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 λ 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 unter einer Abbildung und fordern vorerst *nicht*, dass die Abbildung messbar (d.h. eine Zufallsgrösse) ist.

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

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

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 $\{B_\alpha | \alpha \in I\}$ Borel-Mengen. Dann gelten:*
a) Sei $A \subset B$, dann auch $X^{-1}(A) \subset X^{-1}(B)$
b) $X^{-1}(\cup_I B_\alpha) = \cup_I X^{-1}(B_\alpha)$
c) $X^{-1}(\cap_I B_\alpha) = \cap_I X^{-1}(B_\alpha)$
d) Falls $A \cap B = \phi$, dann auch $X^{-1}(A) \cap X^{-1}(B) = \phi$
e) $X^{-1}(A^c) = [X^{-1}(A)]^c$
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:

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

Definition 2.3 $[\mathcal{B}(\mathbb{R}^n)]$ *Die Borel-σ-Algebra auf $\mathbb{R}^n$ ist diejenige σ-Algebra $\mathcal{B}(\mathbb{R}^n)$, welche vom π-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 $\{\prod_{i=1}^{n} B_i | B_i \in \mathcal{B}(\mathbb{R})\}$. Man beachte, dass $\{\prod_{i=1}^{n} B_i | B_i \in \mathcal{B}(\mathbb{R})\} \subsetneqq \mathcal{B}(\mathbb{R}^n)$; Ausnahmebeispiel:

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

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

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 $f : \mathbb{R}^{n-1} \to \mathbb{R}$ mit kompaktem Träger (vgl Vlsg 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 $X : \Omega \to \mathbb{R}$ mit der Eigenschaft, dass*

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

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

Warum diese mb?

Obige Definition ist übrigens ein Spezialfall von

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). $g : E_1 \to E_2$. 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.

Bemerkungen zu Definition 2.4:

1. Zufallsgrössen nennt man auch Zufallsvariablen.

2. In der Vlsg WTS (weitgehend auch in AS, SM) haben wir uns nicht um die mb gekümmert. Dies wird jetzt anders. Um jedoch die beiden Teile auseinanderzuhalten: für die Anwendungen und Ihre Vorstellungswelt ist eine Zufallsgrösse einfach eine Funktion $X : \Omega \to \mathbb{R}$; damit die Mathematik nicht schiefgeht, fordern wir dazu noch die mb.

Beispiele (Indikatorfunktion(en), konstante Funktionen, einfache Zufallsgrössen):

2.4 Ein paar weitere, verwandte Definitionen

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

Zufallsvektoren werden wir vor allem auch in der Vorlesung SM benutzen: Wenn wir Daten $(x_1, \ldots, x_n)$ haben, so stellen wir uns vor, diese Daten sind Realisationen eines Zufallsvektors $\mathbf{X} = (X_1, \ldots, X_n)$, also $\mathbf{X}(\omega) = (X_1(\omega), \ldots, X_n(\omega)) = (x_1, \ldots, x_n)$, für ein ω, 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 Statistikteil 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-σ-Algebra, Filtration] *Seien $\mathcal{A}$ und $\mathcal{F}$ beide σ-Algebren. Wir sagen, $\mathcal{A}$ ist eine Sub-σ-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-σ-Algebren von $\mathcal{A}$. Wir nennen eine Familie $(\mathcal{A}_t)_{t \in T}$ eine Filtration (in $\mathcal{A}$), wenn

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

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.*

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 $\mathbb{R}$

$$t \to X_t(\omega)$$

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:

2.5 Von Zufallsgrössen erzeugte σ-Algebren

Das folgende Resultat haben wir auf Blatt 2 im Honours-Teil 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 σ-Algebra] *Sei X eine Zufallsgrösse. Die Familie*

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

ist eine σ-Algebra auf Ω. Man nennt sie "die von X erzeugte σ-Algebra".

Beweis Lemma 2.9:

Beispiele zu "von X erzeugte σ-Algebra" 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* $X : \Omega \to \mathbb{R}$ *mit der Eigenschaft, dass* $\{\omega \in \Omega | X(\omega) \leq a\} \in \mathcal{A}$ *für alle reellen* a. *Die geforderte Eigenschaft nennt man Messbarkeit.*

Wie ist das jetzt mit unserer Definition 2.4?

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 Ω 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:

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

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

Beweis von Lemma 2.10 (Fortsetzung)

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.*

b) $\limsup_n X_n$, $\liminf_n X_n$ *sind Zufallsgrössen.*

c) Falls $X(\omega) := \lim_{n\to\infty} X_n(\omega)$ *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

Kleine Bemerkung:

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

Beweis von Lemma 2.12

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 $0 \leq X_1 \leq X_2 \ldots$ sodass $X_n(\omega) \uparrow X(\omega)$ für alle $\omega \in \Omega$.*

Beweis von Lemma 2.13

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 Zufallsgrössen enthält.

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

i) $1 \in \mathbf{H}$ *(konstante Funktion ist drin)*
ii) $\mathbf{1}_A \in \mathbf{H}$ *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.*

Beweis von Theorem 2.14

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 34 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) $P_X(B) := P[X^{-1}(B)], B \in \mathcal{B}(\mathbb{R})$, die Verteilung von X. Das Wort "Verteilung" wird umgangssprachlich jedoch auch allgemeiner benutzt; man kann sagen, X hat die Verteilung $\mathcal{N}(\mu, \sigma^2)$ 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 Exp(λ)-Zufallsgrösse):

X nennen wir diskret, absolut stetig oder singulär 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.

Bemerkungen zur Gleichheit von Zufallsgrössen bzw deren Verteilungen:

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 $F_X : \mathbb{R}^n \to [0,1]$, 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:*

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

Beweis Lemma 2.18

Analog zum eindimensionalen Fall, nennen wir einen Zufallsvektor X

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

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

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

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

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

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

Beweis Lemma 2.20 und Gegenbeispiel zur vermuteten Umkehrung

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

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)

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*

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

Ein paar Bemerkungen und Bilder, um dieses Konstrukt besser kennenzulernen:

Wir listen hier ein paar elementare Eigenschaften von F^{-1} auf, welche wir gleich benutzen werden. Sie beweisen diese Aussagen in Übungsblatt 8.

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

a) Für alle (x,t) gilt $F^{-1}(x) \leq t \Leftrightarrow x \leq F(t)$.

b) F^{-1} ist monoton wachsend und links-stetig.

c) 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 Zufallsgenerator 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).

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 ($\exp(\lambda)$)

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

Zu Lemma 2.24 gibt es auch eine Umkehrung. In der Vorlesung SM sehen wir mit Hilfe von WT-Lemma 2.25 in SM-Kapitel 4 über Testtheorie, dass bei stetiger Teststatistik unter der Nullhypothese der P-Wert eine $U[0,1]$-Verteilung besitzt.

Lemma 2.25 [Verteilung von $F_X(X)$ wenn F_X stetig] *Sei F_X stetig, dann hat $F_X(X)$ eine $U[0,1]$-Verteilung.*

Beweis Lemma 2.25

Ohne Beweis fügen wir noch das Pendant für Vektoren von Satz 2.21 an:

Satz 2.26 [Existenz von $(\Omega, \mathcal{A}, P)$ und $\mathbf{X}$ zu gegebenem F] *Sei $F : \mathbb{R}^n \to [0,1]$ eine n-dimensionale Verteilungsfunktion. Dann existiert ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{A}, P)$ und ein Zufallsvektor $\mathbf{X} := (X_1, \ldots, X_n)$ so, dass $F_{\mathbf{X}} = F$.*

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:

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

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

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

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

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*

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

für alle n und $(t_1, \ldots, t_n)$. 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 $(X_1, \ldots, X_n)$ ist.