Vorlesung 5b

Zufallsvariable mit Dichten

Teil 1 Begriffsbildung, Uniforme Verteilung & Co.

0. Kontinuierlich anstelle von diskret

Bisher im Fokus der Vorlesung:

Diskrete Zufallsvariable.

Sie fallen mit W'keit 1 in eine diskrete
(d.h. endliche oder abzählbar unendliche)
Menge S.

Jetzt wenden wir uns Zufallsvariablen zu,
die kontinuierlich verteilt sind.
Dann ist der Wertebereich *überabzählbar*.

Ein prominentes Beispiel ist der (faire) Münzwurf.

Man kann ihn auffassen als

rein zufällige Wahl eines Elementes aus $\{0,1\}^{\mathbb{N}}$.

Ein weiteres einprägsames Beispiel ist die rein zufällige Wahl

eines Punktes aus dem Einheitsquadrat.

Vgl. dazu das "diskrete Analogon" aus Vorlesung 1a
und den Ausblick auf der dortigen letzen Folie.

Idee bei der rein zufälligen Wahl aus einem Kontinuum:

$\mathbf{P}(X \in A)$ ist gegeben durch den Anteil von A an S

1. Uniforme Verteilung auf dem Einheitsintervall

Eine Zufallsvariable X mit Zielbereich $S = [0, 1]$ heißt

uniform verteilt auf S,

wenn für alle $A \subset S$ mit wohldefiniertem Längenmaß $V(A)$

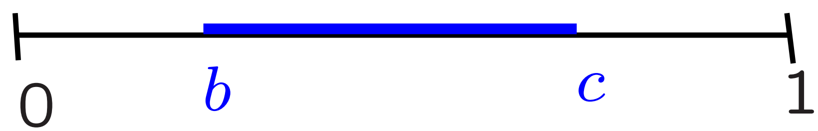
gilt:

$$\mathbf{P}(X \in A) = V(A)$$

(dennn hier ist ja $V(S) = 1$).

Beispiel 1:

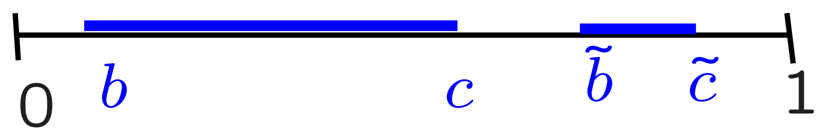
$$A := [b, c] \quad \text{mit } 0 \leq b \leq c \leq 1$$



$$\mathbf{P}(X \in A) = c - b.$$

Beispiel 2:

$A := [b, c] \cup [\tilde{b}, \tilde{c}]$ mit $0 \leq b \leq c < \tilde{b} \leq \tilde{c} \leq 1$



$$\mathbf{P}(X \in A) = (c - b) + (\tilde{c} - \tilde{b}).$$

Beispiel 3:

$A := \{a\}$ mit $0 \le a \le 1$

$$\mathbf{P}(X = a) = a - a = 0.$$

2. Uniforme Verteilung auf einem Rechteck in $\mathbb{R}^2$

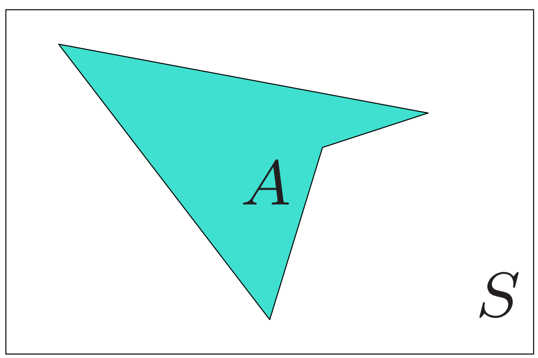
Eine Zufallsvariable X mit Zielbereich

$$S := [0, \ell] \times [0, b] \subset \mathbb{R}^2$$

heißt *uniform verteilt auf* S, wenn

für alle $A \subset S$ mit wohldefiniertem Flächenmaß $V(A)$ gilt:

$$\mathbf{P}(X \in A) = \frac{V(A)}{V(S)} = \frac{V(A)}{\ell \cdot b}.$$



3. Uniforme Verteilung

auf einer Teilmenge des $\mathbb{R}^m$

Definition (Buch S. 12)

Sei S eine Teilmenge des $\mathbb{R}^m$ mit endlichem Inhalt $V(S) > 0$.

Eine Zufallsvariable X mit Zielbereich S heißt

uniform verteilt auf S,

wenn für alle $A \subset S$ mit wohldefiniertem Inhalt $V(A)$ gilt:

$$\mathbf{P}(X \in A) = \frac{V(A)}{V(S)}.$$

$$\mathbf{P}(X \in A) = \frac{V(A)}{V(S)}.$$

Man beachte die Analogie zu

“Anzahl günstige durch Anzahl mögliche Fälle”:

Der zahlenmäßige Anteil von A

an einem endlichen Wertebereich S

wird jetzt ersetzt durch den volumsmäßigen Anteil von A

am (überabzählbar) unendlichen Wertebereich S.

Für diskret uniform verteilte Zufallsvariable hatten wir

$$\mathbf{P}(X = a) = \frac{1}{\#S}.$$

Das anschauliche Analogon dazu ist jetzt:

$$\mathbf{P}(X \in da) = \frac{da}{V(S)}.$$

$$\mathbf{P}(X \in da) = \frac{da}{V(S)}.$$

Der Ausdruck da taucht hier in zwei Bedeutungen auf:

links als infinitesimales Raumstück da

$$\mathbf{P}(X \in da) = \frac{da}{V(S)}.$$

Der Ausdruck da taucht hier in zwei Bedeutungen auf:

links als infinitesimales Raumstück da (um den Punkt a)

und rechts als dessen (infinitesimaler) Inhalt da.

$$\mathbf{P}(X \in da) = \frac{da}{V(S)}.$$

Diese Gleichung bekommt ihre exakte Bedeutung

“unter dem Integral”:

Beispiel: $S = [l, r] \subset \mathbb{R}, \quad l \le b \le c \le r$:

$$\int\limits_{[b,c]} da := \int_b^c da := \int_b^c 1 \, da = c - b.$$

$$\mathbf{P}(X \in da) = \frac{da}{V(S)}.$$

Diese Gleichung bekommt ihre exakte Bedeutung

"unter dem Integral":

$S \subset \mathbb{R}^d$ mit $V(S) < \infty$, $A(\subset S)$ mit Inhalt $V(A)$:

$$\int_A da = \int_A 1\, da = V(A).$$

$$\mathbf{P}(X \in da) = \frac{da}{V(S)}.$$

Diese Gleichung bekommt ihre exakte Bedeutung

“unter dem Integral”:

$S \subset \mathbb{R}^d$ mit $V(S) < \infty$, $A(\subset S)$ mit Inhalt $V(A)$:

$$\int_A \frac{da}{V(S)} = \frac{V(A)}{V(S)}.$$

$$\int_A \mathbf{P}(X \in da) := \mathbf{P}(X \in A)$$

$$\mathbf{P}(X \in da) = \frac{da}{V(S)}, \quad a \in S$$

ist gleichbedeutend mit

$$\mathbf{P}(X \in A) = \frac{V(A)}{V(S)}$$ für alle “messbaren” $A \subset S$.

4. Dichten

Wie im Diskreten begnügen wir uns nicht nur

mit *rein* zufälliger Wahl.

Das Analogon zu den Verteilungsgewichten $\rho(a)$

ist jetzt gegeben durch infinitesimale Gewichte $f(a)\,da$,

wobei $f : S \to \mathbb{R}_+$ eine Funktion ist mit

$$\int_S f(a)\,da = 1.$$

Der wichtigste Fall:

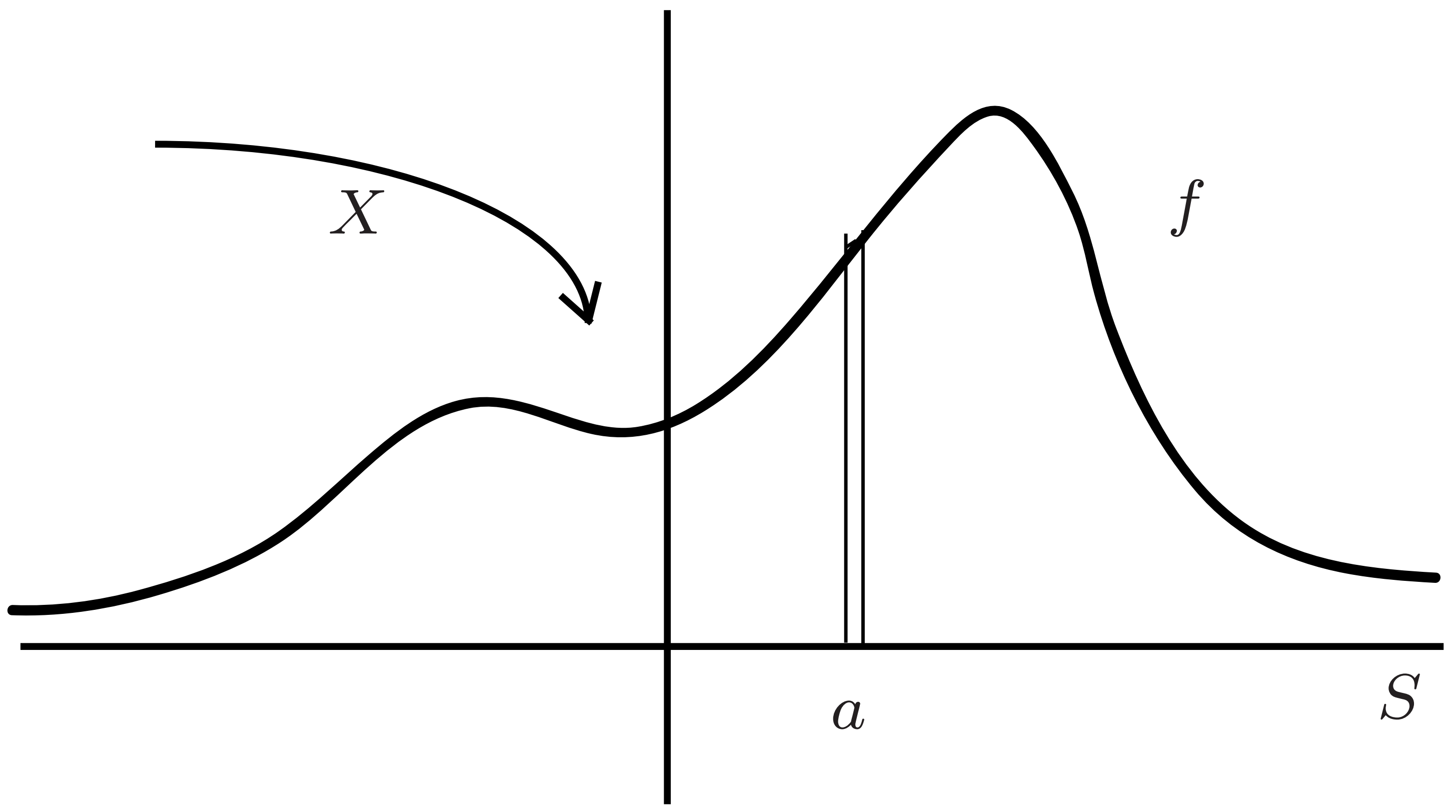
$S \subset \mathbb{R}$ Intervall mit Endpunkten $l,\ r$

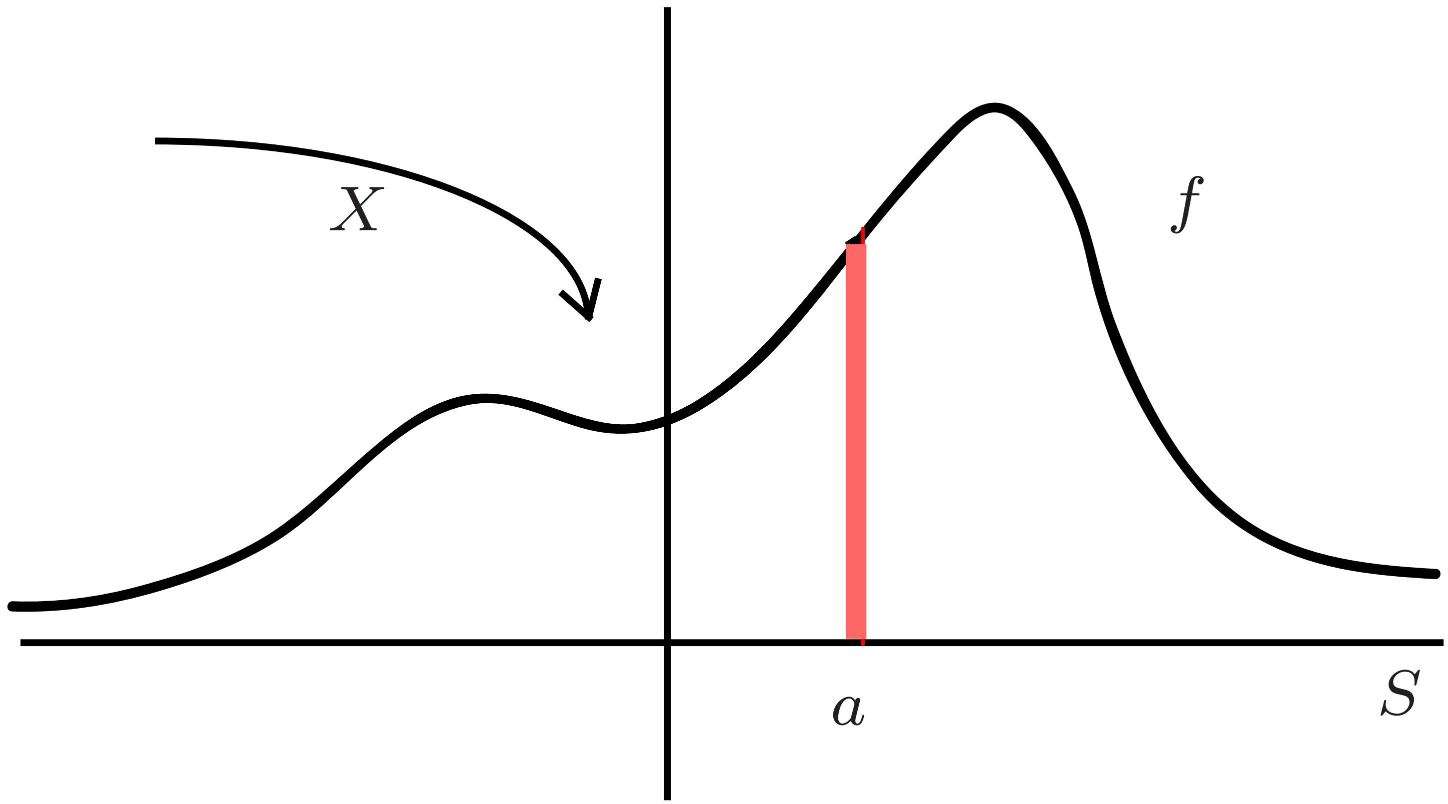
(dabei ist $l = -\infty$ oder $r = \infty$ erlaubt)

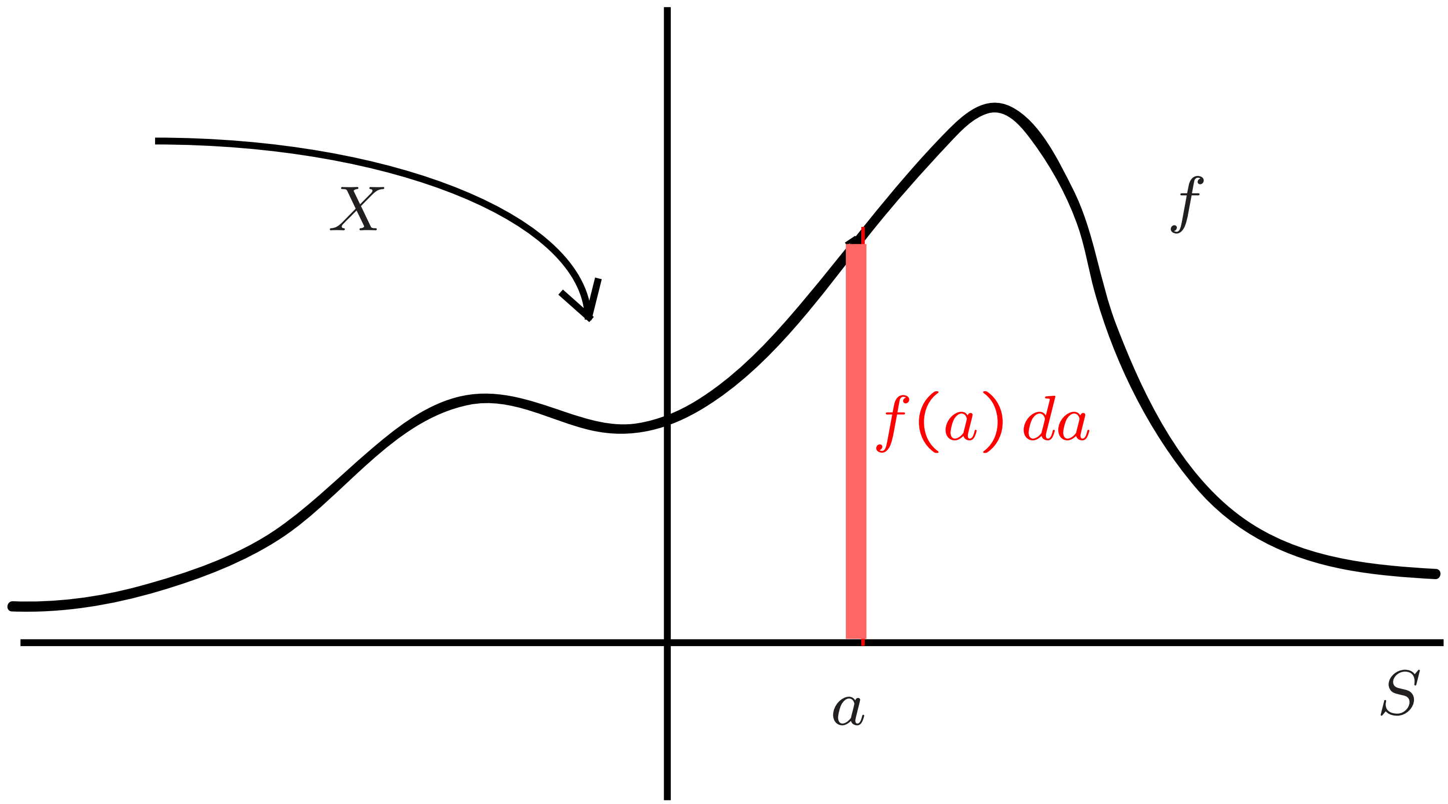
Auf diesen Fall beschränken wir uns im Folgenden.

Sei $f : S \to \mathbb{R}$ nicht-negativ, integrierbar mit

$$\int_l^r f(a)\, da = 1 \; .$$


X
f
a
S


X
f
a
S



X
f
$f(a)\,da$
a
S

Sei X eine Zufallsvariable mit Zielbereich S.

Gilt für alle Intervalle $[b, c] \subset S$ die Gleichung

$$\mathbf{P}\big(X \in [b, c]\big) = \int_b^c f(a)\, da \ ,$$

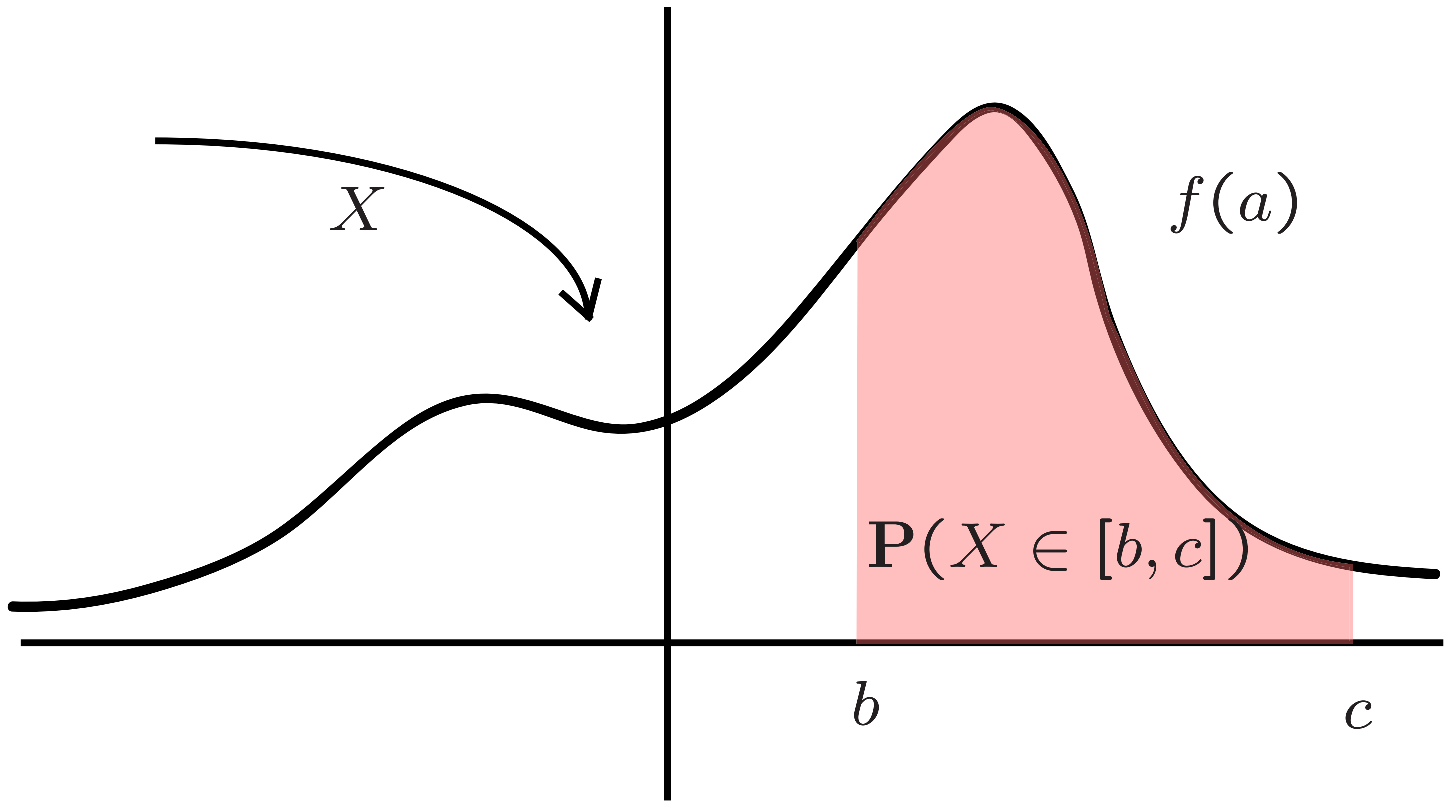
so sagt man, dass

X die *Dichte* $f(a)\, da$ besitzt.

Wir schreiben dann kurz

$$\mathbf{P}(X \in da) = f(a)\, da \ , \quad a \in S \ ,$$

und nennen f *Dichtefunktion* (der Verteilung) von X.

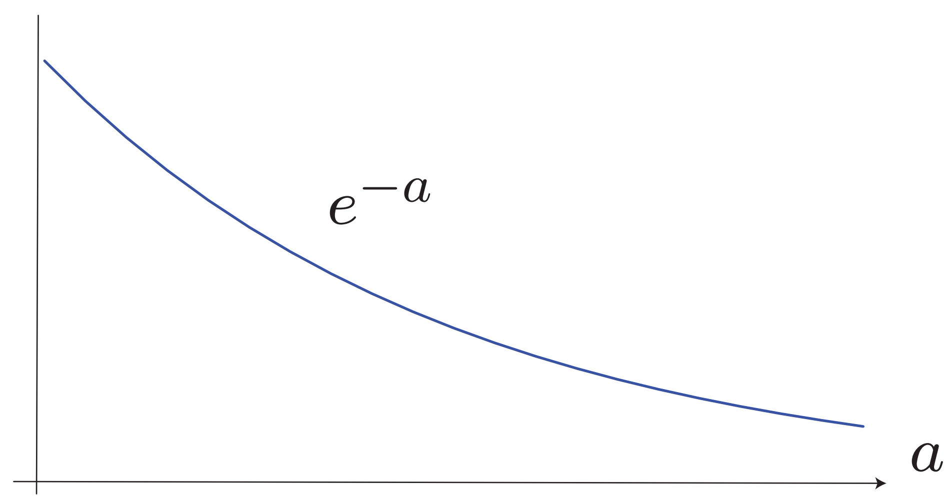

X
$f(a)$
$\mathbf{P}(X \in [b, c])$
b
c

Die Bedingung $\int_S f(a)\, da = 1$ kann auch erfüllt sein,

wenn S unendlichen Inhalt hat.

Man denke an das Beispiel

$$S = [0, \infty); \;\; f(a) = e^{-a}, \quad a \geq 0.$$



Merke:

Für eine Zufallsvariable X mit Dichte $f(a)\,da$

ist für jedes $b \in \mathbb{R}$

$$\mathbf{P}(X = b) = \int_b^b f(a)\,da = 0.$$

Also gilt (mit naheliegender Schreibweise)

für $b \le c \in \mathbb{R}$:

$$\int_{(b,c]} f(a)\,da = \int_{[b,c]} f(a)\,da = \int_b^c f(a)\,da.$$

Tatsächlich hat man (z. B. für stückweise stetiges f) das Integral $\int\limits_A f(a)\,da$ nicht nur für Intervalle A, sondern für eine viel größere Klasse von “messbaren” Mengen’ zur Verfügung.

Außerdem gilt für paarweise disjunkte messbare Mengen $A_1, A_2, \ldots$ die “abzählbare Additivität” des Integrals:

$$\int\limits_{\bigcup A_i} f(a)da = \sum_i \int\limits_{A_i} f(a)\,da.$$

Das hierbei dienliche *Lebesgue-Integral* verallgemeinert den schon aus der Schule bekannten Integralbegriff, sodass Sie “für die Praxis” (und für die in der Vorlesung und den Übungen behandelten Beispiele) nicht umdenken müssen.

Hat die Zufallsvariable X eine Dichte, so gilt

für jede endliche oder abzählbar unendliche Menge S:

$$\mathbf{P}(X \in S) = \sum_{a \in S} \mathbf{P}(X = a) = 0.$$

Insbesondere ist X dann nicht diskret.

Umgekehrt gilt:

Eine diskrete Zufallsvariable besitzt keine Dichte

(im oben definierten Sinn).

5. Verteilungsfunktionen

Wieder sei X eine S-wertige ZV'e mit S Intervall in $\mathbb{R}$.

Die Funktion $F(b) := F_X(b) := \mathbf{P}(X \le b), \quad b \in \mathbb{R}$,

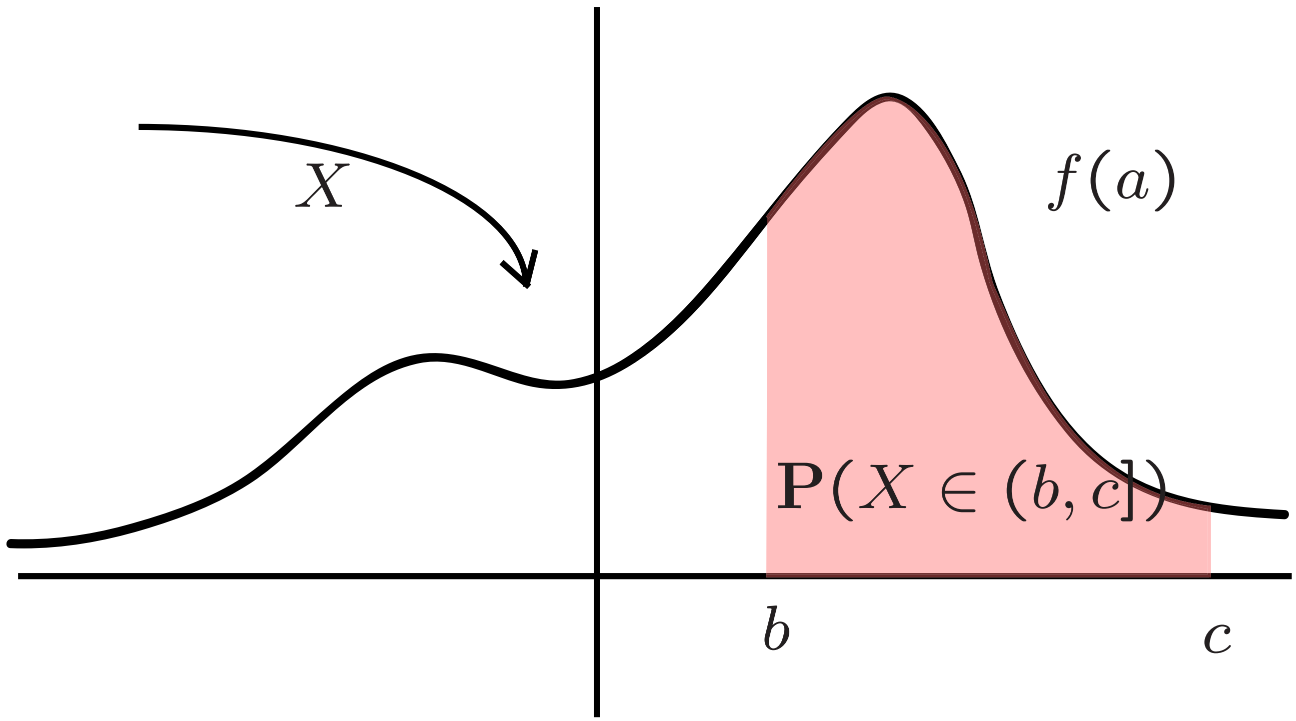
heißt *Verteilungsfunktion* von X.

Hat X die Dichte $f(a)\,da$, so gilt

$$F(b) = \int_{-\infty}^{b} f(a)\,da\ , \quad b \in \mathbb{R}$$

(mit $f(a) := 0$ für $a \notin S$)

Ist f stetig in a, dann ist $f(a) = F'(a)$.



$$\mathbf{P}(X \leq c) - \mathbf{P}(X \leq b) = \mathbf{P}(b < X \leq c)$$

$$F(c) - F(b) = \int_b^c f(a)da$$

Man findet den Hauptsatz der Differential-und Integralrechnung wieder!

Ein hinreichendes Kriterium für die Existenz einer Dichte.

Sei F die Verteilungsfunktion einer reellwertigen Zufallsvariablen X. Hat F keine Sprünge *und* ist F *stückweise stetig differenzierbar**, dann besitzt X eine Dichte.

Denn für jeden Randpunkt a eines der Intervalle gilt: $\mathbf{P}(X = a) = 0$ (ansonsten hätte F in a einen Sprung).
Und innerhalb eines jeden Intervalls gilt nach Voraussetzung der Hauptsatz der Differential-und Integralrechnung, siehe die vorige Folie.

*d.h. es gibt endlich viele disjunkte Intervalle, deren Vereinigung $\mathbb{R}$ ist, so dass F eingeschränkt auf jedes dieser Intervalle eine stetige Ableitung hat.

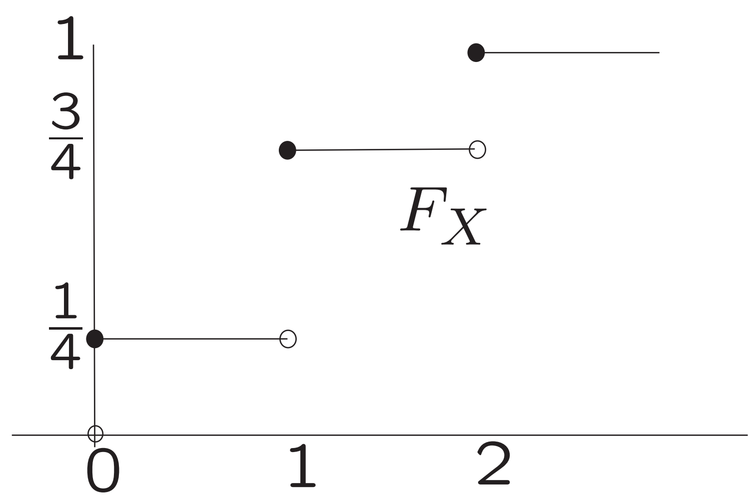
Für *diskrete* reellwertige Zufallsvariable X

ist F_X stückweise konstant, mit Sprüngen der Höhe

$$\mathbf{P}(X = a), \ \ a \in S.$$

Beispiel:

X Binomial$(2, 1/2)$-verteilt



6. Beispiele

A. Eine auf dem Intervall $[0, 2]$

uniform verteilte Zufallsvariable

hat die Dichte $\frac{1}{2}\, da, \;\; 0 \le a \le 2$.

B. Eine in einem endlichen Intervall $S = [l, r]$ uniform verteilte Zufallsvariable

hat die Dichte $\dfrac{1}{r-l}\,da$, $\quad a \in S$.

C. Sei U uniform verteilt auf $[0,1]$.

Gefragt ist nach der Dichte von $X := U^2$.

$$\mathbf{P}(X \le b) = \mathbf{P}(U \le \sqrt{b})$$

$$= \sqrt{b} \overset{!}{=} \int_0^b f(a)\,da, \quad 0 \le b \le 1.$$

$$f(b) = \begin{cases} \frac{1}{2\sqrt{b}}, & 0 < b \le 1 \\ 0 & \text{sonst} \end{cases}$$ ist Dichtefunktion von X.

D. Sei U uniform verteilt auf $[0,2]$.

Gefragt ist nach der Dichte von $X := U^2$.

$$\mathbf{P}(X \leq b) = \mathbf{P}(U \leq \sqrt{b})$$

$$= \frac{1}{2}\sqrt{b} \overset{!}{=} \int_0^b f(a)\,da, \quad 0 < b \leq 4.$$

$$f(b) = \begin{cases} \frac{1}{4}\frac{1}{\sqrt{b}}, & 0 \leq b \leq 4 \\ 0 & \text{sonst} \end{cases}$$ ist Dichtefunktion von X.

E. Sei U uniform verteilt auf $[0,1]$.

Gefragt ist nach der Dichte von $X := -\ln U$.

$$\mathbf{P}(X \leq b) = \mathbf{P}(-\ln U \leq b)$$
$$= \mathbf{P}(U \geq e^{-b}) = 1 - e^{-b} \overset{!}{=} \int_0^b f(a)\, da, \quad b \geq 0.$$

$$f(b) = \begin{cases} e^{-b}, & b \geq 0 \\ 0 & \text{sonst} \end{cases}$$ ist Dichtefunktion von X.