Beweis:

Mit Hilfe des Satzes von der totalen Wahrscheinlichkeit folgt, dass

$$
\begin{aligned}
f_Z(z) = \Pr[Z = z] &= \sum_{x \in W_X} \Pr[X + Y = z \mid X = x] \cdot \Pr[X = x] \\
&= \sum_{x \in W_X} \Pr[Y = z - x] \cdot \Pr[X = x] \\
&= \sum_{x \in W_X} f_X(x) \cdot f_Y(z - x)\,.
\end{aligned}
$$

$\square$

Den Ausdruck $\sum_{x \in W_X} f_X(x) \cdot f_Y(z - x)$ aus Satz 49 nennt man in Analogie zu den entsprechenden Begriffen bei Potenzreihen auch Faltung oder Konvolution der Dichten f_X und f_Y.

Beispiel (Forts.)

Berechne die Dichte von $Z = X + Y$:

$$\Pr[Z = z] = \sum_{x \in W_X} \Pr[X = x] \cdot \Pr[Y = z - x]$$

$$= \sum_{x=1}^{6} \frac{1}{6} \cdot \Pr[Y = z - x] = \sum_{x=\max\{1,z-6\}}^{\min\{6,z-1\}} \frac{1}{36}.$$

Für $2 \leq z \leq 7$ erhalten wir

$$\Pr[Z = z] = \sum_{i=1}^{z-1} \frac{1}{36} = \frac{z-1}{36}\,.$$

Und für $7 < z \leq 12$:

$$\Pr[Z = z] = \frac{13-z}{36}\,.$$

4.3.3 Momente zusammengesetzter Zufallsvariablen

Satz 50 (Linearität des Erwartungswerts)

Für Zufallsvariablen $X_1, \ldots, X_n$ *und* $X := a_1X_1 + \cdots + a_nX_n$ *mit* $a_1, \ldots, a_n \in \mathbb{R}$ *gilt*

$$\mathbb{E}[X] = a_1\mathbb{E}[X_1] + \cdots + a_n\mathbb{E}[X_n]\,.$$

Beweis:

$$\begin{aligned}
\mathbb{E}[X] &= \sum_{\omega\in\Omega}(a_1 \cdot X_1(\omega) + \ldots + a_n \cdot X_n(\omega)) \cdot \Pr[\omega] \\
&= a_1 \cdot \left(\sum_{\omega\in\Omega} X_1(\omega) \cdot \Pr[\omega]\right) + \cdots + a_n \cdot \left(\sum_{\omega\in\Omega} X_n(\omega) \cdot \Pr[\omega]\right) \\
&= a_1 \cdot \mathbb{E}[X_1] + \ldots + a_n \cdot \mathbb{E}[X_n]\,.
\end{aligned}$$

$\square$

Beispiel 51

n betrunkene Seeleute torkeln nach dem Landgang in ihre Kojen. Sie haben völlig die Orientierung verloren, weshalb wir annehmen, dass jede Zuordnung der Seeleute zu den n Betten gleich wahrscheinlich ist (genau ein Seemann pro Bett). Wie viele Seeleute liegen im Mittel im richtigen Bett?

Die Anzahl der Seeleute im richtigen Bett zählen wir mit der Zufallsvariablen X, die als Summe der Zufallsvariablen $X_1, \ldots, X_n$ dargestellt wird, wobei

$$X_i := \begin{cases} 1 & \text{falls Seemann } i \text{ in seinem Bett liegt,} \\ 0 & \text{sonst.} \end{cases}$$

Offenbar gilt $X := X_1 + \cdots + X_n$.

Beispiel 51

Für die Variablen X_i erhalten wir $\Pr[X_i = 1] = \frac{1}{n}$, da jedes Bett von Seemann i mit gleicher Wahrscheinlichkeit aufgesucht wird.

Daraus folgt

$$\mathbb{E}[X_i] = 0 \cdot \Pr[X_i = 0] \; + \; 1 \cdot \Pr[X_i = 1] = \frac{1}{n},$$

und somit

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = \sum_{i=1}^{n} \frac{1}{n} = 1\,.$$

Im Mittel hat also nur ein Seemann sein eigenes Bett aufgesucht.

Satz 52 (Multiplikativität des Erwartungswerts)

Für unabhängige Zufallsvariablen $X_1, \ldots, X_n$ gilt

$$\mathbb{E}[X_1 \cdot \cdots \cdot X_n] = \mathbb{E}[X_1] \cdot \cdots \cdot \mathbb{E}[X_n]\,.$$

Beweis:

Wir beweisen den Fall $n = 2$. Der allgemeine Fall ist analog.

$$\begin{aligned}
\mathbb{E}[X \cdot Y] &= \sum_{x \in W_X} \sum_{y \in W_Y} xy \cdot \Pr[X = x, Y = y] \\
&\overset{\text{Unabh.}}{=} \sum_{x \in W_X} \sum_{y \in W_Y} xy \cdot \Pr[X = x] \cdot \Pr[Y = y] \\
&= \sum_{x \in W_X} x \cdot \Pr[X = x] \sum_{y \in W_Y} y \cdot \Pr[Y = y] \\
&= \mathbb{E}[X] \cdot \mathbb{E}[Y]\,.
\end{aligned}$$

$\square$

Dass für die Gültigkeit von Satz 52 die Unabhängigkeit der Zufallsvariablen wirklich notwendig ist, sieht man beispielsweise am Fall $Y = -X$ für eine Zufallsvariable mit einer von Null verschiedenen Varianz. Dann gilt

$$\mathbb{E}[X \cdot Y] = -\mathbb{E}[X^2] \neq -(\mathbb{E}[X])^2 = \mathbb{E}[X] \cdot \mathbb{E}[Y]\,.$$

Definition 53

Zu einem Ereignis A heißt die Zufallsvariable

$$I_A := \begin{cases} 1 & \text{falls } A \text{ eintritt,} \\ 0 & \text{sonst} \end{cases}$$

Indikatorvariable des Ereignisses A.

Beobachtung:

Für die Indikatorvariable I_A gilt nach Definition

$$\mathbb{E}[I_A] = 1 \cdot \Pr[A] + 0 \cdot \Pr[\bar{A}] = \Pr[A]\,.$$

Ebenso gilt

$$\mathbb{E}[I_{A_1} \cdot \ldots \cdot I_{A_n}] = \Pr[A_1 \cap \ldots \cap A_n],$$

da das Produkt von Indikatorvariablen genau dann gleich 1 ist, wenn alle entsprechenden Ereignisse eintreten.

Beispiel (Forts.)

Wir betrachten wieder das Beispiel der total betrunkenen Matrosen.
Sei A_i das Ereignis, dass der i-te Seemann im richtigen Bett liegt. Mit der Notation der Indikatorvariablen sei $X_i = I_{A_i}$. Dann gilt für beliebige $i, j \in \{1, \ldots, n\}$, $i \neq j$:

$$\mathbb{E}[X_i X_j] = \mathbb{E}[I_{A_i} I_{A_j}] = \Pr[A_i \cap A_j] = \frac{1}{n(n-1)},$$

sowie

$$\mathbb{E}[X_i^2] = 0^2 \cdot \Pr[\bar{A}_i] + 1^2 \cdot \Pr[A_i] = \Pr[A_i] = 1/n.$$

Beispiel (Forts.)

Daraus folgt wegen der Linearität des Erwartungswerts für $X = X_1 + \cdots + X_n$*:*

$$\mathbb{E}[X^2] = \mathbb{E}\left[\sum_{i=1}^{n} X_i^2 + \sum_{i=1}^{n}\sum_{j\neq i} X_iX_j\right]$$

$$= n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n(n-1)} = 2\,.$$

Für die Varianz erhalten wir somit den Wert

$$\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 2 - 1 = 1.$$

Einfacher Beweis für Satz 9 mit Hilfe von Indikatorvariablen:
Zur Erinnerung:

Satz 9 (Siebformel, Prinzip der Inklusion/Exklusion)

Für Ereignisse $A_1, \ldots, A_n$ $(n \geq 2)$ gilt:

$$\begin{aligned}\Pr\left[\bigcup_{i=1}^{n} A_i\right] = \sum_{i=1}^{n} \Pr[A_i] - \sum_{1\leq i_1<i_2\leq n} \Pr[A_{i_1} \cap A_{i_2}] + - \ldots \\ + (-1)^{l-1} \sum_{1\leq i_1<\ldots<i_l\leq n} \Pr[A_{i_1} \cap \ldots \cap A_{i_l}] + - \ldots \\ + (-1)^{n-1} \cdot \Pr[A_1 \cap \ldots \cap A_n]\,.\end{aligned}$$

Beweis:

Zur Erinnerung: Zu Ereignissen $A_1, \ldots, A_n$ wollen wir die Wahrscheinlichkeit $\Pr[B]$ des Ereignisses $B := A_1 \cup \ldots \cup A_n$ ermitteln.

Wir betrachten die Indikatorvariablen $I_i := I_{A_i}$ der Ereignisse $A_1, \ldots, A_n$ und die Indikatorvariable $I_{\bar{B}}$ des Ereignisses $\bar{B}$.

Das Produkt $\prod_{i=1}^{n}(1 - I_i)$ ist genau dann gleich 1, wenn $I_1 = \ldots = I_n = 0$, d.h. wenn B nicht eintritt. Somit gilt $I_{\bar{B}} = \prod_{i=1}^{n}(1 - I_i)$ und wir erhalten:

$$I_{\bar{B}} = 1 - \sum_{1 \le i \le n} I_i + \sum_{1 \le i_1 < i_2 \le n} I_{i_1} I_{i_2} - + \ldots + (-1)^n I_1 \cdot \ldots \cdot I_n\,,$$

also

$$\begin{aligned} I_B &= 1 - I_{\bar{B}} \\ &= \sum_{1 \le i \le n} I_i - \sum_{1 \le i_1 < i_2 \le n} I_{i_1} I_{i_2} + - \ldots + (-1)^{n-1} I_1 \cdot \ldots \cdot I_n\,. \end{aligned}$$

Beweis:

Wegen der Eigenschaften von Indikatorvariablen gilt

$$\Pr[B] = 1 - \Pr[\bar{B}] = 1 - \mathbb{E}[I_{\bar{B}}].$$

Mit Hilfe von Satz 50 „verteilen" wir den Erwartungswert auf die einzelnen Produkte von Indikatorvariablen. Wenn wir nun $\mathbb{E}[I_i]$ durch $\Pr[A_i]$ und allgemein $\mathbb{E}[I_{i_1} \cdot \ldots \cdot I_{i_k}]$ durch $\Pr[A_{i_1} \cap \ldots \cap A_{i_k}]$ ersetzen, haben wir Satz 9 (dieses Mal vollständig) bewiesen. □

Satz 54

Für unabhängige Zufallsvariablen $X_1, \ldots, X_n$ *und* $X := X_1 + \ldots + X_n$ *gilt*

$$\mathrm{Var}[X] = \mathrm{Var}[X_1] + \ldots + \mathrm{Var}[X_n]\,.$$

Beweis:

Wir betrachten nur den Fall $n = 2$ mit den Zufallsvariablen X und Y.

$$\mathbb{E}[(X+Y)^2] = \mathbb{E}[X^2 + 2XY + Y^2] = \mathbb{E}[X^2] + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y^2]$$
$$\mathbb{E}[X+Y]^2 = (\mathbb{E}[X] + \mathbb{E}[Y])^2 = \mathbb{E}[X]^2 + 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[Y]^2$$

Wir ziehen die zweite Gleichung von der ersten ab und erhalten

$$\mathbb{E}[(X+Y)^2] - \mathbb{E}[X+Y]^2 = \mathbb{E}[X^2] - \mathbb{E}[X]^2 + \mathbb{E}[Y^2] - \mathbb{E}[Y]^2\,.$$

Mit Hilfe von Satz 39 folgt die Behauptung. □

Für abhängige Zufallsvariablen $X_1, \ldots, X_n$ gilt Satz 54 im Allgemeinen nicht. Als Beispiel funktioniert wiederum der Fall $X = -Y$:

$$\mathrm{Var}[X+Y] = 0 \neq 2 \cdot \mathrm{Var}[X] = \mathrm{Var}[X] + \mathrm{Var}[Y]\,.$$

5. Wichtige diskrete Verteilungen

Wir diskutieren nun einige wichtige *diskrete* Verteilungen. Bei diesen Verteilungen handelt es sich um Funktionen, die von gewissen *Parametern* abhängen. Eigentlich betrachten wir also immer eine ganze Familie von ähnlichen Verteilungen.

5.1 Bernoulli-Verteilung

Eine Zufallsvariable X mit $W_X = \{0, 1\}$ und der Dichte

$$f_X(x) = \begin{cases} p & \text{für } x = 1, \\ 1 - p & \text{für } x = 0. \end{cases}$$

heißt Bernoulli-verteilt. Den Parameter p nennen wir Erfolgswahrscheinlichkeit.
Eine solche Verteilung erhält man z.B. bei einer einzelnen Indikatorvariablen. Es gilt mit $q := 1 - p$

$$\mathbb{E}[X] = p \text{ und } \mathrm{Var}[X] = pq,$$

wegen $\mathbb{E}[X^2] = p$ und $\mathrm{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = p - p^2$.

Der Name der Bernoulli-Verteilung geht zurück auf den Schweizer Mathematiker Jakob Bernoulli (1654–1705). Wie viele andere Mathematiker seiner Zeit hätte auch Bernoulli nach dem Wunsch seines Vaters ursprünglich Theologe werden sollen. Sein Werk *ars conjectandi* stellt eine der ersten Arbeiten dar, die sich mit dem Teil der Mathematik beschäftigen, den wir heute als Wahrscheinlichkeitstheorie bezeichnen.

5.2 Binomialverteilung

Eine Bernoulli-verteilte Zufallsvariable entspricht der Verteilung *einer* Indikatorvariablen. Häufig betrachtet man jedoch Summen von Indikatorvariablen.

Definition 55

Sei $X := X_1 + \ldots + X_n$ als Summe von n unabhängigen, Bernoulli-verteilten Zufallsvariablen mit gleicher Erfolgswahrscheinlichkeit p definiert. Dann heißt X binomialverteilt mit den Parametern n und p. In Zeichen schreiben wir

$$X \sim \mathrm{Bin}(n, p)\,.$$

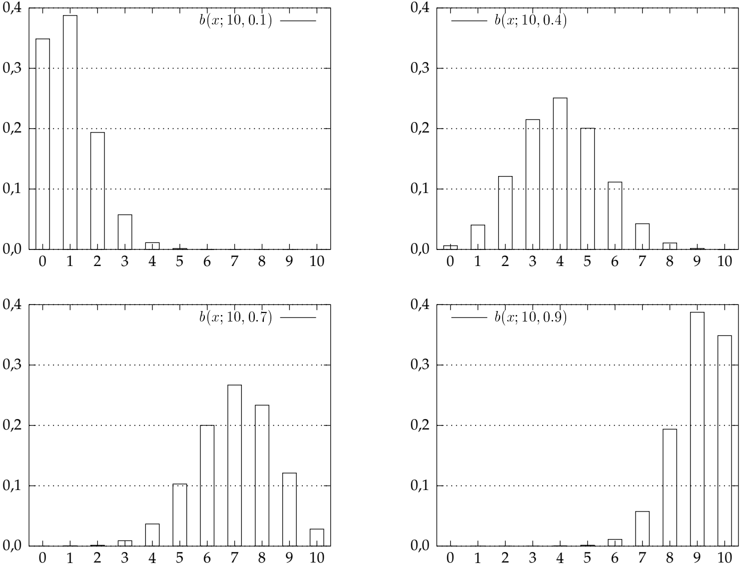
Es gilt $W_X = \{0, \ldots, n\}$. Die Binomialverteilung besitzt die Dichte

$$f_X(x) := b(x; n, p) = \binom{n}{x} p^x q^{n-x}$$

mit $q := 1 - p$. Da die Binomialverteilung eine sehr wichtige Rolle spielt, führen wir für die Dichtefunktion die Abkürzung $b(x; n, p)$ ein.

Mit den Sätzen über Erwartungswert und Varianz von Summen unabhängiger Zufallsvariablen erhalten wir sofort

$$\mathbb{E}[X] = np \quad \text{und} \quad \mathrm{Var}[X] = npq\,.$$



Dichte der Binomialverteilung

Satz 56

Wenn $X \sim \mathrm{Bin}(n_x, p)$ *und* $Y \sim \mathrm{Bin}(n_y, p)$ *unabhängig sind, dann gilt für* $Z := X + Y$*, dass* $Z \sim \mathrm{Bin}(n_x + n_y, p)$.

Beweis:

Die Aussage folgt sofort, wenn man gemäß der Definition der Binomialverteilung X und Y als Summen von Indikatorvariablen darstellt. Z ist dann offensichtlich wieder eine Summe von unabhängigen Indikatorvariablen. $\square$

5.3 Geometrische Verteilung

Definition 57

Die Dichte der geometrischen Verteilung mit Parameter/Erfolgswahrscheinlichkeit $p \in [0,1]$ und $q := 1-p$ ist gegeben durch

$$f_X(i) = pq^{i-1} \quad \text{für } i \in \mathbb{N}\,.$$

Für Erwartungswert und Varianz geometrisch verteilter Zufallsvariablen gilt

$$\mathbb{E}[X] = \frac{1}{p} \quad \text{und} \quad \mathrm{Var}[X] = \frac{q}{p^2},$$

denn es gilt:

$$\mathbb{E}[X] = \sum_{i=1}^{\infty} i \cdot pq^{i-1} = p \cdot \sum_{i=1}^{\infty} i \cdot q^{i-1} = p \cdot \frac{1}{(1-q)^2} = \frac{1}{p}.$$

$\mathbb{E}[X^2]$ ergibt sich gemäß der Formel (siehe DS I)

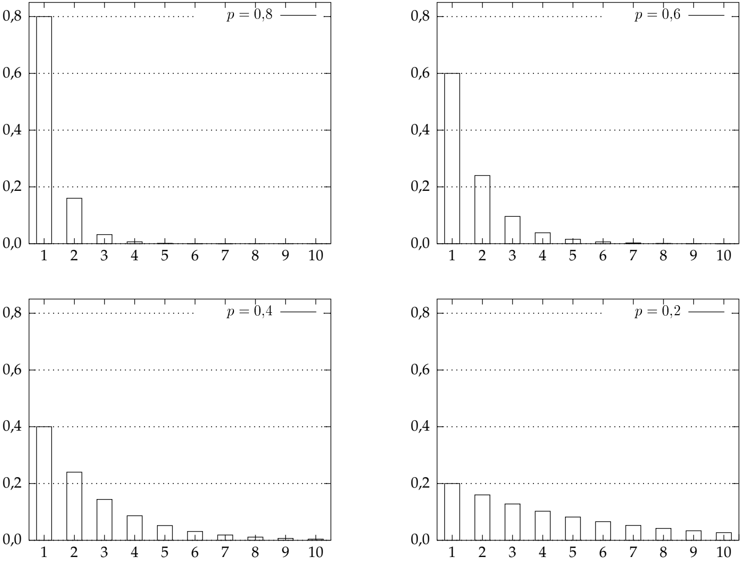
$$\sum_{n\geq 0}\binom{c+n-1}{n} z^n = \frac{1}{(1-z)^c} = (1-z)^{-c}$$

zu

$$\begin{aligned}
\mathbb{E}[X^2] &= \sum_{i=1}^{\infty} i^2 \cdot pq^{i-1}\\
&= p \cdot \left(q \sum_{i=0}^{\infty} (i+2)(i+1) \cdot q^i + \sum_{i=0}^{\infty} (i+1) \cdot q^i \right)\\
&= \frac{q \cdot 2}{p^2} + \frac{1}{p} = \frac{2-p}{p^2}\,,
\end{aligned}$$

und damit

$$\mathrm{Var}[X] = \frac{q}{p^2}\,.$$



Dichte der geometrischen Verteilung

Sei X geometrisch verteilt mit Erfolgswahrscheinlichkeit p. Dann ist $\Pr[X = k]$ die Wahrscheinlichkeit, dass wir bei einem binären Experiment mit Erfolgswahrscheinlichkeit p genau in der k-ten unabhängigen Wiederholung das erste Mal erfolgreich sind.

Wie groß ist die Wahrscheinlichkeit $\Pr[X > y + x \mid X > x]$?

Da bei den ersten x Versuchen kein Erfolg eintrat, stellen wir uns vor, dass das „eigentliche“ Experiment erst ab dem $(x + 1)$-ten Versuch beginnt. Die Zeit bis zum ersten Erfolg bei diesem neuen Experiment nennen wir X'. Damit $X > y + x$ gilt, muss $X' > y$ gelten. Es ist intuitiv, dass X' wieder geometrisch verteilt ist mit Erfolgswahrscheinlichkeit p, dass also für $x, y \in \mathbb{N}$ gilt:

$$\Pr[X > y + x \mid X > x] = \Pr[X' > y]. \tag{6}$$

Formal gilt

$$\Pr[X > x] = \sum_{i=x+1}^{\infty} (1-p)^{i-1} p = (1-p)^x p \cdot \sum_{i=0}^{\infty} (1-p)^i$$
$$= (1-p)^x p \cdot \frac{1}{1-(1-p)} = (1-p)^x \,,$$

sowie

$$\begin{aligned}
\Pr[X > y + x \mid X > x] &= \frac{\Pr[X > y+x, X > x]}{\Pr[X > x]} \\
&= \frac{\Pr[X > y+x]}{\Pr[X > x]} \\
&= (1-p)^{y+x} \cdot (1-p)^{-x} = (1-p)^y \\
&= \Pr[X > y] \,.
\end{aligned}$$

Diese Eigenschaft nennt man *Gedächtnislosigkeit*, da eine geometrisch verteilte Zufallsvariable gewissermaßen vergisst, dass sie schon x Misserfolge hinter sich hat und sich deshalb zum Zeitpunkt $y + x$ genauso verhält wie ursprünglich zur Zeit y.

Warten auf den n-ten Erfolg.

Wir betrachten n unabhängige Zufallsvariablen $X_1, \ldots, X_n$, die jeweils geometrisch verteilt sind mit Parameter p, und bestimmen die Dichte der Zufallsvariablen $Z := X_1 + \cdots + X_n$. Damit bezeichnet Z also die Anzahl der Versuche bis zum n-ten erfolgreichen Experiment (einschließlich).

Falls $Z = z$ ist, so werden also genau n erfolgreiche und $z - n$ nicht erfolgreiche Experimente durchgeführt. Dafür gibt es genau $\binom{z-1}{n-1}$ Möglichkeiten, von denen jede mit Wahrscheinlichkeit $p^n(1-p)^{z-n}$ eintritt. Es gilt also

$$f_Z(z) = \binom{z-1}{n-1} \cdot p^n(1-p)^{z-n}\,.$$

Die Zufallsvariable Z nennt man negativ binomialverteilt mit Ordnung n.

Das Coupon-Collector-Problem

In manchen Branchen legen Firmen den Verpackungen ihrer Produkte oft kleine Bilder oder andere Gegenstände bei, um den Käufer zum Sammeln anzuregen. Wenn es insgesamt n verschiedene solche Beilagen gibt, wie viele Packungen muss man im Mittel erwerben, bis man eine vollständige Sammlung besitzt? Hierbei nehmen wir an, dass bei jedem Kauf jede Beilage mit gleicher Wahrscheinlichkeit auftritt.

Sei

- X die Anzahl der zu tätigenden Käufe, und
- bezeichne Phase i die Schritte vom Erwerb der $(i-1)$-ten Beilage (ausschließlich) bis zum Erwerb der i-ten Beilage (einschließlich).

Sei etwa $n = 4$, und seien die Beilagen mit den Zahlen $1, 2, 3, 4$ identifiziert. Ein Experiment ist z.B.:

$$\underbrace{2}_{1}, \underbrace{2, 1}_{2}, \underbrace{2, 2, 3}_{3}, \underbrace{1, 3, 2, 3, 1, 4}_{4} .$$

Beobachtung:

Phase i endet genau dann, wenn wir eine der $n - i + 1$ Beilagen erhalten, die wir noch nicht besitzen.

Somit ist X_i geometrisch verteilt mit Parameter $p = \frac{n-i+1}{n}$ und es gilt $\mathbb{E}[X_i] = \frac{n}{n-i+1}$.

Damit folgt aber sofort

$$\begin{aligned}\mathbb{E}[X] &= \sum_{i=1}^{n} \mathbb{E}[X_i] \\ &= \sum_{i=1}^{n} \frac{n}{n-i+1} \\ &= n \cdot \sum_{i=1}^{n} \frac{1}{i} = n \cdot H_n,\end{aligned}$$

wobei $H_n := \sum_{i=1}^{n} \frac{1}{i}$ die n-te harmonische Zahl bezeichnet. Da $H_n = \ln n + O(1)$, folgt $\mathbb{E}[X] = n \ln n + O(n)$.