

Manfred Brill

Mathematik für Informatiker

Hanser Verlag

ISBN 3-446-22802-0

©2005

Folien zum Kapitel 15

Wahrscheinlichkeitstheorie
und Statistik

15.1 Beschreibende Statistik

Merkmale

- *Quantitative Merkmale oder Messgrößen*
 $\{0, 1, 3, 5.5, 8\}$
- *Qualitative Merkmale*
 $\{ \text{„ledig“}, \text{„verheiratet“}, \text{„verwitwet“} \}$
- *Rangmerkmal*
„Gehalt besser, gleich oder niedriger“.

Definition

Für $x \in \mathbb{R}$ ist $g(x)$ die Anzahl der Werte in einer Stichprobe, die gleich x sind. $g(x)$ heißt *absolute Häufigkeit* von x , die Funktion g *absolute Häufigkeit* des Merkmals X .

$h(x) = \frac{1}{n}g(x) \in [0; 1]$ wird *relative Häufigkeit* von x genannt, dabei ist n die Größe der Stichprobe.

Die Funktion h heißt *empirische Dichte* oder *Häufigkeitsfunktion* des betrachteten Merkmals.

Beispieldaten

Die letzten 50 Lieferungen eines Lieferanten hatten die folgenden Lieferzeiten in Tagen:

4, 5, 4, 1, 5, 4, 3, 4, 5, 6, 6, 5, 5, 4, 7, 4, 6, 5, 6, 4, 5, 4, 7, 5, 5,
6, 7, 3, 7, 6, 6, 7, 4, 5, 4, 7, 7, 5, 5, 5, 5, 6, 6, 4, 5, 2, 5, 4, 7, 5

Lieferzeiten

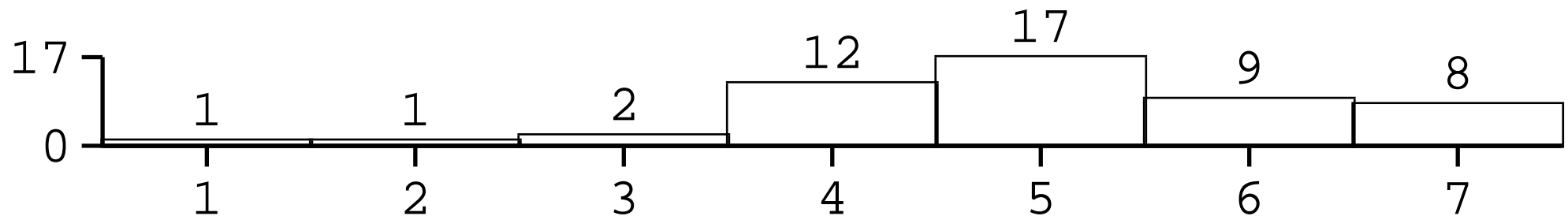


Abbildung 1: Histogramm der absoluten Häufigkeiten

Lieferzeiten

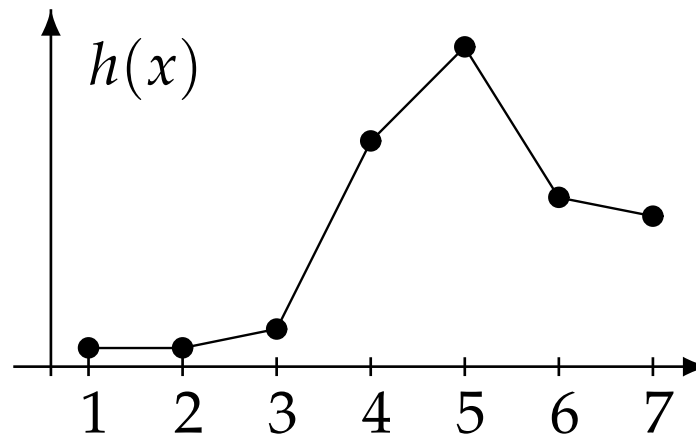


Abbildung 2: Empirische Dichte der Lieferzeiten

Definition

Ist h die empirische Dichte eines Merkmals X und ist K_1, K_2, \dots, K_n eine Partition des Wertebereichs von X , ist die *relative Klassenhäufigkeit* gegeben durch

$$\bar{h}_j = \frac{|\{i \mid x_i \in K_j\}|}{n}, 1 \leq j \leq n.$$

Regeln für die Klassenbildung

- Wählen Sie die Klassengrenzen äquidistant!
- Bestimmen Sie die Höhe der Histogramm-Rechtecke proportional zur Klassenhäufigkeit!
- Bilden Sie bei $n = 100$ mindestens 10 Klassen, 13 Klassen bei $n \sim 1\,000$, 16 Klassen bei $n \sim 10\,000$ (DIN 55 302, Blatt 1).

Beispieldaten

Ein Unternehmen protokolliert die Lebensdauer eines Ersatzteils. Dabei entsteht die folgende Stichprobe mit $n = 30$, die die Lebensdauer der Teile in Stunden enthält:

110, 520, 490, 30, 120, 290, 370, 305, 415, 170, 280, 70, 540, 460, 260,
345, 150, 220, 435, 425, 470, 350, 130, 380, 230, 320, 360, 240, 330, 580.

Relative Klassenhäufigkeiten

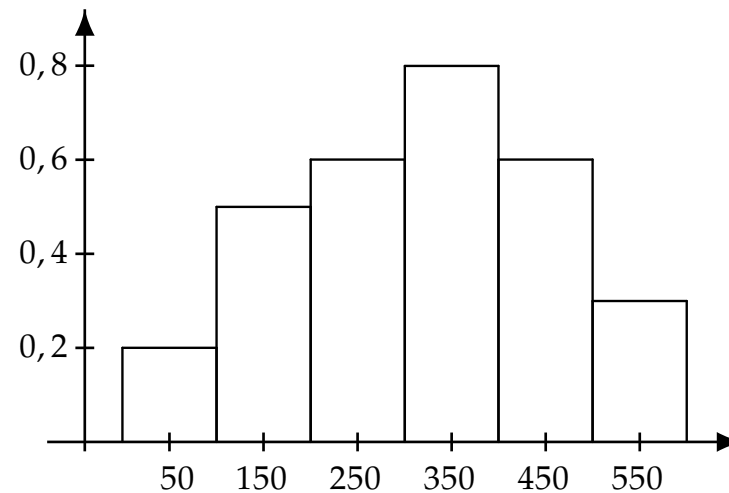


Abbildung 3: Histogramm der relativen Klassenhäufigkeiten der Lebensdauern

Definition

Für eine Stichprobe X und $x \in \mathbb{R}$ ist $G(x)$ die *Summenhäufigkeit* die Anzahl aller Werte x_i mit $x_i \leq x$.

Die Funktion $H(x) = \frac{G(x)}{n}$ heißt *relative Summenhäufigkeit* oder *empirische Verteilungsfunktion*.

Empirische Verteilung

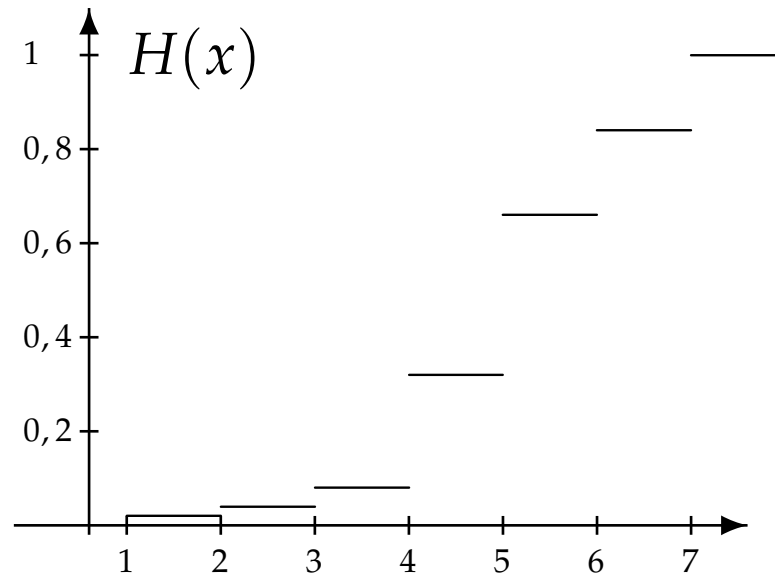


Abbildung 4: Empirische Verteilungsfunktion der Lieferzeiten

Empirische Verteilung

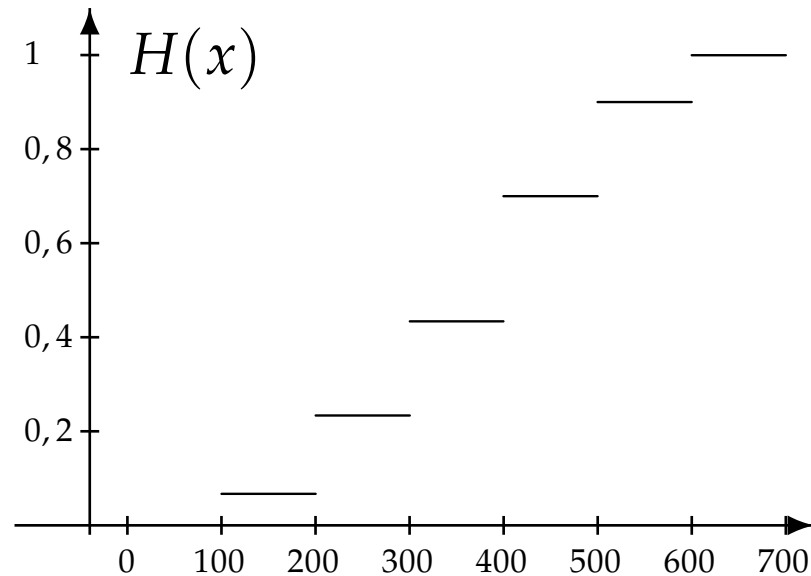


Abbildung 5: Empirische Verteilungsfunktion der Lebensdauern

Definition

Die Ausprägung x_i einer Stichprobe eines Merkmals X , die die größte absolute Häufigkeit aufweist, heißt *Modalwert* x_{mod} .

Definition

Der *Median* oder *Zentralwert* der Stichprobe ist die Zahl x_{med} , für die 50% aller Stichprobenwerte kleiner oder gleich und mindestens 50% größer oder gleich x_{med} sind.

Definition

Das *arithmetische Mittel* einer Stichprobe mit n Werten ist gegeben durch

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i.$$

Definition

Die *Spannbreite* einer Stichprobe ist die Differenz zwischen größtem und kleinstem Stichprobenwert.

Definition

Die *Varianz* der Stichprobe ist

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2$$

Die positive Wurzel $s = +\sqrt{s^2}$ heißt *Standardabweichung* der Stichprobe.

Satz

Sind x_i Werte einer Stichprobe mit n Werten, \bar{x} das arithmetische Mittel und s die Standardabweichung, dann liegen alle Stichprobenwerte im Intervall

$$[\bar{x} - \sqrt{n-1}s; \bar{x} + \sqrt{n-1}s].$$

Kontingenztabelle

h_{ij}		Randhäufigkeiten h_{1i}
Randhäufigkeiten h_{2j}		

Kontingenztabelle

h_{ij}	1	2	3	4	Randhäufigkeiten h_{1i}
2	2	0	3	3	
3	0	1	2	3	
4	0	1	2	1	
5	0	0	0	2	
Randhäufigkeiten h_{2j}					

Kontingenztabelle

h_{ij}	1	2	3	4	Randhäufigkeiten h_{1i}
2	2	0	3	3	8
3	0	1	2	3	6
4	0	1	2	1	4
5	0	0	0	2	2
Randhäufigkeiten h_{2j}	2	2	7	9	

Definition

Sind X und Y zwei quantitative Merkmale und sind weder alle x_i - noch alle y_i -Werte gleich.

$$s_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

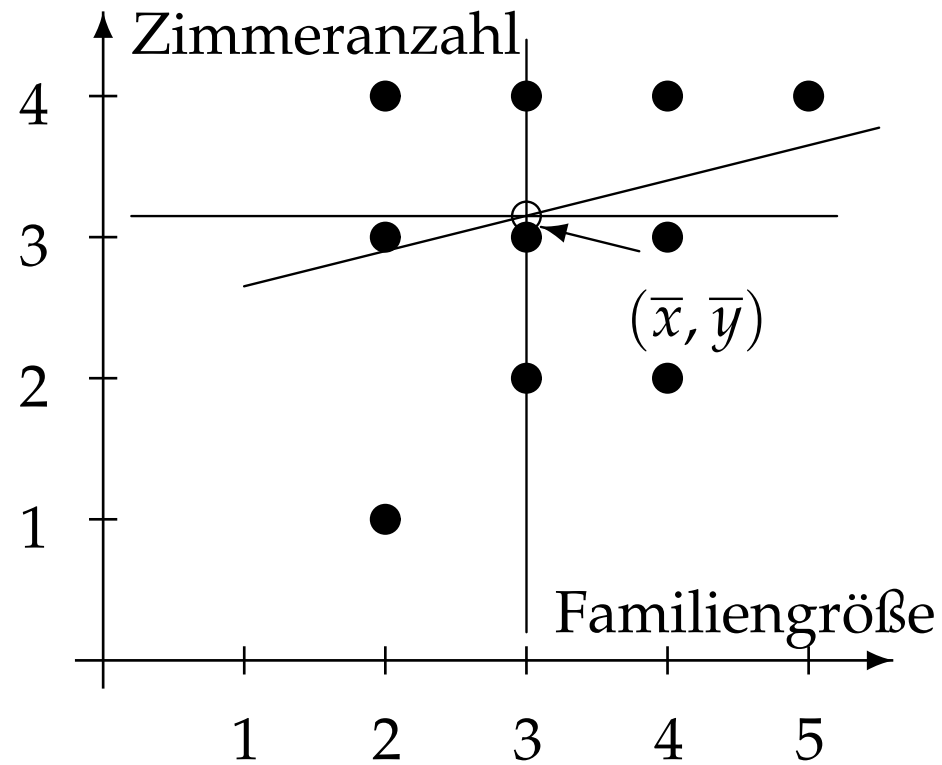
heißt *Stichprobenkovarianz*.

Definition

Der *Bravais-Pearson-Korrelationskoeffizient* ist gegeben durch

$$r = \frac{s_{xy}}{s_x s_y}.$$

Streuungsdiagramm



Definition

Die *bedingte relative Häufigkeit* $f_1(x_i|y_j)$ ist die relative Häufigkeit für x_i bei denjenigen Stichprobenwerten, für die im zweiten Merkmal y_j auftritt.

Definition

Die *bedingte relative Häufigkeit* $f_2(y_j|x_i)$ ist die relative Häufigkeit für y_j bei denjenigen Stichprobenwerten, für die im zweiten Merkmal x_i auftritt.

15.2

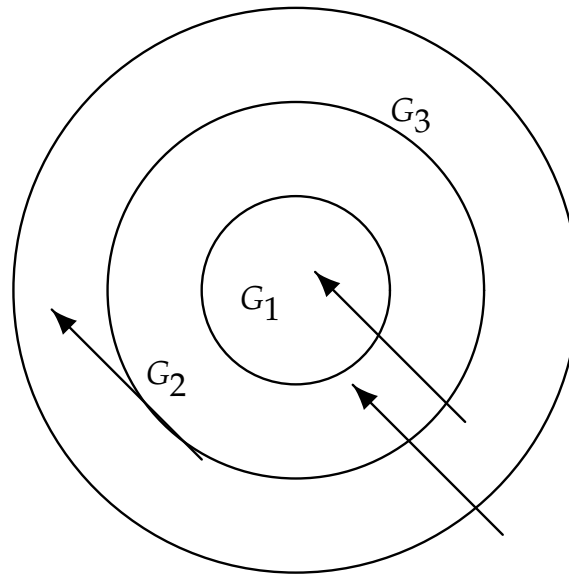
Wahrscheinlichkeitstheorie

Definition

Ω sei eine gegebene Menge. Eine Teilmenge $\mathbb{A} \subseteq \mathbb{P}(\Omega)$ heißt *Ereignisalgebra*, wenn die folgenden Eigenschaften erfüllt sind:

- $\Omega \in \mathbb{A}$;
- $A \in \mathbb{A} \Rightarrow A^c \in \mathbb{A}$;
- Ist (A_i) eine Folge von Elementen von \mathbb{A} , dann ist auch $\bigcup_{i=1}^{\infty} A_i \in \mathbb{A}$.

Geometrische Wahrscheinlichkeiten



Definition

Ist Ω eine Menge und $\mathbb{A} \subseteq \mathbb{P}(\Omega)$ eine Ereignisalgebra. Ein *Wahrscheinlichkeitsmaß* oder kurz eine *Wahrscheinlichkeit* ist eine Abbildung $P : \mathbb{A} \rightarrow [0; 1]$ mit den folgenden Eigenschaften:

■ $P(\Omega) = 1.$

■ Für eine Folge $(A_n) \subseteq \mathbb{A}$ mit $\forall i \neq j A_i \cap A_j = \emptyset$ gilt

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

$P(A)$ heißt *Wahrscheinlichkeit* des Ereignisses A . Das Tripel (Ω, \mathbb{A}, P) wird als *Wahrscheinlichkeitsraum* bezeichnet.

Satz

Ist (Ω, \mathbb{A}, P) ein Wahrscheinlichkeitsraum, dann gilt:

■ $P(\emptyset) = 0.$

■ Für $A \cap B = \emptyset$ gilt $P(A \cup B) = P(A) + P(B);$

für endlich viele Ereignisse A_1, \dots, A_n mit $A_i \cap A_j = \emptyset$ für $i \neq j$ gilt

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i).$$

■ $P(A) + P(A^c) = 1.$

Satz

Ist (Ω, \mathbb{A}, P) ein Wahrscheinlichkeitsraum, dann gilt:

- Für zwei beliebige Ereignisse $A, B \in \mathbb{A}$ gilt
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$
- Für $A \subseteq B$ gilt $P(A) \leq P(B)$, die Wahrscheinlichkeit ist monoton.

Satz

Für eine beliebige Folge $(A_n) \subseteq \mathbb{A}$ von Teilmengen in einem Wahrscheinlichkeitsraum gilt

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i).$$

Definition

Ist (Ω, \mathbb{A}, P) ein Wahrscheinlichkeitsraum und $A \in \mathbb{A}$ ein Ereignis mit $P(A) > 0$.

Die *bedingte Wahrscheinlichkeit von B unter der Bedingung A* ist gegeben durch

$$P(B|A) = \frac{P(B \cap A)}{P(A)}.$$

Satz

Es gilt $P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$.

Satz

Ist (Ω, \mathbb{A}, P) ein Wahrscheinlichkeitsraum und ist $A_i, 1 \leq i \leq n$ eine Partition von Ω , dann gilt

$$P(B) = \sum_{i=1}^n P(B|A_i) \cdot P(A_i).$$

Satz

Ist (Ω, \mathbb{A}, P) ein Wahrscheinlichkeitsraum, $A_i, 1 \leq i \leq n$ eine Partition von Ω und B ein Ereignis mit $P(B) > 0$, dann gilt

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{i=1}^n P(B|A_i) \cdot P(A_i)}.$$

Definition

Sind A und B zwei Ereignisse mit $P(A) > 0$ und gilt $P(B|A) = P(B)$, dann heißt das Ereignis B *unabhängig* vom Ereignis A .

15.3 Zufallsvariable

Definition

Ist (Ω, \mathbb{A}, P) ein Wahrscheinlichkeitsraum, dann heißt eine Abbildung $X : \Omega \rightarrow \mathbb{R}$ *Zufallsvariable*.

Definition

Ist X eine Zufallsvariable auf (Ω, \mathbb{A}, P) , dann heißt die Funktion $F_X : \mathbb{R} \rightarrow [0; 1]$ mit

$$F_X(t) = P(X \leq t) = P(\{\omega \in \Omega \mid X(\omega) \leq t\})$$

Verteilungsfunktion der Zufallsvariablen X .

Satz

Ist F_X die Verteilungsfunktion einer Zufallsvariablen X , dann gilt:

- $\lim_{t \rightarrow -\infty} F_X(t) = 0$ und $\lim_{t \rightarrow \infty} F_X(t) = 1$.
- F_X ist monoton wachsend.
- $P(t_1 < X \leq t_2) = F_X(t_2) - F_X(t_1)$.

Definition

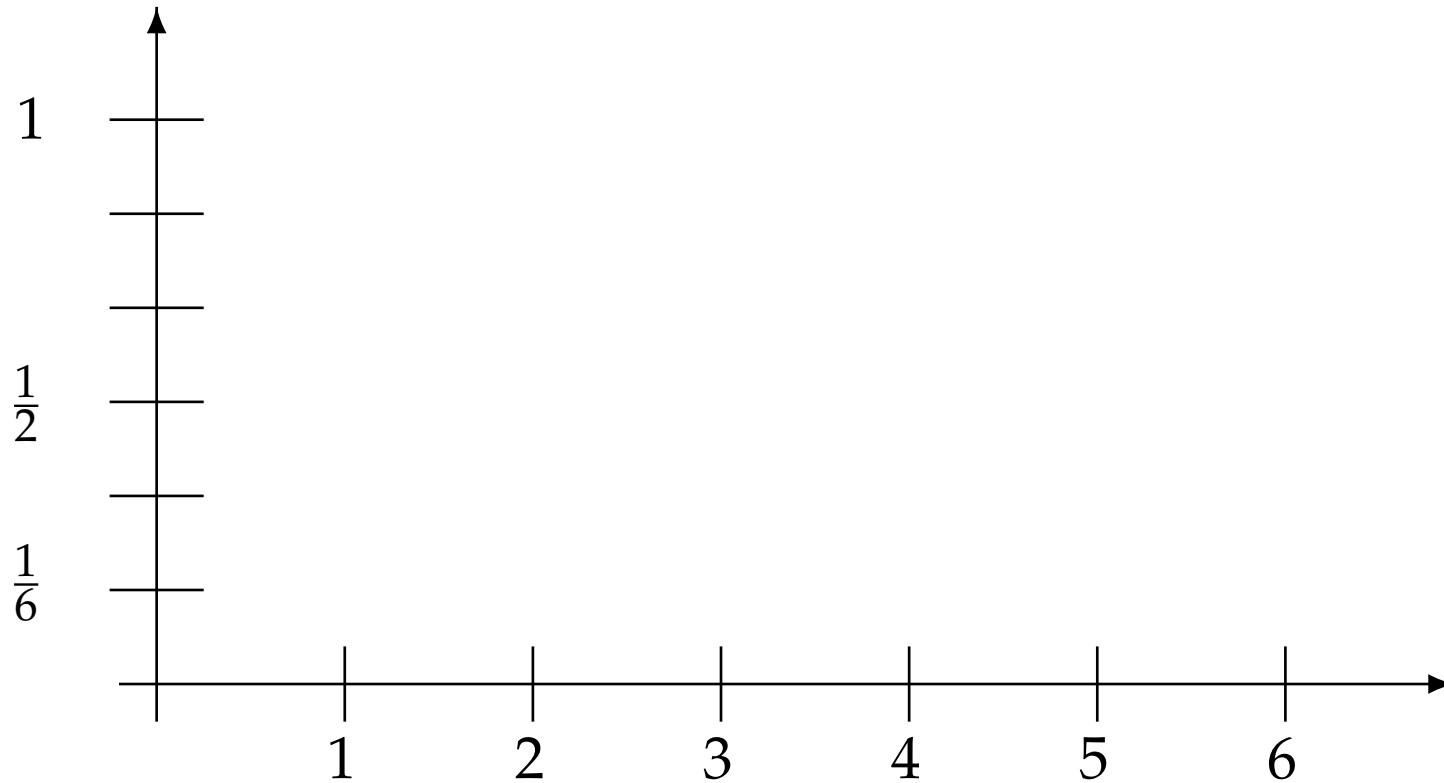
Eine Zufallsvariable heißt *diskret*, wenn sie einen Wertebereich mit *endlich vielen* oder *abzählbar vielen* Elementen besitzt.

Eine diskrete Dichte

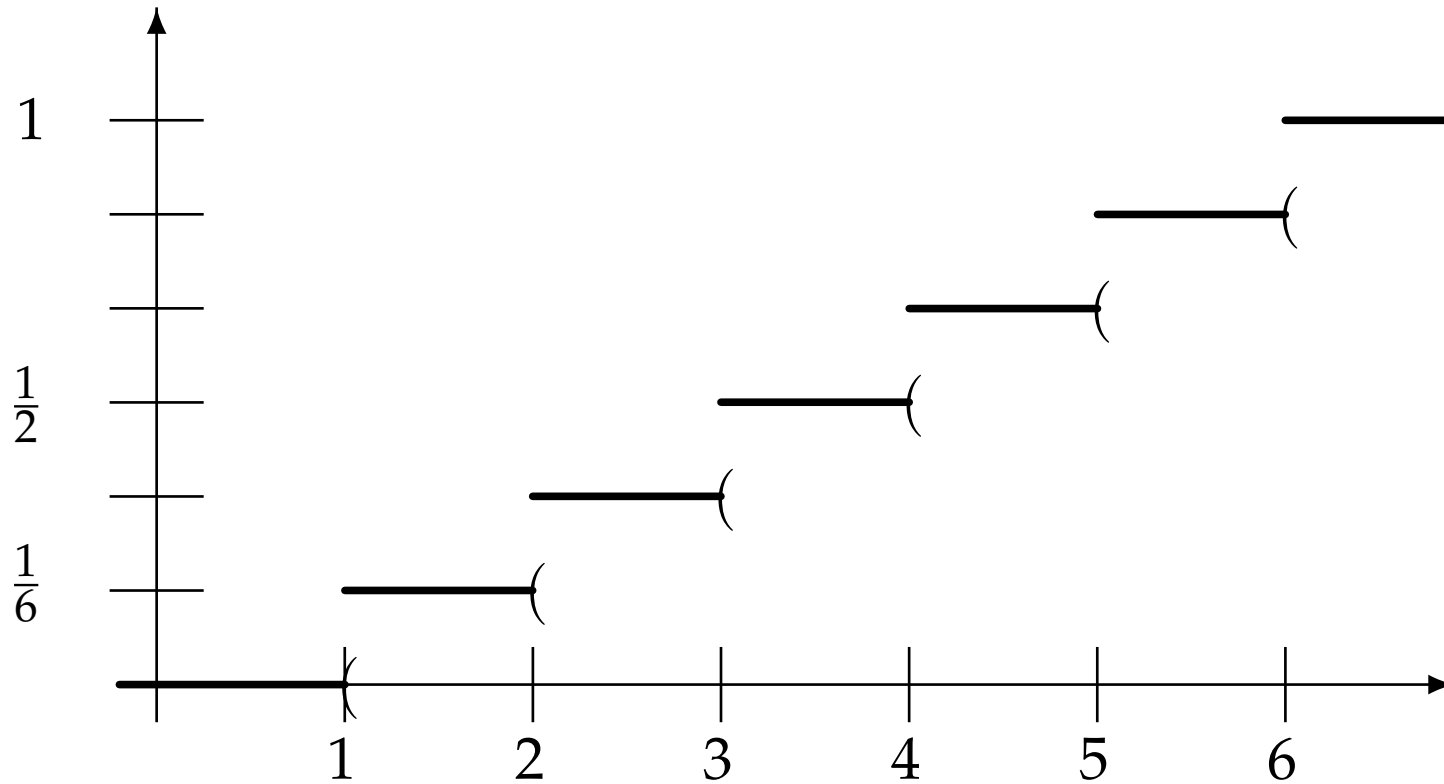
Zeichnen sie die Verteilungsfunktion der diskreten Dichte!

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

Aufgabe



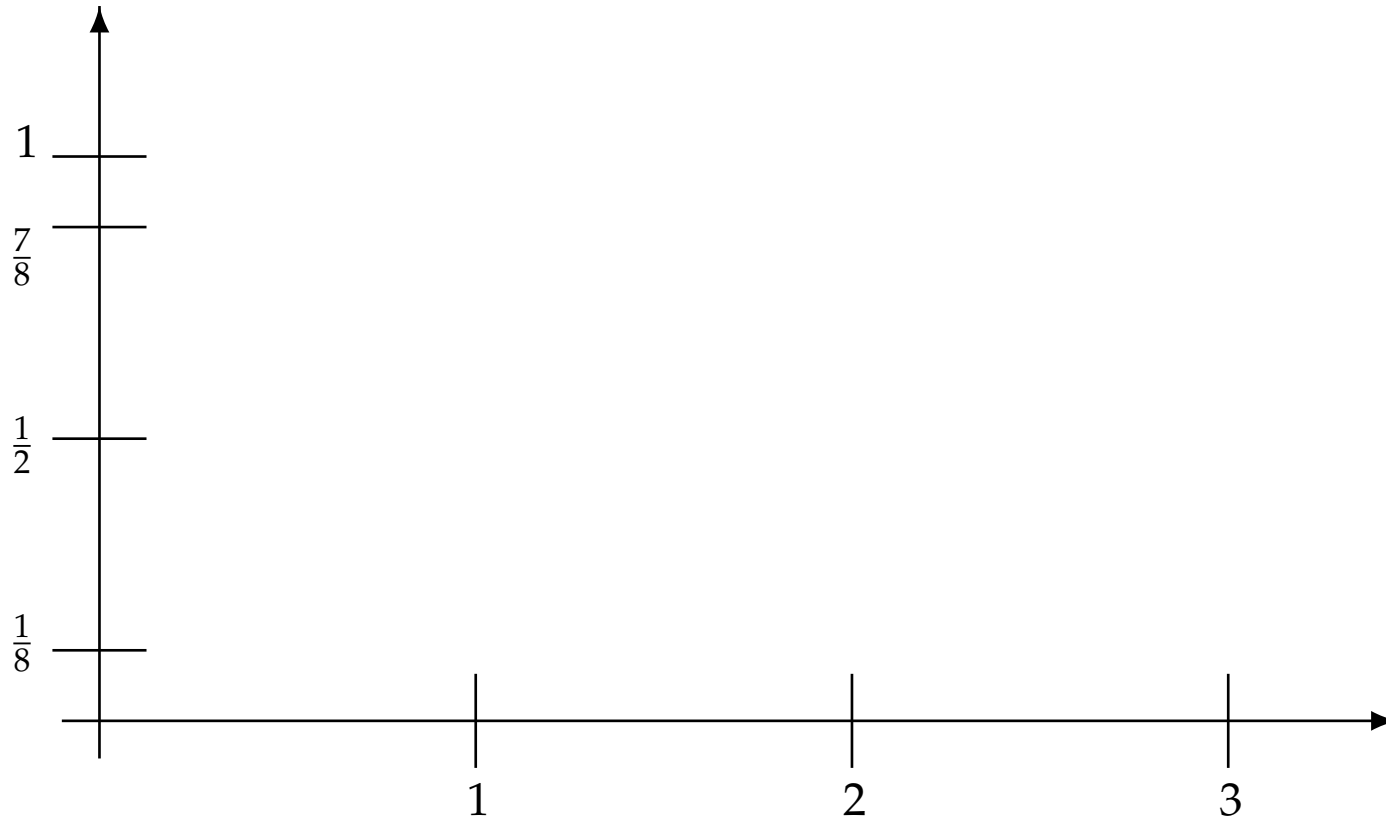
Lösung



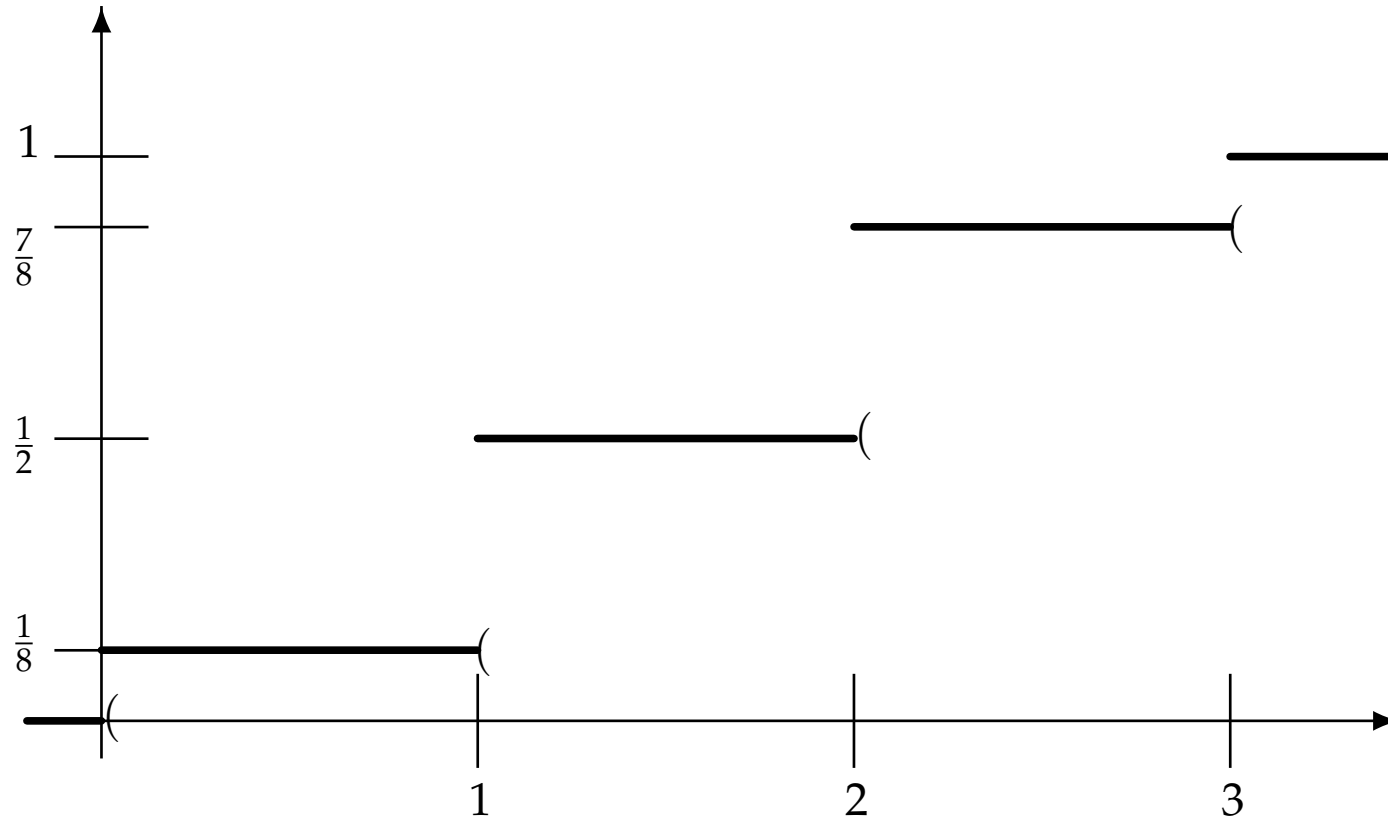
Dreimaliges Werfen einer Münze

x_i	0	1	2	3
p_i	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Verteilungsfunktion



Verteilungsfunktion



Definition

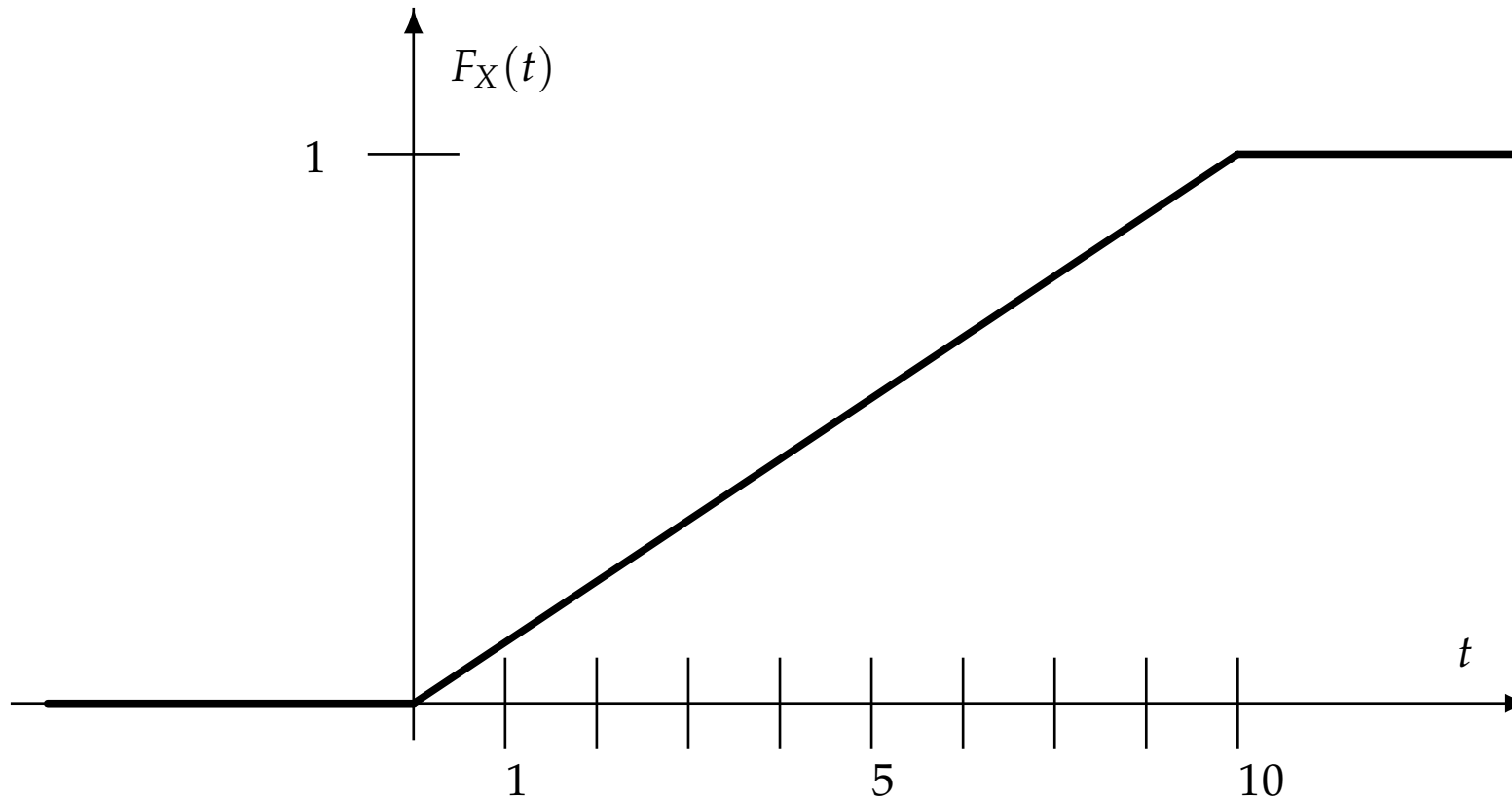
Eine Zufallsvariable X heißt *stetig*, wenn es eine integrierbare Funktion $f_X : \mathbb{R} \rightarrow \mathbb{R}$ mit $f_X(t) \geq 0$ gibt, sodass für die Verteilungsfunktion F_X und jedes $t \in \mathbb{R}$

$$F_X(t) = P(X \leq t) = \int_{-\infty}^t f_X(x) dx$$

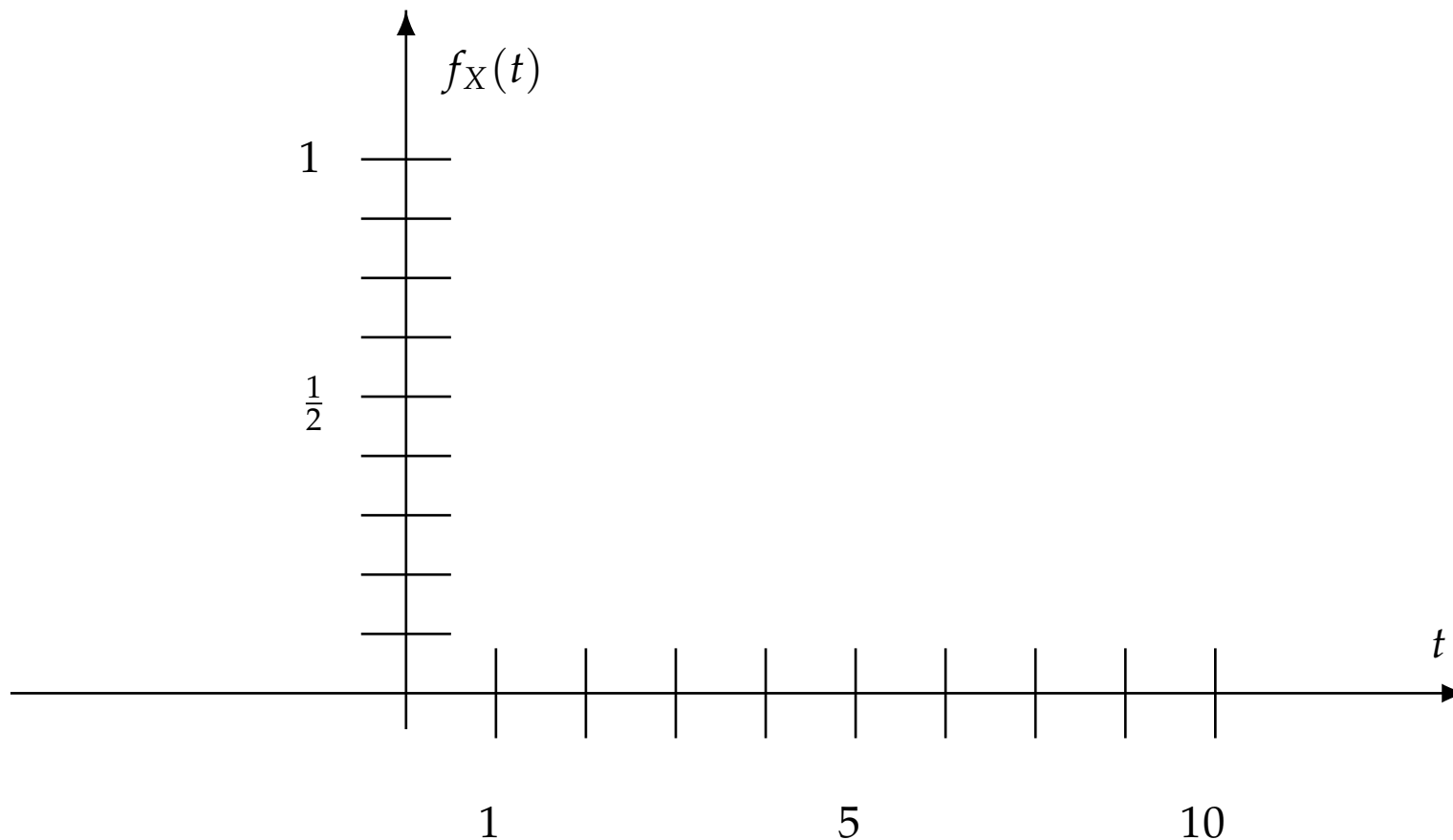
erfüllt ist.

Die Funktion f_X heißt *Wahrscheinlichkeitsdichte* oder kurz *Dichtefunktion* von X .

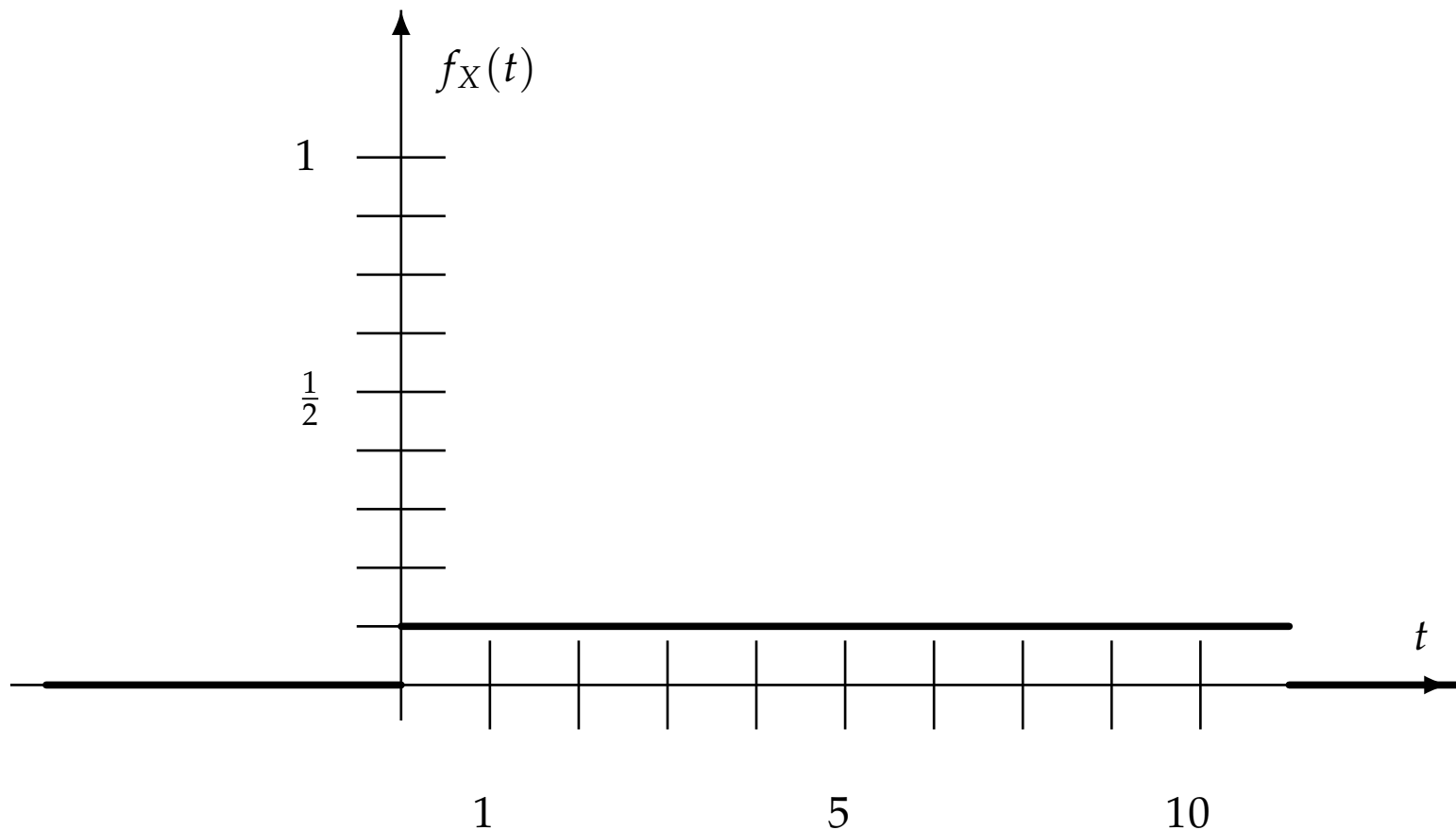
Verteilungsfunktion



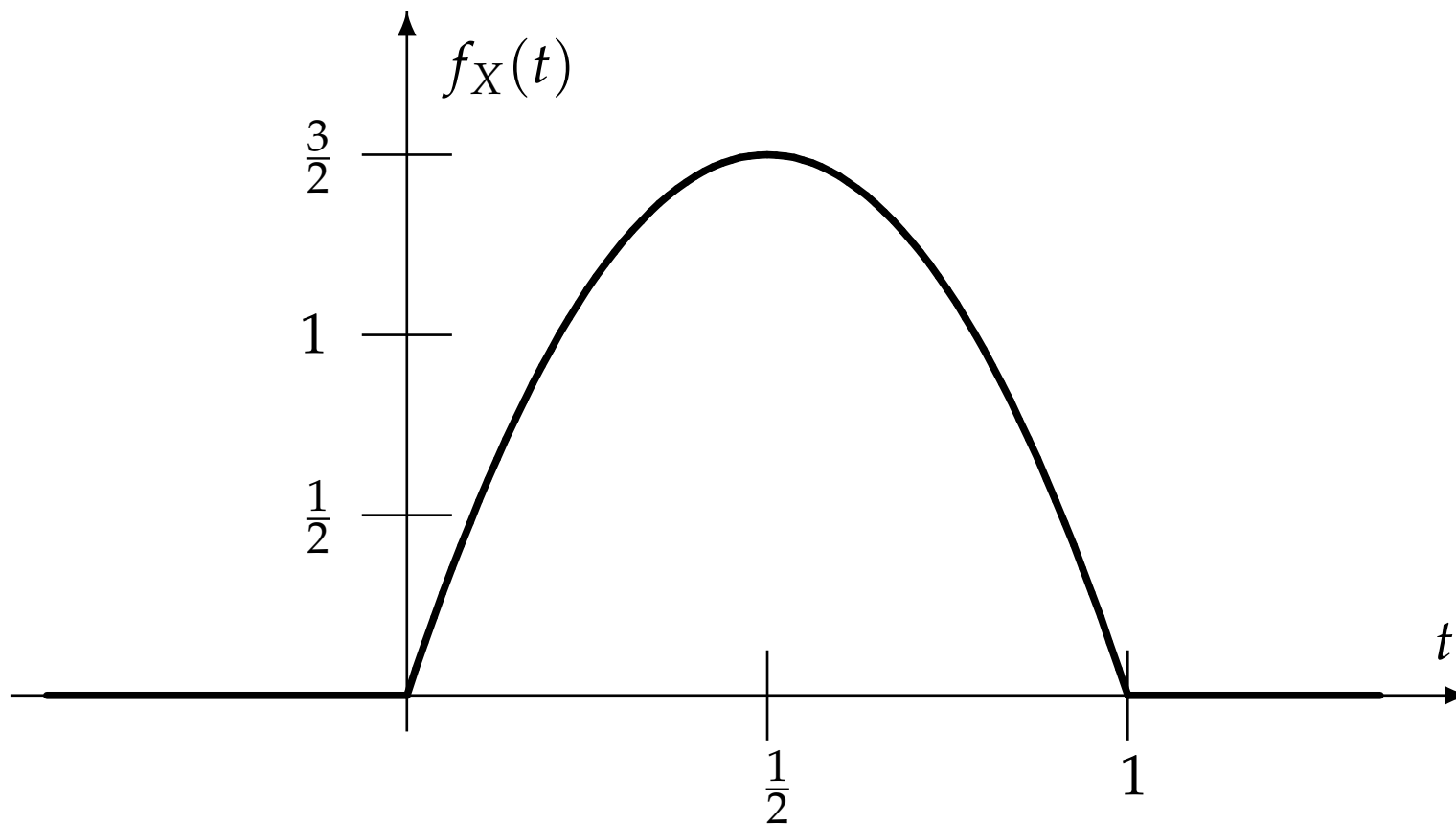
Zugehörige Dichtefunktion



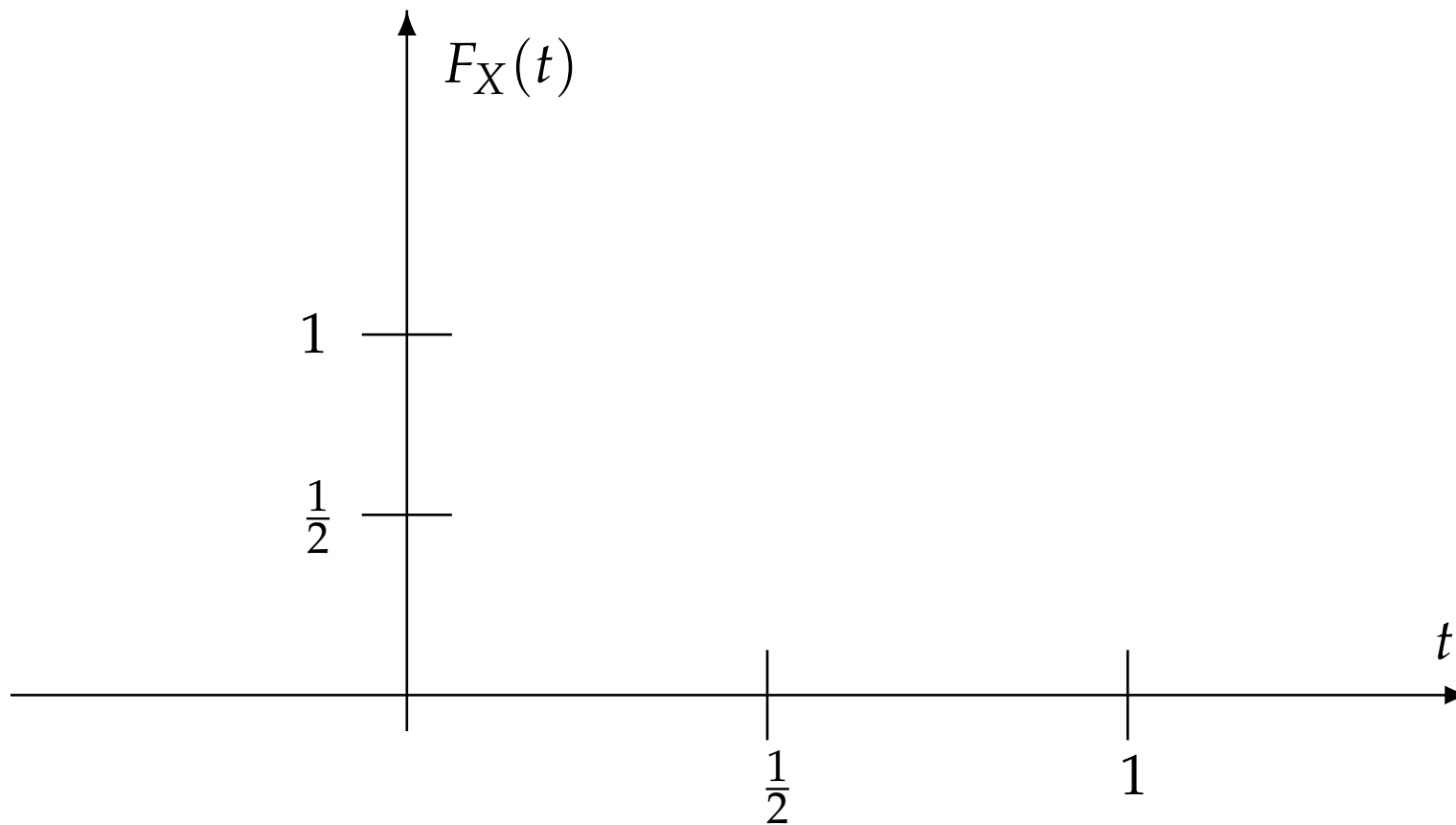
Zugehörige Dichtefunktion



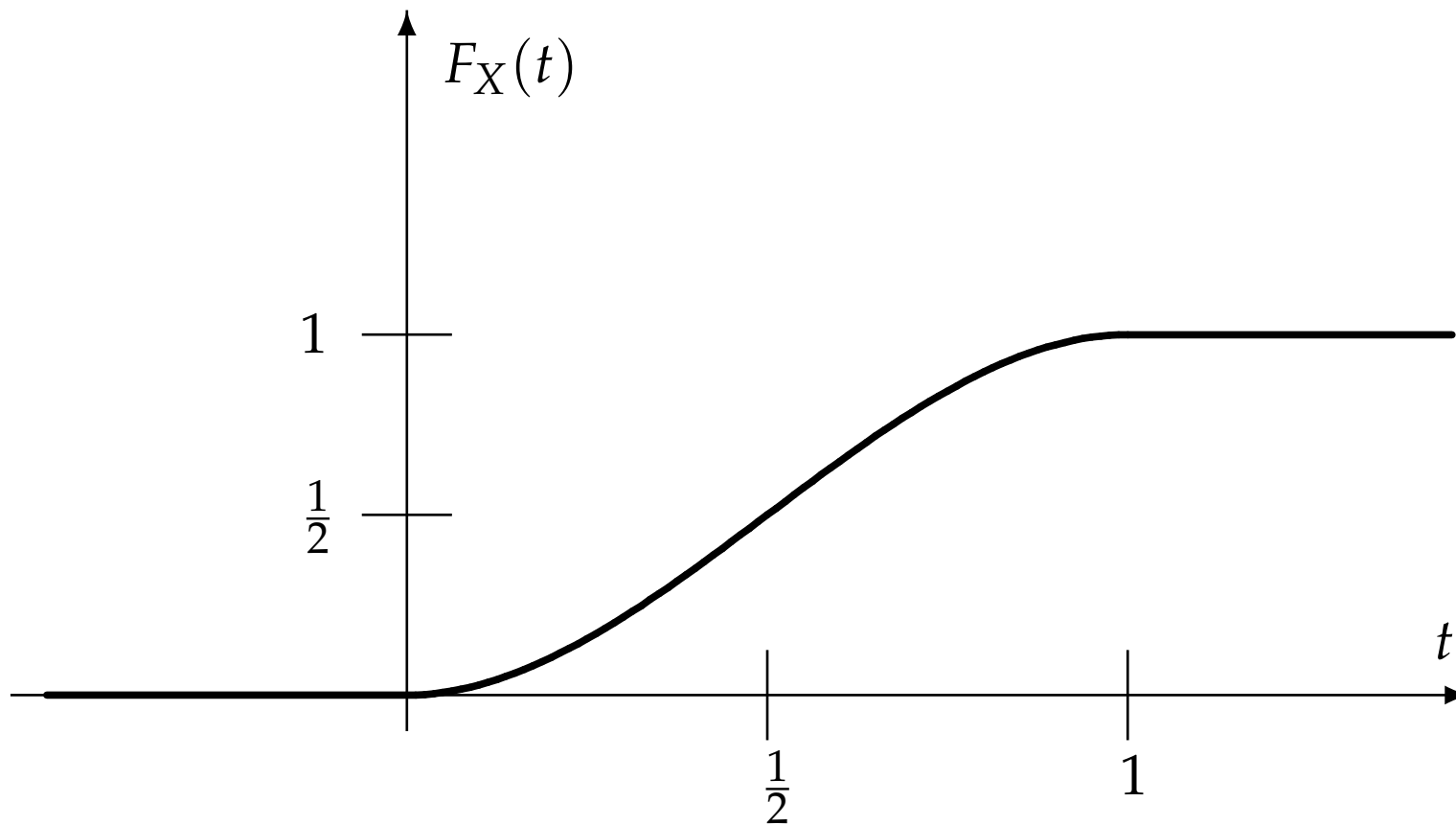
$$f_X(t) = 6t(1 - t)$$



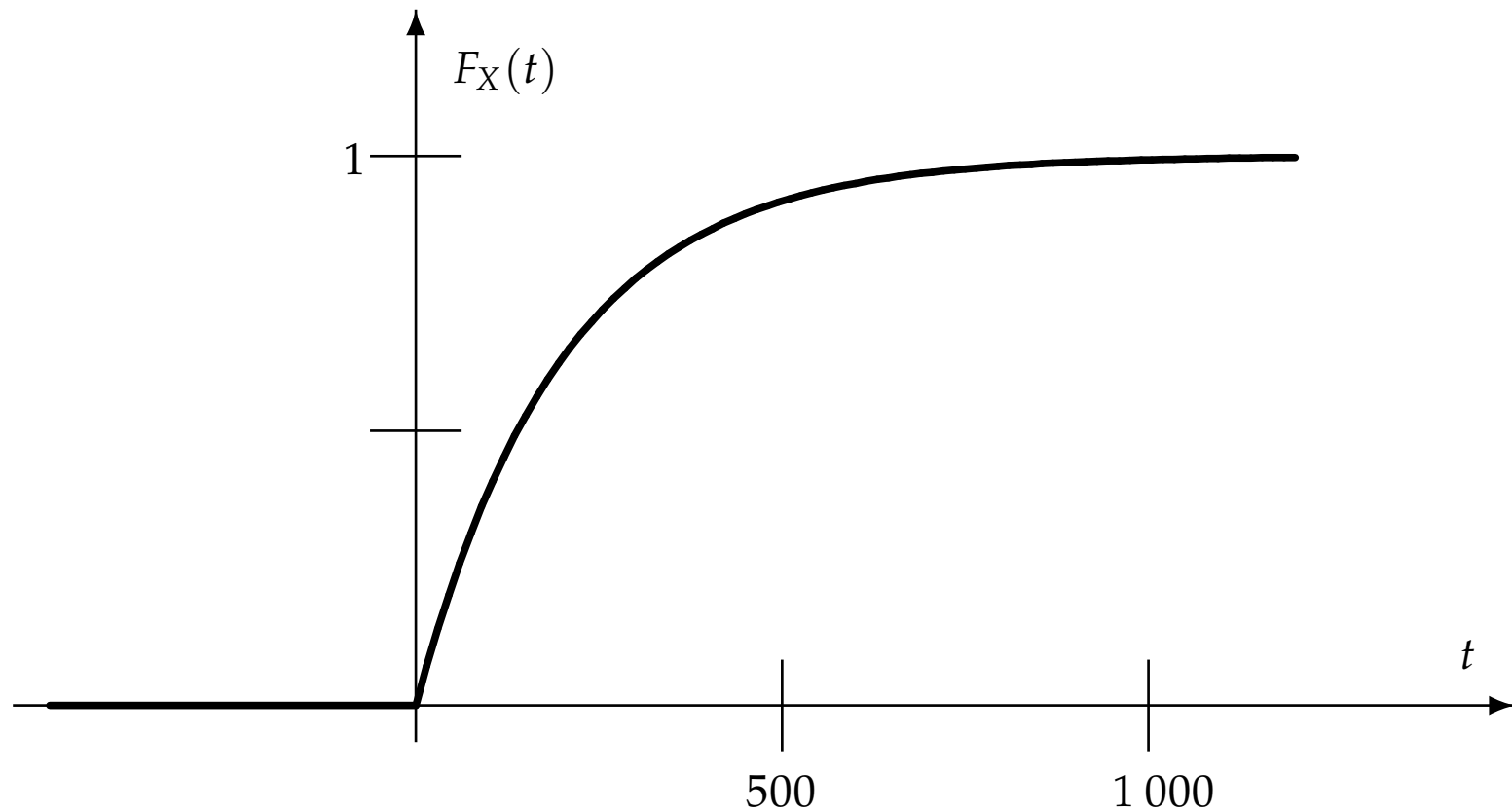
Verteilungsfunktion



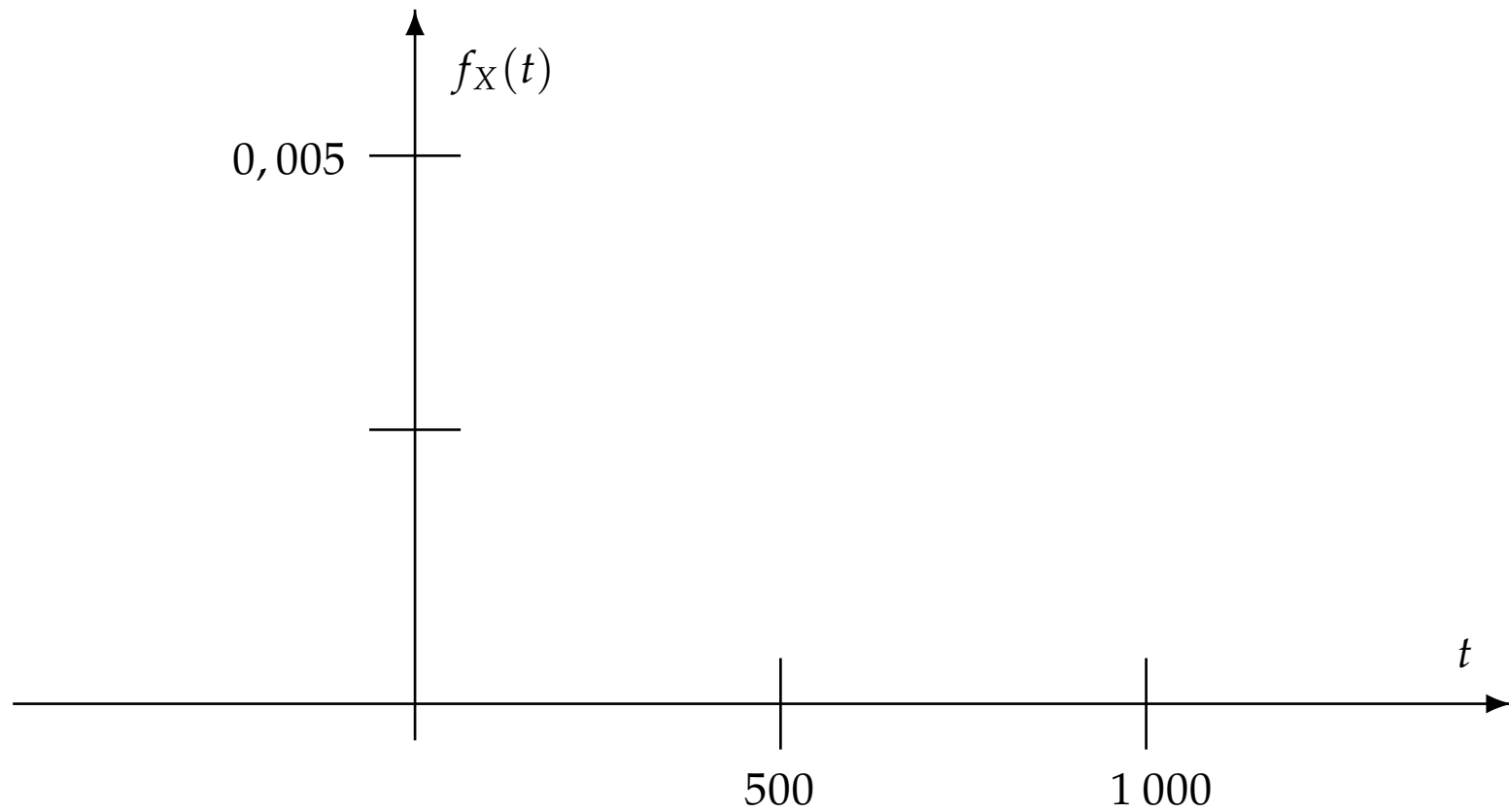
Verteilungsfunktion



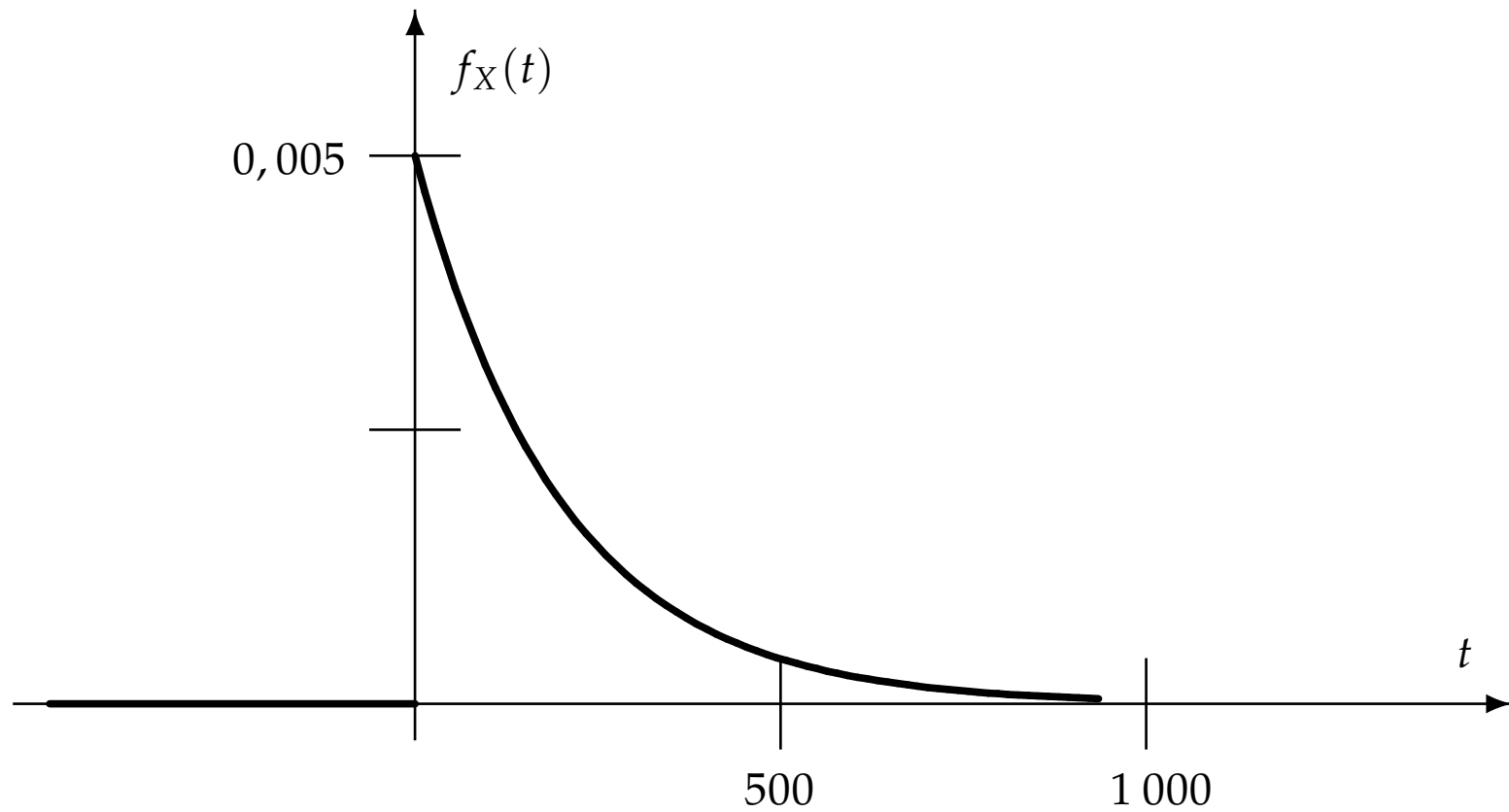
Verteilungsfunktion



Dichtefunktion



Dichtefunktion



Definition

Ist X eine diskrete Zufallsvariable mit der diskreten Dichte (x_i, p_i) und konvergiert die Reihe $\sum_{i=1}^{\infty} x_i p_i$ absolut, dann heißt

$$E(X) = \mu = \sum_{i=1}^{\infty} x_i p_i$$

Erwartungswert der diskreten Zufallsvariablen X .

Definition

Für eine stetige Zufallsvariable mit der Dichtefunktion f_X ist der *Erwartungswert* durch

$$E(X) = \mu = \int_{-\infty}^{\infty} x f_X(x) dx$$

gegeben, falls das Integral existiert.

Definition

Ist X eine diskrete Zufallsvariable mit der diskreten Dichte (x_i, p_i) und konvergiert die Reihe $\sum_{i=1}^{\infty} (x_i - E(X))^2 p_i$ absolut, dann heißt

$$\text{Var}(X) = \sigma^2 = \sum_{i=1}^{\infty} (x_i - E(X))^2 p_i$$

Varianz der Zufallsvariablen X .

$\sigma = \sqrt{\text{Var}(X)}$ heißt *Standardabweichung* von X .

Definition

Für eine stetige Zufallsvariable ist die Varianz durch

$$\text{Var}(X) = \sigma^2 = \int_{-\infty}^{\infty} (x - E(X))^2 f_X dx$$

gegeben, falls das Integral existiert.

Aufgabe

Tabelle 1: Zufallsvariable X

x_i	1	8	18
p_i	$\frac{1}{5}$	$\frac{1}{2}$	$\frac{3}{10}$

Bestimmen Sie Erwartungswert und Varianz der Zufallsvariablen!

Aufgabe

Tabelle 2: Zufallsvariable Y

x_i	7	8	14
p_i	$\frac{1}{5}$	$\frac{1}{2}$	$\frac{3}{10}$

Bestimmen Sie Erwartungswert und Varianz der Zufallsvariablen!

Satz

Für den Erwartungswert einer Zufallsvariablen gilt:

- $E(aX + b) = aE(X) + b$ für $a, b \in \mathbb{R}$.
- Ist $g : \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion, sodass $g(X)$ eine Zufallsvariable ist, dann gilt für eine diskrete Zufallsvariable
 $E(g(X)) = \sum_{i=1}^{\infty} g(x_i) p_i$, falls diese Reihe absolut konvergiert.
Für eine stetige Zufallsvariable gilt
 $E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx$, falls das Integral existiert.
- Sind X und Y zwei Zufallsvariablen, für die $\forall \omega \in \mathbb{A}$
 $X(\omega) \leq Y(\omega)$ erfüllt ist. Dann gilt auch $E(X) \leq E(Y)$.

Satz

Für die Varianz einer diskreten Zufallsvariablen gilt:

- $Var(X) = E(X^2) - E(X)^2$.
- $Var(aX + b) = a^2 Var(X)$ für $a, b \in \mathbb{R}$.

Definition

Sind X und Y zwei Zufallsvariable mit

$$P(\{X \leq t_x\} \cap \{Y \leq t_y\}) = P(\{X \leq t_x\}) \cdot P(\{Y \leq t_y\}),$$

dann nennt man die beiden Zufallsvariablen *unabhängig*.

Satz

Für zwei Zufallsvariable X und Y gilt

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

und

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2E((X - E(X))(Y - E(Y))).$$

Definition

Die *Kovarianz* der Zufallsvariablen X und Y ist definiert durch

$$\text{Cov}(X, Y) = E((X - E(X))(Y - E(Y)))$$

Definition

Sind X und Y zwei Zufallsvariable mit $\text{Cov}(X, Y) = 0$, dann heißen sie *unkorreliert*.

Satz

Sind die Zufallsvariablen X und Y unabhängig, dann gilt

$$E(X \cdot Y) = E(X) \cdot E(Y)$$

und

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y).$$

Ungleichung von Tschebyscheff

Ist X eine Zufallsvariable und $c > 0$, dann gilt

$$P(|X - E(X)| \geq c) \leq \frac{\text{Var}(X)}{c^2}.$$

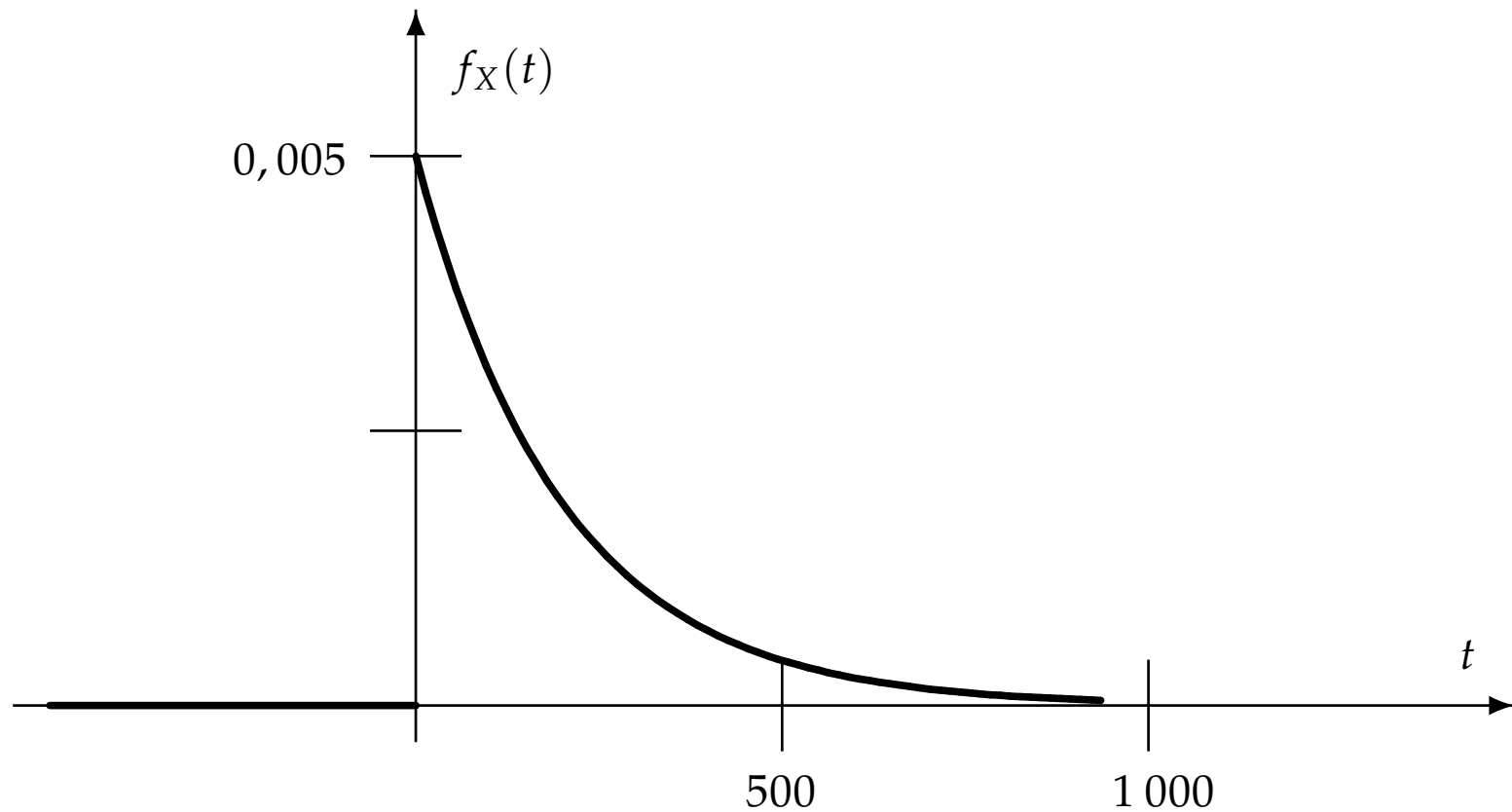
Definition

Ist X eine stetige Zufallsvariable mit Verteilungsfunktion F_X und $p \in [0; 1]$. Dann heißt die Zahl t_p mit

$$F_X(t_p) = P(X \leq t_p) = p$$

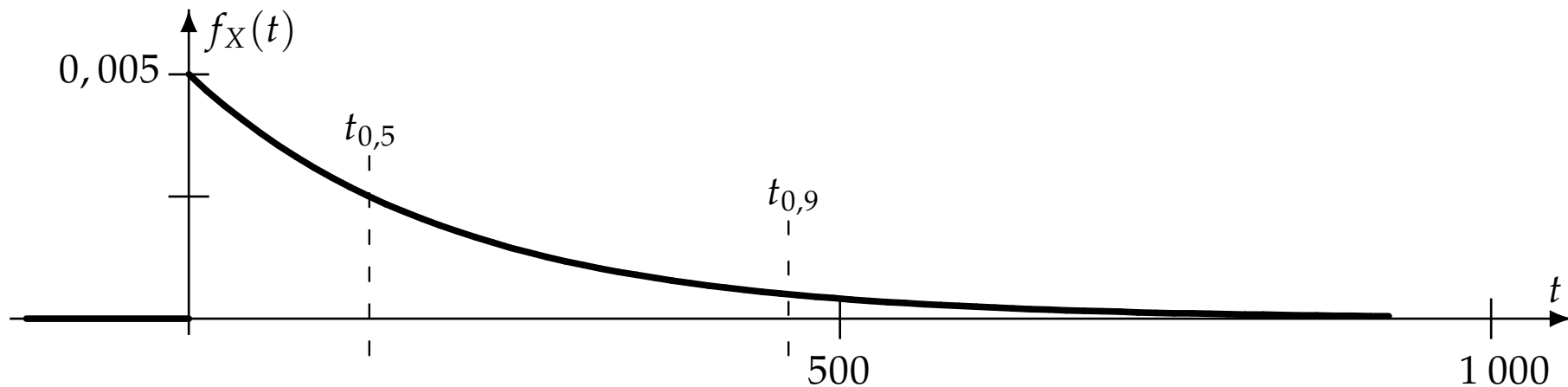
das *Quantil der Ordnung p* oder kurz *p -Quantil* von X .

Dichtefunktion



Quantile

50%- und 90%-Quantil:



15.4 Diskrete und stetige Verteilungen

Definition

Eine diskrete Zufallsvariable X heißt *diskret gleichmäßig verteilt*, wenn ihre n Werte $\{x_1, x_2, \dots, x_n\}$ alle mit der gleichen Wahrscheinlichkeit

$$p = P(X = x_i) = \frac{1}{n}$$

angenommen werden.

Definition

Eine diskrete Zufallsvariable X heißt *Bernoulli-Variable*, falls ihr Wertebereich aus der Menge $\{0, 1\}$ besteht und

$$P(X = 1) = p, P(X = 0) = 1 - p$$

gilt mit $0 \leq p \leq 1$.

Die Zahl p heißt *Erfolgswahrscheinlichkeit*.

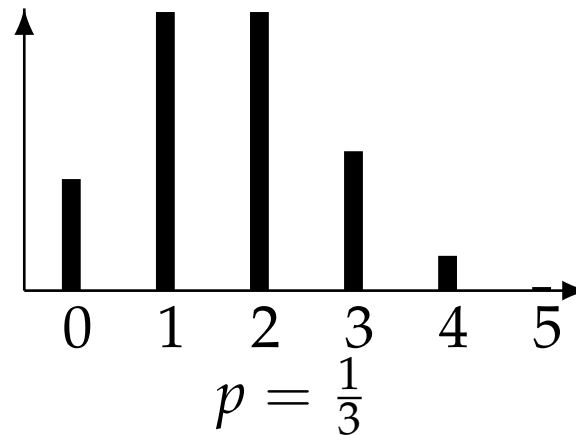
Definition

Eine diskrete Zufallsvariable X mit Wertebereich $\{0, 1, \dots, n\}$ heißt *binomial verteilt*, wenn ihre Einzelwahrscheinlichkeiten durch

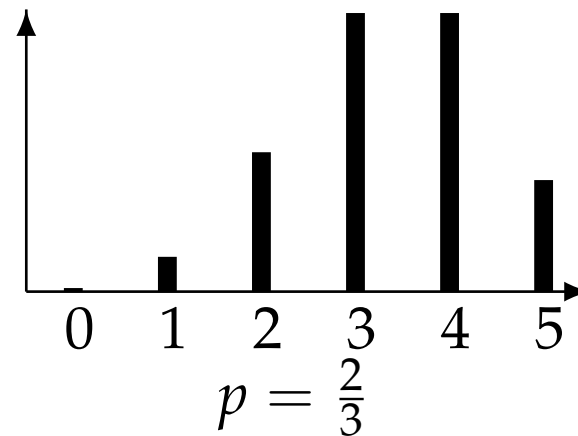
$$p_k = \binom{n}{k} p^k (1 - p)^{n-k}$$

gegeben sind. Man sagt kurz X ist $B(n; p)$ verteilt.

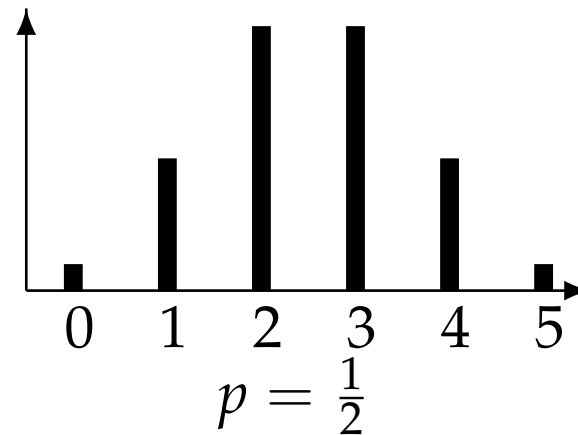
Stabdiagramme für $B(5; p)$



Stabdiagramme für $B(5; p)$



Stabdiagramme für $B(5; p)$



Definition

Eine diskrete Zufallsvariable X mit Wertebereich \mathbb{N} heißt *geometrisch verteilt*, wenn ihre Einzelwahrscheinlichkeiten durch

$$p_k = p(1 - p)^{k-1}, p \in [0; 1]$$

gegeben sind.

Satz

Für eine geometrisch verteilte Zufallsvariable X gilt

$$P(X = j + k | X > k) = P(X = j).$$

Definition

Eine diskrete Zufallsvariable X mit Wertebereich $\{0, 1, \dots, n\}$ heißt *hypergeometrisch verteilt* mit den Parametern N, M und n , wenn ihre Einzelwahrscheinlichkeiten durch

$$p_k = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}, 0 \leq k \leq n$$

gegeben sind. Man sagt kurz X ist $Hyp(N; M; n)$ -verteilt.

Vergleich

k	$Hyp(100; 10; 5)$	$Bi(5; 0, 1)$
0	0,583 75	0,590 49
1	0,339 39	0,328 05
2	0,070 22	0,072 9
3	0,006 38	0,008 1
4	0,000 25	0,000 45
5	$3,3 \cdot 10^{-6}$	10^{-5}

Definition

Eine diskrete Zufallsvariable X mit Wertebereich $\mathbb{N} \cup \{0\}$ besitzt eine *Poisson-Verteilung* mit Parameter $\lambda > 0$, wenn ihre Einzelwahrscheinlichkeiten durch

$$p_k = \frac{\lambda^k}{k!} e^{-\lambda}$$

gegeben sind. Man sagt kurz: X ist $P(\lambda)$ -verteilt.

Definition

Eine stetige Zufallsvariable X mit der Dichtefunktion

$$f_X(t) = \begin{cases} \frac{1}{b-a}, & t \in [a; b], \\ 0, & t \notin [a; b] \end{cases}$$

heißt *gleichverteilt* über dem Intervall $[a; b]$.

Definition

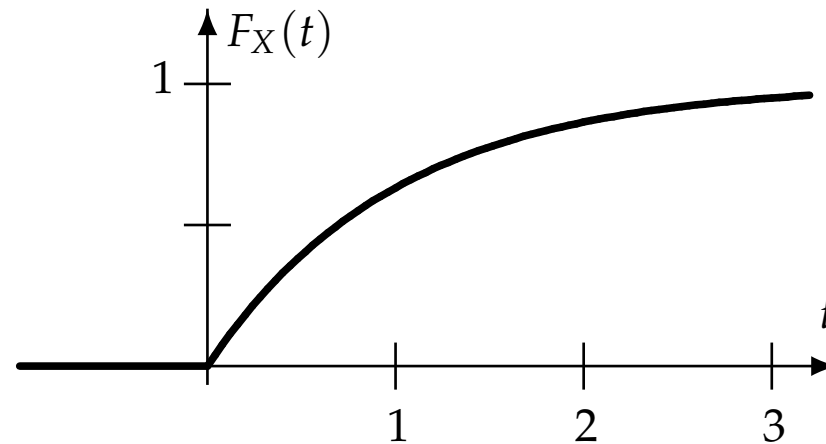
Eine stetige Zufallsvariable X besitzt eine *Exponentialverteilung*, falls sie für $\lambda > 0$ die Dichtefunktion

$$f_X(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0, \\ 0 & t < 0 \end{cases}$$

besitzt. Man sagt kurz: X ist $E(\lambda)$ -verteilt.

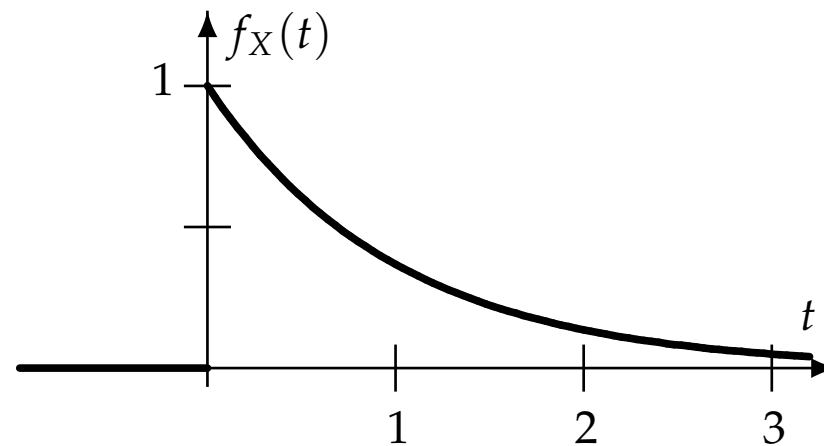
Verteilungsfunktion

Exponentialverteilung für $\lambda = 1$:



Dichtefunktion

Exponentialverteilung für $\lambda = 1$:



Satz

Für eine $E(\lambda)$ -verteilte Zufallsvariable X gilt

$$P(X \leq s + t | X > s) = P(X \leq t).$$

Definition

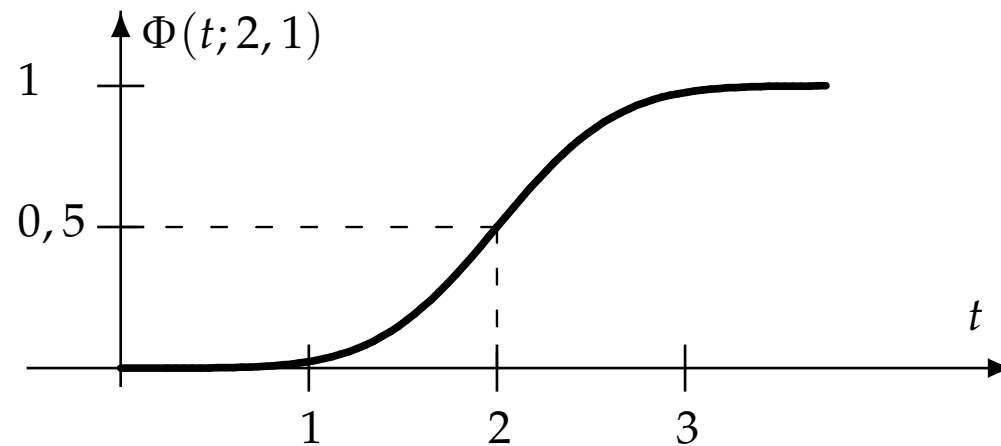
Eine stetige Zufallsvariable X besitzt eine *Normalverteilung*, falls sie für $\sigma > 0$ die Dichtefunktion

$$f_X(t) = \varphi(t; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

besitzt. Man sagt kurz: X ist $N(\mu; \sigma^2)$ -verteilt.

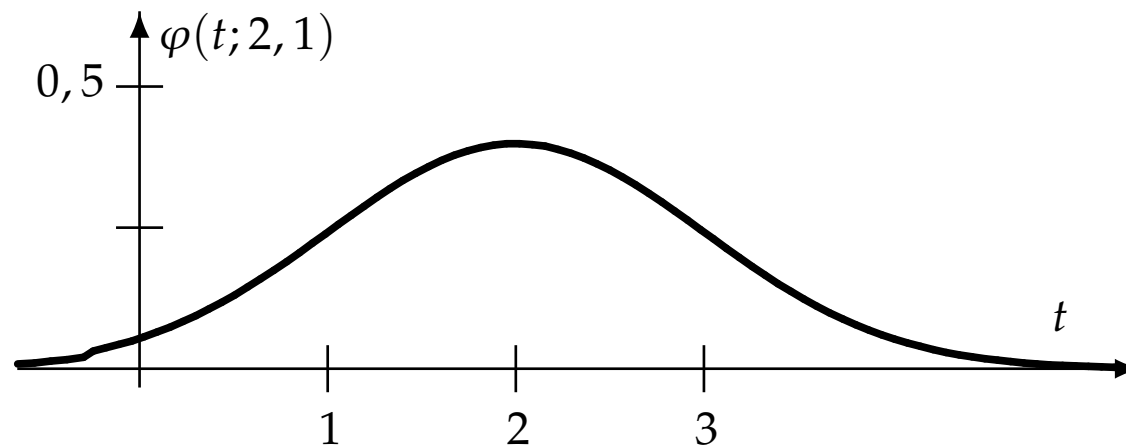
Verteilungsfunktion

Normalverteilung für $\mu = 2, \sigma^2 = 1$:

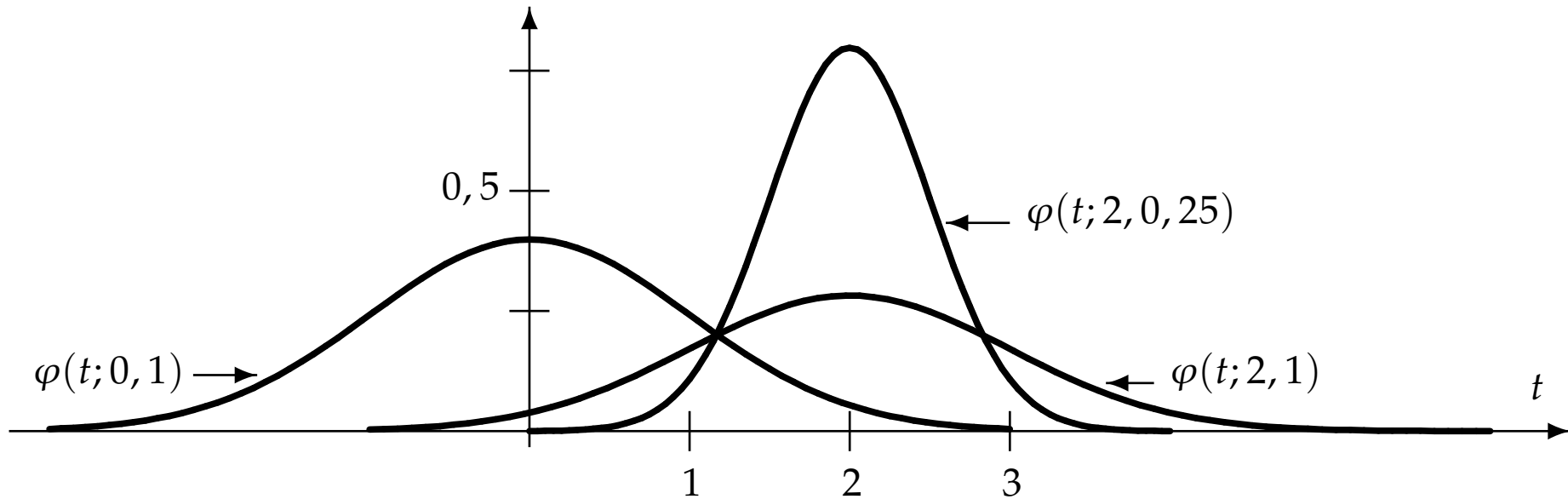


Dichtefunktion

Normalverteilung für $\mu = 2, \sigma^2 = 1$:



Dichtefunktionen



Satz

Ist X $N(\mu; \sigma^2)$ -verteilt, dann ist $Z = \frac{X - \mu}{\sigma}$ eine $N(0; 1)$ -verteilte Zufallsvariable.

Approximation

$$\Phi(t; 0, 1) \approx \left(\frac{1}{2} + \left(\frac{1}{\sqrt{\pi}} - \frac{1}{2} \right) e^{-\frac{t^2}{\sqrt{2\pi}}} \right) \sqrt{1 - e^{-\frac{t^2}{2}}},$$

Verteilungen

Verteilung	$E(X)$	$Var(X)$
$B(n; p)$	$n \cdot p$	$n \cdot p \cdot (1 - p)$
$Hyp(N; M; n)$	$n \frac{M}{N}$	$n \frac{M}{N} \left(\frac{N-M}{N} \frac{N-n}{N-1} \right)$
$P(\lambda)$	λ	λ
Gleichverteilung auf $[a; b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
$E(\lambda)$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$N(\mu; \sigma^2)$	μ	σ^2

15.5 Schätzverfahren in der schließenden Statistik

Satz

Sind $X_i, i \in \mathbb{N}$ unabhängige Zufallsvariablen mit $\mu = E(X_i)$, $\sigma^2 = \text{Var}(X_i)$. Dann erfüllt die Folge der arithmetischen Mittel $\overline{(X_n)}$ das *schwache Gesetz der großen Zahlen*:

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} P(|\overline{X_n} - \mu| \geq \varepsilon) = 0.$$

Definition

Eine Punktschätzung X' für einen Parameter ϑ einer Verteilung X heißt *erwartungstreu*, wenn $E(X') = \vartheta$.

Werden für die Schätzung n Stichprobenwerte verwendet und ist $\lim_{n \rightarrow \infty} E(X') = \vartheta$, dann heißt die Schätzung *asymptotisch erwartungstreu*.

Definition

Eine *Maximum-Likelihood-Schätzung* eines Parameters ϑ einer Wahrscheinlichkeitsverteilung der Zufallsvariablen X ist der Wert für ϑ , für den $L(\vartheta)$ maximal ist.

Verteilungen

Verteilung	Maximum-Likelihood-Schätzer
$B(n; p)$	$p = \bar{x}$
$P(\lambda)$	$\lambda = \bar{x}$
$E(\lambda)$	$\lambda = \frac{1}{\bar{x}}$
$N(\mu; \sigma^2)$	$\mu = \bar{x}$
$N(\mu; \sigma^2)$	bekannter Mittelwert: $\sigma = \sqrt{\frac{1}{n} \sum (x_i - \mu)^2}$
$N(\mu; \sigma^2)$	unbekannter Mittelwert: $\sigma = \sqrt{\frac{1}{n} \sum (x_i - \bar{x})^2}$