Ausarbeitung zum Seminarthema „Unabhängigkeit von Zufallsvariablen“

Seminar: Gegenbeispiele in der Wahrscheinlichkeitstheorie

Pascal Beckedorf

13. März 2013

1 Grundlagen

Seien X_i reelle Zufallsvariablen (ZV), definiert auf dem Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$, $B_i \in \mathfrak{B}(\mathbb{R})$ für $i = 1, \ldots, n$.
Dann heißen zwei ZV X_1 und X_2 paarweise unabhängig, wenn

$$P(X_1 \in B_1, X_2 \in B_2) = P(X_1 \in B_1) \cdot P(X_2 \in B_2)$$

gilt.
ZV $X_1, \ldots, X_n$ heißen unabhängig, wenn

$$P(X_1 \in B_1, \ldots, X_n \in B_n) = P(X_1 \in B_1) \cdot \ldots \cdot P(X_n \in B_n) \tag{1}$$

gilt.
Für den Fall, dass die ZV X_i diskret sind, ist (1) äquivalent zu:

$$P(X_1 = x_1, \ldots, X_n = x_n) = P(X_1 = x_1) \cdot \ldots \cdot P(X_n = x_n). \tag{2}$$

Falls die ZV X_i absolut stetig sind mit gemeinsamer Dichte $f_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$ und Randdichten f_{X_i}, ist (1) äquivalent zu:

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdot \ldots \cdot f_{X_n}(x_n) \text{ f.s.}, \tag{3}$$

wobei $x_i \in \mathbb{R}$ für alle $i = 1, \ldots, n$.
Für die paarweise Unabhängigkeit gelten die Formeln (2) und (3) analog mit $n = 2$.

2 Paarweise unabhängige Zufallsvariablen die nicht unabhängig sind

2.1 Definition und Motivation

Im folgenden Beispiel wird deutlich, dass aus der paarweisen Unabhängigkeit der ZV $X_1, \ldots, X_n$ nicht die Unabhängigkeit der Familie von ZV $(X_1, \ldots, X_n)$ folgt.

2.2 Beispiel

Bestehe Ω aus neun Punkten: allen Permutationen (S_3) der Zahlen 1, 2, 3 und den Tripletts (1, 1, 1), (2, 2, 2), (3, 3, 3). Jedes Element von Ω besitze die Wahrscheinlichkeit $\frac{1}{9}$. Seien X_1, X_2 und X_3 drei ZV, wobei X_i der Zahl an i-ter Stelle eines Tripletts entspricht.
Offensichtlich gilt:

$$\begin{aligned} P(X_i = k) &= \frac{1}{3} \qquad && \forall i = 1, 2, 3, \quad k = 1, 2, 3, \\ P(X_i = k, X_j = l) &= \frac{1}{9} && \forall i, j = 1, 2, 3, \quad k, l = 1, 2, 3. \end{aligned}$$

Somit ist (2) für $n = 2$ erfüllt und X_1, X_2, X_3 sind paarweise unabhängig.
Allerdings ist X_3 durch X_1 und X_2 eindeutig bestimmt, also sind die ZV sind nicht unabhängig und (2) ist nicht erfüllt. So gilt z.B.:

$$P(X_1 = 1, X_2 = 2, X_3 = 3) = \frac{1}{9} \neq \frac{1}{27} = P(X_1 = 1) \cdot P(X_2 = 2) \cdot P(X_3 = 3).$$

2.3 Beispiel

Dieses Beispiel erweitert Beispiel 2.2. Dafür seien (X_4, X_5, X_6), (X_7, X_8, X_9), $\ldots$, Tripletts, die wie (X_1, X_2, X_3) aus Beispiel 2.2 aufgebaut sind. Alle Tripletts seien unabhängig voneinander.
Somit erhalten wir eine unendliche Folge X_k, $k \in \mathbb{N}$ von ZV, die paarweise unabhängig sind. Betrachte zum Beweis X_i und X_j mit $i, j \in \mathbb{N}$ beliebig. Sind X_i und X_j im gleichen Triplett enthalten, so folgt ihre Unabhängigkeit aus Beispiel 2.2. Sind X_i und X_j in verschiedenen Tripletts enthalten, so folgt ihre Unabhängigkeit aus der Unabhängigkeit der Tripletts.
Allerdings sind drei oder mehrere beliebig ausgewählte ZV aus X_k im Allgemeinen nicht unabhängig, da in diesem Fall ein Triplett komplett enthalten sein kann und die Abhängigkeit dann aus Beispiel 2.2 folgt.

2.4 Beispiel

Betrachte die Funktion f mit

$$f(x_1, \ldots, x_n) = \begin{cases} (2\pi)^{-n}(1 - \cos(x_1) \cdot \ldots \cdot \cos(x_n)), & \text{falls } (x_1, \ldots, x_n) \in [0, 2\pi]^n, \\ 0, & \text{sonst.} \end{cases}$$

f ist offensichtlich nicht negativ und als Randdichte f_1 von x_1 ergibt sich:

$$\begin{aligned} f_1(x_1) &= \int\limits_{\mathbb{R}^{n-1}} f(x_1, \ldots, x_n) \, \mathrm{d}x_2 \ldots \mathrm{d}x_n \\ &= (2\pi)^{-n} \int\limits_{[0,2\pi]^{n-2}} \int\limits_0^{2\pi} (1 - \cos(x_1) \cdot \cos(x_2) \cdot \ldots \cdot \cos(x_n)) \, \mathrm{d}x_2 \, \mathrm{d}x_3 \ldots \mathrm{d}x_n \\ &= (2\pi)^{-n} \int\limits_{[0,2\pi]^{n-2}} (2\pi - \cos(x_1) \cdot \sin(2\pi) \cdot \ldots \cdot \cos(x_n)) - (0 - 0) \, \mathrm{d}x_3 \ldots \mathrm{d}x_n \\ &= (2\pi)^{-(n-1)} \int\limits_{[0,2\pi]^{n-2}} 1 \, \mathrm{d}x_3 \ldots \mathrm{d}x_n = (2\pi)^{-(n-1)} \int\limits_{[0,2\pi]^{n-3}} \int\limits_0^{2\pi} 1 \, \mathrm{d}x_3 \, \mathrm{d}x_4 \ldots \mathrm{d}x_n \\ &= (2\pi)^{-(n-2)} \int\limits_{[0,2\pi]^{n-3}} 1 \, \mathrm{d}x_4 \ldots \mathrm{d}x_n = \ldots = (2\pi)^{-1}, \end{aligned}$$

für $x_1 \in [0, 2\pi]$ und $f_1(x_1) = 0$, sonst. Die anderen Randdichten ergeben sich analog.

Außerdem ist

$$\int_{\mathbb{R}^n} f(x_1, \dots, x_n) \, \mathrm{d}x_1 \dots \mathrm{d}x_n \overset{s.o.}{=} \int_0^{2\pi} (2\pi)^{-1} \, \mathrm{d}x_n = 1$$

damit f eine Wahrscheinlichkeitsdichte.
Für die gemeinsame Dichtefunktion g_k von $k \in \{2, \dots, n-1\}$ beliebig aus $x_1, \dots, x_n$ ausgewählten ZV gilt:

$$g_k(x_{i_1}, \dots, x_{i_k}) \overset{s.o.}{=} \begin{cases} (2\pi)^{-k}, & \text{falls } (x_1, \dots, x_k) \in [0, 2\pi]^k, \\ 0, & \text{sonst.} \end{cases} = f_{i_1}(x_{i_1}) \cdot \dots \cdot f_{i_k}(x_{i_k}).$$

Allerdings ist

$$f(x_1, \dots, x_n) = (2\pi)^{-n}(1 - \cos(x_1) \cdot \dots \cdot \cos(x_n)) < (2\pi)^{-n} = f_1(x_1) \cdot \dots \cdot f_n(x_n)$$

für alle $(x_1, \dots, x_n) \in [0, pi/2)^n$ und damit ist (3) nicht erfüllt.
Somit ist jede echte Teilmenge von $(x_1, \dots, x_n)$ unabhängig, $(x_1, \dots, x_n)$ selbst allerdings nicht.

3 Unabhängigkeit von X und Y und Unabhängigkeit von X^2 und Y^2

3.1 Definition und Motivation

Es ist bekannt, dass aus der Unabhängigkeit von X und Y die Unabhängigkeit von X^2 und Y^2 folgt (sogar die Unabhängigkeit von $g(X)$ und $h(Y)$ für alle stetige Funktionen g und h). Hier interessiert uns nun, ob auch die Umkehrung gilt.

3.2 Beispiel

Der Zufallsvektor (X, Y) besitze die Dichte

$$f(x, y) = \begin{cases} \frac{1}{4}(1 + xy), & \text{falls } |x| < 1 \text{ und } |y| < 1, \\ 0, & \text{sonst.} \end{cases}$$

Die zugehörige Randdichte $f_1(x)$ von X lässt sich durch Integration leicht bestimmen:

$$f_1(x) = \begin{cases} \frac{1}{2}, & \text{falls } |x| < 1, \\ 0, & \text{sonst.} \end{cases}$$

$f_2(y)$ von Y ergibt sich analog.

Offensichtlich gilt für $0 < x < 1$ und $0 < y < 1$:

$$f(x,y) > \frac{1}{4} = f_1(x) \cdot f_2(y).$$

Somit gilt (3) nicht und X und Y sind nicht unabhängig.
X^2 und Y^2 nehmen offenbar Werte in $(0,1)$ an. Für $x \in (0,1)$ und $y \in (0,1)$ gilt dann:

$$\begin{aligned} P\Big(X^2 < x, Y^2 < y\Big) &= P\Big(-\sqrt{x} < X < \sqrt{x}, -\sqrt{y} < Y < \sqrt{y}\Big) \\ &= \frac{1}{4} \int\limits_{-\sqrt{x}}^{\sqrt{x}} \int\limits_{-\sqrt{y}}^{\sqrt{y}} (1 + uv)\,\mathrm{d}u\,\mathrm{d}v = \frac{1}{2} \int\limits_{-\sqrt{x}}^{\sqrt{x}} \sqrt{y}\,\mathrm{d}v = \sqrt{x}\sqrt{y} \\ &= P\Big(-\sqrt{x} < X < \sqrt{x}\Big) \cdot P\Big(-\sqrt{y} < Y < \sqrt{y}\Big) = P\Big(X^2 < x\Big) \cdot P\Big(Y^2 < y\Big). \end{aligned}$$

Für $x \notin (0,1)$ oder $y \notin (0,1)$ sind beide Seiten gleich 0. Daher gilt (1) und X^2 und Y^2 sind unabhängig.

4 Unabhängige ZV und ihre charakteristischen Funktionen

4.1 Definition und Motivation

Sei X eine ZV auf $(\Omega, \mathcal{F}, P)$, dann ist die charakteristische Funktion ϕ_X von X definiert durch

$$\phi_X(t) = E(e^{itX}).$$

Für zwei unabhängige ZV X_1 und X_2 und deren charakteristische Funktionen ϕ_1 und ϕ_2 gilt dann:

$$\phi(t) = \phi_1(t) \cdot \phi_2(t), \tag{4}$$

wobei $\phi(t)$ die charakteristische Funktion von $X_1 + X_2$ ist.
Daher stellt sich die Frage, ob aus (4) auch die Unabhängigkeit von X_1 und X_2 folgt.

4.2 Beispiel

Sei (X_1, X_2) ein Zufallsvektor mit Dichte

$$f(x_1, x_2) = \begin{cases} \frac{1}{4}\Big(1 + x_1 x_2 (x_1^2 - x_2^2)\Big), & \text{falls } |x_1| \leq 1 \text{ und } |x_2| \leq 1, \\ 0, & \text{sonst.} \end{cases}$$

Die Randdichte f_1 von X_1 lässt sich durch Integration leicht bestimmen:

$$f_1(x_1) = \begin{cases} \frac{1}{2}, & \text{falls } |x_1| \leq 1, \\ 0, & \text{sonst.} \end{cases}$$

die Randdichte f_2 von X_2 ergibt sich anlog.
Offensichtlich gilt für $0 < x_1 \leq 1$, $0 < x_2 \leq 1$ und $x_1 > x_2$:

$$f(x_1, x_2) > \frac{1}{4} = f_1(x_1) \cdot f_2(x_2).$$

Somit gilt (3) nicht und X_1 und X_2 sind nicht unabhängig.
Für die charakteristischen Funktionen ergibt sich:

$$\phi_1(t) = \phi_2(t) = E(e^{itX_2}) = \int\limits_{\mathbb{R}} e^{itx} f_2(x)\,\mathrm{d}x = \int\limits_{-1}^{1} e^{itx}\frac{1}{2}\,\mathrm{d}x = \frac{1}{2it}(e^{it} - e^{-it}) = \frac{1}{t}\sin t.$$

Durch Integration bestimmt man die Dichte g von $X_1 + X_2$:

$$g(x) = \int\limits_{\mathbb{R}} f(x_1, x - x_1)\,\mathrm{d}x_1.$$

Damit f nicht 0 ist, muss $|x_1| \leq 1$ und $|x - x_1| \leq 1$ gelten. Da für $x = 0$ die beiden Bedingungen gleich sind, bietet sich eine Fallunterscheidung an:

Für $x < 0$ gilt:
$|x - x_1| \leq 1 \Rightarrow x_1 \leq x + 1$ und $x_1 \geq x - 1$.
Aus der ersten Ungleichung, $x_1 \leq x + 1$, und $|x_1| \leq 1$ folgt $x \geq -2$, außerdem verschärft sie die Ungleichung $x_1 \leq 1$. Die zweite Ungleichung ist für $x_1 \geq -1$ immer erfüllt.
Also gilt für die Dichte g:

$$\begin{aligned}
g(x) &= \int\limits_{\mathbb{R}} f(x_1, x - x_1)\,\mathrm{d}x_1 = \int\limits_{-1}^{x+1} f(x_1, x - x_1)\,\mathrm{d}x_1 = \frac{1}{4}\int\limits_{-1}^{x+1} 1 + x_1(x - x_1)(x_1^2 - (x - x_1)^2)\,\mathrm{d}x_1 \\
&= \frac{1}{4}\int\limits_{-1}^{x+1} 1 + x_1(x - x_1)(x_1^2 - (x^2 - 2xx_1 + x_1^2))\,\mathrm{d}x_1 \\
&= \frac{1}{4}\int\limits_{-1}^{x+1} 1 + x_1(x - x_1)(-x^2 + 2xx_1)\,\mathrm{d}x_1 \\
&= \frac{1}{4}\int\limits_{-1}^{x+1} 1 - x^3x_1 + 3x^2x_1^2 - 2xx_1^3\,\mathrm{d}x_1 \\
&= \frac{1}{4}\Big(\big((x+1) - \frac{1}{2}x^3(x+1)^2 + x^2(x+1)^3 - \frac{1}{2}x(x+1)^4\big) - \big(-1 - \frac{1}{2}x^3 - x^2 - \frac{1}{2}x\big)\Big) \\
&= \frac{1}{4}(2 + x).
\end{aligned}$$

Für $x \geq 0$ gilt:
$|x - x_1| \leq 1 \Rightarrow x_1 \leq x + 1$ und $x_1 \geq x - 1$.
Die erste Ungleichung ist für $x_1 \leq 1$ immer erfüllt. Aus der zweiten Ungleichung, $x_1 \geq x - 1$, und $|x_1| \leq 1$ folgt $x \leq 2$, außerdem verschärft sie die Ungleichung $x_1 \geq -1$. Also gilt für die Dichte g:

$$\begin{aligned}
g(x) &= \int\limits_{\mathbb{R}} f(x_1, x - x_1)\,\mathrm{d}x_1 = \int\limits_{x-1}^{1} f(x_1, x - x_1)\,\mathrm{d}x_1 \\
&= \frac{1}{4}\int\limits_{x-1}^{1} 1 - x^3x_1 + 3x^2x_1^2 - 2xx_1^3\,\mathrm{d}x_1 \\
&= \frac{1}{4}\Big(\Big(-1 - \frac{1}{2}x^3 - x^2 - \frac{1}{2}x\Big) - \Big((x-1) - \frac{1}{2}x^3(x-1)^2 + x^2(x-1)^3 - \frac{1}{2}x(x-1)^4\Big)\Big) \\
&= \frac{1}{4}(2 - x).
\end{aligned}$$

Insgesamt ergibt sich:

$$g(x) = \begin{cases} \frac{1}{4}(2 + x), & \text{falls } -2 \leq x < 0, \\ \frac{1}{4}(2 - x), & \text{falls } 0 \leq x \leq 2, \\ 0, & \text{sonst.} \end{cases}$$

Dann erhält man die charakteristische Funktion von $X_1 + X_2$:

$$\begin{aligned}
\phi(t) = E(e^{it(X_1+X_2)}) &= \int\limits_{\mathbb{R}} e^{itx} g(x)\,\mathrm{d}x = \int\limits_{-2}^{0} e^{itx}\frac{1}{4}(2 + x)\,\mathrm{d}x + \int\limits_{0}^{2} e^{itx}\frac{1}{4}(2 - x)\,\mathrm{d}x \\
&= \frac{1}{2}\int\limits_{-2}^{0} e^{itx}\,\mathrm{d}x + \frac{1}{4}\int\limits_{-2}^{0} e^{itx}x\,\mathrm{d}x + \frac{1}{2}\int\limits_{0}^{2} e^{itx}\,\mathrm{d}x - \frac{1}{4}\int\limits_{0}^{2} e^{itx}x\,\mathrm{d}x \\
&= \frac{1}{2}\int\limits_{-2}^{2} e^{itx}\,\mathrm{d}x + \frac{1}{4}\Big(\frac{2}{it}e^{-2it} - \int\limits_{-2}^{0}\frac{1}{it}e^{itx}\,\mathrm{d}x\Big) - \frac{1}{4}\Big(\frac{2}{it}e^{2it} - \int\limits_{0}^{2}\frac{1}{it}e^{itx}\,\mathrm{d}x\Big) \\
&= \frac{e^{2it} - e^{-2it}}{2it} + \frac{e^{-2it}}{2it} - \frac{1 - e^{-2it}}{(2it)^2} - \frac{e^{2it}}{2it} + \frac{e^{2it} - 1}{(2it)^2} \\
&= \frac{e^{-2it} - 2 + e^{2it}}{(2it)^2} = \frac{e^{-2it} - 2e^{-it}e^{it} + e^{2it}}{(2it)^2} = \Big(\frac{e^{-it} - e^{it}}{2it}\Big)^2 \\
&= \Big(\frac{1}{t}\sin t\Big)^2.
\end{aligned}$$

Somit gilt (4), obwohl X_1 und X_2 nicht unabhängig sind.

5 Unabhängige ZV und ihre momenterzeugenden Funktionen

5.1 Definition und Motivation

Sei X eine ZV auf $(\Omega, \mathcal{F}, P)$, dann ist die momenterzeugende Funktion M_X von X definiert durch

$$ {}_X(t) = E(e^{tX}). $$

Für zwei unabhängige ZV X und Y und deren momenterzeugende Funktionen M_X und M_Y gilt für die momenterzeugende Funktion $M_{X+Y}(t)$ von $X+Y$:

$$ M_{X+Y}(t) = M_X(t) \cdot M_Y(t). \tag{5} $$

Daher stellt sich auch hier die Frage, ob aus (5) auch die Unabhängigkeit von X und Y folgt.

5.2 Beispiel

Zur Vereinfachung sei im Folgenden: $p_{i,j} := P(X = i, Y = j)$.
Sei (X, Y) ein Zufallsvektor, der durch die folgenden Wahrscheinlichkeiten definiert ist:

$$ \begin{aligned} p_{1,2} = p_{2,3} = p_{3,1} &= \frac{1}{18}, \\ p_{1,1} = p_{2,2} = p_{3,3} &= \frac{2}{18}, \\ p_{1,3} = p_{2,1} = p_{3,2} &= \frac{3}{18}. \end{aligned} $$

Offensichtlich gilt für alle $k = 1, 2, 3$:

$$ P(X = k) = P(Y = k) = \frac{1}{18} + \frac{2}{18} + \frac{3}{18} = \frac{1}{3}. $$

Für die Wahrscheinlichkeitsverteilung der Summe $X + Y$ ergibt sich:

$$ \begin{aligned} P(X+Y=2) &= \frac{2}{18} = \frac{1}{9}, \\ P(X+Y=3) &= \frac{1}{18} + \frac{3}{18} = \frac{2}{9}, \\ P(X+Y=4) &= \frac{1}{18} + \frac{2}{18} + \frac{3}{18} = \frac{3}{9}, \\ P(X+Y=5) &= \frac{1}{18} + \frac{3}{18} = \frac{2}{9}, \\ P(X+Y=6) &= \frac{2}{18} = \frac{1}{9}. \end{aligned} $$

Betrachte nun die momenterzeugenden Funktionen:

$$M_X(t) = M_Y(t) = E(e^{tY}) = \sum_{y\in\{1,2,3\}} e^{ty} \cdot P(Y = y) = \frac{1}{3}\Big(e^t + e^{2t} + e^{3t}\Big),$$

$$M_{X+Y}(t) = E(e^{t(X+Y)}) = \sum_{z\in\{2,3,4,5,6\}} e^{tz} \cdot P(X + Y = z) = \frac{1}{9}\Big(e^{2t} + 2e^{3t} + 3e^{4t} + 2e^{5t} + e^{6t}\Big).$$

Somit ist (5) erfüllt, obwohl X und Y nicht unabhängig sind, wie man an folgendem Beispiel erkennen kann:

$$P(X = 1, Y = 2) = \frac{1}{18} \neq \frac{1}{9} = P(X = 1) \cdot P(Y = 2).$$

6 Unkorrelierte aber nicht unabhängige ZV

6.1 Definition und Motivation

Seien X und Y reelle Zufallsvariablen auf dem gleichen Wahrscheinlichkeitsraum mit $0 < \operatorname{Var} X, \operatorname{Var} Y < \infty$. Dann ist der Korrelationskoeffizient $\varrho(X, Y)$ definiert als

$$\varrho(X,Y) := \frac{Cov(X,Y)}{\sqrt{\operatorname{Var} X}\sqrt{\operatorname{Var} Y}} = \frac{E(X \cdot Y) - E(X) \cdot E(Y)}{\sqrt{\operatorname{Var} X}\sqrt{\operatorname{Var} Y}}.$$

Man nennt X und Y unkorreliert, wenn der Korrelationskoeffizient gleich Null ist, also wenn $Cov(X, Y) = 0$ gilt.
Es ist leicht zu zeigen, dass aus der Unabhängigkeit von X und Y deren Unkorreliertheit folgt. Mit den folgenden Beispielen wird deutlich, dass die Umkehrung im Allgemeinen nicht gilt.

6.2 Beispiel

Seien X und Y diskrete ZV, die jeweils die Werte $-1, 0, 1$ annehmen. Zur Vereinfachung sei $p_{i,j} := P(X = i, Y = j)$ und die Wahrscheinlichkeiten seien wie folgt gegeben:

$$p_{-1,-1} = p_{-1,0} = p_{-1,1} = p_{0,-1} = p_{0,1} = p_{1,-1} = p_{1,0} = p_{1,1} = \frac{1}{8}, \qquad p_{0,0} = 0.$$

Also gilt für die Wahrscheinlichkeiten von X und Y:

$$P(X = 0) = P(Y = 0) = \frac{1}{4}, \qquad P(X = -1) = P(Y = -1) = P(X = 1) = P(Y = 1) = \frac{3}{8}.$$

Für den Erwartungswert folgt:

$$E(X) = \sum_{x\in\{-1,0,1\}} x \cdot P(X = x) = -\frac{3}{8} + 0 + \frac{3}{8} = 0.$$

Ebenso ergibt sich $E(Y) = 0$ und

$$E(XY) = \sum_{x,y \in \{-1,0,1\}} xy \cdot p_{x,y} = 1 \cdot \frac{1}{8} + 0 + (-1) \cdot \frac{1}{8} + 0 + 0 + 0 + (-1) \cdot \frac{1}{8} + 0 + 1 \cdot \frac{1}{8} = 0.$$

Außerdem gilt

$$E(Y^2) = E(X^2) = \sum_{x \in \{-1,0,1\}} x^2 \cdot P(X = x) = (-1)^2 \cdot \frac{3}{8} + 0 \cdot \frac{1}{4} + 1^2 \cdot \frac{3}{8} = \frac{6}{8},$$

also

$$\operatorname{Var} X = \operatorname{Var} Y = E(Y^2) - (E(Y))^2 = \frac{6}{8} - 0 = \frac{6}{8}.$$

Damit ist $0 < \operatorname{Var} X = \operatorname{Var} Y < \infty$, was die Vorraussetzung für die Existenz des Korrelationkoefizienten ist.
Daher sind X und Y unkorreliert. Allerdings ist (2) für $n = 2$ nicht erfüllt, denn

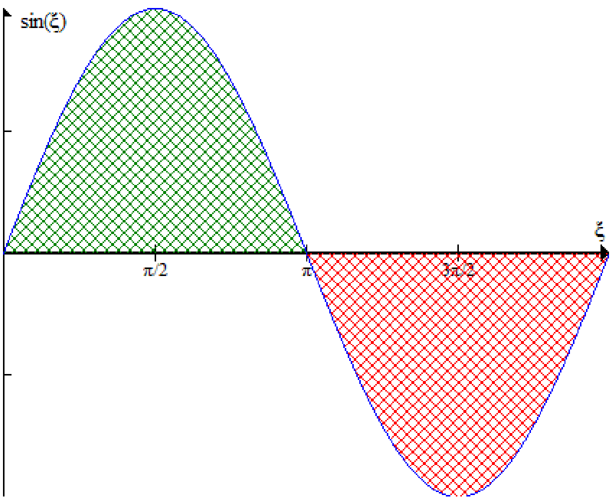
$$P(X = 0, Y = 0) = 0 \neq \frac{1}{16} = P(X = 0) \cdot P(Y = 0).$$

Somit sind X und Y nicht unabhängig, obwohl sie unkorreliert sind.

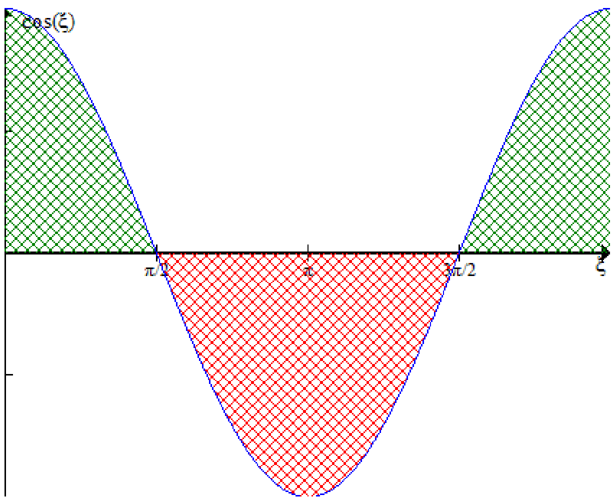
6.3 Beispiel

Sei ξ gleichverteilt auf $(0, 2\pi)$ und X, Y absolut stetige ZV mit $X = \sin \xi$ und $Y = \cos \xi$. Dann gilt $XY = \frac{1}{2} \sin 2\xi$ und die Werte, die die ZV annehmen können, lassen sich grafisch darstellen:

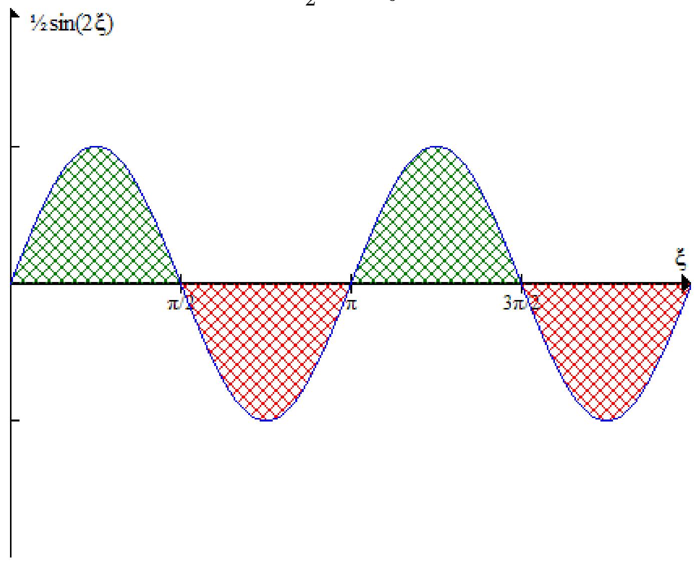
$\sin \xi$



$\cos \xi$



$\frac{1}{2}\sin 2\xi$



Da die Flächen oberhalb und unterhalb der x-Achse jeweils gleich groß sind, ergibt sich für die Erwartungswerte:

$$E(X) = 0, \quad E(Y) = 0, \quad E(XY) = 0.$$

Es lässt sich leicht ablesen, dass die Erwartungswerte von X^2 und Y^2 zwischen 0 und 1 liegen und somit die Bedingung an die Varianz von X und Y erfüllt sind.
Also sind X und Y unkorreliert. Allerdings sind X und Y funktional abhängig ($X^2 + Y^2 = 1$) und daher nicht unabhängig.

6.4 Beispiel

Sei $X_1 \sim \mathcal{N}(0,1)$. Dann ist bekannt, dass $E(X_1) = 0$ und $\mathrm{Var}(X_1) = 1$ gilt. Außerdem ist die momenterzeugende Funktion bekannt:

$$M_{X_1}(t) = e^{\frac{t^2}{2}}.$$

Dadurch lassen sich die Momente berechnen:

$$\begin{aligned}
E(X) &= \frac{d}{dx} M_{X_1}(t)|_{t=0} = te^{\frac{t^2}{2}}|_{t=0} = 0,\\
E(X^2) &= \frac{d^2}{dx^2} M_{X_1}(t)|_{t=0} = \frac{d}{dx} te^{\frac{t^2}{2}}|_{t=0} = e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}}|_{t=0} = 1,\\
E(X^3) &= \frac{d^3}{dx^3} M_{X_1}(t)|_{t=0} = \frac{d}{dx} e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}}|_{t=0} = te^{\frac{t^2}{2}} + 2te^{\frac{t^2}{2}} + t^3 e^{\frac{t^2}{2}}|_{t=0} = 0,\\
E(X^4) &= \frac{d^4}{dx^4} M_{X_1}(t)|_{t=0} = \frac{d}{dx} te^{\frac{t^2}{2}} + 2te^{\frac{t^2}{2}} + t^3 e^{\frac{t^2}{2}}|_{t=0}\\
&= e^{\frac{t^2}{2}} + t^2 e^{\frac{t^2}{2}} + 2e^{\frac{t^2}{2}} + 2t^2 e^{\frac{t^2}{2}} + 3t^2 e^{\frac{t^2}{2}} + t^4 e^{\frac{t^2}{2}}|_{t=0} = 3.
\end{aligned}$$

Also:

$$\begin{aligned}\mathrm{Var}(X_2) &= \mathrm{Var}(X_1^2 - 1) = E((X_1^2 - 1)^2) - (E(X_1^2 - 1))^2 \\ &= E(X_1^4) - 2E(X_1^2) + 1 - (E(X_1^2) - 1)^2 = 3 - 2 + 1 - 0 = 2.\end{aligned}$$

Somit ist die Bedingung an die Varianz von X_1 und X_2 erfüllt.
Wegen $E(X_1) = 0$ und

$$E(X_1 \cdot X_2) = E(X_1^3 - X_1) = E(X_1^3) - E(X_1) = 0 - 0 = 0$$

sind X_1 und X_2 unkorreliert. Allerdings sind die beiden ZV funktional abhängig und damit nicht unabhängig.

7 Unabhängigkeit und bedingte Unabhängigkeit

7.1 Definition und Motivation

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, $A, B, C \in \mathcal{F}, P(B) > 0, P(C) > 0$.
Dann ist die bedingte Wahrscheinlichkeit von A unter der Bedingung B:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Zwei Ereignisse A und B heißen bedingt unabhängig gegeben C, wenn

$$P(A \cap B|C) = P(A|C) \cdot P(B|C) \tag{6}$$

gilt.
Aufgrund der Ähnlichkeit der Namen könnte man eine Verbindung der Begriffe "Unabhängigkeit" und "Bedingte Unabhängigkeit" erwarten. Diese existiert allerdings nicht, wie in den folgenden zwei Beispielen verdeutlicht wird.

7.2 Beispiel

Seien X_1 und X_2 unabhängige, Bernoulli-verteilte ZV mit $p = \frac{1}{3}$. X_1 und X_2 nehmen also jeweils mit Wahrscheinlichkeit $\frac{1}{3}$ den Wert 1 und mit Wahrscheinlichkeit $\frac{2}{3}$ den Wert 0 an. Sei $S_2 = X_1 + X_2$.
Für $S_2 \in \{0,2\}$ existiert nur die Möglichkeit $X_1 = X_2 = 0$ oder $X_1 = X_2 = 1$, also $P\Big((S_2 \in \{0,2\})\Big) = \frac{5}{9}$.

$$P\Big((X_1 = 1) \cap (X_2 = 1) | (S_2 \in \{0,2\})\Big) = \frac{P\Big((X_1 = 1) \cap (X_2 = 1) \cap (S_2 \in \{0,2\})\Big)}{P\Big((S_2 \in \{0,2\})\Big)} = \frac{\frac{1}{9}}{\frac{5}{9}} = \frac{1}{5},$$

$$\begin{aligned} &P\Big((X_1 = 1) | (S_2 \in \{0,2\})\Big) \cdot P\Big((X_2 = 1) | (S_2 \in \{0,2\})\Big) \\ &\quad = \frac{P\Big((X_1 = 1) \cap (S_2 \in \{0,2\})\Big) \cdot P\Big((X_2 = 1) \cap S_2 \in \{0,2\})\Big)}{P\Big((S_2 \in \{0,2\})\Big)^2} = \frac{\frac{1}{9} \cdot \frac{1}{9}}{\Big(\frac{5}{9}\Big)^2} = \frac{1}{25}. \end{aligned}$$

Somit ist (6) nicht erfüllt und die Ereignisse sind nicht bedingt unabhängig, obwohl die ZV bzw. Ereignisse unabhängig sind.

7.3 Beispiel

Seien X_1, X_2 und X_3 unabhängige, Bernoulli-verteilte ZV mit $p = \frac{1}{2}$.
Sei $S_i = X_1 + \ldots + X_i$ für $i = 1, 2, 3$.
Dann sind S_1 und S_3 offensichtlich nicht unabhängig:

$$P\Big(S_1 = 0, S_3 = 3\Big) = 0 \neq \frac{1}{2} \cdot \frac{1}{8} = P\Big(S_1 = 0\Big) \cdot P\Big(S_3 = 3\Big).$$

Für $P(S_2 = k) > 0$ gilt:

$$\begin{aligned} &P\Big(S_1 = i, S_3 = j | S_2 = k\Big) = \frac{P\Big(S_1 = i, S_2 = k, S_3 = j\Big)}{P\Big(S_2 = k\Big)} = \frac{P\Big(S_1 = i, S_2 = k, X_3 = j - k\Big)}{P\Big(S_2 = k\Big)} \\ &= \frac{P\Big(S_1 = i, S_2 = k\Big)}{P\Big(S_2 = k\Big)} P\Big(X_3 = j - k\Big) = P\Big(S_1 = i | S_2 = k\Big) \frac{P\Big(X_3 = j - k\Big) \cdot P\Big(S_2 = k\Big)}{P\Big(S_2 = k\Big)} \\ &= P\Big(S_1 = i | S_2 = k\Big) \frac{P\Big(S_3 = j, S_2 = k\Big)}{P\Big(S_2 = k\Big)} = P\Big(S_1 = i | S_2 = k\Big) \cdot P\Big(S_3 = j | S_2 = k\Big). \end{aligned}$$

Also gilt (6) und die Ereignisse S_1 und S_3 sind bedingt unabhängig, obwohl sie nicht unabhängig sind.