Kapitel 2

Bedingte Wahrscheinlichkeiten und Unabhängigkeit

2.1 Bedingte Wahrscheinlichkeiten

Beispiel 2.1. Wir ziehen zwei Kugeln ohne Zurücklegen aus einer Urne mit $s > 0$ schwarzen und $w > 0$ weißen Kugeln.
Modell: Nummeriere die Kugeln, $1, \ldots, w$ seien weiß, $w + 1, \ldots, w + s$ schwarz,
$\Omega = \{(i, j) : 1 \leq i, j \leq w + s, i \neq j\}$,
P = uniforme Verteilung auf Ω ($|\Omega| = (w + s)(w + s - 1)$),
betrache die Ereignisse

$$A = \{\text{erste Kugel ist weiß}\} = \{(i, j) \in \Omega : i \leq w\},$$
$$B = \{\text{zweite Kugel ist weiß}\} = \{(i, j) \in \Omega : j \leq w\}.$$

Wir ziehen zwei Kugeln ohne Zurücklegen aus einer Urne mit $s > 0$ schwarzen und $w > 0$ weißen Kugeln.

Modell: Nummeriere die Kugeln, $1, \dots, w$ seien weiß, $w + 1, \dots, w + s$ schwarz,
$\Omega = \{(i, j) : 1 \le i, j \le w + s, i \neq j\}$,
P = uniforme Verteilung auf Ω ($|\Omega| = (w + s)(w + s - 1)$); betrachte

$$A = \{\text{erste Kugel ist weiß}\} = \{(i, j) \in \Omega : i \le w\},$$
$$B = \{\text{zweite Kugel ist weiß}\} = \{(i, j) \in \Omega : j \le w\}.$$

Ohne weitere Informationen ist

$$P(B) = \frac{|B|}{|\Omega|} = \frac{w\cancel{(w+s-1)}}{(w+s)\cancel{(w+s-1)}} = \frac{w}{w+s}.$$

$(= P(A))$

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

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

sein (denn es „wurde schon eine weiße Kugel verbraucht“).

Beobachtung und Definition 2.2. *Sei $(\Omega, \mathcal{F}, P)$ W'raum, $A \in \mathcal{F}$ mit $P(A) > 0$.*

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

heißt bedingte Wahrscheinlichkeit von B, gegeben A *(für $B \in \mathcal{F}$).*

$P(\cdot \mid A)$ ist ein Wahrscheinlichkeitsmaß auf $(\Omega, \mathcal{F})$,

man prüft leicht per Inspektion, dass die Eigenschaften aus Definition 1.3, Normierung und σ-Additivität, erfüllt sind.

$$\left(P(\Omega \mid A) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1 \right),$$

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

Seien $B_1, B_2, \ldots$ paarw. disj. Ereign.:

$$P\left(\bigcup_{n=1}^{\infty} B_n \,\Big|\, A\right) = \frac{1}{P(A)} P\left(\left(\bigcup_{n=1}^{\infty} B_n\right) \cap A\right) = \frac{1}{P(A)} P\left(\bigcup_{n=1}^{\infty} (B_n \cap A)\right) = \frac{1}{P(A)} \sum_{n=1}^{\infty} P(B_n \cap A) = \sum_{n=1}^{\infty} P(B_n \mid A)$$

In Beispiel 2.1 ist $P(A) = \frac{w}{w+s}$, $\underbrace{P(A \cap B)}$ $= \frac{|A \cap B|}{|\Omega|} = \frac{w(w-1)}{(w+s)(w+s-1)}$, also ergibt sich tatsächlich $P(B \mid A) = \frac{w-1}{w+s-1}$.

$$\frac{|A \cap B|}{|\Omega|} = \frac{w(w-1)}{(w+s)(w+s-1)}$$

also
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\cancel{w}(w-1)}{\cancel{(w+s)}(w+s-1)} \cdot \frac{\cancel{w+s}}{\cancel{w}} = \frac{w-1}{w+s-1}$$

Bemerkung 2.3 („Natürlichkeit von Definition 2.2"). Nehmen wir an, wir möchten angesichts der Information „A ist eingetreten" das W'maß P revidieren zu einem W'maß $\widetilde{P}$ mit

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

Dann gilt

$$\widetilde{P}(C) = \frac{P(A \cap C)}{P(A)} \quad \big(= P(C \mid A) \big) \quad \text{für alle } C \in \mathcal{F}.$$

Denn: Sei $C \in \mathcal{F}$. $\tilde{P}(C) = \tilde{P}(A \cap C) + \underbrace{\tilde{P}(\overbrace{C \setminus A}^{CA^c})}_{\le \tilde{P}(A^c) = 1 - \tilde{P}(A) = 0}$

Mit Wahl $C = A$ (bzw. $B = A$ in 2.) folgt

$1 \overset{1.}{=} \tilde{P}(A) \overset{2.}{=} c_A \cdot P(A) \Rightarrow c_A = \frac{1}{P(A)}$.

Insges.: $\tilde{P}(C) = \tilde{P}(A \cap C) = c_A P(A \cap C) = \frac{P(A \cap C)}{P(A)}$

Bemerkung 2.4. $P(B \mid A) \neq P(B)$ kann nicht notwendigerweise als „Kausalität" (im Sinne von „A beeinflusst, ob B eintritt") interpretiert werden:
In Beispiel 2.1 ist auch

$$P(A \mid B) = \frac{P(B \cap A)}{P(B)} = \frac{w-1}{w+s-1} \neq P(A),$$

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

Satz 2.5. *I abzählbare Indexmenge, $B_i \in \mathcal{F}$ paarweise disjunkt mit $P\left(\bigcup_{i\in I} B_i\right) = 1$ (und $P(B_i) > 0$ für $i \in I$).*

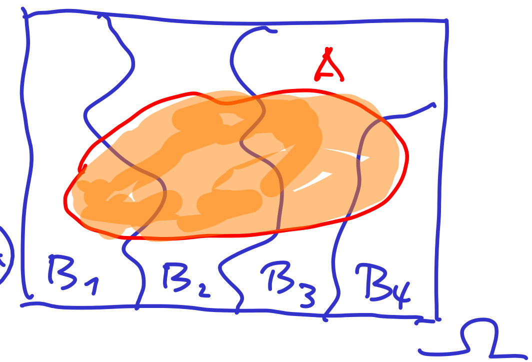
1. *(Formel von der totalen Wahrscheinlichkeit) Für $A \in \mathcal{F}$ gilt*

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

Veranschaulichung am Diagramm:

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

$= P\left(\bigcup_{i\in I}(A\cap B_i)\right) = P\left(A \cap \underbrace{\bigcup_{i\in I} B_i}_{\text{hat } P(..)=1}\right)$ " $P(A)$



2. *(Formel von Bayes[1]) Für $A \in \mathcal{F}$ mit $P(A) > 0$ und jedes $k \in I$ gilt*

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

(denn Nenner rechts = $P(A)$ nach 1.,
Zähler —"— = $\cancel{P(B_k)}\frac{P(A\cap B_k)}{\cancel{P(B_k)}}$)

Bem.: Oft verwendet für $B_1 = B$ und $B_2 = B^c$.

[1] nach Thomas Bayes, 1702–1761; die Arbeit (die eine Frage von Laplace beantwortet) wurde posthum 1763 publiziert

Beispiel 2.6 (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 krank ist?

sei $A = \{\text{Test fällt positiv aus}\}$,

$B = \{\text{Testperson ist krank}\}$

$P(B) = 0{,}02$; $P(A|B) = 0{,}95$; $P(A|B^c) = 0{,}1$

$$P(B|A) = \frac{P(B)P(A|B)}{P(B)P(A|B) + P(B^c)P(A|B^c)} = \ldots \approx 0{,}162$$

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

2.2 Mehrstufige Zufallsexperimente

Wir betrachten folgende Situation:

Wir haben ZVn $X_1, X_2, \dots, X_n$ im Sinn und kennen

1. die Verteilung von X_1,
2. für $2 \le k \le n$ die bedingte Verteilung von X_k, wenn $X_1, X_2, \dots, X_{k-1}$ schon beobachtet wurden.

Frage / Aufgabe : Wie modellieren? (und: wie damit rechnen?)

Beispiel 2.7 (Pólyas Urne[2]). $s, w \in \mathbb{N}_0$, $c \in \{-1, 0, 1, 2, \dots\}$, eine Urne enthält anfangs s schwarze und w weiße Kugeln. Wir greifen in jedem Zug einmal rein zufällig hinein und legen dann die gezogene Kugel zurück zusammen mit c weiteren Kugeln derselben Farbe (im Fall $c = -1$ legen wir die gezogene Kugel nicht zurück). Wir setzen

$$X_i = \mathbf{1}_{\{i\text{-te gezogene Kugel ist schwarz}\}}.$$

[2] nach George Pólya, 1887–1985; siehe F. Eggenberger, G. Pólya, Über die Statistik verketteter Vorgänge, Zeitschr. f. Angew. Math. Mech. 3, 279–290, (1923).

Beobachtung 2.8 (Multiplikationsformel)**.** $A_1, A_2, \ldots, A_n$ *Ereignisse mit*

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

so ist

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1 \cap A_2) \cdots P(A_n \mid A_1 \cap \cdots \cap A_{n-1}).$$

(Beweis per Inspektion, das Produkt rechts teleskopiert)

$$P(\cancel{A_1}) \cdot \underbrace{P(A_2 \mid A_1)}_{\frac{P(\cancel{A_2 \cap} A_1)}{P(\cancel{A_1})}} \cdot \underbrace{P(A_3 \mid A_1 \cap A_2)}_{\frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cancel{\cap} A_2)}} \cdots$$

Satz 2.9 (Konstruktion von W'maßen via bedingte Wahrscheinlichkeiten). *Seien $\Omega_1, \Omega_2, \ldots, \Omega_n$ ($\neq \emptyset$) abzählbar, $p_1 : \Omega_1 \to [0,1]$ Wahrscheinlichkeitsgewichtsfunktion, für $2 \leq k \leq n$ und beliebige $\omega_1 \in \Omega_1, \ldots, \omega_{k-1} \in \Omega_{k-1}$ sei*

$$p_{k|\omega_1,\ldots,\omega_{k-1}} : \Omega_k \to [0,1] \quad \text{W'gewichtsfunktion}$$

setze $\Omega := \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, wir schreiben $X_i : \Omega \to \Omega_i$ mit

$$X_i\big((\omega_1, \ldots, \omega_n)\big) = \omega_i$$

für die i-te Koordinatenprojektion.

Dann ist $p : \Omega \to [0,1]$ mit

$$p\big((\omega_1, \ldots, \omega_n)\big) := p_1(\omega_1) \cdot p_{2|\omega_1}(\omega_2) \cdot p_{3|\omega_1,\omega_2}(\omega_3) \cdot \cdots \cdot p_{n|\omega_1,\ldots,\omega_{n-1}}(\omega_n)$$

eine W'gewichtsfunktion, das W'maß P auf Ω mit Gewichten p erfüllt

1. $P(X_1 = \omega_1) = p_1(\omega_1)$ *für alle* $\omega_1 \in \Omega_1$,

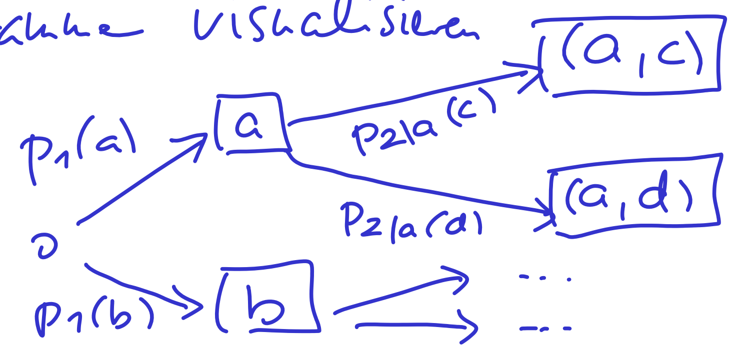
2. $$P(X_k = \omega_k \mid X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1}) = p_{k|\omega_1,\ldots,\omega_{k-1}}(\omega_k)$$

   für $2 \leq k \leq n$, $\omega_1 \in \Omega_1, \ldots, \omega_{k-1} \in \Omega_{k-1}$, sofern $P(X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1}) > 0$.

P ist durch 1. und 2. eindeutig festgelegt.

Man kann dies via Baumdiagramme visualisieren

z.B. $\Omega_1 = \{a, b\}$, $\Omega_2 = \{c, d\}$



$\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_n$, $X_i\big((\omega_1, \ldots, \omega_n)\big) = \omega_i$

$$p\big((\omega_1, \ldots, \omega_n)\big) = p_1(\omega_1) \cdot p_{2|\omega_1}(\omega_2) \cdot p_{3|\omega_1,\omega_2}(\omega_3) \cdot \cdots \cdot p_{n|\omega_1,\ldots,\omega_{n-1}}(\omega_n)$$

Dann gilt: 1. $P(X_1 = \omega_1) = p_1(\omega_1)$, 2. $P(X_k = \omega_k \mid X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1}) = p_{k|\omega_1,\ldots,\omega_{k-1}}(\omega_k)$

Beweis von Satz 2.9.

$$P(X_1 = \omega_1, \ldots, X_k = \omega_k) = \sum_{\omega'_{k+1} \in \Omega_k} \cdots \sum_{\omega'_n \in \Omega_n} P(\{(\omega_1, \ldots, \omega_k, \omega'_{k+1}, \ldots, \omega'_n)\})$$

$$= \sum_{\omega'_{k+1} \in \Omega_{k+1}} \cdots \sum_{\omega'_n \in \Omega_n} p_1(\omega_1) p_{2|\omega_1}(\omega_2) \cdots p_{k|\omega_1,\ldots,\omega_{k-1}}(\omega_k) \cdot p_{k+1|\omega_1,\ldots,\omega_k}(\omega'_{k+1}) \cdots p_{n|\omega_1,\ldots,\omega_k,\omega'_{k+1},\ldots,\omega'_{n-1}}(\omega'_n)$$

$$= p_1(\omega_1) p_{2|\omega_1}(\omega_2) \cdots p_{k|\omega_1,\ldots,\omega_{k-1}}(\omega_k) \cdot \sum_{\omega'_{k+1} \in \Omega_{k+1}} p_{k+1|\omega_1,\ldots,\omega_k}(\omega'_{k+1}) \cdot \underbrace{\cdots}_{= 1 \cdot \ldots \cdot 1} \underbrace{\sum_{\omega'_n \in \Omega_n} p_{n|\omega_1,\ldots,\omega_k,\omega'_{k+1},\ldots,\omega'_{n-1}}(\omega'_n)}_{= 1 \text{ "1"}}$$

□

Mit $k = 1$ folgt 1.

$$P(X_k = \omega_k \mid X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1})$$

$$= \frac{P(X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1}, X_k = \omega_k)}{P(X_1 = \omega_1, \ldots, X_{k-1} = \omega_{k-1})}$$

$$= \ldots = p_{k \mid \omega_1, \ldots, \omega_{k-1}}(\omega_k)$$