## 1.3 Zufallsgrößen und Verteilungsfunktionen

### 1.3.1 Einführung

Vielfach sind die Ergebnisse von Zufallsversuchen Zahlenwerte. Häufig möchte man aber auch in den Fällen, wo dies nicht so ist, Zahlenwerte zur Charakterisierung der Ergebnisse von Zufallssituationen verwenden. Dies geschieht mit Hilfe von Zufallsgrößen $X$, indem jedem Ergebnis $\omega$ aus der Ergebnismenge $\Omega$ eine relle Zahl $X(\omega)$ als Wert der Zufallsgröße zugeordnet wird.

**Definition 1.3.1.** Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum zu einer festen Zufallssituation. Dann heißt eine Abbildung $X : \Omega \mapsto \mathbb{R}$ *Zufallsgröße* oder *Zufallsvariable* über $(\Omega, \mathfrak{A}, P)$, wenn für alle Intervalle $I \subset \mathbb{R}$ gilt:

$$\{\omega \in \Omega : X(\omega) \in I\} \in \mathfrak{A}.$$

Für die Wahrscheinlichkeit $P(\{\omega \in \Omega : X(\omega) \in I\})$ schreiben wir verkürzt $P(X \in I)$.

**Beispiel.** In einer Hühnerhaltung wird das Gewicht von Eiern in Gramm ermittelt. Das Gewicht $\omega$ eines Eies ist eine zufällige positive reelle Zahl. Die Zufallsgröße $X = X(\omega)$ ordnet den Eiern eine der drei Gewichtsklassen 1, 2 oder 3 zu:

$$X(\omega) := \begin{cases} 1, & \omega \leq 40 \\ 2, & 40 < \omega \leq 60 \\ 3, & \omega > 60 \end{cases}.$$

**Beispiel.** Beim Werfen zweier idealer Würfel erhält man die Ergebnismenge

$$\Omega = \big\{(i, j) : i, j \in \{1, 2, 3, 4, 5, 6\}\big\}.$$

Die Augensumme $X(\omega) := i + j$ ist dann eine Zufallsgröße.

**Beispiel.** Es wird die Lebensdauer von $n$ Glühlampen betrachtet, wobei $\omega_i$ die Brenndauer der $i$-ten Glühlampe in Stunden bezeichnet. Haben also

$$\Omega = \{\omega = (\omega_1, \ldots, \omega_n) : \omega_i \geq 0\}$$

als Ergebnismenge. Sowohl $X(\omega) := \omega_k$ (Lebensdauer der $k$-ten Glühlampe) als auch $X(\omega) := \frac{\omega_1 + \ldots + \omega_n}{n}$ (mittlere Lebensdauer der Glühlampen) stellen Zufallsgrößen dar.

### 1.3.2 Diskrete Zufallsgrößen

**Definition 1.3.2.** Eine Zufallsgröße $X$ heißt *diskret*, wenn sie nur endlich oder abzählbar unendlich viele Werte annehmen kann.

**Definition 1.3.3.** Sei $X$ eine diskrete Zufallsgröße mit den Werten $x_1, x_2, \ldots$ (endlich oder abzählbar unendlich viele) und $p_i := P(X = x_i)$. Dann heißt die Zuordnung $x_i \mapsto p_i$ *Wahrscheinlichkeitsfunktion* der Zufallsgröße $X$.

Eine diskrete Zufallsgröße wird vollständig durch ihre Wahrscheinlichkeitsfunktion bestimmt.

**Beispiel.** Beim Werfen eines idealen Würfels ist die Anzahl der möglichen Augenzahlen $x_i$ endlich. Somit kann man die Wahrscheinlichkeitsfunktion als Tabelle darstellen.

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 |
| $p_i$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

### Eigenschaften der Wahrscheinlichkeitsfunktion

- $0 \leq p_i \leq 1$,
- $\sum\limits_i p_i = 1$,
- $P(a \leq X < b) = \sum\limits_{a \leq x_i < b} p_i$.

**Definition 1.3.4.** Sei $X$ eine Zufallsgröße. Dann heißt die Funktion $F(x) := P(X < x)$ *Verteilungsfunktion* von X. Für eine diskrete Zufallsgröße $X$ mit den Werten $x_1, x_2, \ldots$ gilt also

$$F(x) = \sum_{x_i < x} p_i.$$

Bei diskreten Zufallsgrößen ist die Verteilungsfunktion immer eine reine Treppenfunktion. Die Punkte $x_i$ kennzeichnen die Sprungpunkte und die Werte $p_i$ die zugehörigen Sprunghöhen.

### Eigenschaften der Verteilungsfunktion

- $\lim\limits_{x \to -\infty} F(x) = 0$,
- $\lim\limits_{x \to \infty} F(x) = 1$,
- $x_1 < x_2 \Rightarrow F(x_1) \leq F(x_2)$, d.h. $F$ ist monoton wachsend (nicht notwendigerweise streng),
- $\lim\limits_{x \to x_0 - 0} F(x) = F(x_0)$, d.h. $F$ ist linksseitig stetig.

**Bekannte Zahlenreihen** Wir setzen die folgenden drei Zahlenreihen als bekannt voraus und werden sie im Weiteren benutzen:

$$\sum_{i=0}^{\infty} x^i = \frac{1}{1-x} \quad \text{für} \quad |x| < 1, \tag{1.3.1}$$

$$\sum_{i=1}^{\infty} i x^{i-1} = \frac{1}{(1-x)^2} \quad \text{für} \quad |x| < 1, \tag{1.3.2}$$

$$\sum_{i=2}^{\infty} i(i-1) x^{i-2} = \frac{2}{(1-x)^3} \quad \text{für} \quad |x| < 1. \tag{1.3.3}$$

#### 1.3.2.1 Erwartungswert und Varianz

**Beispiel.** Eine neue Maschine zur Produktion elektronischer Bauteile wird über eine Dauer von $n$ Tagen getestet, um herauszufinden, wie viele fehlerfreie Teile sie im Durchschnitt am

Tag liefert. Dabei bezeichne $n_i$ die Anzahl der Tage, an denen genau $i$ funktionstüchtige Teile hergestellt wurden. Es gilt also $\sum_{i=0}^{\infty} n_i = n$ und wir erhalten als Ergebnis

$$\sum_{i=0}^{\infty} i\frac{n_i}{n} = \sum_{i=0}^{\infty} iH_n(X = i) \xrightarrow{n\to\infty} \sum_{i=0}^{\infty} iP(X = i).$$

**Definition 1.3.5.** Ist $p_i = P(X = x_i)$ die Wahrscheinlichkeitsfunktion einer diskreten Zufallsgröße $X$, so wird

$$\mathbf{E}X := \sum_i x_i p_i$$

*Erwartungswert* oder *Mittelwert* der Zufallsgröße $X$ genannt. Der Erwartungswert ist eine endliche reelle Zahl, falls gilt

$$\sum_i |x_i| p_i < \infty.$$

**Beispiel.** Beim Würfeln mit einem idealen Würfel sei $X$ die gewürfelte Augenzahl. Dann ist

$$\mathbf{E}X = \sum_{i=1}^{6} i\frac{1}{6} = 3{,}5,$$

d.h. im Mittel wird die Augenzahl 3,5 erreicht.

Neben dem Erwartungswert für eine Zufallsgröße $X$ kann auch der Erwartungswert von Funktionen $g(X)$ einer Zufallsgröße $X$ betrachtet werden, z.B. für $g(X) = X^2$ oder $g(X) = \sin(X)$.

**Definition 1.3.6.** Ist $p_i = P(X = x_i)$ die Wahrscheinlichkeitsfunktion einer diskreten Zufallsgröße $X$, so wird

$$\mathbf{E}g(X) := \sum_i g(x_i) p_i$$

*Erwartungswert der Funktion* $g(X)$ genannt. Der Erwartungswert einer Funktion ist eine endliche reelle Zahl, falls gilt

$$\sum_i |g(x_i)| p_i < \infty.$$

**Definition 1.3.7.** Die Größe

$$\sigma^2 := \mathbf{D}^2 X := \mathbf{E}(X - \mathbf{E}X)^2$$

heißt *Varianz* (oder *Streuung* oder *Dispersion*) der Zufallsgröße $X$ und gibt die mittlere quadratische Abweichung der Zufallsgröße $X$ von ihrem Erwartungswert $\mathbf{E}X$ an. Die Größe $\sigma := \sqrt{\mathbf{D}^2 X}$ heißt *Standardabweichung* der Zufallsgröße $X$.

**Beispiel.** Wir betrachten nochmals das Würfeln mit einem idealen Würfel (siehe vorhergehendes Beispiel). Haben also $\mathbf{E}X = 3{,}5$. Setzen wir $g(X) := (X - \mathbf{E}X)^2$, so erhalten wir als Varianz der Zufallsgröße $X$

$$\mathbf{D}^2 X = \mathbf{E}(X - \mathbf{E}X)^2 = \mathbf{E}g(X) = \sum_{i=1}^{6} \frac{1}{6} g(i) = \frac{1}{6}\sum_{i=1}^{6} (i - 3{,}5)^2 = 2{,}92.$$

Die Standardabweichung beträgt somit $\sigma = 1{,}71$.

**Hilfssatz 1.3.8.** *Für eine Zufallsgröße $X$ gilt mit $a, b \in \mathbb{R}$:*

$$\mathbf{E}(aX + b) = a\mathbf{E}X + b.$$

*Beweis (für diskrete Zufallsgrößen).*

$$\mathbf{E}(aX + b) = \sum_i (ax_i + b)p_i = \sum_i ax_i p_i + \sum_i bp_i = a \underbrace{\sum_i x_i p_i}_{\mathbf{E}X} + b \underbrace{\sum_i p_i}_{1} = a\mathbf{E}X + b.$$

□

**Hilfssatz 1.3.9.** *Für zwei Funktionen $f(X)$ und $g(X)$ einer Zufallsgröße $X$ gilt:*

$$\mathbf{E}(f(X) + g(X)) = \mathbf{E}f(X) + \mathbf{E}g(X).$$

*Beweis (für diskrete Zufallsgrößen).*

$$\mathbf{E}(f(X) + g(X)) = \sum_i (f(x_i) + g(x_i))p_i = \sum_i f(x_i)p_i + \sum_i g(x_i)p_i = \mathbf{E}f(X) + \mathbf{E}g(X).$$

□

**Hilfssatz 1.3.10.** *Für eine Zufallsgröße $X$ gilt mit $a, b \in \mathbb{R}$:*

$$\mathbf{D}^2(aX + b) = a^2 \mathbf{D}^2 X.$$

*Beweis.*

$$\mathbf{D}^2(aX + b) = \mathbf{E}(aX + b - \underbrace{\mathbf{E}(aX + b)}_{a\mathbf{E}X + b})^2 = \mathbf{E}(aX - a\mathbf{E}X)^2 = a^2\mathbf{E}(X - \mathbf{E}X)^2 = a^2 D^2 X.$$

□

**Satz 1.3.11.** *Für eine Zufallsgröße $X$ gilt:*

$$\mathbf{D}^2 X = \mathbf{E}X^2 - (\mathbf{E}X)^2.$$

*Beweis.*

$$\begin{aligned} \mathbf{D}^2 X &= \mathbf{E}(X - \mathbf{E}X)^2 = \mathbf{E}(X^2 - 2X\mathbf{E}X + (\mathbf{E}X)^2) \\ &= \mathbf{E}X^2 - \underbrace{\mathbf{E}(2X\mathbf{E}X)}_{2(\mathbf{E}X)^2} + (\mathbf{E}X)^2 = \mathbf{E}X^2 - (\mathbf{E}X)^2. \end{aligned}$$

□

**Definition 1.3.12.** Die Größen

$$m_k = \mathbf{E}X^k$$

für $k = 1, 2, \ldots$ heißen *$k$-te Momente* der Zufallsgröße $X$ und die Größen

$$\mu_k = \mathbf{E}(X - \mathbf{E}X)^k$$

für $k = 1, 2, \dots$ heißen *$k$-te zentrale Momente* der Zufallsgröße $X$.

Offensichtlich ist $m_1 = \mathbf{E}X$ und $\mu_2 = \mathbf{E}(X - \mathbf{E}X)^2 = \mathbf{D}^2X$. Weiterhin gilt $\mu_1 = \mathbf{E}(X - \mathbf{E}X) = \mathbf{E}X - \mathbf{E}X = 0$ und nach Satz 1.3.11 ist $\mu_2 = m_2 - m_1^2$. Häufig wird für den Erwartungswert $m_1 = \mathbf{E}X$ ebenfalls das Symbol $\mu$ verwendet.

**Definition 1.3.13.** Eine Folge von Zufallsgrößen $X_1, X_2, \dots, X_n$ heißt *vollständig unabhängig*, wenn sich der zufällige Charakter aller beteiligten Zufallsgrößen nicht beeinflusst.

**Hilfssatz 1.3.14.** *Seien $X_1, \dots, X_n$ $n$ Zufallsgrößen mit endlichen Erwartungswerten $\mathbf{E}X_i$ und endlichen Streuungen $\mathbf{D}^2X_i$. Dann gilt für beliebige reelle Zahlen $a_i$:*

$$\mathbf{E}(a_1X_1 + \dots + a_nX_n) = a_1\mathbf{E}X_1 + \dots + a_n\mathbf{E}X_n.$$

*Falls die Zufallsgrößen $X_1, \dots, X_n$ vollständig unabhängig sind, gilt außerdem:*

$$D^2(a_1X_1 + \dots + a_nX_n) = a_1^2\mathbf{D}^2X_1 + \dots + a_n^2\mathbf{D}^2X_n.$$

**Satz 1.3.15** (Tschebyscheff'sche Ungleichung)**.** *Für alle $\varepsilon > 0$ gilt:*

$$P(|X - \mathbf{E}X| > \varepsilon) < \frac{\mathbf{D}^2X}{\varepsilon^2}.$$

*Setzt man $k := \frac{\varepsilon}{\sqrt{\mathbf{D}^2X}} = \frac{\varepsilon}{\sigma}$, so erhält man:*

$$P(|X - \mathbf{E}X| > k\sigma) < \frac{1}{k^2}.$$

*Beweis (für diskrete Zufallsgrößen).* Sei $M := \{i : |x_i - \mathbf{E}X| > \varepsilon\}$. Dann gilt

$$\mathbf{D}^2X = \sum_i (x_i - \mathbf{E}X)^2 p_i \geq \sum_{i \in M} (x_i - \mathbf{E}X)^2 p_i > \varepsilon^2 \sum_{i \in M} p_i = \varepsilon^2 P(|X - \mathbf{E}X| > \varepsilon)$$

und somit ist

$$P(|X - \mathbf{E}X| > \varepsilon) < \frac{\mathbf{D}^2X}{\varepsilon^2}.$$

□

**Nachweis der schwachen Konvergenz der relativen Häufigkeit** Wir werden nun die Tschebyscheff'sche Ungleichung zum Nachweis der *schwachen Konvergenz* (Konvergenz im Sinne der Wahrscheinlichkeit) der relativen Häufigkeit eines Ereignisses gegen die Wahrscheinlichkeit dieses Ereignisses nutzen.

Dazu realisieren wir $n$ unabhängige Versuche zu einer Zufallssituation mit dem Wahrscheinlichkeitsraum $(\Omega, \mathfrak{A}, P)$. $H_n(A)$ sei die relative Häufigkeit des Eintretens eines Ereignisses $A \in \mathfrak{A}$. Zu zeigen ist nun, dass $H_n(A)$ in einem gewissen Sinn gegen die Wahrscheinlichkeit $p := P(A)$ strebt. Wir setzen

$$X_i = \begin{cases} 1, & \text{wenn } A \text{ im } i\text{-ten Versuch eintritt} \\ 0, & \text{wenn } A \text{ im } i\text{-ten Versuch nicht eintritt} \end{cases}$$

und erhalten somit eine Zufallsgröße

$$\bar{X}_n := H_n(A) = \frac{X_1 + \ldots + X_n}{n}.$$

Aus $P(X_i = 1) = p$ und $P(X_i = 0) = 1 - p$ folgt

$$\begin{aligned}
\mathbf{E}X_i &= 1p + 0(1-p) = p, \\
\mathbf{E}X_i^2 &= 1^2 p + 0^2(1-p) = p, \\
\mathbf{D}^2 X_i &= \mathbf{E}X_i^2 - (\mathbf{E}X_i)^2 = p - p^2 = p(1-p)
\end{aligned}$$

und nach Hilfssatz 1.3.14 gilt

$$\mathbf{E}\bar{X}_n = \frac{1}{n}\mathbf{E}X_1 + \cdots + \frac{1}{n}\mathbf{E}X_n = n \cdot \frac{1}{n}p = p = P(A)$$

und ebenfalls nach Hilfssatz 1.3.14 ist

$$\mathbf{D}^2\bar{X}_n = \frac{1}{n^2}\mathbf{D}^2 X_1 + \cdots + \frac{1}{n^2}\mathbf{D}^2 X_n = n \cdot \frac{1}{n^2}p(1-p) = \frac{p(1-p)}{n}.$$

Die TSCHEBYSCHEFF'sche Ungleichung liefert somit

$$P(|H_n(A) - P(A)| > \varepsilon) = P(|\bar{X}_n - \mathbf{E}\bar{X}_n| > \varepsilon) < \frac{D^2\bar{X}_n}{\varepsilon^2} = \frac{p(1-p)}{n\varepsilon^2}$$

und daraus erhalten wir

$$\lim_{n\to\infty} P(|H_n(A) - P(A)| > \varepsilon) \le \lim_{n\to\infty} \frac{p(1-p)}{n\varepsilon^2} = 0,$$

d.h. $H_n(A)$ strebt für $n \to \infty$ gegen $P(A)$ (Konvergenz im Sinne der Wahrscheinlichkeit):

$$H_n(A) \xrightarrow[n\to\infty]{\text{stoch}} P(A).$$

#### 1.3.2.2 Geometrische Verteilung

Als erste diskrete Wahrscheinlichkeitsverteilung wollen wir die recht einfache geometrische Verteilung und ihre Kenngrößen betrachten und an ihr die Verwendung der obigen Begriffe demonstrieren.

Als Standardbeispiel für die geometrische Verteilung dient ein Automat, der sofort anhält, wenn er ein fehlerhaftes Teil produziert hat, wobei die Qualität der einzelnen Teile von den anderen Teilen unabhängig ist. Wir verwenden die folgenden Bezeichnungen:

Wahrscheinlichkeit $p$ – Wahrscheinlichkeit, dass ein Teil fehlerhaft ist;

Zufallsgröße $X$ – Anzahl der produzierten fehlerfreien Teile;

Ereignis $A_i$ – $i$-tes produziertes Teil ist defekt.

Also ist $P(A_i) = p$ und $P(\overline{A_i}) = 1 - p$. Somit ergibt sich

$$\begin{aligned}
P(X = 0) &= P(A_1) = p, \\
P(X = 1) &= P(\overline{A_1} \cap A_2) = (1-p)p, \\
P(X = 2) &= P(\overline{A_1} \cap \overline{A_2} \cap A_3) = (1-p)^2 p, \\
P(X = 3) &= P(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \cap A_4 = (1-p)^3 p, \\
&\vdots
\end{aligned}$$

und allgemein erhalten wir als Wahrscheinlichkeitsfunktion

$$P(X = i) = p(1-p)^i \qquad (i = 0, 1, 2, \ldots).$$

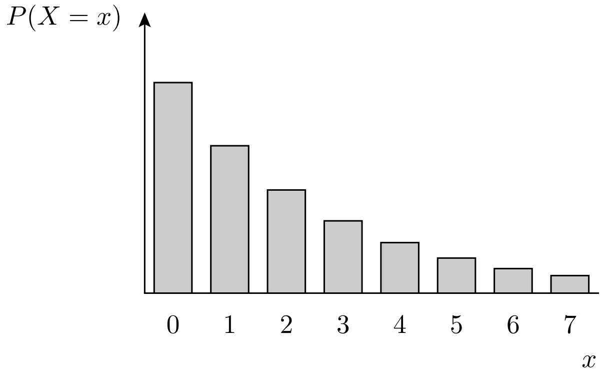
**Definition 1.3.16.** Ein diskrete Zufallsgröße $X$ mit der obigen Wahrscheinlichkeitsfunktion heißt *geometrisch verteilt* mit dem Parameter $p$.

Wie bereits am Anfang des Abschnitts über diskrete Zufallsgrößen erwähnt, muss für diskrete Zufallsgrößen zum einen $0 \leq P(X = x_i) \leq 1$ und zum anderen $\sum_i P(X = x_i) = 1$ gelten. Ersteres ist offensichtlich bei der geometrischen Verteilung erfüllt:

$$0 \leq p(1-p)^i \leq 1.$$

Und auch die zweite Eigenschaft gilt:

$$\sum_{i=0}^{\infty} p(1-p)^i = p \cdot \underbrace{\sum_{i=0}^{\infty} (1-p)^i}_{\substack{\text{geometrische} \\ \text{Reihe}}} = p \cdot \frac{1}{1-(1-p)} = \frac{p}{p} = 1.$$




**Beispiel.** Für das Zahlenbeispiel $p = 0{,}01$ ist die Wahrscheinlichkeit dafür gesucht, dass wenigstens 50 fehlerfreie Teile produziert werden. Wir erhalten

$$\begin{aligned}
P(X \geq 50) &= \sum_{i=50}^{\infty} p(1-p)^i = p(1-p)^{50} \sum_{i=50}^{\infty} (1-p)^{i-50} = p(1-p)^{50} \sum_{j=0}^{\infty} (1-p)^j \\
&= p(1-p)^{50} \frac{1}{1-(1-p)} = (1-p)^{50} = 0{,}99^{50} = 0{,}605.
\end{aligned}$$

**Erwartungswert** Wir berechnen nun den Erwartungswert einer geometrisch verteilten Zufallsgröße:

$$\begin{aligned}
\mathbf{E}X &= \sum_{i=0}^{\infty} iP(X=i) = \sum_{i=0}^{\infty} ip(1-p)^i \\
&= p(1-p)\left(\underbrace{0(1-p)^{-1}}_{i=0} + \sum_{i=1}^{\infty} i(1-p)^{i-1}\right) \\
&= p(1-p)\frac{1}{(1-(1-p))^2} \qquad \text{(nach Formel (1.3.2))} \\
&= \frac{1-p}{p}.
\end{aligned}$$

**Beispiel.** Für $p = 0{,}01$ produziert die Maschine also im Mittel $\frac{0{,}99}{0{,}01} = 99$ fehlerfreie Teile.

Für die Berechnung der Varianz benötigen wir neben $\mathbf{E}X$ auch noch die Größe $\mathbf{E}X^2$:

$$\begin{aligned}
\mathbf{E}X^2 &= \sum_{i=0}^{\infty} i^2 p(1-p)^i = \sum_{i=0}^{\infty} (i^2-i)p(1-p)^i + \underbrace{\sum_{i=0}^{\infty} ip(1-p)^i}_{\mathbf{E}X} \\
&= p(1-p)^2\left(\underbrace{0(1-p)^{-2}}_{i=0} + \underbrace{0(1-p)^{-1}}_{i=1} + \sum_{i=2}^{\infty} i(i-1)(1-p)^{i-2}\right) + \mathbf{E}X \\
&= p(1-p)^2\frac{2}{p^3} + \mathbf{E}X \qquad \text{(nach Formel (1.3.3))} \\
&= \frac{2(1-p)^2}{p^2} + \frac{1-p}{p} = \frac{(p-1)(p-2)}{p^2}.
\end{aligned}$$

**Varianz** Die Varianz einer geometrisch verteilten Zufallsgröße ist somit:

$$\mathbf{D}^2X = \mathbf{E}X^2 - (\mathbf{E}X)^2 = \frac{(p-1)(p-2)}{p^2} - \frac{(1-p)^2}{p^2} = \frac{1-p}{p^2}.$$

#### 1.3.2.3 Binomialverteilung

**Beispiel.** Ein Automat erzeugt nacheinander Teile mit einer Ausschusswahrscheinlichkeit von jeweils $p = 0{,}01$. Die Produktionsqualität ist von Teil zu Teil unabhängig. Gesucht ist die Wahrscheinlichkeit, dass unter $n = 10$ kontrollierten Teilen genau ein fehlerhaftes Teil ist.

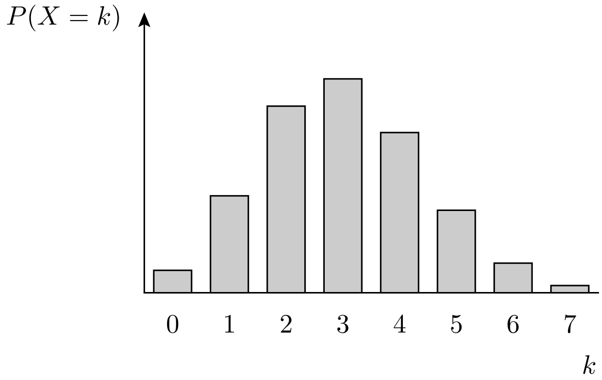
Die Zufallsgröße $X$ sei die Anzahl der Ausschussteile unter den $n$ kontrollierten Teilen. Sie kann also die Werte $0, 1, 2, \ldots, n$ annehmen. Mit $A_i$ bezeichnen wir das Ereignis, dass das $i$-te Teil fehlerhaft ist. Dann gilt:

$$\begin{aligned}
P(X=1) &= P\big((A_1 \cap \overline{A_2} \cap \cdots \cap \overline{A_{10}}) \cup (\overline{A_1} \cap A_2 \cap \cdots \cap \overline{A_{10}}) \cup \cdots \cup (\overline{A_1} \cap \overline{A_2} \cap \cdots \cap A_{10})\big) \\
&= 10p(1-p)^9 = 0{,}091.
\end{aligned}$$

**Verallgemeinerung** Sei $(\Omega, \mathfrak{A}, P)$ ein Wahrscheinlichkeitsraum zu einer Zufallssituation und $A \in \mathfrak{A}$ sei ein festes Ereignis. Wiederholen wir einen dieser Situation entsprechenden Versuch (unabhängig) $n$ mal und bezeichnen wir mit der Zufallsgröße $X$ die Anzahl der Versuche, bei denen $A$ eintritt, dann gilt mit $P(A) = p$:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \qquad (k = 0, 1, \ldots, n). \tag{1.3.4}$$

**Definition 1.3.17.** Eine Zufallsgröße $X$ mit der obigen Wahrscheinlichkeitsfunktion heißt *binomialverteilt* mit den Parametern $n$ (Zahl der Freiheitsgrade) und $p$ (Fehlerrate). Ist $X$ binomialverteilt, so schreiben wir $X \sim \mathbf{B}(n, p)$.




**Beispiel.** Ein idealer Würfel wird $n = 20$ mal geworfen. Gesucht ist die Wahrscheinlichkeit, dass mindestens zweimal eine 6 gewürfelt wird. Die Zufallsgröße $X$ ist die Anzahl der geworfenen Sechsen und $A$ sei das Ereignis, dass eine 6 gewürfelt wird. Dann ist $p = P(A) = \frac{1}{6}$ und $X \sim \mathbf{B}(20, \frac{1}{6})$. Wir erhalten somit als Ergebnis

$$\begin{aligned} P(X \geq 2) &= 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{20}{0}\left(\frac{1}{6}\right)^0\left(\frac{5}{6}\right)^{20} - \binom{20}{1}\left(\frac{1}{6}\right)^1\left(\frac{5}{6}\right)^{19} = 0{,}8696. \end{aligned}$$

**Erwartungswert** Der Erwartungswert für $X \sim \mathbf{B}(n, p)$ beträgt:

$$\begin{aligned} \mathbf{E}X &= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=1}^{n} \frac{n(n-1)(n-2)\cdots(n-k+1)}{(k-1)!} p^k (1-p)^{n-k} \\ &= np \sum_{k=1}^{n} \frac{(n-1)(n-2)\cdots((n-1)-(k-1)+1)}{(k-1)!} p^{k-1}(1-p)^{n-k} \\ &= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1}(1-p)^{n-k} \\ &= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{(n-1)-j} \qquad \text{(binomischer Satz)} \\ &= np(p + (1-p))^{n-1} = np \cdot 1^{n-1} \\ &= np. \end{aligned}$$

Um die Varianz zu berechnen, berechnen wir zunächst $\mathbf{E}X^2$ (analog zu $\mathbf{E}X$):

$$\begin{aligned}
\mathbf{E}X^2 &= \sum_{k=0}^{n} k^2 \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=2}^{n} k(k-1)\binom{n}{k} p^k (1-p)^{n-k} + \mathbf{E}X \\
&= n(1-n)p^2 \sum_{k=2}^{n} \binom{n-2}{k-2} p^{k-2}(1-p)^{n-k} + \mathbf{E}X \\
&= n(n-1)p^2 \sum_{j=0}^{n-2} \binom{n-2}{j} p^j (1-p)^{(n-2)-j} + \mathbf{E}X \\
&= n(n-1)p^2 (p + (1-p))^{n-2} + np = n(n-1)p^2 + np.
\end{aligned}$$

**Varianz** Die Varianz für $X \sim \mathbf{B}(n,p)$ beträgt:

$$\begin{aligned}
\mathbf{D}^2 X &= \mathbf{E}X^2 - (\mathbf{E}X)^2 = n(n-1)p^2 + np - n^2p^2 = np - np^2 \\
&= np(1-p).
\end{aligned}$$

**Beispiel.** Im obigen Würfelbeispiel ist $\mathbf{E}X = 20 \cdot \frac{1}{6} = 3{,}33$. Es werden also bei 20 Würfen im Mittel 3 bis 4 Sechsen gewürfelt. Die Varianz beträgt $\mathbf{D}^2 X = 20 \cdot \frac{1}{6} \cdot \frac{5}{6} = \frac{25}{9} = 2{,}78$. Nun interessiert uns noch, mit welcher Wahrscheinlichkeit die tatsächlich erreichte Anzahl von Sechsen um mehr als 3 vom Mittelwert abweicht. Dazu nutzen wir die TSCHEBYSCHEFF'sche Ungleichung:

$$P(|X - \mathbf{E}X| > 3) < \frac{1}{9}\mathbf{D}^2 X = \frac{1}{9} \cdot \frac{25}{9} = 0{,}31.$$

Somit tritt in durchschnittlich 31 % aller Fälle eine so große Abweichung vom Mittelwert auf. Wenn wir diese Wahrscheinlichkeit exakt berechnen, so erhalten wir:

$$P(X = 0) + P(X = 7) + P(X = 8) + \cdots + P(X = 20) = 0{,}063.$$

Es sind also in Wirklichkeit nur reichlich 6 % aller Fälle, bei denen eine Abweichung von mehr als 3 vom Mittelwert auftritt. Die TSCHEBYSCHEFF'sche Ungleichung liefert insofern eine recht grobe Abschätzung.

**Rekursionsformel** Für eine binomialverteilte Zufallsgröße $X \sim \mathbf{B}(n,p)$ gilt:

$$P(X = k+1) = \frac{n-k}{k+1} \cdot \frac{p}{1-p} \cdot P(X = k).$$

Somit ist $\sum_{k=m}^{n} P(X = k)$ von der unteren Indexgrenze beginnend leicht mit einem Taschenrechner auszuwerten.

#### 1.3.2.4 POISSON-Verteilung

Als Referenzmodell für die POISSON-Verteilung dient eine Telefonzentrale: Innerhalb eines Zeitintervalls der Länge $t$ kommen $X_t$ Anrufe (= Ereignisse) an und es gelten die sogenannten POISSON'schen Voraussetzungen:

- *Stationarität:* Die Wahrscheinlichkeit für das Eintreten einer bestimmten Anzahl von Ereignissen im betrachteten Zeitintervall hängt nur von der Intervalllänge und nicht von der Lage des Intervalls auf der Zeitachse ab.
- *Homogenität:* Die Ereignisfolge ist nachwirkungsfrei, d.h. die Anzahl von Ereignissen im Zeitintervall $[t_0, t_1]$ hat keinen Einfluss auf die Anzahl von Ereignissen in einem späteren Zeitintervall $[t_2, t_3]$, wobei $t_1 < t_2$.
- *Ordinarität:* Die Ereignisse treten für hinreichend kleine Zeitintervalle einzeln auf, d.h. für genügend kleine $\Delta t$ gilt entweder $X_{\Delta t} = 0$ oder $X_{\Delta t} = 1$. Zudem gilt $P(X_{\Delta t} = 1) = \mu \Delta t$ mit $0 < \mu < \infty$. Der Parameter $\mu$ heißt *Intensität*.

Unter diesen Voraussetzungen gilt:

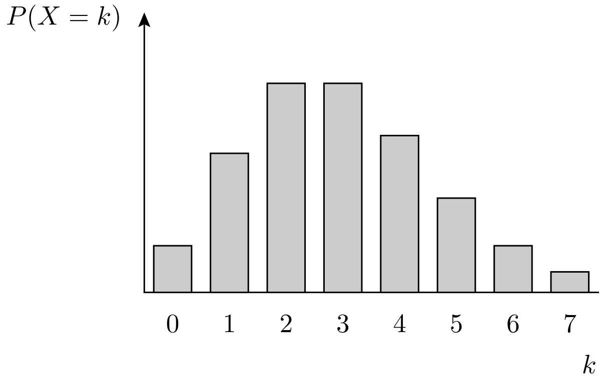
$$P(X_t = k) = \frac{(\mu t)^k}{k!} e^{-\mu t} \qquad (k = 0, 1, 2, \ldots). \tag{1.3.5}$$

Diese Formel erhalten wir, indem wir die beschriebene Verteilung als Grenzfall der Binomialverteilung auffassen: Wir teilen das Zeitintervall der Länge $t$ in $n$ hinreichend kleine Teilintervalle mit Länge $\Delta t = \frac{t}{n}$ und betrachten den Fall $n \to \infty$. Für endliches $n$ ist $X_t$ binomialverteilt ($k$ Teilintervalle mit einem Anruf, $n - k$ Teilintervalle mit null Anrufen) und mit $p = P(X_{\Delta t} = 1) = \mu \Delta t$ erhalten wir aus Formel (1.3.4):

$$\begin{aligned}
P(X_t = k) &= \binom{n}{k} (\mu \Delta t)^k (1 - \mu \Delta t)^{n-k} = \binom{n}{k} \left(\frac{\mu t}{n}\right)^k \left(1 - \frac{\mu t}{n}\right)^{n-k} \\
&= \frac{n!}{k!(n-k)!} \left(\frac{\mu t}{n}\right)^k \left(1 - \frac{\mu t}{n}\right)^{n-k} \\
&= \frac{(\mu t)^k}{k!} \cdot \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n^k}}_{\to 1} \underbrace{\left(1 - \frac{\mu t}{n}\right)^n}_{\to e^{-\mu t}} \underbrace{\left(1 - \frac{\mu t}{n}\right)^{-k}}_{\to 1} \\
&\xrightarrow[n \to \infty]{} \frac{(\mu t)^k}{k!} e^{-\mu t}.
\end{aligned}$$

**Definition 1.3.18.** Eine Zufallsgröße $X_t$ mit der Wahrscheinlichkeitsfunktion (1.3.5) heißt POISSON-*verteilt* und wir schreiben $X_t \sim \pi_{\mu t}$. Oft wird $\lambda := \mu t$ gesetzt. Ist $X_t$ POISSON-verteilt mit dem Parameter $\lambda$, so schreiben wir $X_t \sim \pi_\lambda$.

Die Wahrscheinlichkeiten $\pi_\lambda(k) = P(X_t = k)$ findet man für übliche Parameterwerte $\lambda$ in Tabellen.




**Beispiel.** Wir betrachten die Anzahl der eingehenden Anrufe in einer Telefonzentrale. Wir rechnen in Minuten und setzen $\mu = \frac{1}{3}$, d.h. mit einer Wahrscheinlichkeit von $\frac{1}{3}$ kommt innerhalb einer Minute genau ein Anruf an. Gesucht ist die Wahrscheinlichkeit dafür, dass in einer Viertelstunde wenigstens 3 und höchstens 7 Anrufe ankommen. Unsere Zufallsgröße $X_t$ ist also die Anzahl der innerhalb von $t = 15$ Minuten eingehenden Anrufe und $\lambda = \mu t = 5$. Da $X_t \sim \pi_5$, erhalten wir:

$$P(3 \leq X_t \leq 7) = \sum_{k=3}^{7} \pi_5(k) = \sum_{k=3}^{7} \frac{(15\mu)^k}{k!} e^{-15\mu} = 0{,}742.$$

**Rekursionsformel** Für eine POISSON-verteilte Zufallsgröße $X_t \sim \pi_\lambda$ gilt:

$$P(X_t = k+1) = \frac{\lambda^{k+1}}{(k+1)!} e^{-\lambda} = \frac{\lambda}{k+1} \cdot \frac{\lambda^k}{k!} e^{-\lambda} = \frac{\lambda}{k+1} P(X_t = k).$$

Somit ist $\sum_{k=m}^{n} P(X_t = k)$ von der unteren Indexgrenze beginnend leicht mit einem Taschenrechner auszuwerten.

**Erwartungswert** Der Erwartungswert für $X_t \sim \pi_\lambda$ beträgt:

$$\mathbf{E}X_t = \sum_{k=0}^{\infty} k \frac{\lambda^k}{k!} e^{-\lambda} = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} = \lambda e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

Somit gibt der Parameter $\mu = \frac{\lambda}{t} = \frac{\mathbf{E}X_t}{t}$ die mittlere Ereignisanzahl pro Zeiteinheit an. Um die Varianz zu berechnen, benötigen wir noch $\mathbf{E}X_t^2$:

$$\begin{aligned}
\mathbf{E}X_t^2 &= \sum_{k=0}^{\infty} k^2 \frac{\lambda^k}{k!} e^{-\lambda} = \sum_{k=0}^{\infty} k(k-1) \frac{\lambda^k}{k!} e^{-\lambda} + \mathbf{E}X_t = \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!} + \mathbf{E}X_t \\
&= \lambda^2 e^{-\lambda} \sum_{j=0}^{\infty} \frac{\lambda^j}{j!} + \mathbf{E}X_t = \lambda^2 E^{-\lambda} e^{\lambda} + \mathbf{E}X_t = \lambda^2 + \lambda.
\end{aligned}$$

**Varianz** Die Varianz für $X_t \sim \pi_\lambda$ beträgt:

$$\mathbf{D}^2 X_t = \mathbf{E}X_t^2 - (\mathbf{E}X_t)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda.$$

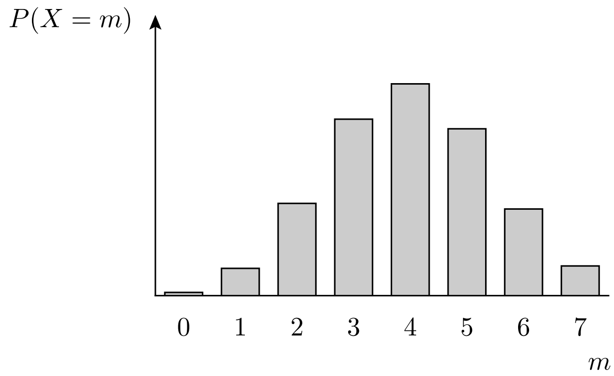
#### 1.3.2.5 Hypergeometrische Verteilung

Als Referenzmodell dient die bereits bekannte Urne mit $N$ Kugeln, von denen $M$ Kugeln schwarz und $N - M$ Kugeln weiß sind. Wir ziehen ohne Zurücklegen $n$ Kugeln, wobei unsere Zufallsgröße $X$ die Anzahl der entnommenen schwarzen Kugeln ist. Dann gilt:

$$P(X = m) = \frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}},$$

wobei $\max(0, n - (N - M)) \leq m \leq \min(n, M)$ ist.

**Definition 1.3.19.** Eine Zufallgröße $X$ mit der obigen Wahrscheinlichkeitsfunktion heißt *hypergeometrisch verteilt* und wir schreiben $X \sim \mathbf{H}(n, N, M)$.




**Erwartungswert** Der Erwartungswert für $X \sim \mathbf{H}(n, N, M)$ beträgt:

$$\mathbf{E}X = n\frac{M}{N}.$$

**Varianz** Die Varianz für $X \sim \mathbf{H}(n, N, M)$ beträgt:

$$\mathbf{D}^2X = n\frac{M}{N}\left(1 - \frac{M}{N}\right)\frac{N-n}{N-1}.$$

### 1.3.3 Stetige Zufallsgrößen

Im Abschnitt über Wahrscheinlichkeitsräume haben wir bereits die Brenndauer einer Glühlampe und die Reichweite eines Fahrzeugs bei begrenztem Treibstoffvorrat als Beispiele für stetige Zufallsgrößen betrachtet. Da stetige Zufallsgrößen überabzählbar unendlich viele Werte besitzen und somit deren Werte ganze Intervalle der reellen Achse ausfüllen können, ist es nicht mehr möglich, die Wahrscheinlichkeit für jeden einzelnen Wert in einer Wahrscheinlichkeitsfunktion auszudrücken. Jedoch kann man mit Hilfe sogenannter Dichtefunktionen die Verteilung der Wahrscheinlichkeitsmasse auf der reellen Achse angeben und so die Wahrscheinlichkeit dafür charakterisieren, dass der Wert der Zufallsgröße in einem gegebenen Intervall liegt.