

Stochastik für die Naturwissenschaften

Dr. C.J. Luchsinger

4. ZUFALLSGRÖSSE X

4.1 Grundlagen

Drei einleitende Beispiele; Beispiel 4.1.2.A und B und 4.1.4.A im Buch

Definition 4.1 [Zufallsgrösse X (random variable)] *Eine Zufallsgrösse X ist eine Funktion*

$$X : \Omega \rightarrow \mathbb{R}.$$

Zufallsgrössen (Z.G.) nennt man auch Zufallsvariablen (Z.V.).

4.2 Diskrete Zufallsgrössen

1. Würfel: $\Omega = \{1, 2, 3, 4, 5, 6\}$. X soll die Zufallsgrösse sein, welche die Augenzahl angibt. Wir wählen $X(i) = i$ für alle $1 \leq i \leq 6$. Falls Y die Zufallsgrösse sein soll, welche das Quadrat der Augenzahl angibt, so wählen wir $Y(i) = i^2$ für alle $1 \leq i \leq 6$.

2. Münzwurf: $\Omega = \{k, z\}$. Ich gewinne 1.- falls es Kopf (k) gibt und verliere 1.- falls es Zahl (z) gibt. X_1 ist mein Konto nach dem *ersten* Münzwurf: $X_1(k) = 1, X_1(z) = -1$.

3. Tierversuch mit *einer* Maus: $\Omega := \{S, D\}$ ("survives, dies"). $U_1(S) = 1$ falls Maus überlebt, $U_1(D) = 0$ falls Maus stirbt. U_1 ist die Anzahl Mäuse, welche nach Versuch noch leben (0 oder 1).

3'. Mehrere Mäuse: gleiches Experiment mit 5 Mäusen, unabhängige Versuche. U_i ist Anzahl überlebende Mäuse im i 'ten Versuch, also an i 'ter Maus (0 oder 1).

$$W := \sum_{i=1}^5 U_i$$

ist die Anzahl noch lebende Mäuse, sobald wir alle 5 Experimente abgeschlossen haben.

Wir werden jetzt das bisher Erreichte mit dem dritten Kapitel kombinieren, also mit dem "P" verbinden. Angenommen, die Überlebenswahrscheinlichkeit ist 10 %, so können wir

$$P[U_1 = 1] =$$

in Beispiel 3 angeben. Man sagt: "Die Wahrscheinlichkeit, dass U_1 gleich 1 ist 0.1."

Mathematisch sauber ist dies

$$P[U_1 = 1] := P[\{\omega | U_1(\omega) = 1\}] = P[\{S\}] = 0.1,$$

weil P auf Teilmengen von Ω (hier $\{S\}$) operiert. Wir müssen aber nicht immer so in die Niederungen der Definitionen absteigen.

Betrachten wir ein Beispiel, welches ein bisschen komplizierter ist (3'): Mit Überlebenswahrscheinlichkeit 10 % haben wir

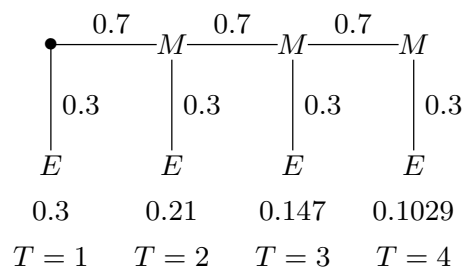
$$P[W = 4] = 0.00045.$$

Wie kommt man auf diese Zahl? Vielleicht hilft es, wenn man die Formel dazu anschaut (wir müssen diese aber ausgiebig besprechen):

$$\binom{5}{4} 0.1^4 0.9^1.$$

Beispiel 4.1.5.A im Buch - keine Binomialverteilung: Wenn ich einen Auskunftsdienst anrufe, komme ich erfahrungsgemäss mit einer Wahrscheinlichkeit von 30% durch (“Your call is important to us”). Ich rufe mehrmals hintereinander an, bis ich Erfolg habe. Spätestens nach fünf Versuchen gebe ich aber auf. Wie ist die Zufallsgrösse $T =$ “Anzahl Anrufe” verteilt?

Es ist klar, dass T die Werte 1 bis 5 annehmen kann. Die zugehörigen Wahrscheinlichkeiten für 1 bis 4 Anrufe entnimmt man am besten einem Baumdiagramm (E bedeutet Erfolg, M Misserfolg), die Wahrscheinlichkeit für $T = 5$ ist dann die Ergänzung auf die Summe 1.



Die Verteilung von T ist also durch die folgende Tabelle gegeben:

t	1	2	3	4	5
$P(T = t)$	0.3	0.21	0.147	0.1029	0.2401

Daraus liest man beispielsweise sofort ab:

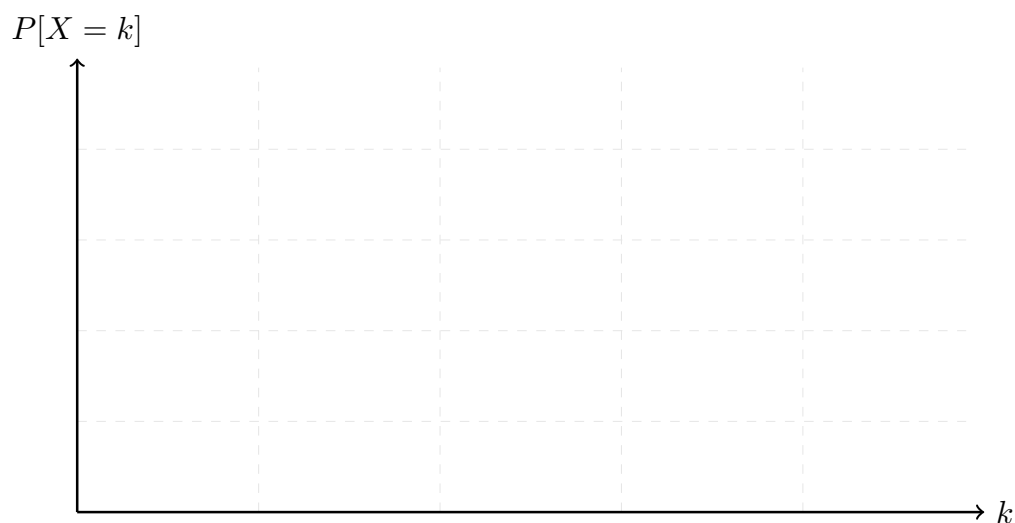
$$P(2 \leq T \leq 4) = 0.21 + 0.147 + 0.1029 = 0.4599 .$$

☒

Definition 4.2 [Diskrete Zufallsgrösse] Wenn eine Zufallsgrösse X nur endlich viele oder abzählbar unendlich viele Werte annehmen kann, nennen wir sie *diskret*.

Die bisherigen Beispiele in 4.2 waren alle diskret und sogar endlich. In Kapitel 2 entspricht das den diskreten Merkmalswerten. Dort ging es um **Daten**, jetzt kommt die **Modellwelt** in Kapiteln 4-7!

Bevor wir uns der Verteilungsfunktion zuwenden, wollen wir vom Publikum wissen, welche Verteilungen bereits bekannt sind. Es kommen im Verlauf der Vorlesung immer wieder neue Verteilungen vor. Merke (frei nach Mark Twain): “Wer nur einen Hammer hat, für den sieht die ganze Welt wie ein Nagel aus.” Deshalb: viele Verteilungen kennenlernen!



Definition 4.3 [Wahrscheinlichkeitsfunktion $P[X = x]$ (probability function) und Verteilungsfunktion F (Cumulative Distribution Function CDF)] Die Wahrscheinlichkeitsfunktion einer diskreten Zufallsgrösse X ist definiert als

$$(f(x)) := P[X = x](:= P[\{\omega | X(\omega) = x\}]).$$

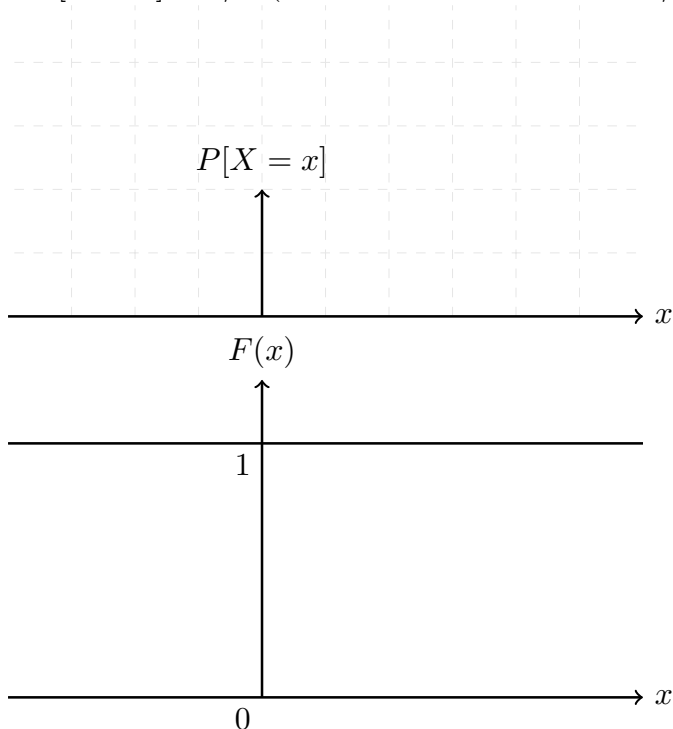
Dabei durchläuft x alle möglichen Werte der Zufallsgrösse X . X ist die Z.G, x ein konkreter Wert, den Sie zuerst als fest denken. Die Verteilungsfunktion F einer Zufallsgrösse X (diskret oder stetig) ist definiert als

$$F(x) := P[X \leq x](:= P[\{\omega | X(\omega) \leq x\}]).$$

Wenn wir uns an die beschreibende Statistik in Kapitel 2 erinnern, dann entspricht das Stabdiagramm der Wahrscheinlichkeitsfunktion, die kumulierte Summenhäufigkeit der Verteilungsfunktion. Es gilt zudem offenbar (letztes Gleichheitszeichen nur bei diskret)

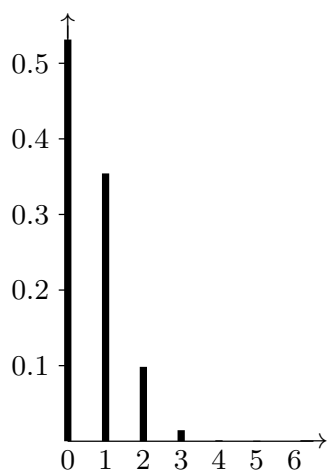
$$F(x) = P[X \leq x] = \sum_{x_i \leq x} P[X = x_i].$$

Beispiel von letzter Seite: diskrete Gleichverteilung: wenn X genau n Werte annehmen kann: $P[X = x] = 1/n$ (zB Würfel mit $n = 6$ und $1/6$).

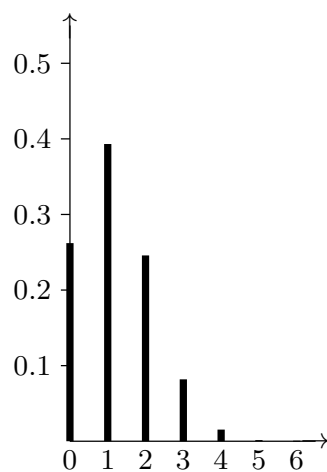


Welche Eigenschaften besitzen offenbar Verteilungsfunktionen?

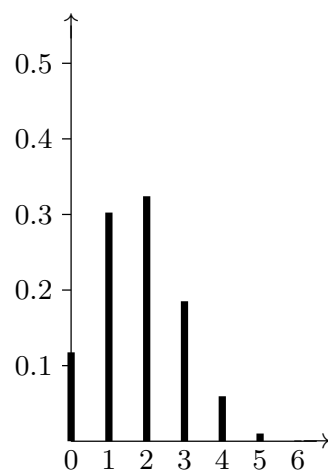
Wir betrachten die Stabdiagramme der Binomialverteilung für die Fälle $n = 6, p = 0.1, 0.2, 0.3, 0.4, 0.5$ und 0.6 :



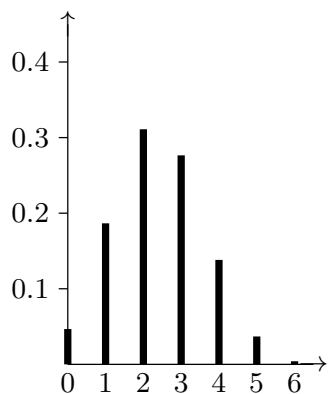
$n = 6, p = 0.1$



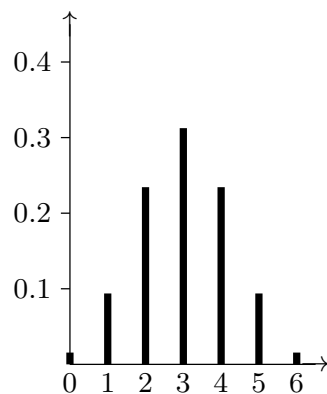
$n = 6, p = 0.2$



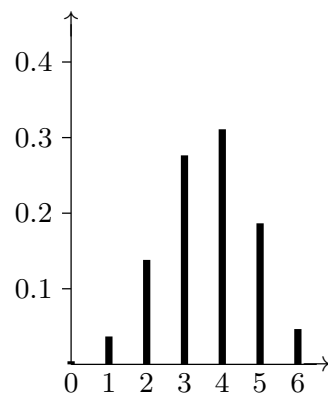
$n = 6, p = 0.3$



$n = 6, p = 0.4$



$n = 6, p = 0.5$



$n = 6, p = 0.6$

Aufgaben zur Tabelle vor einer Seite - diese Tabelle finden Sie nochmals in Kapitel 6. Ziel ist jetzt: Tabellen lesen lernen. Ab jetzt heisst " $X \sim \text{Bin}(n, p)$ ", dass X $\text{Bin}(n, p)$ -verteilt ist; das Zeichen \sim nennt man "Tilde".

$$X \sim \text{Bin}(6, 0.2): P[X \in [2, 4]] =$$

$$X \sim \text{Bin}(8, 0.2): P[X \in [2, 4]] =$$

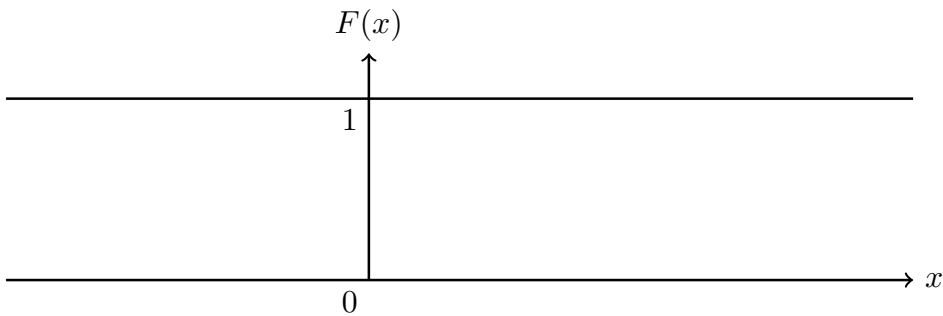
$$X \sim \text{Bin}(8, 0.8): P[X \in [3, 4]] =$$

Kompakte Schreibweise: anstatt $X \sim \text{Bin}(n, p)$ und dann $P[X \in [3, 4]]$ schreibt man kurz $P[\text{Bin}(n, p) \in [3, 4]]$.

Aufgabe 4–2: Eine diskrete Zufallsgrösse X ist durch die folgende Tabelle gegeben

x_i	-3	-1	0	2	4	6
p_i	0.2	0.1	0.1	0.15	0.25	a

a) Wie gross ist a ? b) Zeichnen Sie den Graphen der Verteilungsfunktion von X . Berechnen Sie ferner c) $P(0 < X \leq 4)$, d) $P(\pi \leq X \leq 2\pi)$, e) $P(X^2 < 10)$.



Aufgabe 4–3: Eine diskrete Zufallsgrösse X ist durch die folgende Tabelle gegeben:

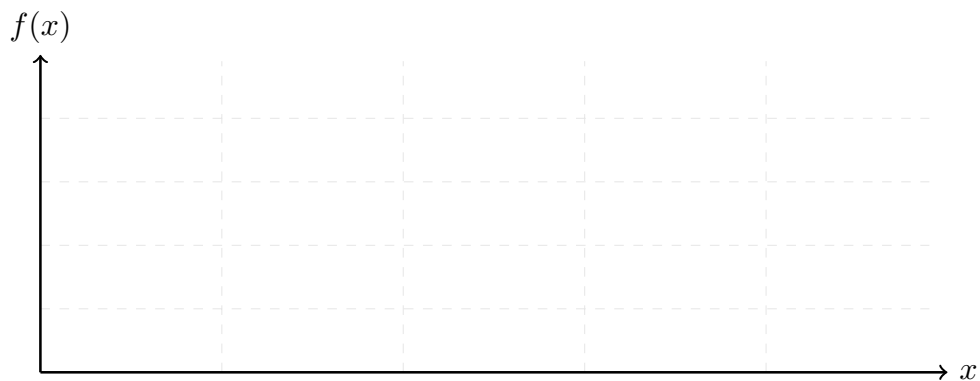
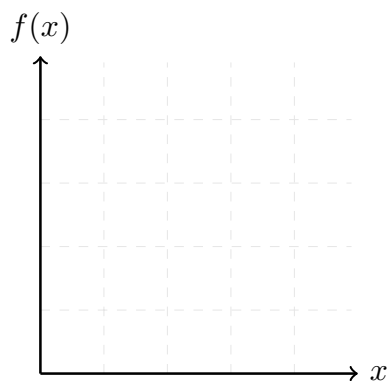
x_i	-2	-1	0	1	4
p_i	0.1	0.2	0.3	0.1	0.3

a) Stellen Sie die entsprechende Tabelle für $Z = X^2 - 1$ auf.
 b) Berechnen Sie $P(3 \leq Z \leq 8)$.

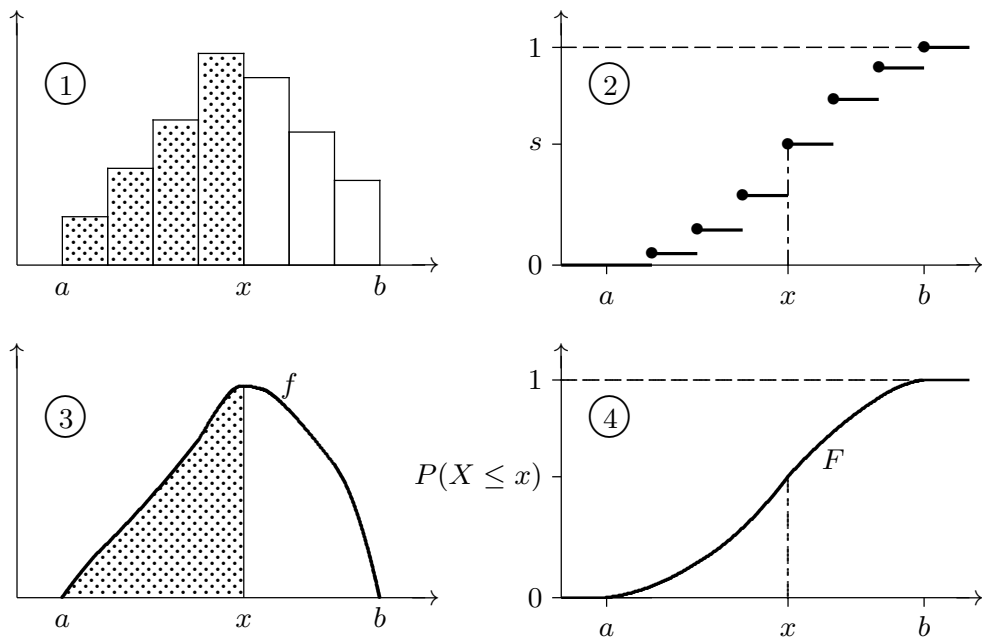
4.3 Stetige Zufallsgrößen

Die kommenden Seiten sind zentral, aber schwer zu verstehen. Eine gute Strategie für die Studierenden ist es, die Resultate vorerst einfach zu akzeptieren, die Beispiele sorgfältig anzuschauen und Übungen dazu gut zu lösen.

Motivation anhand Hund (uniform) und dann Normalverteilung.



Motivation anhand Buch (4.3.2)



Definition 4.4 [Diskrete und stetige Zufallsgrösse] Wenn eine Zufallsgrösse X nur Werte auf einer abzählbaren Teilmenge von \mathbb{R} annehmen kann, nennen wir sie diskret. Wir nennen eine Zufallsgrösse Y stetig, wenn ihre Verteilungsfunktion F_Y sich darstellen lässt als

$$F_Y(a) := P[Y \leq a] = P[Y \in (-\infty, a]] = \int_{-\infty}^a f(u) du,$$

wobei f eine nichtnegative (das heisst $f(x) \geq 0$ für alle $x \in \mathbb{R}$), stückweise stetige Funktion auf \mathbb{R} ist (indirekte Definition, seltsam). Diese Funktion f erfüllt

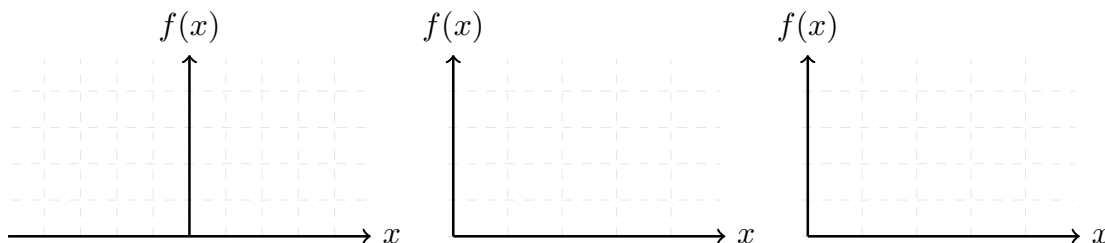
$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

Wir nennen f Dichte oder Dichtefunktion der Zufallsgrösse Y . Der Index Y in F_Y unterstreicht, dass dieses F zu Y gehört. Salopp gilt: diskret ist auf einzelnen Punkten und stetig ist auf durchgezogenen Intervallen.

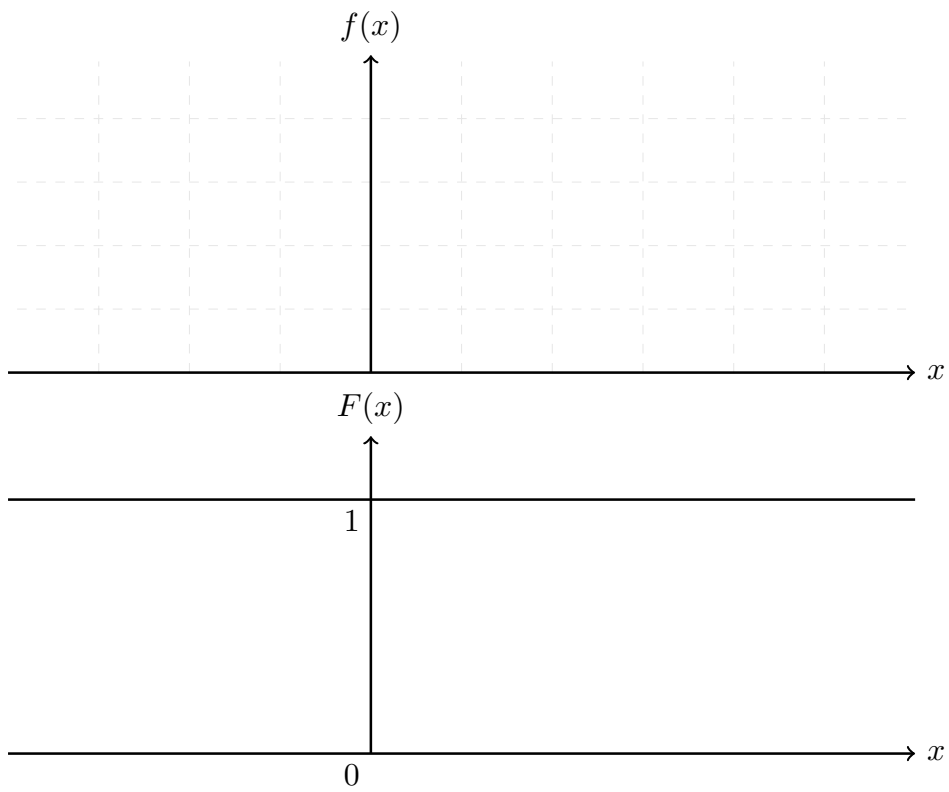
Die Dichtefunktion $f(x)$ der Normalverteilung ist übrigens (ausführlich in Buch und Skript Kapitel 5 und 6):

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

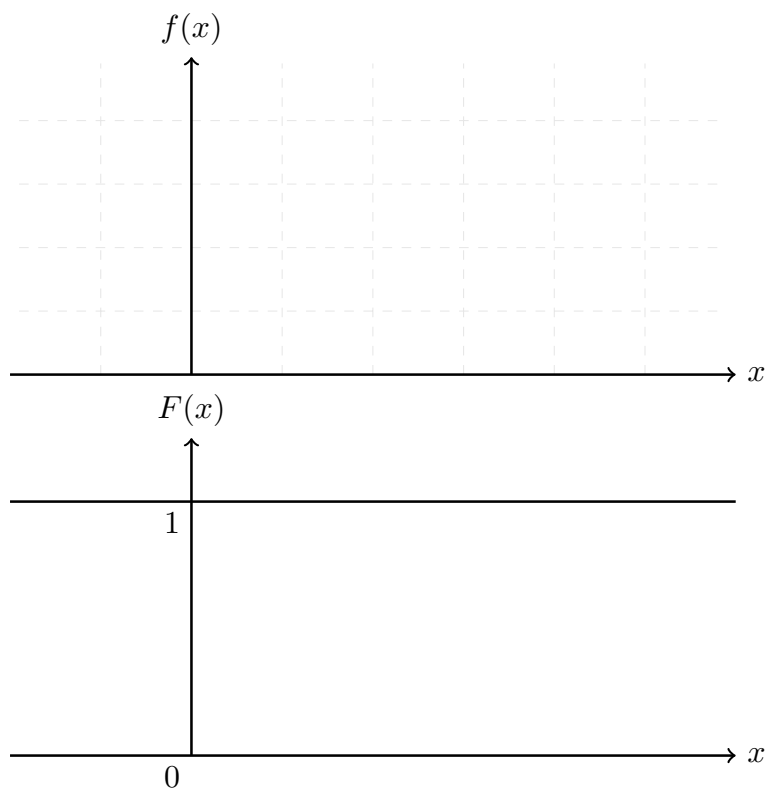
Wir werden später sehen, dass μ der sogenannte Erwartungswert ist (ein Lagemass) und σ^2 die sogenannte Varianz (ein Streuungsmass). Bilder dazu:



Skizzieren Sie: Verteilungsfunktion der Uniform-Zufallsgröße $U[-1, 0.5]$ (auch Rechteck- oder Gleichverteilung, "gleiche Verteilung" ist was anderes...)



Skizzieren Sie: Dichte und Verteilungsfunktion der Normalverteilung (Gauss)-Zufallsgrösse $\mathcal{N}(2, 4)$



Wir repetieren kurz den Hauptsatz der Differential- und Integralrechnung (MAT182, Kapitel 11):

Wir wollen im Folgenden 4 Eigenschaften von **stetigen** Zufallsgrössen herausarbeiten und später am Beispiel der Normalverteilung veranschaulichen.

1. Für $a < b$ haben wir:

$$\begin{aligned} P[a < X \leq b] &= P[X \leq b] - P[X \leq a] \\ &= F(b) - F(a) \\ &= \int_a^b f(x)dx. \end{aligned}$$

2. Für stetige Zufallsgrössen gilt:

$$P[X = x_0] = \int_{x_0}^{x_0} f(x)dx = 0.$$

3. Wegen 2. gilt bei stetig insbesondere für $a < b$:

$$\begin{aligned} P[a \leq X \leq b] &= P[a < X \leq b] \\ &= P[a < X < b] \\ &= P[a \leq X < b] \\ &= P[X \leq b] - P[X \leq a] \\ &= F(b) - F(a) = \int_a^b f(x)dx. \end{aligned}$$

4. Wenn F differenzierbar ist (mit gesundem Menschenverstand auch in den meisten anderen Fällen) gilt:

$$F'(x) = f(x).$$

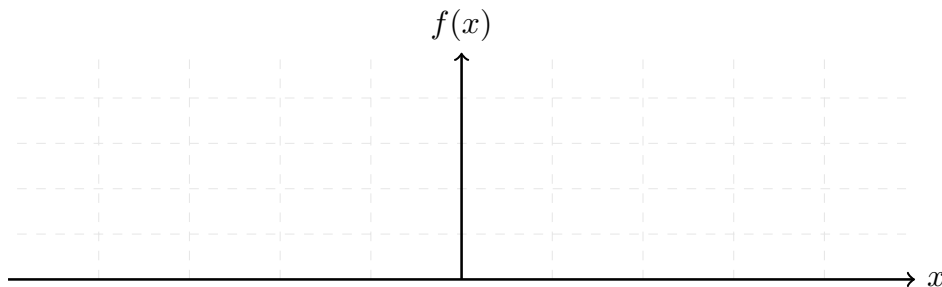
Weshalb betrachtet man die **Verteilungsfunktion** F einer Zufallsgrösse? 4 Gründe:

1. Wir können Dank “Fläche unter Kurve ist Wahrscheinlichkeit” W’keiten schneller ablesen.
2. Wir können im Fall von stetigen Zufallsgrössen mit Hilfe von $F' = f$ einfach die Dichte berechnen.
3. In Kapitel 7 werden wir mit Hilfe der Verteilungsfunktion den CLT (Central Limit Theorem; Zentraler Grenzwertsatz) formulieren können.
4. In der Statistik kann man zum Beispiel mit sogenannten QQ-Plots (Quantil vs Quantil-Plots) Verteilungen miteinander vergleichen **noch anpassen, eigenes Buch/Skript (mehr dazu in Stahel 11.2 - Prüfungstoff)**.

Lose Bemerkung: Falls es in einer Aufgabe oder Frage heisst ”Wie ist die Verteilung von X ?”, so ist diese Frage ungenau. Antworten können dann sein:

1. Die Dichte ist...
2. Die Verteilungsfunktion sieht folgendermassen aus...
3. Die Verteilung ist $\exp(\lambda)$.

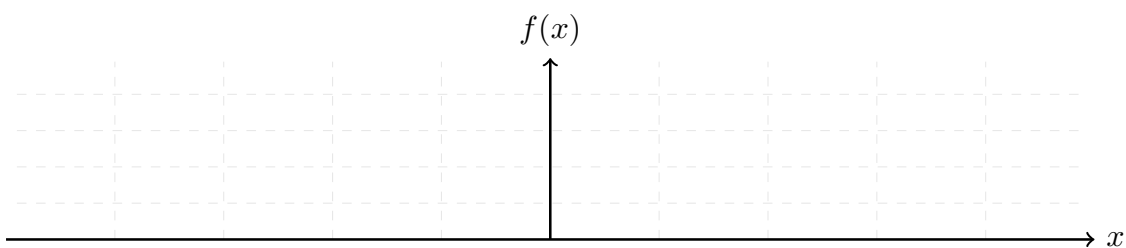
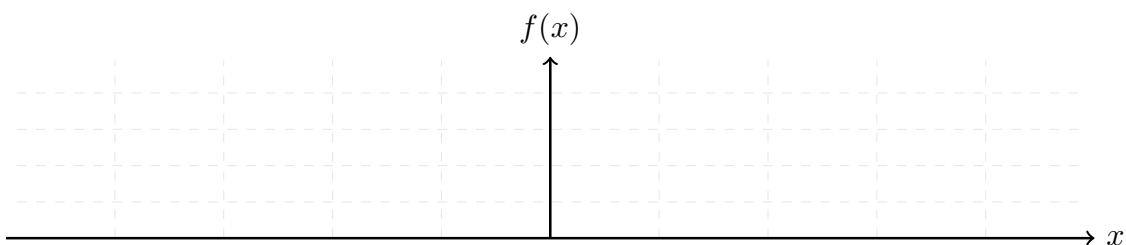
Aufgaben zur Tabelle der Standardnormalverteilung - diese Tabelle finden Sie nochmals in Kapitel 6.



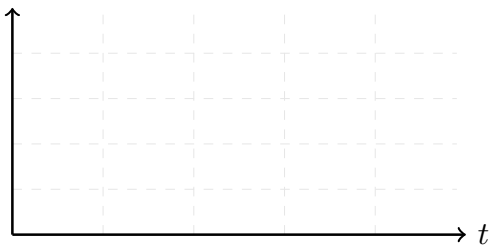
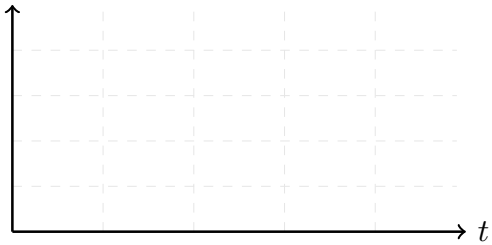
Werte der Verteilungsfunktion der Standardnormalverteilung $\mathcal{N}(0, 1)$

x	$F(x)$	x	$F(x)$	x	$F(x)$	x	$F(x)$
-4.00	0.0000	-1.10	0.1357	0.02	0.5080	1.12	0.8686
-3.90	0.0000	-1.08	0.1401	0.04	0.5160	1.14	0.8729
-3.80	0.0001	-1.06	0.1446	0.06	0.5239	1.16	0.8770
-3.70	0.0001	-1.04	0.1492	0.08	0.5319	1.18	0.8810
-3.60	0.0002	-1.02	0.1539	0.10	0.5398	1.20	0.8849
-3.50	0.0002	-1.00	0.1587	0.12	0.5478	1.22	0.8888
-3.40	0.0003	-0.98	0.1635	0.14	0.5557	1.24	0.8925
-3.30	0.0005	-0.96	0.1685	0.16	0.5636	1.26	0.8962
-3.20	0.0007	-0.94	0.1736	0.18	0.5714	1.28	0.8997
-3.10	0.0010	-0.92	0.1788	0.20	0.5793	1.30	0.9032
-3.00	0.0013	-0.90	0.1841	0.22	0.5871	1.32	0.9066
-2.90	0.0019	-0.88	0.1894	0.24	0.5948	1.34	0.9099
-2.80	0.0026	-0.86	0.1949	0.26	0.6026	1.36	0.9131
-2.70	0.0035	-0.84	0.2005	0.28	0.6103	1.38	0.9162
-2.60	0.0047	-0.82	0.2061	0.30	0.6179	1.40	0.9192
-2.50	0.0062	-0.80	0.2119	0.32	0.6255	1.42	0.9222
-2.45	0.0071	-0.78	0.2177	0.34	0.6331	1.44	0.9251
-2.40	0.0082	-0.76	0.2236	0.36	0.6406	1.46	0.9279
-2.35	0.0094	-0.74	0.2296	0.38	0.6480	1.48	0.9306
-2.30	0.0107	-0.72	0.2358	0.40	0.6554	1.50	0.9332
-2.25	0.0122	-0.70	0.2420	0.42	0.6628	1.55	0.9394
-2.20	0.0139	-0.68	0.2483	0.44	0.6700	1.60	0.9452
-2.15	0.0158	-0.66	0.2546	0.46	0.6772	1.65	0.9505
-2.10	0.0179	-0.64	0.2611	0.48	0.6844	1.70	0.9554
-2.05	0.0202	-0.62	0.2676	0.50	0.6915	1.75	0.9599
-2.00	0.0228	-0.60	0.2743	0.52	0.6985	1.80	0.9641
-1.95	0.0256	-0.58	0.2810	0.54	0.7054	1.85	0.9678
-1.90	0.0287	-0.56	0.2877	0.56	0.7123	1.90	0.9713
-1.85	0.0322	-0.54	0.2946	0.58	0.7190	1.95	0.9744
-1.80	0.0359	-0.52	0.3015	0.60	0.7257	2.00	0.9772
-1.75	0.0401	-0.50	0.3085	0.62	0.7324	2.05	0.9798
-1.70	0.0446	-0.48	0.3156	0.64	0.7389	2.10	0.9821
-1.65	0.0495	-0.46	0.3228	0.66	0.7454	2.15	0.9842
-1.60	0.0548	-0.44	0.3300	0.68	0.7517	2.20	0.9861
-1.55	0.0606	-0.42	0.3372	0.70	0.7580	2.25	0.9878
-1.50	0.0668	-0.40	0.3446	0.72	0.7642	2.30	0.9893
-1.48	0.0694	-0.38	0.3520	0.74	0.7704	2.35	0.9906
-1.46	0.0721	-0.36	0.3594	0.76	0.7764	2.40	0.9918
-1.44	0.0749	-0.34	0.3669	0.78	0.7823	2.45	0.9929
-1.42	0.0778	-0.32	0.3745	0.80	0.7881	2.50	0.9938
-1.40	0.0808	-0.30	0.3821	0.82	0.7939	2.60	0.9953
-1.38	0.0838	-0.28	0.3897	0.84	0.7995	2.70	0.9965
-1.36	0.0869	-0.26	0.3974	0.86	0.8051	2.80	0.9974
-1.34	0.0901	-0.24	0.4052	0.88	0.8106	2.90	0.9981
-1.32	0.0934	-0.22	0.4129	0.90	0.8159	3.00	0.9987
-1.30	0.0968	-0.20	0.4207	0.92	0.8212	3.10	0.9990
-1.28	0.1003	-0.18	0.4286	0.94	0.8264	3.20	0.9993
-1.26	0.1038	-0.16	0.4364	0.96	0.8315	3.30	0.9995
-1.24	0.1075	-0.14	0.4443	0.98	0.8365	3.40	0.9997
-1.22	0.1112	-0.12	0.4522	1.00	0.8413	3.50	0.9998
-1.20	0.1151	-0.10	0.4602	1.02	0.8461	3.60	0.9998
-1.18	0.1190	-0.08	0.4681	1.04	0.8508	3.70	0.9999
-1.16	0.1230	-0.06	0.4761	1.06	0.8554	3.80	0.9999
-1.14	0.1271	-0.04	0.4840	1.08	0.8599	3.90	1.0000
-1.12	0.1314	-0.02	0.4920	1.10	0.8643	4.00	1.0000
		0.00	0.5000				

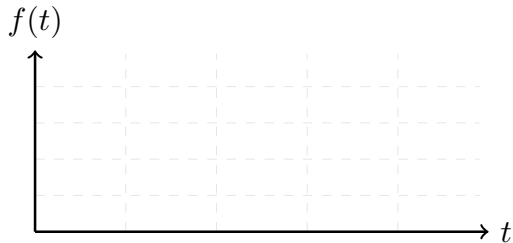
Aufgaben zur Tabelle der Standardnormalverteilung $\mathcal{N}(0, 1)$:



Die Exponentialverteilung I (Biochemie, Radioaktivität, Wartezeiten)



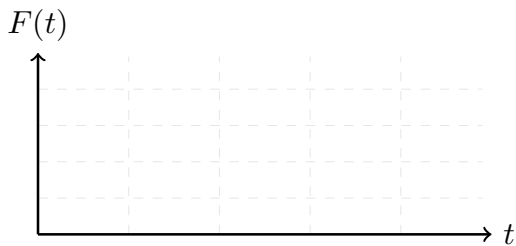
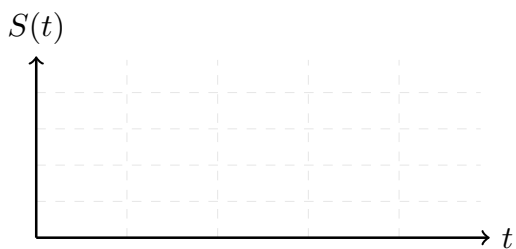
Die Exponentialverteilung II: X sei Z.G., welche die Zeit misst, bis ein gegebenes Isotop zerfällt. Von der Physik wissen wir einerseits, dass diese Zeit gedächtnislos ist. Die Mathematik sagt uns andererseits, dass wir damit eine $\exp(\lambda)$ -Z.G. benutzen *müssen* (vgl Kapitel 6). Untersuchen wir die Exponentialverteilung $\exp(\lambda)$. Wir untersuchen dazu Dichte-, Verteilungs- und Survivalfunktion; zur **Dichtefunktion**:



diese muss 3 Bedingungen erfüllen: 1. nichtnegativ:

2. stückweise stetig:

3. $\int_{-\infty}^{\infty} f(t)dt = 1$. Übungen/Prüfung: Oft sinnvoll, zuerst $F(a)$ berechnen, weil man $F(a)$ später braucht und dann schauen, ob $\lim_{a \rightarrow \infty} F(a) = \int_{-\infty}^{\infty} f(t)dt = 1$: Mit $a > 0$:
 $F(a) = \int_{-\infty}^a f(t)dt =$

Verteilungsfunktion:**Survivalfunktion:**

Man beachte: $f(t) = \lambda \cdot S(t)$: Bis auf die Konstante λ ist der Verlauf, das Bild, also gleich.

Und wie sieht die Survivalfunktion bei uns Menschen aus? Dazu drei Links:

* Die Dichte:

https://www.demogr.mpg.de/de/news_events_6123/news_pressemitteilungen_4630/presse/sterbealter_gleichen_sich_an_4883/

* Survivalfunktion seit 16. Jh:

<https://www.kommunikation.uzh.ch/static/unimagazin/archiv/2-98/aelter.html>

* Survivalfunktion seit 1900:

https://www.researchgate.net/figure/Abbildung-4-Ueberlebenskurven-von-Frauen-verschieden-fig2_280810795

Man beachte: Kindersterblichkeit hat massiv abgenommen, Frauen sterben nicht mehr bei der Niederkunft, offenbar gibt es eine Grenze um die 100 Jahre, Verteilung wird enger.

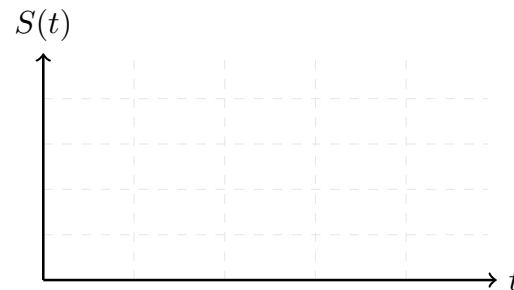
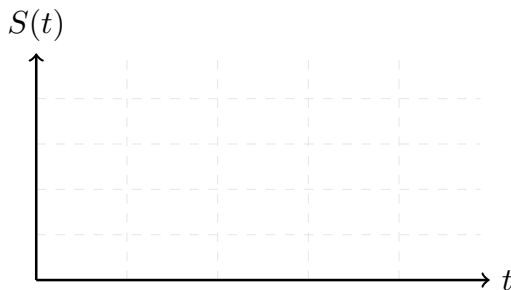
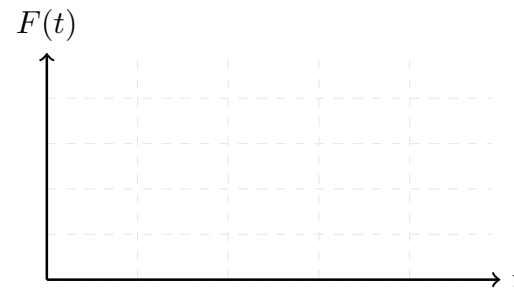
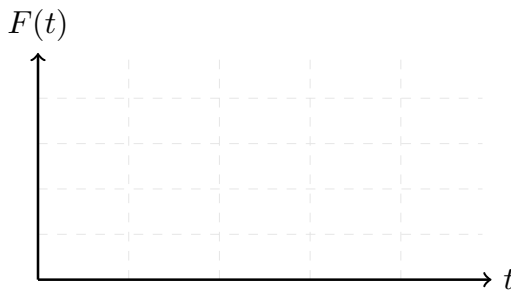
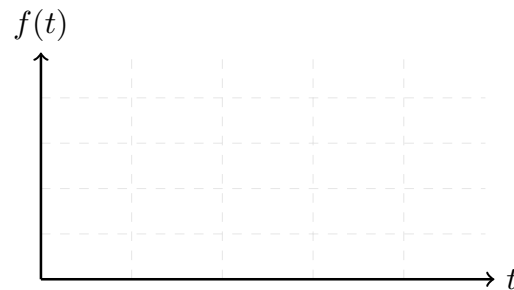
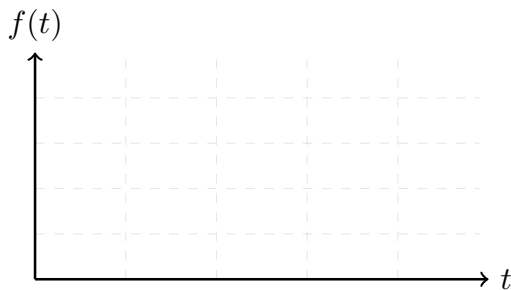
Die Exponentialverteilung III: Das Nachfolgende ist nicht zu verwechseln mit HAA (**H**och **A**ktive **A**bfälle) - diese sind sicher langlebig, aber nicht alles was langlebig ist, ist auch HAA (zB U-238). In welchem Gestein in den Schweizer Alpen kommt U-238 auch prominent vor?

Für Nachfolgendes: Das Schwein heisst Schwein weil es stinkt wie ein Schwein!

kurzlebig

vs

langlebig

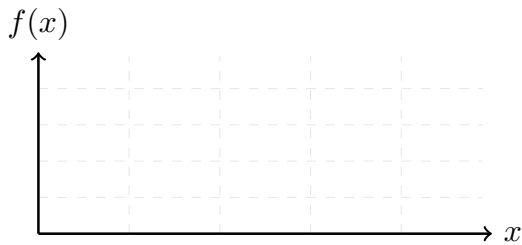


Das λ nennt man Zerfallsrate. Die Namenswahl ist intuitiv sinnvoll: eine Rate, eine Kadenz ist etwas in der Art “10 pro Sekunde” (Schraubenproduktion). Dann dauert es $1/10$ Sekunde, bis die nächste Schraube kommt. Deshalb wird $1/\lambda$ in Kapitel 5 der Erwartungswert einer Exponentialverteilung sein (man kann hier auch von Übergangszeit oder Wartezeit sprechen). Hohe Kadenz heisst, es dauert nicht lange bis zum nächsten Ereignis, deshalb der Kehrwert.

Aufgabe 4–23 a) Bestimmen Sie die Zahl c so, dass

$$f(x) = \begin{cases} 0 & x < 0 \\ c(1+x)^{-3} & x \geq 0 \end{cases}$$

die Dichtefunktion einer Zufallsvariablen X wird. b) Skizzieren Sie den Graphen von f , und stellen Sie die Wahrscheinlichkeit $p = P(0.5 \leq X \leq 1)$ graphisch dar. c) Berechnen Sie p .

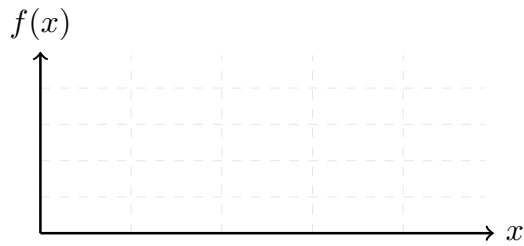


[Hier ist es besser, $F(a)$ *nicht* auf Vorrat zu berechnen]

Aufgabe 4–24 a) Wie muss man die Zahl a wählen, damit

$$f(x) = \begin{cases} 0 & x < 1 \\ ax^{-2} & x \geq 1 \end{cases}$$

eine Dichtefunktion wird? Die zugehörige stetige Zufallsgrösse nennen wir X . b) Skizzieren Sie den Graphen von f , und stellen Sie die Wahrscheinlichkeit $p = P(1.5 \leq X \leq 2)$ graphisch dar. c) Bestimmen Sie die Verteilungsfunktion von X . d) Berechnen Sie p .



[Hier ist es besser, $F(a)$ auf Vorrat zu berechnen]

4.4 Unabhängigkeit von Zufallsgrössen

Anschaulich bedeutet " X unabhängig von Y ", dass X und Y von unabhängigen Mechanismen erzeugt werden. Mathematisch führt man die Unabhängigkeit von Zufallsgrössen auf die Unabhängigkeit von Ereignissen zurück, das heisst: $P[A \cap B] = P[A]P[B]$ (Idee der Faktorisierung). In der folgenden Definition ist ein Komma als "und" zu lesen und damit als Schnittmenge: $\{X_1 \leq x_1, X_2 \leq x_2\}$ heisst $\{X_1 \leq x_1\} \cap \{X_2 \leq x_2\}$.

Definition 4.5 [Unabhängigkeit von Zufallsgrössen; Idee der Faktorisierung]

Zufallsgrössen X_1, X_2, \dots, X_n nennen wir unabhängig voneinander, wenn

$$\begin{aligned} P[X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n] \\ = P[X_1 \leq x_1]P[X_2 \leq x_2] \dots P[X_n \leq x_n] \end{aligned}$$

für beliebige x_1, \dots, x_n .

Jargon: unabhängig, identisch verteilt = independent and identically distributed (**iid**)

Warnung: " X_1 und X_2 sind identisch verteilt " heisst nicht, dass $X_1(\omega) = X_2(\omega)$ für alle $\omega \in \Omega$ (gleicher Wert) oder dass die X_i uniform verteilt sind, sondern lediglich, dass die Verteilung gleich ist; zum Beispiel beide $\text{Be}(p)$ oder beide $\text{Exp}(\lambda)$.

Wichtig:

1. Lesen Sie jetzt das komplette Kapitel im Buch selber durch.
2. Lösen Sie danach mindestens 5 Aufgaben hinten im Kapitel und vergleichen Sie mit den Lösungen am Schluss des Buches. Bei Bedarf lösen Sie mehr Aufgaben.
3. Gehen Sie in die Übungsstunde. Drucken Sie das Übungsblatt dazu *vorher* aus, lesen Sie *vorher* die Aufgaben durch und machen sich erste Gedanken dazu (zum Beispiel, wie man sie lösen könnte).

4. Dann lösen Sie das Übungsblatt: zuerst immer selber probieren, falls nicht geht: Tipp von Mitstudi benutzen, falls immer noch nicht geht: Lösung von Mitstudi anschauen, 1 Stunde warten, versuchen, aus dem Kopf heraus wieder zu lösen, falls immer noch nicht geht: Lösung von Mitstudi abschreiben (und verstehen - also sollte man insbesondere keine Fehler abschreiben!).
5. Lösen Sie die entsprechenden Prüfungsaufgaben im Archiv.

Lessons learnt:

In 3Blue1Brown

<https://www.3blue1brown.com/lessons/gaussian-integral>

<https://www.3blue1brown.com/lessons/binomial-distributions>

<https://www.3blue1brown.com/lessons/pdfs>

Binomialverteilung verstehen, sonst auswendig: $P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$ wo unabhängige Ereignisse, zählen # Erfolge, Erfolgsw'keit p

Z.G. diskret falls nur endlich viele oder abzählbar unendlich viele Werte annehmen kann (vgl beschreibende Statistik; stetig ist aber komplizierter, siehe weiter unten!)

Wahrscheinlichkeitsfunktion diskret: $P[X = x]$; Verteilungsfunktion diskret oder stetig $F(x) := P[X \leq x]$. Von Kapitel 2: Stabdiagramm entspricht der Wahrscheinlichkeitsfunktion, kummulierte Summenhäufigkeit der Verteilungsfunktion. 2 Fehlerquellen bei Verteilungsfunktion: auf ganz \mathbb{R} definiert und "Böppeli oben ausgemalt". Verteilungsfunktionen nehmen Werte an in $[0, 1]$, sind monoton wachsend und rechtsstetig ("Böppeli oben ausgemalt").

Vorsicht bei Beträgen und Quadraten in $P[X^2 < 10]$, $P[|X| < 10]$ wegen negativen Beiträgen zur Gesamtwahrscheinlichkeit.

Vorsicht Transformationen (bei diskreten ZG): $Z := X^2 - 1$: die p_i nicht mittransformieren (falsch: $p_i^2 - 1$), aber schauen, ob neue Reihenfolge und ob zwei transformierte Werte neu zusammenfallen (Aufgabe 4-3).

bei stetig: “Fläche unter Kurve ist Wahrscheinlichkeit.”

Y stetig, wenn Verteilungsfunktion F_Y sich darstellen lässt als $F_Y(a) := P[Y \leq a] = P[Y \in (-\infty, a]] = \int_{-\infty}^a f(u)du$, wo f (3 Bedingungen): nichtnegativ, stückweise stetig, $\int_{-\infty}^{\infty} f(u)du = 1$. f heisst Dichtefunktion. Salopp gilt wie in Kapitel 2: diskret auf einzelnen Punkten, stetig auf durchgezogenen Intervallen.

stetig: $P[a < X \leq b] = P[X \leq b] - P[X \leq a] = F(b) - F(a) = \int_a^b f(x)dx$

stetig: $P[X = x_0] = \int_{x_0}^{x_0} f(x)dx = 0$ und damit $P[a \leq X \leq b] = P[a < X \leq b] = P[a < X < b] = P[a \leq X < b] = P[X \leq b] - P[X \leq a] = F(b) - F(a) = \int_a^b f(x)dx$.

F differenzierbar: $F'(x) = f(x)$.

Unabh Z.G. Definition “Faktorisierung”:

$$P[X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n] = P[X_1 \leq x_1]P[X_2 \leq x_2] \dots P[X_n \leq x_n].$$

Klickerfragen zum Aufwärmen:

Frage 1: Welche der folgenden Aussagen über eine Zufallsgrösse X ist korrekt?

- (a) Eine Zufallsgrösse ist immer eine diskrete Variable.
- (b) Eine Zufallsgrösse ist eine feste Konstante.
- (c) Eine Zufallsgrösse ist eine Funktion, die jedem Element des Ergebnisraums eine reelle Zahl zuordnet.

Frage 2: Welche der folgenden Zufallsgrössen ist diskret?

- (a) Anzahl der geworfenen Sechsen in 10 Würfeln
- (b) Temperatur eines zufällig gewählten Tages
- (c) Grösse eines Menschen in cm

Frage 3: Gegeben ist eine Zufallsgrösse mit den Werten $X = 1, 2, 3$ und den Wahrscheinlichkeiten $P[X = 1] = 0.2$, $P[X = 2] = 0.5$, $P[X = 3] = 0.3$. Wie gross ist $P(X \leq 2)$?

- (a) 0.7
- (b) 0.5
- (c) 0.3

Frage 4: Gegeben ist eine Zufallsgrösse mit den Werten $X = 1, 2, 3$ und den Wahrscheinlichkeiten $P[X = 1] = 0.2$, $P[X = 2] = 0.5$, $P[X = 3] = 0.3$. Wie gross ist $P(X \neq 2)$?

- (a) 0.7
- (b) 0.5
- (c) 0.3

Frage 5: Welche der folgenden Aussagen über stetige Zufallsgrössen ist falsch?

- (a) Die Wahrscheinlichkeitsfunktion einer stetigen Zufallsgrösse gibt Wahrscheinlichkeiten direkt an.
- (b) Eine stetige Zufallsgrösse kann unendlich viele Werte in einem Intervall annehmen.
- (c) Die Wahrscheinlichkeit $P(X = x)$ für eine bestimmte Zahl x ist immer null.

Frage 6: Welche der folgenden Aussagen über die Verteilungsfunktion $F(x)$ ist falsch?

- (a) $F(x)$ gibt die Wahrscheinlichkeit an, dass $X \leq x$ ist.
- (b) $F(x)$ ist immer monoton wachsend.
- (c) $F(x)$ kann auch abnehmen.

Frage 7: Die Exponentialverteilung wird oft verwendet zur Modellierung von:

- (a) Anzahl der Sechsen in 10 Würfeln
- (b) Lebensdauern von Geräten
- (c) Körpergrössen von Personen

Frage 8: Welche der folgenden Eigenschaften hat die Dichtefunktion einer stetigen Zufallsgrösse?

- (a) Sie ist immer grösser als 1.
- (b) Die Fläche unter der Kurve ist 1.
- (c) Die Dichtefunktion kann negativ sein.

Frage 9: Ein diskreter Wahrscheinlichkeitsraum muss immer die Eigenschaft erfüllen:

- (a) Die Summe aller Wahrscheinlichkeiten ist 0.
- (b) Die Summe aller Wahrscheinlichkeiten ist 1.
- (c) Mindestens eine Wahrscheinlichkeit muss grösser als 1 sein.

Frage 10: Wie gross ist die Wahrscheinlichkeit, dass eine standardnormalverteilte Zufallsgrösse X zwischen -1 und 1 liegt?

- (a) Etwa 0.5
- (b) Etwa 0.68
- (c) Etwa 0.95

Frage 11: Ein fairer Münzwurf wird 3-mal durchgeführt.

Wie gross ist die Wahrscheinlichkeit, genau 1-mal Kopf zu werfen?

- (a) $\frac{1}{2}$
- (b) $\frac{1}{8}$
- (c) $\frac{3}{8}$

$$P(X = 1) = \binom{3}{1} \times \left(\frac{1}{2}\right)^1 \times \left(\frac{1}{2}\right)^2 = \frac{3}{8}$$

Frage 12: Welche der folgenden Aussagen zur Normalverteilung ist richtig?

- (a) Sie ist immer symmetrisch um ihren Mittelwert.
- (b) Sie ist nur für ganze Zahlen definiert.
- (c) Ihre Dichtefunktion kann auch negative Werte annehmen.

Frage 13: Die Standardnormalverteilung hat welchen Erwartungswert und welche Standardabweichung?

Zur Erinnerung: μ = Erwartungswert, σ = Standardabweichung, σ^2 = Varianz.

- (a) $\mu = 1, \sigma = 0$
- (b) $\mu = 1, \sigma = 1$
- (c) $\mu = 0, \sigma = 1$

Lösungen zu Klickerfragen:

Frage 1: (c) Eine Zufallsgrösse ist eine Funktion, die jedem Element des Ergebnisraums eine reelle Zahl zuordnet; Frage 2: (a) Anzahl der geworfenen Sechsen in 10 Würfeln. Diskrete Zufallsgrössen nehmen nur abzählbare Werte an (z. B. 0, 1, 2, ...). Temperatur und Körpergrösse sind stetige Zufallsgrössen; Frage 3: (a) $P(X \leq 2) = P(X = 1) + P(X = 2) = 0.2 + 0.5 = 0.7$; Frage 4: (b) $P(X \neq 2) = P(X = 1) + P(X = 3) = 0.2 + 0.3 = 0.5$; Frage 5: (a) Die Wahrscheinlichkeitsfunktion einer stetigen Zufallsgrösse gibt Wahrscheinlichkeiten direkt an. Bei stetigen Zufallsgrössen gibt es keine Wahrscheinlichkeitsfunktion, sondern eine Dichtefunktion; Frage 6: (c) $F(x)$ kann auch abnehmen. Eine Verteilungsfunktion ist immer monoton wachsend, da mit wachsendem x die Wahrscheinlichkeit einer immer grösseren Ereignismenge bestimmt wird ($A \subset B \rightarrow P[A] \leq P[B]$); Frage 7: (b) Lebensdauern von Geräten; Frage 8: (b) Die Fläche unter der Kurve ist 1; Frage 9: (b) Die Summe aller Wahrscheinlichkeiten ist 1; Frage 10: (b) Etwa 0.68; Frage 11: (c) $\frac{3}{8}$. Die Binomialverteilung für $n = 3$, $k = 1$, $p = 0.5$; Frage 12: (a) Sie ist immer symmetrisch um ihren Mittelwert; Frage 13: (c) $\mu = 0, \sigma = 1$.