Statistik für Informatiker, SS 2019

1 Grundlagen aus der Wahrscheinlichkeitstheorie

1.1 Ereignisse, Zufallsvariablen und Wahrscheinlichkeiten

1.1.3 Der Fall mit Dichte

Matthias Birkner

http://www.staff.uni-mainz.de/birkner/StatInfo19/

JOHANNES GUTENBERG UNIVERSITÄT MAINZ

6.5.2019

Zufallsvariablen mit Dichten sind ein kontinuierliches Analogon zu Zufallsvariablen mit Gewichten. In vielen Situationen ist eine Modellierung eines zufälligen Werts X als „allgemeine“ reelle Zahl angemessen, d.h. die Annahme, dass der Wertebereich S diskret ist, ist zu „eng“.

Zufallsvariablen mit Dichten sind ein kontinuierliches Analogon zu Zufallsvariablen mit Gewichten. In vielen Situationen ist eine Modellierung eines zufälligen Werts X als „allgemeine" reelle Zahl angemessen, d.h. die Annahme, dass der Wertebereich S diskret ist, ist zu „eng".

(Auch wenn man argumentieren könnte, dass die Menge der im Rechner mit gegebener Genauigkeit darstellbaren Werte prinzipiell diskret ist, ist es oft „unpraktisch", sich immer auf eine konkrete Diskretisierung festlegen zu müssen.)

Beispiel 1.22 (Approximation der Exponentialverteilung durch reskalierte geometrisch verteilte ZVn)

Sei $W \sim \mathrm{Geom}_p$ mit $p \ll 1$, $X := pW$ (hat Werte in $p\mathbb{Z}_+ \subset \mathbb{R}$). Für jedes feste $K \in \mathbb{N}$ ist

$$P(W \geq K) = \sum_{j=K}^{\infty} P(W = j) = \sum_{j=K}^{\infty} p(1-p)^j = (1-p)^K p \sum_{\ell=0}^{\infty} (1-p)^\ell$$

$$= (1-p)^K p \frac{1}{1-(1-p)} = (1-p)^K$$

(≈ 1, wenn p sehr klein)

Beispiel 1.22 (Approximation der Exponentialverteilung durch reskalierte geometrisch verteilte ZVn)

Sei $W \sim \mathrm{Geom}_p$ mit $p \ll 1$, $X := pW$ (hat Werte in $p\mathbb{Z}_+ \subset \mathbb{R}$). Für jedes feste $K \in \mathbb{N}$ ist

$$P(W \geq K) = \sum_{j=K}^{\infty} P(W = j) = \sum_{j=K}^{\infty} p(1-p)^j = (1-p)^K p \sum_{\ell=0}^{\infty} (1-p)^\ell$$
$$= (1-p)^K p \frac{1}{1-(1-p)} = (1-p)^K$$

(≈ 1, wenn p sehr klein), andererseits ist für (festes) $x > 0$

$$P(X \geq x) = P(pW \geq x) = P\big(W \geq \lceil x/p \rceil\big)$$
$$= (1-p)^{\lceil x/p \rceil} \approx (1-p)^{x/p} \approx e^{-x}$$

(und zwar “egal, wie klein” p ist).

Bsp. 1.22, Forts.

$$P(pW \geq x) = (1-p)^{\lceil x/p \rceil} \approx (1-p)^{x/p} \approx e^{-x}$$

Interpretation z.B.: via Wartezeiten auf sehr fein zeitdiskretisiertem Gitter

Bsp. 1.22, Forts.

$$P(pW \geq x) = (1-p)^{\lceil x/p \rceil} \approx (1-p)^{x/p} \approx e^{-x}$$

Interpretation z.B.: via Wartezeiten auf sehr fein zeitdiskretisiertem Gitter

Frage also: Gibt es eine reellwertige ZV X, für die obiges ($P(X \geq x) = e^{-x}$) als Identität gilt?

Bsp. 1.22, Forts.

$$P(pW \geq x) = (1-p)^{\lceil x/p \rceil} \approx (1-p)^{x/p} \approx e^{-x}$$

Interpretation z.B.: via Wartezeiten auf sehr fein zeitdiskretisiertem Gitter

Frage also: Gibt es eine reellwertige ZV X, für die obiges ($P(X \geq x) = e^{-x}$) als Identität gilt?

Ja, wie wir sehen werden.

Bsp. 1.22, Forts.

$$P(pW \geq x) = (1-p)^{\lceil x/p \rceil} \approx (1-p)^{x/p} \approx e^{-x}$$

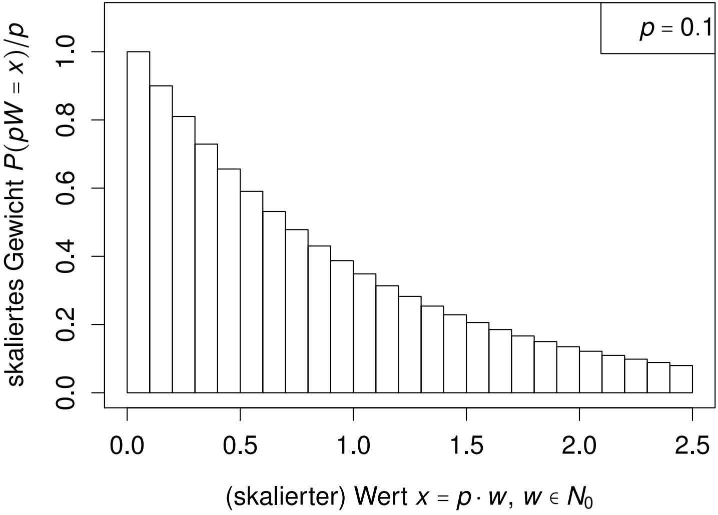
Interpretation z.B.: via Wartezeiten auf sehr fein zeitdiskretisiertem Gitter

Frage also: Gibt es eine reellwertige ZV X, für die obiges ($P(X \geq x) = e^{-x}$) als Identität gilt?

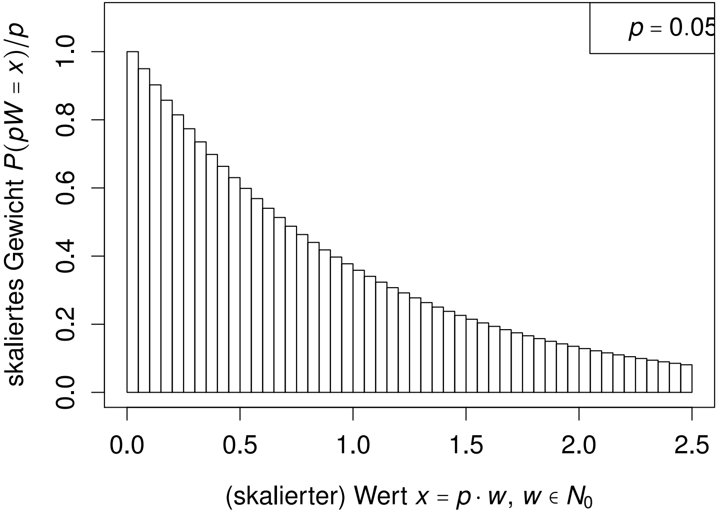
Ja, wie wir sehen werden.

In den folgenden Bildern skalieren wir die Höhe der Balken so, dass die Fläche des Balkens bei $x = pw$ gerade $P(pW = x) = p \cdot P(pW = x)/p$ entspricht.

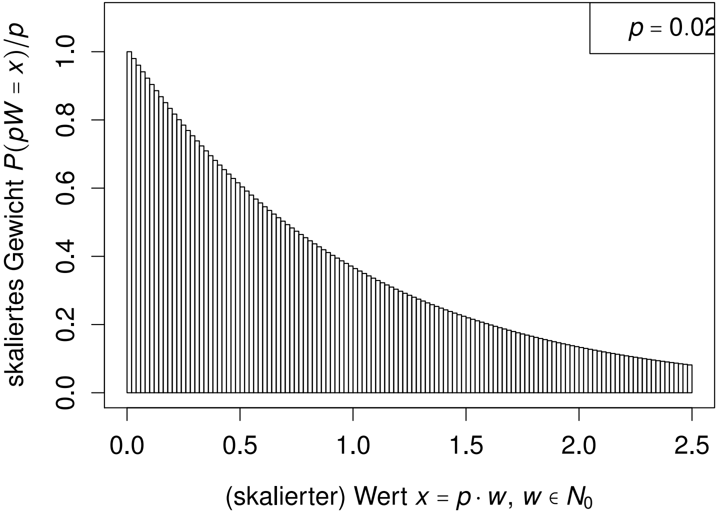
Bsp. 1.22: Illustration



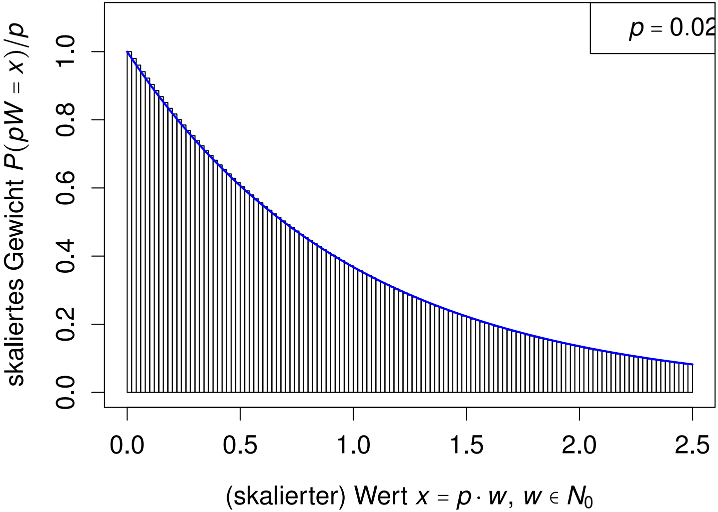
Bsp. 1.22: Illustration



Bsp. 1.22: Illustration



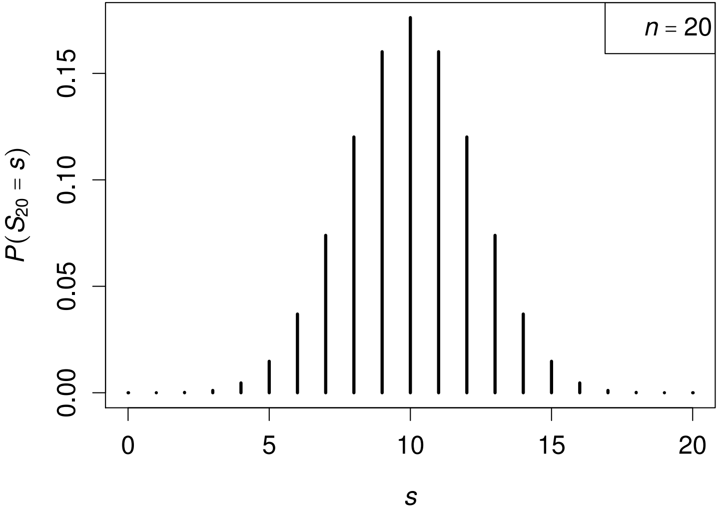
Bsp. 1.22: Illustration (die blaue Kurve ist e^{-x})



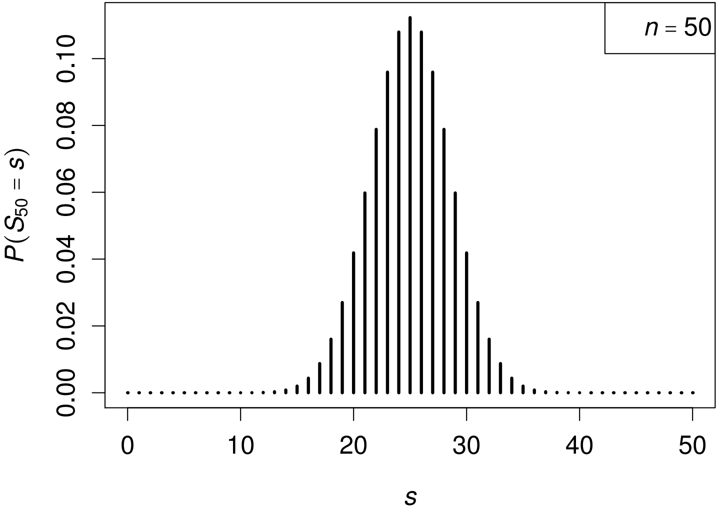
Beispiel 1.23 (Approximation einer Normalverteilung durch reskalierte binomialverteilte ZVn)

Betrachten wir für $S_n \sim \mathrm{Bin}_{n,1/2}$ (und verschiedene n) die Verteilungsgewichte, so beobachten wir, dass sich die „Form stabilisiert“.

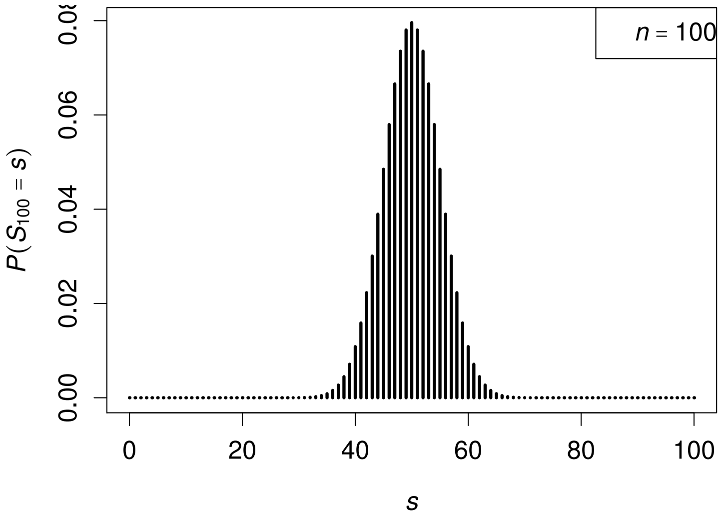
Bsp. 1.23: Illustration



Bsp. 1.23: Illustration



Bsp. 1.23: Illustration



Beispiel 1.23 (Approximation einer Normalverteilung durch reskalierte binomialverteilte ZVn)

Betrachten wir für $S_n \sim \mathrm{Bin}_{n,1/2}$ (und verschiedene n) die Verteilungsgewichte, so beobachten wir, dass sich die „Form stabilisiert".

Zentrieren und stauchen wir:

$$X_n := \frac{S_n - n/2}{\sqrt{n}}$$

Beispiel 1.23 (Approximation einer Normalverteilung durch reskalierte binomialverteilte ZVn)

Betrachten wir für $S_n \sim \mathrm{Bin}_{n,1/2}$ (und verschiedene n) die Verteilungsgewichte, so beobachten wir, dass sich die „Form stabilisiert“.

Zentrieren und stauchen wir:

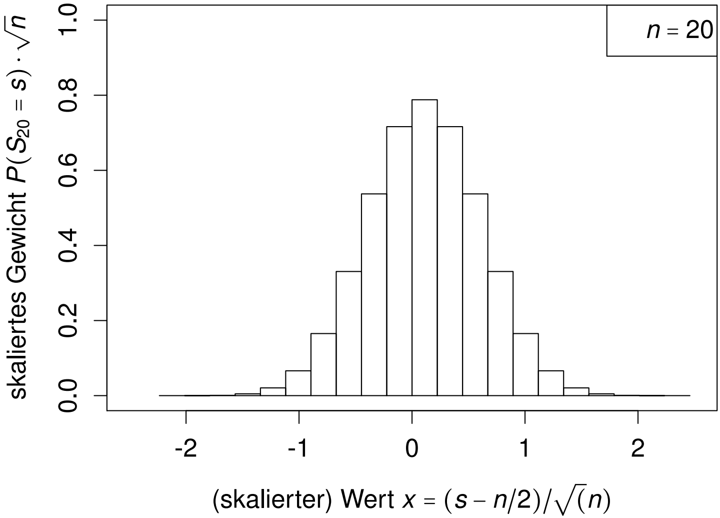
$$X_n := \frac{S_n - n/2}{\sqrt{n}}$$

und skalieren die „Balken“ so, dass die Fläche des Balkens bei $(s - n/2)/\sqrt{n}$ gerade

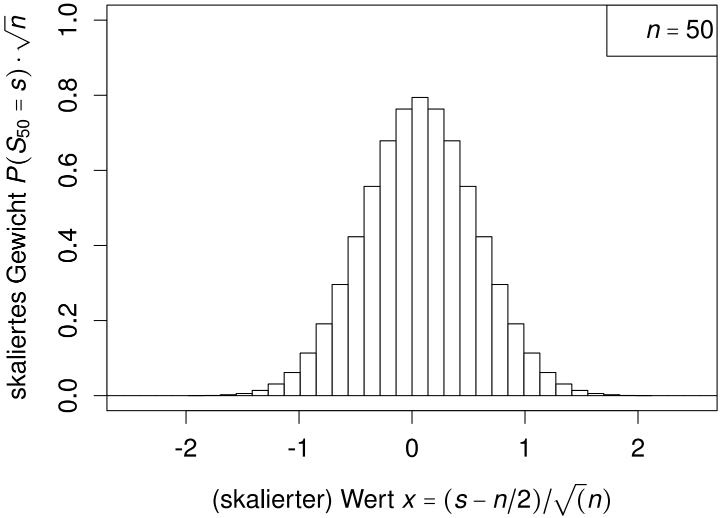
$$P(S_n = s) = \frac{1}{\sqrt{n}} \times \sqrt{n} P(S_n = s)$$

entspricht.

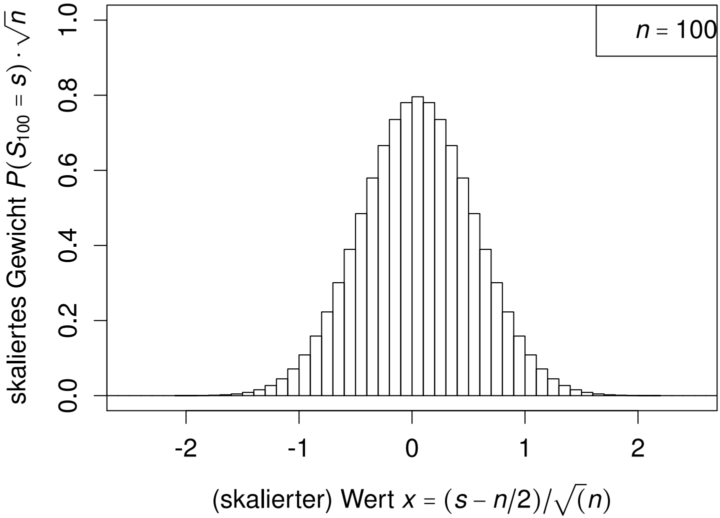
Bsp. 1.23: Illustration, 2



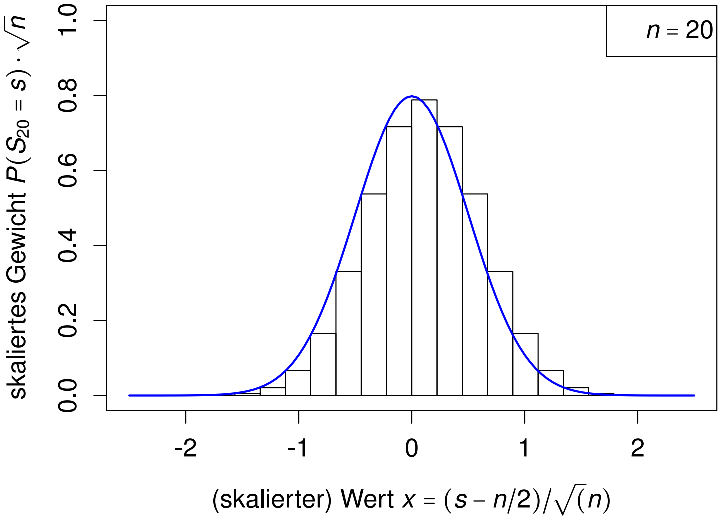
Bsp. 1.23: Illustration, 2



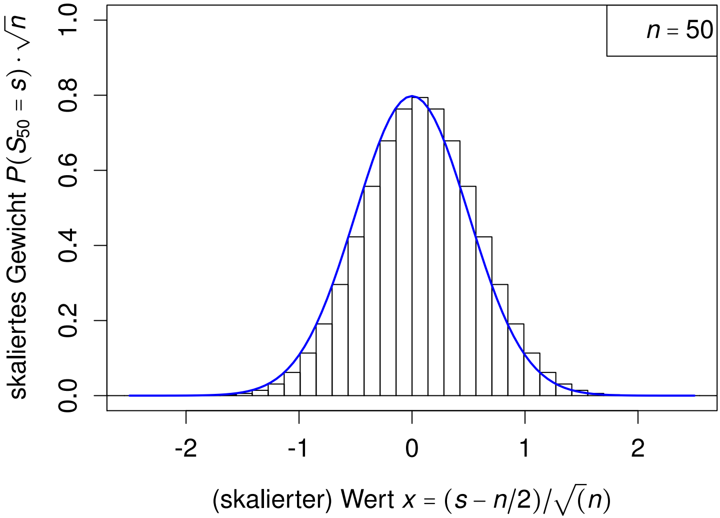
Bsp. 1.23: Illustration, 2



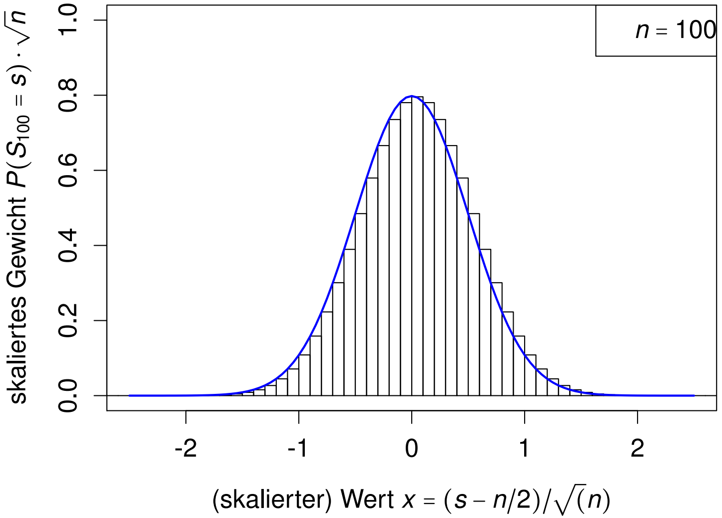
Bsp. 1.23: Illustration, 2



Bsp. 1.23: Illustration, 2



Bsp. 1.23: Illustration, 2



Definition 1.24

Sei X eine Zufallsvariable mit Werten in einem Intervall $S = [a, b] \subset \mathbb{R}$ mit $-\infty \leq a < b \leq \infty$ (im Fall $a > -\infty, b = \infty$ meinen wir $S = [a, \infty)$, etc.) und sei $f : S \to [0, \infty)$ eine integrierbare Funktion mit

$$\int_a^b f(x)\, dx = 1.$$

Definition 1.24

Sei X eine Zufallsvariable mit Werten in einem Intervall $S = [a, b] \subset \mathbb{R}$ mit $-\infty \le a < b \le \infty$ (im Fall $a > -\infty, b = \infty$ meinen wir $S = [a, \infty)$, etc.) und sei $f : S \to [0, \infty)$ eine integrierbare Funktion mit

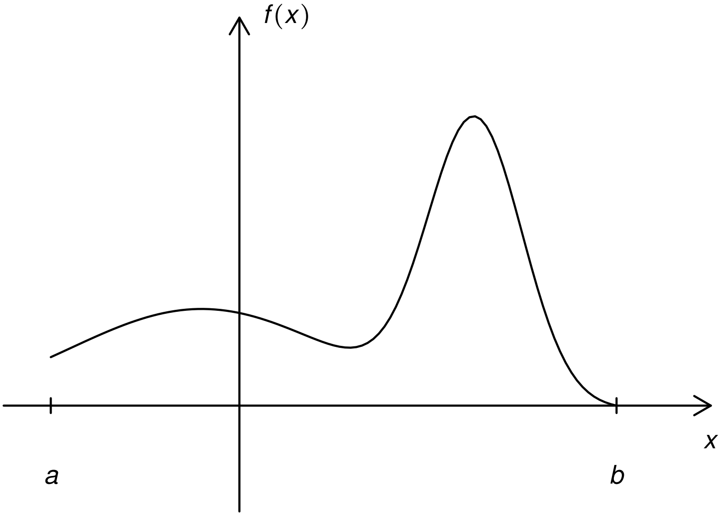
$$\int_a^b f(x)\,dx = 1.$$

X besitzt die *Dichte* (auch: *Wahrscheinlichkeitsdichte*) f, wenn gilt

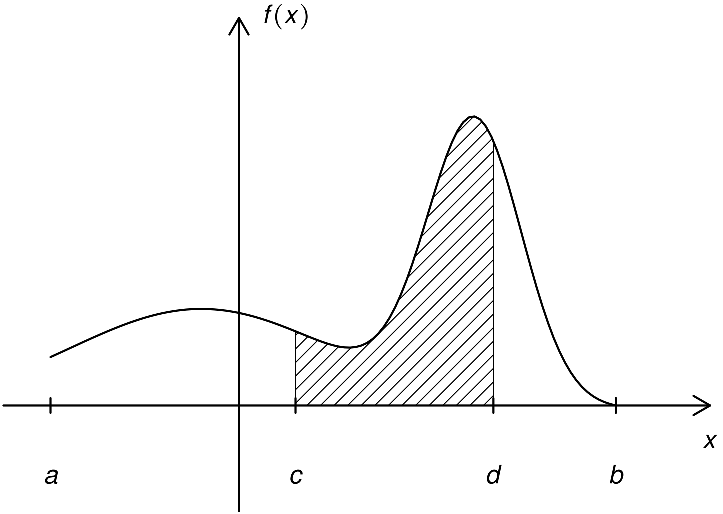
$$P\big(X \in [c, d]\big) = \int_c^d f(x)\,dx \quad \text{für jedes Teilintervall } [c, d] \subset S.$$

Wir notieren oft auch f_X für die Dichte einer ZV X (um den Bezug zu X zu betonen, speziell wenn wir mehrere ZVn zugleich ins Auge fassen).

X hat Dichte f



X hat Dichte f, so ist $P\big(X \in [c,d]\big) = \int_c^d f(x)\,dx$



Interpretation der Dichte: X ZV mit Dichte f_X, für $x \in \mathbb{R}$ und kleines $\delta > 0$ ist

$$P\big(X \in [x, x+\delta]\big) = \int_x^{x+\delta} f_X(a)\, da \approx \delta f_X(x)$$

(wörtlich zumindest für Stetigkeitsstellen x von f_X)

Interpretation der Dichte: X ZV mit Dichte f_X, für $x \in \mathbb{R}$ und kleines $\delta > 0$ ist

$$P\big(X \in [x, x+\delta]\big) = \int_x^{x+\delta} f_X(a)\, da \approx \delta f_X(x)$$

(wörtlich zumindest für Stetigkeitsstellen x von f_X), also

$$f_X(x) = \lim_{\delta \downarrow 0} \frac{1}{\delta} P\big(X \in [x, x+\delta]\big)$$

Interpretation der Dichte: X ZV mit Dichte f_X, für $x \in \mathbb{R}$ und kleines $\delta > 0$ ist

$$P\big(X \in [x, x+\delta]\big) = \int_x^{x+\delta} f_X(a)\,da \approx \delta f_X(x)$$

(wörtlich zumindest für Stetigkeitsstellen x von f_X), also

$$f_X(x) = \lim_{\delta \downarrow 0} \frac{1}{\delta} P\big(X \in [x, x+\delta]\big)$$

Man formuliert dies gelegentlich auch mit „infinitesimalen Größen“ als

$$P(X \in dx) = f_X(x)\,dx$$

(Dieser suggestive Ausdruck erhält einen Sinn im Sinne der „Standard-Analysis“, wenn man auf beiden Seiten x über ein Intervall $[c, d]$ integriert, dann erhält man Def. 1.24).

Bemerkung. Für eine ZV X mit Dichte f_X ist es – im Gegensatz zum Fall mit Gewichten – nicht besonders sinnvoll, nach der Wahrscheinlichkeit von Ereignissen $\{X = x\}$ für feste Punkte $x \in \mathbb{R}$ zu fragen, es gilt dann nämlich immer

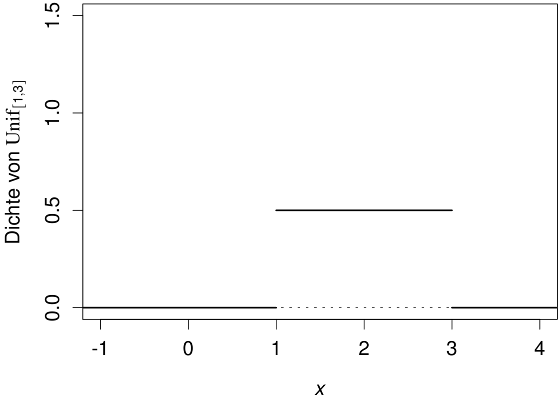
$$\begin{aligned} P(X = x) &= \lim_{\delta \downarrow 0} P\big(X \in [x, x+\delta]\big) \\ &= \lim_{\delta \downarrow 0} \int_x^{x+\delta} f_X(a)\, da = \int_x^x f_X(a)\, da = 0. \end{aligned}$$

Beispiel 1.25 (Einige „klassische“ eindimensionale Verteilungen mit Dichte)

1. (uniforme Verteilung) $a, b \in \mathbb{R}$, $a < b$.
$\mathrm{Unif}_{[a,b]}$ mit Dichte $\frac{1}{b-a} 1_{[a,b]}(x)$

Beispiel 1.25 (Einige „klassische“ eindimensionale Verteilungen mit Dichte)

1. (uniforme Verteilung) $a, b \in \mathbb{R}$, $a < b$.
$\mathrm{Unif}_{[a,b]}$ mit Dichte $\frac{1}{b-a} 1_{[a,b]}(x)$

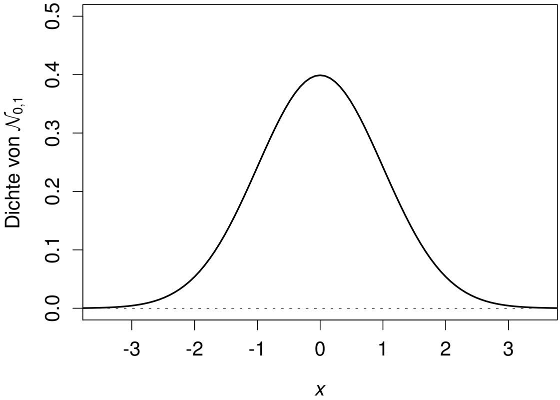


Beispiel 1.25

2. $\mu \in \mathbb{R}$, $\sigma > 0$. $\mathcal{N}_{\mu,\sigma^2}$ mit Dichte $\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2} \right)$ heißt Normalverteilung mit Mittelwert μ und Varianz σ^2.
$\mathcal{N}_{0,1}$ heißt die *Standardnormalverteilung*.

Beispiel 1.25

2. $\mu \in \mathbb{R}$, $\sigma > 0$. $\mathcal{N}_{\mu,\sigma^2}$ mit Dichte $\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ heißt Normalverteilung mit Mittelwert μ und Varianz σ^2.
$\mathcal{N}_{0,1}$ heißt die *Standardnormalverteilung*.

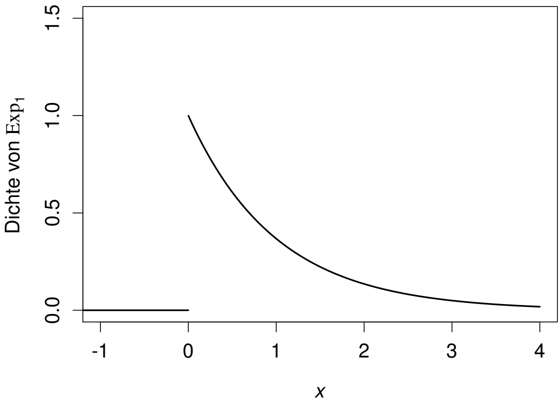


Beispiel 1.25

3. (Exponentialverteilung[en]) $\theta > 0$,
Exp_θ hat Dichte $\theta e^{-\theta x} 1_{[0,\infty)}(x)$

Beispiel 1.25

3. (Exponentialverteilung[en]) $\theta > 0$,
Exp_θ hat Dichte $\theta e^{-\theta x} 1_{[0,\infty)}(x)$



Definition 1.26 (Verteilungsfunktion)

Für eine Zufallsvariable X mit Werten in $\mathbb{R}$ (bzw. in einer Teilmenge $S \subset \mathbb{R}$) heißt die Funktion

$$F_X(x) := P(X \leq x), \quad x \in \mathbb{R} \tag{1}$$

die *Verteilungsfunktion* von X.

Definition 1.26 (Verteilungsfunktion)

Für eine Zufallsvariable X mit Werten in $\mathbb{R}$ (bzw. in einer Teilmenge $S \subset \mathbb{R}$) heißt die Funktion

$$F_X(x) := P(X \le x), \quad x \in \mathbb{R} \tag{1}$$

die *Verteilungsfunktion* von X.

Wenn X mit Wertebereich $S \subset \mathbb{R}$ die Dichte f_X besitzt, so gilt offenbar

$$F_X(x) = \int_{-\infty}^{x} f_X(a)\, da \tag{2}$$

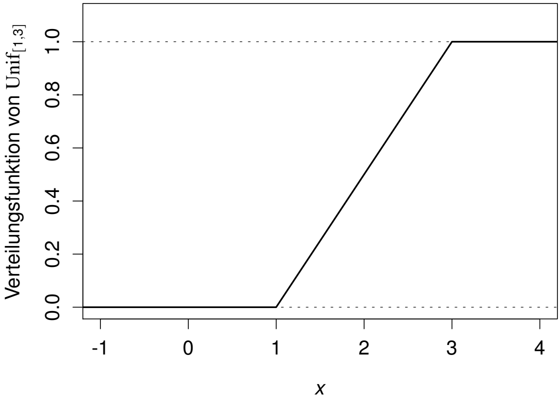
(mit Setzung $f_X(a) = 0$ für $a \notin S$, dem Wertebereich von X).

Beispiel 1.25 (Fortsetzung)

1. (uniforme Verteilung) $a, b \in \mathbb{R}$, $a < b$. $\mathrm{Unif}_{[a,b]}$ mit Dichte $\frac{1}{b-a} 1_{[a,b]}(x)$,
Verteilungsfunktion $\max\left\{\min\left\{\frac{x-a}{b-a}, 1\right\}, 0\right\}$

Beispiel 1.25 (Fortsetzung)

1. (uniforme Verteilung) $a, b \in \mathbb{R}$, $a < b$. $\mathrm{Unif}_{[a,b]}$ mit Dichte $\frac{1}{b-a} 1_{[a,b]}(x)$,
Verteilungsfunktion $\max\left\{\min\left\{\frac{x-a}{b-a}, 1\right\}, 0\right\}$



Beispiel 1.25 (Fortsetzung)

2. Die Verteilungsfunktion der *Standardnormalverteilung* $\mathcal{N}_{0,1}$,

$$\Phi(x) := F_{\mathcal{N}_{0,1}}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

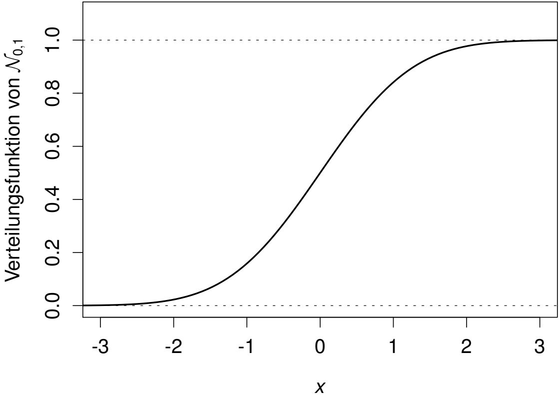
ist tabelliert bzw. in vielen Computerprogrammen implementiert (z.B. `pnorm` in R)

Beispiel 1.25 (Fortsetzung)

2. Die Verteilungsfunktion der *Standardnormalverteilung* $\mathcal{N}_{0,1}$,

$$\Phi(x) := F_{\mathcal{N}_{0,1}}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

ist tabelliert bzw. in vielen Computerprogrammen implementiert (z.B. `pnorm` in R)

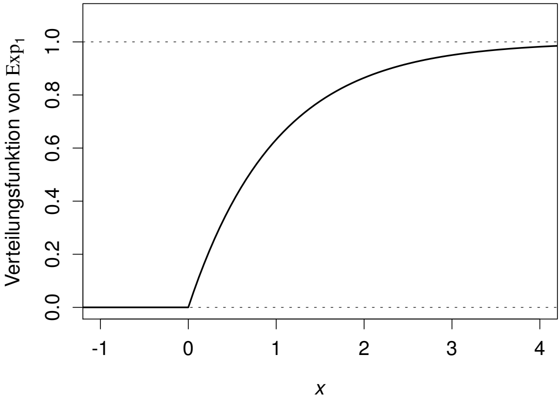


Beispiel 1.25 (Fortsetzung)

3. (Exponentialverteilung[en]) $\theta > 0$, Exp_θ hat Dichte $\theta e^{-\theta x} 1_{[0,\infty)}(x)$,
Verteilungsfunktion $F_{\mathrm{Exp}_\theta}(x) = (1 - e^{-\theta x}) 1_{[0,\infty)}(x)$

Beispiel 1.25 (Fortsetzung)

3. (Exponentialverteilung[en]) $\theta > 0$, Exp_θ hat Dichte $\theta e^{-\theta x} 1_{[0,\infty)}(x)$,
Verteilungsfunktion $F_{\mathrm{Exp}_\theta}(x) = (1 - e^{-\theta x}) 1_{[0,\infty)}(x)$



Bemerkung 1.27

1. Die Dichte / Verteilungsfunktion von X hängt nur von der Verteilung von X ab:

wenn $Y =^d X$ („Gleichheit in Verteilung“), also $P(X \in B) = P(Y \in B)$ für alle B gilt, so hat Y dieselbe (offenbar).

Wir sprechen daher oft auch kurz von der Dichte bzw. Verteilungsfunktion einer Wahrscheinlichkeitsverteilung auf $\mathbb{R}$, ohne die zugehörige ZV explizit zu machen.

Bemerkung 1.27

1. Die Dichte / Verteilungsfunktion von X hängt nur von der Verteilung von X ab:

wenn $Y =^d X$ („Gleichheit in Verteilung"), also $P(X \in B) = P(Y \in B)$ für alle B gilt, so hat Y dieselbe (offenbar).

Wir sprechen daher oft auch kurz von der Dichte bzw. Verteilungsfunktion einer Wahrscheinlichkeitsverteilung auf $\mathbb{R}$, ohne die zugehörige ZV explizit zu machen.

2. Wenn X Dichte f_X und Verteilungsfunktion F_X besitzt, so ist

$$\frac{d}{dx} F_X(x) = \frac{d}{dx} \int_{-\infty}^{x} f_X(a)\, da = f_X(x)$$

(zumindest an Stetigkeitspunkten von f_X)

Bemerkung 1.27

3. X ZV mit Werten in $S \subset \mathbb{R}$ mit Verteilungsfunktion F_X, $c < d$, so ist

$$P\big(X \in (c,d]\big) = P(X \le d) - P(X \le c) = F(d) - F(c)$$

Bemerkung 1.27

3. X ZV mit Werten in $S \subset \mathbb{R}$ mit Verteilungsfunktion F_X, $c < d$, so ist

$$P\big(X \in (c,d]\big) = P(X \le d) - P(X \le c) = F(d) - F(c)$$

(und falls $P(X = c) = 0$, z.B. weil X eine Dichte besitzt, so ist natürlich auch $P\big(X \in [c,d]\big) = P(X = c) + P\big(X \in (c,d]\big) = F(d) - F(c)$).

Bemerkung 1.27

3. X ZV mit Werten in $S \subset \mathbb{R}$ mit Verteilungsfunktion F_X, $c < d$, so ist

$$P\big(X \in (c,d]\big) = P(X \le d) - P(X \le c) = F(d) - F(c)$$

(und falls $P(X = c) = 0$, z.B. weil X eine Dichte besitzt, so ist natürlich auch $P\big(X \in [c,d]\big) = P(X = c) + P\big(X \in (c,d]\big) = F(d) - F(c)$).

Für $B = \bigcup_{i=1}^{n} (c_i, d_i]$ mit $c_1 < d_1 < c_2 < d_2 < \cdots < c_{n-1} < d_n$ ist (mit Eigenschaft (A) aus Def. 1.7)

$$P(X \in B) = \sum_{i=1}^{n} P\big(X \in (c_i, d_i]\big) = \sum_{i=1}^{n} \big(F_X(d_i) - F_X(c_i)\big)$$

Bemerkung 1.27

3. X ZV mit Werten in $S \subset \mathbb{R}$ mit Verteilungsfunktion F_X, $c < d$, so ist

$$P\big(X \in (c,d]\big) = P(X \le d) - P(X \le c) = F(d) - F(c)$$

(und falls $P(X = c) = 0$, z.B. weil X eine Dichte besitzt, so ist natürlich auch $P\big(X \in [c,d]\big) = P(X = c) + P\big(X \in (c,d]\big) = F(d) - F(c)$).

Für $B = \bigcup_{i=1}^{n} (c_i, d_i]$ mit $c_1 < d_1 < c_2 < d_2 < \cdots < c_{n-1} < d_n$ ist (mit Eigenschaft (A) aus Def. 1.7)

$$P(X \in B) = \sum_{i=1}^{n} P\big(X \in (c_i, d_i]\big) = \sum_{i=1}^{n} \big(F_X(d_i) - F_X(c_i)\big)$$

(und „allgemeine" Mengen $B \subset \mathbb{R}$ können auf diese Weise approximiert werden).

In diesem Sinne „weiß F_X alles" über die Verteilung von X.

Bemerkung 1.27

4. (Bezug zum diskreten Fall). Sei X ZV mit (höchstens) abzählbarem Wertebereich $S = \{x_1, x_2, \dots\} \subset \mathbb{R}$ und Gewichten $\rho_X(x_n)$ wie in in Def. 1.11, d.h.

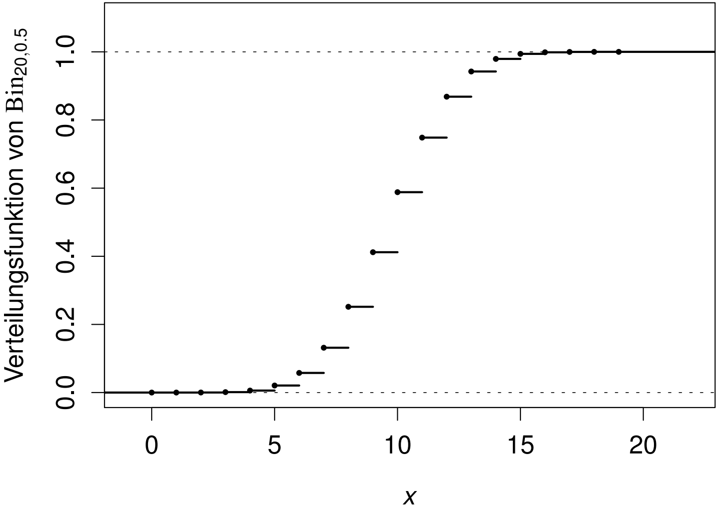
$$P(X \in B) = \sum_{n:\, x_n \in B} \rho_X(x_n),$$

dann ergibt sich als Verteilungsfunktion

$$F_X(x) = \sum_{n:\, x_n \le x} \rho_X(x_n).$$

(Diese ist stückweise konstant mit [höchstens] abzählbar vielen Sprüngen.)

Beispiel: Verteilungsfunktion einer Binomialverteilung



Bemerkung 1.27

5. Stets ist F_X nicht-fallend und rechtsstetig (wenn X eine Dichte besitzt, so ist F_X stetig) mit

$$\lim_{x \to \infty} F_X(x) = 1, \qquad \lim_{x \to -\infty} F_X(x) = 0.$$

[Die Eigenschaften folgen aus Prop. 1.8:
Für $x < y$ ist $\{X \le x\} \subset \{X \le y\}$ und somit $F_X(x) \le F_X(y)$; seien $x_n \ge x$ mit $x_n \searrow x$, so gilt $\{X \le x_n\} \searrow \{X \le x\}$ für $n \to \infty$ und daher auch
$F_X(x_n) = \searrow F_X(x)$, d.h. F_P ist rechtsstetig. Analog gilt
$F_P(x_n) = P(X_n \le x_n) \nearrow P(\mathbb{R}) = 1$ für $x_n \nearrow \infty$ und
$F_X(x_n) = P(X \le x_n) \searrow P(X \in \varnothing) = 0$ für $x_n \searrow -\infty$]

Bemerkung 1.27

5. Stets ist F_X nicht-fallend und rechtsstetig (wenn X eine Dichte besitzt, so ist F_X stetig) mit

$$\lim_{x\to\infty} F_X(x) = 1, \qquad \lim_{x\to-\infty} F_X(x) = 0.$$

[Die Eigenschaften folgen aus Prop. 1.8:
Für $x < y$ ist $\{X \le x\} \subset \{X \le y\}$ und somit $F_X(x) \le F_X(y)$; seien $x_n \ge x$ mit $x_n \searrow x$, so gilt $\{X \le x_n\} \searrow \{X \le x\}$ für $n \to \infty$ und daher auch
$F_X(x_n) =\searrow F_X(x)$, d.h. F_P ist rechtsstetig. Analog gilt
$F_P(x_n) = P(X_n \le x_n) \nearrow P(\mathbb{R}) = 1$ für $x_n \nearrow \infty$ und
$F_X(x_n) = P(X \le x_n) \searrow P(X \in \varnothing) = 0$ für $x_n \searrow -\infty$]

6. Umgekehrt gibt es zu jeder Funktion $F : \mathbb{R} \to [0,1]$ mit den Eigenschaften aus Def. 1.26, 5. eine ZV X mit $F_X = F$. (Wir kommen darauf zurück, siehe Beob. 1.29 unten.)

Definition 1.28

Die (verallgemeinerte) inverse Funktion von F_X,

$$F_X^{-1}(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad t \in [0,1]$$

(mit Setzung $\inf \emptyset = +\infty$) heißt auch die *Quantilfunktion* von X.

(Beachte, dass die so definierte Funktion F_X^{-1} linksstetig ist.)

Definition 1.28

Die (verallgemeinerte) inverse Funktion von F_X,

$$F_X^{-1}(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad t \in [0,1]$$

(mit Setzung $\inf \varnothing = +\infty$) heißt auch die *Quantilfunktion* von X.

(Beachte, dass die so definierte Funktion F_X^{-1} linksstetig ist.)

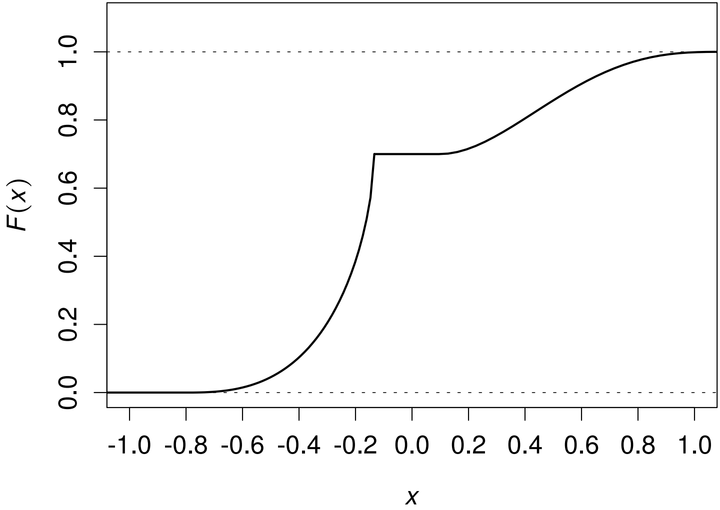
Mit dieser Definition ergibt sich für $x \in \mathbb{R}$, $t \in [0,1]$ die Beziehung

$$F_X^{-1}(t) \leq x \iff t \leq F_X(x).$$

(In der Literatur gibt es leicht verschiedene Definitionen der „Quantilfunktion“, man prüfe ggfs. jeweils die verwendete Konvention.)

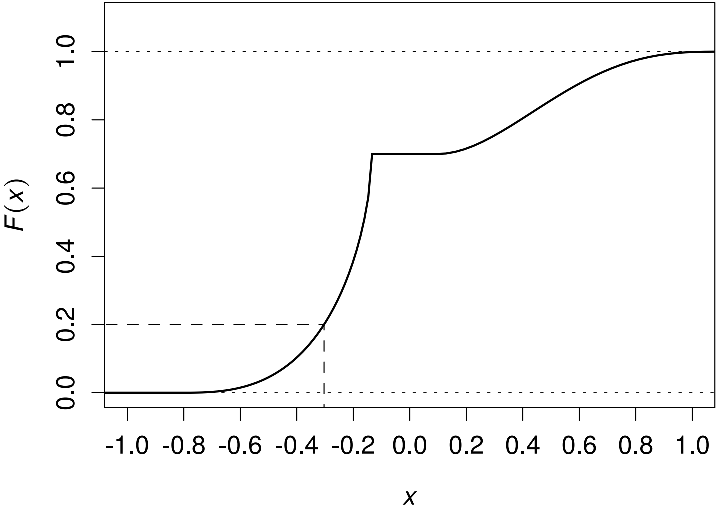
Beispiel: Ablesen von F^{-1}

$$F_X^{-1}(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad t \in [0,1]$$



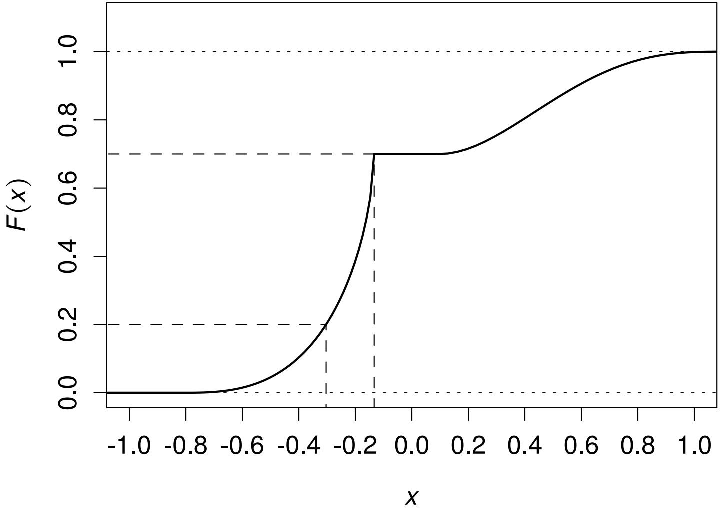
Beispiel: Ablesen von F^{-1}

$$F_X^{-1}(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad t \in [0, 1]$$

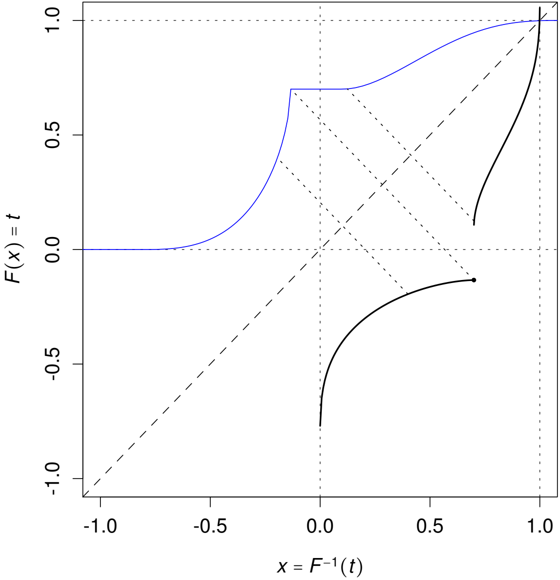


Beispiel: Ablesen von F^{-1}

$$F_X^{-1}(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\}, \quad t \in [0, 1]$$



Erinnerung: Inverse via Spiegelung an der Diagonalen



Beobachtung 1.29 (Erzeugung reeller ZVn mit vorgegebener Verteilung)

Sei $F : \mathbb{R} \to [0, 1]$ *eine Verteilungsfunktion ,*

$$F^{-1}(t) := \inf\left\{x \in \mathbb{R} : F(x) \geq t\right\}, \quad t \in [0, 1]$$

die inverse Verteilungsfunktion oder Quantilfunktion (aus Def. 1.28), U reelle ZV, $U \sim \mathrm{Unif}_{[0,1]}$*, dann hat*

$$X := F^{-1}(U)$$

die Verteilungsfunktion $F_X = F$*.*

Beobachtung 1.29 (Erzeugung reeller ZVn mit vorgegebener Verteilung)

Sei $F : \mathbb{R} \to [0,1]$ *eine Verteilungsfunktion ,*

$$F^{-1}(t) := \inf\left\{x \in \mathbb{R} : F(x) \geq t\right\}, \quad t \in [0,1]$$

die inverse Verteilungsfunktion oder Quantilfunktion (aus Def. 1.28), U reelle ZV, $U \sim \mathrm{Unif}_{[0,1]}$, *dann hat*

$$X := F^{-1}(U)$$

die Verteilungsfunktion $F_X = F$.

Beweis. Es gilt $F^{-1}(t) \leq x \iff t \leq F(x)$, somit ist für $x \in \mathbb{R}$

$$\begin{aligned} P(X \leq x) &= P(F^{-1}(U) \leq x) = P(U \leq F(x)) \\ &= P(0 \leq U \leq F(x)) = F(x) - 0 = F(x). \end{aligned}$$

Beispiel 1.30

1. Exp_θ hat Verteilungsfunktion
$F_{\mathrm{Exp}_\theta}(x) = (1 - e^{-\theta x})1_{[0,\infty)}(x)$ mit inverser Funktion
$F^{-1}_{\mathrm{Exp}_\theta}(t) = -\frac{1}{\theta}\log(1-t)$, also ist $-\dfrac{1}{\theta}\log(1-U) \sim \mathrm{Exp}_\theta$
(und natürlich ebenso $-\dfrac{1}{\theta}\log(U)$)

Beispiel 1.30

1. Exp_θ hat Verteilungsfunktion
$F_{\mathrm{Exp}_\theta}(x) = (1 - e^{-\theta x})1_{[0,\infty)}(x)$ mit inverser Funktion
$F^{-1}_{\mathrm{Exp}_\theta}(t) = -\frac{1}{\theta}\log(1-t)$, also ist $-\dfrac{1}{\theta}\log(1-U) \sim \mathrm{Exp}_\theta$
(und natürlich ebenso $-\dfrac{1}{\theta}\log(U)$)

2. $p(k), k \in \mathbb{N}_0$ Wahrscheinlichkeitsgewichte,
$F(x) = \sum_{k\,:\,k\le x} p(k)$ zugehörige Verteilungsfunktion (vgl. Bem. 1.27, 4.), so hat

$$X := \min\left\{n \in \mathbb{N}_0 : \sum_{k=0}^{n} p(k) \ge U\right\}$$

die Gewichte $(p(k))_{k\in\mathbb{N}_0}$.

(Dies ist etwa eine Möglichkeit, eine Poisson-verteilte ZV zu simulieren.)

Bericht: Verteilungen in R.

R verwendet folgende Namenskonvention: Wenn `name` für eine Verteilung steht, so ist

- `dname` die Dichte- bzw. Gewichtsfunktion von `name`
- `pname` die Verteilungsfunktion von `name` (“`p`” steht für “probability distribution function”)
- `qname` die Quantilfunktion von `name`
- `rname` simuliert gemäß `name` (“`r`” steht für “random”)

Beispiele:

- Uniforme Verteilung $\mathrm{Unif}_{[a,b]}$:
  `[d|p|q|r]unif(...,min=`a`,max=`b`)`
- Normalverteilung $\mathcal{N}_{\mu,\sigma^2}$:
  `[d|p|q|r]norm(...,mean=`μ`,sd=`σ`)`
  (beachte: R parametrisiert mit σ, nicht mit σ^2)
- Exponentialverteilung Exp_λ:
  `[d|p|q|r]exp(...,rate=`λ`)`
- Poissonverteilung Poi_λ:
  `[d|p|q|r]pois(...,lambda=`λ`)`
- Binomialverteilung $\mathrm{Bin}_{n,p}$:
  `[d|p|q|r]binom(...,size=`n`,prob=`p`)`

(Siehe auch die Online-Hilfe in R.)

Transformation von Dichten

Beobachtung 1.31

X habe Dichte f_X, sei $a > 0$, $b \in \mathbb{R}$, so hat $Y := aX + b$ die Dichte

$$f_Y(y) = \frac{1}{a} f_X\big((y - b)/a\big)$$

Transformation von Dichten

Beobachtung 1.31

X habe Dichte f_X, sei $a > 0$, $b \in \mathbb{R}$, so hat $Y := aX + b$ die Dichte

$$f_Y(y) = \frac{1}{a} f_X\big((y-b)/a\big)$$

denn

$$P(Y \le y) = P\big(X \le \tfrac{y-b}{a}\big) = \int_{-\infty}^{(y-b)/a} f_X(x)\,dx$$
$$= \int_{-\infty}^{y} f_X\big(\tfrac{z-b}{a}\big)\frac{1}{a}dz$$

wir substituieren $x = (z-b)/a$ ($\Leftrightarrow z = ax + b$), $\frac{dx}{dz} = 1/a$.

(Beachte: $Y \in [y, y+\delta] \iff X \in \big[\frac{y-b}{a}, \frac{y-b}{a} + \frac{\delta}{a}\big]$.)

[Ein Analogon zu Beob. 1.11 für den Fall mit Dichte]

Beispiel 1.32

1. $X \sim \mathcal{N}_{0,1}$, $\mu \in \mathbb{R}$, $\sigma > 0$, so ist $Y := \sigma X + \mu \sim \mathcal{N}_{\mu,\sigma^2}$
Insbesondere gilt $\mathcal{N}_{\mu,\sigma^2}((-\infty, x]) = \Phi((x-\mu)/\sigma)$

Beispiel 1.32

1. $X \sim \mathcal{N}_{0,1}$, $\mu \in \mathbb{R}$, $\sigma > 0$, so ist $Y := \sigma X + \mu \sim \mathcal{N}_{\mu,\sigma^2}$
Insbesondere gilt $\mathcal{N}_{\mu,\sigma^2}((-\infty, x]) = \Phi((x-\mu)/\sigma)$

2. $X \sim \mathrm{Exp}_1$, $a > 0$, so hat $Y := aX$ die Dichte $\frac{1}{a}e^{-x/a}$ (d.h. $Y \sim \mathrm{Exp}_{1/a}$)

[ggf. Details an der Tafel]

Transformation von Dichten, 2

Proposition 1.33 (Allgemeine Dichtetransformation im Fall $\mathbb{R}^1$)

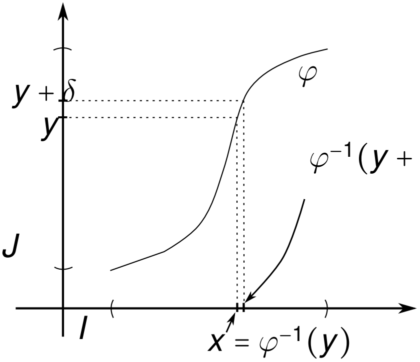
X reelle ZV mit Dichte f_X, d.h. $F_X(x) = \int_{-\infty}^{x} f_X(z)\,dz$, $I \subset \mathbb{R}$ (möglicherweise unbeschränktes) offenes Intervall mit $P(X \in I) = 1$, $J \subset I$, $\varphi : I \to J$ stetig differenzierbar, bijektiv. Dann hat $Y := \varphi(X)$ die Dichte

$$f_Y(y) = \begin{cases} \dfrac{f_X\big(\varphi^{-1}(y)\big)}{\big|\varphi'\big(\varphi^{-1}(y)\big)\big|}, & y \in J, \\ 0, & y \notin J. \end{cases}$$

[Bem.: Wenn φ nicht bijektiv ist, so wird $\varphi(X)$ i.A. keine Dichte besitzen, z.B. $X \sim \mathcal{N}_{0,1}$, $\varphi(x) = 1_{(0,\infty)}(x)$, so ist $\varphi(X) \sim \mathrm{Ber}_{1/2}$.]

Skizze zur Dichtetransformation (Prop. 1.33):

$Y = \varphi(X)$ hat Dichte $f_X\big(\varphi^{-1}(y)\big)/\varphi'\big(\varphi^{-1}(y)\big)$



$$\varphi^{-1}(y+\delta) \approx \varphi^{-1}(y) + \delta \cdot (\varphi^{-1})'(y)$$
$$= \varphi^{-1}(y) + \delta \frac{1}{\varphi'(\varphi^{-1}(y))}$$

[Rechnung ggf. an der Tafel]

Beispiel 1.34

1. $U \sim \mathrm{Unif}_{[0,1]}$, so hat $X := -\log(U)$ Dichte e^{-x} für $x \geq 0$ (wie wir in Bsp. 1.30 gesehen haben)

2. $U \sim \mathrm{Unif}_{[0,1]}$, $n \in \mathbb{N}$, so hat $Y := U^n$ Dichte $f_Y(y) = n^{-1} y^{1/n-1}$ (für $0 \leq y \leq 1$)

[Rechnung ggfs. an der Tafel]

Dichten im mehrdimensionalen Fall

Das multivariate Analogon zu Def. 1.24 ist folgende

Definition 1.35

Sei $f : \mathbb{R}^d \to \mathbb{R}_+$ eine (geeignet) integrierbare* Funktion mit

$$\int_{\mathbb{R}^d} f(x)\,dx = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_d)\,dx_1 \cdots dx_d = 1.$$

Eine Zufallsvariable X mit Werten in $\mathbb{R}^d$ besitzt die *Dichte* f, wenn für (geeignete) Teilmengen $A \subset \mathbb{R}^d$ gilt

$$\begin{aligned} P(X \in A) &= \int_A f(x)\,dx \\ &:= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} 1_A(x_1, \ldots, x_d) f(x_1, \ldots, x_d)\,dx_1 \cdots dx_d. \end{aligned}$$

---

*Wir denken hier z.B. an ein d-fach iteriertes Riemann-Integral.

Dichten im mehrdimensionalen Fall

Analog zum 1-dimensionalen Fall besitzt die Dichte f_X einer d-dimensionalen ZV X die Interpretation

$$P\Big(X \in [x_1, x_1 + \delta_1] \times [x_2, x_2 + \delta_2] \times \cdots \times [x_d + \delta_d]\Big)$$
$$\approx \delta_1 \delta_2 \cdots \delta_d \cdot f_X\big((x_1, \ldots, x_d)\big)$$

für $(x_1, \ldots, x_d) \in \mathbb{R}^d$ und $0 < \delta_1, \delta_2, \ldots, \delta_d \ll 1$.

Beispiel 1.36

1. Uniforme oder Laplace-Verteilung auf einem beschränkten Gebiet $S \subset \mathbb{R}^d$: X mit Dichte $f_X(x) = \frac{1}{\mathrm{vol}(S)} 1_S(x)$ erfüllt für $A \subset S$ (geeignet*)

$$P(X \in A) = \int_A \frac{1}{\mathrm{vol}(S)} 1_S(x)\, dx = \frac{\int_A 1\, dx}{\int_S 1\, dx} = \frac{\mathrm{vol}(A)}{\mathrm{vol}(S)}.$$

(Z.B. der zufällig gewählte Punkt Z aus Kapitel 0 kann als uniform auf $S = [0,1]^2$ verteilt modelliert werden.)

---

* in dem Sinne, dass ein d-dimensionales „Volumen“ $\mathrm{vol}(A)$ definierbar ist

Beispiel 1.36 (Forts.)

2. Die 2-dimensionale Standard-Normalverteilung hat Dichte $f(x,y) = \dfrac{1}{2\pi} \exp\big(-\tfrac{1}{2}(x^2+y^2)\big)$

Beispiel 1.36 (Forts.)

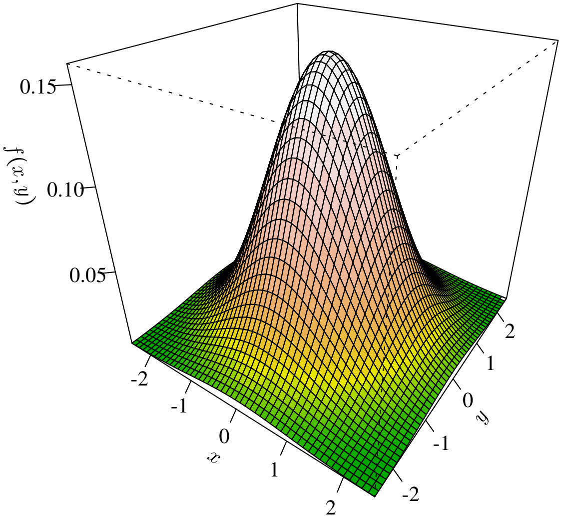
2. Die 2-dimensionale Standard-Normalverteilung hat Dichte $f(x,y) = \dfrac{1}{2\pi} \exp\Big(-\tfrac{1}{2}(x^2+y^2)\Big)$

3. Allgemeiner: Die d-dimensionale Standard-Normalverteilung hat Dichte

$$f(x_1,\ldots,x_d) = \frac{1}{(2\pi)^{d/2}} \exp\Big(-\frac{1}{2}(x_1^2+\cdots+x_d^2)\Big)$$

Die 2-dimensionale Standard-Normalverteilung hat Dichte

$$f(x,y) = \frac{1}{2\pi} \exp\left(-\tfrac{1}{2}(x^2 + y^2) \right)$$



Die 2-dimensionale Standard-Normalverteilung hat Dichte

$$f(x,y) = \frac{1}{2\pi} \exp\Big(-\tfrac{1}{2}(x^2 + y^2)\Big)$$