

# Primer WaSt2

## 1. Grundlagen der Wahrscheinlichkeitstheorie

In diesem Kapitel werden die statistischen Grundlagen, die für die Behandlung von stochastischen Prozessen benötigt werden, kurz aufgeführt. Für eine detailliertere Behandlung sei auf die einschlägigen Statistik-Lehrbücher verwiesen.

### 1.1. Zufallsvariablen

Grundlegend für die Beschreibung von stochastischen Prozessen ist der Begriff der Zufallsvariablen. Wir betrachten zunächst nur univariate Zufallsvariablen.

Eine (univariate) Zufallsvariable ist eine Funktion  $X: \Omega \rightarrow \mathbb{R}$ , wobei ein Wahrscheinlichkeitsraum  $(\Omega, E, P)$  gegeben ist ( $\Omega$  = Menge der experimentellen Resultate (*sample space*),  $E$  = Menge der Ereignisse, Ereignis = Teilmenge von  $\Omega$ ,  $P$  = Wahrscheinlichkeitsmass auf  $E$ ). Eine Realisierung von  $X$  wird mit  $x$  bezeichnet. Also ist  $X$  eine Zufallsvariable, aber  $x$  ist einfach eine Zahl.

Der Wertemenge  $S$  einer Zufallsvariablen  $X$  (die man auch *Zustandsraum* nennt) ist gegeben durch  $S=X(\Omega)$ . Sie bezeichnet die Menge aller möglichen experimentellen Resultate und ist eine Teilmenge von  $\mathbb{R}$ .

Durch eine Zufallsvariable wird also jedem möglichen Ausgang eines Zufallsexperiments, d. h. jedem Element von  $\Omega$ , eine Zahl zugeordnet.

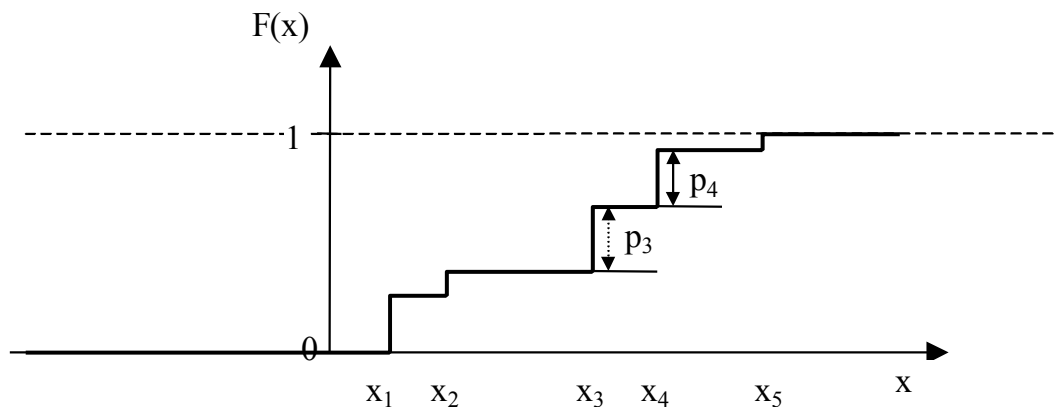
**Beispiel:** Wir betrachten das gleichzeitige Werfen von zwei Münzen. Die Menge  $\Omega$  der experimentellen Resultate ist  $\Omega=\{\text{Zahl-Zahl, Zahl-Kopf, Kopf-Zahl, Kopf-Kopf}\}$ . Eine Zufallsvariable  $X$  kann nun definiert werden durch die Zuordnung

Zahl-Zahl	$\rightarrow$	0
Zahl-Kopf	$\rightarrow$	1
Kopf-Zahl	$\rightarrow$	2
Kopf-Kopf	$\rightarrow$	3

Jedes Elementarereignis wird damit durch eine Zahl repräsentiert. Ereignisse werden Mengen von Zahlen (diskrete Mengen oder Intervalle) zugeordnet. Z. B. wird das Ereignis  $e$ ="gleiche Münzen"= $\{\text{Zahl-Zahl, Kopf-Kopf}\}$  der Menge  $\{0,3\}$  zugeordnet. Eine Realisierung einer Zufallsvariable ist immer eine Zahl.

Jedes Element  $e$  der Menge  $E$  der interessierenden Ereignisse ist dann äquivalent zu einer Menge von Zahlen. Dies kann eine abzählbare Menge sein, oder eine überabzählbare, etwa ein Intervall. Aus Konsistenzgründen muss man verlangen, dass die Menge

$$\{\omega \in \Omega \mid X(\omega) \leq x\}$$



**Fig. 1.1:** Verteilungsfunktion  $F(x)$  für eine diskrete Zufallsvariable.

ein Ereignis ist, d. h. zur Menge  $E$  gehört. Diese Menge (bzw. dieses Ereignis) wird auch kurz durch die Schreibweise  $[X \leq x]$  bezeichnet.

Eine Zufallsvariable ist charakterisiert durch ihre *kumulative Verteilungsfunktion*, oder kurz *Verteilungsfunktion*. Sie ist gegeben durch

$$F(x) := P[X \leq x] \quad (1.1)$$

$F(x)$  ist monoton steigend mit  $0 \leq F(x) \leq 1$ , mit  $F(-\infty)=0$ , und  $F(\infty)=1$ .

### 1.1.1. Diskrete Zufallsvariablen

Für einen diskreten Zustandsraum  $S=\{x_1, x_2, \dots\}$  ist die Wahrscheinlichkeit  $p_k$  eines Wertes  $x_k$  ( $k \geq 0$ ) definiert durch

$$p_k = P[X = x_k] \quad (1.2)$$

Im Englischen heisst diese Funktion die *probability mass function*. Die Grösse  $p_k$  ist einheitenlos; sie gibt direkt eine Wahrscheinlichkeit an.

**Beispiel:** Wir betrachten einen idealen Würfel. Der Zustandsraum ist gegeben durch  $S=\{1, 2, 3, 4, 5, 6\}$ . Die Wahrscheinlichkeiten  $p_k$ ,  $k=1, \dots, 6$  sind gegeben durch

$$p_k = \frac{1}{6}, \quad k=1, \dots, 6.$$

Es gilt für die Verteilungsfunktion

$$F(x) = \sum_{k: x_k \leq x} p_k \quad (1.3)$$

Die Verteilungsfunktion hat also einen stufigen Verlauf. Die Sprünge treten genau an den Stellen  $x_k$  auf (siehe Fig. 1.1).

Es gilt:

$$\begin{aligned}
 p_k &\geq 0, \quad k \geq 0 \\
 \sum_k p_k &= 1
 \end{aligned}
 \tag{1.4}$$

Die Summation läuft dabei über alle  $k$ , was eine unendliche Anzahl sein kann.

### Beispiele:

- Münzwerfen.  $\Omega=\{0,1\}$ , wobei Kopf der 0 entspreche, und Zahl der 1.

$$p_0=p_1=0.5.$$

- Werfen von zwei Münzen.  $\Omega=\{0,1,2,3\}$ . Dabei entspricht 0: Kopf-Kopf, 1: Kopf-Zahl, 2: Zahl-Kopf, 3: Zahl-Zahl.

$$P_k = 0.25, k=0,1,2,3.$$

- Werfen eines Würfels:  $p_k=1/6, k=1,\dots,6$ .

### 1.1.2. Stetige Zufallsvariablen

Stetige Zufallsvariablen haben einen Wertebereich, der nicht durch einzelne diskrete Werte, sondern durch ein Intervall gekennzeichnet ist.

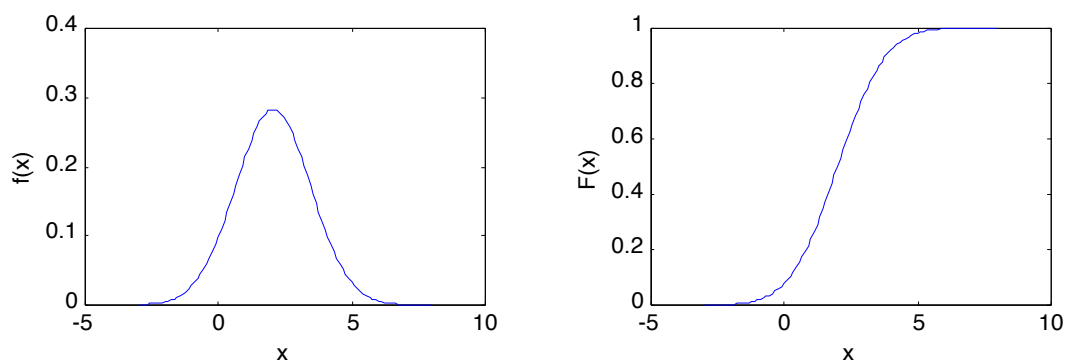
**Definition (stetige Zufallsvariable):** Eine Zufallsvariable mit Verteilungsfunktion  $F$  heisst *stetig*, wenn es eine stetige Funktion  $f$  gibt, sodass

$$F(x) = \int_{-\infty}^x f(u) du$$

für alle  $x \in (-\infty, \infty)$ .

Die Funktion  $f(x)$  heisst *Wahrscheinlichkeitsdichte* (im Englischen *probability density function pdf*). Es gilt:

$$f(x) = \frac{\partial}{\partial x} F(x) \tag{1.5}$$



**Fig. 1.2:** Verteilungsfunktion  $F(x)$  und Wahrscheinlichkeitsdichte  $f(x)$  für eine stetige Zufallsvariable.

In Fig. 1.2 ist der Zusammenhang von  $f(x)$  und  $F(x)$  schematisch gezeigt.

Für eine Wahrscheinlichkeitsdichte  $f$  gilt immer:

$$f(x) \geq 0 \quad \text{für alle } x$$

$$\int_{x=-\infty}^{\infty} f(x) dx = 1 \quad (1.6)$$

Ist man an der Wahrscheinlichkeit interessiert, dass  $X$  im Intervall  $a, b$  liegt, so kann man dies durch Integration wie folgt bestimmen:

$$P(a \leq X \leq b) = \int_a^b f(x) dx = F(b) - F(a)$$

Wenn man streng nach der mathematischen Definition geht, hat eine Zufallsvariable keine physikalische Einheit. Andererseits repräsentieren Zufallsvariablen oft physikalische Größen. Z. B. kann  $X$  die (zufällige) Lebensdauer eines Produktes sein, was die Einheit einer Zeit hat. Man kann deshalb oft "Zufallsvariablen mit Einheiten" definieren.

Sei  $[X]$  die Einheit von  $X$ . Die Einheit der Dichte ist dann  $1/[X]$ , d. h. wenn  $X$  die Einheit  $s$  hat, dann hat die Dichte  $f(x)$  die Einheit  $1/s$ . Für kleine  $dx$  gibt die Grösse  $f(x) \cdot dx$  die Wahrscheinlichkeit an, dass  $X$  im Bereich  $[x, x+dx]$  liegt. Diese Grösse ist dimensionslos (wie es sich für eine Wahrscheinlichkeit auch gehört).

**Beispiel: Gleichverteilte Zufallszahl:** Wir betrachten eine Zufallszahl  $X$ , die eine Länge bezeichne, also die Einheit  $m$  hat. Diese Zufallszahl sei gleichverteilt im Intervall  $[0, 2m]$ , d.h.  $f(x) = \text{const.}$  Aus der Normierungsbedingung (1.6) folgt zwingend

$$f(x) = \frac{1}{2m}$$

Denn damit ergibt sich

$$\int_{x=0}^{2m} f(x) dx = 1$$

### 1.1.3. Funktionen von Zufallsvariablen

Sei  $Y = g(X)$  eine Funktion einer Zufallsvariablen  $X$ . Es könnte z.B. sein:

$$Y = aX + b$$

$$Y = X^2$$

Dann ist  $Y$  selbst wieder eine Zufallsvariable. Wie kann die Verteilungsfunktion von  $Y$  berechnet werden? In der Regel ist die Berechnung von  $F_Y(Y)$  nicht einfach. Wir werden das Vorgehen anhand einiger Beispiele erläutern.

**Beispiel (Lineare Transformation):** Sei  $X$  eine Zufallsvariable mit Verteilungsfunktion  $F_X(\cdot)$ . Sei  $Y = aX + b$  mit Konstanten  $a$  und  $b$ . Die Verteilungsfunktion von  $Y$  ist dann

$$\begin{aligned}
F_Y(y) &= P[Y \leq y] \\
&= P[aX + b \leq y] \\
&= P\left[X \leq \frac{y-b}{a}\right] \\
&= F_X\left(\frac{y-b}{a}\right)
\end{aligned} \tag{1.7}$$

Wenn  $X$  stetig ist mit der Dichtefunktion  $f_X$ , dann ist  $Y$  auch stetig, hat also eine Dichtefunktion  $f_Y$ , die durch Differentiation von  $F_Y$  berechnet werden kann:

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right) \tag{1.8}$$

**Beispiel (Quadrierung):** Sei  $X$  eine stetige Zufallsvariable, und  $Y=X^2$ . Die Verteilungsfunktion von  $Y$  ergibt sich aus

$$\begin{aligned}
F_Y(y) &= P[Y \leq y] \\
&= P[X^2 \leq y] \\
&= P[-\sqrt{y} \leq X \leq \sqrt{y}] \\
&= F_X(\sqrt{y}) - F_X(-\sqrt{y})
\end{aligned} \tag{1.9}$$

Die Dichtefunktion kann durch Differenzieren gewonnen werden:

$$f_Y(y) = \frac{1}{2\sqrt{y}} (f_X(\sqrt{y}) + f_X(-\sqrt{y}))$$

## 1.2. Erwartungswert und Varianz

### 1.2.1. Erwartungswert

Das Konzept der Erwartungswerts ist eines der grundlegenden Konzepte der Wahrscheinlichkeitstheorie. Intuitiv ist der Erwartungswert der Wert, "dessen Realisierung man im Mittel erwartet", auch wenn diese Erklärung sehr unpräzise ist.

**Definition (Erwartungswert für diskrete Zufallsvariable):** Der *Erwartungswert* einer diskreten Zufallsvariable ist definiert durch

$$E(X) = \sum_k x_k P(X = x_k) = \sum_k x_k p_k \tag{1.10}$$

Der Erwartungswert ist eine gewichtete Summe aller möglichen Realisierungen. Die Gewichtung ist umso stärker, je höher die Wahrscheinlichkeit der entsprechenden Realisierung ist.

Es ist zu beachten, dass  $E(X)$  nicht unbedingt ein Element des Zustandsraums der Zufallsvariablen  $X$  ist.

**Beispiel:** Wir werfen eine ideale Münze. Wenn wir Zahl werfen, gewinnen wir 1 Franken. Bei Kopf gewinnen wir 2 Franken. Das Spiel kostet jedoch eine Gebühr  $G$ . Was ist die höchste Gebühr, die wir zu zahlen bereit sind?

Die Zufallsvariable  $X$  bezeichne den Gewinn beim Werfen der Münze. Die Menge der möglichen Realisierungen (Zustandsraum) ist  $S=\{1,2\}$ . Es gilt:  $p_1=0.5$ ,  $p_2=0.5$ . Der Erwartungswert ist gegeben durch

$$E(X) = \sum_{k=1}^2 p_k x_k = 0.5 \cdot 1 + 0.5 \cdot 2 = 1.5$$

Dies entspricht dem erwarteten (mittleren) Gewinn pro Spiel. Die höchste Gebühr, die wir zu zahlen bereit sind, ist also 1.5 Franken.

**Beispiel:** Wir werfen zwei Würfel. Sei  $X$  die Summe der beiden geworfenen Würfelzahlen. Der Zustandsraum ist  $S=\{2,3,\dots,12\}$ . Der Erwartungswert ist

$$\begin{aligned} E(X) &= \sum_{k=2}^{12} p_k x_k \\ &= \frac{1}{36} \cdot 2 + \frac{2}{36} \cdot 3 + \frac{3}{36} \cdot 4 + \frac{4}{36} \cdot 5 + \frac{5}{36} \cdot 6 + \\ &\quad + \frac{6}{36} \cdot 7 + \frac{5}{36} \cdot 8 + \frac{4}{36} \cdot 9 + \frac{3}{36} \cdot 10 + \frac{2}{36} \cdot 11 + \frac{1}{36} \cdot 12 \\ &= 7 \end{aligned}$$

**Definition (Erwartungswert für stetige Zufallsvariable):** Der Erwartungswert einer stetigen Zufallsvariable ist definiert durch

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \quad (1.11)$$

Dies ist die Übertragung von Gl. (1.10) auf stetige Zufallsvariablen.

### 1.2.2. Erwartungswert von Funktionen von Zufallsvariablen

Oft betrachtet man Funktion  $Y=g(X)$  einer Zufallsvariablen  $X$ . Wie wir oben gesehen haben, ist dies wieder eine Zufallsvariable. Wie ist deren Erwartungswert?

Im Prinzip kann man die Verteilungsfunktion von  $Y$  berechnen, und daraus dann den Erwartungswert. Dies ist im Allgemeinen aber schwierig. Für die Erwartungswertbildung kann man einen einfacheren Weg gehen. Dies ist die Aussage des nächsten Satzes:

**Satz 1.1 (Erwartungswert von Funktion einer Zufallsvariablen):**

Für diskrete Zufallszahlen gilt:

$$E(Y) = \sum_k p_X(x_k) g(x_k)$$

Für stetige Zufallsvariablen gilt:

$$E(Y) = \int_{-\infty}^{\infty} f_X(x) g(x) dx$$

Beweis: ---.

Man kann sich Satz 1.1 für den diskreten Fall mit folgender Argumentation klarmachen:  $X$  kann die Werte  $x_k$  annehmen. Dann nimmt  $Y$  die Werte  $g(x_k)$  an. Die Wahrscheinlichkeit für ein  $x_k$  ist gerade  $p_k$ . Dies ist unabhängig davon, ob man nun gerade  $X=x_k$  oder  $Y=g(x_k)$  betrachtet. Damit kann man den Erwartungswert von  $E(Y)$  wie oben ausrechnen.

Den stetigen Fall kann man sich als Grenzfall der diskreten Version für immer feinere Diskretisierungen vorstellen.

**Beispiel:** Sei  $f(x)$  die Wahrscheinlichkeitsdichte einer stetigen Zufallszahl  $X$ . Der Erwartungswert von  $X$  ist dann gegeben durch

$$E(X) = \int x f(x) dx$$

Der Erwartungswert von  $X^2$  ist gegeben durch

$$E(X^2) = \int x^2 f(x) dx$$

Analog ist der Erwartungswert von  $X^n$  gegeben durch

$$E(X^n) = \int x^n f(x) dx$$

Den Erwartungswert  $E(X^n)$  nennt man auch das *n-te Moment* von  $X$ .

Eine Ausformung von Satz 1.1 ist der folgende Satz:

**Satz 1.2 (Linearität des Erwartungswerts):** Die Erwartungswertbildung ist eine lineare Operation, d.h. es gilt für beliebige Zufallszahlen  $X, Y$  und Konstanten  $a, b$ :

$$E(aX + b) = aE(X) + b$$

$$E(X + Y) = E(X) + E(Y)$$

### 1.2.3. Varianz

Neben dem Erwartungswert ist die Varianz eine besonders wichtige Grösse einer Zufallsvariablen.

**Definition (Varianz):** Die Varianz einer Zufallsvariablen  $X$  ist

$$Var(X) = E\left((X - E(X))^2\right)$$

Die Varianz bezeichnet also den Erwartungswert des quadrierten Abstandes einer Zufallsvariablen von seinem Mittelwert. Die Wurzel aus der Varianz heisst *Standardabweichung* und wird mit  $\sigma$  bezeichnet:

$$\sigma_X = \sqrt{Var(X)}$$

## 1.3. Multivariate Zufallsvariablen

Genau, wie man sich bei zwei Ereignissen  $A$  und  $B$  fragen kann: Wie wahrscheinlich ist es, dass  $A$  und  $B$  auftritt, kann man sich fragen wie wahrscheinlich ist es, dass eine Zufallsvari-

able  $X_1$  einen Werte  $x_1$  und eine andere  $X_2$  einen Wert  $x_2$  annimmt.<sup>1</sup> Wir können  $X_1$  und  $X_2$  zu einem Vektor kombinieren. Dies führt uns auf folgende Definition:

**Definition (Multivariate Zufallszahl):** Seien  $X_i, i=1,2,\dots,n$ , Zufallsvariablen mit Zustandsraum  $S_i$ . Der Vektor  $X=(X_1,X_2,\dots,X_n)$  heisst *multivariate Zufallsvariable*. Die Menge aller Werte, die  $X$  annehmen kann, heisst *Zustandsraum* von  $X$ . Er ist gegeben durch  $S=S_1\times S_2\times\dots\times S_n$ .

Eine multivariate Zufallsvariable mit zwei zugrunde liegenden Variablen heisst *bivariat*.

**Beispiel 1:** wir betrachten ein Zufallsexperiment, bei ein Würfel 4 mal geworfen wird. Die folgenden Zufallsvariablen werden definiert:

$X_1$  = Augenzahl des ersten Wurfs

$X_2$  = Augenzahl des zweiten Wurfs

$X_3$  = Augenzahl des dritten Wurfs

$X_4$  = Augenzahl des vierten Wurfs

Dann ist  $(X_1,X_2,X_3,X_4)$  eine multivariate Zufallszahl.

**Beispiel 2:** wir betrachten ein Zufallsexperiment, bei dem zwei Würfel geworfen werden. Die folgenden Zufallsvariablen werden definiert:

$X_1$  = Maximum der zwei Zahlen

$X_2$  = Minimum der zwei gewürfelten Zahlen

$X_3$  = Summe der beiden Zahlen

$X_4$  = Betrag der Differenz der beiden Zahlen

Dann ist  $(X_1,X_2,X_3,X_4)$  eine multivariate Zufallsvariable,  $(X_3,X_4)$  ist eine bivariate Zufallsvariable. Man beachte, dass im zweiten Beispiel  $S_1=\{1,2,\dots,6\}$  ist und  $S_3=\{1,2,\dots,12\}$ . Meist betrachten man allerdings mehrdimensionale Zufallsvariablen bei denen, wie im Beispiel 1,  $S_1=S_2=\dots=S_n$  gilt.

**Definition (Multivariate Verteilungsfunktion):** Die Funktion

$$F(x_1,x_2,\dots,x_n) = P[X_1 \leq x_1, \dots, X_n \leq x_n]$$

heisst *multivariate (kumulative) Verteilungsfunktion* der Zufallsvariablen Vektor  $X=(X_1,X_2,\dots,X_n)$ .

Die Konzepte der Wahrscheinlichkeitsdichte für stetige Zufallsvariablen bzw. der Wahrscheinlichkeiten für diskrete Zufallsvariablen übertragen sich analog. Man definiert eine multivariate Wahrscheinlichkeitsdichte (für stetige Zufallsvariablen)

$$f(X_1 = x_1, X_2 = x_2, \dots, X_n = x_n) = f(x_1, x_2, \dots, x_n)$$

bzw. eine multivariate Wahrscheinlichkeit (für diskrete multivariate Zufallsvariablen)

---

<sup>1</sup> Im Falle von stetigen ZV werden die Wahrscheinlichkeiten natürlich zu Dichten.



$$p(x_1, x_2, \dots, x_n) \quad x_i \in S_i$$

wobei  $S_i$  der Zustandsraum für die Zufallsvariable  $X_i$  ist ( $i=1, \dots, n$ ).

### 1.3.1. Abhängigkeit von Zufallszahlen

Wenn man multivariate Zufallsvariablen betrachtet, stösst man automatisch auf das Thema der Abhängigkeit bzw. Unabhängigkeit von Zufallsvariablen.

**Definition (Unabhängige Zufallszahlen):** Zufallsvariablen  $X_1, X_2, \dots, X_n$  heissen *unabhängig*, wenn gilt

$$F(x_1, x_2, \dots, x_n) = F_{X_1}(x_1)F_{X_2}(x_2) \dots F_{X_n}(x_n)$$

oder (im diskreten Fall)

$$p(x_1, x_2, \dots, x_n) = p_{X_1}(x_1)p_{X_2}(x_2) \dots p_{X_n}(x_n)$$

oder (im stetigen Fall)

$$f(x_1, x_2, \dots, x_n) = f_{X_1}(x_1)f_{X_2}(x_2) \dots f_{X_n}(x_n).$$

Im Fall von unabhängigen Zufallszahlen faktorisiert also die gemeinsame Verteilungsfunktion wie auch die gemeinsame Wahrscheinlichkeitsdichte.

In vielen Anwendungen werden wir Zufallsvariablen begegnen, die unabhängig sind und die gleiche Verteilungsfunktion besitzen. Solche Zufallsvariablen nennt man unabhängig gleichverteilte Zufallszahlen oder *iid-Variablen* (nach dem Englischen *independent and identically distributed*). Wir haben dann:

$$f(x_1, x_2, \dots, x_n) = f(x_1)f(x_2) \dots f(x_n)$$

### 1.3.2. Bedingte Wahrscheinlichkeiten

Die Wahrscheinlichkeit eines zufälligen Ereignisses A kann sich ändern, wenn bereits bekannt ist, dass ein anderes Ereignis B eingetreten ist. Die Wahrscheinlichkeit von A unter der Bedingung, dass schon B eingetreten ist, wird mit

$$P(A|B)$$

bezeichnet und heisst “bedingte Wahrscheinlichkeit”.

Es gilt:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{für } P(B) \neq 0.$$

Auch das Konzept der bedingten Wahrscheinlichkeit, überträgt sich auch auf Zufallsvariablen und deren Dichten. So ist  $f(x_1 | x_2, \dots, x_n)$  die Wahrscheinlichkeitsdichte, dass die ZV  $X_1$  den Wert  $x_1$  annimmt, wenn die anderen Zufallsvariablen die Werte  $x_2, \dots, x_n$  angenommen haben. Es gilt

$$f(x_1 | x_2, \dots, x_n) = f(x_1, x_2, \dots, x_n) / f(x_2, \dots, x_n)$$

Im Falle iid verteilter ZV gilt:

$$f(x_1 | x_2, \dots, x_n) = f(x_1)f(x_2)\dots f(x_n) / f(x_2)\dots f(x_n) = f(x_1)$$

Die Kenntnisse der Werte  $x_2$  bis  $x_n$  ändert nichts an der Wahrscheinlichkeitsverteilung für  $x_1$ . Nehmen wir an, dass alle ZV denselben Zustandsraum haben, so gilt:

$$p(x_1) = \sum_j p(x_1 | x_j) p(x_j)$$

Dies bezeichnet man, als den Satz der totalen Wahrscheinlichkeit. Im kontinuierlichen Fall wird aus der Summe ein Integral und man erhält.

$$f(x) = \int f(x | x') f(x') dx'$$

### 1.3.3. Erwartungswerte multivariater Zufallsvariablen

Häufig treten Summen oder Produkte verschiedener (nicht unbedingt unabhängiger) Zufallsvariablen auf. Der folgende Satz gibt zwei Regeln zur Berechnung der Erwartungswerte.

**Satz 1.3 (Summe und Produkt von Zufallsvariablen):** Seien  $X_1, X_2, \dots, X_n$  Zufallsvariablen. Dann gilt:

$$E(X_1 + X_2 + \dots + X_n) = E(X_1) + E(X_2) + \dots + E(X_n)$$

Falls die Zufallsvariablen ausserdem unabhängig sind, gilt:

$$E(X_1 \cdot X_2 \cdot \dots \cdot X_n) = E(X_1) \cdot E(X_2) \cdot \dots \cdot E(X_n).$$

Man beachte, dass der erste Teil des Satzes (Erwartungswert einer Summe) auch für abhängige Zufallszahlen gilt! Dies ist nicht mehr der Fall für das Produkt.

### 1.3.4. Randverteilungen

**Definition (Randverteilung):** Die Funktion

$$F_{X_i}(x_i) = P[X_i \leq x_i]$$

heisst *Randverteilung* von  $X_i$ .

Die Randverteilung charakterisiert die Verteilung bzgl. der Variablen  $X_i$  ungeachtet der Werte der anderen Variablen  $X_j$  mit  $j \neq i$ .

Einer Randverteilung entspricht wiederum eine Dichtefunktion  $f_{X_i}(x_i)$  (im stetigen Fall) bzw. ein Wahrscheinlichkeitsvektor  $p_{X_i}(x_i)$  (im diskreten Fall), die durch Aufintegrieren bzw. Aufsummieren der übrigen Variablen entstehen. Zum Beispiel ist die Randverteilung von  $X_1$  gegeben durch

$$f_{X_1}(x) = \int_{x_2} \int_{x_3} \dots \int_{x_n} f(x, x_2, \dots, x_n) dx_2 dx_3 \dots dx_n$$

bzw.

$$p_{X_1}(x) = \sum_{x_2} \sum_{x_3} \dots \sum_{x_n} p(x, x_2, \dots, x_n).$$

## 1.4. Wichtige Zufallszahlen

### 1.4.1. Binomialverteilung $\text{Bin}(n, p)$

Die Binomialverteilung tritt auf, wenn ein Zufallsexperiment öfter wiederholt wird. Bei jedem Versuch möge ein bestimmtes Ereignis  $A$  eintreten können oder nicht. Die Wahrscheinlichkeit, dass  $A$  eintritt, sei mit  $p$  bezeichnet. Die einzelnen Versuche seien unabhängig voneinander. Die betrachtete Zufallsvariable  $X$  sei die Anzahl des Eintretens von  $A$  bei  $n$  Versuchen. Der Zustandsraum von  $X$  ist dann  $\{0, 1, \dots, n\}$ .

Es gilt:

$$P[X = k] \equiv \text{Bin}(n, p) = \binom{n}{k} p^k (1-p)^{n-k}$$

mit

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

$X$  nennt man dann binomialverteilt, oder  $X \sim \text{Bin}(n, p)$ .

Der Erwartungswert einer binomialverteilten Zufallsvariable ist  $np$ :  $X$  kann man sehen als eine Summe von  $n$  unabhängigen Zufallsvariablen  $X_1, \dots, X_n$ , wobei  $X_i$  jeweils die Werte 0 oder 1 annehmen kann, mit den dazugehörigen Wahrscheinlichkeiten  $(1-p)$  und  $p$ . Dann ist

$$E(X) = E(X_1 + X_2 + \dots + X_n) = E(X_1) + \dots + E(X_n)$$

Die Erwartungswerte der Variablen  $X_i$  ( $i=1, \dots, n$ ) sind gerade  $p$ , so dass man erhält:

$$E(X) = n \cdot p.$$

Die Varianz einer  $\text{Bin}(n, p)$ -verteilten Zufallszahl ist gegeben durch

$$\text{Var}(X) = n \cdot p \cdot (1-p).$$

### 1.4.2. Die Poissonverteilung $P(\lambda T)$

Die Poissonverteilung kann als Grenzfall der Binomialverteilung für sehr grosse  $n$  und kleine  $p$  gesehen werden. Sie tritt z.B. dann auf, wenn ein Prozess in einem festgelegten Zeitintervall  $[0, T]$  beobachtet wird, und die Anzahl bestimmter Ereignisse in diesem Zeitbereich gezählt wird. Ein Beispiel ist etwa die Anzahl von Ausfällen einer Maschine in einem bestimmten Zeitintervall, oder die Anzahl von Glühbirnen, die innerhalb eines festgelegten Zeitintervalls kaputtgehen. Die Zufallsvariable  $X$  bezeichnet also die Anzahl von Ereignissen. Der Zustandsraum ist diskret und besteht aus der Menge  $\{0, 1, 2, 3, \dots\}$ .

Die Poissonverteilung ist spezifiziert durch ein Zeitintervall  $T$  und eine *Rate*  $\lambda$ , wobei  $\lambda$  die mittlere Anzahl von Ereignissen pro Zeit kennzeichnet und die Einheit  $1/\text{Zeit}$  hat. Für die Anzahl von Ereignissen ist nur das Produkt  $\lambda T$  wichtig.

Die Wahrscheinlichkeit dafür, dass genau  $k$  Ereignisse gezählt werden, ist

$$P[X = k] \equiv P(\lambda T) = \frac{(\lambda T)^k}{k!} e^{-\lambda T}.$$

Der Erwartungswert und die Varianz der Poissonverteilung ist gegeben durch

$$E(X) = \lambda T$$

$$Var(X) = \lambda T.$$

Bemerkungen:

1. In der klassischen statistischen Literatur wird das Argument der Poissonverteilung (d.h. das Argument der Funktion  $P(\cdot)$ ) meist mit  $\lambda$  bezeichnet, wobei dies eine dimensionslose Zahl ist, und genau dem Produkt  $\lambda T$  von oben entspricht. In diesem Skript wird mit  $\lambda$  aber immer eine Rate bezeichnet, also eine Grösse der Einheit 1/Zeit. Das Argument der Poissonverteilung ist dann durch  $\lambda T$  bezeichnet, was wieder eine dimensionslose Grösse ist. Man lasse sich durch diesen Unterschied nicht verwirren.
2. Man kann die Poissonverteilung als Grenzfall der Binomialverteilung interpretieren. Dazu teilt man das betrachtete Zeitintervall  $T$  in  $n$  kleine Stücke  $dt$  mit  $T = n \cdot dt$ . Nun setzt man weiter voraus, dass in einem dieser kleinen Intervalle höchstens ein Ereignis stattfinden kann. Die Wahrscheinlichkeit dafür sei durch  $\lambda \cdot dt$  gegeben. Dann ist die Anzahl der Ereignisse auf dem ganzen Zeitintervall eine Zufallsvariable, die  $\text{Bin}(T/dt, \lambda \cdot dt)$ -verteilt ist. Man kann nun den Grenzfall immer kleinerer  $dt$  berechnen:  $dt \rightarrow 0$ , wobei aber immer  $n \cdot dt = T$  ist. Dann ergibt sich genau die oben angeführte Poissonverteilung  $P(\lambda T)$ .

### 1.4.3. Gleichverteilte Zufallszahlen $U(a,b)$

Eine gleichverteilte Zufallszahl  $U(a,b)$  auf dem Intervall  $[a,b]$  hat folgende Wahrscheinlichkeitsdichte:

$$f(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{1}{b-a} & \text{für } a < x < b \\ 0 & \text{für } x > b \end{cases} \quad (1.12)$$

Die Wahrscheinlichkeit für eine Realisierung im Intervall  $[x, x+h] \subseteq [a,b]$  ist

$$P[x < X \leq x+h] = \frac{h}{b-a} \quad (1.13)$$

unabhängig von  $x$ . Daher kommt die Bezeichnung “gleichverteilt”.

Der Erwartungswert der Gleichverteilung ist

$$E(X) = \frac{a+b}{2}.$$

### 1.4.4. Exponentialverteilte Zufallszahlen $\text{Exp}(\lambda)$

Ein wichtiger Typ von Zufallszahlen sind die exponentialverteilten Zufallszahlen  $\text{Exp}(\lambda)$ . Eine exponentialverteilte Zufallszahl hat die Wahrscheinlichkeitsdichte

$$f(x) = \begin{cases} 0 & \text{für } x < 0 \\ \lambda e^{-\lambda x} & \text{für } x \geq 0 \end{cases} \quad (1.14)$$

Die Verteilungsfunktion ergibt sich zu

$$F(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 - e^{-\lambda x} & \text{für } x \geq 0 \end{cases} \quad (1.15)$$

Die Einheit des Parameters  $\lambda$  ist  $1/[X]$ . Wenn also  $X$  eine Zeit bezeichnet mit Einheit  $s$ , dann hat  $\lambda$  die Einheit  $1/s$ . Dies sieht man z.B. daran, dass das Argument der Exponentialfunktion  $-\lambda x$  ist, dies muss aber dimensionslos sein. Die Wahrscheinlichkeitsdichte  $f(x)$  hat die Einheit  $1/[X]$ , wie es sein muss.

Der Erwartungswert der Exponentialverteilung ist

$$E(X) = \frac{1}{\lambda}$$

(Herleitung als Übungsaufgabe).

Exponentialverteilte Zufallszahlen werden oft verwendet, um Lebensdauern zu modellieren. Die Lebensdauer als stochastische Variable wird dann als exponentialverteilt angenommen. Der Parameter  $\lambda$  wird dann gesetzt als Kehrwert der mittleren Lebensdauer, denn dann ist  $E(X)$  gerade die mittlere Lebensdauer.

#### 1.4.5. Normalverteilte Zufallszahlen $N(\mu, \sigma^2)$

Eine besonders wichtige Verteilung ist die Normalverteilung oder Gauss-Verteilung. Sie ist gegeben durch

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Dabei kann  $x$  alle Werte von  $-\infty$  bis  $\infty$  annehmen. Eine solche Zufallszahl wird mit  $N(\mu, \sigma^2)$  bezeichnet. Die  $N(0,1)$ -Verteilung heisst *Standardnormalverteilung*.

Die Verteilungsfunktion als Integral über  $f(x)$  kann nicht geschlossen ausgedrückt werden. Sie ist aber tabelliert und in den meisten Programmpaketen als Funktion vorhanden.

.