Ein Handzettel für heutige Vorlesung finden Sie auf der Seite der Vorlesung auf meiner Website, siehe Tagebuch.

In der letzten Vorlesungen haben wir gesehen: Bedingte W-keit $\mathbb{P}_B$ und überschrankung von σ-Algebren $\mathscr{F}_B$.

3 Bedingte W-keiten, Unabhängigkeit und Produktmaße (fortsetzung)

(Kapitel 3 in Bovier Skript)

Lemma 1. *(Bayes'sche Formel) Seien $A, B \in \mathscr{F}$ $\mathbb{P}(A) > 0$, $\mathbb{P}(B) > 0$. Dann*

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

Beweis. Trivial aus $\mathbb{P}(A|B)\mathbb{P}(B) = \mathbb{P}(A \cap B) = \mathbb{P}(B|A)\mathbb{P}(A)$. □

Die Bayes-Formel erlaubt man, die „Wahrscheinlichkeit von Ursachen" abzuleiten. Typischerweise ist B ein beobachtetes Ereignis und A eine Hypothese. $\mathbb{P}(B|A)$ ist die bedingte Wahrscheinlichkeit, dass B passiert, sobald wir wissen, dass A passiert ist.

Wir können auch vollständige Hypothesenfamilien $(A_n)_n \subseteq \mathscr{F}$ betrachten mit $\cup_n A_n = \Omega$ $A_i \cap A_j \neq \emptyset$. Dann geschieht dies:

$$\mathbb{P}_B(A_n) = \mathbb{P}(A_n|B) = \mathbb{P}(B|A_n)\frac{\mathbb{P}(A_n)}{\mathbb{P}(B)} = \mathbb{P}(B|A_n)\frac{\mathbb{P}(A_n)}{\sum_n \mathbb{P}(B|A_n)\mathbb{P}(A_n)}$$

Wir müssen es also nur die Größe $(\mathbb{P}(A_k))_k$ und $(\mathbb{P}(B|A_k))_k$ wissen. $(\mathbb{P}(A_k))_k$ sind die Apriori-Wahrscheinlichkeiten für die ``Ursachen" und $(\mathbb{P}_B(A_k))_k$ sind ihre Aposteriori-Wahrscheinlichkeiten, sobald wir die Ereignis B beobachtet haben.

Manchmal sind bedingte Erwartungen oder W-keiten nicht ganz so intuitiv. Wir können die Hilfe einer strengen Formalisierung verwenden, um logische Fehler zu vermeiden.

Beispiel. Test auf einer Krankheit.

- Sei $x = 1\%$ die Anzahl von Leute die krank sind.
- Sei $y = 90\%$ die chancen, dass ein Test korrekt die Krankenheit diagnostiziert.
- Sei $z = 5\%$, die W-keit eine falsch-positive diagnose.

Frage: Falls jemand an die Krankheit positiv getested ist, wie hoch ist die W-keit, dass dar tatsächlich krank ist?

- $A =$ „krank sein", $\mathbb{P}(A) = x = 0.01$
- $B =$ „positiv sein", $\mathbb{P}(B) = ?$

$$\mathbb{P}(B) = \mathbb{P}(B \cap A) + \mathbb{P}(B \cap A^c) = \underbrace{\mathbb{P}(B|A)}_{y}\underbrace{\mathbb{P}(A)}_{x} + \underbrace{\mathbb{P}(B|A^c)}_{z}\underbrace{\mathbb{P}(A^c)}_{1-x}$$

$$= 0.9 \times 0.01 + 0.05 \times 0.99 = 0.0585$$

Bayes-Formel:

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(B|A)\mathbb{P}(A)}{\mathbb{P}(B)} = \frac{x \times y}{x \times y + z \times (1-x)} = \frac{0.009}{0.0585} = 0.015 \approx 15\%.$$

Falls positiv getestet, hat man nur $\approx 15\% \ll 90\%$ W-keit krank zu sein!

Die Bayes-Formel ist die Grundlage eines vollständigen Forschungsfeldes zum automatisierten ``Denken" (Maschinelles Lernen). Siehe z. B. Wikipedia über Bayesian inference.

Was ist **Unabhängigkeit**? Intuitiv, man würde sagen, dass zwei Ereignisse A, B unabhängig sind, falls das eintreten einer keinen Einfluss auf das eintreten der anderen hast. Keine mehr Information.

Wie formalisiert es? Die bedingte Wahrscheinlichkeit erlaubt es uns, das grundlegende Konzept der Unabhängigkeit zu betrachten.

Wir wollen sagen dass A, B sind unabhängig, falls

$$\mathbb{P}_B(A) = \mathbb{P}(A), \qquad \mathbb{P}_A(B) = \mathbb{P}(B)$$

geht.

Definition 2. *Zwei Ereignisse $A, B \in \mathcal{F}$ mit $\mathbb{P}(A) > 0$ und $\mathbb{P}(B) > 0$ heißen unabhängig, falls*

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

Allgemeiner, heißen n Ereignisse $A_1, \ldots, A_n$ (mit $\mathbb{P}(A_k) > 0$ für $k = 1, \ldots, n$) unabhängig, falls $\forall m \leqslant n$, $1 \leqslant i_1 < \cdots < i_m \leqslant n$,

$$\mathbb{P}(\cap_{k=1}^m A_{i_k}) = \Pi_{k=1}^m \mathbb{P}(A_{i_k}).$$

Bemerkung. Ereignisse $A_1, \ldots, A_n$ sind paarweise unabhängig, falls nur

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\mathbb{P}(A_j)$$

für alle i, j gilt. Paarweise unabhängig Ereignisse konnten nicht wirklich unabhängig sein. Denken Sie an ein Gegenbeispiel. (Drei ereignisses sind ausreichend)

Bemerkung 3. Wir können die Begriff von unabhängigkeit auch von Mengensystemen betrachten. Ist dasselbe.

1 Unabhängige Z.V.

Um Unabhängigkeit von Z.V. zu reden, müssen wir schauen wie „viel Information haben".

Jede Zufallsvariable gibt eine bestimmte „Ansicht" über den Wahrscheinlichkeitsraum. Es kann keine Ergebnisse unterscheiden, für die es den gleichen Wert ergibt. Sie generiert also nur eine Teilmenge aller möglichen Ereignisse.

Definition 4. *Sei $(\Omega, \mathcal{F})$ ein Messraum, $X : \Omega \to \mathbb{R}$ eine Z.V.. Definiere $\sigma(X)$ die* ***kleinste*** *unter-σ-Algebra von $\mathcal{F}$ s.d. X bzg. $\sigma(X)$ messbar ist. $\sigma(X)$ heißt die von X erzeugte σ-Algebra.*

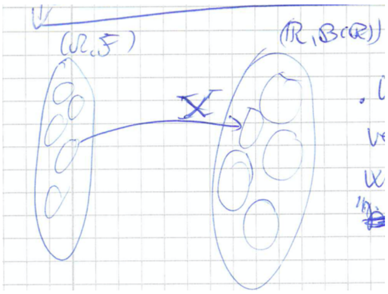
$$\sigma(X) = \cap\{\mathcal{G} \subseteq \mathcal{F} \ \sigma\text{-algebra s.d. } X \ \mathcal{G}\text{-messbar ist}\}.$$

Bemerkung. $\sigma(X)$ ist die kleinste σ-Algebra, die $X^{-1}(B)$, $\forall B \in \mathcal{B}(\mathbb{R})$ enhält. Aber

$$X^{-1}(\mathcal{B}(\mathbb{R})) = \{X^{-1}(B), \forall B \in \mathcal{B}(\mathbb{R})\}$$

ist eine σ-Algebra und $X^{-1}(\mathscr{B}(\mathbb{R})) \subseteq \sigma(X)$, dann offenlich gilt

$$\sigma(X) = X^{-1}(\mathscr{B}(\mathbb{R})).$$



Definition 5. *Sei $\mathscr{G}, \mathscr{G}'$ σ-unteralgebren von $\mathscr{F}$ m, dann sie sind unabhängig, falls $\forall A \in \mathscr{G}, B \in \mathscr{G}'$ (mit $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$) gilt*

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

Definition 6. *Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein W-raum und X_1, X_2 zwei Z.V. X_1 und X_2 heißen unabhängig, falls $\forall A \in \sigma(X_1), B \in \sigma(X_2)$ mit $\mathbb{P}(A) > 0, \mathbb{P}(B) > 0$, gilt*

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

(oder $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$). D.h. X_1, X_2 sind unabhängig falls $\sigma(X_1), \sigma(X_2)$ unabhängig sind.

Bemerkung. Da $\sigma(X) = X^{-1}(\mathscr{B}(\mathbb{R}))$, X_1 und X_2 sind unab. falls $\forall B_1, B_2 \in \mathscr{B}(\mathbb{R})$

$$\mathbb{P}(\{\omega \in \Omega | X_1(\omega) \in B_1, X_2(\omega) \in B_2\}) = \mathbb{P}(\{\omega \in \Omega | X_1(\omega) \in B_1\})\mathbb{P}(\{\omega \in \Omega | X_2(\omega) \in B_2\})$$

Wir schreiben auch kurtz

$$\mathbb{P}(X_1 \in B_1, X_2 \in B_2) = \mathbb{P}(X_1 \in B_1)\mathbb{P}(X_2 \in B_2).$$

Lemma 7. *Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein W-raum, X_1, X_2 unab. Z.V. Seien g_1, g_2 messbare Funktionen von $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ nach $(\mathbb{R}, \mathscr{B}(\mathbb{R}))$ mit*

$$\int_\Omega |g_i(X_i)| \mathrm{d}\mathbb{P} < \infty, \qquad i = 1, 2.$$

Dann,

$$\int_\Omega g_1(X_1) g_2(X_2) \mathrm{d}\mathbb{P} = \Big(\int_\Omega g_1(X_1) \mathrm{d}\mathbb{P}\Big)\Big(\int_\Omega g_2(X_2) \mathrm{d}\mathbb{P}\Big).$$

Beweis. $(g_i(X_i)(\omega) = (g_i \circ X_i)(\omega) = g_i(X_i(\omega))$ die Verkettung) Dann $g_i(X_i) : (\Omega, \sigma(X_i)) \to (\mathbb{R}, \mathscr{B}(\mathbb{R}))$ messbar ist, weil

$$(g_i(X_i))^{-1}(\mathscr{B}(\mathbb{R})) = X_i^{-1}(g_i^{-1}(\mathscr{B}(\mathbb{R}))) \subseteq X_i^{-1}(\mathscr{B}(\mathbb{R})) = \sigma(X_i) \subseteq \mathscr{F}.$$

[Schritt 1] Seien g_i Indikatorfunktionen $g_i = \mathbb{1}_{A_i}$, $A_i \in \mathscr{B}(\mathbb{R})$. Für $A_i \in \mathscr{B}(\mathbb{R})$:

$$\int_\Omega \mathbb{1}_{A_i}(X_i(\omega)) \mathrm{d}\mathbb{P}(\omega) = \int_\Omega \mathbb{1}_{X_i \in A_i}(\omega) \mathrm{d}\mathbb{P}(\omega) = \mathbb{P}(X_i \in A_i), \quad i = 1, 2.$$

und

$$\int_\Omega \mathbb{1}_{A_1}(X_1)\mathbb{1}_{A_2}(X_2) \mathrm{d}\mathbb{P} = \mathbb{P}(X_1 \in A_1, X_2 \in A_2) \underset{\text{unab.}}{=} \mathbb{P}(X_1 \in A_1)\mathbb{P}(X_2 \in A_2) = \int_\Omega \mathbb{1}_{A_1}(X_1) \mathrm{d}\mathbb{P} \int_\Omega \mathbb{1}_{A_2}(X_2) \mathrm{d}\mathbb{P}.$$

[Schritt 2] Wegen Linearität, gilt fur alle einfache Funktionen g_i.

[Schritt 3] Seien nun $g_i \geqslant 0$, wählen wir $(h_n^{(i)})_n$ s.d. $h_n^{(i)} \nearrow g_i$ punktweise (d.h. $h_n^{(i)} \leqslant h_{n+1}^{(i)}$, $h_n^{(i)} \leqslant g_i$ und $h_n^{(i)} \to g_i$ punktweise). Dann Ergebnis folgt aus Monotone Konvergenz.

[Schritt 4] Für allgemeine g_i: $g_i = (g_i)_+ - (g_i)_-$ (zerlegung) und Linearität. □

Bemerkung 8. Die Aussage gilt „Umgekehrt". Falls für alle messbare Funktionen g_1, g_2

$$\int_\Omega g_1(X_1) g_2(X_2) \mathrm{d}\mathbb{P} = \left(\int_\Omega g_1(X_1) \mathrm{d}\mathbb{P}\right)\left(\int_\Omega g_2(X_2) \mathrm{d}\mathbb{P}\right)$$

dann sind X_1, X_2 unabhängig. (Klar!)

Definition 9. *Sei $(\Omega, \mathscr{F}, \mathbb{P})$ ein W-raum. X_1, X_2 Z.V. heißen <u>unkorreliert</u>, falls*

$$\int X_1 X_2 \mathrm{d}\mathbb{P} = \int X_1 \mathrm{d}\mathbb{P} \int X_2 \mathrm{d}\mathbb{P}.$$

Bemerkung. Unabhängigkeit $\Rightarrow$ Unkorreliertheit. ($\not\Leftarrow$ nur für Gauß'sche **Zufallsvektoren**, siehe später)

Notierung.

$$\mathbb{E}(X) := \int X \mathrm{d}\mathbb{P}, \qquad \textit{(Erwartungswert)}$$

$$\mathrm{Var}(X) := \mathbb{E}[(X - \mathbb{E}(X))^2], \qquad \textit{(Varianz)}$$

$$\mathrm{Cov}(X, Y) := \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))]. \qquad \textit{(Covarianz)}$$

$$X, Y \text{ Unkorreliert} \Leftrightarrow \mathrm{Cov}(X, Y) = 0.$$

2 Produkträume

Wir wollen schauen wie man unabhängige Z.V. konstruiren kann.

Seien $(\Omega_1, \mathscr{F}_1, \mathbb{P}_1)$ und $(\Omega_2, \mathscr{F}_2, \mathbb{P}_2)$ zwei W-raume, und $X_1 : \Omega_1 \to \mathbb{R}$, $X_2 : \Omega_2 \to \mathbb{R}$ messbare Funktionen. (d.h. Z.V.)

<u>Ziel</u>: Konstruiere $(\Omega, \mathscr{F}, \mathbb{P})$ und $\hat{X}_1 : \Omega \to \mathbb{R}$ und $\hat{X}_2 : \Omega \to \mathbb{R}$ Z.V. auf $(\Omega, \mathscr{F})$, s.d. diese Z.V. unabhängig sind bzg. $\mathbb{P}$ und s.d. $\hat{X}_i$ und X_i gleich verteilt sind für $i = 1, 2$. Anders gesagt

$$\mathbb{P}(\hat{X}_1 \in A, \hat{X}_2 \in B) = \mathbb{P}_1(X_1 \in A) \mathbb{P}_2(X_2 \in B), \qquad A, B \in \mathscr{B}(\mathbb{R}).$$

Definition 10.

- $$\Omega = \Omega_1 \times \Omega_2 = \{\omega = (\omega_1, \omega_2) | \omega_1 \in \Omega_1, \omega_2 \in \Omega_2\}$$

  heißt das <u>Produktraum</u> von Ω_1 und Ω_2.

- *$\mathscr{F} = \mathscr{F}_1 \otimes \mathscr{F}_2$ ist die kleinste σ-Algebra, die alle Menge der Form (Rechtecke)*

  $$C = A \times B = \{\omega = (\omega_1, \omega_2) | \omega_1 \in A, \omega_2 \in B\} \subseteq \Omega$$

  $A \in \mathscr{F}_1$, $B \in \mathscr{F}_2$ enthält, d.h.

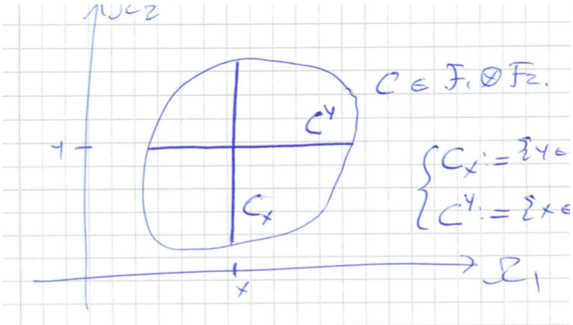
  $$\mathscr{F}_1 \otimes \mathscr{F}_2 = \sigma(\mathscr{F}_1 \times \mathscr{F}_2), \qquad \mathscr{F}_1 \times \mathscr{F}_2 = \{A \times B : A \in \mathscr{F}_1, B \in \mathscr{F}_2\} \subseteq \mathscr{P}(\Omega_1 \times \Omega_2)$$

  $\mathscr{F}_1 \otimes \mathscr{F}_2$ heißt die <u>Produkt σ-Algebra</u> von $\mathscr{F}_1$ und $\mathscr{F}_2$.

Bemerkung. $\mathscr{F}_1 \times \mathscr{F}_2 = \{C = A \times B\}$ ist durchschnittstabil, weil

$$(A_1 \times B_1) \cap (A_2 \times B_2) = (A_1 \cap A_2) \times (B_1 \cap B_2).$$

Wir müssen $\mathbb{P}$ auf $(\Omega_1 \times \Omega_2, \mathscr{F}_1 \otimes \mathscr{F}_2)$ definieren, s.d. „$\mathscr{F}_1 \otimes \Omega_2$“ und „$\Omega_1 \otimes \mathscr{F}_2$“ unabhängig sind.



Falls $C \in \mathscr{F}_1 \otimes \mathscr{F}_2$ und $f: \Omega \to \mathbb{R}$ dann $C_x \subseteq \Omega_2$, $C^y \subseteq \Omega_1$, $f_x: \Omega_2 \to \mathbb{R}$, $f^y: \Omega_1 \to \mathbb{R}$,

$$C_x := \{y \in \Omega_2 | (x,y) \in C\}, \qquad x \in \Omega_1, \qquad\qquad C^y := \{x \in \Omega_1 | (x,y) \in C\}, \qquad y \in \Omega_2.$$

$$f_x(y) := f(x,y), \qquad x \in \Omega_1 \qquad\qquad f^y(x) := f(x,y), \qquad y \in \Omega_2$$

Lemma 11. *Es gilt*

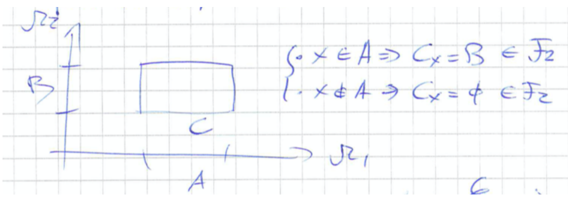
a) $\forall C \in \mathscr{F}_1 \otimes \mathscr{F}_2$, $x \in \Omega_1$, $y \in \Omega_2$ *dann,* $C_x \in \mathscr{F}_2$, $C^y \in \mathscr{F}_1$

b) $\forall f: \Omega_1 \times \Omega_2 \to \mathbb{R}$ *messbar dann,* $\forall x \in \Omega_1, y \in \Omega_2$ $f_x: (\Omega_2, \mathscr{F}_2) \to \mathbb{R}$, $f^y: (\Omega_1, \mathscr{F}_1) \to \mathbb{R}$ *messbar sind.*

Beweis. Sei $x \in \Omega_1$, z.z. $C_x \in \mathscr{F}_2$. (a) Sei

$$\mathscr{C}_x = \{C \in \mathscr{F}_1 \otimes \mathscr{F}_2 | C_x \in \mathscr{F}_2\} \subseteq \mathscr{F}_1 \otimes \mathscr{F}_2$$

Wir muss zeigen dass $\mathscr{C}_x = \mathscr{F}_1 \otimes \mathscr{F}_2$. Betrachten wir die Mengen $C = A \times B$ für $A \in \mathscr{F}_1$ und $B \in \mathscr{F}_2$. Diese sind in $\mathscr{C}_x$:



Es ist leicht zu prüfen, dass $\mathscr{C}_x$ eine σ-Algebra ist (Übung!).

$$\mathscr{F}_1 \otimes \mathscr{F}_2 \underset{\text{def.}}{=} \sigma(\{A \times B : A \in \mathscr{F}_1, B \in \mathscr{F}_2\}) \subseteq \mathscr{C}_x \underset{\text{def.}}{\subseteq} \mathscr{F}_1 \otimes \mathscr{F}_2 \Rightarrow \mathscr{C}_x = \mathscr{F}_1 \otimes \mathscr{F}_2.$$

(b) Sei $M \in \mathscr{B}(\mathbb{R})$, dann

$$f_x^{-1}(M) = \{y \in \Omega_2 | f_x(y) \in M\} = \{y \in \Omega_2 | f(x,y) \in M\}$$

$$= \{y \in \Omega_2 | (x,y) \in f^{-1}(M)\} = (f^{-1}(M))_x \underset{\text{Teil (a)}}{\in} \mathscr{F}_2$$

wegen (a). Denn f_x ist $\mathscr{F}_2$-messbar. $\square$

Satz 12. *Seien* $\mathbb{P}_1, \mathbb{P}_2$ *W-masse auf* $(\Omega_1, \mathscr{F}_1)$ *und* $(\Omega_2, \mathscr{F}_2)$

a) $\exists!\ \mathbb{P} = \mathbb{P}_1 \otimes \mathbb{P}_2$ *W-masse auf* $(\Omega, \mathscr{F}_1 \otimes \mathscr{F}_2)$ *s.d.*

$$(\mathbb{P}_1 \otimes \mathbb{P}_2)(A \times B) = \mathbb{P}_1(A)\mathbb{P}_2(B) \qquad \forall A \in \mathscr{F}_1, B \in \mathscr{F}_2.$$

b) *Falls* $C \in \mathscr{F}_1 \otimes \mathscr{F}_2$, *dann*

$$\mathbb{P}_1 \otimes \mathbb{P}_2(C) = \int_{\Omega_1} \mathbb{P}_2(C_x) \mathrm{d}\mathbb{P}_1(x) = \int_{\Omega_2} \mathbb{P}_1(C^y) \mathrm{d}\mathbb{P}_2(y).$$

(I forgot to record the second part of the lecture…)
