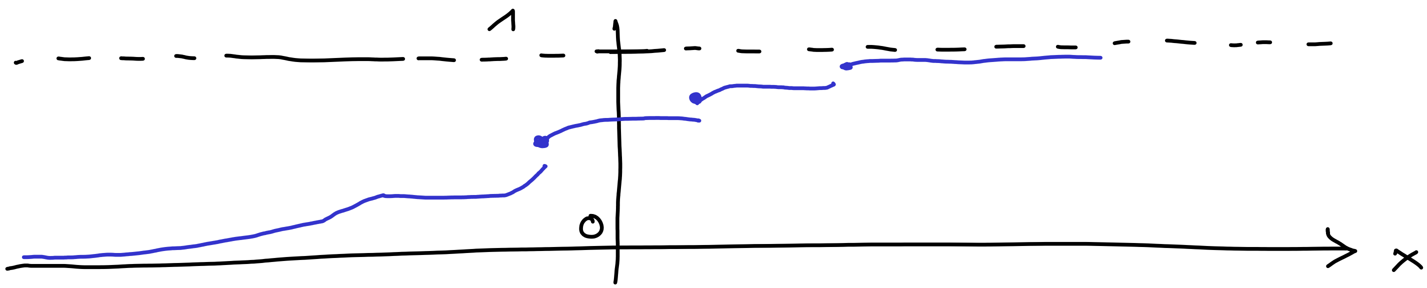
Beobachtung und Definition 1.11. Für ein Wahrscheinlichkeitsmaß P auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ ist die Funktion $F_P : \mathbb{R} \to [0,1]$, $F_P(x) := P\big((-\infty, x]\big)$ nicht-fallend und rechtsstetig mit

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

F_p heißt die *Verteilungsfunktion* von P.



Die Eigenschaften folgen aus Lemma 1.5, (1.5):
Für $x < y$ ist $(-\infty, x] \subset (-\infty, y]$ und somit $F_P(x) = P((-\infty, x]) \le P((-\infty, y]) = F_P(y)$;
seien $x_n \ge x$ mit $x_n \searrow x$, setze $A_n := (-\infty, x_n]$, $A := (-\infty, x]$, so gilt $A_n \searrow A$ für $n \to \infty$ und daher auch $F_P(x_n) = P(A_n) \searrow P(A) = F_P(A)$, d.h. F_P ist rechtsstetig.
Analog gilt $F_P(x_n) = P(A_n) \nearrow P(\mathbb{R}) = 1$ für $x_n \nearrow \infty$ und $F_P(x_n) = P(A_n) \searrow P(\varnothing) = 0$ für $x_n \searrow -\infty$.
Analog betrachtet man im Fall $d > 1$ für $(x_1, x_2, \dots, x_d) \in \mathbb{R}^d$ die Funktion

$$F_P(x_1, x_2, \dots, x_d) := P\Big((-\infty, x_1] \times (-\infty, x_2] \times \dots (-\infty, x_d]\Big),$$

die entsprechende Eigenschaften besitzt.

Bericht 1.12. Umgekehrt definiert jede Funktion $F : \mathbb{R} \to [0,1]$ (bzw. $F : \mathbb{R}^d \to [0,1]$) mit den Eigenschaften aus Def. 1.11 ein (eindeutiges) Wahrscheinlichkeitsmaß P mit $F = F_P$.

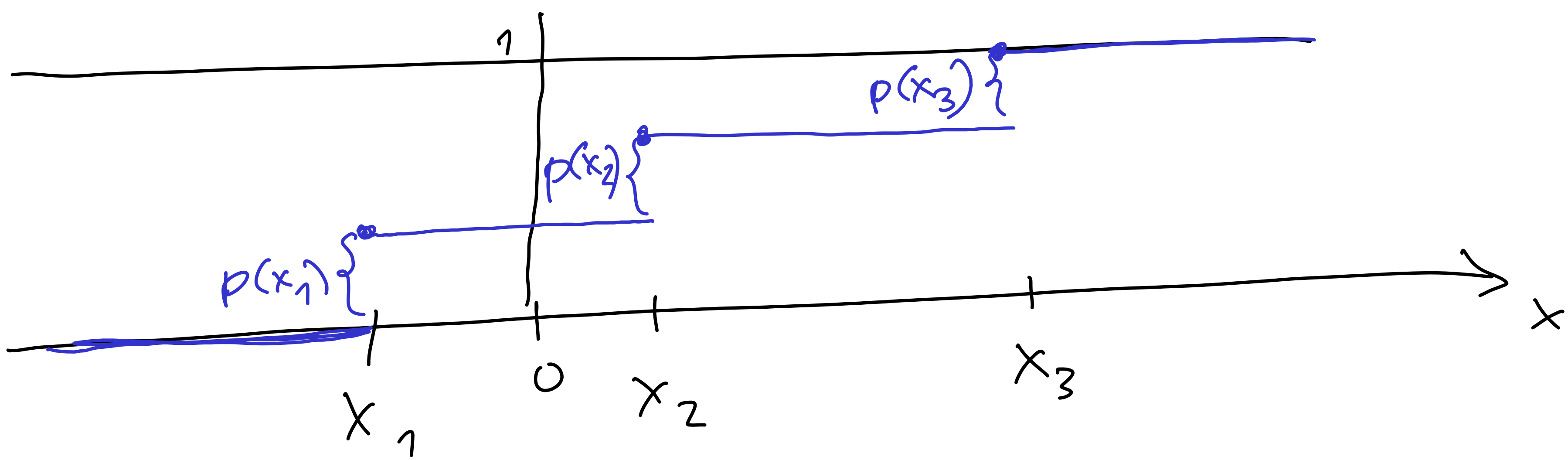
Bemerkung 1.13. 1 (Bezug zum diskreten Fall). Sei eine (höchstens) abzählbare Menge $\Omega' = \{x_1, x_2, \dots\} \subset \mathbb{R}$ (z.B. $\Omega' = \mathbb{N}$, $\Omega' = \mathbb{Z}$, ...) und ein diskretes W'maß P auf $(\Omega', 2^{\Omega'})$ mit Gewichten $p(\cdot)$ (wie in Bsp. 1.4) gegeben. Wir können P auch als W'maß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ auffassen via

$$P(A) := \sum_{n : x_n \in A} p(x_n),$$

dann ergibt sich als Verteilungsfunktion

$$F_P(x) = \sum_{n : x_n \leq x} p(x_n).$$

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



2. Sei P Wahrscheinlichkeitsmaß auf $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ mit Verteilungsfunktion F_P. Die (verallgemeinerte) inverse Funktion von F_P,

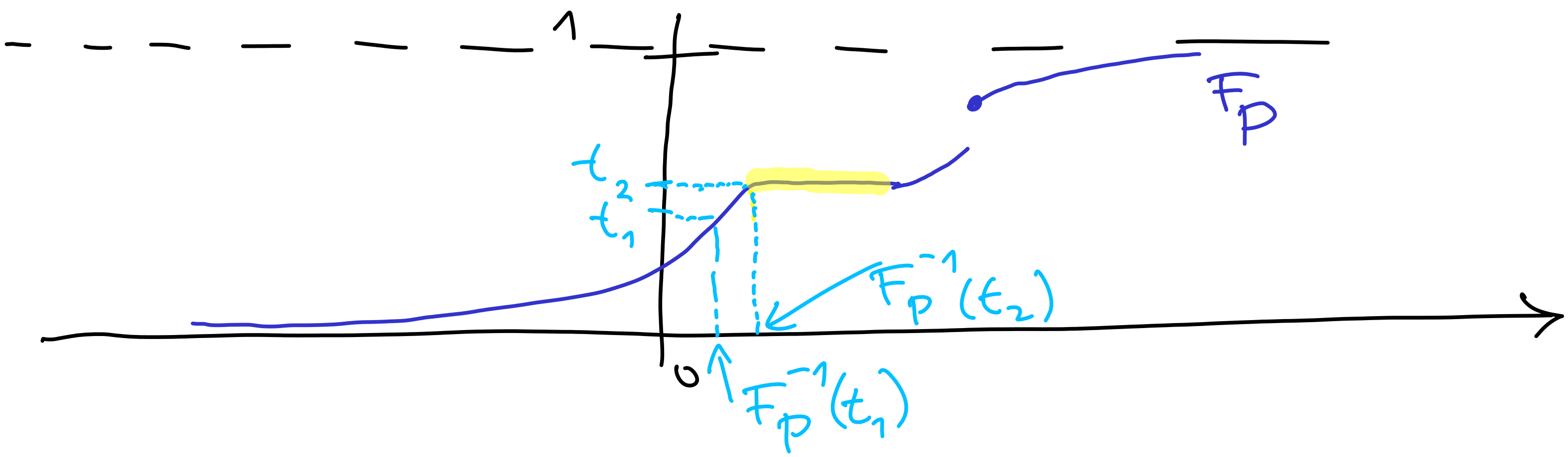
$$F_P^{-1}(t) := \inf\{x \in \mathbb{R} : F(x) \geq t\},$$

heißt auch die *Quantilfunktion* von P.

(Beachte, dass die so definierte Funktion F_P^{-1} linksstetig ist. Mit dieser Definition ergibt sich für $x \in \mathbb{R}$, $t \in [0,1]$ die Beziehung

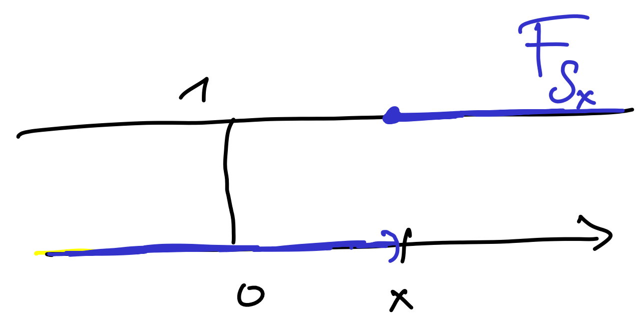
$$\boxed{F_P^{-1}(t) \leq x \iff t \leq F_P(x).}$$

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



Beispiel (Dirac-Maß)**.**

$$\delta_x(A) = \begin{cases} 1, & x \in A, \\ 0 & x \notin A. \end{cases}$$



(Im Fall $\Omega = \mathbb{R}$ ist $F_{\delta_x}(y) = 1(x \leq y)$.)

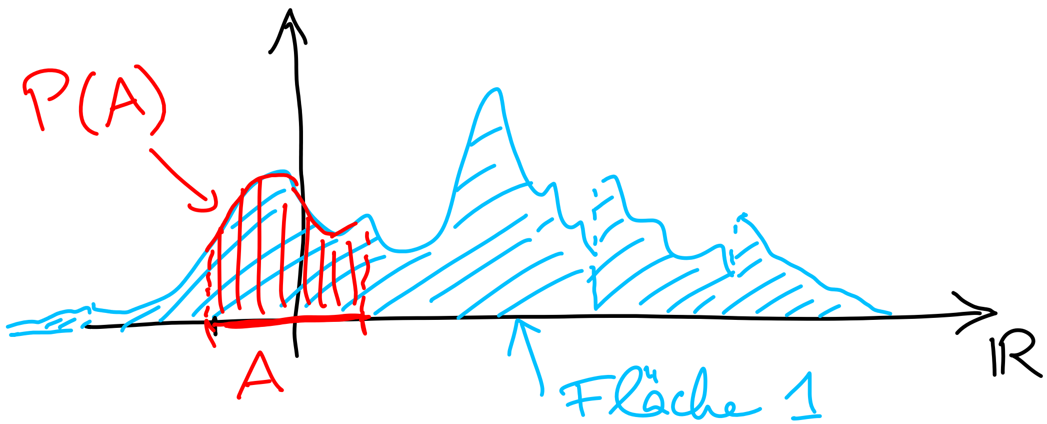
Beispiel 1.14 (Maße mit Dichten auf $\mathbb{R}$ (bzw. auf $\mathbb{R}^d$)). Sei $f_P : \mathbb{R} \to \mathbb{R}_+$ integrierbar[4] mit

$$\int_{\mathbb{R}} f_P(x)\,dx = 1.$$

Dann definiert[5]

$$P(A) := \int_A f_P(x)\,dx$$

ein Wahrscheinlichkeitsmaß, die Funktion f_P heißt die *Dichte* (auch: Wahrscheinlichkeitsdichte) von P.



Die Verteilungsfunktion

$$F_P(x) = P\big((-\infty, x]\big) = \int_{-\infty}^{x} f_P(y)\,dy$$

ist dann (zumindest an Stetigkeitsstellen von f_P) differenzierbar mit

$$\frac{d}{dx} F_P(x) = f_P(x)$$

[4] In einem mit den Vorkenntnissen der Hörer verträglichen Sinn: Wir werden nur Beispiele betrachten, in denen f_P wenigstens stückweise stetig ist, so dass man hier durchaus an das Riemann-Integral (oder auch ganz salopp an die „Fläche unter der Kurve") denken kann. Für einen Bericht zum Lebesgue-Integral z.B. [G, Tatsache 1.14].

[5] Wiederum hängt es vom verwendeten Integralbegriff ab, für welche Mengen A das Integral $\int_A f_P(x)\,dx$ sinnvoll definiert ist. Man verliert an dieser Stelle wenig, wenn man bei A etwa an eine endliche Vereinigung von Intervallen denkt.

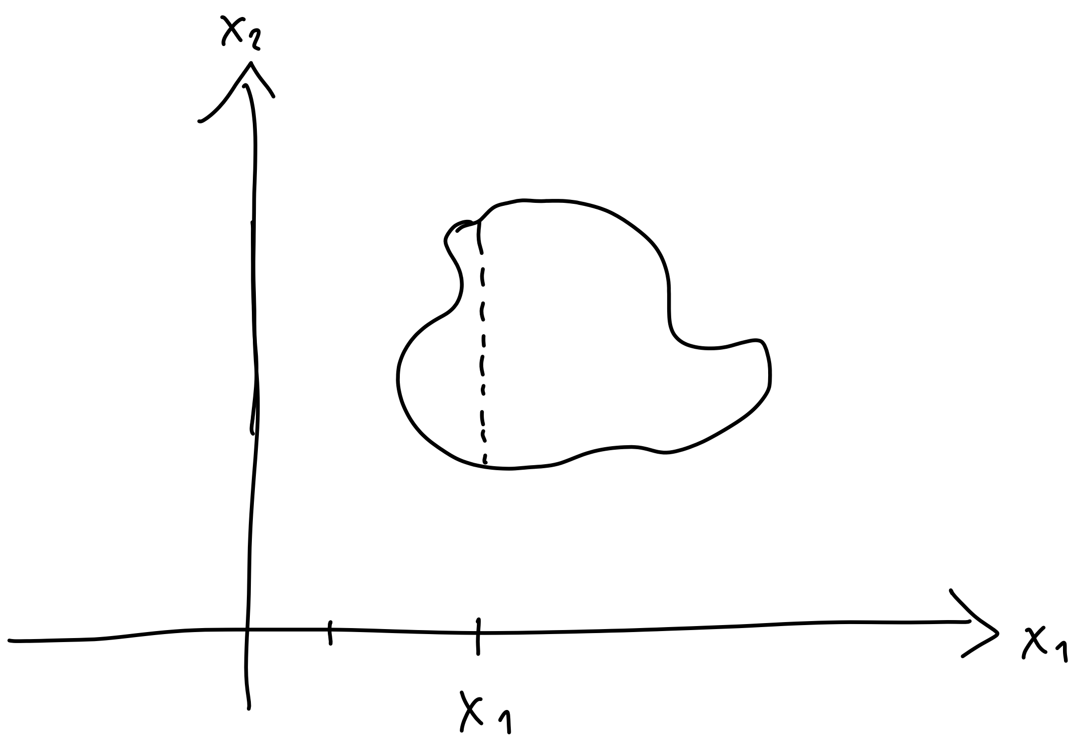
Analog definiert man für (geeignet) integrierbares $f_P : \mathbb{R}^d \to \mathbb{R}_+$ mit $\int_{\mathbb{R}^d} f_P(x)\,dx = 1$ durch

$$P(A) := \int_A f_P(x)\,dx$$

ein W'maß P auf $\mathbb{R}^d$ mit Dichte f_P. Wir denken an dieser Stelle wiederum an geeignet „gutartige" Funktionen f_P und Teilmengen $A \subset \mathbb{R}^d$, für die das Integral beispielsweise als iteriertes Riemann-Integral wohldefiniert ist, d.h.

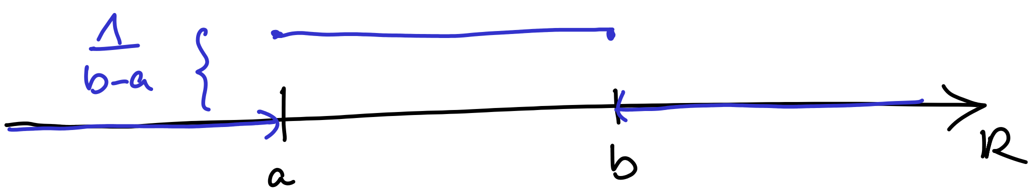
$$\int_A f_P(x)\,dx = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \mathbf{1}_A(x_1,\ldots,x_d) f_P(x_1,\ldots,x_d)\,dx_1 \cdots dx_d.$$

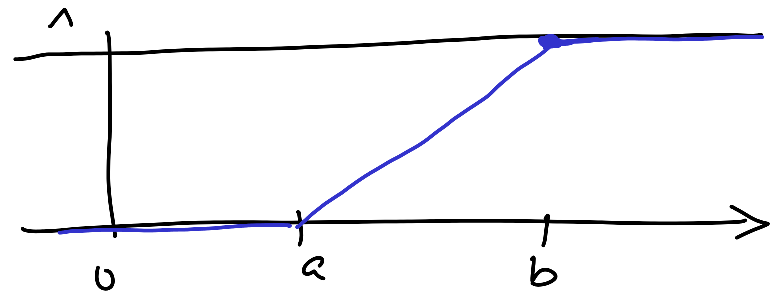
Dies wird für die im Rahmen der Vorlesung betrachteten Beispiele genügen.



Beispiel 1.15 („Klassische“ eindimensionale Verteilungen mit Dichte)**.**

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



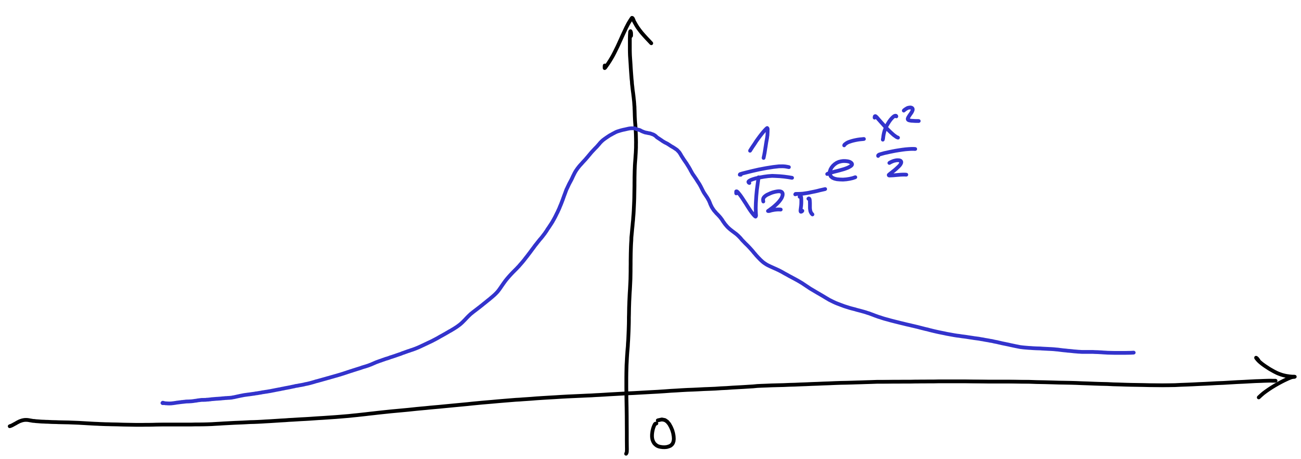


2. (Normalverteilung[en]) $\mu \in \mathbb{R}$, $\sigma > 0$. $\mathcal{N}_{\mu,\sigma^2}$ mit Dichte $\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$ heißt Normalverteilung mit Mittelwert μ und Varianz σ^2.

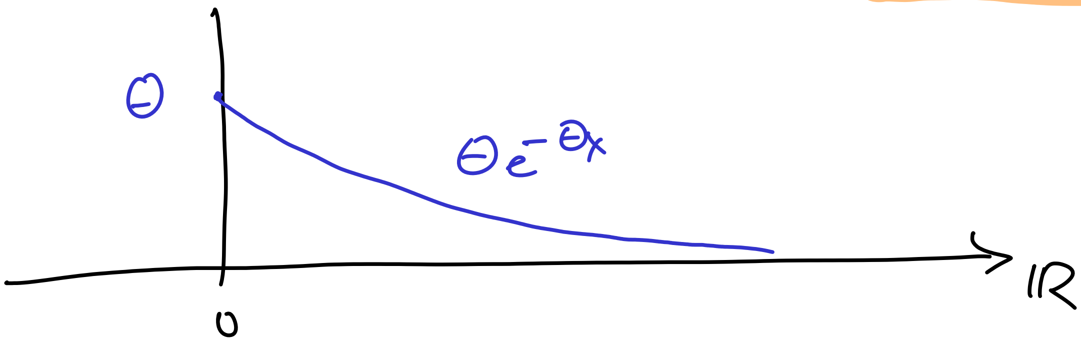
   $\mathcal{N}_{0,1}$ heißt die *Standardnormalverteilung*, die Verteilungsfunktion

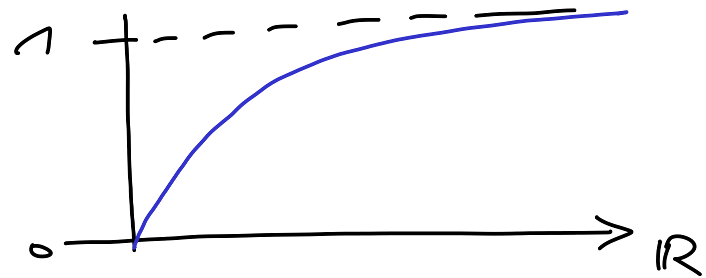
$$\Phi(x) := F_{\mathcal{N}_{0,1}}(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz$$

   ist tabelliert bzw. in vielen Computerprogrammen implementiert und es ist $\mathcal{N}_{\mu,\sigma^2}((-\infty, x]) = \Phi((x-\mu)/\sigma)$ (Übung).



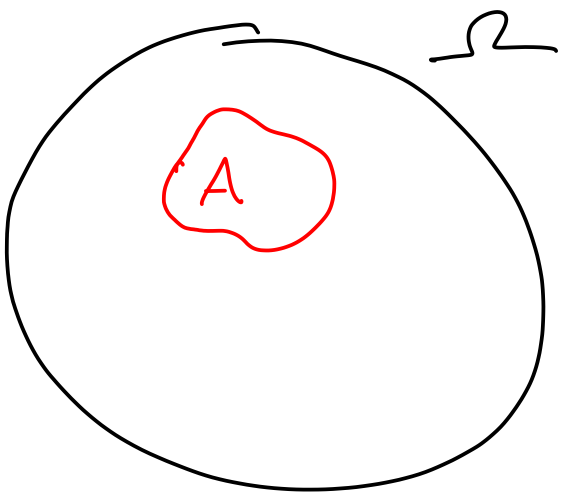
3. (Exponentialverteilung[en]) $\theta > 0$, Exp_θ hat Dichte $\theta e^{-\theta x}\mathbf{1}_{[0,\infty)}(x)$, Verteilungsfunktion $(1-e^{-\theta x})\mathbf{1}_{[0,\infty)}(x)$





Beispiel 1.16. Laplace-Verteilung auf einem beschränkten Gebiet $\Omega \subset \mathbb{R}^d$: Für $A \subset \Omega$ (geeignet[6]) ist

$$P(A) := \frac{\mathrm{vol}(A)}{\mathrm{vol}(\Omega)} = \frac{\int_A 1\,dx}{\int_\Omega 1\,dx}.$$



Dichte $\dfrac{\mathbb{1}_A(x)}{\mathrm{vol}(\Omega)}$

Indikatorfunktion:

$$\mathbb{1}_A(x) = \begin{cases} 1, & \text{falls } x \in A \\ 0, & \text{falls } x \notin A \end{cases}$$

[6]in dem Sinne, dass ein „Volumen" $\mathrm{vol}(A)$ definierbar ist

1.3 Klassische diskrete Verteilungen

Beispiel 1.17 (Urnenmodelle)**.** Eine Urne enthalte n (nummerierte) Kugel, wir ziehen zufällig k ($\le n$) heraus.

1. mit Zurücklegen, mit Beachtung der Reihenfolge:

$$\Omega_1 = \{(\omega_1, \dots, \omega_k) : \omega_1, \dots, \omega_k \in \{1, 2, \dots, n\}\} = \{1, 2, \dots, n\}^k,$$
$$P_1(\{(\omega_1, \dots, \omega_k)\}) = \frac{1}{n^k} \quad \left(= \frac{1}{|\Omega_1|}, \text{ vgl. auch Beispiel 1.8} \right)$$

2. ohne Zurücklegen, mit Beachtung der Reihenfolge:

$$\Omega_2 = \{(\omega_1, \dots, \omega_k) : \omega_1, \dots, \omega_k \in \{1, 2, \dots, n\}, \omega_i \neq \omega_j \text{ für } i \neq j\},$$
$$P_2(\{(\omega_1, \dots, \omega_k)\}) = \frac{1}{n \cdot (n-1) \cdots (n-k+1)} = \frac{(n-k)!}{n!} \quad \left(= \frac{1}{|\Omega_2|} \right)$$

3. ohne Zurücklegen, ohne Beachtung der Reihenfolge:

$$\Omega_3 = \{A \subset \{1, 2, \dots, n\} : |A| = k\},$$
$$P_3(\{A\}) = \frac{1}{\binom{n}{k}} = \frac{k!(n-k)!}{n!} \quad \left(= \frac{1}{|\Omega_3|}, \text{ denn es gibt } \binom{n}{k} \text{ versch. } k\text{-elementige Teilmengen} \right)$$

4. mit Zurücklegen, ohne Beachtung der Reihenfolge:

$$\Omega_4 = \{(\ell_1, \ell_2, \dots, \ell_n) \in \mathbb{N}_0^n : \ell_1 + \ell_2 + \dots + \ell_n = k\}$$

(ℓ_i gibt an, wie oft Kugel i gezogen wurde)

Für $(\ell_1, \ell_2, \dots, \ell_n) \in \Omega_4$ gibt es

$$\binom{k}{\ell_1, \ell_2, \dots, \ell_n} := \frac{k!}{\ell_1! \cdot \ell_2! \cdot \dots \cdot \ell_n!} \qquad \text{„Multinomialkoffizient“}$$

verschiedene $\omega = (\omega_1, \dots, \omega_k) \in \Omega_1$ mit Anz. $\omega = (\omega_1, \dots, \omega_k) \in \Omega_1$

mit : $\big|\{1 \le j \le k : \omega_j = i\}\big| = \ell_i \quad \text{für } i = 1, \dots, n.$ ist $\binom{k}{\ell_1, \ell_2, \dots, \ell_n}$

$$P_4\big(\{(\ell_1, \dots, \ell_n)\}\big) = \binom{k}{\ell_1, \ell_2, \dots, \ell_n}\Big(\frac{1}{n}\Big)^k, \quad (\ell_1, \dots, \ell_n) \in \Omega_4$$

$$\underbrace{\binom{k}{\ell_1}}_{\text{Wahlen für 1er}}\underbrace{\binom{k-\ell_1}{\ell_2}}_{\dots \text{ für 2er}}\binom{k-\ell_1-\ell_2}{\ell_3} \cdots \underbrace{\binom{k-\ell_1-\dots-\ell_{n-1}}{\ell_n}}_{\text{Wahlen für }n\text{-er}}$$

Bemerkung 1.18. Es gilt

$$|\Omega_4| = \binom{n+k-1}{k} \quad \left(= \binom{n+k-1}{n-1} \right)$$

Ein "Zähltrick": Lege k Kugeln und $n-1$ "Trennstäbe" – also insgesamt $n+k-1$ Objekte – in eine Reihe:

$$\underbrace{\bigcirc\cdots\bigcirc}_{\ell_1 \text{ Kugeln}}\Big|\underbrace{\bigcirc\bigcirc\cdots\bigcirc}_{\ell_2 \text{ Kugeln}}\Big|\underbrace{\;}_{\ell_3 = 0}\Big|\cdots\Big|\underbrace{\bigcirc\bigcirc\cdots\bigcirc}_{\ell_{n-1} \text{ Kugeln}}\Big|\underbrace{\bigcirc\cdots\bigcirc}_{\ell_n \text{ Kugeln}}$$

Insbesondere ist die Verteilung auf Ω_4 aus Beispiel 1.17, 4. nicht die uniforme.

Die uniforme Verteilung auf dem Ω_4 aus Beispiel 1.17, 4. heißt auch die „Bose-Einstein-Verteilung", die in Beispiel 1.17, 4. betrachtete Verteilung heißt die „Maxwell-Boltzmann-Verteilung".

Beispiel. Eine Hörsaalreihe habe n Plätze, darauf nehmen m $(\le n/2)$ Männer und $n-m$ Frauen rein zufällig Platz. Die Wahrscheinlichkeit, dass keine zwei Männer nebeneinander sitzen

$$= \frac{\binom{n-m+1}{m}}{\binom{n}{m}}$$

(Zur Zählung: Setze m Männer auf n−m+1 Plätze, setze je rechts davon eine Frau, fülle „Restplätze" mit Frauen auf)

Beispiel 1.19 (Hypergeometrische Verteilung)**.** Eine Urne enthalte n Kugeln, davon s schwarze und w weiße ($s + w = n$), ziehe k-mal ohne Zurücklegen,

$$\mathrm{Hyp}_{s,w,k}\big(\{\ell\}\big) = \frac{\binom{s}{\ell}\binom{w}{k-\ell}}{\binom{s+w}{k}}, \quad \ell = 0, 1, \dots, k$$

ist die W'keit, genau ℓ schwarze Kugeln zu ziehen.

Beispiel 1.20 (p-Münzwurf)**.** 1. $\Omega = \{0,1\}$, $\mathrm{Ber}_p(\{1\}) = p = 1 - \mathrm{Ber}_p(\{0\})$ mit einem $p \in [0,1]$ („Bernoulli-Verteilung"[7])

2. n-facher p-Münzwurf (mit $p \in [0,1]$): $\Omega = \{0,1\}^n$,

$$\mathrm{Ber}_p^{\otimes n}\big(\{(\omega_1,\ldots,\omega_n)\}\big) = p^{|\{i \le n : \omega_i = 1\}|}(1-p)^{|\{i \le n : \omega_i = 0\}|}$$

$$\sum_{\omega_1=0}^{1}\sum_{\omega_2=0}^{1}\cdots\sum_{\omega_n=0}^{1} \mathrm{Ber}_p^{\otimes n}(\{(\omega_1,\ldots,\omega_n)\}) = \ldots = \Big(\sum_{\omega_1=0}^{1}\mathrm{Ber}_p(\{\omega_1\})\Big)^n = 1$$

3. Binomialverteilung (zum Parameter n und p, $n \in \mathbb{N}$, $p \in [0,1]$):

$$\mathrm{Bin}_{n,p}\big(\{k\}\big) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k \in \{0,1,\ldots,n\}$$

(dies ist die W'keit, beim n-fachen Münzwurf genau k Erfolge zu beobachten)

Beispiel 1.21 (Geometrische Verteilung)**.** $p \in (0,1)$, $\Omega = \mathbb{N}_0$,

$$\mathrm{Geom}_p\big(\{k\}\big) = p(1-p)^k, \qquad k \in \mathbb{N}_0$$

ist die W'keit, bei wiederholtem p-Münzwurf genau k Misserfolge vor dem ersten Erfolg zu beobachten

Beachte: Manche Autoren betrachten die geometrische Verteilung auf $\mathbb{N}$ (statt auf $\mathbb{N}_0$), dann ist das Gewicht $p(1-p)^{k-1}$ und die Interpretation „k Würfe (einschließlich) bis ersten Erfolg."

Beispiel 1.22 (Negative Binomialverteilung, auch Pascal-Verteilung genannt)**.** Für $r \in (0,\infty)$, $p \in (0,1)$ ist

$$\mathrm{NegBin}_{r,p}\big(\{k\}\big) = \binom{-r}{k}(-1)^k p^r (1-p)^k, \qquad k \in \mathbb{N}_0$$

wobei $\binom{-r}{k} := \frac{(-r)(-r-1)\cdots(-r-k+1)}{k!}$ $\Big(= (-1)^k\binom{r+k-1}{k}$ für $r \in \mathbb{N}\Big)$.

[7] nach Jakob Bernoulli, 1654–1705