

# Wahrscheinlichkeitstheorie

Dr. C.J. Luchsinger

## 2 Zufallsgrößen

Bevor wir uns den Zufallsgrößen zuwenden (2.3), wollen wir noch kurz 2 Themen vorholen: Allgemeine Bemerkungen zu Abbildungen und Mengen (2.1) und Bemerkungen zu  $\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)$  und  $\lambda$  auf  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  (2.2):

### 2.1 Allgemeine Bemerkungen zu Abbildungen und Mengen

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum. Wir untersuchen erstmal das Urbild einer Menge und fordern vorerst *nicht*, dass die Abbildung messbar (d.h. eine Zufallsgrösse) ist. *(von der oben Abbildung)*

**Definition 2.1 [Urbild einer Menge]** Sei  $X$  eine Funktion von  $\Omega$  nach  $\mathbb{R}$ . Das Urbild unter einer Abbildung  $X$  von  $B \in \mathcal{P}(\mathbb{R})$  ist die folgende Teilmenge von  $\Omega$ :

$$X^{-1}(B) := \{\omega \in \Omega \mid X(\omega) \in B\}.$$

Man beachte, dass wir erst am Schluss dieser Definitions-Kette einen mathematisch exakten Ausdruck haben!

Wir wollen die Abbildung  $X^{-1}$  genauer untersuchen; die Abbildung  $X^{-1}$  erhält Teilmengen, Vereinigungen, Schnitte, Disjunktheit und Komplementbildung (vergleiche auch mit Honours-Aufgabe auf Blatt 2):

**Lemma 2.2 [ $X^{-1}$  und Mengenoperationen]** Seien  $A, B$  sowie  $\{B_\alpha \mid \alpha \in I\}$  Borel-Mengen. Dann gelten:

- Sei  $A \subset B$ , dann auch  $X^{-1}(A) \subset X^{-1}(B)$
- $X^{-1}(\cup_I B_\alpha) = \cup_I X^{-1}(B_\alpha)$
- $X^{-1}(\cap_I B_\alpha) = \cap_I X^{-1}(B_\alpha)$
- Falls  $A \cap B = \emptyset$ , dann auch  $X^{-1}(A) \cap X^{-1}(B) = \emptyset$
- $X^{-1}(A^c) = [X^{-1}(A)]^c$

Vorsicht:  $\cup_{\alpha \in I} B_\alpha \notin \mathcal{B}(\mathbb{R})$  möglich da  $I$  überabzählbar hier erlaubt!

Je nach Zeit; Beweis von Teilen von Lemma 2.2 in VlsG; sonst/Rest in den

Übungen:

$$a) X^{-1}(A) = \{w \mid X(w) \in A\} \stackrel{A \subset B}{\subset} \{w \mid X(w) \in B\} = X^{-1}(B)$$

$$b) \subseteq: \text{Sei } w \in X^{-1}\left(\bigcup_{\alpha} B_{\alpha}\right), \text{ d.h. } X(w) \in \bigcup B_{\alpha}, \text{ d.h. } \exists \alpha_0 \in I: X(w) \in B_{\alpha_0} \\ \Rightarrow w \in X^{-1}(B_{\alpha_0}) \Rightarrow w \in \bigcup_{\alpha} X^{-1}(B_{\alpha}).$$

$$\supseteq: \text{Sei } w \in \bigcup_{\alpha} X^{-1}(B_{\alpha}) \Rightarrow \exists \alpha_0 \in I: X(w) \in B_{\alpha_0} \subseteq \bigcup_{\alpha} B_{\alpha} \Rightarrow \\ w \in X^{-1}\left(\bigcup_{\alpha} B_{\alpha}\right).$$

$$c) \subseteq: \text{Sei } w \in X^{-1}\left(\bigcap_{\alpha} B_{\alpha}\right), \text{ d.h. } X(w) \in \bigcap B_{\alpha}; \text{ d.h. } X(w) \in B_{\alpha} \forall \alpha \in I \\ \text{d.h. } w \in X^{-1}(B_{\alpha}) \forall \alpha \in I \text{ d.h. } w \in \bigcap_{\alpha \in I} X^{-1}(B_{\alpha}).$$

$$\supseteq: \text{Sei } w \in \bigcap_{\alpha} X^{-1}(B_{\alpha}), \forall w \in X^{-1}(B_{\alpha}) \forall \alpha \in I;$$

$$\forall \alpha \in I: X(w) \in B_{\alpha} \forall \alpha \in I \text{ d.h. } X(w) \in B_{\alpha} \forall \alpha \in I \\ \text{d.h. } X(w) \in \bigcap_{\alpha \in I} B_{\alpha} \text{ d.h. } w \in X^{-1}\left(\bigcap_{\alpha \in I} B_{\alpha}\right)$$

$$d) \text{Z-Ansatz: Sei } w \in X^{-1}(A) \cap X^{-1}(B) \neq \emptyset, \text{ d.h. } \\ X(w) \in A \ \& \ X(w) \in B. \Rightarrow A \cap B \neq \emptyset.$$

$$e) \text{Sei } w \in X^{-1}(A^c) \text{ d.h. } X(w) \notin A$$

$$X^{-1}(A^c) = \{w \mid X(w) \notin A\}$$

$$= \{w \mid X(w) \in A^c\}$$

$$= \{w \mid X(w) \in A\}^c$$

$$= \left(X^{-1}(A)\right)^c$$

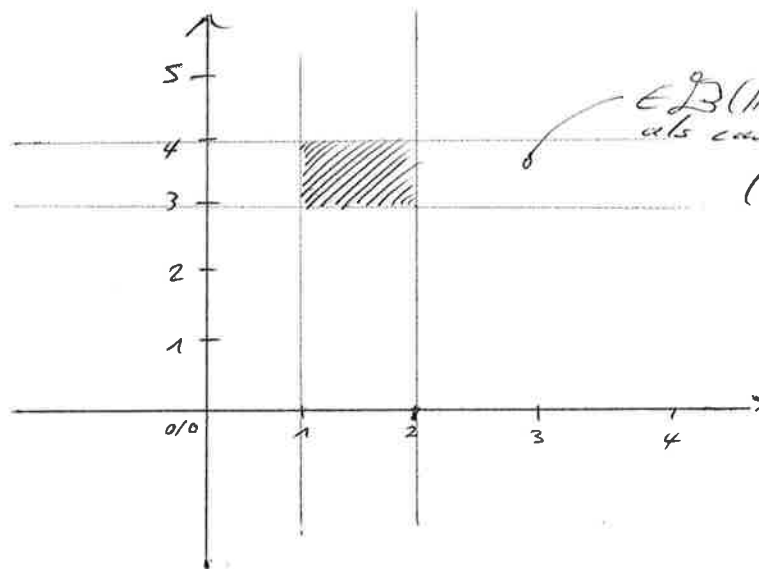
## 2.2 Bemerkungen zu $\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n)$ und $\lambda$ auf $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$

**Definition 2.3** [ $\mathcal{B}(\mathbb{R}^n)$ ] Die Borel- $\sigma$ -Algebra auf  $\mathbb{R}^n$  ist diejenige  $\sigma$ -Algebra  $\mathcal{B}(\mathbb{R}^n)$ , welche vom  $\pi$ -System der Rechtecke (Cartesische Produkte von Intervallen)

$$\prod_{i=1}^n (a_i, b_i]$$

erzeugt wird.

Ein weiterer Erzeuger von  $\mathcal{B}(\mathbb{R}^n)$  ist die Menge  $\{\prod_{i=1}^n B_i \mid B_i \in \mathcal{B}(\mathbb{R})\}$ . Man beachte, dass  $\{\prod_{i=1}^n B_i \mid B_i \in \mathcal{B}(\mathbb{R})\} \subsetneq \mathcal{B}(\mathbb{R}^n)$ ; Ausnahmebeispiel:  $n = 2$ ;  $[1, 2] \times [3, 4]$



Wir wollen noch das Lebesgue-Mass  $\lambda$  auf  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  einführen. Wir unterscheiden bei der Bezeichnung von  $\lambda$  nicht nach Dimension! Auf  $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$  ist  $\lambda$  das einzige  $\sigma$ -endliche Mass derart, dass  $(a_i \leq b_i, \text{ für } 1 \leq i \leq n)$

$$\lambda\left(\prod_{i=1}^n [a_i, b_i]\right) = \prod_{i=1}^n (b_i - a_i).$$

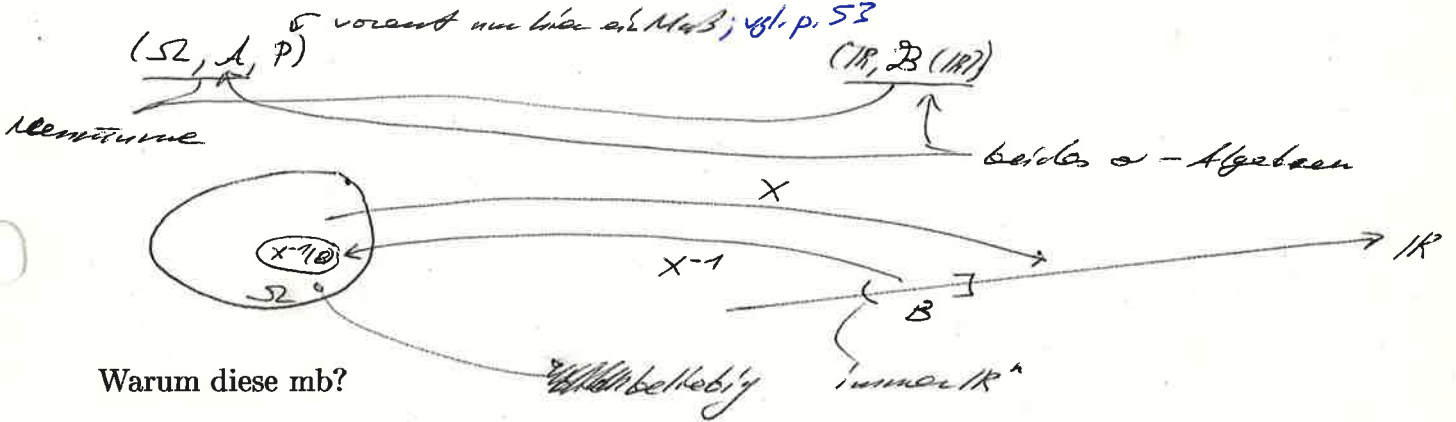
Man sieht an dieser Formel bereits, dass eine Gerade in der Ebene Lebesgue-Mass 0 hat und ebenso eine Ebene im  $\mathbb{R}^3$ . Allgemein Hyperebenen im  $\mathbb{R}^n$ ; auch der Graph von stetigen Funktionen  $f : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$  mit kompaktem Träger (vgl V1sg Reelle Analysis).

### 2.3 Zufallsgrösse

**Definition 2.4 [Zufallsgrösse  $X$  auf  $(\Omega, \mathcal{A}, P)$ ]** Eine Zufallsgrösse auf  $(\Omega, \mathcal{A}, P)$  ist eine Funktion  $X : \Omega \rightarrow \mathbb{R}$  mit der Eigenschaft, dass

$$X^{-1}(B) \in \mathcal{A} \tag{mb}$$

für alle  $B \in \mathcal{B}(\mathbb{R})$ . Die geforderte Eigenschaft nennt man Messbarkeit (mb).



Warum diese mb?

• wollen nochmal  $P$  einführen

•  $P: \mathcal{A} \rightarrow [0, 1] \oplus$

•  $\forall B \in \mathcal{B}(\mathbb{R}) \quad P[X \in B], B \in \mathcal{B}(\mathbb{R})$  ← "fast alle"

•  $P[X \in B] = P[\{\omega \mid X(\omega) \in B\}] = P[X^{-1}(B)]$

Obige Definition ist übrigens ein Spezialfall von

Wichtigste  
of Def. 2.1

•  $\mu \in \mathcal{A}$  sein wegen  $\oplus$ , auf  $\mathcal{P}(\mathbb{R})$  geht gar nicht  
Bauch - Kinn Paroli

**FS14 Definition 2.5 [messbare Abbildung, Borel-Funktion]** Es seien  $(E_1, \mathcal{E}_1)$  und  $(E_2, \mathcal{E}_2)$  Messräume (vgl. Definition 1.24).  $g : E_1 \rightarrow E_2$ .  $g$  heisst  $\mathcal{E}_1 - \mathcal{E}_2$ -messbar, wenn

$$g^{-1}(A) \in \mathcal{E}_1$$

für alle  $A \in \mathcal{E}_2$ . Falls  $(E_i, \mathcal{E}_i) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$  für  $i \in \{1, 2\}$ , nennt man  $g$  eine Borel-Funktion.

Für LeserInnen auf der Suche nach den ganz grossen Zusammenhängen der höheren Mathematik: wenn Sie den Begriff des Messraumes mit der Topologie und die messbare Abbildung mit der stetigen Funktion vergleichen, werden Sie gewisse Analogien entdecken.

$f: X \rightarrow Y, \forall y \in Y$   
 $f^{-1}(y)$  offen in  $X$ .

$A, X \in T$   
 $U, V \in T \Rightarrow U \cap V \in T$   
 $U, V \in T \Rightarrow U \cup V \in T$   
sel. Indexmenge  
"offen" falls  $\in T$



## 2.4 Ein paar weitere, verwandte Definitionen

**Definition 2.6** [*n*-dimensionaler Zufallsvektor] *Ein n-dimensionaler Zufallsvektor*  $\mathbf{X} = (X_1, \dots, X_n)$  *ist eine Funktion*  $\Omega \rightarrow \mathbb{R}^n$  *derart, dass jede Koordinate eine Zufallsgrösse ist.*

Zufallsvektoren werden wir vor allem auch in der Vorlesung SM benutzen: Wenn wir Daten  $(x_1, \dots, x_n)$  haben, so stellen wir uns vor, diese Daten sind Realisationen eines Zufallsvektors  $\mathbf{X} = (X_1, \dots, X_n)$ , also  $\mathbf{X}(\omega) = (X_1(\omega), \dots, X_n(\omega)) = (x_1, \dots, x_n)$ , für ein  $\omega$ , welches gerade eingetreten ist ("Zustand der Welt"). Das geschieht im *Modellierungsschritt*, wo man auch die Verteilung der Zufallsgrösse wählen muss (je mehr Zufallsgrössen Sie kennen, desto besser können Sie das). Wir haben diesen Schritt im Statistikeil der WTS auch schon gemacht. In der Informatik heisst Modellierung übrigens etwas anderes! In der WT werden wir Zufallsvektoren entweder in obigem Sinn benutzen (zufälliges Element im  $\mathbb{R}^n$ ) oder als endliche Folge von Zufallsgrössen. Mathematisch ist es beide Male das gleiche.

**Definition 2.7** [Sub- $\sigma$ -Algebra, Filtration] *Seien*  $\mathcal{A}$  *und*  $\mathcal{F}$  *beide*  $\sigma$ -*Algebren. Wir sagen,*  $\mathcal{A}$  *ist eine Sub- $\sigma$ -Algebra von*  $\mathcal{F}$  *(geschrieben als*  $\mathcal{A} \subseteq \mathcal{F}$ *), wenn*  $\forall A \in \mathcal{A}$  *gilt*

$$A \in \mathcal{A} \Rightarrow A \in \mathcal{F}.$$

*Sei*  $T$  *eine geordnete Menge und seien*  $(\mathcal{A}_t)_{t \in T}$  *Sub- $\sigma$ -Algebren von*  $\mathcal{A}$ . *Wir nennen eine Familie*  $(\mathcal{A}_t)_{t \in T}$  *eine Filtration (in*  $\mathcal{A}$ *), wenn*

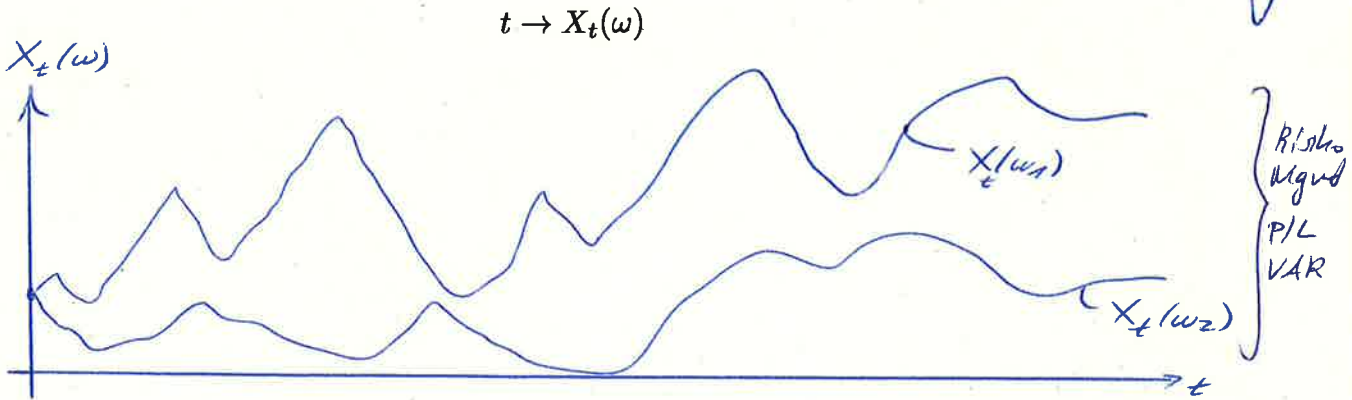
$$s \leq t \Rightarrow \mathcal{A}_s \subseteq \mathcal{A}_t. \quad (\text{Isotonie})$$

**Definition 2.8** [ $(\mathcal{A}_t)_{t \in T}$ -adaptierter Stochastischer Prozess] *Sei*  $T$  *eine geordnete Indexmenge. Wir nennen*  $(\Omega, \mathcal{A}, P, (\mathcal{A}_t)_{t \in T}, (X_t)_{t \in T})$  [kurz  $(X_t)_{t \in T}$ ] *einen*  $(\mathcal{A}_t)_{t \in T}$ -*adaptierten stochastischen Prozess* [kurz "stochastischen Prozess"], *wenn für alle*  $t \in T$  *gilt, dass*  $X_t$   $\mathcal{A}_t$ - $\mathcal{B}(\mathbb{R})$ -*messbar ist.*

d.h. kurz z. w. z. s.:

- $X_t \quad \forall t \in T$  eine Zufallsgrösse (also  $\mathcal{A}_t$ - $\mathcal{B}(\mathbb{R})$ -mb.)
- $(\mathcal{A}_t)_{t \in T}$  Filtration

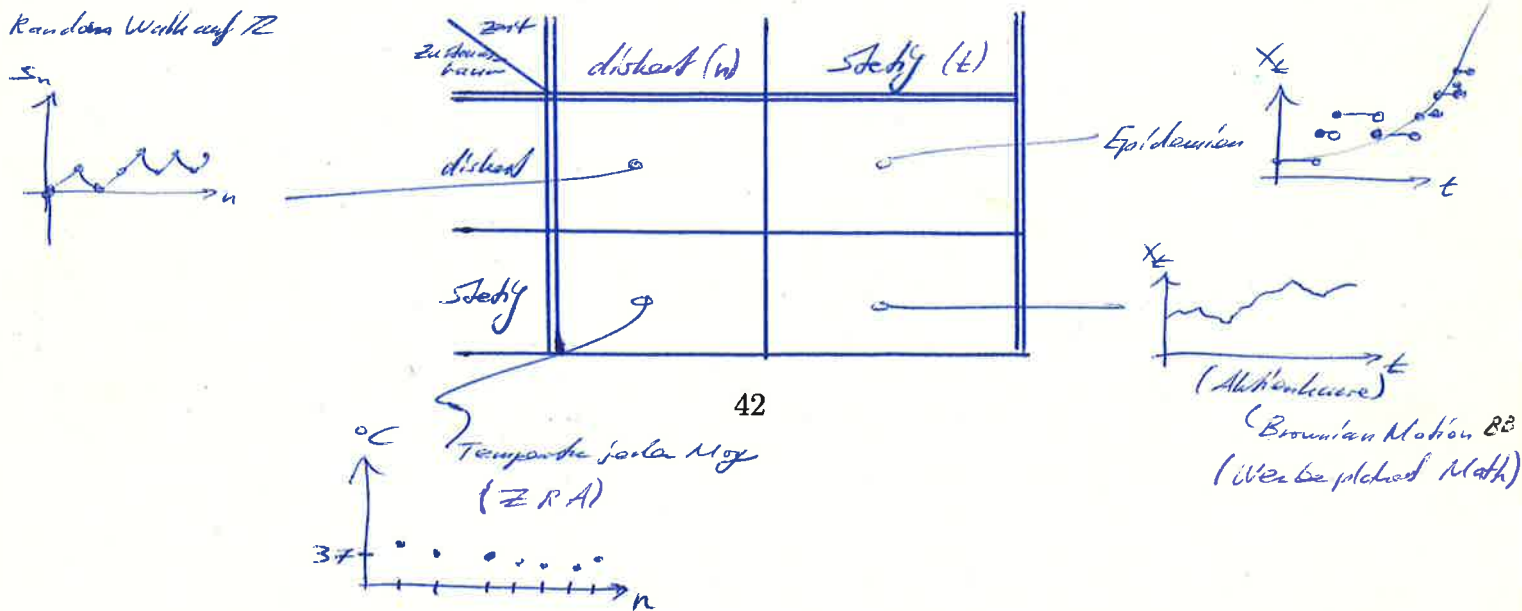
Wie bei der Definition der Zufallsgrösse, vernebelt die strenge mathematische Definition den Blick auf das Wesentliche:  $T$  ist die Zeitmenge (diskret oder stetig je nach Modellierungsgegenstand) und der Zustandsraum ist allgemein  $\mathbb{R}$ . Für jedes feste  $\omega \in \Omega$  nennen wir die Abbildung von  $T$  nach  $\mathbb{R}$



Pfad (Trajektorie, Realisierung) des Prozesses. Man nennt deswegen stochastische Prozesse auch *zufällige Funktionen*. In der Funktionalanalysis werden Sie die Analysis in einer Verallgemeinerung betreiben, dass Sie (deterministische) Funktionen (zB im  $C[[0, \infty]]$ ) wie Punkte in einem Raum begreifen werden. In dieser Abstraktion werden dann viele Resultate bewiesen, welche für die Stochastischen Prozesse ebenfalls gebraucht werden können. In der höheren Stochastik sind deshalb gute Kenntnisse in Funktionalanalysis sehr wichtig.

In den Anwendungen (vgl VlsG AS) kann man dann je nach Modellierungsgegenstand also die Zeit (zB diskret oder stetig) und den Zustandsraum (diskret oder stetig) frei wählen; für's erste gibt es die folgenden 4 (= 2 \* 2) Möglichkeiten:

(da zu noch deterministische Analoge wegen LLN; sobald n gross)



- Sub- $\sigma$ -Algebra:  $\Omega = \{1, 2, 3, 4, 5, 6\}$   
 $\mathcal{F} := \mathcal{P}(\Omega)$   
 $\mathcal{A} := \{\emptyset, \Omega, \{2, 3\}, \{1, 4, 5, 6\}\}$   
 $\Rightarrow \mathcal{A} \subseteq \mathcal{F}$ .

- geordnete Menge  $T$ :  $\exists (\leq)$ :  $t \leq t$  (refl.)  
 $s \leq t \wedge t \leq s \Rightarrow s = t$  (antisym.)  
 $s \leq t, t \leq u \Rightarrow s \leq u$  (transitiv)

- Familie: Menge (mit Band (Zerammungseigenschaft)); oben  $\mathcal{I}$  (=geordn. Menge)  
 mathematisch (z.B. Zeit)

wie sehr fiktiv zu sein. Pkw, Altko, Wasserband

- $\geq \mathcal{B}$ : Random - Walk auf  $\mathbb{Z}$  (vgl. Vlog 45)
  - $X_i = \begin{cases} -1 & \text{mit Wkt } 1/2 \\ +1 & \text{--- " ---} \end{cases}$
  - $(X_i)_{i=1}^n$  i.i.d. (egal)
  - also  $(X_i) \sim \text{Be}(\frac{1}{2})$  i.i.d.
  - $n \in \mathbb{N}$ ,  $T = \{0, 1, 2, \dots, n\}$
  - $S_0 = 0$ ;  $S_n = \sum_{i=1}^n X_i$

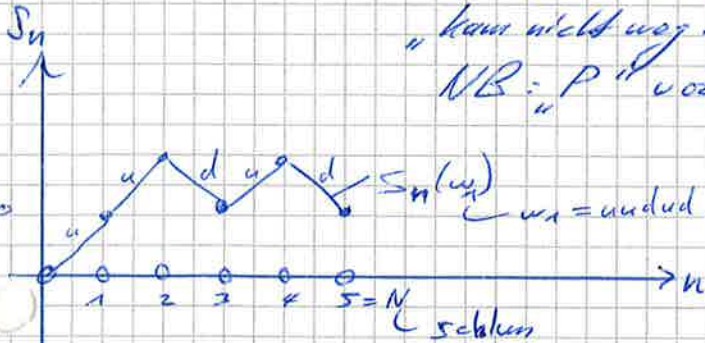
$N=0$ :  $\Omega = \{e\}$ ;  $\mathcal{A} = \{\emptyset, \Omega\}$

$P[\{e\}] = 1$ ;  $\mathcal{A}_0 = \mathcal{A}$

$T = \{0\}$ ;  $\{S_n\}_{n \in T}$

"kann nicht weg vom Fleck"

NB: "P" vorerst nicht interessant; Fokus auf  $(\mathcal{A}_t)_{t \in T}$



$N=1$ :  $\Omega = \{u, d\}$ ;  $\mathcal{A} = \{\emptyset, \Omega, \{u\}, \{d\}\}$ ;  $T = \{0, 1\}$   
 $\mathcal{A}_0 = \{\Omega, \emptyset\}$ ;  $\mathcal{A}_1 = \mathcal{A} \Rightarrow \mathcal{A}_0 \subseteq \mathcal{A}_1$

~~...~~  $S_0 = 0 \forall w \in \Omega \Rightarrow$  mb, p. 40

$S_1(u) = +1, S_1(d) = -1$

$\{w \mid \frac{1}{2}(w) \in \mathcal{B}\} = \begin{cases} \emptyset \text{ falls } -1 \notin \mathcal{B}, +1 \notin \mathcal{B} \\ \{u\} \text{ falls } -1 \notin \mathcal{B}, +1 \in \mathcal{B} \\ \{d\} \text{ falls } -1 \in \mathcal{B}, +1 \notin \mathcal{B} \\ \Omega \text{ falls } -1 \in \mathcal{B}, +1 \in \mathcal{B} \end{cases}$

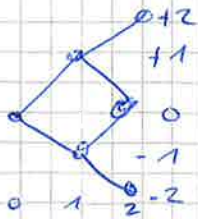
$\Rightarrow$  immer  $\in \mathcal{A}_1 = \text{mb}$ .

(kleiner gelut nicht!  
 (kann man P-1 beschreiben)

$S_1 \in \mathcal{A}_1 - \mathcal{B}(P) = \text{mb}!$

FS 14

$N=2$   $\Omega = \{uu, ud, du, dd\}$ ;  $\mathcal{A} = \{ \emptyset, \Omega, \{uu\}, \{ud\}, \{du\}, \{dd\}, \{uu, ud\}, \{uu, du\}, \{uu, dd\}, \{ud, du\}, \{ud, dd\}, \{du, dd\}, \{uu, ud, du\}, \{uu, ud, dd\}, \{uu, du, dd\}, \{ud, du, dd\} \}$ ;  $|\mathcal{A}| = 16 = 2^4$



$T = \{0, 1, 2\}$

-  $\mathcal{A}_0 = \{ \Omega, \emptyset \}$ ;  $\mathcal{A}_1 = \{ \Omega, \emptyset, \{uu, ud\}, \{du, dd\} \}$ ;  $\mathcal{A}_2 = \mathcal{A}$   
 wobei  $\{uu, ud\}$  ~~...~~  $\{uu\}$  oder  $\{ud\}$  ~~...~~  $\{uu, ud\}$  nach 1 Schritt weiter ist erst, ob 1. Schritt  $\leq u$

FS 10

- $S_0(w) = 0 \forall w \in \Omega \Rightarrow \text{arb.}$
- $S_1 \mathcal{A}_1 - \mathcal{B}(\mathbb{R})$  mb?

$S_1(uu) = S_1(ud) = 1$

$S_1(du) = S_1(dd) = -1$

$\{w \mid S_1(w) \in \mathcal{B}\} = \begin{cases} \emptyset & \text{falls } -1 \notin \mathcal{B}, +1 \notin \mathcal{B} \\ \{uu, ud\} & \text{falls } -1 \notin \mathcal{B}, +1 \in \mathcal{B} \\ \{du, dd\} & \text{falls } -1 \in \mathcal{B}, +1 \notin \mathcal{B} \\ \Omega & \text{falls } -1 \in \mathcal{B}, +1 \in \mathcal{B} \end{cases}$

$\Rightarrow$  immer  $\in \mathcal{A}_1 \Rightarrow \text{arb.}$

-  $S_2 \mathcal{A}_2 - \mathcal{B}(\mathbb{R})$  mb? ; Kurzwegweis: ja:  $\mathcal{A}_2 = \mathcal{F}(\Omega)$

$N=3$  in Aufgabe; Monous ; wollen nicht einfach  $\mathcal{A}_i = \mathcal{F}(\Omega)$  ange

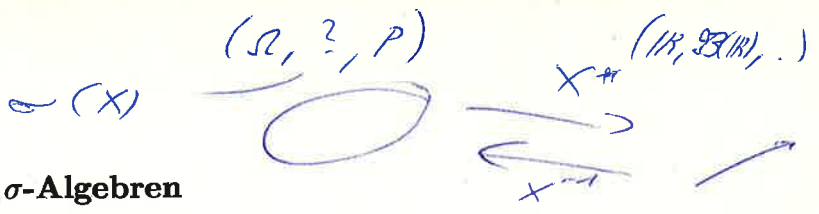
$\Rightarrow$  wird kompliziert; wenn man immer  $(\mathcal{A}_k)_{k \in \mathbb{N}}$  so angeben? Nein; Terminalzeitung analog Lemma 2.9 geht immer (später)! (vgl. Diss Luder p. 22)

$(\mathcal{A}_k)_{k \in \mathbb{N}}$  steht für Informationsmenge in Finanz/Verlässigkeitswelt

2. Absatz a) bei  $N=2$  &  $n=1$  habe ich Informations  $(\mathcal{A}_i)_{i=0}^1$ ; & ich habe Info, welches Ergebnis eintritt!

deswegen  $\mathcal{A}_i = \mathcal{F}(\Omega)$ .

e)  $E[S_m \mid \mathcal{A}_n] = S_n$  <sup>z.B.</sup> wenn  $m > n$  &  $(S_n)_n$  R.W.  $\Rightarrow$  M.G. "anschaulich, mündlich"; Wertelieferung ~~...~~ R.W.



## 2.5 Von Zufallsgrößen erzeugte $\sigma$ -Algebren

*Blatt 2, Honours*

Das folgende Resultat haben wir ~~in Aufgabe 12~~ bereits zu Fuss bewiesen; wir werden es jetzt mit neu erlernten Begriffen und Resultaten eleganter formulieren und beweisen können:

**Lemma 2.9 [von  $X$  erzeugte  $\sigma$ -Algebra]** Sei  $X$  eine Zufallsgröße. Die Familie

$$\sigma(X) := \{X^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R})\}$$

ist eine  $\sigma$ -Algebra auf  $\Omega$ . Man nennt sie "die von  $X$  erzeugte  $\sigma$ -Algebra".

**Beweis Lemma 2.9:** (d.h.  $\sigma(X)$  ist  $\sigma$ -Algebra)

•  $\Omega = X^{-1}(\mathbb{R}) \in \sigma(X)$ ; da  $\mathbb{R} \in \mathcal{D}(\mathbb{R})$

• Sei  $X^{-1}(A) \in \sigma(X) \Rightarrow A^c \in \mathcal{D}(\mathbb{R}) \Rightarrow X^{-1}(A^c) \in \sigma(X)$ ;  
 aber  $X^{-1}(A^c) = (X^{-1}(A))^c \in \sigma(X)$ .  
*(Lemma 2.2e)*

• Sei  $X^{-1}(B_i) \in \sigma(X) \forall i, \Rightarrow \bigcup B_i \in \mathcal{D}(\mathbb{R}) \Rightarrow$   
 $X^{-1}(\bigcup B_i) \in \sigma(X)$ ; aber

$X^{-1}(\bigcup B_i) = \bigcup X^{-1}(B_i) \in \sigma(X)$ .  
*(Lemma 2.2b)*

□

NB: •  $\sigma(X)$  ist  $\sigma$ -Algebra für jede Funktion  $X$  von  $\Omega \rightarrow \mathbb{R}$   
 (vgl. ~~Blatt~~ Honours A2)

•  $X$  ist eine Zufallsgröße genau dann wenn  $\sigma(X) \in \mathcal{A}$ .  $\oplus$

(d.h.  $\sigma$ -Algebra, und  $X$  ist b.)

Beispiele zu "von X erzeugte  $\sigma$ -Algebra"

Sei  $A \in \mathcal{A}$ ; beachte  $\sigma(\mathbb{1}_A)$ !

$$\sigma(\mathbb{1}_A) = \{ \mathbb{1}_A^{-1}(B) \mid B \in \mathcal{B}(\mathbb{R}) \} = \{ \emptyset, A, A^c, \Omega \}$$

FS 12

Sei  $X = \sum_{i=1}^n a_i \mathbb{1}_{A_i}$  mit  $(a_i \neq a_j; i \neq j)$ ; dann (vgl. p. 40a)

einfache ZB

p. 6 werden von Mengensystem erzeugte  $\sigma$ -Algebren

$$\sigma(X) = \sigma(\{A_1, \dots, A_n\}) = \left\{ \bigcup_{i \in I} A_i, I \subseteq \{1, \dots, n\} \right\};$$

(natürliche Familie  $\mathcal{B} = \{a_i\} \rightarrow$  filtert  $A_i$  aus)  
 egal wie  $a_i$ ! Warum wenn  $a_i$  nicht mehr verschieden?

$\Downarrow$   
 $\sigma(X)$  wird "einfach" ~~über~~  $A_i$ ;  
~~Wahl~~  $a_i = a_j \Rightarrow$   
 $a_i \neq a_j \forall i \neq j$   
 $A_i \notin \sigma(X), A_j \notin \sigma(X)$ ;  
 $A_i \cup A_j \in \sigma(X)$ .

$\Rightarrow \sigma(X)$  hängt von Anzahl verschiedener Werte ab die  $X$  annehmen kann, nicht von konkreten Werten.

FS 08

Jemand hat  $(\Omega, \mathcal{F}, P)$  & ein kompliziertes  $X$ .  
 (für die Konstruktion)  
 Man will nicht  $\mathcal{A} = \mathcal{F}(\Omega)$  nehmen. Was kann man machen? Nehme einfach

$$\mathcal{A}_1 = \sigma(X), \text{ dann } \sigma(X) \in \mathcal{A}_1; \text{ also } \oplus \text{ p. 43 erfüllt!}$$

Analog Vgl SM-Häufigkeit:  $44(X_1, \dots, X_n)$  (vgl. Def. 2.6)  $\rightarrow$  Blatt

v. a.  $(X_t)_{t \geq 0}$  stoch. Prozess;  $\mathcal{B}^B; a_i$ , wie  $(A_t)_{t \geq 0}$ ?

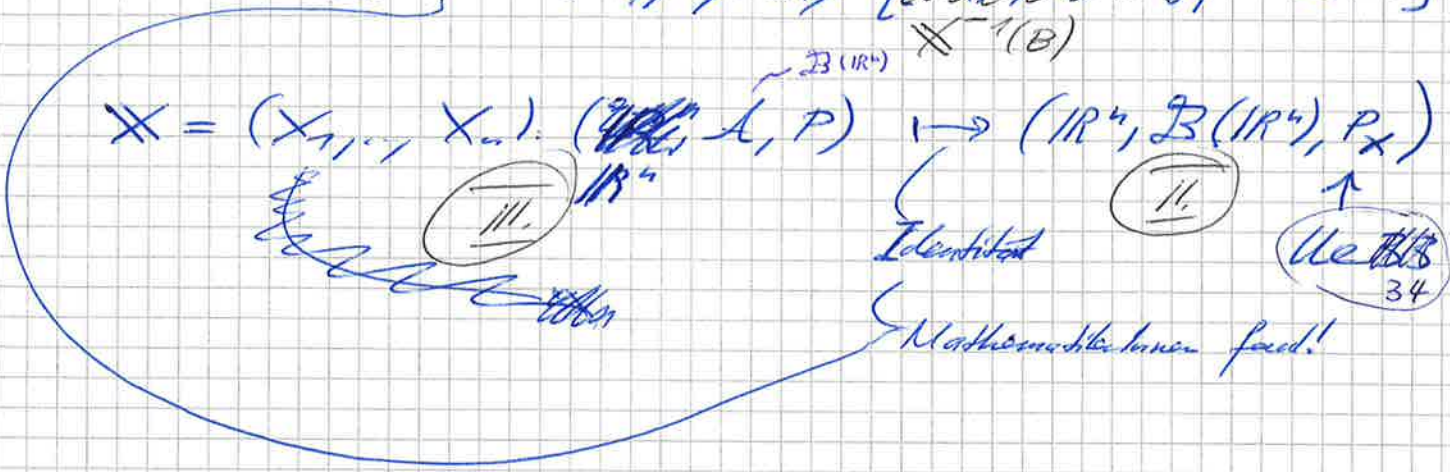
$$(A_t)_{t \geq 0} := \sigma\{X_s \mid s \geq 0\}.$$

Wie wandelt man Algebraisch VLK in SM? (vgl. Def. 2.6), p. 44u

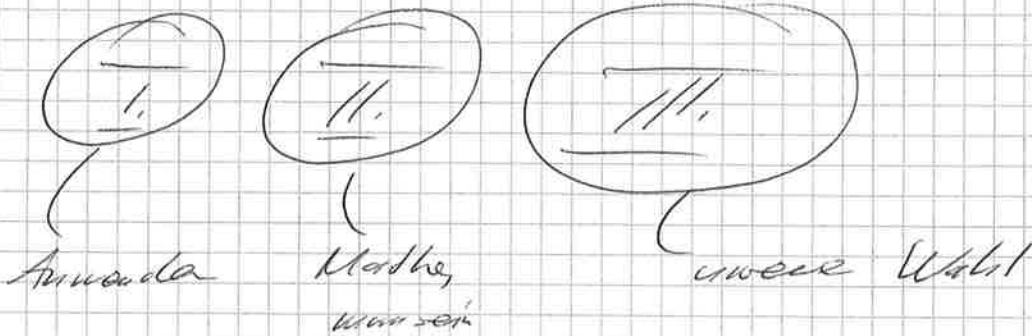
- Daten  $(x_1, \dots, x_n)$
- Modell (z. B.) zB  $MVN_n(\mu, \Sigma)$ :  $(x_1, \dots, x_n)$   
 $\searrow$  je zB

• Wie  $(\Omega, \mathcal{A}, P)$   $\leftarrow$  das benutzen wir

• Teilfall:  $\Omega = \mathbb{R}^{n \times 2}$   
 $\mathcal{A} = \sigma(x_1, \dots, x_n) = \{ \{ \text{Matrix } B \} / B \in \mathcal{B}(\mathbb{R}^{n \times 2}) \}$   
 $\neq \mathcal{B}(\mathbb{R}^n)$



$\Rightarrow$  in SM alles OK mit Daten/Zufallsvektor, wenn wenn  $n \rightarrow \infty$ !



## kleine Dissonanz und deren Auflösung

Wir haben in der WTS die Zufallsgrößen folgendermassen definiert:

**WTS-Definition 2.1** [Zufallsgrösse  $X$  auf  $(\Omega, \mathcal{A}, P)$ ] Eine Zufallsgrösse auf  $(\Omega, \mathcal{A}, P)$  ist eine Funktion  $X : \Omega \rightarrow \mathbb{R}$  mit der Eigenschaft, dass  $\{\omega \in \Omega \mid X(\omega) \leq a\} \in \mathcal{A}$  für alle reellen  $a$ . Die geforderte Eigenschaft nennt man Messbarkeit.

Wie ist das jetzt mit unserer Definition 2.4?

" $\Rightarrow$ " aus Def. 2.4 folgt WTS-Def. 2.1 da

$$\{\omega \in \Omega \mid X(\omega) \leq a\} = \{\omega \in \Omega \mid X(\omega) \in (-\infty, a]\} \in \mathcal{B}(\mathbb{R}).$$

" $\Leftarrow$ " Sei  $\mathcal{G}$  die Familie der Borel-Mengen ~~...~~ wo  $X^{-1}(B) \in \mathcal{A}$ .

$$\mathcal{G} := \{B \in \mathcal{B}(\mathbb{R}) \mid X^{-1}(B) \in \mathcal{A}\}. \text{ (dabei wo es schon bestimmt)}$$

• Ul. HA:  $\mathcal{G}$  ist eine  $\sigma$ -Algebra auf  $\mathbb{R}$  & enthält alle Intervalle.

•  $\Rightarrow \mathcal{G} \equiv \mathcal{B}(\mathbb{R})$  ged.

~~...~~

• WTS-Def. 2.1. ist einfacher!

• statt  $\{X \leq a\}$ , geht auch  $\{X > a\}$ ,  $\{X \geq a\}$ ,  $\{X < a\}$  (Ul. HA)

$$\{(x,y) \mid x+y < t\} = \bigcup_{r \in \mathbb{Q}} \left( \{(x,y) \mid x < r\} \cap \{(x,y) \mid y < t-r\} \right)$$

- a)  $\supset$
- b)  $\subset$

a) Sei  $(x,y)$  so, d.h.  $\exists r \in \mathbb{Q} : x < r$  &  $y < t-r$

$$\Rightarrow \underline{\underline{x+y < t}} \quad \square$$

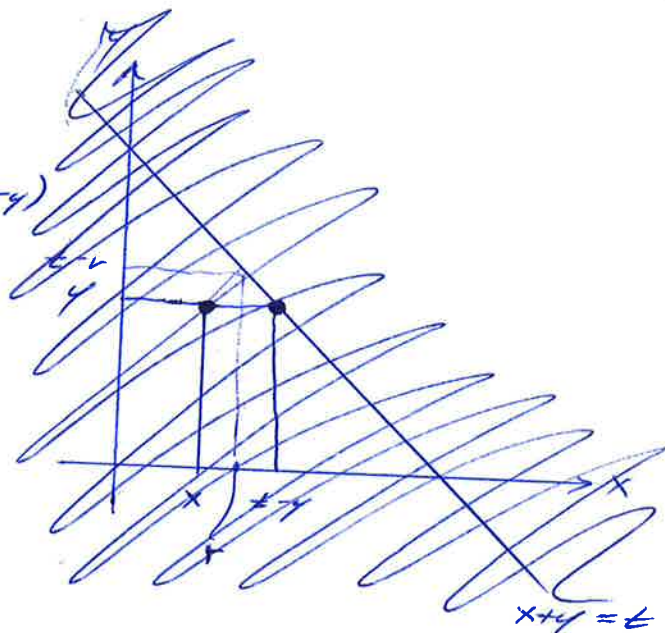
b) Sei  $(x,y)$  so, d.h.

$$x+y < t \Rightarrow \exists r \in \mathbb{Q} : r \in (x, t-y)$$

$\Rightarrow x < r$  &  $r < t-y$ , d.h.:

$$\underline{\underline{x < r}} \text{ & } \underline{\underline{y < t-r}}$$

qed.



## 2.6 Algebraische Verknüpfungen, Limiten und Transformationen von Zufallsgrößen

Falls Sie Wahrscheinlichkeitsrechnung und Statistik auf Gymnasial- oder Fachhochschulstufe unterrichten oder Service-Veranstaltungen für Nicht-MathematikerInnen halten, so werden Sie ohne Bedenken zum Beispiel Summen von Zufallsgrößen bilden. Definition 2.4 beinhaltet aber, wie bereits gesagt, 2 Teile. Der erste Teil ist unproblematisch: Zufallsgrößen sind Abbildungen von  $\Omega$  nach  $\mathbb{R}$ . Also wird man Summen und andere algebraische Verknüpfungen und Limiten punktweise definieren. Aber sind das dann noch Zufallsgrößen? Haben wir auch die Messbarkeit? Die folgenden Lemmata bejahen dies:

$$X+Y=Z \Leftrightarrow Z(\omega) = X(\omega) + Y(\omega) \quad \forall \omega \in \Omega$$

**Lemma 2.10 [Algebraische Operationen]** Seien  $X$  und  $Y$  Zufallsgrößen. Dann gilt:

a)  $aX + bY$  ist eine Zufallsgröße, wo  $a, b \in \mathbb{R}$ ; damit wird die Menge aller Zufallsgrößen zu einem (reellen) Vektorraum!

b)  $\max\{X, Y\}$  und  $\min\{X, Y\}$  sind Zufallsgrößen

c)  $XY$  ist eine Zufallsgröße

d) Falls für jedes  $\omega \in \Omega$  gilt, dass  $Y(\omega) \neq 0$ , so ist auch  $X/Y$  eine Zufallsgröße

e)  $X^+, X^-, |X|$  sind Zufallsgrößen.

**Beweis von Lemma 2.10**

a) wir zeigen  $(X+Y)$  &  $(aX)$  sind Z.G.; wir benutzen WTS-Def. der Veränderl. 2.1.  
 $\forall t \in \mathbb{R}$ :

$$\bullet \{X+Y < t\} = \bigcup_{t \in \mathbb{R}} \underbrace{\{X < t\}}_{\in \mathcal{A}} \cap \underbrace{\{Y < t - X\}}_{\in \mathcal{A}} \quad \otimes$$

$$\underbrace{\hspace{10em}}_{\in \mathcal{A}}$$

46

$$\bullet a > 0; \forall t \in \mathbb{R}: \{aX \leq t\} = \left\{ X \leq \frac{t}{a} \right\}$$

$$\bullet a < 0; \forall t \in \mathbb{R}: \{aX \leq t\} = \left\{ X \geq \frac{t}{a} \right\} \Rightarrow aX \text{ ist Z.G.}$$

Beweis von Lemma 2.10 (Fortsetzung)

$$b) \{\max\{x, y\} \leq t\} = \{x \leq t\} \cap \{y \leq t\}$$

$$\{\min\{x, y\} \leq t\} = \{x \leq t\} \cup \{y \leq t\}$$

c) Vorbereitung:  $X^2$  ist ZB: sei  $t \geq 0$ ,

$$\{X^2 \leq t\} = \{-\sqrt{t} \leq X \leq \sqrt{t}\} = \underbrace{\{X \leq \sqrt{t}\} \setminus \{X < -\sqrt{t}\}}_{\in A}$$

$$x \cdot y = \frac{1}{2} \left[ \underbrace{(x+y)^2}_{a)} - \underbrace{x^2}_{b)} - \underbrace{y^2}_{c)} \right]$$

Vorbereitung (.)<sup>2</sup>

$$d) \left(\frac{1}{y}\right) \text{ ist ZB. } \left\{ \frac{1}{y} \leq t \right\} = \begin{cases} \{y < 0\} \cup \{y \geq \frac{1}{t}\} & \text{falls } t > 0 \\ \{y \leq \frac{1}{t}\} & \text{falls } t < 0 \end{cases}$$

$\Rightarrow$  Z.B.

$$\Rightarrow X/Y = X \left(\frac{1}{Y}\right) \text{ ist ZB wegen c)}$$

FS.12

$$e) X^+ = \max\{X, 0\}$$

$$X^- = -\min\{X, 0\} \text{ (hochgeklappter neg. Teil)}$$

$$X = X^+ - X^- \text{ (damit } X \text{ Differenz zweier nichtneg. ZB)}$$

$$|X| = X^+ + X^-$$

$X^+, X^-$  &  $|X|$  sind ZB wegen b) & d).

**Lemma 2.11 [Folgen, Limiten, Summen von Zufallsgrößen]** Sei  $(X_i)_{i \geq 1}$  eine Folge von Zufallsgrößen. Dann gilt:

- a)  $\sup_n X_n, \inf_n X_n$  sind Zufallsgrößen  $\sup_n X_n(\omega) ; \forall \omega \in \Omega$  punktweise definiert
- b)  $\limsup_n X_n, \liminf_n X_n$  sind Zufallsgrößen
- c) Falls  $X(\omega) := \lim_{n \rightarrow \infty} X_n(\omega)$  für alle  $\omega \in \Omega$  existiert, dann ist auch  $X$  eine Zufallsgröße.
- d) Falls  $X(\omega) := \sum_{n=1}^{\infty} X_n(\omega)$  für alle  $\omega \in \Omega$  existiert, dann ist auch  $X$  eine Zufallsgröße.

Beweis von Lemma 2.11

$\{\omega | X_n(\omega) \leq t\}$ ; fixiere  $\omega$

a)  $\{\sup_n X_n \leq t\} = \bigcap_n \{X_n \leq t\} \Rightarrow \sup X_n$  ist Z.G.  
 $\{\inf_n X_n \geq t\} = \bigcap_n \{X_n \geq t\} \Rightarrow \inf X_n$  ist Z.G.

b)  $\limsup_n X_n \stackrel{A1}{=} \inf_k \sup_{m \geq k} X_m \Rightarrow \limsup X_n$  ist Z.G.  
 $\liminf_n X_n \stackrel{A1}{=} \sup_k \inf_{m \geq k} X_m \Rightarrow \liminf X_n$  ist Z.G.

$(-1)^n + \frac{1}{n}$   
 monoton fallend  
 sup.

c) falls  $\lim X_n \exists \Rightarrow \exists \limsup$  & dieser ist gleich  $\lim X_n \Rightarrow$  auch  $\lim X_n$  ist Z.G. wegen d.)

d)  $X(\omega) := \lim_{N \rightarrow \infty} \sum_{n=1}^N X_n(\omega) = \lim_{N \rightarrow \infty} S_N(\omega)$  ist Z.G. wegen c)  
 Lemma 2.10d)

warum  $\lim X_n = +\infty$  möglich  
 - U9 AS vorgeben  
 - Abbildung Werteset ist (fast) O.k. w.B.)

**Kleine Bemerkung:** wegen w.B. &  $X = \infty$ ; Forderung wird, dass  $P[X \in (-\infty, \infty)] = 1$ .

• Bsp AS:



$\rightarrow T_3 := \min\{n | X_n = 3\}$  wo  $X_0 = 3$  &  $X_n =$  Zustand zu Zeit  $n$ .

Rückkehrzeit nach 3:  $P[T_3 = 1] = 1/2$

48

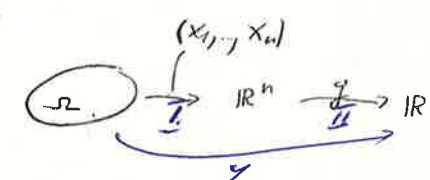
$P[T_3 = 2] = 1/4$  .....  $P[T_3 = k] = 0$   
 $P[T_3 = \infty] = 1/4$

• Lsg (mb):  $P[T_3 \in \mathbb{R}] = 1$ ;  $\{T_3 = \infty\} \in \mathcal{A}$ ,  $\{T_3 = -\infty\} \in \mathcal{A}$ ,  $\{T_3 \in \emptyset\} \in \mathcal{A} \cup \mathcal{B} \in \mathcal{B}(\mathbb{R})$ .

• Vorsicht:  $P[|X| < \infty] = 1$  &  $E[|X|] = \infty$  möglich; vgl. Kapitel 4 ( $P[|X| = \infty] > 0 \Rightarrow E[|X|] = \infty$ )

**Lemma 2.12 [Transformationen von Zufallsgrößen]** Sei  $(X_1, \dots, X_n)$  ein Zufallsvektor und sei  $g: \mathbb{R}^n \rightarrow \mathbb{R}$  eine Borel-Funktion ( $g^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$  für alle  $B \in \mathcal{B}(\mathbb{R})$ ). Dann ist auch  $Y := g(X_1, \dots, X_n)$  eine Zufallsgröße.

*über Zufallsbilder*  
**Beweis von Lemma 2.12** Habe  $(\Omega, \mathcal{A}, P)$



• Zeigen zuerst  $\{(x_1, \dots, x_n) \in \tilde{B}\} \in \mathcal{A} \quad \forall \tilde{B} \in \tilde{\mathcal{B}}(\mathbb{R}^n)$ . I.

Sei  $\tilde{B} = \prod_{i=1}^n B_i$  ( $B_i \in \mathcal{B}(\mathbb{R})$ ; kartesisches Produkt), dann

*speziell*

$$\{(x_1, \dots, x_n) \in \tilde{B}\} = \{x_1 \in B_1, x_2 \in B_2, \dots, x_n \in B_n\} = \bigcap_{i=1}^n \{x_i \in B_i\}$$

(Kommutativ und  $\cap$ )  
 $\in \mathcal{A}$  (Def 2.6)  
 $\in \mathcal{A}$

$\{\tilde{B} \in \mathcal{B}(\mathbb{R}^n) \mid \tilde{X}^{-1}(\tilde{B}) \in \mathcal{A}\}$  ist  $\sigma$ -Algebra (gleiche Schritte wie Beweis von Lemma 2.4) & erzeugt  $\mathcal{B}(\mathbb{R}^n) \rightarrow$  ist sogar gleich  $\mathcal{B}(\mathbb{R}^n)$ .  
 (da  $\mathbb{R}^n = \prod_{i=1}^n \mathbb{R}$ ) *also  $\forall \tilde{B} \in \mathcal{B}(\mathbb{R}^n)$ .*

• Sei jetzt  $B \in \mathcal{B}(\mathbb{R})$ ;  $\{Y \in B\} = \{(x_1, \dots, x_n) \in g^{-1}(B)\}$ , aber  $g^{-1}(B) \in \mathcal{B}(\mathbb{R}^n)$   
 da  $g$  Borel-Funktion. ~~...~~  $\in \mathcal{A}$  wegen I.

Wir haben die einfachen Zufallsgrößen bereits eingeführt. Sie werden ihrem Namen völlig gerecht insofern, dass sie *endliche* Linearkombinationen von *Indikatorfunktionen* sind - und Indikatorfunktionen sind wirklich einfach zu handhaben. Wir werden die einfachen Zufallsgrößen bei der Definition von Erwartungswerten benutzen. Dazu wird Lemma 2.13 benutzt:

**Lemma 2.13 [Approximation nichtnegativer Zufallsgrößen durch eine Folge einfacher Zufallsgrößen]** Sei  $X$  eine nichtnegative Zufallsgröße. Dann gibt es eine monoton wachsende Folge einfacher Zufallsgrößen  $0 \leq X_1 \leq X_2 \dots$  sodass  $X_n(\omega) \uparrow X(\omega)$  für alle  $\omega \in \Omega$ .

**Beweis von Lemma 2.13**  $X_n$  definieren wir (clever!); 2 Minuten lernen

• 
$$X_n = \sum_{k=1}^{n \cdot 2^n} \frac{(k-1)}{2^n} \mathbb{1} \left( \frac{k-1}{2^n} < X \leq \frac{k}{2^n} \right) + n \cdot \mathbb{1}(X > n)$$

• ist einfache ZB!

• auf  $\left\{ \frac{k-1}{2^n} < X \leq \frac{k}{2^n} \right\}$  nimmt  $X_n$  den Wert  $\frac{k-1}{2^n}$  an, also den unteren Wert; damit haben wir

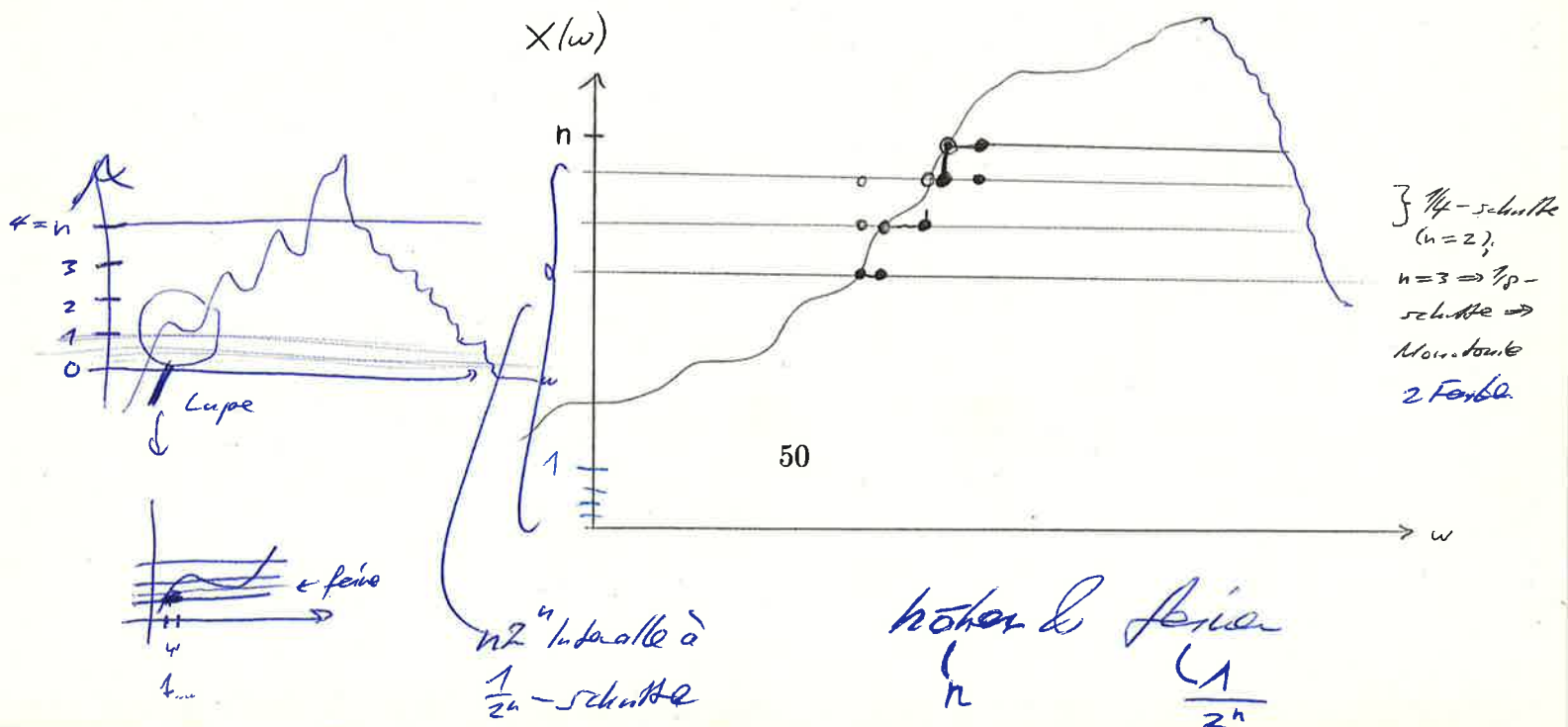
also, was größer n auf n unte

a)  $X_n \leq X$  (i)

b) Monotonie (ii)

•  $X=0 \Rightarrow X_n=0 \forall n$

c)  $X_n(\omega) \nearrow X(\omega)$  (iii)



Wir kommen jetzt - was die Beweistechnik anbelangt (!) - zu einem Analogon von Satz 1.6 [Monoton-Lemma für Mengen]. Wir werden das nachfolgende Monoton-Lemma für Zufallsgrößen dann einsetzen, wenn wir beweisen wollen, dass eine bestimmte Menge von Zufallsgrößen alle (nichtnegativen) Zufallsgrößen enthält.



**Theorem 2.14 [Monoton-Lemma für Zufallsgrößen]** Sei  $\mathcal{G}$  ein  $\pi$ -System, welches  $\mathcal{A}$  erzeugt.  $\mathbf{H}$  sei eine Menge von Funktionen auf  $\Omega$  derart, dass

- i)  $1 \in \mathbf{H}$  (konstante Funktion ist drin)
- ii)  $1_A \in \mathbf{H}$  für alle  $A \in \mathcal{G}$
- iii)  $\mathbf{H}$  ist ein Vektorraum
- iv) Falls  $X_n \in \mathbf{H}$  für alle  $n$  und  $\sup_n X_n(\omega) < \infty$  für alle  $\omega \in \Omega$ , dann gehört auch  $\sup_n X_n$  zu  $\mathbf{H}$ .

Dann beinhaltet  $\mathbf{H}$  alle (!) Zufallsgrößen. *U.M.E.:  $H = ?$  wenn  $\mathcal{A} = \{\emptyset, \Omega, A, A^c\}$*   
*vgl. p. 44:  $\sum_{i=1}^2 c_i 1_{A_i}$ ;  $A_1 = A$ ,  $A_2 = A^c$   $\Rightarrow c_1 1_A + c_2 1_{A^c} = c_1 1_A + c_2 (1 - 1_A) = (c_1 - c_2) 1_A + c_2$   $c_i \in \mathbb{R}$*   
 Beweis von Theorem 2.14 *mit Monoton-Lemma f. Mengen (Satz 1.6)*

- $\mathcal{M} = [A \in \mathcal{A} \mid 1_A \in \mathbf{H}]$  (Indes schon OK)
- $\mathcal{G} \subseteq \mathcal{M}$  wegen ii)
- $\rightarrow \Omega \in \mathcal{M}$  wegen i)
- $\rightarrow A, B \in \mathcal{M}, A \subseteq B \Rightarrow 1_A, 1_B \in \mathbf{H} \stackrel{iii)}{\Rightarrow} 1_B - 1_A \in \mathbf{H}; 1_B - 1_A = 1_{B \setminus A} \in \mathbf{H} \Rightarrow B \setminus A \in \mathcal{M}$
- $\rightarrow A_n \nearrow \in \mathcal{M} \Rightarrow 1_{A_n} \in \mathbf{H} \forall n; \sup 1_{A_n} < \infty \stackrel{iv)}{\Rightarrow} \sup 1_{A_n} \in \mathbf{H} \Rightarrow \bigcup A_n \in \mathcal{M}$

$\mathcal{M}$  ist Dynkin-System; vgl. p. 9 (d-System)

- Satz 1.6:  $\mathcal{A} = \sigma(\mathcal{G}) = \mathcal{D}(\mathcal{G}) \subseteq \mathcal{M}$ ; sofort auch  $= \mathcal{M}$ , da  $\mathcal{A} \subseteq \mathcal{A}$
- $\mathbf{H}$  beinhaltet alle Indikatoren  $\&$  iii); damit alle einfachen ~~Zufallsgrößen~~  $Z \in \mathcal{G}$ .
- wegen iv) ist  $\mathbf{H}$  abgeschlossen bzgl. monoton wachsender L.kanten  $\rightarrow$  wegen Lemma 2.13 beinhaltet  $\mathbf{H}$  alle nicht-negativen Zufallsgrößen auf  $(\Omega, \mathcal{A}, P)$ .
- $X^+, X^- \in \mathbf{H} \Rightarrow X = X_{51}^+ - X^- \in \mathbf{H}$  wegen iii).  $\Rightarrow$  alle Z.B.!

(Reichhaltigkeit von  $\mathbf{H}$  ist abhängig von Reichhaltigkeit von  $\mathcal{A}$ )

Beweisfortsetzung:

**Bemerkung 2.15** [zu **Karr-Theorem 2.22**] Im Buch von Karr hat es hier noch ein Theorem 2.22; es ist falsch:  $1_{A^c}$  ist zum Beispiel nicht in  $H$ , wenn man  $S = \{A\}$  wählt!

## 2.7 Verteilungen, Verteilungsfunktionen und vorgegebene Verteilungen

### 2.7.1 Verteilungen und Verteilungsfunktionen im Fall von Zufallsgrößen ( $n=1$ )

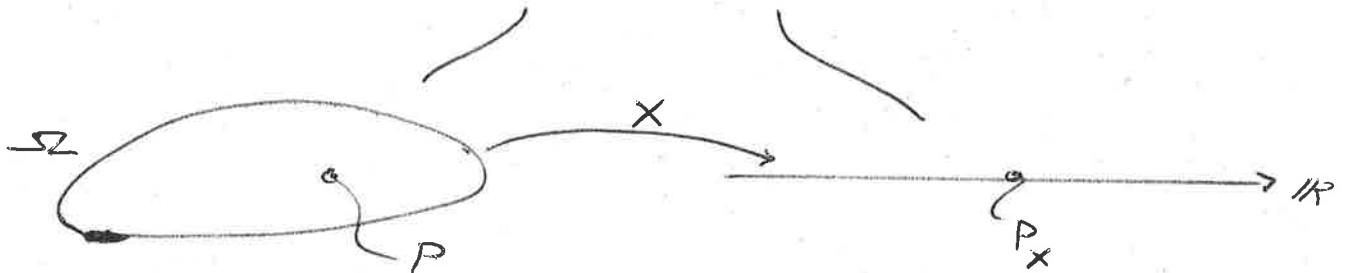
Wir haben in Kapitel 1 (Wahrscheinlichkeit) nach der Wahrscheinlichkeit  $P$  auch die Verteilungsfunktion einer Wahrscheinlichkeit  $F_P$  auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  kennengelernt. Mit Definition 1.14 und Satz 1.15 haben wir eine 1-1-Beziehung zwischen den  $P$  und den  $F_P$  erhalten. Wir haben damit viel Vorarbeit geleistet, welche uns jetzt ein zügiges Vorgehen erlaubt, um diese Konzepte mit dem  $X$  zu verbinden. In Kapitel 1 hatten wir die Zufallsgrößen höchstens zu Motivationszwecken benutzt.

Wir werden in 2.7.3 lernen, dass wir zu jeder vorgegebenen Verteilungsfunktion bzw. Wahrscheinlichkeit immer auch eine Zufallsgröße mit ebendieser Verteilung konstruieren können. Zu jeder Zufallsgröße erhalten wir aber auch kanonisch eine Wahrscheinlichkeit auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ :

Sei  $(\Omega, \mathcal{A}, P)$  ein Wahrscheinlichkeitsraum. Sei  $X$  eine Zufallsgröße auf diesem Wahrscheinlichkeitsraum. In **Aufgabe 33** haben Sie gezeigt: durch

$$P_X(B) := P[X^{-1}(B)] := P[\{\omega | X(\omega) \in B\}]$$

wird eine Wahrscheinlichkeit auf  $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$  definiert. Wir haben damit zusammengefasst folgende 2 Wahrscheinlichkeitsräume  $(\Omega, \mathcal{A}, P)$  und  $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ :



Nebenbei: in der allgemeinen Masstheorie spricht man von einem *Bildmass*;  $P$  wird durch  $X$  abgebildet; man hat dann in diesem Sinn für ein  $B \in \mathcal{B}(\mathbb{R})$  folgende Schreibweise

$$X(P)[B] := P[X^{-1}(B)].$$

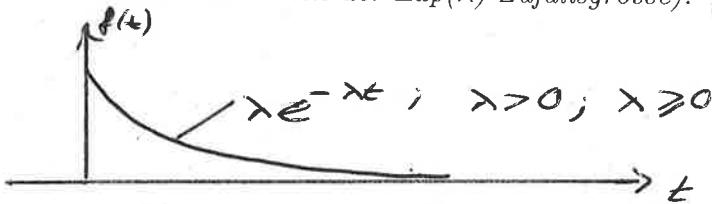
Es folgen noch einige Bezeichnungen, die sich eingebürgert haben:

**Definition 2.16 [Verteilung, Verteilungsfunktion, Survivalfunktion]** Sei  $X$  eine Zufallsgrösse. Dann nennt man

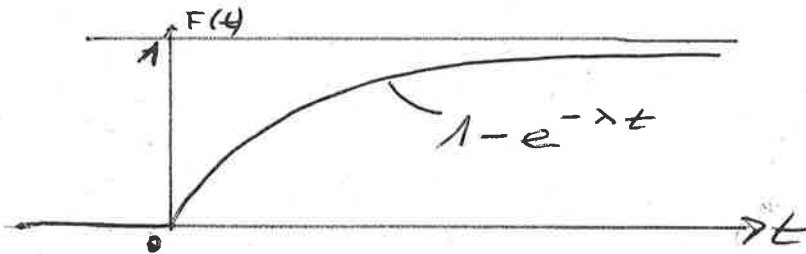
a)  $P_X(B) := P[X^{-1}(B)]$  die Verteilung von  $X$ . Das Wort "Verteilung" wird umgangssprachlich jedoch auch allgemeiner benutzt; man kann sagen,  $X$  hat die Verteilung  $N(\mu, \sigma^2)$  oder hat diese Dichte oder Verteilungsfunktion statt das  $P_X(B)$  anzugeben.

b)  $F_X(t) := P_X((-\infty, t]) := P[X \leq t]$  die Verteilungsfunktion von  $X$ . Englisch: Cumulative Distribution Function (CDF).

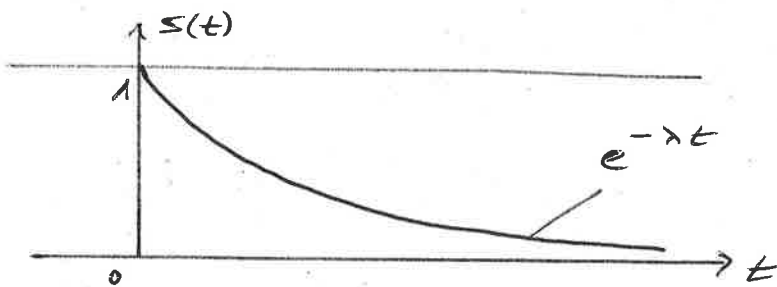
c)  $S_X(t) := 1 - F_X(t) := P[X > t]$  nennt man selbsterklärend die Survivalfunktion (am Besten motiviert mit der  $\text{Exp}(\lambda)$ -Zufallsgrösse): (Wann zerfällt ein Atom)



(Dichte)



(Verteilungsfunktion)



$$T_{1/2} = \frac{1}{\lambda}$$

(Survivalfunktion)

(Anzahl Atome, die noch leben (exp) bzw. W'keit dass ausgew. Atom noch lebt.)

$X$  nennen wir diskret, absolut stetig oder singular stetig, falls das  $P_X$  von der jeweiligen Art ist. Ebenso spricht man im Fall von absolut stetigen Zufallsgrössen  $X$  von der Dichtefunktion  $f_X$ , falls  $f_X = f_P$  und  $f_P$  die Dichtefunktion von  $P_X$  ist.

$\hookrightarrow$  p 25; Korollar 1.21:  $P$  absolut stetig  $\iff \exists f_P \geq 0: \int_{-\infty}^{\infty} f_P = 1$   
 $F_P(t) = \int_{-\infty}^t f_P(s) ds$  (besü X!, Kapitel 1)

Bemerkungen zur Gleichheit von Zufallsgrößen bzw deren Verteilungen:

•  $X \stackrel{D}{=} Y \stackrel{\text{def}}{\iff} F_X(t) = F_Y(t) \quad \forall t \in \mathbb{R}$   
 (gleich in Verteilung)

• vgl. Bemerkung Seite 24:  $f_X(t) = f_Y(t) \quad \forall t \in \mathbb{R}$  muss nicht gelten; nur bis auf Lebesgue-Maßnahmen gleich (bei absolut stetigen ZV)

•  $X \stackrel{\text{a.s.}}{=} Y \stackrel{\text{def}}{\iff} P[\{\omega \mid X(\omega) = Y(\omega)\}] = 1.$

•  $X \stackrel{\text{a.s.}}{=} Y \implies X \stackrel{D}{=} Y$  (vgl. Analysis)

•  $X \stackrel{D}{=} Y \not\Rightarrow X \stackrel{\text{a.s.}}{=} Y$ ; Gegenbsp.

$X \sim N(0,1); Y(\omega) := -X(\omega) \quad \forall \omega \in \Omega.$

$\implies Y \sim N(0,1)$  (gespiegeltes  $X$ )

$\implies X \stackrel{D}{=} Y$ ; aber

$P[\{\omega \mid X(\omega) = Y(\omega)\}] = P[\{\omega \mid X(\omega) = 0\}] = 0 \neq 1.$

$\implies X \not\stackrel{\text{a.s.}}{=} Y$

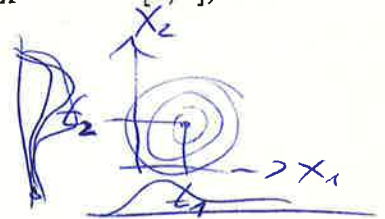
• Ähnlich, aber bitte nicht verwechseln mit Konjugation, Kap. 5

## 2.7.2 Verteilungen und Verteilungsfunktionen im Fall von Zufallsvektoren

**Definition 2.17** [Gemeinsame Verteilungsfunktion; engl. Joint CDF] Sei  $X = (X_1, \dots, X_n)$  ein Zufallsvektor. Dann definieren wir:

- a) Die Verteilung von  $X$  ist die Wahrscheinlichkeit  $P_X(B) := P[X \in B]$  auf  $\mathbb{R}^n$ .
- b) Die Gemeinsame Verteilungsfunktion von  $X$  ist die Funktion  $F_X : \mathbb{R}^n \rightarrow [0, 1]$ , welche folgendermassen definiert ist:

$$F_X(t_1, \dots, t_n) := P[X_1 \leq t_1, \dots, X_n \leq t_n].$$



Auch hier kann man zeigen, dass  $P_X$  durch  $F_X$  eindeutig festgelegt ist (kleine HA). Wir zeigen jetzt, dass wir aus der gemeinsamen Verteilungsfunktion von  $X$  immer die Randverteilungen herausdestillieren können:

**Lemma 2.18** [Gewinnen von  $F_{X_i}$  aus  $F_X$ ] Sei  $X$  ein Zufallsvektor. Dann gilt für alle  $t, i$ :

$$F_{X_i}(t) = \lim_{t_j \rightarrow \infty, j \neq i} F_X(t_1, \dots, t_{i-1}, t, t_{i+1}, \dots, t_n).$$

**Beweis Lemma 2.18** Wenn  $t_j \rightarrow \infty \forall j \neq i$ , dann

$$\{X_1 \leq t_1, \dots, X_{i-1} \leq t_{i-1}, X_i \leq t, X_{i+1} \leq t_{i+1}, \dots, X_n \leq t_n\} \xrightarrow{w/} \{X_i \leq t\}$$

& damit gilt Lemma 2.18 wegen Satz 1.9b); (Monotonie).

Platz kann

Analog zum eindimensionalen Fall, nennen wir einen Zufallsvektor  $X$

\* **diskret**, falls es eine höchstens abzählbare Menge  $C \subseteq \mathbb{R}^n$  gibt, sodass  $P[X \in C] = 1$ . /

\* **absolut stetig**, falls es eine Funktion  $f_X : \mathbb{R}^n \rightarrow \mathbb{R}_+$  (gemeinsame Dichte - im Gegensatz zu den Rand-Dichten) gibt, sodass

$$P[X_1 \leq t_1, \dots, X_n \leq t_n] = \int_{-\infty}^{t_1} \dots \int_{-\infty}^{t_n} f_X(u_1, \dots, u_n) du_1 \dots du_n. \quad (\otimes)$$

*(d.h. muss als (inkludiert) Integral darstellbar sein)*

Im Fall der diskreten Zufallsvektoren gilt (Vorsicht beim vermuteten absolut stetigen Pendant):

**Lemma 2.19 [diskreter Zufallsvektor und diskrete Einzelkomponenten]** Ein Zufallsvektor ist genau dann diskret, wenn alle seine Einzelkomponenten diskrete Zufallsgrößen sind.

**Beweis Lemma 2.19**



" $\Rightarrow$ " Sei  $C \subseteq \mathbb{R}^n$  <sup>höchstens</sup> abzählbar mit  $P[X \in C] = 1$ .

$\Rightarrow P[X_i \in C_i] = 1$  wo  $C_i := \{a_j \mid a \in C\}$   
 $a = (a_1, \dots, a_n)$

Sei  $\forall i$  <sup>abzählbare</sup>  $C_i \subseteq \mathbb{R} : P[X_i \in C_i] = 1 \Rightarrow$

$P[X \in C] = 1$  wo  $C = \prod_{j=1}^n C_j$   
 wegen A 2.1)

# Beispiel

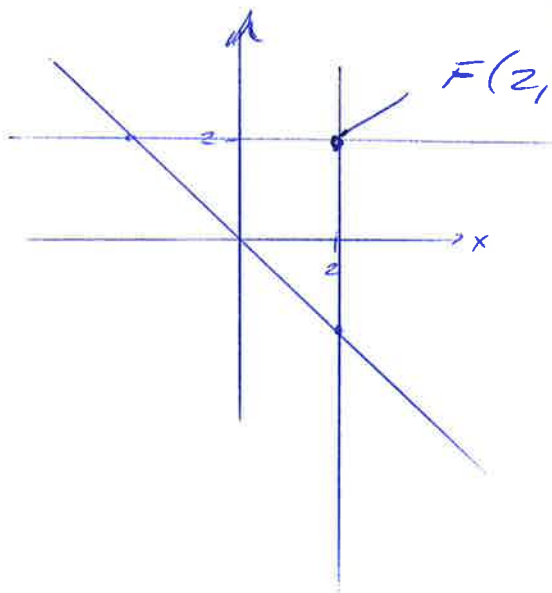
$$X \sim \mathcal{N}(0, 1)$$

$$Y(\omega) := -X(\omega) \quad \forall \omega \in \Omega \quad (\text{ger. stetig})$$

$$Y \sim \mathcal{N}(0, 1)$$

$X, Y$  je absolut stetig

$(X, Y)$  nicht 2-dim abs. stetig:



$$F(2, 2) \stackrel{95\%}{\neq} \int_{-\infty}^2 \int_{-\infty}^2 f_{(X, Y)}(u, v) \, du \, dv \quad (2 \doteq 1.96)$$

nicht als Lebesgue-Integral darstellbar, da man auf Nullmenge  $\neq \emptyset$ : Fo III, S 7, 54

(Gerade  $x = -y$  ist Nullmenge in  $\mathbb{R}^2$ )

**Lemma 2.20 [absolut stetiger Zufallsvektor und absolut stetige Einzelkomponenten]** Sei  $X = (X_1, \dots, X_n)$  ein absolut stetiger Zufallsvektor. Dann gilt für alle Einzelkomponenten, dass auch diese absolut stetig sind und es gilt:

$$f_{X_i}(t) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_X(u_1, \dots, u_{i-1}, t, u_{i+1}, \dots, u_n) du_1 \dots du_{i-1} du_{i+1} \dots du_n.$$

analog Würfeln:  $P[X=6] = \sum_{j=1}^6 P[X=6, Y=j]$  (unabh.  $\rightarrow$  nötig!); Bild 

**Beweis Lemma 2.20 und Gegenbeispiele zur vermuteten Umkehrung**

• Von Lemma 2.18 &  $\otimes$ .

$$P[X_i \leq t] = \lim_{\substack{t_j \rightarrow \infty \\ j \neq i}} P[X_1 \leq t_1, \dots, X_i \leq t, \dots, X_n \leq t_n]$$

$$\stackrel{\otimes}{=} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_X(y_1, \dots, y_{i-1}, t, y_{i+1}, \dots, y_n) dy_1 \dots dy_n$$

$$\rightarrow \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f_X(y_1, \dots, y_{i-1}, t, y_{i+1}, \dots, y_n) dy_1 \dots dy_{i-1} dy_{i+1} \dots dy_n \right] dt$$

$f_X \geq 0 \Rightarrow$   
Kettenregel  
vertauscht  
Analyse

$\rightarrow$  fastig, weise p 54 finden!

~~$\Rightarrow$  Hauptsatz Diff. folgt Lemma 2.20~~  
 ~~$\frac{d}{dt} \int_{-\infty}^t f(s) ds = f(t)$~~   
 Probe:  $f(x) = e^{-x}$   
 $\int_{-\infty}^t e^{-s} ds = -e^{-s} \Big|_{-\infty}^t = -e^{-t} - (-\infty) = e^{-t}$   
 $\frac{d}{dt} e^{-t} = -e^{-t} = -f(t)$   
 $\Rightarrow$  absolut stetig

- allg.: um etwas über Dichten zu zeigen:  $F_X$  nehmen;  $f_X$  durch  $F_X$  ableiten
- Bsp. Seite 55:  $X \sim N(0,1) \sim Y$ ;  $(X, Y)$  nicht absolut stetig!  $F_0$  III, 37, 54

In Karr folgen an dieser Stelle noch die wichtigsten Verteilungen. Wir haben dies in der WTS (weitere in SM, AS) besprochen. Lesen Sie vielleicht nochmals WTS-Kapitel 4 durch. Auch die Transformation von stetigen Zufallsgrößen wurde bereits in WTS-Kapitel 2 (2.6) besprochen. Karr geht noch unvollständig auf die mehrdimensionalen Transformationen ein; dies machen wir nur einmal in der Vlsg SM.

### 2.7.3 Vorgegebene Verteilungen

Bisher haben wir Zufallsgrößen  $X$  einfach als gegeben betrachtet. Wir haben Sätze geschrieben der Art: "sei  $X$  standardnormalverteilt", also eine  $\mathcal{N}(0, 1)$ -Zufallsgröße. Dies konnten wir zu Recht machen, denn es gilt

*dan, oder p geben Anwandern!*  
**Satz 2.21** [Existenz von  $(\Omega, \mathcal{A}, P)$  und  $X$  zu gegebenem  $F$ ] Sei  $F$  eine Verteilungsfunktion auf  $\mathbb{R}$ . Dann existiert ein Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  und eine Zufallsgröße  $X$  so, dass  $F_X = F$ .

**Beweis Satz 2.21** (handschriftlich bereits in 2.5 behandelt)

• Wähle  $(\Omega, \mathcal{A}) := (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ .

• Satz 1.17:  $\exists! P: F_P = F$

•  $X(\omega) = \omega$  ist Zufallsgröße

$$F_X(t) := P[X \leq t] = P[(-\infty, t]] = F_P(t) = F(t)$$

$$\{\omega \mid X(\omega) \leq t\} = \{\omega \mid \omega \leq t\} = (-\infty, t]$$

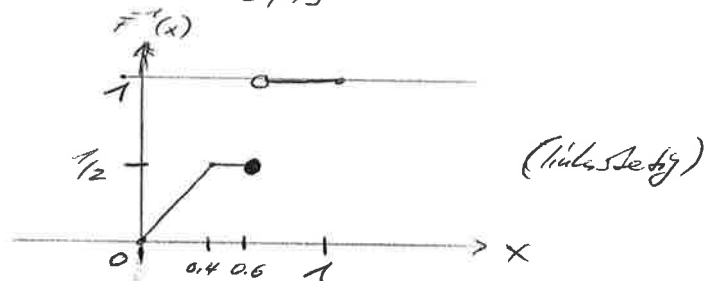
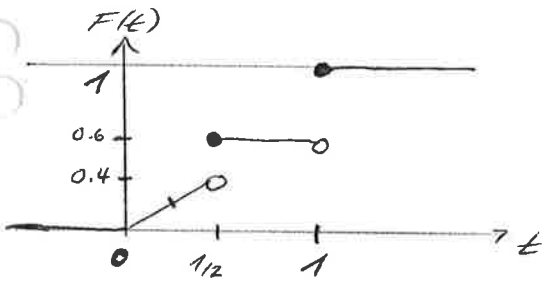
Satz 2.21 ist jedoch nur eine Existenzaussage, wir wollen noch eine *konstruktive* Methode herleiten. Dazu führen wir Konzepte ein, welche in der Vorlesung SM in Kapitel 4 auch verwendet werden.

**Definition 2.22 [Inverse von  $F$  oder Quantil-Funktion von  $F$ ]** Wir definieren die Inverse (oder Quantil-Funktion) einer Verteilungsfunktion  $F$  als

$$F^{-1}(x) := \inf\{t : F(t) \geq x\}, \quad x \in (0, 1).$$

Paar Bemerkungen und Bilder, um dieses Konstrukt besser kennenzulernen:

- $F$  darf Sprünge haben & flache Abschnitte ( $F' = 0$  über ganze Intervalle)
- $F^{-1}(x)$  existiert (Aufbau einer Menge)
- $F^{-1}$  springt wo  $F$  flach & vice versa
- $F$  muss offensichtlich nicht bijektiv sein von  $\mathbb{R} \rightarrow [0, 1]$



- $F^{-1}(0.43) := \inf_t \{t \mid F(t) \geq 0.43\} = \inf_t \{t \mid t \geq 0.5\} = 0.5$  [zwischen den]
- $F^{-1}(0.6) := \inf_t \{t \mid F(t) \geq 0.6\} = \inf_t \{t \mid t \geq 0.5\} = 0.5$  [wo flach]
- $F^{-1}(0.1) := \inf_t \{t \mid F(t) \geq 0.1\} = \inf_t \{t \mid t \cdot \frac{4}{5} \geq 0.1\} = \inf_t \{t \geq \frac{5}{40}\} = \frac{5}{40}$  [normaler Wert]

Wir listen hier ein paar elementare Eigenschaften von  $F^{-1}$  auf, welche wir gleich benutzen werden. Sie beweisen diese Aussagen in Übungsblatt 9.

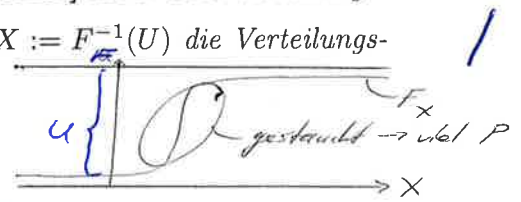
**Lemma 2.23 [elementare Eigenschaften von  $F^{-1}$ ]** Sei  $F^{-1}$  die Inverse von  $F$ . Dann gelten:

- Für alle  $(x, t)$  gilt  $F^{-1}(x) \leq t \Leftrightarrow x \leq F(t)$ .
- $F^{-1}$  ist monoton wachsend und links-stetig.
- Falls  $F$  stetig ist, dann gilt  $F(F^{-1}(x)) = x$  für alle  $x \in (0, 1)$ .

Jetzt kommt - im Gegensatz zu Satz 2.21 - eine konstruktive Methode. Dazu ein paar Vorbemerkungen:

1. Im Statistik-Paket R (vgl VlsG WTS) und in anderen Statistik-Paketen haben Sie alle wichtigen Zufallsgrößen bereits implementiert.
2. Man kann sich aber einerseits fragen, wie die EntwicklerInnen von R dies gemacht haben und andererseits kann es auch sein, dass Sie in einer anderen Programmierumgebung dies zu Fuss selber machen müssen.
3. Was jetzt folgt berücksichtigt *nicht* allfällige algorithmische Probleme wegen der Rechengeschwindigkeit.
4. Im Semesterapparat finden Sie das Buch "Introduction to Stochastic Calculus Applied to Finance" von Lamberton/Lapeyre. Dort hat es in Kapitel 8 weitere Angaben zur Programmierung von Zufallsgrößen, Zufallsvektoren und -Prozessen.
5. Wir setzen im Folgenden voraus, dass Sie eine  $U[0, 1]$ -Zufallsgröße bereits besitzen. Diese erhalten Sie in meist genügender Präzision zum Beispiel, in dem Sie einen Zufalls-generator für natürliche Zahlen von 1 bis  $N$  haben und dann das Resultat durch  $N$  teilen.
6. Die jetzt folgende Methode "Quantil-Transformation" (eher Statistik) nennt man auch "Inverse Distribution Function"-Methode (eher Stochastik).

*Anwendung!*  
**Lemma 2.24 [Inverse Distribution Function-Methode]** Sei  $F$  eine Verteilungsfunktion auf  $\mathbb{R}$ . Sei  $U$  eine  $U[0, 1]$ -Zufallsgröße. Dann hat  $X := F^{-1}(U)$  die Verteilungsfunktion  $F$ .

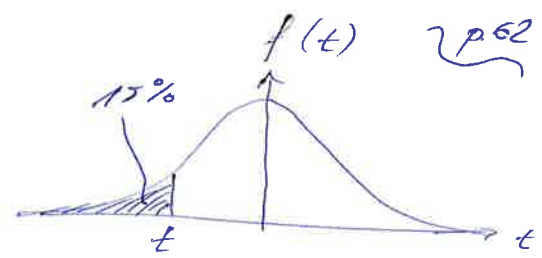


**Beweis von Lemma 2.24 und Beispiel ( $\exp(\lambda)$ )**

- $F^{-1}$  ist monoton wegen Lemma 2.23 b)  $\rightarrow$  wegen  $A_{\mathbb{R}}$  auch Borel  $\Rightarrow F^{-1}(u)$  ist also ZG
- $F_X(t) := P[X \leq t] = P[F^{-1}(U) \leq t] \stackrel{L. 2.23 a)}{=} P[U \leq F(t)] = F(t)$   $\checkmark$   $U \sim U[0, 1]$   $ged$
- $FS 14$
- wollen aus  $U[0, 1]$  eine  $\exp(\lambda)$  machen:  $F_X = \int_{(0, \infty)} (t) \cdot \lambda \cdot e^{-\lambda t} \rightarrow F^{-1}(u) = -\frac{\log(u)}{\lambda}$ ;  $u \in (0, 1)$   $\leftarrow$  Bilden
- $\Rightarrow -\frac{\log(u)}{\lambda} \sim \exp(\lambda)$ , wo  $U \sim U[0, 1]$ ; Bild dazu!

Sie sehen anhand des Beispiels auch, dass Sie entweder  $F^{-1}$  in geschlossener Form bereits haben müssen oder eine gute numerische Approximation kennen.

- ~~Beide~~
- $n = 100$  ;  $H_0: Y \geq \mu_0 = 70g$  ;  $t$ -Test:
- $N(\mu, \sigma^2)$  ;  $H_1: Y < \mu_0 = 70g$   
 $\hookrightarrow$  je unbekannt



- Teststatistik:

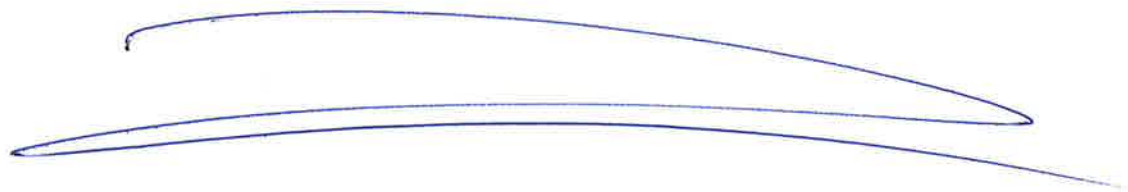
$$T = \frac{\sqrt{n}(\bar{X} - \mu_0)}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}} \stackrel{H_0}{\sim} t_{n-1}$$

z.B.

- Was ist der P-Wert?  $t$ -Resultat:  
 $\hookrightarrow$  Realisation

$$P[T \leq t] = F_T(t) ; \text{ also fest}$$

$$\text{das Zufallsgeschehen} \rightarrow F_T(T) \sim U[0, 1]$$





Wir schliessen dieses Kapitel ab mit Bemerkungen zu Folgen von Zufallsgrössen. Wir brauchen dies in Kapitel 5 und vor allem in der VlsG AS.

Nehmen wir einmal an, wir haben mathematisch sauber ein  $(\Omega, \mathcal{A}, P)$  und eine Folge von Zufallsgrössen  $(X_n)_{n \geq 0}$  konstruiert. Dann muss wegen der Stetigkeit von  $P$  sicher gelten:

$$P[X_1 \leq t_1, \dots, X_n \leq t_n] = \lim_{t \rightarrow \infty} P[X_1 \leq t_1, \dots, X_n \leq t_n, X_{n+1} \leq t].$$

Falls wir zu einer gegebenen Folge von Verteilungsfunktionen  $(F_n)_{n \geq 0}$  eine Folge von Zufallsgrössen  $(X_n)_{n \geq 0}$  konstruieren wollen, müssen wir also sicher fordern, dass

$$\lim_{t \rightarrow \infty} F_{n+1}(t_1, \dots, t_n, t) = F_n(t_1, \dots, t_n).$$

In der Tat ist diese *Konsistenzbedingung* auch genügend. Dies ist ein fundamentales Resultat aus der Masstheorie und lautet (Beweis in A.N. Sirjaev: Wahrscheinlichkeit)

*das ist das beste; geben Anweisung an! AS*  
**Satz 2.27 [Satz von Kolmogorov über die Existenz stochastischer Prozesse]**  
 Für alle  $n$  gelte, dass  $F_n$  eine Verteilungsfunktion auf  $\mathbb{R}^n$  ist. Es gelte zudem die *Konsistenzbedingung*

$$\lim_{t \rightarrow \infty} F_{n+1}(t_1, \dots, t_n, t) = F_n(t_1, \dots, t_n)$$

für alle  $n$  und  $(t_1, \dots, t_n)$ . Dann gibt es einen Wahrscheinlichkeitsraum  $(\Omega, \mathcal{A}, P)$  und eine Folge von Zufallsgrössen  $(X_n)_{n \geq 0}$  so, dass  $F_n$  für alle  $n$  die Verteilungsfunktion von  $(X_1, \dots, X_n)$  ist.