

Spezielle stetige Verteilungen

- Nun: Vorstellung spezieller parametrischer Verteilungsfamilien von *stetigen* Verteilungen
- In Verallgemeinerung des **Trägers** diskreter Verteilungen:
Träger $T(X)$ einer stetigen Verteilung als „**Bereich positiver Dichte**“.
- Wegen Möglichkeit, Dichtefunktionen abzuändern, etwas genauer:

$$T(X) := \{x \in \mathbb{R} \mid \text{es gibt eine Dichtefunktion } f_X \text{ von } X \text{ und ein } \epsilon > 0 \\ \text{mit } (f_X(t) > 0 \text{ für alle } t \in [x - \epsilon, x]) \\ \text{oder } (f_X(t) > 0 \text{ für alle } t \in [x, x + \epsilon])\}$$

Stetige Gleichverteilung

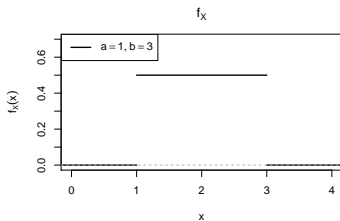
- Einfachste stetige Verteilungsfamilie:
Stetige Gleichverteilung auf Intervall $[a, b]$
- Modellierung einer stetigen Verteilung, in der alle Realisationen in einem Intervall $[a, b]$ als „gleichwahrscheinlich“ angenommen werden.
- Verteilung hängt von den beiden Parametern $a, b \in \mathbb{R}$ mit $a < b$ ab.
- Dichtefunktion f_X einer gleichverteilten Zufallsvariablen X kann auf Intervall $[a, b]$ konstant zu $\frac{1}{b-a}$ gewählt werden.
- Träger der Verteilung: $T(X) = [a, b]$
- Symbolschreibweise für stetige Gleichverteilung auf $[a, b]$: $X \sim \text{Unif}(a, b)$

Stetige GleichverteilungUnif(a, b)

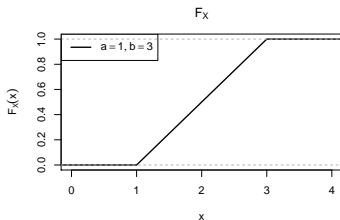
Parameter:

 $a, b \in \mathbb{R}$ mit $a < b$ Träger: $T(X) = [a, b]$ Dichtefunktion: $f_X : \mathbb{R} \rightarrow \mathbb{R}$;

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion: $F_X : \mathbb{R} \rightarrow \mathbb{R}$;

$$F_X(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } a \leq x \leq b \\ 1 & \text{für } x > b \end{cases}$$



Momente: $E(X) = \frac{a+b}{2}$

$$\gamma(X) = 0$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

$$\kappa(X) = \frac{9}{5}$$

Normalverteilung

- Verteilung entsteht als Grenzverteilung bei Durchschnittsbildung vieler (unabhängiger) Zufallsvariablen (später mehr!) \rightsquigarrow Einsatz für Näherungen
- Familie der Normalverteilungen hat Lageparameter $\mu \in \mathbb{R}$, der mit Erwartungswert übereinstimmt, und Streuungsparameter $\sigma^2 > 0$, der mit Varianz übereinstimmt, Standardabweichung ist dann $\sigma := +\sqrt{\sigma^2}$.
- Verteilungsfunktion von Normalverteilungen schwierig zu handhaben, Berechnung muss i.d.R. mit Software/Tabellen erfolgen.

- Wichtige Eigenschaft der Normalverteilungsfamilie:

Ist X normalverteilt mit Parameter $\mu = 0$ und $\sigma^2 = 1$, dann ist $aX + b$ für $a, b \in \mathbb{R}$ normalverteilt mit Parameter $\mu = b$ und $\sigma^2 = a^2$.

\rightsquigarrow Zurückführung allgemeiner Normalverteilungen auf den Fall der **Standardnormalverteilung (Gauß-Verteilung)** mit Parameter $\mu = 0$ und $\sigma^2 = 1$, Tabellen/Algorithmen für Standardnormalverteilung damit einsetzbar.

- Dichtefunktion der Standardnormalverteilung: φ , Verteilungsfunktion: Φ .
- Träger aller Normalverteilungen ist $T(X) = \mathbb{R}$.
- Symbolschreibweise für Normalverteilung mit Parameter μ, σ^2 : $X \sim N(\mu, \sigma^2)$

Normalverteilung

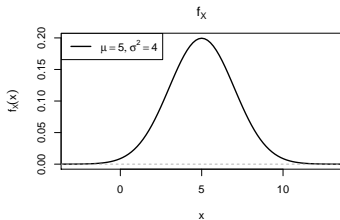
$$N(\mu, \sigma^2)$$

Parameter:

$$\mu \in \mathbb{R}, \sigma^2 > 0$$

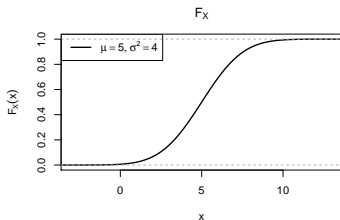
Träger: $T(X) = \mathbb{R}$ Dichtefunktion: $f_X : \mathbb{R} \rightarrow \mathbb{R}$;

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma} \varphi\left(\frac{x-\mu}{\sigma}\right)$$



Verteilungsfunktion:

$$F_X : \mathbb{R} \rightarrow \mathbb{R}; F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

Momente: $E(X) = \mu$ $\text{Var}(X) = \sigma^2$ $\gamma(X) = 0$ $\kappa(X) = 3$

Exponentialverteilung

- Beliebte Verteilungsfamilie zur Modellierung von **Wartezeiten**.
- Verteilung entsteht als Grenzverteilung der geometrischen Verteilung (Anzahl Fehlversuche vor erstem Erfolg bei wiederholter, unabhängiger Ausführung eines Bernoulli-Experiments) bei Erfolgswahrscheinlichkeit $p \rightarrow 0$.
- Da die Anzahl X der benötigten Versuche für $p \rightarrow 0$ offensichtlich immer größere Werte annehmen wird, wird statt der *Anzahl* der benötigten Versuche die *Zeit* zur Durchführung der benötigten Versuche modelliert, und mit $p \rightarrow 0$ zugleich die *pro Zeiteinheit* durchgeführten Versuche n des Bernoulli-Experiments so erhöht, dass $p \cdot n =: \lambda$ konstant bleibt.
- Einziger Parameter der resultierenden Exponentialverteilung ist damit die als „erwartete Anzahl von Erfolgen pro Zeiteinheit“ interpretierbare Größe $\lambda > 0$.
- Ist X exponentialverteilt mit Parameter λ , so erhält man $F_X(x)$ aus der Verteilungsfunktion der geometrischen Verteilung für $x \geq 0$ gemäß

$$F_X(x) = \lim_{n \rightarrow \infty} 1 - \left(1 - \frac{\lambda}{n}\right)^{n \cdot x} = \lim_{n \rightarrow \infty} 1 - \left(1 + \frac{-\lambda \cdot x}{n \cdot x}\right)^{n \cdot x} = 1 - e^{-\lambda x}.$$

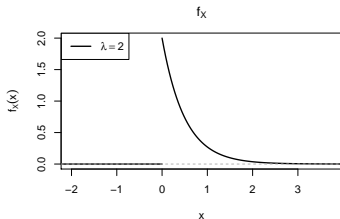
- Träger der Exponentialverteilungsfamilie ist $\mathbb{R}_+ := \{x \in \mathbb{R} \mid x \geq 0\}$.
- Symbolschreibweise für Exponentialverteilung mit Parameter λ : $X \sim \text{Exp}(\lambda)$

ExponentialverteilungExp(λ)

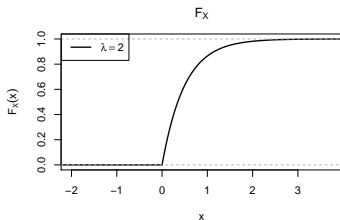
Parameter:

 $\lambda > 0$ Träger: $T(X) = \mathbb{R}_+$ Dichtefunktion: $f_X : \mathbb{R} \rightarrow \mathbb{R}$;

$$f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$$

Verteilungsfunktion: $F_X : \mathbb{R} \rightarrow \mathbb{R}$;

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 - e^{-\lambda x} & \text{für } x \geq 0 \end{cases}$$

Momente: $E(X) = \frac{1}{\lambda}$ $\text{Var}(X) = \frac{1}{\lambda^2}$ $\gamma(X) = 2$ $\kappa(X) = 9$

Verwendung spezieller Verteilungen

- Übliche Vorgehensweise zur Berechnung von (Intervall-)Wahrscheinlichkeiten für Zufallsvariablen X : Verwendung der Verteilungsfunktion F_X
- Problem bei einigen der vorgestellten Verteilungen:
Verteilungsfunktion F_X schlecht handhabbar bzw. nicht leicht auszuwerten!
- Traditionelle Lösung des Problems: *Vertafelung* bzw. *Tabellierung* der Verteilungsfunktionswerte, Ablesen der Werte dann aus Tabellenwerken.
- Lösung nicht mehr zeitgemäß: (kostenlose) PC-Software für alle benötigten Verteilungsfunktionen verfügbar, zum Beispiel Statistik-Software **R** (<http://www.r-project.org>)
- **Aber:** In Klausur keine PCs verfügbar, daher dort Rückgriff auf Tabellen.
- Problematische Verteilungsfunktionen (bisher) sind die der Standardnormalverteilung, Binomialverteilung sowie Poisson-Verteilung.
- Tabellen oder Tabellenausschnitte zu diesen Verteilungen werden in Klausur (sofern benötigt) zur Verfügung gestellt!
- Auch das Bestimmen von Quantilen ist für diese Verteilungen nicht ohne Hilfsmittel möglich und muss mit Hilfe weiterer Tabellen oder auf Grundlage der tabellierten Verteilungsfunktionswerte erfolgen.

Ausschnitt aus Tabelle für $\Phi(x)$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936

Ausschnitt aus Tabelle für $F_{B(n,p)}(x)$

n	x	$p = 0.05$	$p = 0.10$	$p = 0.15$	$p = 0.20$	$p = 0.25$	$p = 0.30$	$p = 0.35$	$p = 0.40$
1	0	0.9500	0.9000	0.8500	0.8000	0.7500	0.7000	0.6500	0.6000
1	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600
2	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400
2	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160
3	1	0.9928	0.9720	0.9392	0.8960	0.8438	0.7840	0.7182	0.6480
3	2	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360
3	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296
4	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752
4	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208
4	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744
4	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778
5	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370
5	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826
5	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130
5	4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898
5	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467
6	1	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333
6	2	0.9978	0.9842	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443
6	3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208
6	4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590
6	5	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9982	0.9959
6	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

R-Befehle für spezielle Verteilungen

- Verteilungsfunktionen können sofort nach dem Start von **R** mit den folgenden Befehlen ausgewertet werden:

Verteilung von X	Parameter	F_X an Stelle x mit R
$B(n, p)$	size= n , prob= p	pbinom(x , size, prob)
Geom(p)	prob= p	pgeom(x , prob)
Pois(λ)	lambda= λ	ppois(x , lambda)
Unif(a, b)	min= a , max= b	punif(x , min, max)
$N(\mu, \sigma^2)$	mean= μ , sd= $\sqrt{\sigma^2}$	pnorm(x , mean, sd)
Exp(λ)	rate= λ	pexp(x , rate)

- Ersetzt man in den Befehlen den ersten Buchstaben p durch d (z.B. dnorm), so erhält man den Wert der Dichtefunktion bzw. Wahrscheinlichkeitsfunktion an der Stelle x .
- Ersetzt man in den Befehlen den ersten Buchstaben p durch q (z.B. qnorm) und x durch p , so erhält man das (bzw. ein) p -Quantil der zugehörigen Verteilung.
- Ersetzt man schließlich in den Befehlen den ersten Buchstaben p durch r (z.B. rnorm) und x durch $n \in \mathbb{N}$, so erhält man n (Pseudo-)Zufallszahlen zur zugehörigen Verteilung.

Hinweise zur Tabellennutzung

- Bezeichnet $F_{B(n,p)}$ für $n \in \mathbb{N}$ und $p \in (0, 1)$ die Verteilungsfunktion der $B(n, p)$ -Verteilung, so gilt (!)

$$F_{B(n,1-p)}(x) = 1 - F_{B(n,p)}(n - x - 1)$$

für alle $n \in \mathbb{N}$, $p \in (0, 1)$, $x \in \{0, \dots, n - 1\}$. Daher werden Tabellen zur Binomialverteilung nur für $p \in (0, 0.5]$ erstellt, und die benötigten Werte für $p \in [0.5, 1)$ mit obiger Formel aus den Werten für $p \in (0, 0.5]$ gewonnen.

- Wegen der Symmetrie der Standardnormalverteilung um 0 gilt nicht nur $\varphi(x) = \varphi(-x)$ für alle $x \in \mathbb{R}$, sondern auch (vgl. Folie 216)

$$\Phi(x) = 1 - \Phi(-x) \quad \text{für alle } x \in \mathbb{R} .$$

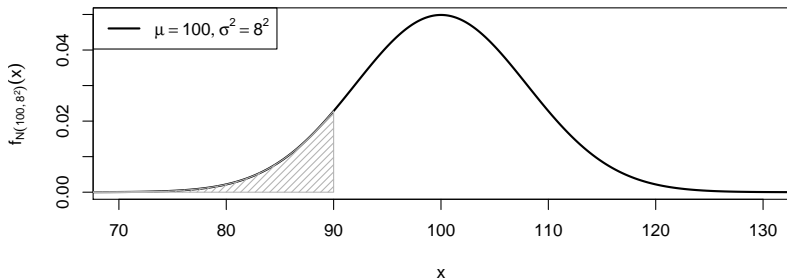
Daher werden Tabellen für $\Phi(x)$ in der Regel nur für $x \in \mathbb{R}_+$ erstellt.

- Zur Bestimmung von Quantilen darf in der Klausur ein beliebiger Wert des Intervalls, in dem das Quantil laut Tabelle liegen muss, eingesetzt werden; eine lineare Interpolation ist zwar sinnvoll, aber nicht nötig!
- Generell gilt: Ist ein Wert nicht tabelliert, wird stattdessen ein „naheliegender“ Wert aus der Tabelle eingesetzt.

Beispiel: Für fehlenden Wert $F_{B(4,0.28)}(2)$ wird $F_{B(4,0.3)}(2)$ eingesetzt.

Beispiel: Arbeiten mit Normalverteilungstabelle

- Frage:** Mit welcher Wahrscheinlichkeit nimmt eine $N(100, 8^2)$ -verteilte Zufallsvariable Werte kleiner als 90 an? (Wie groß ist die schraffierte Fläche?)

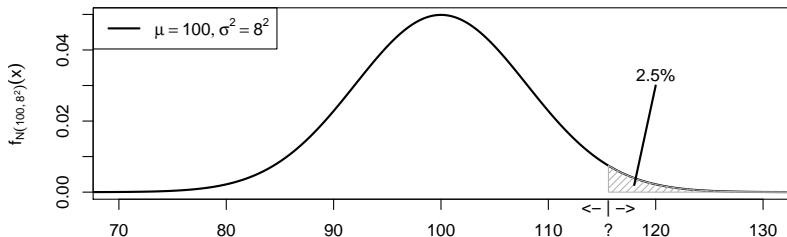


- Antwort:** Ist $X \sim N(100, 8^2)$, so gilt:

$$\begin{aligned} P\{X < 90\} &= F_{N(100, 8^2)}(90) = \Phi\left(\frac{90 - 100}{8}\right) \\ &= \Phi(-1.25) = 1 - \Phi(1.25) = 1 - 0.8944 = 0.1056 \end{aligned}$$

↪ Die gesuchte Wahrscheinlichkeit ist $0.1056 = 10.56\%$.

- **Frage:** Welchen Wert x überschreitet eine $N(100, 8^2)$ -verteilte Zufallsvariable nur mit 2.5% Wahrscheinlichkeit? (Welche linke Grenze x führt bei der schraffierten Fläche zu einem Flächeninhalt von 0.025?)



- **Antwort:** Ist $X \sim N(100, 8^2)$, so ist das 97.5%- bzw. 0.975-Quantil von X gesucht. Mit

$$F_X(x) = F_{N(100, 8^2)}(x) = \Phi\left(\frac{x - 100}{8}\right)$$

erhält man

$$\begin{aligned} \Phi\left(\frac{x - 100}{8}\right) &\stackrel{!}{=} 0.975 \Leftrightarrow \frac{x - 100}{8} = \Phi^{-1}(0.975) = 1.96 \\ &\Rightarrow x = 8 \cdot 1.96 + 100 = 115.68 \end{aligned}$$

Beispiel: Arbeiten mit Statistik-Software R

- Beantwortung der Fragen (noch) einfacher mit Statistik-Software **R**:
- **Frage:** Mit welcher Wahrscheinlichkeit nimmt eine $N(100, 8^2)$ -verteilte Zufallsvariable Werte kleiner als 90 an?
- **Antwort:**

```
> pnorm(90, mean=100, sd=8)
[1] 0.1056498
```
- **Frage:** Welchen Wert x überschreitet eine $N(100, 8^2)$ -verteilte Zufallsvariable nur mit 2.5% Wahrscheinlichkeit?
- **Antwort:**

```
> qnorm(0.975, mean=100, sd=8)
[1] 115.6797
```

oder alternativ

```
> qnorm(0.025, mean=100, sd=8, lower.tail=FALSE)
[1] 115.6797
```

Inhaltsverzeichnis

(Ausschnitt)

10 Mehrdimensionale Zufallsvariablen

- Borelsche sigma-Algebra
- Diskrete Zufallsvektoren
- Stetige Zufallsvektoren
- Randverteilungen
- (Stochastische) Unabhängigkeit
- Bedingte Verteilungen
- Momente zweidimensionaler Zufallsvektoren
- Momente höherdimensionaler Zufallsvektoren

Mehrdimensionale Zufallsvariablen/Zufallsvektoren I

- Im Folgenden: *Simultane Betrachtung mehrerer* (endlich vieler) Zufallsvariablen über *demselben* Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) .
- Ist $n \in \mathbb{N}$ die Anzahl der betrachteten Zufallsvariablen, so fasst man die n Zufallsvariablen X_1, \dots, X_n in einem n -dimensionalen Vektor $\mathbf{X} = (X_1, \dots, X_n)'$ zusammen.
- Damit ist \mathbb{R}^n der Wertebereich der Abbildung $\mathbf{X} : \Omega \rightarrow \mathbb{R}^n$, als σ -Algebra über \mathbb{R}^n wählt man die n -dimensionale Borelsche σ -Algebra \mathcal{B}^n , in der alle karthesischen Produkte von n Elementen aus \mathcal{B} enthalten sind.
- Insbesondere enthält \mathcal{B}^n alle endlichen und abzählbar unendlichen Teilmengen von \mathbb{R}^n sowie alle karthesischen Produkte von n Intervallen aus \mathbb{R} .
- Damit lassen sich die meisten bekannten Konzepte eindimensionaler Zufallsvariablen leicht übertragen.
- Ähnlich zur Situation bei mehrdimensionalen Merkmalen in der deskriptiven Statistik werden viele Darstellungen im Fall $n > 2$ allerdings schwierig.

Mehrdimensionale Zufallsvariablen/Zufallsvektoren II

Definition 10.1 (Zufallsvektor, Mehrdimensionale Zufallsvariable)

Seien (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum, $n \in \mathbb{N}$,

$$\mathbf{X} : \Omega \rightarrow \mathbb{R}^n : \mathbf{X}(\omega) = (X_1(\omega), \dots, X_n(\omega))'$$

eine $\mathcal{F} - \mathcal{B}^n$ -messbare Abbildung. Dann heißen \mathbf{X} **n -dimensionale Zufallsvariable** bzw. **n -dimensionaler Zufallsvektor** über (Ω, \mathcal{F}, P) und die gemäß Definition 8.3 gebildete Bildwahrscheinlichkeit

$$P_{\mathbf{X}} : \mathcal{B}^n \rightarrow \mathbb{R}; \mathbf{B} \mapsto P(\mathbf{X}^{-1}(\mathbf{B}))$$

(gemeinsame) **Wahrscheinlichkeitsverteilung** oder kürzer **(gemeinsame) Verteilung** von \mathbf{X} . $(\mathbb{R}^n, \mathcal{B}^n, P_{\mathbf{X}})$ ist damit ebenfalls ein Wahrscheinlichkeitsraum. Liegt nach Durchführung des Zufallsexperiments (Ω, \mathcal{F}, P) das Ergebnis $\omega \in \Omega$ vor, so heißt der zugehörige Wert $\mathbf{x} = \mathbf{X}(\omega)$ die **Realisierung** oder **Realisation** von \mathbf{X} .

Mehrdimensionale Zufallsvariablen/Zufallsvektoren III

- Wie im eindimensionalen Fall sind Kurzschreibweisen (zum Beispiel) der Form

$$P\{X_1 \leq x_1, \dots, X_n \leq x_n\} := P_{\mathbf{X}}((-\infty, x_1] \times \dots \times (-\infty, x_n])$$

für $x_1, \dots, x_n \in \mathbb{R}$ geläufig.

- Auch hier legen die Wahrscheinlichkeiten solcher Ereignisse die Verteilung des n -dimensionalen Zufallsvektors bereits eindeutig fest, und man definiert analog zum eindimensionalen Fall die gemeinsame Verteilungsfunktion

$$F_{\mathbf{X}} : \mathbb{R}^n \rightarrow \mathbb{R}; F_{\mathbf{X}}(x_1, \dots, x_n) := P_{\mathbf{X}}((-\infty, x_1] \times \dots \times (-\infty, x_n]) .$$

- Gemeinsame Verteilungsfunktionen mehrdimensionaler Zufallsvariablen sind allerdings für den praktischen Einsatz im Vergleich zur eindimensionalen Variante relativ unbedeutend und werden daher hier nicht weiter besprochen.

Diskrete Zufallsvektoren I

- Ist analog zum eindimensionalen Fall

$$\mathbf{X}(\Omega) := \{\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n \mid \mathbf{x} = \mathbf{X}(\omega) \text{ für (mindestens) ein } \omega \in \Omega\}$$

endlich oder abzählbar unendlich bzw. existiert (wiederum etwas allgemeiner) eine endliche oder abzählbar unendliche Menge $\mathbf{B} \subseteq \mathbb{R}^n$ mit $P(\{\mathbf{X} \in \mathbf{B}\}) = 1$, so nennt man auch solche Zufallsvektoren „diskret“.

- Mit Hilfe einer (mehrdimensionalen) Wahrscheinlichkeitsfunktion $p_{\mathbf{X}}$ mit $p_{\mathbf{X}}(\mathbf{x}) := P_{\mathbf{X}}(\{\mathbf{x}\})$ für $\mathbf{x} \in \mathbb{R}^n$ können Wahrscheinlichkeiten $P\{\mathbf{X} \in \mathbf{A}\}$ für Ereignisse $\mathbf{A} \in \mathcal{B}^n$ wiederum durch Aufsummieren der Punktwahrscheinlichkeiten aller Trägerpunkte \mathbf{x}_i mit $\mathbf{x}_i \in \mathbf{A}$ berechnet werden, also durch:

$$P\{\mathbf{X} \in \mathbf{A}\} = \sum_{\mathbf{x}_i \in \mathbf{A} \cap T(\mathbf{X})} p_{\mathbf{X}}(\mathbf{x}_i) \quad \text{für alle } \mathbf{A} \in \mathcal{B}^n$$

Diskrete Zufallsvektoren II

Definition 10.2 (Diskreter Zufallsvektor)

Seien (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum, $n \in \mathbb{N}$, \mathbf{X} ein n -dimensionaler Zufallsvektor über (Ω, \mathcal{F}, P) und $\mathbf{B} \subseteq \mathbb{R}^n$ endlich oder abzählbar unendlich mit $P(\{\mathbf{X} \in \mathbf{B}\}) = 1$. Dann nennt man

- ▶ \mathbf{X} einen **diskreten Zufallsvektor**,
- ▶ $p_{\mathbf{X}} : \mathbb{R}^n \rightarrow [0, 1]$; $p_{\mathbf{X}}(\mathbf{x}) := P_{\mathbf{X}}(\{\mathbf{x}\})$ die **(gemeinsame) Wahrscheinlichkeitsfunktion** von \mathbf{X} ,
- ▶ $T(\mathbf{X}) := \{\mathbf{x} \in \mathbb{R}^n \mid p_{\mathbf{X}}(\mathbf{x}) > 0\}$ den **Träger** von \mathbf{X} sowie alle Elemente $\mathbf{x} \in T(\mathbf{X})$ **Trägerpunkte** von \mathbf{X} und deren zugehörige Wahrscheinlichkeitsfunktionswerte $p_{\mathbf{X}}(\mathbf{x})$ **Punktwahrscheinlichkeiten**.

Stetige Zufallsvektoren I

- Zweiter wichtiger Spezialfall (wie im eindimensionalen Fall):
stetige n -dimensionale Zufallsvektoren **X**
- Wiederum gilt $P_{\mathbf{X}}(\mathbf{B}) = 0$ *insbesondere* für alle endlichen oder abzählbar unendlichen Teilmengen $\mathbf{B} \subseteq \mathbb{R}^n$.
- Auch hier ist die definierende Eigenschaft die Möglichkeit zur Berechnung spezieller Wahrscheinlichkeiten als Integral über eine (nun mehrdimensionale) Dichtefunktion.
- In Verallgemeinerung der Berechnung von *Intervallwahrscheinlichkeiten* im eindimensionalen Fall müssen nun *Wahrscheinlichkeiten von Quadern* als (Mehrfach-)Integral über eine Dichtefunktion berechnet werden können.

Stetige Zufallsvektoren II

Definition 10.3 (Stetiger Zufallsvektor, (gemeinsame) Dichtefunktion)

Seien (Ω, \mathcal{F}, P) ein Wahrscheinlichkeitsraum, $n \in \mathbb{N}$ und \mathbf{X} ein n -dimensionaler Zufallsvektor über (Ω, \mathcal{F}, P) . Gibt es eine nichtnegative Abbildung $f_{\mathbf{X}} : \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$P_{\mathbf{X}}(\mathbf{A}) = \int_{a_1}^{b_1} \cdots \int_{a_n}^{b_n} f_{\mathbf{X}}(t_1, \dots, t_n) dt_n \cdots dt_1 \quad (5)$$

für alle Quader $\mathbf{A} = (a_1, b_1] \times \cdots \times (a_n, b_n] \subset \mathbb{R}^n$ mit $a_1 \leq b_1, \dots, a_n \leq b_n$, so heißt der Zufallsvektor \mathbf{X} **stetig**. Jede nichtnegative Abbildung $f_{\mathbf{X}}$ mit der Eigenschaft (5) heißt **(gemeinsame) Dichtefunktion** von \mathbf{X} .

Notationen im Spezialfall $n = 2$

- Im Folgenden wird (auch für weitere Anwendungen) regelmäßig der Spezialfall $n = 2$ betrachtet.
- Zur Vereinfachung der Darstellung (insbesondere zur Vermeidung doppelter Indizes) sei der betrachtete Zufallsvektor \mathbf{X} dann mit $\mathbf{X} = (X, Y)'$ oder $\mathbf{X} = (X, Y)$ statt $\mathbf{X} = (X_1, X_2)'$ bezeichnet.
- Stetige 2-dimensionale Zufallsvektoren $\mathbf{X} = (X, Y)$ werden in der Regel durch die Angabe einer gemeinsamen Dichtefunktion

$$f_{X,Y} : \mathbb{R}^2 \rightarrow \mathbb{R}; (x, y) \mapsto f_{X,Y}(x, y)$$

spezifiziert.

- Ist $\mathbf{X} = (X, Y)$ ein zweidimensionaler diskreter Zufallsvektor mit „wenigen“ Trägerpunkten, stellt man die gemeinsame Verteilung — analog zu den Kontingenztabelle der deskriptiven Statistik — gerne in Tabellenform dar.

Beispiel: (Gemeinsame) Wahrscheinlichkeitsfunktion

bei zweidimensionaler diskreter Zufallsvariable

- Ist $\mathbf{X} = (X, Y)$ zweidimensionale diskrete Zufallsvariable mit endlichem Träger, $A := T(X) = \{x_1, \dots, x_k\}$ der Träger von X und $B := T(Y) = \{y_1, \dots, y_l\}$ der Träger von Y , so werden die Werte der Wahrscheinlichkeitsfunktion $p_{\mathbf{X}}$ auch mit

$$p_{ij} := p_{(X,Y)}(x_i, y_j) \quad \text{für } i \in \{1, \dots, k\} \text{ und } j \in \{1, \dots, l\}$$

bezeichnet und wie folgt tabellarisch dargestellt:

$X \setminus Y$	y_1	y_2	\dots	y_l
x_1	p_{11}	p_{12}	\dots	p_{1l}
x_2	p_{21}	p_{22}	\dots	p_{2l}
\vdots	\vdots	\vdots	\ddots	\vdots
x_k	p_{k1}	p_{k2}	\dots	p_{kl}

Randverteilungen I

- Wie in der deskriptiven Statistik lassen sich die Verteilungen der einzelnen Zufallsvariablen eines n -dimensionalen Zufallsvektors auch aus der gemeinsamen Verteilung gewinnen.
- Analog zu den „Randhäufigkeiten“ erhält man so die **Randverteilungen** der einzelnen Komponenten des Zufallsvektors.
- Ist \mathbf{X} diskreter n -dimensionaler Zufallsvektor mit Wahrscheinlichkeitsfunktion $p_{\mathbf{X}}$, so erhält man für $j \in \{1, \dots, n\}$ die Wahrscheinlichkeitsfunktion p_{X_j} zur j -ten Komponente X_j durch:

$$p_{X_j}(x) = \sum_{\substack{\mathbf{x}_i = (x_{i,1}, \dots, x_{i,n}) \\ x_{i,j} = x}} p_{\mathbf{X}}(\mathbf{x}_i)$$

Randverteilungen II

- Ist \mathbf{X} stetiger n -dimensionaler Zufallsvektor mit gemeinsamer Dichtefunktion $f_{\mathbf{X}}$, so erhält man für $j \in \{1, \dots, n\}$ eine Dichtefunktion $f_{X_j} : \mathbb{R} \rightarrow \mathbb{R}$ zur j -ten Komponente X_j durch:

$$f_{X_j}(x) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{(n-1)\text{-mal}} f_{\mathbf{X}}(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_n) dx_n \cdots dx_{j+1} dx_{j-1} \cdots dx_1$$

- Für $\mathbf{X} = (X, Y)$ erhält man also eine Randdichtefunktion f_X zu X durch

$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$

sowie eine Randdichtefunktion f_Y zu Y durch

$$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx .$$

Fortsetzung Beispiel (zweidimensional, diskret)

Ergänzung der gemeinsamen Wahrscheinlichkeitstabelle um Randverteilungen

- Ist $A = T(X) = \{x_1, \dots, x_k\}$ der Träger von X und $B = T(Y) = \{y_1, \dots, y_l\}$ der Träger von Y , so erhält man für $i \in \{1, \dots, k\}$ als Zeilensummen

$$p_{i\cdot} := p_X(x_i) = \sum_{j=1}^l p_{(X,Y)}(x_i, y_j) = \sum_{j=1}^l p_{ij}$$

sowie für $j \in \{1, \dots, l\}$ als Spaltensummen

$$p_{\cdot j} := p_Y(y_j) = \sum_{i=1}^k p_{(X,Y)}(x_i, y_j) = \sum_{i=1}^k p_{ij}$$

und damit insgesamt die folgende ergänzte Tabelle:

$X \setminus Y$	y_1	y_2	\cdots	y_l	$p_{i\cdot}$
x_1	p_{11}	p_{12}	\cdots	p_{1l}	$p_{1\cdot}$
x_2	p_{21}	p_{22}	\cdots	p_{2l}	$p_{2\cdot}$
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots
x_k	p_{k1}	p_{k2}	\cdots	p_{kl}	$p_{k\cdot}$
$p_{\cdot j}$	$p_{\cdot 1}$	$p_{\cdot 2}$	\cdots	$p_{\cdot l}$	1

(Stochastische) Unabhängigkeit von Zufallsvariablen I

- Die Komponenten X_1, \dots, X_n eines n -dimensionalen Zufallsvektors werden genau dann **stochastisch unabhängig** genannt, wenn alle Ereignisse der Form

$$\{X_1 \in B_1\}, \dots, \{X_n \in B_n\} \quad \text{für } B_1, \dots, B_n \in \mathcal{B}$$

stochastisch unabhängig sind:

Definition 10.4

Seien $n \in \mathbb{N}$ und X_1, \dots, X_n Zufallsvariablen über demselben Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) . Dann heißen die Zufallsvariablen X_1, \dots, X_n **(stochastisch) unabhängig**, wenn für alle $B_1, \dots, B_n \in \mathcal{B}$ gilt:

$$P\{X_1 \in B_1, \dots, X_n \in B_n\} = \prod_{i=1}^n P\{X_i \in B_i\} = P\{X_1 \in B_1\} \cdot \dots \cdot P\{X_n \in B_n\}$$

(Stochastische) Unabhängigkeit von Zufallsvariablen II

- Man kann weiter zeigen, dass n diskrete Zufallsvariablen X_1, \dots, X_n über einem gemeinsamen Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) genau dann stochastisch unabhängig sind, wenn für den (in diesem Fall ebenfalls diskreten) Zufallsvektor $\mathbf{X} = (X_1, \dots, X_n)'$ bzw. die zugehörigen Wahrscheinlichkeitsfunktionen für alle $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ gilt:

$$p_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n p_{X_i}(x_i) = p_{X_1}(x_1) \cdot \dots \cdot p_{X_n}(x_n)$$

- Insbesondere sind X und Y im Fall $n = 2$ mit $\mathbf{X} = (X, Y)$ bei endlichen Trägern $A = T(X) = \{x_1, \dots, x_k\}$ von X und $B = T(Y) = \{y_1, \dots, y_l\}$ von Y unabhängig, falls $p_{ij} = p_i \cdot p_j$ gilt für alle $i \in \{1, \dots, k\}$ und $j \in \{1, \dots, l\}$.

(Stochastische) Unabhängigkeit von Zufallsvariablen III

- Weiterhin sind n stetige Zufallsvariablen X_1, \dots, X_n über einem gemeinsamen Wahrscheinlichkeitsraum (Ω, \mathcal{F}, P) stochastisch unabhängig, wenn der Zufallsvektor $\mathbf{X} = (X_1, \dots, X_n)'$ stetig ist und eine gemeinsame Dichte $f_{\mathbf{X}}$ bzw. Randdichten f_{X_1}, \dots, f_{X_n} existieren, so dass für alle $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ gilt:

$$f_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n f_{X_i}(x_i) = f_{X_1}(x_1) \cdot \dots \cdot f_{X_n}(x_n)$$

- Insbesondere sind stetige Zufallsvariablen X und Y im Fall $n = 2$ mit $\mathbf{X} = (X, Y)$ genau dann unabhängig, wenn es Dichtefunktionen f_X von X , f_Y von Y sowie $f_{X,Y}$ von (X, Y) gibt mit

$$f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y) \quad \text{für alle } x, y \in \mathbb{R}.$$