

### 3 Darstellungen von Wahrscheinlichkeitsmassen

Die drei wichtigsten Darstellungsformen für W-Masse für die Modellierung und Untersuchung sind Zähldichten, Riemann-Dichten und Verteilungsfunktionen.

#### Begriffe

*Die drei Begriffe **Wahrscheinlichkeitsverteilung**, **Verteilungsfunktion** und **Wahrscheinlichkeitsdichte** stehen in engem Zusammenhang miteinander; (...)*

Eine **Wahrscheinlichkeitsverteilung** gibt an, mit welcher Wahrscheinlichkeit eine Zufallsgrösse  $X$  bestimmte Werte annimmt. Im Falle einer **diskreten** Zufallsgrösse kann man zum Beispiel die Wahrscheinlichkeit, dass  $X$  den Wert 5 annimmt, als  $P(X = 5)$  notieren. Im Falle einer **kontinuierlichen** Zufallsgrösse ist der Begriff der Wahrscheinlichkeitsverteilung dagegen problematisch, denn die Wahrscheinlichkeit, dass eine Zufallsgröße auf unendlich viele Stellen genau einen bestimmten reellen Wert annimmt, ist 0. Um eine kontinuierliche Wahrscheinlichkeitsverteilung zu bestimmen, fragt man nicht, mit welcher Wahrscheinlichkeit  $X$  einen bestimmten Wert annimmt, sondern mit welcher Wahrscheinlichkeit  $X$  in ein bestimmtes Intervall fällt. Zum Beispiel kann man  $P(5 \leq X < 6)$  angeben. Mit einem nach unten unendlichen Intervall erhält man eine kumulative Verteilungsfunktion; mit einem infinitesimalen Intervall eine Wahrscheinlichkeitsdichte. (...)

Eine **kumulative Verteilungsfunktion**  $F(x) = P(X \leq x)$  gibt die Wahrscheinlichkeit an, dass die Zufallsvariable  $X$  einen Wert kleiner oder gleich  $x$  annimmt. Man kann Verteilungsfunktionen sowohl für diskrete wie auch für kontinuierliche Wahrscheinlichkeitsverteilungen angeben. *In der Regel spricht man einfach von einer **Verteilungsfunktion**. Die explizite Kennzeichnung als **kumulativ** kann helfen, Verwechslungen mit der Wahrscheinlichkeitsdichte zu vermeiden.*

Eine **Wahrscheinlichkeitsdichte**  $f(x)$  wird üblicherweise nur im Falle kontinuierlicher Wahrscheinlichkeitsverteilungen verwendet. Man erhält sie als Ableitung der kumulativen Verteilungsfunktion,  $f(x) = dF(x)/dx$ . Durch Multiplikation von  $f(x)$  mit dem Differential  $dx$  erhält man die Wahrscheinlichkeit, dass  $X$  einen Wert aus einem infinitesimal schmalen Intervall annimmt,

$$dx f(x) = P(x \leq X < x+dx).$$

Durch Integration über beliebige Intervalle  $[a, b]$  erhält man daraus endliche Wahrscheinlichkeiten

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

[<http://de.wikipedia.org/wiki/Wahrscheinlichkeitsverteilung> (14. Dez. 04)]

### 3.1 Diskrete W-Masse und Zähl-dichten

Def: Zähl-dichte

Für  $f(\omega) = P(\{\omega\})$  heisst  $f(\omega)$  Zähl-dichte (Z-Dichte)

Eigenschaften:

- $f(\omega) \geq 0$
- $\sum_{\omega \in \Omega} f(\omega) = 1$
- $P(\{\omega\}) = f(\omega)$

Für jeden endlichen oder abzählbar unendlichen Merkmalraum  $\Omega$  können W-Masse durch entsprechende Z-Dichten angegeben werden. Die Werte der Z-Dichte dürfen nicht negativ sein und müssen die Summe 1 besitzen.

Def: Binomial-Z-Dichte

$$f(k) = b(n, p; k) := \binom{n}{k} p^k q^{n-k} \text{ über } \Omega = \{0, 1, \dots, n\}$$

Def: Binomial-Verteilung

$$B(n, p) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k}$$

Def: Geometrische Z-Dichte

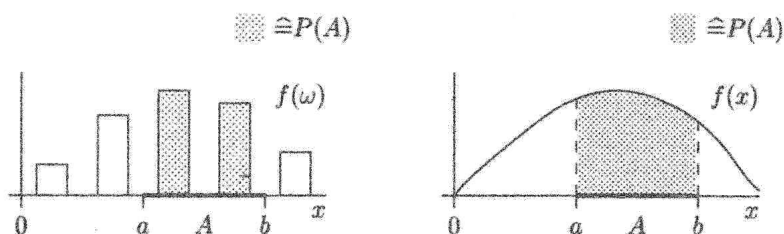
$$f(k) = (1-q)q^k \text{ über } \Omega = \{0, 1, \dots, n\}$$

### 3.2 Stetige W-Masse und Riemann-Dichten

Die wesentliche Idee zur Angabe der Wahrscheinlichkeit eines Ereignisses, z.B. eines Intervalls  $A=(a,b]$ , besteht darin, in Analogie zum Aufsummieren von Z-Dichten

$\left( P(\{A\}) = \sum_{\omega \in A} f(\omega) \right)$  stattdessen über eine gewisse Dichtefunktion  $f$  zu integrieren. Man

erhält dann die Darstellung  $P(A) = \int_a^b f(x) dx$ .



Def: Riemann-Dichte

Eine Riemann-integrierbare Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R} \text{ mit } f(x) \geq 0 \ (x \in \mathbb{R}) \text{ und } \int_{-\infty}^{+\infty} f(x) dx = 1$$

heisst Riemann-Dichte (R-Dichte) oder auch stetige Dichte. Jede R-Dichte definiert eindeutig ein W-Mass  $P$  über  $(\mathbb{R}, \mathcal{B})$  mit der Eigenschaft

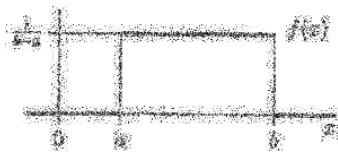
$$P((a,b]) = P([a,b]) = \int_a^b f(x) dx \ (a \leq b), \ P(\{a\}) = 0.$$

Wichtige **Beispiele** für R-Dichten:

Def: Rechteck-Verteilung

Ist  $f(x) = 1 / (b - a)$  für  $a < x < b$ ,  $a, b \in \mathbb{R}$ , und  $f(x) = 0$  sonst, dann ist  $f$  eine R-Dichte über  $\Omega = \mathbb{R}$ . Das zugehörige W-Mass heisst stetige Gleichverteilung oder – wegen der Form von  $f$  – Rechteck-Verteilung, kurz  $\mathfrak{R}(a, b)$ .

Bei der  $\mathfrak{R}(a, b)$ -Verteilung hat jedes Intervall  $[c, d] \subset [a, b]$  eine Wahrscheinlichkeit, die seinem Längenanteil in  $[a, b]$  entspricht, nämlich  $(d-c)/(b-a)$ .



Def: Exponential-Verteilung  $\text{Exp}(\alpha)$

Ist  $\alpha > 0$  und

$$f(x) = \alpha e^{-\alpha x} 1_{(0, \infty)}(x) = \begin{cases} \alpha e^{-\alpha x} & : x > 0, \\ 0 & : x \leq 0, \end{cases}$$

dann ist  $f$  eine R-Dichte. Das zugehörige W-Mass heisst Exponential-Verteilung.



Die Exponentialverteilung ist eine typische Lebensdauerverteilung. So ist beispielsweise die Lebensdauer von elektronischen Geräten häufig annähernd exponentialverteilt.

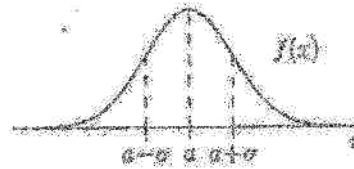
Def: Normalverteilung  $N(a, \sigma^2)$

Für jeden Wert  $a \in \mathbb{R}$  und  $\sigma > 0$  ist

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-a)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

eine R-Dichte. Das zugehörige W-Mass heisst Normalverteilung mit dem „Mittelwert“  $a$  und der „Streuung“  $\sigma$ , kurz  $N(a, \sigma^2)$ . Im Spezialfall  $a = 0$  und  $\sigma = 1$  spricht man von der Standard-Normalverteilung  $N(0,1)$ . In diesem Fall wird die R-Dichte als  $\varphi(x)$  geschrieben:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$$



### 3.3 Verteilungsfunktionen

In der bisherigen Darstellung von W-Massen durch R-Dichten besteht ein gewisser Nachteil darin, dass für jedes Ereignis ein Integral ausgewertet werden muss. Man kann dies vermeiden, indem man das unbestimmte Integral, die sog. Stammfunktion, heranzieht.

Def: Verteilungsfunktion

Ist  $P$  ein beliebiges W-Mass über  $(\mathbb{R}, \mathcal{B})$  dann heisst die Abbildung  $F: \mathbb{R} \rightarrow \mathbb{R}$  mit  $F(x) := P((-\infty, x])$ ,  $x \in \mathbb{R}$ ,

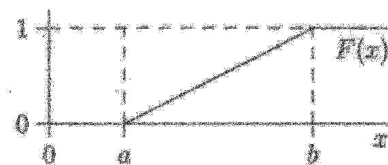
die Verteilungsfunktion (VF) von  $P$ . Aus obiger Definition erhält man die Darstellung  $P((a, b]) = F(b) - F(a)$ ,  $a, b \in \mathbb{R}$ ,  $a < b$ .

Für ein W-Mass  $P$  über  $\mathbb{R}$  mit R-Dichte  $f$  gilt:

$$F(x) = \int_{-\infty}^x f(t) dt \quad \text{und} \quad P((a, b]) = \int_a^b f(t) dt = F(b) - F(a).$$

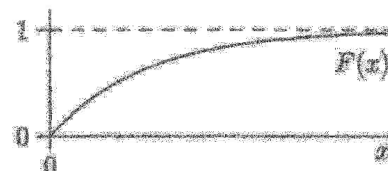
Verteilungsfunktion für die Rechteck-Verteilung:

$$F(x) = \begin{cases} 0, & x \leq a, \\ (x-a)/(b-a), & a \leq x \leq b, \\ 1, & a \geq b \end{cases}$$



Verteilungsfunktion der Exponential-Verteilung:

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-ax}, & x \geq 0 \end{cases}$$



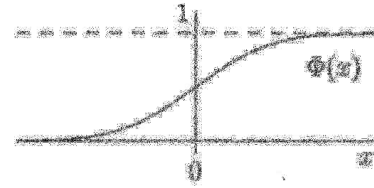
Verteilungsfunktion der Normalverteilung:

Für die Verteilungsfunktion  $\Phi$  der  $N(0,1)$ -Verteilung gilt:

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt, \quad x \in \mathbb{R}$$

Daraus folgt für die  $N(a, \sigma^2)$ -Verteilung:

$$F_{a, \sigma^2}(x) = \int_{-\infty}^x \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{t-a}{\sigma}\right)^2} dt = \Phi\left(\frac{x-a}{\sigma}\right)$$



Empirische Verteilungsfunktion:

$$\hat{F}_n^x(x) := \frac{1}{n} \sum_{i=1}^n 1_{[x_i, \infty)}(x), \quad x \in \mathbb{R}$$