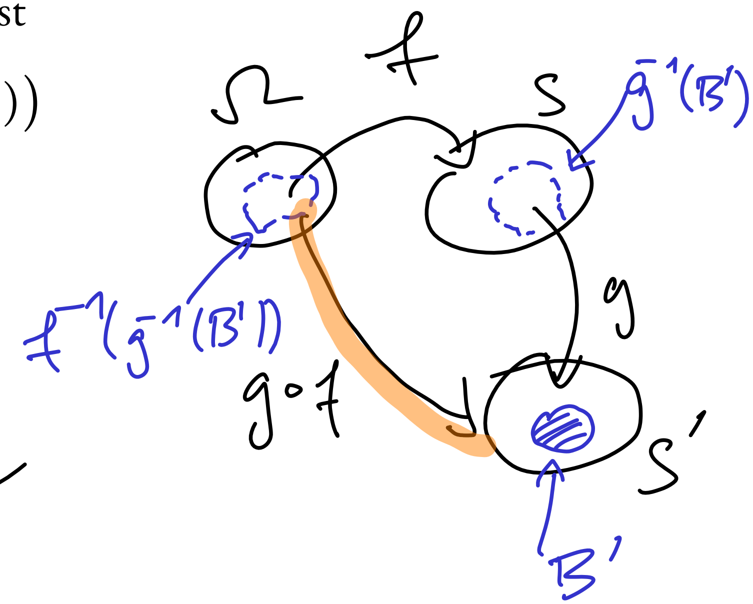
Beobachtung 1.32 (Verkettung messbarer Abbildungen, Funktionen von Zufallsvariablen).
$(S, \mathcal{S})$, $(S', \mathcal{S}')$, messbare Räume, $f : \Omega \to S$, $g : S \to S'$ messbar, dann ist

$$g \circ f : \Omega \to S', \quad g \circ f(\omega) = g\big(f(\omega)\big)$$

$(\mathcal{F}$-$\mathcal{S}$-)messbar. Für $B' \in \mathcal{S}'$ ist

$$(g \circ f)^{-1}(B') = f^{-1}\big(\underbrace{g^{-1}(B')}_{\in \mathcal{S}}\big) \in \mathcal{F}$$

✓



Insbesondere ist für eine S-wertige ZV X auf $(\Omega, \mathcal{F}, P)$

$$g(X) := g \circ X$$

wiederum eine Zufallsvariable. (mit Wertebereich S')

(Interpretation: Wir werten die Funktion g an der zufälligen Stelle X aus.)

(Erinnerung: Es muss gelten $\{g(X) \in B'\} \overset{!}{\in} \mathcal{F}$,)
d.h. $g \circ X$ muss $(\mathcal{F}-\mathcal{S}')$-messbar sein.

Beobachtung und Definition 1.33. *1. (Verteilung) X ZV (auf einem W'raum $(\Omega, \mathcal{F}, P)$) mit Werten in S,*

$$P_X(B) := P\big(\{X \in B\}\big), \quad B \in \mathcal{S}$$

definiert ein W'maß auf $(S, \mathcal{S})$ (Übung), wir nennen P_X die Verteilung *von X (unter P) und schreiben auch $\mathscr{L}_P(X) := P_X$, oft auch nur $\mathscr{L}(X)$, wenn P fixiert ist oder aus dem Kontext klar.*

(Das $\mathscr{L}$ erinnert an English "law" bzw. Französisch ≪loi≫, d.h. „Gesetz".)

2. μ ein W'maß auf S, wir schreiben $X \sim \mu$, wenn $\mathscr{L}_P(X) = \mu$.

3. X und Y ZVn mit Werten in S heißen identisch verteilt, *wenn $\mathscr{L}_P(X) = \mathscr{L}_P(Y)$ (man schreibt dies auch als $X \stackrel{d}{=} Y$)*

($\stackrel{d}{=}$ für "equal in distribution")

4. $X_1, X_2, \dots, X_d$ (reelle) ZVn,

$$Y := \big(X_1, X_2, \dots, X_d\big)$$

(Y hat Wertebereich $\mathbb{R}^d$)

so heißt $\mathscr{L}_P(Y)$ die gemeinsame Verteilung *der $X_1, X_2, \dots, X_d$, $\mathscr{L}_P(X_i)$ heißt die i-te* Randverteilung *(oder* Marginalverteilung*) von Y.*

Schreibweise. Wir kürzen (auch im Folgenden) oft ab

$$P(X \in B) := P\big(\{X \in B\}\big), \quad P(X = x) := P\big(\{X = x\}\big),$$
$$P(X_1 \in B_1, X_2 \in B_2) := P\big(\{X_1 \in B_1\} \cap \{X_2 \in B_2\}\big)$$

etc.

Beispiel 1.34. 1. Für $A \in \mathcal{F}$ ist $\mathscr{L}_P(\mathbf{1}_A) = \mathrm{Ber}_{P(A)}$.

$\left(\mathbb{1}_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \notin A \end{cases} \right)$

Wertebereich $\{0,1\} = S$

$P(\{\mathbb{1}_A = 1\}) = P(A)$

2. Sei $\Omega = \{0,1\}^n$, $P = \mathrm{Ber}_p^{\otimes n}$ (aus Bsp. 1.20, 2.), $X_i(\omega) = \omega_i$, $Y := X_1 + X_2 + \dots + X_n$, so ist

$$\mathscr{L}_P(X_i) = \mathrm{Ber}_p, \ \mathscr{L}_P(Y) = \mathrm{Bin}_{n,p}$$

$\Big(P(\{\underbrace{(\omega_1, \dots, \omega_n)}_{=\omega}\}) = p^{\#\{i \le n:\, \omega_i = 1\}} (1-p)^{\#\{i \le n:\, \omega_i = 0\}} \Big)$

$P(X_i = 1) = P(X_i^{-1}(\{1\})) = P(\{\omega \in \Omega : \omega_i = 1\})$

$= \sum_{\omega_1, \omega_2, \dots, \omega_{i-1}, \omega_{i+1}, \dots, \omega_n = 0}^{1} P(\{(\omega_1, \dots, \omega_{i-1}, 1, \omega_{i+1}, \dots)\}) = p$

3. (Augensumme beim zweifachen Münzwurf, vgl. auch Bsp. 1.4 1. (b) und Bsp. 1.27)
W_1 und W_2 das Ergebnis des ersten bzw. des zweiten Wurfs beim zweimaligen Werfen eines fairen Würfels und

$$X := W_1 + W_2 \text{ die Augensumme}$$

(z.B. formalisiert via $\Omega = \{1,2,\dots,6\}^2$ mit $p\big((\omega_1,\omega_2)\big) = 1/36$ für $(\omega_1,\omega_2) \in \Omega$, $W_i\big((\omega_1,\omega_2)\big) = \omega_i$, $i = 1,2$), so hat X Wertebereich $S_X = \{2,3,\dots,12\}$ und Verteilung $\mathscr{L}(X) =: \mu$ mit

$$\mu(\{x\}) = P(X = x) = \sum_{(w_1,w_2):\, w_1+w_2=x} P(W_1 = w_1, W_2 = w_2)$$
$$= \frac{\#\{(w_1,w_2) \in \{1,2,\dots,6\}^2 : w_1 + w_2 = x\}}{36} = \frac{6 - |7-x|}{36}, \quad x \in S_X$$

$P(Y = k) = P(X_1 + \dots + X_n = k)$, $0 \le k \le n$

$= P(\{\omega \in \Omega : \omega_1 + \dots + \omega_n = k\})$

$= \binom{n}{k} p^k (1-p)^{n-k}$

Bemerkung 1.35. Die Randverteilungen legen (i.A.) nicht die gemeinsame Verteilung fest, z.B.: Wir haben eine faire Münze M_1 und zwei gezinkte Münzen M_2, M_3, wobei

$$P(M_2 = K) = \frac{3}{4}, \; P(M_3 = K) = \frac{1}{4}.$$

Wir werfen erst M_1, wenn M_1 K (Kopf) zeigt, so werfen wir dann M_2, sonst M_3.
Sei

$$X_i = \text{Resultat des } i\text{-ten Wurfs}, \; i = 1, 2$$

Die gemeinsame Verteilung von (X_1, X_2) ist

((X_1, X_2) hat Wertebereich $\{K, Z\} \times \{K, Z\}$)

$X_1 \backslash X_2$	K	Z	
K	$\frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}$	$\frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$	$\frac{1}{2}$
Z	$\frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$	$\frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	

$P(X_2 = K)$ " $P(X_1 = K, X_2 = K)$ $+ P(X_1 = Z, X_2 = K)$ "

also $P(X_1 = K) = P(X_2 = K) = \frac{1}{2}$ und dieselben Randverteilungen ergeben sich, wenn man zwei Mal M_1 wirft, aber die gemeinsame Verteilung wäre eine andere, nämlich:

	K	Z
K	$\frac{1}{2} \cdot \frac{1}{2}$	$\frac{1}{2} \cdot \frac{1}{2}$
Z	$\frac{1}{4}$	$\frac{1}{4}$

Bemerkung 1.36 (Kanonisches Modell für eine ZV)**.** Man kann eine Zufallsvariable X mit Wertebereich S und Verteilung μ stets in „kanonischer Weise“ mit der Wahl

$$\Omega = S \text{ und } P = \mu \text{ und geeignetem } \mathcal{F} \subset 2^S$$

(im diskreten Fall kann man $\mathcal{F} = 2^S$ wählen) als

$$X = \mathrm{Id}_S \text{ auf dem W'raum } (S, \mathcal{F}, \mu)$$

formulieren.

Man kann daher genausogut die mathematische Modellierung eines Zufallsphänomens mit der Formulierung geeigneter Zufallsvariablen samt Verteilung beginnen
(dies ist der in dem Buch von G. Kersting und A. Wakolbinger [KW] beschrittene Weg).

Definition 1.37. X ZV mit Werten in $\mathbb{R}$

$$F_X : \mathbb{R} \to [0,1], \quad F_X(x) = P(X \le x), \; x \in \mathbb{R}$$

heißt die *Verteilungsfunktion* von X. (Strenggenommen: die Verteilungsfunktion von $\mathscr{L}_P(X)$)

Beobachtung 1.38 (Erzeugung reeller ZVn mit vorgegebener Verteilung)**.** *Sei $F : \mathbb{R} \to [0,1]$ eine Verteilungsfunktion,*

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

die inverse Verteilungsfunktion oder Quantilfunktion (aus Bem. 1.13), U reelle ZV, $U \sim \mathrm{Unif}_{[0,1]}$, dann hat

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

die Verteilungsfunktion $F_X = F$.

Bew.: Erinnerung $F^{-1}(t) \le x \iff t \le F(x)$,

vgl. Bem. 1.13

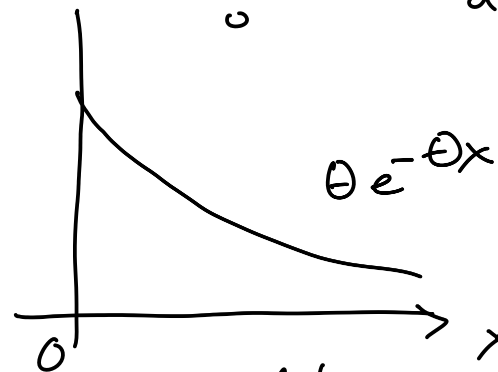
Sei $x \in \mathbb{R}$:

$$P(X \le x) = P(F^{-1}(U) \le x) = P(U \le F(x))$$
$$= P(0 \le U \le F(x)) = F(x) - 0 = F(x)$$

✓

Beispiel 1.39. Exp_θ hat Verteilungsfunktion , $\theta > 0$

$$F_{\mathrm{Exp}_\theta}(x) = (1 - e^{-\theta x})\mathbf{1}_{[0,\infty)}(x) = \int_0^x \theta e^{-\theta y}\, dy$$



Dichte von Exp_θ

mit inverser Funktion

$$F^{-1}_{\mathrm{Exp}_\theta}(t) = -\frac{1}{\theta}\log(1-t)$$

also ist $-\frac{1}{\theta}\log(1-U) \sim \mathrm{Exp}_\theta$ (und natürlich ebenso $-\frac{1}{\theta}\log(U)$)

$(U \sim \mathrm{Unif}_{[0,1]})$

denn

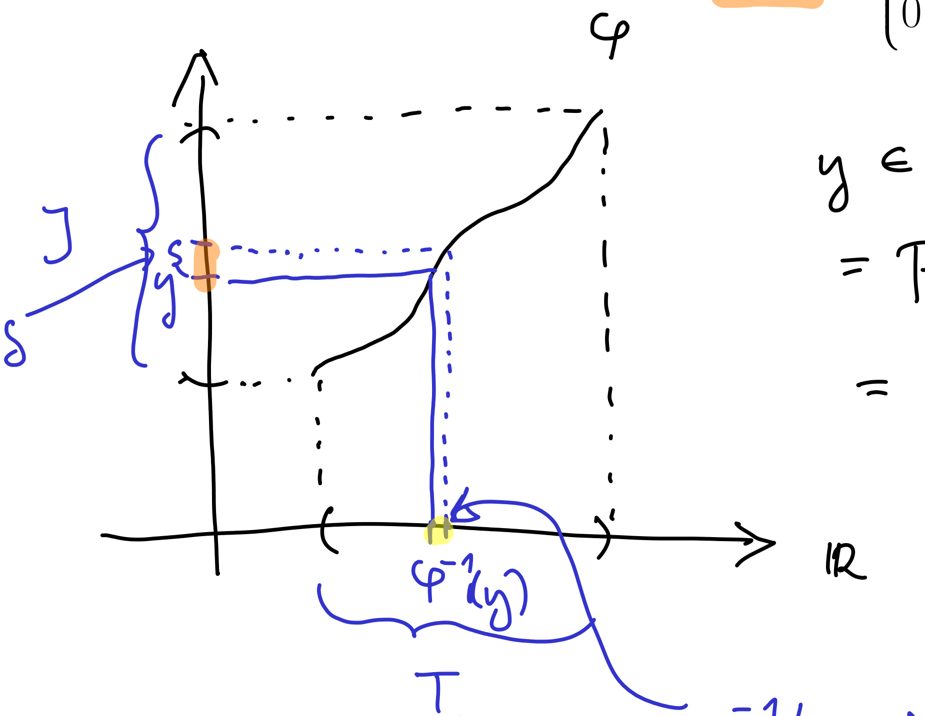
$$1 - U \overset{d}{=} U$$

$$\Big(P(1-U \leq u) = P(U \geq 1-u) = \underbrace{P(U \leq 1)}_{=1} - \underbrace{P(U < 1-u)}_{=1-u} = u \Big)$$

Beobachtung 1.40 (Dichtetransformation im Fall $\mathbb{R}^1$). *X reelle ZV mit Dichte f_X, d.h. $F_X(x) = \int_{-\infty}^{x} f_X(z)\,dz$, $I \subset \mathbb{R}$ (möglicherweise unbeschränktes) offenes Intervall mit $P(X \in I) = 1$, $J \subset \mathbb{R}$, $\varphi : I \to J$ stetig differenzierbar, bijektiv.*

Dann hat $Y := \varphi(X)$ die Dichte

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



$$y \in J: \quad P(Y \le y) = P(\varphi(X) \le y)$$
$$= P(X \le \varphi^{-1}(y))$$
$$= \int_{-\infty}^{\varphi^{-1}(y)} f_X(x)\,dx = \int_{-\infty}^{y} f_X(\varphi^{-1}(z)) \frac{1}{\varphi'(\varphi^{-1}(z))}\,dz$$

Substituiere

$$x = \varphi^{-1}(z), \text{ also } \frac{dx}{dz} = \frac{1}{\varphi'(\varphi^{-1}(z))}$$

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

Dichte $\frac{1}{\sqrt{2\pi}} e^{-x^2/2}$

Beispiel 1.41. $X \sim \mathcal{N}_{0,1}$, $\mu \in \mathbb{R}$, $\sigma > 0$, so ist $Y := \sigma X + \mu \sim \mathcal{N}_{\mu,\sigma^2}$ (Übung).

$$(2\pi\sigma^2)^{-1/2} \exp\left(-(y-\mu)^2/2\sigma^2\right)$$

Bericht 1.42 (Dichtetransformation im $\mathbb{R}^d$). X $\mathbb{R}^d$-wertige ZV mit Dichte f_X, $I \subset \mathbb{R}^d$ offen mit $P(X \in I) = 1$, $J \subset \mathbb{R}^d$ offen, $\varphi : I \to J$ bijektiv, stetig differenzierbar mit Ableitung

$$\varphi'(x) = \left(\frac{\partial \varphi_i}{\partial x_j}(x)\right)_{i,j=1}^{d} \qquad (\text{„Jacobi-Matrix“}),$$

dann hat $Y := \varphi(X)$ die Dichte

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

Beweise finden sich in Analysis-Lehrbüchern, z.B. G. Kersting und M. Brokate, *Maß und Integral*, S. 107, H. Heuser, *Analysis, Teil 2*, Satz 205.2 (“Substitutions-Regel”), O. Forster, *Analysis 3*, Kap. 9, Satz 1 (“Transformationsformel”).

Wir betrachten hier nur folgende Heuristik (im Fall $d = 2$): Lokal sieht

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \varphi(x) = \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix}$$

„aus wie“

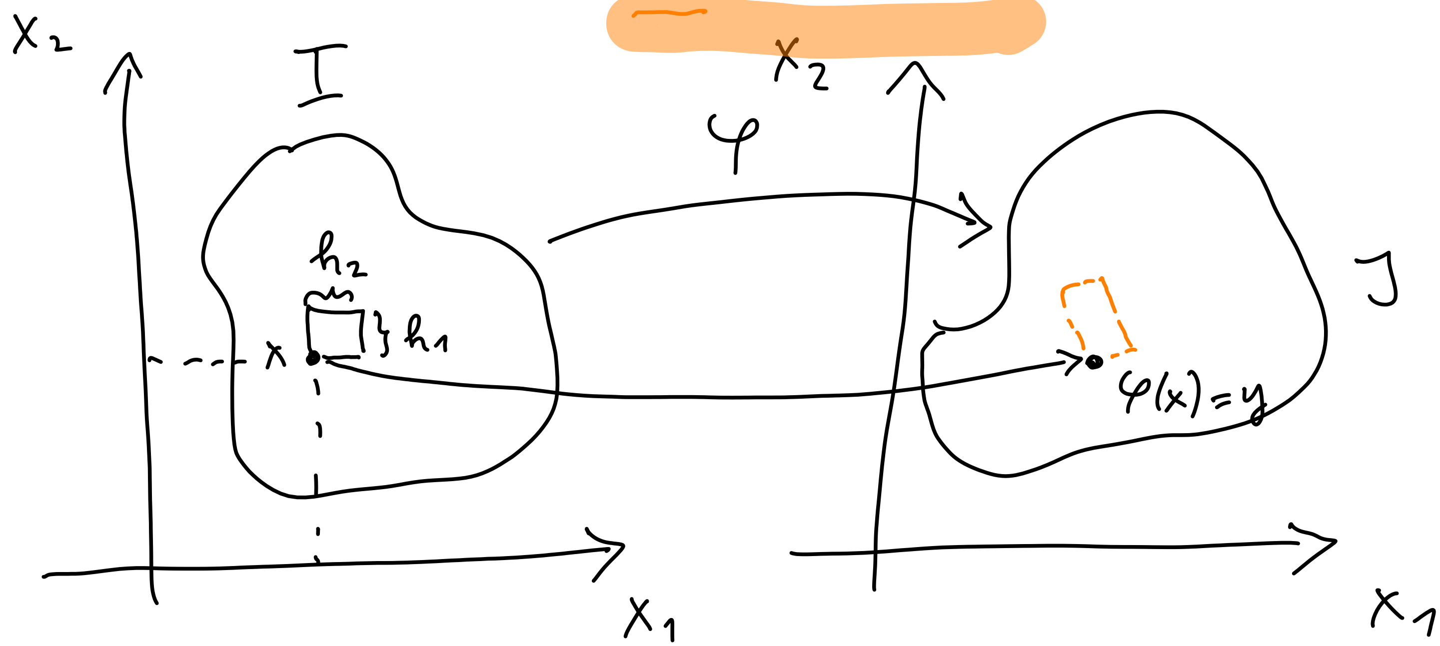
$$\varphi(x') \approx \varphi(x) + \varphi'(x) \cdot (x' - x) = \varphi(x) + \begin{pmatrix} \frac{\partial}{\partial x_1}\varphi_1(x) & \frac{\partial}{\partial x_2}\varphi_1(x) \\ \frac{\partial}{\partial x_1}\varphi_2(x) & \frac{\partial}{\partial x_2}\varphi_2(x) \end{pmatrix} \cdot \begin{pmatrix} x_1' - x_1 \\ x_2' - x_2 \end{pmatrix}$$

(plus Terme, die $O(\|x' - x\|^2)$ sind), also:

$$\text{die Fläche der Größe } h_1 \cdot h_2 \text{ „rund um } x\text{“}$$

wird auf

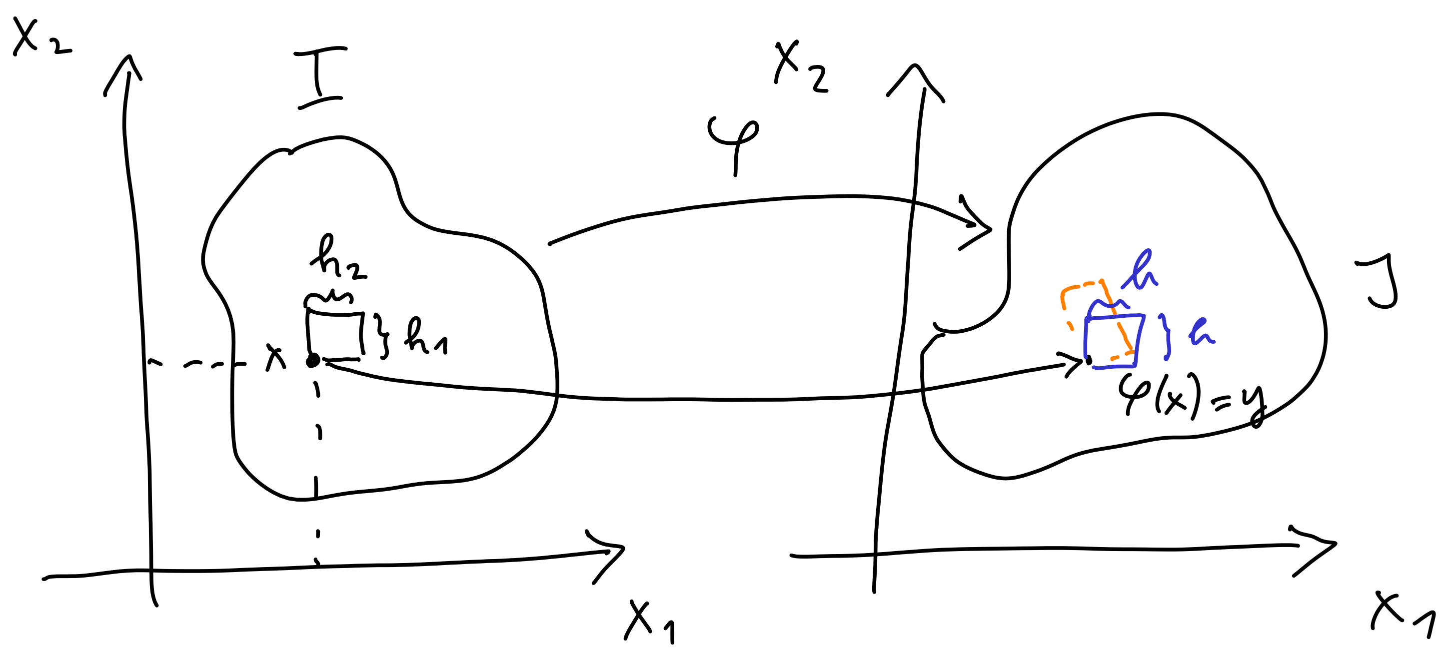
$$\approx \text{Fläche } h_1 \cdot h_2 \cdot |\det \varphi'(x)| \text{ „rund um } y\text{“ abgebildet.}$$



$$\text{Die Fläche der Größe } h_1 \cdot h_2 \text{ „rund um } x\text{“}$$

wird auf

$$\approx \text{Fläche } h_1 \cdot h_2 \cdot |\det \varphi'(x)| \text{ „rund um } y\text{“ abgebildet.}$$



Wenden wir dies auf $Y = \varphi(X)$ an, so bedeutet das anschaulich: Für $y = \varphi(x) \in J$ (und sehr kleines $h > 0$) ist

$$\begin{aligned} f_Y(y)h^2 &\approx \mathbb{P}\big(Y \text{ nimmt Wert in einem Quadrat der Fläche } h^2 \text{ mit „Aufpunkt“ } y \text{ an}\big) \\ &\approx \mathbb{P}\big(X \text{ nimmt Wert in einem Quader der Fläche } h^2/|\det \varphi'(x)| \text{ mit „Aufpunkt“ } x \text{ an}\big) \\ &\approx f_X(x)\frac{h^2}{|\det \varphi'(x)|} = \frac{f_X(\varphi^{-1}(y))}{|\det \varphi'(\varphi^{-1}(y))|}h^2. \end{aligned}$$

Literaturverzeichnis

[KW] G. Kersting und A. Wakolbinger, *Elementare Stochastik*, Birkhäuser 2010

[G] H.-O. Georgii, *Stochastik*, De Gruyter 2015

[H] N. Henze, *Stochastik: Eine Einführung mit Grundzügen der Maßtheorie*, Springer 2019

[Pf] J. Pfanzagl, *Elementare Wahrscheinlichkeitsrechnung*, de Gruyter 1991

[Kr] U. Krengel, Einführung in die Wahrscheinlichkeitstheorie und Statistik, Vieweg 2000

[DH] H. Dehling und B. Haupt, Einführung in die Wahrscheinlichkeitstheorie und Statistik, Springer 2003

[GS] C.M. Grinstead und J.L. Snell, *Introduction to Probability*, AMS 1997.
`http://www.dartmouth.edu/~chance/teaching_aids/books_articles/probability_book/book.html`

[P] J. Pitman, *Probability*, Springer 1997

[F] W. Feller, *An introduction to probability theory and its applications*, Vol I und II, Wiley 1968, 1971

[R] The R Project for Statistical Computing, `http://www.r-project.org/`