

Statistik für Informatiker, SS 2021

1 Grundlagen aus der Wahrscheinlichkeitstheorie

1.3 Bedingte Wahrscheinlichkeiten, Unabhängigkeit, gemeinsame Verteilung

Matthias Birkner

~~18.~~ Mai 2021

19.

Ein Beispiel zur Motivation

Beispiel 1.37

Wir ziehen zwei Kugeln ohne Zurücklegen aus einer Urne mit $s > 0$ schwarzen und $w > 0$ weißen Kugeln. Wir stellen uns die Kugeln nummeriert vor, Nr. $1, \dots, w$ seien weiß, $w+1, \dots, w+s$ schwarz. Sei X_i die Nr. der Kugel im i -ten Zug ($= 1, 2$).

$X = (X_1, X_2)$ ist uniform verteilt auf $S = \{(i, j) :$

$A = \{\text{erste gez. Kugel weiß}\} = \{X_1 \leq w\}$ $\begin{matrix} 1 \leq i, j \leq w+s, \\ i \neq j \end{matrix}$

$B = \{\text{zweite} \quad \text{---} \text{---} \text{---}\} = \{X_2 \leq w\}$

$(= \{X \in \{(i, j) \in S : j \leq w\})$

Beispiel 1.37, Forts.

$$A = \{\text{erste Kugel ist weiß}\}, \quad B = \{\text{zweite Kugel ist weiß}\}$$

Ohne weitere Informationen ist

$$\begin{aligned} P(B) &= P(X \in \{(i, j) \in S : j \leq w\}) = \frac{\#\{(i, j) \in S : j \leq w\}}{\#S} \\ &= \frac{w \cdot (w+s-1)}{(w+s) \cdot (w+s-1)} = \frac{w}{w+s} \quad (= P(A)) \end{aligned}$$

Beispiel 1.37, Forts.

$$A = \{\text{erste Kugel ist weiß}\}, \quad B = \{\text{zweite Kugel ist weiß}\}$$

Ohne weitere Informationen ist

$$\begin{aligned} P(B) &= P(X \in \{(i, j) \in S : j \leq w\}) \\ &= \frac{\#\{(i, j) \in S : j \leq w\}}{\#S} = \frac{w(w + s - 1)}{(w + s)(w + s - 1)} = \frac{w}{w + s}. \end{aligned}$$

(und übrigens $= P(A)$)

Nehmen wir an, wir haben den ersten Zug beobachtet und gesehen, dass A eingetreten ist. Mit dieser Information sollte die W'keit von B

$$\frac{w-1}{w+s-1} < \frac{w}{w+s}$$

Beobachtung und Definition 1.38

Sei $(\Omega, \mathcal{F}, P)$ W'raum, $A \in \mathcal{F}$ mit $P(A) > 0$. Für $B \in \mathcal{F}$

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

heißt bedingte Wahrscheinlichkeit von B , gegeben A .

$P(\cdot | A)$ ist ein Wahrscheinlichkeitsmaß auf Ω , d.h. die Eigenschaften aus Forderung (1.2) Normierung und σ -Additivität sind erfüllt. ✓

Wir lassen $P(B | A)$ undefiniert, wenn $P(A) = 0$.

$$P(\emptyset | A) = \frac{P(\emptyset \cap A)}{P(A)} = \frac{0}{P(A)} = 0$$

Seien $A_1, A_2, \dots$ paarw. disj.,

$$P\left(\bigcup_{n=1}^{\infty} A_n | A\right) = \frac{P\left(\left(\bigcup_{n=1}^{\infty} A_n\right) \cap A\right)}{P(A)}$$

$$\sum_{n=1}^{\infty} P(A_n | A)$$

$$\sum_{n=1}^{\infty} \frac{P(A_n \cap A)}{P(A)}$$

$$\bigcup_{n=1}^{\infty} (A_n \cap A)$$

$$P\left(\left(\bigcup_{n=1}^{\infty} A_n\right) \cap A\right) / P(A)$$

In Beispiel 1.37 ist

$$A = \{\text{erste Kugel ist weiß}\}, \quad B = \{\text{zweite Kugel ist weiß}\}$$

$$P(A) = \frac{w}{w+s}$$

$$P(A \cap B) = P(X_1 \leq w, X_2 \leq w) = \frac{w(w-1)}{(w+s)(w+s-1)}$$

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

$$= \frac{\frac{w(w-1)}{(w+s)(w+s-1)}}{\frac{w}{w+s}} = \frac{w-1}{w+s-1}$$

Bemerkung 1.39 („Natürlichkeit von Definition 1.38“)

Nehmen wir an, wir möchten angesichts der Information „ A ist eingetreten“ das W'maß P revidieren zu einem W'maß $\tilde{P}$ mit

1. $\tilde{P}(A) = 1$ (d.h. A ist sicher unter $\tilde{P}$) und
2. $\tilde{P}(B) = c_A P(B)$ für $B \subset A$ mit einem $c_A > 0$
(d.h. Teilereignisse von A erhalten bis auf Normierung ihr altes Gewicht).

Dann gilt $\tilde{P}(C) = \frac{P(A \cap C)}{P(A)}$ für jedes $C \in \mathcal{F}$,

denn

$$\tilde{P}(C) = \tilde{P}(A \cap C) + \underbrace{\tilde{P}(A^c \cap C)}_{\leq \tilde{P}(A^c) = 1 - \tilde{P}(A) = 0} = c_A \cdot P(C).$$

Mit Wahl $C = A$ folgt

$$1 = \tilde{P}(A) = c_A \cdot P(A) \Rightarrow c_A = \frac{1}{P(A)}$$

✓

Beachte:

(Stochastische) Abhängigkeit und Kausalität sind nicht dasselbe

Bemerkung 1.40

$P(B | A) \neq P(B)$ kann nicht notwendigerweise als „Kausalität“ (im Sinne von „ A beeinflusst, ob B eintritt“) interpretiert werden:

So ist in Beispiel 1.37 ist auch

$$P(A | B) = \frac{P(B \cap A)}{P(B)} = \frac{w-1}{w+s-1} \neq P(A), \quad \text{mit } P(A) = \frac{w}{w+s}$$

aber es passt nicht zu unserer Vorstellung, dass der 2. Zug den 1. Zug beeinflusst.

Satz 1.41

Sei $(\Omega, \mathcal{F}, P)$ ein Wahrscheinlichkeitsraum, I eine (höchstens abzählbare) Indexmenge, $B_i \in \mathcal{F}$ paarweise disjunkt mit $P(\cup_{i \in I} B_i) = 1$ und $P(B_i) > 0$ für $i \in I$.

1. (Formel von der totalen Wahrscheinlichkeit)

Für $A \in \mathcal{F}$ gilt

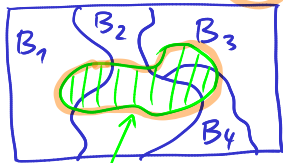
$$P(A) = \sum_{i \in I} P(B_i) P(A | B_i). \checkmark$$

denn

$$\sum_{i \in I} P(B_i) P(A | B_i)$$

$$= \sum_{i \in I} \cancel{P(B_i)} \cdot \frac{P(A \cap B_i)}{\cancel{P(B_i)}}$$

$$= \sum_{i \in I} P(A \cap B_i) = \underbrace{P(B_i)}_{\text{disj.}} (A \cap B_i)_{\text{paarw.}} = P\left(\bigcup_{i \in I} (A \cap B_i)\right) = P(A)$$



Fläche 1

Ω

Satz 1.41 (Fortsetzung)

$B_i \in \mathcal{F}$ paarw. disjunkt mit $P(\cup_{i \in I} B_i) = 1$ und $P(B_i) > 0$ für $i \in I$

2. (Formel von Bayes)

Für $A \in \mathcal{F}$ mit $P(A) > 0$ und jedes $k \in I$ gilt

$$P(B_k | A) = \frac{P(B_k)P(A | B_k)}{\sum_{i \in I} P(B_i)P(A | B_i)}$$

Satz 1.41 (Fortsetzung)

$B_i \in \mathcal{F}$ paarw. disjunkt mit $P(\bigcup_{i \in I} B_i) = 1$ und $P(B_i) > 0$ für $i \in I$

2. (Formel von Bayes)

Für $A \in \mathcal{F}$ mit $P(A) > 0$ und jedes $k \in I$ gilt

$$P(B_k | A) = \frac{P(B_k)P(A | B_k)}{\sum_{i \in I} P(B_i)P(A | B_i)}$$



Insbesondere gilt für Ereignisse A, B mit $P(A) > 0$

$$P(B | A) = \frac{P(B)P(A | B)}{P(B)P(A | B) + P(B^c)P(A | B^c)}$$

(bedr.
 $B_1 = B,$
 $B_2 = B^c$)

denn Zähler $\cancel{P(B_k)} \cdot \frac{P(A \cap B_k)}{\cancel{P(B_k)}}$,
Nenner = $P(A)$

Beispiel 1.42 (Verrauschter Übertragungskanal)

Ein (einzelnes) Bit X (aus $\{0, 1\}$, es gelte $P(X = 1) = a \in (0, 1)$) wird über einen fehleranfälligen Kanal gesendet, der jede 1 mit W'keit f_1 und jede 0 mit W'keit f_0 flippt, sei Y das empfangene Bit.

Beispiel 1.42 (Verrauschter Übertragungskanal)

Ein (einzelnes) Bit X (aus $\{0, 1\}$, es gelte $P(X = 1) = a \in (0, 1)$) wird über einen fehleranfälligen Kanal gesendet, der jede 1 mit W'keit f_1 und jede 0 mit W'keit f_0 flippt, sei Y das empfangene Bit.

Dann ist

$$\begin{aligned}P(Y = 0 \mid X = 1) &= f_1, & P(Y = 1 \mid X = 1) &= 1 - f_1, \\P(Y = 1 \mid X = 0) &= f_0, & P(Y = 0 \mid X = 0) &= 1 - f_0,\end{aligned}$$

also

$$\begin{aligned}P(Y = 1) &= P(X = 1)P(Y = 1 \mid X = 1) + P(X = 0)P(Y = 1 \mid X = 0) \\&= a(1 - f_1) + (1 - a)f_0,\end{aligned}$$

Beispiel 1.42 (Fortsetzung)

Ein (einzelnes) Bit X (aus $\{0, 1\}$, es gelte $P(X = 1) = a \in (0, 1)$) wird über einen fehleranfälligen Kanal gesendet, der jede 1 mit W'keit f_1 und jede 0 mit W'keit f_0 flippt, sei Y das empfangene Bit.

Analog ist

$$\begin{aligned} P(Y = 0) &= P(X = 0)P(Y = 0 \mid X = 0) + P(X = 1)P(Y = 0 \mid X = 1) \\ &= (1 - a)(1 - f_0) + af_1 \end{aligned}$$

Beispiel 1.42, Forts.

$$\mathcal{B}_1 = \{X=1\}, \\ \mathcal{B}_2 = \{X=0\}$$

$$\begin{aligned} P(X=1 \mid Y=1) \\ &= \frac{P(X=1)P(Y=1 \mid X=1)}{P(X=1)P(Y=1 \mid X=1) + P(X=0)P(Y=1 \mid X=0)} \\ &= \frac{a(1-f_1)}{a(1-f_1) + (1-a)f_0} \end{aligned}$$

Beispiel 1.42, Forts.

$$\begin{aligned} P(X = 1 \mid Y = 1) &= \frac{P(X = 1)P(Y = 1 \mid X = 1)}{P(X = 1)P(Y = 1 \mid X = 1) + P(X = 0)P(Y = 1 \mid X = 0)} \\ &= \frac{a(1 - f_1)}{a(1 - f_1) + (1 - a)f_0} \end{aligned}$$

$$\begin{aligned} P(X = 0 \mid Y = 0) &= \frac{P(X = 0)P(Y = 0 \mid X = 0)}{P(X = 0)P(Y = 0 \mid X = 0) + P(X = 1)P(Y = 0 \mid X = 1)} \\ &= \frac{(1 - a)(1 - f_0)}{(1 - a)(1 - f_0) + af_1} \end{aligned}$$

Beispiel 1.42 (Fortsetzung)

Ein (einzelnes) Bit X (aus $\{0, 1\}$, es gelte $P(X = 1) = a \in (0, 1)$) wird über einen fehleranfälligen Kanal gesendet, der jede 1 mit W'keit f_1 und jede 0 mit W'keit f_0 flippt, sei Y das empfangene Bit.

Z.B. für die konkreten Werte $a = 0.3$, $f_1 = 0.05$, $f_0 = 0.1$ ergibt sich $P(Y = 1) = 0.355$

$$P(X = 1 \mid Y = 1) \approx 0.803, \quad P(X = 0 \mid Y = 0) \approx 0.977.$$

Beispiel 1.43 (Medizinische Reihenuntersuchung)

Eine Krankheit

- ▶ komme bei 2% der Bevölkerung vor („Prävalenz 2%“),
- ▶ ein Test schlage bei 95% der Kranken an („Sensitivität 95%“),
- ▶ aber auch bei 10% der Gesunden („Spezifität 90%“).

Eine zufällig gewählte Person wird mit positivem Resultat getestet. Wie wahrscheinlich ist es, dass sie tatsächlich krank ist?

Sei

$A := \{\text{Test fällt positiv aus}\}, \quad B := \{\text{getestete Person ist krank}\},$

also $P(B) = 0,02, P(A | B) = 0,95, P(A | B^c) = 0,1$

Sei

$A := \{\text{Test fällt positiv aus}\}, \quad B := \{\text{getestete Person ist krank}\},$

also $P(B) = 0,02, P(A | B) = 0,95, P(A | B^c) = 0,1$, somit

$$\begin{aligned} P(B | A) &= \frac{P(B)P(A | B)}{P(B)P(A | B) + P(B^c)P(A | B^c)} \\ &= \frac{0,02 \cdot 0,95}{0,02 \cdot 0,95 + 0,98 \cdot 0,1} \approx 0,162 \end{aligned}$$

Sei

$A := \{\text{Test fällt positiv aus}\}, \quad B := \{\text{getestete Person ist krank}\},$

also $P(B) = 0,02, P(A | B) = 0,95, P(A | B^c) = 0,1$, somit

$$\begin{aligned} P(B | A) &= \frac{P(B)P(A | B)}{P(B)P(A | B) + P(B^c)P(A | B^c)} \\ &= \frac{0,02 \cdot 0,95}{0,02 \cdot 0,95 + 0,98 \cdot 0,1} \approx 0,162 \end{aligned}$$

(wir sehen: geringe „positive Korrektheit“)

Sei

$A := \{\text{Test fällt positiv aus}\}, \quad B := \{\text{getestete Person ist krank}\},$

also $P(B) = 0,02, P(A | B) = 0,95, P(A | B^c) = 0,1$, somit

$$\begin{aligned} P(B | A) &= \frac{P(B)P(A | B)}{P(B)P(A | B) + P(B^c)P(A | B^c)} \\ &= \frac{0,02 \cdot 0,95}{0,02 \cdot 0,95 + 0,98 \cdot 0,1} \approx 0,162 \end{aligned}$$

(wir sehen: geringe „positive Korrektheit“, andererseits

$$\begin{aligned} P(B | A^c) &= \frac{P(B)P(A^c | B)}{P(B)P(A^c | B) + P(B^c)P(A^c | B^c)} \\ &= \frac{0,02 \cdot 0,05}{0,02 \cdot 0,05 + 0,98 \cdot 0,9} \approx 0,0011 \end{aligned}$$

(recht hohe „negative Korrektheit“).

Demnach: Ein negatives Testergebnis schließt die Krankheit mit hoher W'keit aus, ein positives Testergebnis sollte eher als „der Fall sollte weiter beobachtet / untersucht werden“ interpretiert werden als „die Testperson ist krank“.

Demnach: Ein negatives Testergebnis schließt die Krankheit mit hoher W'keit aus, ein positives Testergebnis sollte eher als „der Fall sollte weiter beobachtet / untersucht werden“ interpretiert werden als „die Testperson ist krank“.

S.a. Gerd Gigerenzer, *Das Einmaleins der Skepsis*, Berlin Verlag, 2002, der auch einlädt, den Sachverhalt anschaulich anhand einer „Vierfelder-Tafel“ bezogen auf eine Gesamtpopulation der Größe 1000 zu betrachten:

	krank	gesund	Σ
pos. getestet	19	98	117
neg. getestet	1	882	883
Σ	20	980	1000

Beobachtung 1.44 (Multiplikationsformel)

$A_1, A_2, \dots, A_n \in \mathcal{F}$ Ereignisse (auf einem W'raum $(\Omega, \mathcal{F}, P)$ erklärt) mit

$$P(A_1 \cap \dots \cap A_{n-1}) > 0,$$

so ist

$$\begin{aligned} &P(A_1 \cap A_2 \cap \dots \cap A_n) \\ &= P(A_1) \underbrace{P(A_2 | A_1)} \underbrace{P(A_3 | A_1 \cap A_2)} \dots P(A_n | A_1 \cap \dots \cap A_{n-1}). \end{aligned}$$

Denn: R.S.

$$= \cancel{P(A_1)} \cdot \frac{\cancel{P(A_1 \cap A_2)}}{\cancel{P(A_1)}}, \frac{\cancel{P(A_1 \cap A_2 \cap A_3)}}{\cancel{P(A_1 \cap A_2)}} \dots \dots \frac{\cancel{P(A_1 \cap \dots \cap A_n)}}{\cancel{P(A_1 \cap \dots \cap A_{n-1})}}$$

Betrachten wir z.B. in der Situation von Bsp. 1.37 (Ziehen ohne Zurücklegen aus Urne mit s schwarzen und w weißen Kugeln) drei Züge, sei

$A_1 = \{\text{erste Kugel ist weiß}\}$, $A_2 = \{\text{zweite Kugel ist weiß}\}$,
 $A_3 = \{\text{dritte Kugel ist schwarz}\}$, so ist

$$P(A_1) = \frac{w}{s+w}, \quad P(A_2 | A_1) = \frac{w-1}{s+w-1},$$

$$P(A_3 | A_1 \cap A_2) = \frac{s}{s+w-2},$$

Betrachten wir z.B. in der Situation von Bsp. 1.37 (Ziehen ohne Zurücklegen aus Urne mit s schwarzen und w weißen Kugeln) drei Züge, sei

$A_1 = \{\text{erste Kugel ist weiß}\}$, $A_2 = \{\text{zweite Kugel ist weiß}\}$,
 $A_3 = \{\text{dritte Kugel ist schwarz}\}$, so ist

$$P(A_1) = \frac{w}{s+w}, \quad P(A_2 | A_1) = \frac{w-1}{s+w-1},$$

$$P(A_3 | A_1 \cap A_2) = \frac{s}{s+w-2},$$

also

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= \frac{w}{s+w} \cdot \frac{w-1}{s+w-1} \cdot \frac{s}{s+w-2} \\ &= \frac{w(w-1)s}{(s+w)(s+w-1)(s+w-2)}. \end{aligned}$$

Gemeinsame Verteilung und Randverteilungen

Definition 1.45 (Gemeinsame Verteilung und Marginalverteilungen)

Seien $X_1, X_2, \dots, X_d$ ZVn (die auf einem gemeinsamen W'raum $(\Omega, \mathcal{F}, P)$ definiert sind), X_i habe Werte in S_i ,

$$X := (X_1, X_2, \dots, X_d)$$

die Produkt-Zufallsvariable (mit Werten in $S_1 \times \dots \times S_d$), so heißt $\mathcal{L}(X)$ die *gemeinsame Verteilung* der $X_1, X_2, \dots, X_d$

($\mathcal{L}(X)$ ist ein Wahrscheinlichkeitsmaß auf dem Wertebereich von X , d.h. auf dem Produktraum $S_1 \times \dots \times S_d$, vgl. Definition 1.10).

$\mathcal{L}(X_i)$ heißt die *i-te Randverteilung* (oder *Marginalverteilung*) von X .

Beispiel

Fassen wir die Häufigkeitstabelle (normiert auf Gesamtsumme = 1) aus Bsp. 1.43 als Verteilungstafel der gemeinsamen Verteilung von

X_1 („Testergebnis“ mit Werten in $S_1 = \{\text{positiv}, \text{negativ}\}$) und X_2 („Gesundheitszustand“ mit Werten in $S_2 = \{\text{krank}, \text{gesund}\}$) auf:

$X_1 \backslash X_2$	krank	gesund	
positiv	0.019	0.098	0.117
negativ	0.001	0.882	0.883
	0.02	0.98	1

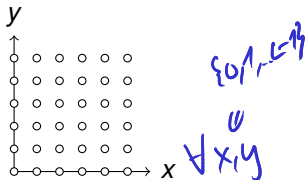
Beispiel 1.46

1. Sei $X = (X_1, Y_1)$ uniform verteilt auf dem (diskreten) Quadrat $Q_L := \{0, 1, \dots, L-1\}^2 = \{(x, y) \in \mathbb{N}_0^2 : \max\{x, y\} < L\}$ (mit $L \in \mathbb{N}$), d.h.

$$P(X_1 = x, Y_1 = y) = \frac{1}{\#Q_L} = \frac{1}{L^2} \quad \text{für } x, y \in \{0, 1, \dots, L-1\}.$$

Für $x \in \{0, 1, \dots, L-1\}$:

$$\begin{aligned} P(X_1 = x) &= \sum_{y=0}^{L-1} P(X_1 = x, Y_1 = y) \\ &= L \cdot \frac{1}{L^2} = \frac{1}{L} \end{aligned}$$

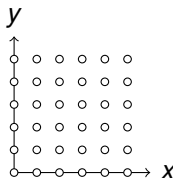


$= P(Y_1 = x)$, d.h. X_1 und Y_1 sind uniform verteilt
 und $P(X_1 = x, Y_1 = y) = \frac{1}{L^2} = \frac{1}{L} \cdot \frac{1}{L} = P(X_1 = x) \cdot P(Y_1 = y)$

Beispiel 1.46

1. Sei $X = (X_1, Y_1)$ uniform verteilt auf dem (diskreten) Quadrat $Q_L := \{0, 1, \dots, L-1\}^2 = \{(x, y) \in \mathbb{N}_0^2 : \max\{x, y\} < L\}$ (mit $L \in \mathbb{N}$), d.h.

$$P(X_1 = x, Y_1 = y) = \frac{1}{\#Q_L} = \frac{1}{L^2} \quad \text{für } x, y \in \{0, 1, \dots, L-1\}.$$



Wir sehen hier:

$P(X_1 = x, Y_1 = y) = P(X_1 = x) \cdot P(Y_1 = y)$, d.h. die gemeinsamen Gewichte erhält man als *Produkt* der Randverteilungs-Gewichte. Man sagt dazu auch: X_1 und Y_1 sind *unabhängig* (siehe Def. 1.51 unten für den allgem. Fall).

Beispiel 1.46 (Fortsetzung)

2. Sei $X = (X_2, Y_2)$ uniform verteilt auf dem (diskreten) Dreieck

$D_L := \{(x, y) \in \mathbb{N}_0^2 : x + y < L\}$, d.h.

$$(\#D_L = 1 + 2 + \dots + L$$

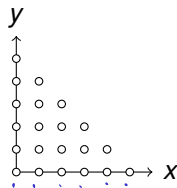
$$P(X_2 = x, Y_2 = y) = \frac{1}{\#D_L} = \frac{2}{L(L+1)} \quad \text{für } (x, y) \in D_L = \frac{L(L+1)}{2}$$

$$P(X_2 = x) = \sum_{y=0}^{L-1-x} P(X_2 = x, Y_2 = y)$$

$$= \sum_{y=0}^{L-1-x} \frac{2}{L(L+1)}$$

$$= \frac{2(L-x)}{L(L+1)} = \frac{2}{L(L+1)}$$

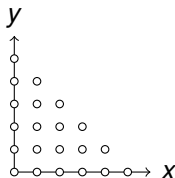
$$Y_2 \stackrel{d}{=} X_2$$



Beispiel 1.46 (Fortsetzung)

2. Sei $X = (X_2, Y_2)$ uniform verteilt auf dem (diskreten) Dreieck $D_L := \{(x, y) \in \mathbb{N}_0^2 : x + y < L\}$, d.h.

$$P(X_2 = x, Y_2 = y) = \frac{1}{\#D_L} = \frac{2}{L(L+1)} \quad \text{für } (x, y) \in D_L$$



Wir sehen in diesem Fall:

$P(X_2 = x, Y_2 = y) \neq P(X_2 = x) \cdot P(Y_2 = y)$, die Koordinaten X_2 und Y_2 sind nicht unabhängig.

Beispiel 1.46 (Fortsetzung)

3. Sei $X = (X_3, Y_3)$ uniform verteilt auf der (diskreten, verschobenen) Nebendiagonale $N_L := \{(x, y) \in \mathbb{N}_0^2 : x + y = L\}$, d.h.

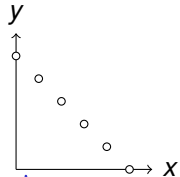
$$P(X_3 = x, Y_3 = y) = \frac{1}{\#N_L} = \frac{1}{L} \quad \text{für } (x, y) \in N_L$$

$$P(X_3 = x) = \sum_{y=0}^{L-1} P(X_3 = x, Y_3 = y)$$

$$= \frac{1}{L}$$

$$= \begin{cases} \frac{1}{L}, & y = L - x \\ 0, & \text{sonst} \end{cases}$$

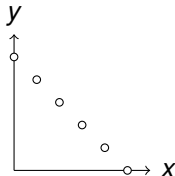
$$= P(Y_3 = y) \quad , \quad x, y \in \{0, 1, \dots, L-1\}$$



Beispiel 1.46 (Fortsetzung)

3. Sei $X = (X_3, Y_3)$ uniform verteilt auf der (diskreten, verschobenen) Nebendiagonale $N_L := \{(x, y) \in \mathbb{N}_0^2 : x + y = L\}$, d.h.

$$P(X_3 = x, Y_3 = y) = \frac{1}{\#N_L} = \frac{1}{L} \quad \text{für } (x, y) \in N_L$$

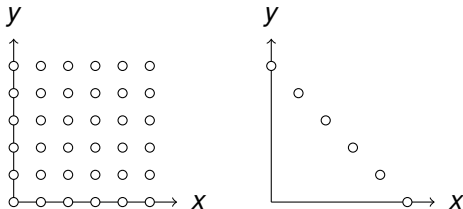


Wir sehen in diesem Fall:

$P(X_3 = x, Y_3 = y) \neq P(X_3 = x) \cdot P(Y_3 = y)$, die Koordinaten X_3 und Y_3 sind nicht unabhängig.

(Es gilt hier sogar $Y_3 = L - X_3$, d.h. es gibt einen deterministischen Zusammenhang.)

Zudem sehen wir an Beispiel 1.46, 1. und 3. zusammen:
Die Randverteilungen legen die gemeinsame Verteilung nicht fest.



So gilt $\mathcal{L}(X_1) = \mathcal{L}(X_3)$ und $\mathcal{L}(Y_1) = \mathcal{L}(Y_3)$
(man schreibt dies auch als $X_1 =^d X_3$, $Y_1 =^d Y_3$,
„Gleichheit in Verteilung“), aber

$$\mathcal{L}((X_1, Y_1)) \neq \mathcal{L}((X_3, Y_3)).$$