} \frac{1}{P(S_i \leq c)} \sum_{s=0}^{c} p(0,5 - \delta) \cdot P(S_i = s) \\
&= p(0,5 - \delta)
\end{aligned}$$

Analog zu 3. und 4. zeigt man:

5. $P(A_i = 4 | S_i = s) = p \cdot \left(0,5 + \delta \cdot \frac{P(S_i \leq c)}{1 - P(S_i \leq c)}\right)$ für $s = c+1, \ldots, n-1$

6. $P(A_i = 4 | S_i > c) = p \cdot \left(0,5 + \delta \cdot \frac{P(S_i \leq c)}{1 - P(S_i \leq c)}\right)$

7. Nun gilt, mit dem Satz von der totalen Wahrscheinlichkeit, und mit 4. und 6.

$$\begin{aligned}
P(A_i = 4 | X_i = 1) &= \frac{1}{p} \cdot P(A_i = 4) \\
&= \frac{1}{p} \cdot (P(A_i = 4 | S_i \leq c) \cdot P(S_i \leq c) + P(A_i = 4 | S_i > c) \cdot P(S_i > c)) \\
&= \frac{1}{p} \cdot \left(p \cdot (0,5 - \delta) \cdot P(S_i \leq c) + p \cdot (0,5 + \delta \cdot \frac{P(S_i \leq c)}{1 - P(S_i \leq c)}) \cdot P(S_i > c)\right) \\
&= 0,5 \cdot 1 - \delta \cdot P(S_i \leq c) + \delta \cdot P(S_i \leq c) \\
&= 0,5
\end{aligned}$$

was ja auf jeden Fall gelten soll. Man rechnet ganz ähnlich nach, dass gilt:

8. $P(A_i = 0 | X_i = 1) = 0,2$ und $P(A_i = 12 | X_i = 1) = 0,3$

   was natürlich auch gelten soll.

   Nun ist nur noch zu zeigen, dass <u>nicht</u> für alle $x \in \{0, \ldots, n\}$ gilt

$$E(\sum A_i | X = x) = 5,6 \cdot x$$

9. Es gilt für alle $b \neq 0$

$$\begin{aligned}
P(A_i = b | X = x) &= \frac{P(A_i = b \text{ und } X = x)}{P(X = x)} \\
&\overset{b \neq 0}{=} \frac{P(A_i = b \text{ und } X_i = 1 \text{ und } X = x)}{P(X = x)} \\
&\overset{\text{n. Definition von } S_i}{=} \frac{P(A_i = b \text{ und } X_i = 1 \text{ und } S_i = x - 1)}{P(X = x)} \\
&\overset{b \neq 0}{=} \frac{P(A_i = b \text{ und } S_i = x - 1)}{P(X = x)}
\end{aligned}$$

Dieser Bruch wird nun mit $P(S_i = x-1)$ erweitert (Trick), dann ergibt sich

$$\begin{aligned} P(A_i = b|X = x) &= \frac{P(A_i = b \text{ und } S_i = x-1)}{P(S_i = x-1)} \cdot \frac{P(S_i = x-1)}{P(X = x)} \\ &= P(A_i = b|S_i = x-1) \cdot \frac{\binom{n-1}{x-1} p^{x-1}(1-p)^{(n-1)-(x-1)}}{\binom{n}{x} p^x (1-p)^{n-x}} \end{aligned}$$

Dabei wurde ausgenutzt, dass S_i wieder binomialverteilt ist, denn die X_j mit $j \neq i$ sind stochastisch unabhängig. Nach Kürzen ergibt sich wegen 1.

$$\begin{aligned} P(A_i = b|X = x) &= P(A_i = b|S_i = x-1) \cdot \frac{x}{np} \\ &= \frac{x}{n} \cdot P(A_i = b|X_i = 1 \text{ und } S_i = x-1) \end{aligned}$$

Also gilt mit Berücksichtigung der Fallunterscheidung bei der Festlegung der Verteilung von A_i:

$$\begin{aligned} P(A_i = 4|X = x) &= \frac{x}{n} \cdot P(A_i = 4 \text{ und } X_i = 1 \text{ und } S_i = x-1) \\ &= \begin{cases} \frac{x}{n} \cdot (0,5 - \delta) & \text{für } x = 1, \ldots, c+1 \\ \frac{x}{n} \cdot (0,5 + \delta \cdot \dfrac{P(S_i \leq c)}{1 - P(S_i \leq c)}) & \text{für } x = c+2, \ldots, n \end{cases} \end{aligned}$$

Ferner

$$P(A_i = 12|X = x) = \frac{x}{n} \cdot P(A_i = 12|X_i = 1 \text{ und } S_i = x-1) = \frac{x}{n} \cdot 0,3$$

Damit gilt für $i = 1, \ldots, n$:

$$\begin{aligned} E(A_i|X = x) &= 4 \cdot P(A_i = 4|X = x) + 12 \cdot P(A_i = 12|X = x) \\ &= 12 \cdot \frac{x}{n} \cdot 0,3 + 4 \cdot \frac{x}{n} \cdot \begin{cases} 0,5 - \delta & \text{für } x = 1, \ldots, c+1 \\ 0,5 + \delta \cdot \dfrac{P(S_i \leq c)}{1 - P(S_i \leq c)} & \text{für } x = c+2, \ldots, n \end{cases} \end{aligned}$$

Das ist für alle i gleich, es gilt

$$\begin{aligned} E(\sum_{i=1}^{n} A_i|X = x) &= \sum_{i=1}^{n} E(A_i|X = x) \\ &= n \cdot E(A_1|X = x) \\ &= x \cdot \left(\underbrace{12 \cdot 0,3 + 4 \cdot 0,5}_{=5,6} + 4 \cdot \begin{cases} -\delta & \text{für } x = 1, \ldots, c+1 \\ +\delta \cdot \dfrac{P(S_i \leq c)}{1 - P(S_i \leq c)} & \text{für } x = c+2, \ldots, n \end{cases} \right) \end{aligned}$$

Also ist hier der Faktor mit dem x multipliziert wird, nicht gleich $5,6$ für alle x, sondern für $x \leq c+1$ kleiner und für $x \geq c+2$ größer.