KAPITEL 6

Diskrete Verteilungen

Nun werden wir verschiedene Beispiele von diskreten Zufallsvariablen betrachten.

1. Gleichverteilung

DEFINITION 6.1. Eine Zufallsvariable $X : \Omega \to \mathbb{R}$ heißt *gleichverteilt* (oder *Laplace-verteilt*) auf einer endlichen Menge $\{y_1, \ldots, y_n\} \subset \mathbb{R}$, wenn

$$\mathbb{P}[X = y_i] = \frac{1}{n} \text{ für alle } i = 1, \ldots, n.$$

BEMERKUNG 6.2. Eine Zufallsvariable ist also gleichverteilt, wenn sie n Werte annehmen kann und die Wahrscheinlichkeiten dieser n Werte gleich sind.

BEMERKUNG 6.3. Für den Erwartungswert dieser Zufallsvariable gilt

$$\mathbb{E}X = \frac{y_1 + \ldots + y_n}{n}.$$

Dies ist das arithmetische Mittel von $y_1, \ldots, y_n$.

BEMERKUNG 6.4. Definition 6.1 funktioniert nur für endliches n. Eine Zufallsvariable kann nicht unendlich viele Werte mit gleicher Wahrscheinlichkeit annehmen. Hätte jeder Wert die gleiche, strikt positive Wahrscheinlichkeit $p > 0$, so wäre die Summe aller Wahrscheinlichkeiten unendlich. Hätte jeder Wert Wahrscheinlichkeit 0, so wäre die Summe aller Wahrscheinlichkeiten 0. Die Summe sollte aber 1 sein. In beiden Fällen ergibt sich ein Widerspruch. Eine Gleichverteilung (im obigen Sinne) auf einer unendlichen Menge gibt es also nicht.

2. Bernoulli-Experimente und die Binomialverteilung

DEFINITION 6.1. Ein *Bernoulli-Experiment* ist ein Zufallsexperiment mit zwei Ausgängen:

$$0 \quad (\text{“Misserfolg”}) \qquad \text{und} \qquad 1 \quad (\text{“Erfolg”}).$$

Die Wahrscheinlichkeit von “Erfolg” bezeichnen wir mit $p \in [0, 1]$. Die Wahrscheinlichkeit von “Misserfolg” ist dann $q := 1 - p$.

DEFINITION 6.2. Eine Zufallsvariable X heißt *Bernoulli-verteilt* mit Parameter $p \in [0, 1]$, falls

$$\mathbb{P}[X = 1] = p, \quad \mathbb{P}[X = 0] = 1 - p.$$

BEZEICHNUNG. $X \sim \text{Bern}(p)$.

DEFINITION 6.3. Ein *n-faches Bernoulli-Experiment* ist ein Bernoulli-Experiment, das n-mal unabhängig voneinander ausgeführt wurde.

Beispiel 6.4. Wir können eine (faire oder unfaire) Münze n-mal werfen und zum Beispiel "Kopf" als "Erfolg" auffassen.

Beispiel 6.5. Wir können einen Würfel n-mal werfen. Fassen wir eine 6 als einen "Erfolg" auf, so erhalten wir ein n-faches Bernoulli-Experiment mit Erfolgswahrscheinlichkeit $p = \frac{1}{6}$.

Bemerkung 6.6. Die Grundmenge eines n-fachen Bernoulli-Experiments ist $\Omega = \{0,1\}^n$. Wegen der Unabhängikeit der einzelnen Experimente ist die Wahrscheinlichkeit eines Ausgangs $(a_1, \dots, a_n) \in \Omega$ gegeben durch

$$p(a_1, \dots, a_n) = p^k (1-p)^{n-k},$$

wobei $k = \sum_{i=1}^n a_i$ die Anzahl der Einsen unter $a_1, \dots, a_n$ ist. Für $p \neq 1/2$ sind die Ausgänge nicht gleichwahrscheinlich.

Satz 6.7. *Sei X die Anzahl der "Erfolge" in einem n-fachen Bernoulli-Experiment mit Erfolgswahrscheinlichkeit p. Dann gilt:*

$$\mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k} \text{ für alle } k = 0, 1, \dots, n. \tag{6.1}$$

Beweis. Sei $k \in \{0, \dots, n\}$. Wir betrachten das Ereignis $\{X = k\}$. Es besteht aus allen Ausgängen $(a_1, \dots, a_n) \in \{0,1\}^n$ mit genau k Einsen. Es gibt genau $\binom{n}{k}$ solche Ausgänge. Jeder dieser Ausgänge hat Wahrscheinlichkeit von jeweils $p^k(1-p)^{n-k}$. Für die Wahrscheinlichkeit des Ereignisses $\{X = k\}$ ergibt sich somit Formel (6.1). □

Bemerkung 6.8. Eine Zufallsvariable X, die (6.1) erfüllt, heißt *binomialverteilt* mit Parametern $n \in \mathbb{N}$ und $p \in [0,1]$.

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

Bemerkung 6.9. Die Summe der Wahrscheinlichkeiten aller Werte einer diskreten Zufallsvariable sollte 1 ergeben. Dies ist bei der Binomialverteilung der Fall, denn

$$\sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} = (p + (1-p))^n = 1.$$

Dabei haben wir die binomische Formel benutzt, daher die Bezeichnung "Binomialverteilung".

Satz 6.10. *Für $X \sim Bin(n, p)$ gilt $\mathbb{E}X = np$. In Worten: Die erwartete Anzahl von "Erfolgen" in einem n-fachen Bernoulli-Experiment mit Erfolgswahrscheinlichkeit p ist gleich np.*

Beweis. Wir definieren die Zufallsvariablen $X_1, \dots, X_n : \{0,1\}^n \to \mathbb{R}$ wie folgt:

$$X_i(a_1, \dots, a_n) = a_i, \text{ wobei } (a_1, \dots, a_n) \in \{0,1\}^n.$$

In Worten:

$$X_i = \begin{cases} 1, & \text{falls Experiment } i \text{ ein "Erfolg" ist,} \\ 0, & \text{falls Experiment } i \text{ ein "Misserfolg" ist.} \end{cases}$$

Da die Erfolgswahrscheinlichkeit in jedem Experiment gleich p ist, gilt

$$\mathbb{P}[X_i = 1] = p, \quad \mathbb{P}[X_i = 0] = 1 - p \text{ für alle } i = 1, \ldots, n.$$

Für den Erwartungswert von X_i gilt somit:

$$\mathbb{E}X_i = p \cdot 1 + (1-p) \cdot 0 = p \text{ für alle } i = 1, \ldots, n.$$

Die Anzahl der "Erfolge" im n-fachen Bernoulli-Experiment ist gegeben durch

$$X = X_1 + \ldots + X_n.$$

Aus der Additivität des Erwartungswerts folgt, dass $\mathbb{E}X = \mathbb{E}X_1 + \ldots + \mathbb{E}X_n = np$. □

3. Poisson-Verteilung

Definition 6.1. Eine Zufallsvariable X hat *Poisson-Verteilung* mit Parameter (auch Intensität genannt) $\lambda > 0$, wenn

$$\mathbb{P}[X = k] = e^{-\lambda} \frac{\lambda^k}{k!} \text{ für alle } k = 0, 1, 2, \ldots. \tag{6.2}$$

Bezeichnung. $X \sim \mathrm{Poi}(\lambda)$.

Bemerkung 6.2. Die Wahrscheinlichkeiten in (6.2) summieren sich zu 1, denn

$$\sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1.$$

Satz 6.3. *Für $X \sim Poi(\lambda)$ gilt $\mathbb{E}X = \lambda$.*

Beweis. Wir verwenden die Definition des Erwartungswertes:

$$\mathbb{E}X = \sum_{k=0}^{\infty} k \cdot \mathbb{P}[X = k] = \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda \cdot \lambda^{k-1}}{(k-1)!} = e^{-\lambda} \lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda.$$

Dabei haben wir $m = k - 1$ gesetzt. □

Die Poisson-Verteilung entsteht als Grenzwert der Binomialverteilung. Das wird im folgenden Satz beschrieben.

Satz 6.4 (Poisson-Grenzwertsatz). *Sei $p_n \in (0, 1)$ eine Folge mit*

$$\lim_{n \to \infty} np_n = \lambda \in (0, \infty). \tag{6.3}$$

Sei S_n eine Zufallsvariable mit $S_n \sim Bin(n, p_n)$. Für jedes $k = 0, 1, 2, \ldots$ gilt dann

$$\lim_{n \to \infty} \mathbb{P}[S_n = k] = e^{-\lambda} \frac{\lambda^k}{k!}. \tag{6.4}$$

Beispiel 6.5. Man stelle sich S_n vor, als die Anzahl der "Erfolge" in einem n-fachen Bernoulli-Experiment mit einem sehr großen n und einer sehr kleinen Erfolgswahrscheinlichkeit

$$p_n \approx \frac{\lambda}{n}.$$

Der Poisson-Grenzwertsatz besagt, dass die Anzahl der "Erfolge" in einem solchen Experiment approximativ Poisson-verteilt ist. Beispiele von Zufallsvariablen, die approximativ Poisson-verteilt sind:

(1) Anzahl der Schäden, die einer Versicherung gemeldet werden (viele Versicherungsverträge, jeder Vertrag erzeugt mit einer sehr kleinen Wahrscheinlichkeit einen Schaden).
(2) Anzahl der Druckfehler in einem Buch (viele Buchstaben, jeder Buchstabe kann mit einer sehr kleinen Wahrscheinlichkeit ein Druckfehler sein).
(3) Anzahl der Zugriffe auf einen Webserver (viele User, jeder User greift mit einer sehr kleinen Wahrscheinlichkeit zu).

Beweis von Satz 6.4. Sei $k \in \mathbb{N}_0$ fest. Da S_n binomialverteilt ist, gilt:

$$\begin{aligned}
\mathbb{P}[S_n = k] &= \binom{n}{k} \cdot p_n^k \cdot (1-p_n)^{n-k} \\
&= \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{k!} \cdot p_n^k \cdot (1-p_n)^{n-k} \\
&= \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} \cdot \frac{(np_n)^k}{k!} \cdot \left(1 - \frac{np_n}{n}\right)^n \cdot (1-p_n)^{-k} \\
&\underset{n\to\infty}{\longrightarrow} 1 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1.
\end{aligned}$$

Dabei haben wir benutzt, dass $\lim_{n\to\infty} p_n = 0$. Dies folgt aus der Annahme (6.3). Außerdem haben wir die folgende Formel benutzt:

$$\lim_{n\to\infty} \left(1 - \frac{\lambda_n}{n}\right)^n = e^{-\lambda},$$

für jede Folge λ_n mit $\lim_{n\to\infty} \lambda_n = \lambda \in (0,\infty)$. In unserem Fall war $\lambda_n = np_n$. □

Beispiel 6.6. Im Hörsaal befinden sich $n = 100$ Personen. Betrachte das Ereignis

$A =$ "mindestens eine Person im Hörsaal hat heute Geburtstag".

Bestimme die Wahrscheinlichkeit von A.

Wir werden zwei Lösungen präsentieren. Die erste Lösung ist exakt, die zweite approximativ.

Lösung 1 (exakt). Wir nummerieren die Personen mit $1, \ldots, n$. Wir betrachten das Ereignis "Person i hat heute Geburtstag" als "Erfolg" im i-ten Bernoulli-Experiment. Die Wahrscheinlichkeit von "Erfolg" ist für jede Person i gegeben durch

$$p := \mathbb{P}[\text{Person } i \text{ hat heute Geburtstag}] = \frac{1}{365}.$$

Dabei können wir die Geburtstage der verschiedenen Personen als unabhängig betrachten. Es handelt sich also um ein $n = 100$-faches Bernoulli-Experiment mit Erfolgswahrscheinlichkeit $p = \frac{1}{365}$. Die Anzahl der Personen im Hörsaal, die heute Geburtstag haben, ist $\text{Bin}(100, \frac{1}{365})$-verteilt.
Für die Wahrscheinlichkeit von A^c erhalten wir:

$$\mathbb{P}[A^c] = \mathbb{P}[\text{keine Person im Hörsaal hat heute Geburtstag}] = (1-p)^n.$$

Es folgt, dass

$$\mathbb{P}[A] = 1 - (1-p)^n = 1 - \left(1 - \frac{1}{365}\right)^{100} \approx 0.239933.$$

Lösung 2 (approximativ). Die Anzahl der Personen im Hörsaal, die heute Geburtstag haben, ist binomialverteilt mit $n = 100$ und $p = \frac{1}{365}$. Die Wahrscheinlichkeit p ist sehr klein, die Anzahl der Personen n ist sehr groß. Deshalb benutzen wir die Poisson-Approximation:

$$\text{Bin}\left(100, \frac{1}{365}\right) \approx \text{Poi}\left(\frac{100}{365}\right).$$

Somit ist die Anzahl der Personen, die heute Geburtstag haben, approximativ Poisson-verteilt mit Parameter $\lambda = np = \frac{100}{365}$. Für die Wahrscheinlichkeit von A^c erhalten wir aus der Formel (6.2) mit $k = 0$:

$$\mathbb{P}[A^c] = \mathbb{P}[\text{keiner im Hörsaal hat heute Geburtstag}] \approx e^{-\lambda} \cdot \frac{\lambda^0}{0!} = e^{-\lambda}.$$

Es folgt, dass

$$\mathbb{P}[A] \approx 1 - e^{-\lambda} = 1 - e^{-\frac{100}{365}} \approx 0.239647.$$

4. Geometrische Verteilung

Wir betrachten ein Bernoulli-Experiment mit Erfolgswahrscheinlichkeit $p \in (0,1]$, das unendlich oft und unabhängig wiederholt wird. Die Grundmenge ist dann die Menge aller unendlichen Folgen aus Nullen und Einsen:

$$\Omega = \{0,1\}^\infty \stackrel{def}{=} \{(a_1, a_2, \dots) : a_i \in \{0,1\} \text{ für alle } i \in \mathbb{N}\}.$$

Diese Menge ist überabzählbar. Solche Experimente werden wir später genauer betrachten. Nun legen wir eine Zufallsvariable $T : \Omega \to \mathbb{R}$ fest:

$$T(a_1, a_2, \dots) \stackrel{def}{=} \min\{n \in \mathbb{N} : a_n = 1\}.$$

Die Zufallsvariable T ist somit die Wartezeit auf den ersten "Erfolg".

Beispiel 6.1. Gehen die Experimente wie folgt aus:

$$0,0,0,1,0,0,1,0,1,1,0,\dots,$$

so erhalten wir $T = 4$.

Was für eine Verteilung hat nun T?

Satz 6.2. *Es gilt*

$$\mathbb{P}[T = k] = (1-p)^{k-1}p \text{ für alle } k = 1,2,\dots. \tag{6.5}$$

Bemerkung 6.3. Eine Zufallsvariable T, die (6.5) mit einem $p \in (0,1]$ erfüllt, heißt *geometrisch verteilt* mit Parameter p.

Bezeichnung. $T \sim \text{Geo}(p)$.

Bemerkung 6.4. Die Wahrscheinlichkeiten in (6.5) summieren sich zu 1, denn

$$\sum_{k=1}^{\infty}(1-p)^{k-1}p = p \cdot \sum_{k=1}^{\infty}(1-p)^{k-1} = p \cdot \frac{1}{1-(1-p)} = 1.$$

Dabei haben wir eine geometrische Reihe summiert, daher die Bezeichnung "geometrische Verteilung".

Beweis von Satz 6.2. Wir benutzen die Notation $0 =$ "Misserfolg" und $1 =$ "Erfolg". Sei $k \in \mathbb{N}$ fest. Damit das Ereignis $\{T = k\}$ eintritt, müssen die ersten k Experimente so ausgehen:

$$0, 0, \dots, 0, 1.$$

Dabei ist es egal, wie alle anderen Experimente ausgehen. Wegen der Unabhängigkeit der einzelnen Bernoull-Experimente gilt für die Wahrscheinlichkeit davon:

$$\mathbb{P}[T = k] = (1-p) \cdot \ldots \cdot (1-p) \cdot p = (1-p)^{k-1}p.$$

□

Satz 6.5. *Für $T \sim$ Geo(p) gilt $\mathbb{E}T = \frac{1}{p}$. In Worten: Die durchschnittliche Wartezeit auf den ersten Erfolg in einem unendlich oft wiederholten Bernoulli-Experiment mit Erfolgswahrscheinlichkeit p ist gleich $\frac{1}{p}$.*

Beweis. Wiederum wird die Definition des Erwartungswerts benutzt:

$$\mathbb{E}T = \sum_{k=1}^{\infty} k \cdot \mathbb{P}[T = k] = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p \sum_{k=1}^{\infty} kq^{k-1},$$

wobei $q = 1 - p$. Die Summe auf der rechten Seite können wir wie folgt berechnen:

$$\sum_{k=1}^{\infty} kq^{k-1} = \left(\sum_{k=1}^{\infty} q^k\right)' = \left(\sum_{k=0}^{\infty} q^k\right)' = \left(\frac{1}{1-q}\right)' = \frac{1}{(1-q)^2}.$$

Es folgt, dass

$$\mathbb{E}T = p \cdot \frac{1}{(1-q)^2} = p \cdot \frac{1}{p^2} = \frac{1}{p}.$$

□

Beispiel 6.6. Wir würfeln mit einem fairen Würfel so lange, bis eine 1 kommt. Wie lange müssen wir im Durchschnitt warten?

Lösung. Die Wartezeit T ist Geo($\frac{1}{6}$)-verteilt. Der Erwartungswert von T ist:

$$\mathbb{E}T = \frac{1}{1/6} = 6.$$

Dieses Ergebnis steht im Einklang mit der Intuition: Im Durchschnitt ist jeder sechste Wurf eine 1, deshalb brauchen wir im Durchschnitt 6 Würfe, um eine 1 zu würfeln.

Bemerkung 6.7 (zum folgenden Satz). Angenommen, wir haben 100 Mal gewürfelt ohne auch ein einziges Mal eine 1 zu erzielen. Dies ist zwar sehr unwahrscheinlich, jedoch nicht unmöglich. Die Frage ist nun, wie lange müssen wir jetzt noch warten, bis eine 1 kommt? Man könnte meinen, dass aufgrund dessen, dass die 1 schon sehr lange überfällig ist, diese nun sehr bald kommen muss. Das ist jedoch nicht der Fall, da der Würfel kein Gedächtnis hat und von der Geschichte der bereits ausgeführten Würfe nichts weiß. Die Anzahl der Würfe, die nun noch benötigt werden, bis eine 1 kommt, ist nach wie vor geometrisch verteilt mit Parameter $\frac{1}{6}$. Diese Eigenschaft wird nun im folgenden Satz beschrieben.

Satz 6.8 (Gedächtnislosigkeit der geometrischen Verteilung). *Sei $T \sim Geo(p)$, dann gilt:*

$$\mathbb{P}[T-n>k|T>n]=\mathbb{P}[T>k] \text{ für alle } n,k\in\mathbb{N}.$$

Beweis. Sei $m \in \mathbb{N}$. Zuerst berechnen wir die Wahrscheinlichkeit, dass $T > m$. Dieses Ereignis tritt genau dann ein, wenn die ersten m Experimente "Misserfolge" sind. Die Ausgänge der anderen Experimente sind dabei egal. Wegen der Unabhängigkeit der einzelnen Experimente hat dieses Ereignis Wahrscheinlichkeit $(1-p)^m$. Man kann auch direkt vorgehen:

$$\mathbb{P}[T>m]=\sum_{i=m+1}^{\infty}\mathbb{P}[T=i]=\sum_{i=m+1}^{\infty}p(1-p)^{i-1}=(1-p)^m\sum_{i=1}^{\infty}p(1-p)^{i-1}=(1-p)^m.$$

Nun erhalten wir mit der Definition der bedingten Wahrscheinlichkeit, dass

$$\mathbb{P}[T-n>k|T>n]=\frac{\mathbb{P}[T>n+k,T>n]}{\mathbb{P}[T>n]}=\frac{\mathbb{P}[T>n+k]}{\mathbb{P}[T>n]}=\frac{(1-p)^{n+k}}{(1-p)^n}=(1-p)^k.$$

Dies stimmt mit $\mathbb{P}[T>k]$ überein. □

5. Negative Binomialverteilung

Wir betrachten wieder ein unendlich oft wiederholtes Bernoulli-Experiment mit Erfolgswahrscheinlichkeit $p \in (0,1]$. Für $r \in \mathbb{N}$ bezeichnen wir mit T_r die Wartezeit auf den r-ten "Erfolg".

Beispiel 6.1. Gehen die Experimente wie folgt aus:

$$0,0,0,1,0,0,1,0,1,1,0,\ldots,$$

so erhalten wir $T_1 = 4$, $T_2 = 7$, $T_3 = 9$, $T_4 = 10$. Dabei steht 0 für "Misserfolg" und 1 für "Erfolg".

Wir haben bereits gezeigt, dass $T_1 \sim \text{Geo}(p)$. Wie ist nun T_r für ein allgemeines $r \in \mathbb{N}$ verteilt?

Satz 6.2. *Für jedes $r \in \mathbb{N}$ ist die Wartezeit T_r auf den r-ten "Erfolg" in einem unendlich oft wiederholten Bernoulli-Experiment mit Erfolgswahrscheinlichkeit p wie folgt verteilt:*

$$\mathbb{P}[T_r=k]=\binom{k-1}{r-1}p^r(1-p)^{k-r} \text{ für alle } k=r,r+1,\ldots. \tag{6.6}$$

Bemerkung 6.3. Die Mindestanzahl an Experimenten, die man benötigt, um r "Erfolge" zu erzielen, ist r. Daher ist der kleinste mögliche Wert von T_r gleich r.

Definition 6.4. Eine Zufallsvariable T_r, die (6.6) mit einem $r \in \mathbb{N}$ und einem $p \in (0,1]$ erfüllt, heißt *negativ binomialverteilt* mit Parametern r und p.

Bezeichnung. $T_r \sim \mathrm{NB}(r,p)$.

Bemerkung 6.5. Die geometrische Verteilung ist ein Spezialfall der negativen Binomialverteilung: $\mathrm{Geo}(p) = \mathrm{NB}(1,p)$.

Bemerkung 6.6. Negative Binomialverteilung wird auch Pascal- oder Polya-Verteilung genannt.

Beweis von Satz 6.2. Seien $r \in \mathbb{N}$ und $k \geq r$ fest. Das Ereignis $\{T_r = k\}$ tritt genau dann ein, wenn die beiden folgenden Ereignisse eintreten:

$A =$ "Das k-te Experiment ist ein "Erfolg"",

$B =$ "In den Experimenten $1,\ldots,k-1$ werden genau $r-1$ "Erfolge" erzielt".

Die Geschichte der Bernoulli-Experimente muss also wie folgt aussehen:

$$\underbrace{\underset{1}{?}\underset{2}{?}\underset{3}{?}\ldots\ldots\ldots\underset{k-1}{?}}_{r-1 \text{ "Erfolge"}}\underset{k}{1},$$

wobei das Fragezeichen für 0 oder 1 steht und genau $r-1$ Fragezeichen Einsen sein sollen. Die Wahrscheinlichkeit von A ist p. Die Wahrscheinlichkeit von B berechnen wir mit Hilfe der Binomialverteilung:

$$\mathbb{P}[B] = \binom{k-1}{r-1} p^{r-1}(1-p)^{(k-1)-(r-1)} = \binom{k-1}{r-1} p^{r-1}(1-p)^{k-r}.$$

Dabei sind A und B unabhängig. Es folgt:

$$\mathbb{P}[T_r = k] = \mathbb{P}[A] \cdot \mathbb{P}[B] = p \cdot \binom{k-1}{r-1} p^{r-1}(1-p)^{k-r} = \binom{k-1}{r-1} p^{r}(1-p)^{k-r}.$$

□

Bemerkung 6.7. Wir zeigen noch, dass die Summe aller Wahrscheinlichkeiten in (6.6) gleich 1 ist:

$$\sum_{k=r}^{\infty} \binom{k-1}{r-1} p^{r}(1-p)^{k-r} = 1.$$

Eigentlich folgt das aus Satz 6.2. Wir geben aber einen direkten Beweis. Wir nehmen uns die Binomische Formel zur Hilfe:

$$(1+x)^{\alpha} = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \quad |x| < 1.$$

Diese Formel gilt für alle $\alpha \in \mathbb{R}$. Dabei muss α nicht unbedingt ganz und nicht unbedingt positiv sein. Der Binomialkoeffizient $\binom{\alpha}{k}$ ist definiert durch

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\ldots(\alpha-k+1)}{k!}, \qquad \alpha \in \mathbb{R}, k \in \mathbb{N}_0.$$

Wir setzen $\alpha = -r$ und $-x$ anstatt von x in die Formel ein:

$$\begin{aligned}(1-x)^{-r} &= \sum_{k=0}^{\infty} \binom{-r}{k} (-x)^k \\ &= \sum_{k=0}^{\infty} \frac{(-r)(-r-1)\cdot \ldots \cdot (-r-k+1)}{k!} \cdot (-1)^k x^k \\ &= \sum_{k=0}^{\infty} \frac{r(r+1)\cdot \ldots \cdot (r+k-1)}{k!} \cdot x^k \\ &= \sum_{k=0}^{\infty} \binom{r+k-1}{k} x^k.\end{aligned}$$

Mit $m = k + r$ erhalten wir dann

$$(1-x)^{-r} = \sum_{m=r}^{\infty} \binom{m-1}{m-r} x^{m-r} = \sum_{m=r}^{\infty} \binom{m-1}{r-1} x^{m-r}.$$

Somit erhalten wir schließlich mit $x = 1 - p$:

$$\sum_{k=r}^{\infty} \binom{k-1}{r-1} p^r (1-p)^{k-r} = p^r \sum_{m=r}^{\infty} \binom{m-1}{r-1} x^{m-r} = p^r (1-x)^{-r} = 1.$$

Die Verteilung heißt "negative Binomialverteilung", da wir in der Binomischen Formel für α einen negativen Wert eingesetzt haben.

Satz 6.8. *Für* $T_r \sim NB(r,p)$ *gilt* $\mathbb{E}T_r = \frac{r}{p}$.

Beweisidee. Auf einen "Erfolg" muss man im Durchschnitt $\frac{1}{p}$ Experimente warten. Auf r "Erfolge" wartet man dementsprechend im Durchschnitt $\frac{r}{p}$ Experimente. □