1 General

1.1 Potenzen und Wurzeln

- $a^1 = a$ • $a^0 = 1$
- $a^{-n} = \frac{1}{a^n}$
- $a^{\frac{1}{n}} = \sqrt[n]{a}$ • $a^{\frac{m}{n}} = \sqrt[n]{a^m}$
- $a^m a^n = a^{m+n}$
- $\frac{a^m}{a^n} = a^{m-n}$
- $(a^m)^n = a^{mn}$
- $a^n b^n = (ab)^n$
- $\frac{a^n}{b^n} = \left(\frac{a}{b}\right)^n$
- $a^b = e^{b \ln a}$
- $e^{a \ln b} = b^a$
- $e^{a+b} = e^a \cdot e^b$
- $e^0 = 1$

1.2 Logarithmen

$y = \log_a x \iff a^x = x$

- $a^{\log_a x} = x$
- $\log_a a^x = x$
- $\log_a a = 1$ • $\log_a 1 = 0$
- $\log(uv) = \log(u) + \log(v)$
- $\log(\frac{u}{v}) = \log(u) - \log(v)$
- $\log(u^r) = r \log(u)$
- $\log_a x = \frac{\log_b x}{\log_b a}$
- $\ln(a \cdot b) = \ln(a) + \ln(b)$
- $\ln(1) = 0$

1.3 Trigonometrische Funktionen

- **Sinussatz:** $\frac{a}{\sin\alpha} = \frac{b}{\sin\beta} = \frac{c}{\sin\gamma} = 2r$
- **Cosinussatz:** $a^2 = b^2 + c^2 - 2bc\cos\alpha$
- **Tangens:** $\tan(z) := \frac{\sin(z)}{\cos(z)}, z \notin \{\frac{\pi}{2} + \pi k\}$
- **Cotangens:** $\cot(z) := \frac{\cos(z)}{\sin(z)}, z \notin \{\pi k\}$
- $\exp(iz) = \cos(x) + i\sin(z)$
- $\cos(z) = \cos(-z)$ • $\sin(-z) = -\sin(z)$
- $\sin(z) = \frac{e^{iz} - e^{-iz}}{2i}$ • $\cos(z) = \frac{e^{iz} + e^{-iz}}{2}$
- $\sin(z+w) = \sin(z)\cos(w) + \cos(z)\sin(w)$ • $\cos(z+w) = \cos(z)\cos(w) - \sin(z)\sin(w)$
- $\cos(z)^2 + \sin(z)^2 = 1$
- $\sin(2z) = 2\sin(z)\cos(z)$ • $\cos(2z) = \cos(z)^2 - \sin(z)^2 = 1 - 2\sin^2(z) = 2\cos^2(z) - 1$
- $\tan a = \frac{\sin a}{\cos a}$
- $1 + \tan^2 a = \frac{1}{\cos^2 a}$

1.3.1 Winkel

deg	rad	sin	cos	deg	rad	sin	cos
0	0	0	1	30	$\frac{\pi}{6}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$
45	$\frac{\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	60	$\frac{\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$
90	$\frac{\pi}{2}$	1	0	120	$\frac{2\pi}{3}$	$\frac{\sqrt{3}}{2}$	$\frac{-1}{2}$
135	$\frac{3\pi}{4}$	$\frac{\sqrt{2}}{2}$	$\frac{-\sqrt{2}}{2}$	150	$\frac{5\pi}{6}$	$\frac{1}{2}$	$\frac{-\sqrt{3}}{2}$
180	π	0	-1	210	$\frac{7\pi}{6}$	$\frac{-1}{2}$	$\frac{-\sqrt{3}}{2}$
225	$\frac{5\pi}{4}$	$\frac{-\sqrt{2}}{2}$	$\frac{-\sqrt{2}}{2}$	240	$\frac{4\pi}{3}$	$\frac{-\sqrt{3}}{2}$	$\frac{-1}{2}$
270	$\frac{3\pi}{2}$	-1	0	300	$\frac{5\pi}{3}$	$\frac{-\sqrt{3}}{2}$	$\frac{1}{2}$
315	$\frac{7\pi}{4}$	$\frac{-\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	330	$\frac{11\pi}{6}$	$\frac{-1}{2}$	$\frac{\sqrt{3}}{2}$

$\arcsin(0) = 0$	$\arccos(0) = \frac{\pi}{2}$	$\arctan(0) = 0$
$\arcsin(1) = \frac{\pi}{2}$	$\arccos(0) = 0$	$\arctan(1) = \frac{\pi}{4}$
$\arcsin(-1) = -\frac{\pi}{2}$	$\arccos(0) = \pi$	$\arctan(-1) = -\frac{\pi}{4}$
		$\arctan(\sqrt{3}) = \frac{\pi}{3}$

1.4 Sonstiges

- **Mitternacht:** $x_{1/2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$
- $\sqrt{i} = \frac{\sqrt{2}}{2} + i\frac{\sqrt{2}}{2}$

1.5 Komplexe Zahlen

- **Imaginary Number:** $i, i^2 = -1$
- **Complex Number:** $z, z = x + iy, \ x, y \in \mathbb{R}$
- **Set:** $\mathbb{C} = \{x + iy | x, y \in \mathbb{R}\}$
- **Konjugate:** $\overline{z} = x - iy, \ z = x + iy$
  - ⋆ $z \cdot \overline{z} = x^2 \cdot y^2 = |z|^2$
  - ⋆ $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$
  - ⋆ $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$
- **Betrag:** $|z|$ Distanz zwischen z und Origin
  - ⋆ $|z| = \sqrt{x^2 + y^2} = \sqrt{z \cdot \overline{z}}$
  - ⋆ $|z| = |\overline{z}|$
  - ⋆ $|z_1 + z_2| \leq |z_1| + |z_2|$
  - ⋆ $|z_1 \cdot z_2| = |z_1||z_2|$
  - ⋆ $|\frac{z_1}{z_2}| = \frac{|z_1|}{|z_2|}$
- **Euler:** $e^{i\gamma} = \cos\gamma + i\sin\gamma$
  - ⋆ $|e^{i\gamma}| = 1$

1.5.1 Arithmetic

Für $z_1 = a + ib = re^{i\gamma}, z_2 = b + id = se^{i\delta}$:

- $z_1 + z_2 = (a + ib) + (c + id) = (a + c) + i(b + d)$
- $z_1 \cdot z_2 = (a + ib) \cdot (c + id) = (ac - bd) + i(ad + bc)$
  - ⋆ $z_1 z_2 = rse^{i(\gamma+\delta)}$
- $\frac{z_1}{z_2} = \frac{z_1 \cdot \overline{z_2}}{z_2 \cdot \overline{z_2}} = \frac{z_1 \cdot \overline{z_2}}{|z_2|}$
  - ⋆ $\frac{z_1}{z_2} = \frac{r}{s} e^{i(\gamma - \delta)}$
- $\sqrt[n]{z_1} = z_2 \implies z_1 = z_2^n = r^n e^{in\delta} \stackrel{!}{=} re^{i\gamma}$
  - ⋆ $s = \sqrt[n]{r}$
  - ⋆ $n\gamma = \gamma + 2\pi k, k = 0 \dots n-1$

1.5.2 Polar Coordinates

$z = x + iy \iff z = r(\cos\gamma + i\sin\gamma) \stackrel{Euler}{=} z = re^{i\gamma}$

- $x = r\cos\gamma$ • $y = r\sin\gamma$
- $r = |z|$
- $\gamma = \arccos\frac{x}{r} = \arcsin\frac{y}{r}$
- $\arg z = |z| \in [0, 2\pi[\implies r$ ist eindeutig bestimmt

1.6 Rechenregeln Ableitung

- **f + g:** $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.
- **f · g:** $(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$.
- $\mathbf{\frac{f}{g}}$: $\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}, \ g(x_0) \neq 0$.
- **g ∘ f:** $(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0)$

1.7 Rechenregel Integral

- **Partiell:** $\int_a^b f(x)g'(x)\mathrm{d}x = f(x)g(x)|_a^b - \int_a^b f'(x)g(x)\mathrm{d}x$
- **Substitution:** $\int_{\phi(a)}^{\phi(b)} f(x)\mathrm{d}x = \int_a^b f(\phi(t))\phi'(t)\mathrm{d}t$

1.8 Bekannte Reihen

Reihe		Wert	konv.	div.
Geometrische Reihe		$q \in \mathbb{C}$		
$\sum_{k=0}^{\infty} aq^k$	$a + aq +$	$\frac{a}{1-q}$	$\lvert g\rvert < 1$	$\lvert q\rvert \geq 1$
$\sum_{k=0}^{\infty}(k+1)q^k$	$1 + 2q +$	$\frac{1}{(1-q)^2}$		
Harmonische Reihe				
$\sum_{k=1}^{\infty} \frac{1}{k}$		∞		
$\sum_{k=1}^{\infty} \frac{1}{k^2}$		$\frac{\pi^2}{6}$		
$\sum_{k=1}^{\infty} \frac{1}{k^4}$		$\frac{\pi^4}{90}$		
$\sum_{k=1}^{\infty} \frac{1}{k^a}$			$a > 1$	$a \leq 1$
Alternierende Harmon. Reihe				
$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$		$\ln 2$		
$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2}$		$\frac{\pi^2}{12}$		
$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^4}$		$\frac{\pi^4}{720}$		
$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$	$1 - \frac{1}{3} + \frac{1}{5} -$	$\frac{\pi}{4}$		
Teleskopreihe				
$\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$		1		
Exponentialfunktion $z \in \mathbb{C}$, konv. abs.				
$\sum_{k=0}^{\infty} \frac{z^k}{k!}$	$1 + z + \frac{z^2}{2!} +$	$\exp z$		
$\sum_{k=0}^{\infty} \frac{(-a)^k}{k!}$		$\frac{1}{e^a}$		
$\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{(2k+1)!}$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} -$	$\sin x$		
$\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k}}{(2k)!}$	$1 - \frac{x^2}{2} + \frac{x^4}{4!} -$	$\cos x$		
$\sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}$	$x + \frac{x^3}{3!} + \frac{x^5}{5!} +$	$\sinh x$		
$\sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$	$1 + \frac{x^2}{2} + \frac{x^4}{4!} +$	$\cosh x$		

1.9 Spezielle Summen

$\sum_{k=1}^{n} k$	$1 + 2 + \dots + n$	$\frac{n(n+1)}{2}$
$\sum_{k=1}^{n} k^2$	$1 + 4 + \dots + n^2$	$\frac{n(n+1)(2n+1)}{6}$
$\sum_{k=1}^{n} k^3$	$1 + 8 + \dots n^3$	$\left(\sum_{k=1}^{\infty} k\right)^2$
$\sum_{k=1}^{n}(2k-1)$	$1 + 3 + \dots + (2n-1)$	n^2
$\sum_{k=1}^{n}(2k-1)^2$	$1 + \dots + (2n-1)^2$	$\frac{n(2n-1)(2n+1)}{3}$
$\sum_{k=0}^{n-1} q^k$	$1 + q + \dots = \frac{q^n - 1}{q-1}$	$\frac{1-q^n}{1-q}, \ q \notin \{0, 1\}$

1.10 Funktionen und deren Grenzwert

Funktion	Grenzwert	Bedingung
$\lim_{n\to\infty} a^n$	0	$\lvert a\rvert < 1$
$\lim_{n\to\infty} \sqrt[n]{a}$	1	$a > 0$
$\lim_{n\to\infty} \sqrt[n]{n^a}$	1	$a > 0$
$\lim_{n\to\infty} \sqrt[n]{n}$	1	
$\lim_{n\to\infty} \frac{\log_a n}{n}$	0	$a > 1$
$\lim_{n\to\infty} \frac{n^k}{a^n}$	0	$a > 1$
$\lim_{n\to\infty} \frac{a^n}{n!}$	0	
$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k}$	∞	
$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$	e	
$\lim_{n\to\infty} \left(1 + \frac{a}{n}\right)^n$	e^a	
$\lim_{n\to\infty} \left(1 - \frac{1}{n}\right)^n$	$\frac{1}{e}$	
$\lim_{n\to 0} \frac{\sin n}{n}$	1	
$\lim_{n\to 1} \frac{\ln n}{n-1}$	1	
$\lim_{n\to\infty} \frac{n^m}{\exp(an)}$	0	$m \in \mathbb{R}, a > 0$
$\lim_{n\to 0} \frac{\exp(n)-1}{n}$	1	
$\lim_{n\to 0} \frac{\ln(1+n)}{n}$	1	
$\lim_{n\to\infty} \frac{\ln n}{n^a}$	0	$a > 0$
$\lim_{n\to 0} \frac{a^n - 1}{n}$	$\ln a$	$a > 0$
$\lim_{n\to 0}(n^a \ln n)$	0	$a > 0$

1.11 Auf- und Ableitungen

$f'(x)$	$f(x)$	$F(x)$
sax^{s-1}	ax^s	$\frac{a}{s+1}x^{s+1}$
$\frac{1}{x}$	$\ln(ax)$	$x\ln(ax) - x$
$\frac{1}{x+a}$	$\ln(x+a)$	$(x+a)\ln(x+a) + x$
ae^{ax}	e^{ax}	$\frac{1}{a}e^{ax}$
$\frac{-a}{x^2}$	$\frac{a}{x}$	$a\ln(\lvert x\rvert)$
$\ln(a)ba^{bx}$	a^{bx}	$\frac{a^{bx}}{\ln(a)b}$
$\frac{1}{2\sqrt{x}}$	$\sqrt{x}$	$\frac{2}{3}x^{3/2}$
$\frac{1}{x\ln(a)}$	$\log_a(x)$	$\frac{x(\ln(x)-1)}{\ln(a)}$
$\cos(x)$	$\sin(x)$	$-\cos(x)$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$	$x\arcsin(x) + \sqrt{1-x^2}$
$2\sin(x)\cos(x)$	$\sin^2(x)$	$\frac{x - \sin(x)\cos(x)}{2}$
$-\sin(x)$	$\cos(x)$	$\sin(x)$
$\frac{-1}{\sqrt{1-x^2}}$	$\arccos(x)$	$x\arccos(x) - \sqrt{1-x^2}$
$-2\sin(x)\cos(x)$	$\cos^2(x)$	$\frac{x + \sin(x)\cos(x)}{2}$
$\frac{1}{\cos^2(x)} = 1 + \tan^2(x)$	$\tan(x)$	$-\ln(\lvert\cos(x)\rvert)$
$\frac{1}{1+x^2}$	$\arctan(x)$	$x\arctan(x) - \frac{\ln(x^2+1)}{2}$
$2a(\cos(ax))^2 - a$	$\sin(ax)\cos(ax)$	$-\frac{1}{4a}\cos(2ax) = \frac{-(\cos^2(ax))}{2a}$
	$\sin(x)\sin^2(x)$	$\frac{-\sin^2(x)-2}{3}\cos(x)$
	$\cos(x)\cos^2(x)$	$\frac{\sin(x)\cos^2(x)+2\sin(x)}{3}$
	$\sin(x)\cos^2(x)$	$\frac{-\cos^3(x)}{3}$
	$\sin^2(x)\cos(x)$	$\frac{\sin^3(x)}{3}$
	$\sin(x)\sin(2x)$	$\frac{-\sin(3x)+3\sin(x)}{6}$
	$\cos(x)\cos(2x)$	$\frac{\sin(3x)+3\sin(x)}{6}$
$\frac{2x}{(x^2+1)^2}$	$\frac{x^2}{x^2+1}$	$x - \arctan(x)$
$\frac{-4x}{(x^2+1)^3}$	$\frac{1}{(1+x^2)^2}$	$\frac{x}{2x^2+2} + \frac{\arctan(x)}{2}$
	$\frac{u'(x)}{u(x)}$	$\ln\lvert u(x)\rvert$
	$u'(x) \cdot u(x)$	$\frac{1}{2}(u(x)^2)$
$\frac{-x}{\sqrt{a^2-x^2}}$	$\sqrt{a^2 - x^2}$	$\frac{a^2\arcsin(\frac{x}{\lvert a\rvert}) + x\sqrt{a^2-x^2}}{2}$
$\frac{x}{\sqrt{x^2+a^2}}$	$\sqrt{a^2 + x^2}$	$\frac{a^2\ln(\lvert\sqrt{x^2+a^2}+x\rvert) + x\sqrt{a^2+x^2}}{2}$
$x^x(1 + \ln x)$	x^x	
$e^{\ln\lvert x\rvert}(\ln(\lvert x\rvert) + 1)$	$\lvert x\rvert^x = e^{x\ln x}$	
$\frac{-2}{(x+a)^3}$	$\frac{1}{(x+a)^2}$	$\frac{-1}{x+a}$
$\cosh(x)$	$\sinh(x)$	$\cosh(x)$
$\frac{1}{\sqrt{1+x^2}}$	$\operatorname{arcsinh}(x)$	
$\sinh(x)$	$\cosh(x)$	$\sinh(x)$
$\frac{1}{\sqrt{x-1}\sqrt{x+1}}$	$\operatorname{arccosh}(x)$	
$\frac{1}{\cosh^2(x)} = 1 - \tanh^2(x)$	$\tanh(x)$	$\ln(e^{2x} + 1) - x$
$\frac{1}{2x+2} - \frac{1}{2x-2}$	$\operatorname{arctanh}(x)$	

1.12 Kombinatorik

	geordnet	ungeordnet
mit zurücklegen	n^k	$\binom{n+k-1}{k}$
ohne zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k} = \frac{n!}{k!(n-k)!}$

Part I

Wahrscheinlichkeiten

2 Wahrscheinlichkeit

2.1 Grundbegriffe

- **Ereignisraum Ω:** Menge aller möglichen Ergebnisse des Zufallsexperiments
  - ⋆ $\Omega \neq \emptyset$
- **Elementarereignisse $w \in \Omega$:** Element aus Ereignisraum
- **Potenzmenge 2^{Ω}:** von Ω ist die Menge aller Teilmengen von Ω
- **Prinzipielles Ereignis $A \subseteq \Omega$:** Set von Elementarereignissen
  - ⋆ A tritt ein, falls das realisierte Elementarereignis $w \in A$
- **Beobachtbare Ereignisse $\mathcal{F}$:** Teilmenge der Potenzmenge
  - ⋆ **Diskreter Fall:** Dann ist oft $\mathcal{F} = 2^{\Omega}$
  - ⋆ **Ω überabzählbar:**
    - ∗ $\mathcal{F} \subset 2^{\Omega}$ und
    - ∗ $\mathcal{F}$ ist eine σ-Algebra
- **σ-Algebra** ist $\mathcal{F}$ falls

(i) $\Omega \in \mathcal{F}$

(ii) $\forall A \in \mathcal{F}$ ist auch Komplement $A^{\complement} \in \mathcal{F}$

(iii) $\forall$ Folge $(A_n)_{n\in\mathbb{N}}$ mit $A_n \in \mathcal{F}$ für alle $n \in \mathbb{N}$ ist auch $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

- **Wahrscheinlichkeitsmass P:** Abbildung $P : \mathcal{F} \to [0,1]$ mit folgenden Axiomen:

A0) $P[A] \geq 0 \quad \forall A \in \mathcal{F}$

A1) $P[\Omega] = 1$

A2) $P\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} P[A_i]$ für disjunkte Ereignisse A_i

Daraus folgt
  - ⋆ $P[A^{\complement}] = 1 - P[A]$
  - ⋆ $P[\emptyset] = 0$
  - ⋆ $A \subseteq B \implies P[A] \leq P[B]$
  - ⋆ **Additionsregel:** $P[A \cup B] = P[A] + P[B] - P[A \cap B]$

2.2 Diskrete Wahrscheinlichkeitsräume

- Ω endlich oder abzählbar und $\mathcal{F} = 2^{\Omega}$

2.2.1 Laplace Raum

- Falls:
  - ⋆ $\Omega = \{\omega_1, \ldots, \omega_N\}$ endlich mit $|\Omega| = N$ und
  - ⋆ Alle ω_i gleich wahrscheinlich, also $p_i = 1/N$

  dann:
  - ⋆ Ω einen **Laplace Raum** und
  - ⋆ P ist die *diskrete Gleichverteilung*
- $P[A] = \frac{\text{Anz. Elementarereignisse in } A}{\text{Anz. Elementarereignisse in } \Omega} = \frac{|A|}{|\Omega|}$

2.3 Bedingte Wahrscheinlichkeit

- Falls für A, B Ereignisse und $P[A] > 0$

  Dann $P[B|A] := \frac{P[B\cap A]}{P[A]}$
- Bei fixierter Bedingung A ist $P[\cdot|A]$ wieder ein Wahrscheinlichkeitsmass auf $(\Omega, \mathcal{F})$
  - ⋆ I.e. $P[\cdot|A]$ erfüllt die Axiome A0), A1), A2
- **Multiplikationsregel:** $P[A \cap B] = P[B \mid A] \cdot P[A]$

2.3.1 Satz der totalen Wahrscheinlichkeit

- Falls: $A_1, \ldots, A_n$ eine Zerlegung von Ω in paarweise disjunkte Ereignisse
  - ⋆ D.h. $\bigcup_{i=1}^{n} A_i = \Omega$ und $A_i \cap A_k = \emptyset \, \forall i \neq k$

  Dann: $P[B] = \sum_{i=1}^{n} P[B \mid A_i] \cdot P[A_i]$

2.3.2 Satz von Bayes:

- Falls
  - ⋆ $A_1, \ldots, A_n$ eine Zerlegung von Ω mit $P[A_i] > 0$ für $i = 1 \ldots n$ und
  - ⋆ B ein Ereignis mit $P[B] > 0$

  Dann gilt für jedes k
  - ⋆ $P[A_k \mid B] = \frac{P[B|A_k]\cdot P[A_k]}{\sum_{i=1}^{n} P[B|A_i]\cdot P[A_i]}$
  - ⋆ $P[A \mid B] = \frac{P[A\cap B]}{P[B]} = \frac{P[B|A]\cdot P[A]}{P[B|A]\cdot P[A] + P[B|A^C]\cdot P[A^C]}$

2.4 Unabhängigkeit von Ereignissen

2.4.1 Unabhängigkeit von 2 Ereignissen

- A und B sind stochastisch unabhängig, falls eins von:
  - ⋆ $P[A \cap B] = P[A] \cdot P[B]$
  - ⋆ $P[A] = 0$ oder $P[B] = 0$
  - ⋆ $P[A] \neq 0, P[B \mid A] = P[B]$
  - ⋆ $P[B] \neq 0, P[A \mid B] = P[A]$

2.4.2 Allgemeine Unabhängigkeit

- Falls für jede endliche Teilfamilie die Produktformel gilt
  - ⋆ D.h. für ein $m \in \mathbb{N}$ und $\{k_1, \ldots, k_m\} \subseteq \{1, \ldots, n\}$ gilt immer $P\left[\bigcap_{i=1}^{m} A_{k_i}\right] = \prod_{i=1}^{m} P[A_{k_i}]$

  Dann sind $A_1, \ldots, A_n$ stochastisch unabhängig
- **Paarweise Unabhängig:** Falls die definition nur für alle paare gilt
  - ⋆ Allgemeine Unabhängigkeit $\implies$ Paarweise Unabhängigkeit

3 Diskrete Zufallsvariablen und Verteilungen

3.1 Grundbegriffe

- **Zufallsvariable (ZV):** $X : \Omega \to \mathbb{R}$
  - ⋆ Falls Σ endlich/abzählbar dann ist Wertebereich $\mathcal{W}(X)$ auch endlich
- **Verteilungsfunktion (VF)** von X: $F_X : \mathbb{R} \to [0,1]$
  - ⋆ $F_X(t) := P[X \leq t] := P[\{\omega \mid X(\omega) \leq t\}]$
- **Gewichtsfunktion** von X: $p_X : \mathcal{W}(X) \to [0,1]$
  - ⋆ $p_X(x_k) := P[X = x_k] = P[\{\omega \mid X(\omega) = x_k\}], k = 1, 2, \ldots$
  - ⋆ $0 \leq p_X(x_k) \leq 1, \forall x_k \in \mathcal{W}(X)$
  - ⋆ $\sum_{x_k\in\mathcal{W}(X)} p_X(x_k) = 1$
- **Indikatorfunktion:** Sein A eine Ereignis, dann ist $Y = I_A$ die Indikatorfunktion von A:

  $Y(\omega) = \begin{cases} 1 & \text{für } \omega \in A \\ 0 & \text{für } \omega \notin A \end{cases}$
  - ⋆ Es gilt $P[Y = 1] = P[A]$

3.1.1 Erwartungswert

- **Erwartungswert (EW):** $E[X] := \sum_{x_k\in\mathcal{W}(X)} x_k p_X(x_k)$
  - ⋆ Definiert falls $\sum_{x_k\in\mathcal{W}(X)} |x_k| p_X(x_k) < \infty$
- Sei $Y = g(X), g : \mathbb{R} \to \mathbb{R}$, dann: $E[Y] = E[g(X)] = \sum_{x_k\in\mathcal{W}(X)} g(x_k) p_X(x_k)$
- X, Y ZV und EW existieren, dann:
  - ⋆ **Monotonie:** $\forall \omega \; X(\omega) \leq Y(\omega) \implies E[X] \leq E[Y]$
  - ⋆ **Linearität:** Für bel. $a, b \in \mathbb{R}$ gilt $E[aX + b] = aE[X] + b$
  - ⋆ Falls $\mathcal{W}(X) = \mathbb{N}_0$ so gilt $E[X] = \sum_{j=1}^{\infty} P[X \geq j] = \sum_{l=0}^{\infty} P[X > l]$
- **Bedingter EW:** $E[X_1 \mid X_2](x_2) = \sum_{x_1\in\mathbb{R}} x_1 p_{X_1|X_2}(x_1 \mid x_2) =: h(x_2)$
  - ⋆ $h : \mathcal{W}(X_2) \to \mathbb{R}$
  - ⋆ $E[E[X_1 \mid X_2]] = E[X_1]$

3.1.2 Varianz

- **Varianz:** $Var[X] := E[(X - E[X])^2]$
  - ⋆ Definiert falls $E[X^2] < \infty$
  - ⋆ $Var[X] = \sum_{x_k\in\mathcal{W}(X)} (x_k - E[X])^2 p_X(x_k)$
  - ⋆ Alternative definition $Var[X] = E[X^2] - (E[X])^2$
  - ⋆ $Var[aX + b] = a^2 Var[X]$
- **Standardabweichung:** $\sigma(X) := \sqrt{Var[X]}$
- **Kovarianz:** $Cov(X, Y) := E[XY] - E[X]E[Y] = E[(X - E[X])(Y - E[Y])]$
- $Var[X + Y] = Var[X] + Var[Y] + 2Cov(X, Y)$
- $Cov(X, X) = Var[X]$
- **Unkorreliert:** Falls $Cov(X, Y) = 0$
- **Paarweise Unkorreliert:** Falls alle Paare X_i, X_j mit $i \neq j$ unkorreliert sind

3.2 Mehrere Zufallsvariable

- **Gemeinsame Verteilungsfunktion:**
  - ⋆ $F : \mathbb{R}^n \to [0,1]$
  - ⋆ $F(x_1, \ldots, x_n) := P[X_1 \leq x_1, \ldots, X_n \leq x_n]$ für $X_1, \ldots, X_n$ ZV
- **Gemeinsame Gewichtsfunktion:**
  - ⋆ $p : \mathbb{R}^n \to [0,1]$
  - ⋆ $p(x_1, \ldots, x_n) := P[X_1 = x_1, \ldots, X_n = x_n]$ für $X_1, \ldots, X_n$ diskrete ZV
- **(Verteilungsfunktion der) Randverteilung:**
  - ⋆ Sofern X und Y gemeinsame VF F haben
  - ⋆ $F_X : \mathbb{R} \to [0,1]$
  - ⋆ $F_X(x) := P[X \leq x] = \lim_{y\to\infty} P[X \leq x, Y < y] = \lim_{y\to\infty} F(x, y)$
- **Gewichtsfunktion der Randverteilung:**
  - ⋆ $p_X : \mathcal{W}(X) \to [0,1]$
  - ⋆ $p_x(x) = P[X = x] = \sum_{y_j\in\mathcal{W}(Y)} P[X = x, Y = y_j] = \sum_{y_j\in\mathcal{W}(Y)} p(x, y_j)$
- ZV $X_1, \ldots, X_n$ heissen **unabhängig**, falls $F(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n)$ oder falls ZV diskret $p(x_1, \ldots, x_n) = p_{X_1}(x_1) \ldots p_{X_n}(x_n)$

3.3 Funktionen von mehreren ZV

- Seien $X_1, \ldots, X_n$ diskrete ZV und $g : \mathbb{R}^n \to \mathbb{R}$, dann ist $Y = g(X_1, \ldots, X_n)$ eine diskret ZV
- Seien $X_1, \ldots, X_n$ diskret ZV mit endlichen EW und $Y = a + \sum_{l=1}^{n} b_l X_l$. Dann ist $E[Y] = a + \sum_{l=1}^{n} b_l E[X_l]$
- **Summenformel für Varianzen:** $Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var[X_i] + 2\sum_{i<j} Cov(X_i, X_j)$
- Falls $X_1, \ldots, X_n$ paarweise unkorreliert dann $Var[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} Var[X_i]$
- Seien $X_1, \ldots, X_n$ diskret, unabhängige ZV mit endlichen EW, dann: $E[\Pi_{i=1}^{n} X_i] = \Pi_{i=1}^{n} E[X_i]$
- unabhängig $\implies$ paarweise unabhängig $\implies$ unkorreliert

3.4 Summe von zwei ZV

- Seien X, Y diskret ZV mit gemeinsamer GF $p(x, y)$, dann ist $Z = X + Y$ diskret. GF von Z: $p_Z(z) = P[Z = z] = \sum_{x_k\in\mathcal{W}(X)} P[X = x_k, Y = z - x_k] = \sum_{x_k\in\mathcal{W}(X)} p(x_k, z - x_k)$
- Sind X, Y unabhängig ZV, so ist $p(x, y) = p_X(x) p_Y(y)$ und damit: $p_Z(z) = \sum_{x_k\in\mathcal{W}(X)} p_x(x_k) p_Y(z - x_k) = p_X * p_Y$

3.5 Bedingte Verteilungen

- Seien X, Y diskrete ZV mit gemeinsamer GF $p(x, y)$. Die bedingte GF von X, gegeben $Y = y$ ist: $p_{X|Y}(x \mid y) := P[X = x \mid Y = y] = \frac{P[X=x, Y=y]}{P[Y=y]} = \frac{p(x,y)}{p_Y(y)}$
  - ⋆ Falls $p_Y(y) > 0$ sonst 0
- Es gilt $\sum_{x_k\in\mathcal{W}(X)} p_{X|Y}(x_k \mid y) = \frac{1}{p_Y(y)} \sum_{x_k\in\mathcal{W}(X)} p(x_k, y) = 1$
- Analog für Verteilungsfunktion

3.6 Wichtige diskrete Verteilungen

3.6.1 Diskrete Gleichverteilung

- Sei X eine ZV $\mathcal{W}(X) = \{x_1, \ldots, x_N\}$ dann ist $p_X(x_k) = P[X = x_k] = \frac{1}{N}, \; k = 1, \ldots, N$

3.6.2 Bernoulli Verteilung

- $X \sim Be(p)$
- $p_X(x) = p^x(1-p)^{1-x}, \; x \in \{0,1\} = \mathcal{W}(X)$
- $E[X] = p$
- $Var[X] = p(1-p)$

3.6.3 Binomial Verteilung

- $X \sim Bin(n, p), \implies Bin(1, p) = Be(p)$
- $p_X(k) = P[X = k] = \binom{n}{k} p^k (1-p)^{n-k}, \; k = 0, \ldots, n$
- $\mathcal{W}(X) = \{0, \ldots, n\}$
- $E[X] = \sum_{i=1}^{n} E[Y_i] = np$
- $Var[X] = \sum_{n=1}^{n} Var[Y_i] = np(1-p)$
- $p(k+1; n) = \frac{p}{1-p} \frac{n-k}{k+1} p(k; n)$

3.6.4 Geometrische Verteilung

- $X \sim Geom(p)$
- $X = \inf\{i \in \mathbb{N} \mid A_i \text{ tritt ein}\} = \inf\{i \in \mathbb{N} \mid Y_i = 1\}$

- $p_X(k) = P[X = k] = p(1-p)^{k-1},\ k = 1, 2, \ldots$
- $P[X > k] = (1-p)^k$
- **Gedächtnislosigkeit:** $P[X > k + n | X > n] = P[X > k]$
- $\mathcal{W}(X) = \mathbb{N}$
- $E[X] = \sum_{l=0}^{\infty} P[X > l] = \sum_{l=0}^{\infty} (1-p)^l = \frac{1}{p}$
- $Var[X] = \frac{1-p}{p^2}$

3.6.5 Poisson Verteilung

- $X \sim \mathcal{P}(\lambda),\ \lambda \in (0, \infty)$
- $p_X(k) = e^{-\lambda} \frac{\lambda^k}{k!},\ k = 0, 1, \ldots$
- Sein $X_n \sim Bin(n, p_n)$ und $np_n = \lambda$. Lassen wir $n \to \infty$, so geht $p_n = \frac{\lambda}{n} = 0$. Weiter ist $\lim_{n\to\infty} P[X_n = k] = e^{-\lambda} \frac{\lambda^k}{k!} = P[X = k]$, wenn $X \sim \mathcal{P}(\lambda)$
  - ⋆ Gilt auch für $np_n \to \lambda$ für $n \to \infty$
  - ⋆ **Faustregel:** Approximation ist brauchbar für $np^2 \leq 0.05$
- $E[X] = \sum_{k=0}^{\infty} k p_X(k) = \sum_{k=1}^{\infty} k e^{-\lambda} \frac{\lambda^k}{k!} = \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} = \lambda \sum_{j=1}^{\infty} e^{-\lambda} \frac{\lambda^j}{j!} = \lambda$
- $Var[X] = \lambda$
- Für unabhängige ZV $X \sim \mathcal{P}(\lambda), Y \sim \mathcal{P}(\mu)$ is $X + Y \sim \mathcal{P}(\lambda + \mu)$

4 Allgemeine Zufallsvariablen

4.1 Grundbegriffe

- VF F_X hat Eigenschaften:
  - ⋆ Wachsend: $F_X(s) \leq F_X(t)$ für $s \leq t$
  - ⋆ Rechtsstetig: $F_X(u) \to F_X(t)$ für $u \to t$ mit $u > t$
  - ⋆ $\lim_{t\to-\infty} F_X(t) = 0$, $\lim_{t\to\infty} F_X(t) = 1$
- **Dichtefunktion (DF):** Eine ZV X mit VF $F_X(t)$ heisst (absolut)stetig mit Dichte(funktion) $f_X : \mathbb{R} \to [0, \infty)$, falls: $F_X(t) = \int_{-\infty}^{t} f_X(s) \mathrm{d}s\ \forall t \in \mathbb{R}$
  - ⋆ DF f_X hat Eigenschaften:
    - ∗ $f_X \geq 0$ und $f_X = 0$ ausserhalb von $\mathcal{W}(X)$
    - ∗ $\int_{-\infty}^{\infty} f_X(s) \mathrm{d}s = 1$
- $P[a < X \leq b] = P[X \leq b] - P[X \leq a] = F_X(b) - F_X(a) = \int_a^b f_X(s) \mathrm{d}s$.
- Allgemeiner für Mengen $B \subseteq \mathbb{R}$: $P[X \in B] = \int_B f_X(s) \mathrm{d}s$
- Dichtefunktion = Ableitung der Verteilungsfunktion
- **Erwartungswert:**
  - ⋆ Ist ZV X stetig mit DF $f_X(x)$ dann: $E[X] = \int_{-\infty}^{\infty} x f_X(x) \mathrm{d}x$ sofern das Integral absolut konvergiert.
  - ⋆ ZV X und ZV $Y = g(x)$. Ist X stetig mit DF $f_X(x)$, dann: $E[Y] = E[g(x)] = \int_{-\infty}^{\infty} g(x) f_X(x) \mathrm{d}x$ sofern das Integral abs. konvergiert.
  - ⋆ Für beliebige unabhängige ZV $X_1, \ldots, X_n$ ist $E[X_1 \cdot \cdots \cdot X_n] = E[X_1] \cdot \cdots \cdot E[X_n]$
  - ⋆ Für beliebige ZV $X_1, \ldots, X_n$ ist $E[X_1 + \cdots + X_n] = E[X_1] + \cdots + E[X_n]$
  - ⋆ **Bedingter EW:** $E[X_1 \mid X_2](x_2) = \int_{-\infty}^{\infty} x_1 f_{X_1|X_2}(x_1 \mid x_2) \mathrm{d}x_1 =: h(x_2)$
    - ∗ $h : \mathcal{W}(X_2) \to \mathbb{R}$
    - ∗ $E[E[X_1 \mid X_2]] = E[X_1]$
- **i.i.d.:** Haben ZVs $X_1, \ldots, X_n$ die selbe VF und sind unabhängig nennt man sie i.i.d.
- **Momente:** Die Momente von X sind definiert als $m_n := E[X^n], n \in \mathbb{N}$ unter der Voraussetzung, dass $E[|X|^n] < \infty$
  - ⋆ $m_1 = E[X]$
- **Zentralmomente:** Für $n \in \mathbb{N}$ und falls $E[|X|^n] < \infty$ ist, kann man das n-te Zentralmoment von X definieren als $\mu_n := E[(X - E[X])^n]$.
  - ⋆ $\mu_2 = E[(X - E[X])^2] = Var[X]$
- **Momenterzeugende Funktion** einer ZV X ist $M_X(t) := E[e^{tX}], t \in \mathbb{R}$
  - ⋆ Wohldefiniert in $[0, \infty]$, kann aber $+\infty$ werden

4.2 Wichtige stetige Verteilungen

4.2.1 Gleichverteilung

- ZV $X \sim \mathcal{U}(a, b)$ hat $\mathcal{W}(X) = [a, b]$
- DF $f_X(t) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq t \leq b \\ 0 & \text{sonst} \end{cases}$
- VF $F_X(t) = \begin{cases} 0 & \text{für } t < a \\ \frac{t-a}{b-a} & \text{für } a \leq t \leq b \\ 1 & \text{für } t > b \end{cases}$
- $E[X] = \int_a^b x \frac{1}{b-a} \mathrm{d}x = \frac{a+b}{2}$
- $Var[X] = E[X^2] - (E[X])^2 = \frac{1}{12}(b-a)^2$

4.2.2 Exponentialfunktion

- ZV $X \sim Exp(\lambda)$ für $\lambda > 0$ und $\mathcal{W}(X) = [0, \infty)$
- DF $f_X(t) = \begin{cases} \lambda e^{-\lambda t} & \text{für } t \geq 0 \\ 0 & \text{für } t < 0 \end{cases}$
- VF $F_X(t) = \int_{-\infty}^{t} f_X(s) \mathrm{d}s = \begin{cases} 1 - e^{-\lambda t} & \text{für } t \geq 0 \\ 0 & \text{für } t < 0 \end{cases}$
- $E[X] = \int_0^{\infty} \lambda x e^{-\lambda x} \mathrm{d}x = \frac{1}{\lambda}$
- $Var[X] = \int_0^{\infty} (x - \frac{1}{\lambda})^2 \lambda e^{-\lambda x} \mathrm{d}x = \frac{1}{\lambda^2}$
- **Gedächtnislosigkeit:** $P[X > t + s \mid X > s] = P[X > t]$

4.2.3 Normalverteilung

- ZV $X \sim \mathcal{N}(\mu, \sigma^2), \mu \in \mathbb{R}, \sigma^2 > 0$ hat $\mathcal{W}(X) = \mathbb{R}$,
- DF $f_X(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}},\ t \in \mathbb{R}$
- $E[X] = \mu,\ Var[X] = \sigma^2$
- **Standard-Normalverteilung:** Falls $\mathcal{N}(0, 1)$
  - ⋆ **DF:** $\varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}$
  - ⋆ **VF:** $\Phi(t) = \int_{-\infty}^{t} \varphi(s) \mathrm{d}s = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{s^2}{2}} \mathrm{d}s$
- $X \sim \mathcal{N}(\mu, \sigma^2)$, so ist $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1)$, also: $F_X(t) = P[X \leq t] = P[\frac{X-\mu}{\sigma} \leq \frac{t-\mu}{\sigma}] = \Phi(\frac{t-\mu}{\sigma})$
- Falls $X_1, \ldots, X_n$, unabhängig und $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ dann $\sum_{i=1}^n X_i \sim \mathcal{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$
- Fläche zwischen $[\mu - \sigma, \mu + \sigma] \approx \frac{2}{3}$

4.3 Gemeinsame Verteilungen, unabhängige ZV

- **Gemeinsame VF:** Analog zu diskret
- **Gemeinsame DF:** Falls die gemeinsame VF F von $X_1, \ldots, X_n$ sich schreiben lässt als $F(x_1, \ldots, x_n) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_n} f(t_1, \ldots, t_n) \mathrm{d}t_n \ldots \mathrm{d}t_1$ für eine Funktion $f : \mathbb{R}^n \to [0, \infty)$, so heisst $f(x_1, \ldots, x_n)$ die gemeinsame Dichte von $X_1, \ldots, X_n$. Es gilt:
  - ⋆ $f(x_1, \ldots, x_n) \geq 0$ und $= 0$ ausserhalb von $\mathcal{W}(X_1, \ldots, X_n)$
  - ⋆ $\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, \ldots, x_n) \mathrm{d}x_n \ldots \mathrm{d}x_1 = 1$
  - ⋆ $P[(X_1, \ldots, X_n) \in A] = \int_A f(x_1, \ldots, x_n) \mathrm{d}x_n \ldots \mathrm{d}x_1$ für $A \subseteq \mathbb{R}^n$
- **Randverteilung:** Analog zu diskret
- **Randverteilungsdichte:** Falls X und Y gemeinsame Dichte $f(x, y)$, so haben auch die Randverteilungen von X und Y Dichten $f_X, f_Y : \mathbb{R} \to [0, \infty)$. $f_X(x) = \int_{-\infty}^{\infty} f(x, y) \mathrm{d}y$

4.3.1 Unabhängigkeit

- Die ZV $X_1, \ldots, X_n$ heissen Unabhängig falls: $F(x_1, \ldots, x_n) = F_{X_1}(x_1) \ldots F_{X_n}(x_n)$ bzw. $f(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n)$

4.4 Funktionen und Transformationen von ZV

- Sei $Z = X + Y$ dann $F_Z(z) = P[Z \leq z] = P[X + Y \leq z]$. Setzen wir $A_z := \{(x, y) \in \mathbb{R}^2 \mid x + y \leq z\}$ so ist $F_Z(z) = P[(X, Y) \in A_z] = \int_{A_z} \int f(x, y) \mathrm{d}y \mathrm{d}x = \int_{-\infty}^{\infty} \int_{-\infty}^{z-x} f(x, y) \mathrm{d}y \mathrm{d}x$. Mit der Variablentransformation $v = x + y$ wird $y = v - x, \mathrm{d}y = \mathrm{d}v$ und $F_Z(z) = \int_{-\infty}^{\infty} \int_{-\infty}^{z} f(x, v - x) \mathrm{d}v \mathrm{d}x = \int_{-\infty}^{z} \int_{-\infty}^{\infty} f(x, v - x) \mathrm{d}x \mathrm{d}v$. Also hat Z auch eine DF: $f_Z(z) = \frac{d}{dz} F_Z(z) = \int_{-\infty}^{\infty} f(x, z - x) \mathrm{d}x = \int_{-\infty}^{\infty} f(z - y, y) \mathrm{d}y$.
  - ⋆ Sind zusätzlich X und Y unabhängig, so ist $f(x, y) = f_X(x) f_Y(y)$, sprich: $f_Z(z) = \int_{-\infty}^{\infty} f_x(x) f_Y(z - x) \mathrm{d}x = \int_{-\infty}^{\infty} f_X(z - y) f_Y(y) \mathrm{d}y =: (f_X * f_Y)(z)$.

4.4.1 ZV Transformationen

- Sei X ZV mit F_X und f_X. Für eine Funktion $g : \mathbb{R} \to \mathbb{R}$ betrachten wir die neue ZV $Y = g(X)$. Wie sehen dann VF und DF von Y aus?
- **Allgemeiner Ansatz:** $F_Y(t) = P[Y \leq t] = P[g(X) \leq t] = \int_{A_g} f_X(s) \mathrm{d}x$ mit Menge $A_g := \{s \in \mathbb{R} \mid g(s) \leq t\}$. Dann versuchen rechte Seite auszurechnen.
- z.B. für Affine Transformationen $g(x) = ax + b, a > 0, b \in \mathbb{R}$: $F_Y(t) = P[Y \leq t] = P[aX + b \leq t] = P[X \leq \frac{t-b}{a}] = F_X(\frac{t-b}{a})$ und damit nach der Kettenregel $f_Y(t) = \frac{d}{dt} F_Y(t) = \frac{1}{a} f_X(\frac{t-b}{a})$.
- Analog bis auf Vorzeichen geht das auch mit $a < 0$
- Sei ZV F stetige und streng monoton wachsende VF, mit Umkehrfunktion F^{-1}. Falls $X \sim \mathcal{U}(0, 1)$ und $Y = F^{-1}(X)$ dann hat Y die VF F.
  - ⋆ Dieser Satz erlaubt die explizite Konstruktion einer ZV Y mit gewünschter VF F.

5 Ungleichungen und Grenzwertsätze

5.1 Ungleichungen

- **Markov Ungleichung:** Sei ZV X und $g : \mathcal{W}(X) \to [0, \infty)$ eine wachsende Funktion. $\forall c \in \mathbb{R}$ mit $g(c) \geq 0$ gilt dann $P[X \geq c] \leq \frac{E[g(X)]}{g(c)}$
- **Chebyshev-Ungleichung:** Sei ZV Y mit endlicher Var. Für jedes $b > 0$ gilt dann $P[|Y - E[Y]| \geq b] \leq \frac{Var[Y]}{b^2}$
- **Chernoff Schranke:**
  - ⋆ Seien $X_1, \ldots, X_n$ i.i.d. ZV, für welche $M_X(t)$ für alle $t \in \mathbb{R}$ endlich ist. Für jedes $b \in \mathbb{R}$ gilt dann: $P[S_n \geq b] \leq \exp(\inf_{t\in\mathbb{R}}(n \log M_x(t) - tb))$
  - ⋆ Seien $X_1, \ldots X_n$ unabhängig mit $X_i \sim Be(p_i)$ und $S_n = \sum_{i=1}^n X_i$. Sei $\mu_n := E[S_n] = \sum_{i=1}^n p_i$ und $\delta > 0$. Dann gilt: $P[S_n \geq (1 + \delta)\mu_n] \leq (\frac{e^\delta}{(1+\delta)^{1+\delta}})^{\mu_n}$

5.2 Grenzwertsätze

- **Schwaches Gesetz der grossen Zahlen:** Sein $X_1, X_2, \ldots$ eine Folge von unabhängigen ZV (oder paarweise unkorreliert), die alle den gleichen EW $E[X_i] = \mu$ und die Var $Var[X_i] = \sigma^2$ haben. Sei: $\overline{X}_n = \frac{1}{n} S_n = \frac{1}{n} \sum_{i=1}^n X_i$. Dann konvergiert $\overline{X}_n$ für $n \to \infty$ gegen μ.
  - ⋆ D.h. $P[|\overline{X}_n - \mu| > \varepsilon] \underset{n\to\infty}{\to} 0, \forall \varepsilon > 0$.
- **Starkes Gesetz der grossen Zahlen:** Seien $X_1, X_2, \ldots$ i.i.d. und ihr EW $\mu = E[X_i]$ sei endlich. Für: $\overline{X}_n = \frac{1}{n} S_n = \frac{1}{n} \sum_{i=1}^n X_i$, dann gilt: $\overline{X}_n \underset{n\to\infty}{\to} \mu$ *P-fastsicher.*
  - ⋆ D.h. $P[\{\omega \in \Omega \mid \overline{X}_n(\omega) \underset{n\to\infty}{\to} \mu\}] = 1$.
- **Zentraler Grenzwertsatz (ZGS):** Sei $X_1, X_2, \ldots$ eine Folge von i.i.d. ZV mit $E[X] = \mu$ und $Var[X] = \sigma^2$. Für die Summe $S_n = \sum_{i=1}^n X_i$ gilt dann $\lim_{n\to\infty} P[\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x] = \Phi(x)\ \forall x \in \mathbb{R}$.
  - ⋆ Es gilt $E[S_n] = n\mu$ und $Var[S_n] = n\sigma^2$. Also gilt $S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}}$, $E[S_n^*] = 0$ und $Var[S_n^*] = 1$.
  - ⋆ $P[S_n^* \leq x] \approx \Phi(x)$ für n gross
  - ⋆ $S_n^* \overset{\text{approx}}{\sim} \mathcal{N}(0, 1)$ für n gross
  - ⋆ $S_n \overset{\text{approx}}{\sim} \mathcal{N}(n\mu, n\sigma^2)$
  - ⋆ $\overline{X}_n \overset{\text{approx}}{\sim} \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$
  - ⋆ Nützliche Umformung: $P[|S_n^*| \leq x] \approx \Phi(x) - \Phi(-x) = 2\Phi(x) - 1$
  - ⋆ **Kontinuitätskorrektur:** Bietet besser approximation für Normalverteilung durch "Zentrierung" der Stäbe des Histogramms durch hinzufügen der Korrektur $+\frac{1}{2}$
    - ∗ $P[S_n^* \leq \frac{b - np}{\sqrt{np(1-p)}}] \implies P[S_n^* \leq \frac{b + \frac{1}{2} - np}{\sqrt{np(1-p)}}]$

Part II

Statistik

6 Grundideen

- Man fasst die Daten $x_1, \dots, x_n$ auf als Realisierungen $X_1(\omega), \dots, X_n(\omega)$ von ZV $X_1, \dots, X_n$, und sucht dann Aussagen über die Verteilung von $X_1, \dots, X_n$.
  - ⋆ $x_1, \dots, x_n$ sind Daten (in der Regel Zahlen)
  - ⋆ $X_1, \dots, X_n$ sind der generierende Mechanismus (ZV)
  - ⋆ **Stichprobe:** ZV $X_1, \dots, X_n$
  - ⋆ **Stichprobenumfang:** Anzahl Stichproben n
- Wir haben einen Datensatz aus einer Stichprobe und suchen dafür ein Modell welches durch einen ggf. hochdimensionalen parameter $\vartheta \in \Theta$ beschreibbar ist. P_ϑ beschreibt das Wahrscheinlichkeitsmass abhängig von ϑ. Wir versuchen nun über die Daten Rückschlüsse auf Θ zu ziehen.
- Parametrische Statistische Analyse:

1. Aufgrund von Daten/Graphen eine Idee für die Wahl einer geeigneten Modelierung zu finden
2. **Wahl eines Modells:** Paremetermenge Θ und Familie $(P_\vartheta)_{\vartheta\in\Theta}$ von Modellen spezifizieren
3. Schätzung der Parameter durch Benutzung eines Schätzers
4. Überprüfen ob gewähltes ϑ bzw. Modell P_ϑ gut passt durch geeignete statistische Tests
5. Aussagen über Zuverlässigkeit der Schätzungen machen. Ev. Bereich in Θ wählen statt einzelner parameter ϑ (Konfidenzbereich)

7 Schätzer

- Seien $X_1, \dots, X_n$ eine Stichprobe, für die wir ein Modell suchen. Wir haben also einen Parameterraum Θ und für jedes $\vartheta \in \Theta$ einen Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P_\vartheta)$. Meistens ist $\Omega \in \mathbb{R}^m$, und wir suchen dann für die Parameter $\vartheta_1, \dots, \vartheta_n$ Schätzer $T_1, \dots, T_n$ aufgrund unsere Stichprobe. Solche Schätzer sind Zufallsvariablen der Form $T_j = t_j(X_1, \dots, X_n)$, wobei wir die Schätzfunkton $t_j : \mathbb{R}^n \to \mathbb{R}$ noch geeignet wählen/finden müssen. Einsetzen von Daten $x_i = X_i(\omega), i = 1, \dots, n$ liefert dann Schätzwert $T_j(\omega) = t_j(x_1, \dots, x_n)$ für $\omega_j, j = 1, \dots, m$. Der Kürze halber schreiben wir oft auch $T = (T_1, \dots, T_m)$ und $\vartheta = (\vartheta_1, \dots, \vartheta_m)$.
- Schätzer: ZV, Schätzwert: Zahl

7.1 Grundbegriffe

- **Erwartungstreu:** Schätzer T ist erwartungstreu für ϑ, falls gilt $E_\vartheta[T] = \vartheta$
  - ⋆ **Bias:** $E_\vartheta[T] - \vartheta$
- **Mittlere Quadratische Schätzfehler (MSE):** Definiert als $\mathrm{MSE}_\vartheta[T] := E_\vartheta[(T-\vartheta)^2]$
  - ⋆ $\mathrm{MSE}_\vartheta[T] = E_\vartheta[(T-\vartheta)^2] = \mathrm{Var}_\vartheta[T] + (E_\vartheta[T] - \vartheta)^2$
- **Konsistent:** Folge von Schätzern $T^{(n)}, n \in \mathbb{R}$, heisst konsistent für ϑ, falls $T^{(n)}$ für $n \to \infty$ gegen ϑ konvergiert. D.h. $\forall \vartheta \in \Theta$ gilt: $\lim_{n\to\infty} P_\vartheta[|T^{(n)} - \vartheta| > \varepsilon] = 0$

7.2 Maximum-Likelihood-Methode (ML-Methode)

- **Likelihood-Funktion:** $L(x_1, \dots, x_n; \vartheta) :=$ $\begin{cases} p(x_1, \dots, x_n; \vartheta) = \Pi_{i=1}^n p_X(x_i; \vartheta) & \text{diskreten Fall} \\ f(x_1, \dots, x_n; \vartheta) = \Pi_{i=1}^n f_X(x_i; \vartheta) & \text{stetigen Fall} \end{cases}$
  - ⋆ **Log-Likelihood:** Often brauchen wir $\log L(x_1, \dots, x_n; \vartheta)$ damit, wenn ZV i.i.d. sind, rechnen wir mit Summen und nicht Produkten.
- **Maximum-Likelihood-Schätzer (ML-Schätzer)** T_{ML} für ϑ maximiert $\vartheta \mapsto L(X_1, \dots, X_n; \vartheta)$ als Funktion von ϑ.
  - ⋆ Statt zu maximieren sucht man meistens nur Nullstellen der Ableitung nach ϑ.
    - ∗ Falls ϑ hochdimensional ist, leiten wir für jeden Parameter ϑ_i einzeln ab um das Resultat für ϑ_i zu erhalten
  - ⋆ Falls Funktion nicht differenzierbar muss man anders vorgehen

7.3 Momentenschätzer (MM)

- Seien $X_1, \dots, X_n$ i.i.d. Für eine Funktion $h : \Theta \to \mathbb{R}^d$ wollen wir die d Grössen $h_j(\vartheta), j = 1, \dots, d$ schätzen. Um MM $h(\vartheta)$ zu bestimmen:

1) Für $j = 1, \dots, d$ berechnen wir in jedem Modell P_ϑ das j-te Moment $m_j(\vartheta) = E_\vartheta[X^j]$ als Funktion von ϑ.
2) Für $j = 1, \dots, d$ definieren wir das j-Te *empirische Mittel* als $\widetilde{m}_j(x_1, \dots, x_n) := \frac{1}{n}\sum_{i=1}^n x_i^j$
3) Betrachten System von d Gleichungen $\widetilde{m}_j(x_1, \dots, x_n) = m_j(\vartheta), j = 1, \dots, d$ und lösen Gleichungssystem für d Unbekannten $h_1(\vartheta), \dots h_d(\vartheta)$.

   - ⋆ Falls eindeutige Lösung existiert, nennen wir diese $t_{MM}(x_1, \dots, x_n)$, $t_{MM} : \mathbb{R}^n \to \mathbb{R}^d$.

7.4 Verteilungsaussagen

- Verteilungsaussagen über Schätzer sind nicht einfach.
- Falls Schätzer von der Form $\sum_{i=1}^n Y_i$, wobei Y_i i.i.d., kann man ZGS benutzen. Einfaches Bsp: Falls $T = \frac{1}{n}\sum_{i=1}^n Y_i$ folgt für grosse n, dass $\sum_{i=1}^n Y_i \overset{\text{approx}}{\sim} \mathcal{N}(nE_\vartheta[Y_i], n\mathrm{Var}_\vartheta[Y_i])$
- Seien $X_1, \dots, X_n$ i.i.d. $\sim \mathcal{N}(\mu, \sigma^2)$ Dann gilt:

1) $\overline{X}_n \sim \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$, und $\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathcal{N}(0,1)$
2) $\frac{n-1}{\sigma^2}S^2 = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \overline{X}_n)^2 \sim \mathcal{X}_{n-1}^2$
3) $\overline{X}_n$ und S^2 sind unabhängig
4) $\frac{\overline{X}_n - \mu}{\frac{S}{\sqrt{n}}} = \frac{\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}}{\frac{S}{\sigma}} = \frac{\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{1}{n-1}\frac{n-1}{\sigma^2}S^2}} \sim t_{n-1}$

7.4.1 $\mathcal{X}^2$-Verteilung mit n Freiheitsgraden

- Diese gehört zu einer stetigen ZV Y mit DF $f_Y(y) = \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} y^{\frac{n}{2}-1} e^{-\frac{1}{2}y}, \ y \geq 0.$
  - ⋆ $\Gamma(z) := \int_0^\infty t^{z-1}e^{-t}dt \quad - \Gamma(z) := (z-1)!$
- ⋆ $E[\mathcal{X}_n^2] = n$ ⋆ $var[\mathcal{X}_n^2] = 2n$
- Entsteht wie folgt: Sind die ZV $X_1, \dots, X_n$ i.i.d. und $\sim \mathcal{N}(0,1)$, so ist die Summe $Y := \sum_{i=1}^n X_i^2 \sim \mathcal{X}_n^2$
  - ⋆ $E[Y] = n$ ⋆ $Var[Y] = 2n$
- Für $n = 2$ erhält man $\sim Exp(\frac{1}{2})$

7.4.2 t-Verteilung mit n Freiheitsgraden

- Diese gehört zu einer stetigen ZV Z mit DF $f_Z(z) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})}\left(1 + \frac{z^2}{n}\right)^{-\frac{n+1}{2}}, \ z \in \mathbb{R}$
- Entsteht wie folgt: Sind X und Y unabhängig mit $X \sim \mathcal{N}(0,1)$ und $Y \sim \mathcal{X}_n^2$ so ist $Z := \frac{X}{\sqrt{\frac{1}{n}Y}}$ t-verteilt mit n Freiheitsgraden
- Für $n = 1$ ist das eine Cauchy-Verteilung
- Für $n \to \infty$ erhält man $\sim \mathcal{N}(0,1)$

8 Tests

8.1 Grundbegriffe

- Wir haben eine Vermutung wo in Θ der richtige (aber unbekannte) Parameter ϑ liegen könnte, und wollten diese mit Hilfe der Daten überprüfen ("testen"). Das Grundproblem ist, eine Entscheidung zwischen zwei konkurrierenden Modellklassen zu treffen: ⋆ **Hypothese** (H_0) $\Theta_0 \subseteq \Theta$ ⋆ **Alternative** (H_A) $\Theta_A \subseteq \Theta$, wobei $\Theta_0 \cap \Theta_A = \emptyset$
- Hypothese/Alternative heissen:
  - ⋆ **Einfach:** falls Θ_0/Θ_A aus einzelnem Wert ϑ_0/ϑ_A bestehen
  - ⋆ **Zusammengesetzt:** Ansonsten
- **Test:** Eine Entscheidungsregel, die zu Daten $x_1, \dots, x_n$ 0 oder 1 zurück gibt
  - ⋆ **1:** Hypothese H_0 wird abgelehnt
  - ⋆ **0:** Hypothese H_0 wird akzeptiert

  Konkreter: Man hat eine Abbildung $t : \mathbb{R}^n \to \mathbb{R}$ und einen **kritischen Bereich** oder **Verwerfungsbereich** $K \subseteq \mathbb{R}$. Die ZV $T = t(X_1, \dots, X_n)$ heisst **Teststatistik**. Die Entscheidungsregel ist $I_{\{t(x_1,\dots,x_n)\in K\}}$
- **Fehler 1. Art:** Hypothese zu unrecht abgelehnt: Passiert für $\vartheta \in \Theta_0$ und $T \in K$. $P_\vartheta[T \in K]$ für $\vartheta \in \Theta_0$ ist die WS für Fehler 1. Art.
- **Fehler 2. Art:** Hypothese zu unrecht nicht verworfen: Passiert für $\vartheta \in \Theta_A$ und $T \notin K$. $P_\vartheta[T \notin K] = 1 - P_\vartheta[T \in K]$ für $\vartheta \in \Theta_A$ ist die WS für Fehler 2. Art.
- **Signifikanzniveau:** Grundsätzlich möchte man $\vartheta \mapsto P_\vartheta[T \in K]$ auf Θ_0 möglichst klein und auf Θ_A möglichst gross haben. Da diese meist stetig ist und Θ_0 und Θ_A meist nebeneinander liegen ist ein gleichzeitiges min-/maximieren von Θ_0/Θ_A nicht möglich. Deshalb gibt es folgendes zweistufiges Vorgehen:

1) Wähle **Signifikanzniveau** $\alpha \in (0,1)$ und sorge für $\sup_{\vartheta\in\Theta_0} P_\vartheta[T \in K] \leq \alpha$
   - ∗ D.h. man kontrolliere WS für Fehler 1. Art durch α
   - ∗ α wird kleiner $\implies$ WS für Fehler 1. Art wird kleiner $\implies$ Verwerfungsbereich wird kleiner $\implies$ Fehler 2. Art wird grösser $\implies$ Macht des Tests wird grösser
   - ∗ $1 - \alpha$ ist die Wahrscheinlichkeit, dass H_0 unter P_{ϑ_0} beibehalten wird
2) Versuche **Macht** des Tests $\beta : \Theta_A \to [0,1], \beta(\vartheta) := P_\vartheta[T \in K]$ möglichst gross zu bekommen, soweit möglich
   - ∗ D.h. WS für Fehler 2. Art möglich klein

- ⋆ Dieses asymmetrische Vorgehen macht es schwieriger die Hypothese zu verwerfen als beizubehalten. Ein seriöser Test wird deshalb immer die Negation der gewünschten Aussage benutzen.

8.2 Konstruktion von Tests

8.2.1 Likelihood-Quotient (LQ)

Sein $L(x_1, \dots, x_n)$ die Likelihood-Funktion. Für $\vartheta_0 \in \Theta_0$ und $\vartheta_A \in \Theta_A$ betrachten wir den LQ: $R(x_1, \dots, x_n; \vartheta_0, \vartheta_A) := \frac{L(x_1,\dots,x_n;\vartheta_A)}{L(x_1,\dots,x_n;\vartheta_0)}$. Ist der LQ gross bedeutet dies, dass die Beobachtungen $x_1, \dots, x_n$ als Resultate im Modell P_{ϑ_A} deutlich wahrscheinlicher sind als im Modell P_{ϑ_0}. Es liegt deshalb nahe $T := R(X_1, \dots, X_n; \vartheta_0, \vartheta_A)$ und $K := (c, \infty)$ zu wählen, wenn man ϑ_0 gegen ϑ_A testen will

- D.h. R gross $\implies$ wir verwerfen H_0
- Häufig nimmt man nicht den ganzen Quotienten als Test, sondern nur den Teil, welcher R gross mach
- Sind Hypothese und Alternative beide einfach, so ist dieser Test optimal.

8.2.2 Neyman-Pearson-Lemma

Seien $\Theta_0 = \{\vartheta_0\}$ und $\Theta_A = \{\vartheta_A\}$. Wie oben sei $T := R(X_1, \dots, X_n; \vartheta_0, \vartheta_A)$ und $K = (c, \infty)$ sowie $\alpha^* := P_{\vartheta_0}[T \in K] = P_{\vartheta_0}[T > c]$. Der LQ-Test mit T und K ist dann im folgenden Sinn optimal: Jeder andere Test mit Signifikanzniveau $\alpha \leq \alpha^*$ hat kleiner Macht, bzw. eine grössere WS für Fehler 2. Art.

- **Formaler:** Ist (T', K') einer anderer Test mit $P_{\vartheta_0}[T' \in K'] \leq \alpha^* = P_{\vartheta_0}[T \in K]$, so gilt auch $P_{\vartheta_A}[P' \in K'] \leq P_{\vartheta_A}[T \in K]$

8.2.3 Test bei zusammengesetzten ϑ_0/ϑ_A

Bei zusammengesetzten ϑ_0/ϑ_A betrachtet man den **verallgemeinterten LQ** $R(x_1, \dots, x_n) := \frac{\sup_{\vartheta\in\Theta_A} L(x_1,\dots,x_n;\vartheta)}{\sup_{\vartheta\in\Theta_0} L(x_1,\dots,x_n;\vartheta)}$ oder auch $\widetilde{R}(x_1, \dots, x_n) := \frac{\sup_{\vartheta\in\Theta_A\cup\Theta_0} L(x_1,\dots,x_n;\vartheta)}{\sup_{\vartheta\in\Theta_0} L(x_1,\dots,x_n;\vartheta)}$. Und wählt als Teststatistik $T_0 := R(X_1, \dots, X_n)$ bzw. $\widetilde{T} := \widetilde{R}(X_1, \dots, X_n)$ mit $K_0 = (c_0, \infty)$. c_0 bzw. K_0

wählt man so, dass das gewünschte Signifikanzniveau α eingehalten wird.

8.3 Beispiele

8.3.1 z-Test (Test für EW bei bekannter Var)

Sind $X_1,\dots,X_n$ i.i.d. $\sim\mathcal{N}(\mu,\sigma^2)$ unter P_ϑ mit bekannter Var σ^2, und wir wollen die Hypothese $H_0:\vartheta=\vartheta_0$ testen. Mögliche Alternativen H_A sind $\vartheta>\vartheta_0$ oder $\vartheta<\vartheta_0$ (**einseitig**) oder $\vartheta\neq\vartheta_0$ (**zweiseitig**). Die Teststatistik ist $T:=\frac{\overline{X}_n-\vartheta_0}{\frac{\sigma}{\sqrt{n}}}\sim\mathcal{N}(0,1)$ unter P_{ϑ_0}. Der kritische Bereich K ist von der Form $(c_>,\infty)$, $(-\infty,c_<)$, bzw. $(-\infty,-c_{\neq})\cup(c_{\neq},+\infty)$. Die Konstanten $c_>,c_<,c_{\neq}$ bestimmt man zum gewählten Niveau mit Hilfe der Verteilung von T unter P_{ϑ_0}. Zum Beispiel liefert die Bedingung $\alpha=P_{\vartheta_0}[T\in K_>]=P_{\vartheta_0}[T>c_>]=1-P_{\vartheta_0}[T\le c_>]=1-\Phi(c_>)$, dass $c_>=\Phi^{-1}(1-\alpha)=:z_{1-\alpha}$ das sogenannte $(1-\alpha)$-**Quantil** der $\mathcal{N}(0,1)$ Verteilung sein muss; Für $\vartheta>\vartheta_0$ verwirft man H_0, falls $\overline{X}_n>\vartheta_0+z_{1-\alpha}\frac{\sigma}{\sqrt{n}}$ ist. Analog ist $c_<=z_\alpha=-z_{1-\alpha}$ und $c_{\neq}=z_{1-\frac{\alpha}{2}}$, wobei wir die Symmetrie der $\mathcal{N}(0,1)$ Verteilung ausnutzen; Es gilt nämlich $\alpha=P_{\vartheta_0}[t<c_<]=P_{\vartheta_0}[T>-c_<]$ für $-c_<=z_{1-\alpha}$ und $\alpha=P_{\vartheta_0}[T\in K_{\neq}]=P_{\vartheta_0}[T<-c_{\neq}]+P_{\vartheta_0}[T>c_{\neq}]=\Phi(-c_{\neq})+1-\Phi(c_{\neq})=2-2\Phi(c_{\neq}))$

8.3.2 t-Test (Test für EW bei unbekannter Var)

Sind $X_1,\dots,X_n$ i.i.d. $\sim\mathcal{N}(\mu,\sigma^2)$ unter $P_{\vec{\vartheta}}$ wobei $\vec{\vartheta}=(\mu,\sigma^2)$ und die Var σ^2 ist unbekannt. Wir wollen die Hypothese $\mu=\mu_0$ testen. Explizit: $\Theta_0=\{\mu_0\}\times(0,\infty)=\{\vec{\vartheta}=(\mu,\sigma^2)\mid\mu=\mu_0\}$. Die Teststatistik ist $T:=\frac{\overline{X}_n-\mu_0}{\frac{S}{\sqrt{n}}}\sim t_{n-1}$ unter P_{ϑ_0}. Wir ersetzen unbekannte Var durch Schätzer $S^2=\frac{1}{n-1}\sum_{i=1}^n(X_i-\overline{X}_n)^2$ für σ^2. Der Kritische Bereich hat eine der drei Formen aus dem letzten Beispiel; die kritischen Werte hier sind $c_>=t_{n-1,1-\alpha}$, bzw. $c_<=t_{n-1,\alpha}=-t_{n-1,1-\alpha}$, bzw $c_{\neq}=t_{n-1,1-\frac{\alpha}{2}}$. Hier bezeichnen wir mit $t_{m,\gamma}$ das γ-**Quantil** einer t_m-Verteilung, d.h. denjenigen Wert $t_{m,\gamma}$, für den gilt $P[X\le t_{m,\gamma}]=\gamma$ für $X\sim t_m$. Diese Werte findet man in Tabellen.

8.3.3 Gepaarter Zweistichproben-Test bei $\mathcal{N}$

Sind $X_1,\dots,X_n$ i.i.d. $\sim\mathcal{N}(\mu_X,\sigma^2)$ und $Y_1,\dots,Y_m$ i.i.d $\sim\mathcal{N}(\mu_Y,\sigma^2)$ unter P_ϑ; insbesondere ist $m=n$ und die Varianz σ^2 bei beiden Stichproben dieselbe. In dieser Situation kann man Tests über den Vergleich von μ_X und μ_Y auf den Fall nur einer Strichprobe zurückführen; die Differenzen $Z_i=X_i-Y_i$ sind nämlich under P_ϑ i.i.d. $\sim\mathcal{N}(\mu_X-\mu_Y,2\sigma^2)$. Damit kann man t-Test/z-Test in leicht angepasster Form benutzen.

- Es reicht zu fordern, dass Z_i i.i.d. $\sim\mathcal{N}(\mu_X-\mu_Y,2(1-\varrho)\sigma^2)$ mit bekannter Korrelation $\varrho\in(-1,+1)$
- Der Fall $\varrho=0$ entspricht Unabhängigkeit

8.3.4 Ungepaarter Zweistichproben-Test bei $\mathcal{N}$

Sind $X_1,\dots,X_n$ i.i.d. $\sim\mathcal{N}(\mu_X,\sigma^2)$ und $Y_1,\dots,Y_m$ i.i.d. $\sim\mathcal{N}(\mu_Y,\sigma^2)$ under P_ϑ, wobei die Var in beiden Fällen dieselbe ist, aber $m\neq n$ sein kann. Will man einen Vergleich über μ_X und μ_Y testen, so kann man nicht mehr paarweise Differenzen bilden. Diesen Test muss man auch benutzen, falls zufällig $m=n$ ist, aber die Daten nicht natürlich gepaart sind. Wie nehmen immer noch an, dass $X_1,\dots,X_n$ und $Y_1,\dots,Y_m$ unabhängig sind.

- **Ungepaarter Zweistrichproblen-z-Test:** Ist σ^2 bekannt, so ist $T:=\frac{(\overline{X}_n-\overline{Y}_m)-(\mu_X-\mu_Y)}{\sigma\sqrt{\frac{1}{n}+\frac{1}{m}}}\sim\mathcal{N}(0,1)$ under jedem P_ϑ. $\mu_X-\mu_Y$ ist durch Hypothese H_0 bekannt. Die kritischen Werte für den Verwerfungsbereich sind, wie oben, geeignete Quantile der $\mathcal{N}(0,1)$-Verteilung.
- **Ungepaarter Zweistichproben-t-Test:** Ist σ^2 unbekannt, brauchen wir: $S_X^2:=\frac{1}{n-1}\sum_{i=1}^n(X_i-\overline{X}_n)^2$, $S_Y^2:=\frac{1}{m-1}\sum_{j=1}^m(Y_j-\overline{Y}_n)^2$. Mit $S^2:=\frac{1}{m+n-2}((n-1)S_X^2+(m-1)S_Y^2)=\frac{1}{m+n-2}(\sum_{i=1}^n(X_i-\overline{X}_n)^2+\sum_{j=1}^m(Y_j-\overline{Y}_m)^2)$ Ist dann die Teststatistik $T:=\frac{(\overline{X}_n-\overline{Y}_m)-(\mu_X-\mu_Y)}{S\sqrt{\frac{1}{n}+\frac{1}{m}}}\sim t_{n+m-2}$ unter jedem P_ϑ. Der Rest geht analog wie oben.

8.4 Der P-Wert

- Sei $X_1,\dots,X_n$ eine Stichprobe. Wir wollen eine Hypothese H_0 gegen eine Alternative H_A testen. Bei einem Test zum Niveau α wird die Hypothese abgelehnt, wenn $T(\omega)=t(x_1,\dots,x_n)$ einen Wert in K ergibt, und wir haben $P_{\vartheta_0}[T\in K]\le\alpha$. Alternativ kann man auch den sogenannten **p-Wert** berechnen. Sei $\vec{x}=(x_1,\dots,x_n)$ der Vektor unserer Daten und $\vec{X}=(X_1,\dots,X_n)$ unsere Stichprobe. Wir brauchen zuerst für jedes Niveau $\alpha\in[0,1]$ einen Test $\phi_\alpha(\vec{x})=I_{\{t(\vec{x})\in K_\alpha\}}$ mit $T=t(\vec{X})$ und kritischen Bereich K_α mit $P_\vartheta[T\in K_\alpha]\le\alpha\ \forall\vartheta\in\Theta_0$. Zudem muss gelten, dass diese Tests kompatibel sind in dem Sinn, dass aus $\alpha'<\alpha$ immer folgt, dass $\phi_{\alpha'}\le\phi_\alpha$, d.h. falls man die Hypothese auf einem Niveau α nicht verwirft, so kann man sie auf einem kleineren Niveau α' auch nicht verwerfen. Dann definieren wir $\pi(\vec{x}):=\inf\{\alpha\in[0,1]\mid\phi_\alpha(\vec{x})=1\}$, und der p-Wert ist dann gegeben durch; p-Wert $=\pi(\vec{X})$. Wie die Stichprobe ist der p-Wert eine ZV weil er von den Daten abhängt.
- Ist der Test zweiseitig betrachtet man $|T|\le t$ resp. $|T|\ge t$
- Ist der Test einseitig betrachtet man $T\le t$ resp. $T\ge t$
- t ist dabei der realisierte Wert der Stichprobe

8.5 Test Rezept

1) Wahl des Modells
2) Formulierung von Hypothese und Alternativ
3) Bestimmung der Teststatistik T und der Form des kritischen Bereichs K; das kann aus einer Herleitung via verallgemeinerten LQ-Test oder direkt aus einem Statistikbuch stammen.
4) Festlegung des Niveaus α liefert die Grenze für den kritische Bereich K; dazu braucht man die Verteilung von T unter P_ϑ für alle $\vartheta\in\Theta_0$ (exakt oder approximativ).
5) Berechnen der Teststatistik $T(\omega)$ aus den Daten; ist $T(\omega)\in K$, so wird die Hypothese abgelehnt, andernfalls wird die Hypothese nicht verworfen.

5') Berechnen von Teststatistik $T(\omega)$ und entsprechendem realisiertem p-Wert(ω) aus den Daten; ist letzterer $\le\alpha$, so wird die Hypothese abgelehnt, andernfalls nicht.

9 Konfidenzbereiche

Wir suchen aus einer Familie $(P_\vartheta)_{\vartheta\in\Theta}$ von Modellen eines, das zu unseren Daten $x_1,\dots,x_n$ passt. Ein Schätzer für ϑ gibt uns dabei einen einzelnen zufälligen möglichen Parameterwert. Weil es schwierig ist, mit diesem einen Wert dem richtigen Parameter zu treffen, suchen wir nun stattdessen eine Teilmenge des Parameterbereichs, die hoffentlich den wahren Parameter enthält.

- **Konfidenzbereiche (KB):** für ϑ zu Daten $x_1,\dots,x_n$ ist eine Menge $C(x_1,\dots,x_n)\subseteq\Theta$; in den meisten Fällen ist das ein Intervall, dessen Endpunkte von $x_1,\dots,x_n$ abhängen. Ersetzen wir die Daten $x_1,\dots,x_n$ durch die sie generierende ZV so ist $\tilde{C}:=C(X_1,\dots,X_n)$ also eine zufällige Teilmenge von Θ, mit Realisierung $\tilde{C}(\omega)=C(X_1(\omega),\dots,X_n(\omega))$ bei einem festen ω. Ein solches C heisst KB zum Niveau $1-\alpha$, falls gilt: $P_\vartheta[C(X_1,\dots,X_n)\ni\vartheta]\ge1-\alpha\quad\forall\vartheta\in\Theta$, d.h. in jedem Modell erwischt man den Parameter mit grosser WS.
- **KB für σ^2** : zum Niveau $1-\alpha$
  - ⋆ μ **unbekannt:** $$C(X_1,\dots,X_n)=\left[\frac{(n-1)S^2}{\chi^2_{n-1,1-\frac{\alpha}{2}}},\frac{(n-1)S^2}{\chi^2_{n-1,\frac{\alpha}{2}}}\right]$$
  - ⋆ μ **bekannt:** TODO: to be found
- **KB für μ:** zum Niveau $1-\alpha$
  - ⋆ σ **unbekannt:** $C(X_1,\dots,X_n)=\left[\overline{X}_n-t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}},\overline{X}_n+t_{n-1,1-\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right]$
  - ⋆ σ **bekannt:** $C(X_1,\dots,X_n)=\left[\overline{X}_n-z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}},\overline{X}_n+z_{1-\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$

9.1 Konfidenzbereiche - Test - Dualität

- Sei zuerst $C(X_1,\dots,X_n)$ ein KB für ϑ zum Niveau $1-\alpha$. Um die Hypothese $H_0:\vartheta=\vartheta_0$ zu testen, definieren wir einen Test durch $I_{\{\vartheta_0\notin C(X_1,\dots,X_n)\}}$, d.h. wie lehnen H_0 genau dann ab, wenn ϑ_0 nicht in $C(X_1,\dots,X_n)$ liegt. Dann ist wegen $\Theta_0=\{\vartheta_0\}$ für jedes $\vartheta\in\Theta_0$: $P_\vartheta[\vartheta_0\notin C(X_1,\dots,X_n)]=1-P_{\vartheta_0}[C(X_1,\dots,X_n)\ni\vartheta_0]\le\alpha$, so dass der Test gerade α als Signifikanzniveau hat. Aus einem KB für ϑ erhalten wir also eine ganze Familie von Tests, nämlich einen für jede einfache Hypothese $\Theta_0=\{\vartheta_0\}$ mit einen $\vartheta_0\in\Theta$.
- Sei nun umgekehrt für jede einfache Hypothese $\Theta_0=\{\vartheta_0\}$ eine Test zum Niveau α gegeben; für jedes ϑ_0 haben wir also einen kritischen Bereich K_{ϑ_0}, so dass $H_0:\vartheta=\vartheta_0$ genau dann abgelehnt wird, wenn $(X_1,\dots,X_n)\in K_{\vartheta_0}$ ist. Zudem gilt wegen Niveau α: $P_{\vartheta_0}[(X_1,\dots,X_n)\in K_{\vartheta_0}]\le\alpha\ \forall\vartheta_0 in\Theta$. Nun definieren wir eine Teilmenge $C(X_1,\dots,X_n)$ von Θ durch die Bedingung $\vartheta\in C(X_1,\dots,X_n):\iff(X_1,\dots,X_n)\notin K_\vartheta$. Dann ist das ein KB zum Niveau $1-\alpha$, denn für jedes $\vartheta\in\Theta$ gilt $P_\vartheta[C(X_1,\dots,X_n)\ni\vartheta]=P_\vartheta[(X_1,\dots,X_n)\notin K_\vartheta]=1-P_\vartheta[(X_1,\dots,X_n)\in K_\vartheta]\ge1-\alpha$.

Part III

Appendix

10 Sonstiges

10.1 Erwartungswert

- **Gegeben:** EW von $f(x,y) = \begin{cases} \frac{1}{2} & 0 \le x \le y \le 2 \\ 0 & \text{else} \end{cases}$
- **Gesucht:** $E[XY] = \int_0^2 \int_0^y \frac{1}{2} xy \, \mathrm{d}x\mathrm{d}y = \int_0^2 \frac{1}{4} x^2 y |_0^y = \int_0^2 \frac{1}{4} y^3 \, \mathrm{d}y = \frac{1}{16} y^4 |_0^2 = 1$

10.2 Umformungen

- $P[|X-\mu| \le 1] = P[-1 \le X - \mu \le 1] = P[-1+\mu \le X \le 1+\mu] = P[X \le 1+\mu] - P[X \le -1+\mu]$
- $\frac{1}{n}\sum_{i=1}^n (X_i - \overline{X}_n)^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - (\overline{X}_n)^2$

$$E[Z_1^4] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^4 e^{-x^2/2}\,dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^3 x e^{-x^2/2}\,dx$$
$$= \frac{1}{\sqrt{2\pi}} \left(-x^3 e^{-x^2/2}\right)\Big|_{-\infty}^{\infty} + \frac{3}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2}\,dx = 3 \qquad (1P)$$

-

10.3 Likelihood Quotient

- ⋆ **Gegeben:**

$$L(x_1,\dots,x_n;\vartheta) = \Pi_{i=1}^n p_X(x_i;\vartheta) = \vartheta^{\sum_{i=1}^n x_i}(1-\vartheta)^{n-\sum_{i=1}^n x_i}$$

⋆ **Gesucht:**

$$R(x_1,\dots,x_n;\vartheta_0,\vartheta_A) = \frac{L(x_1,\dots,x_n;\vartheta_A}{L(x_1,\dots,x_n;\vartheta_0} = \left(\frac{\vartheta_A}{\vartheta_0}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-\vartheta_A}{1-\vartheta_0}\right)^{n-\sum_{i=1}^n x_i} = \left(\frac{\vartheta_A(1-\vartheta_0)}{\vartheta_0(1-\vartheta_A)}\right)^{\sum_{i=1}^n x_i} \left(\frac{1-\vartheta_A}{1-\vartheta_0}\right)^n$$

∗ Sei $\vartheta_0 = \frac{1}{2}, \vartheta_A \ge \frac{1}{2}$, so ist R gross, falls $\sum_{i=1}^n x_i$ gross ist $\implies T := \sum_{i=1}^n X_i = S_n$

- ⋆ **Gegeben:**

$$L(x_1,\dots,x_n;\vartheta) = \Pi_{i=1}^n f_X(x_i;\vartheta) = (2\pi\sigma^2)^{\frac{-n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\vartheta)^2\right)$$

⋆ **Gesucht:**

$$R(x_1,\dots,x_n;\vartheta_0,\vartheta_A) = \frac{L(x_1,\dots,x_n;\vartheta_A}{L(x_1,\dots,x_n;\vartheta_0} = \exp\left(-\frac{1}{2\sigma^2}\left(\sum_{i=1}^n (x_i-\vartheta_A)^2 - \sum_{i=1}^n (x_i-\vartheta_0)^2\right)\right) = \text{const.}(\sigma,\vartheta_0,\vartheta_A)\exp\left(\frac{1}{2\sigma^2}(\vartheta_A-\vartheta_0)\sum_{i=1}^n x_i\right)$$

⋆ Für ϑ_0, ϑ_A beliebig, R ist gross wenn $(\vartheta_A - \vartheta_0)\sum_{i=1}^n x_i$ gross ist $\implies T := \sum_{i=1}^n X_i$

10.4 Likelihood Schätzer

- **Gesucht:** Maximum-Likelihood Schätzer für λ für $X_1,\dots,X_n$, i.i.d. $X_i \sim \text{Poisson}(\lambda)$ ist $T = \frac{1}{n}\sum_{i=1}^n X_i$
- **Gesucht:** Maximum-Likelihood Schätzer für p für $X_1,\dots,X_n$, i.i.d. $X_i \sim \text{Be}(p)$ ist $T = \frac{1}{n}\sum_{i=1}^n X_i$

10.5 Mehrere ZV

- **Gegeben:** $Z \sim \text{Exp}(\lambda), Z \sim \text{Exp}(\mu)$
- **Gesucht:** $P[Z > Y] = \int\int_{\{(y,z)\in\mathbb{R}^2 | z > y\}} \lambda e^{-\lambda y}\mu e^{-\mu z}\mathrm{d}z\mathrm{d}y = \int_0^\infty \lambda e^{\lambda y}\int_y^\infty \mu e^{-\mu z}\mathrm{d}z\mathrm{d}y = \int_0^\infty \lambda e^{-\lambda y}e^{-\mu y}\mathrm{d}y = \frac{\lambda}{\lambda+\mu}$

10.6 Proof Gedächnislos

Beweise: $Y \sim Exp(p) \iff P[Y > s] = P[Y > s+t \mid Y > t], \quad s,t \in \mathbb{R}_+$:

⇒: Wir nehmen an, dass $Y \sim Exp(p)$. Dann

$$P[Y > s+t \,|\, Y > t] = \frac{e^{-p(s+t)}}{e^{-pt}} = e^{-ps} = P[Y > s].$$

⇐: Wir nehmen an, dass Y gedächtnislos ist. Dann, für $s,t \ge 0$

$$P[Y > s] = P[Y > s+t \,|\, Y > t] = \frac{P[Y > s+t, Y > t]}{P[Y > t]} = \frac{P[Y > s+t]}{P[Y > t]}$$

so

$$P[Y > s+t] = P[Y > s]\,P[Y > t].$$

Sei $g(s) := P[Y > s]$, dann

$$g(s+t) = g(s)\,g(t). \tag{1}$$

g ist stetig und monoton auf $\mathbb{R}_+$, so g ist differenzierbar auf $\mathbb{R}_+$, und, für $\delta > 0$

$$\lim_{h\to 0}\frac{g(s+\delta+h)-g(s+\delta)}{h} \overset{(1)}{=} g(s)\lim_{h\to 0}\frac{g(\delta+h)-g(\delta)}{h} = g(s)\,c_\delta,$$

mit $c_\delta < 0$, weil g strikt antiton ist. Wir nützen (1) noch einmal und haben

$$g'(s+\delta) = g(s)\,c_\delta = g(s+\delta)\,\frac{c_\delta}{g(\delta)},$$

das zusammen mit $g(0) = 1$ bringt

$$g(s+\delta) = e^{\frac{c_\delta}{g(\delta)}(s+\delta)} = e^{-\lambda(s+\delta)}, \quad \lambda > 0,$$

so dass (nehmen $\delta \to 0$) Y kann nur Exponentialverteilung haben.

10.7 Hypergeometrische Verteilung

In einer Urne seien n Gegenstände, davon r vom Typ 1 und $n-r$ vom Typ 2. Man zieht ohne Zurücklegen m der Gegenstände; die Zufallsvariable X beschreibe die Anzahl der Gegenstände vom Typ 1 in dieser Stichprobe vom Umfang m. Dann hat X eine *hypergeometrische Verteilung* mit Parametern $n \in \mathbb{N}$ und $m, r \in \{1,\dots,n\}$. Der Wertebereich von X ist $\mathcal{W}(X) = \{0,1,\dots,\min(m,r)\}$, und X hat die Gewichtsfunktion

$$p_X(k) = \frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}} \qquad \text{für } k \in \mathcal{W}(X).$$

Zur Herleitung der Gewichtsfunktion nehmen wir an, dass alle der insgesamt $\binom{n}{m}$ möglichen Stichproben gleich wahrscheinlich sind. $X = k$ heisst, dass wir aus den r Gegenständen vom Typ 1 gerade k und aus den restlichen $n-r$ die noch fehlenden $m-k$ erwischen müssen; die Anzahl der Stichproben mit $X = k$ ist also $\binom{r}{k}\binom{n-r}{m-k}$. Daraus ergibt sich $p_X(k) = P[X = k]$ wie in der obigen Formel.

10.8 Momenteschätzer

- Für ein Experiment mit $n = 5$ unf $f_T(t) = C^2 e^{-Ct}1_{[0\le t]}$, bestimme den Momenteschätzer:

Um den Momentenschätzer $\hat{C}^{MOM}$ zu bestimmen, setzen wir das empirische Mittel gleich dem theoretischen Erwartungswert $\mathbb{E}[T]$. Der empirische Mittelwert ist $\frac{1}{5}\sum_{i=1}^5 T_i$, mit realisiertem Wert

$$\frac{1}{5}\sum_{i=1}^5 T_i(\omega) = \frac{5480}{5} = 1096.$$

Weiter berechnet sich der Erwartungswert mit Hilfe der Substitution $u := Ct$, $du = Cdt$ und dem Hinweis als

$$\mathbb{E}[T] = \int_{\mathbb{R}} t f_T(t)\,dt = \int_0^\infty tC^2 t\exp(-Ct)\,dt = \int_0^\infty (Ct)^2\exp(-Ct)\,dt \qquad (0.5P)$$
$$\overset{\text{Sub}}{=} \frac{1}{C}\int_0^\infty u^2\exp(-u)\,du \qquad (0.5P)$$
$$\overset{\text{Hinweis}}{=} \frac{1}{C}\left(-u^2\exp(-u) - 2u\exp(-u) - 2\exp(-u)\right)\Big|_0^\infty.$$

Da $\lim_{u\to\infty} u^p\exp(-u) = 0$ für jedes $p > 0$ (0.5P), ergibt sich durch Einsetzen der Grenzen $\mathbb{E}[T] = \frac{2}{C}$ (0.5P). Somit erhalten wir $\hat{C}^{MOM}$ durch

$$\frac{1}{5}\sum_{i=1}^5 T_i = \frac{2}{C} \qquad (1P) \qquad \Longrightarrow \qquad \hat{C}^{MOM} = \frac{10}{\sum_{i=1}^5 T_i} = \hat{C}^{MLE}.$$

Der realisierte Schätzwert ist also wiederum

$$\hat{C}^{MOM}(\omega) = \hat{C}^{MLE}(\omega) = \frac{1}{548}(\approx 0.00182). \qquad (1P)$$

- More MM:

Wir wählen ϑ_1 so, dass das s-te Moment $m_s(\vartheta) = E_\vartheta[X^s]$ von X unter P_ϑ und das s-te empirische Mittel $\tilde{m}_s(x_1,\dots,x_n) = \frac{1}{n}\sum_{i=1}^n x_i^s$ von $x_1,\dots,x_n$ übereinstimmen. Für $s = 1$ gilt

$$m_1(\vartheta_1) = \tilde{m}_1(x_1,\dots,x_n) \Leftrightarrow \frac{(\vartheta_1-1)\vartheta_2}{\vartheta_1-2} = \tilde{m}_1(x_1,\dots,x_n)$$
$$\Leftrightarrow \vartheta_1 = \frac{\vartheta_2 - 2\tilde{m}_1(x_1,\dots,x_n)}{\vartheta_2 - \tilde{m}_1(x_1,\dots,x_n)} = 1 - \frac{\tilde{m}_1(x_1,\dots,x_n)}{\vartheta_2 - \tilde{m}_1(x_1,\dots,x_n)}$$
$$\Leftrightarrow \vartheta_1 = 1 - \frac{\frac{1}{n}\sum_{i=1}^n x_i}{\vartheta_2 - \frac{1}{n}\sum_{i=1}^n x_i}.$$

Dies führt zum Momentenschätzer

$$T_{\text{MM}} = 1 - \frac{\frac{1}{n}\sum_{i=1}^n X_i}{\vartheta_2 - \frac{1}{n}\sum_{i=1}^n X_i} = 1 - \frac{\bar{X}_n}{\vartheta_2 - \bar{X}_n}.$$

10.9 Test Nicht Normalverteilt

Zur Belegung der Aussage wählt man als Nullhypothese

$$H_0 : p = 0.4 \qquad \text{(0.5P)}$$

und als Alternativhypothese die zu belegende Aussage, also

$$H_A : p < 0.4. \qquad \text{(0.5P)}$$

Der Verwerfungsbereich ist $\{x : x \leq c\}$, wobei $c \in \mathbb{N}_0$ so gross wie möglich mit $P_{0.4}[X_{100} \leq c] \leq 0.05$ ist. (0.5P) Zur Berechnung von $P_{0.4}[X_{100} \leq c]$ verwendet man die Normalapproximation. Unter Berücksichtigung der Kontinuitätskorrektur ergibt sich

$$\begin{aligned} P_{0.4}[X_{100} \leq c] &= P_{0.4}\left[\frac{X_{100} - 40}{\sqrt{0.4 \times 0.6 \times 100}} \leq \frac{c - 40}{\sqrt{0.4 \times 0.6 \times 100}}\right] \\ &\approx \Phi\left(\frac{c - 40 + 0.5}{\sqrt{0.4 \times 0.6 \times 100}}\right). \end{aligned}$$

Somit benötigen wir

$$\Phi\left(\frac{c - 39.5}{\sqrt{0.4 \times 0.6 \times 100}}\right) \leq 0.05,$$

beziehungsweise

$$1 - \Phi\left(\frac{c - 39.5}{\sqrt{0.4 \times 0.6 \times 100}}\right) = \Phi\left(\frac{39.5 - c}{\sqrt{0.4 \times 0.6 \times 100}}\right) \geq 0.95, \qquad \text{(1P)}$$

oder

$$\begin{aligned} c \leq 39.5 - \Phi^{-1}(0.95)\sqrt{0.4 \times 0.6 \times 100} &\overset{\Phi^{-1}(0.95)\approx 1.645}{\approx} 39.5 - 1.645\sqrt{24} \qquad \text{(0.5P)} \\ &\approx 39.5 - 1.6 \cdot 5 = 31.5. \qquad \text{(0.5P)} \end{aligned}$$

Somit ist der maximale approximative Verwerfungsbereich $\{x : x \leq 31\}$.

10.10 MinStuff

a) Seien H_1, H_2 die Lebensdauern der entsprechenden Halbleiter. Nach Voraussetzung sind H_1, H_2 i.i.d. und Exp(λ)-verteilt mit $\lambda = 1/60$. Sei T die Zeit, nach der die Kontrolllampe aufleuchtet; also ist $T = \min\{H_1, H_2\}$ (0.5P). Die Verteilungsfunktion von T ist gegeben durch

$$\begin{aligned} F_T(t) &= P[T \leq t] = P[\min\{H_1, H_2\} \leq t] \\ &= 1 - P[\min\{H_1, H_2\} > t] = 1 - P[H_1 > t, H_2 > t] \qquad \text{(0.5P)} \\ &= 1 - P[H_1 > t]P[H_2 > t] \qquad \text{(0.5P)} \\ &= (1 - \exp(-2\lambda t))1_{[0,\infty)}(t). \qquad \text{(0.5P)} \end{aligned}$$

d.h. T ist wieder exponentialverteilt mit Parameter $2\lambda = 1/30$.

b) Nach einer Zeit $T_i, i = 1, 2, \ldots$, wird das i-te Bauteil ersetzt, wobei ein Bauteil aus zwei Halbleitern besteht. Nach Teil a) sind $T_1, T_2, \ldots$ i.i.d. und Exp(2λ)-verteilt. Sei A das Ereignis, dass innerhalb dreier Jahre (1095 Tage) mehr als 35 Ersatzbauteile benötigt werden. Also ist

$$A = \{S < 1095\}, \qquad \text{(1P)}$$

wobei $S := T_1 + \ldots + T_{36}$. Es gilt $E[S] = 36E[T_1] = 36 \cdot 30 = 1080$ (0.5P) und wegen Unabhängigkeit $\mathrm{Var}(S) = 39\mathrm{Var}(T_1) = 36 \cdot 30^2 = 32400$ (0.5P). Nach dem zentralen Grenzwertsatz folgt somit

$$\begin{aligned} P[A] = P[S < 1095] &= P\left[\frac{S - \mathbb{E}[S]}{\sqrt{\mathrm{Var}(S)}} < \frac{1095 - \mathbb{E}[S]}{\sqrt{\mathrm{Var}(S)}}\right] \\ &= P\left[\frac{S - 1080}{180} < \frac{15}{180}\right] \qquad \text{(0.5P)} \\ &\approx \Phi(1/12) \approx \Phi(0.08) \ \text{(0.5P)} = 0.5319. \qquad \text{(0.5P)} \end{aligned}$$

Wurde (dem Hinweis folgend) angenommen, dass $T_i \sim \mathcal{N}(30, 900)$ $i = 1, \ldots, 36$, so ergibt sich, dass $S \sim \mathcal{N}(1080, 32400)$ und durch direkte Rechnung, dass $P[A] = \Phi(1/12) \approx 0.5319$.

10.11 χ Verteilung

Für $Y = \sum_{k=1}^{\nu} Z_k^2$, $\quad Z_i \sim \mathcal{N}(0,1)$ i.i.d. find $E[Y], Var[Y]$:

Für ganzzahlige Werte von ν kann Y geschrieben werden als $Y = \sum_{k=1}^{\nu} Z_k^2$ mit $Z_1, \ldots, Z_\nu$ i.i.d. $\mathcal{N}(0,1)$.

Daraus folgt einerseits

$$E[Y] = \sum_{k=1}^{\nu} E[Z_k^2] = \nu \cdot 1 = \nu. \qquad \text{(1P)}$$

Andererseits ist wegen der Unabhängigkeit und identischen Verteilung der Z_k

$$E[Y^2] = E[(Z_1^2 + Z_2^2 + \cdots + Z_\nu^2)^2] = \nu E[Z_1^4] + \nu(\nu - 1)(E[Z_1^2])^2, \qquad \text{(0.5P)}$$

oder alternativ

$$\mathrm{Var}[Y] = \mathrm{Var}\left[\sum_{k=1}^{\nu} Z_k^2\right] = \sum_{k=1}^{\nu} \mathrm{Var}\left[Z_1^2\right] = \nu\left(E[Z_1^4] - E[Z_1^2]^2\right). \qquad \text{(0.5P)}$$

$$\begin{aligned} E[Z_1^4] &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} Y x^4 e^{-x^2/2}\, dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} Y x^3 x e^{-x^2/2}\, dx \\ &= \frac{1}{\sqrt{2\pi}} \left(-x^3 e^{-x^2/2}\right)\Big|_{-\infty}^{\infty} + \frac{3}{\sqrt{2\pi}} \int_{-\infty}^{\infty} Y x^2 e^{-x^2/2}\, dx = 3 \qquad \text{(1P)} \end{aligned}$$

erhalten wir schliesslich $E[Y^2] = 3\nu + \nu(\nu - 1) = \nu^2 + 2\nu$, und daraus

$$\mathrm{Var}[Y] = (\nu^2 + 2\nu) - \nu^2 = 2\nu, \qquad \text{(0.5P)}$$

beziehungsweise

$$\mathrm{Var}[Y] = \nu(3 - 1) = 2\nu. \qquad \text{(0.5P)}$$

10.12 Grenzwert Überschreitung

(1.5 Punkte) Wie wahrscheinlich ist es, dass man mit 16 unabhängigen Wasserproben eine Grenzwertüberschreitung nachweisen kann, wenn die wahre Ammoniumkonzentration tatsächlich über dem Grenzwert und zwar bei 205 μgNH$_4$-N/ℓ liegt?
Aus a) folgt: Die Nullhypothese kann verworfen werden, falls der Mittelwert aller Messungen grösser als 204.1 ist,

$$\overline{X} > 204.1.$$

Um die Wahrscheinlichkeit zu berechnen, dass eine Grenzwertüberschreitung nachgewiesen werden kann (H_0 verworfen werden kann), geht man wieder zu einer standardisierten normalverteilten Zufallsvariable über. Mit $\mu_A = 205$ und $\sigma = 10$ erhält man

$$\begin{aligned} P_{\mu_A}[\overline{X} > 204.1] \quad (0.5\text{P}) &= P_{\mu_A}\left[\frac{\overline{X}-\mu_A}{\sigma/\sqrt{n}} > \frac{204.1-\mu_A}{\sigma/\sqrt{n}}\right] \\ &= P_{\mu_A}\left[\frac{\overline{X}-\mu_A}{\sigma/\sqrt{n}} > -0.36\right] \qquad (0.5\text{P}) \\ &= P[Z > -0.36]. \end{aligned}$$

Dies entspricht also der Wahrscheinlichkeit, dass eine normalverteilte Zufallsvariable Z mit Varianz 1,

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

einen Wert grösser als -0.36 annimmt. Diese Wahrscheinlichkeit ist wegen der Symmetrie der Normalverteilung gleich

$$P[Z \leq 0.36] = 0.6406, \qquad (0.5\text{P})$$

wie aus der Tabelle entnommen werden kann. Die gesuchte Wahrscheinlichkeit (die Macht des Tests an der Stelle $\mu = \mu_A = 205$) ist also ungefähr 64%.

10.13 Gemeinsame Dichte und Randdichte

$P \sim \text{Uniform}(0,1), H|P \sim \text{Uniform}(0,P)$, finde: • Gemeinsame Dichte
Nach Aufgabenstellung sind $P \sim \text{U}(0,1)$ und $H|P \sim \text{U}(0,P)$, d.h. die bedingte Dichte von H gegeben P ist durch

$$f_{H|P}(h|p) = \frac{1}{p}1_{(0,p)}(h) \qquad (0.5\text{P})$$

gegeben. Also verschwindet $f_{H|P}$ ausserhalb des Dreiecks

$$D := \{\,(p,h)\,|\,0 < h < p \leq 1\,\}.$$

Für die gemeinsame Dichte von P und H ergibt sich:

$$f_{P,H}(p,h) = f_{H|P}(h|p)f_P(p) \; (0.5\text{P}) = \frac{1}{p}1_{(0,p)}(h)1_{[0,1]}(p). \qquad (0.5\text{P})$$

Der Bereich D, wo $f_{H,P}$ positiv ist, ist in Abbildung 1 skizziert. (1P)

• Randdichte von H

Die Randdichte von H berechnet sich für $0 < h \leq 1$ als Integral über die gemeinsame Dichte, also

$$\begin{aligned} f_H(h) &= \int_{-\infty}^{\infty} f_{P,H}(p,h)\,dp \qquad (0.5\text{P}) \\ &= \int_0^1 \frac{1}{p}1_{(0,p)}(h)\,dp = \int_h^1 \frac{1}{p}\,dp \qquad (0.5\text{P}) \\ &= -\log h. \qquad (0.5\text{P}) \end{aligned}$$

• $E[H]$
Mit partieller Integration folgt

$$\begin{aligned} E[H] &= -\int_0^1 h\log h\,dh \qquad (0.5\text{P}) \\ &= -\frac{h^2}{2}\log h\Big|_0^1 + \int_0^1 \frac{h^2}{2}\frac{1}{h}\,dh = 0 + \int_0^1 \frac{h^2}{2}\frac{1}{h}\,dh \qquad (0.5\text{P}) \\ &= \frac{h^2}{4}\Big|_0^1 = \frac{1}{4}. \qquad (0.5\text{P}) \end{aligned}$$

Andererseits ist $E[P] = 1/2$, also $E[H] = E[P]/2$. (0.5P)
Alternativ: Es gilt $E[H] = E[E[H|P]]$. (0.5P) Laut Aufgabenstellung ist H gegeben P $\text{U}(0,P)$-verteilt, so dass $E[H|P] = \frac{1}{2}P$ (0.5P) ist und damit

$$E[H] = \frac{1}{2}E[P] \; (0.5\text{P}) = \frac{1}{4}. \qquad (0.5\text{P})$$