# Wahrscheinlichkeitstheorie für Informatikstudien (Kapitel #8)

Siegfried Hörmann

TU Graz

WS 2020/21

# Kursaufbau

1. Wahrscheinlichkeitsräume.
2. Laplace-Experimente.
3. Wichtige diskrete Verteilungen.
4. Zufallsvariablen.
5. Zufallsvektoren.
6. Momente.
7. Bedingte Wahrscheinlichkeiten.
8. Unabhängigkeit.
9. Konvergenz von Zufallsvariablen.
10. Der zentrale Grenzwertsatz.

# Unabhängigkeit

8.1. Unabhängigkeit von Ereignissen.

8.2. Unabhängigkeit von Zufallsvariablen.

8.3. Unabhängigkeit und Momente.

## Unabhängigkeit von Ereignissen

Wir kommen nun zu einem der wichtigsten Konzepte in der Wahrscheinlichkeitstheorie (und ganz besonders wichtig in Statistik!).

Sei $(\Omega, \mathcal{A}, P)$ ein W-Raum. Wir bezeichnen zwei Ereignisse $A$ und $B$ in $\mathcal{A}$ als unabhängig, falls $P(A \cap B) = P(A)P(B)$.

Kurz: $A \perp\!\!\!\perp B$.

Mit bedingten Wahrscheinlichkeiten kann man dies auch so ausdrücken $P(A|B) = \frac{P(A \cap B)}{P(B)} = P(A)$.

Die Information des Ereignisses $B$ hat keinen Einfluss auf die Wahrscheinlichkeit von $A$.

## Unabhängigkeit von Ereignissen

Unabhängigkeit ist wiederum in einem Laplace-Modell leicht zu veranschaulichen, in welchem wir 2 Experimente durchführen.

Wir führen 2 (unabhängige) Laplace Experimente durch. Wie groß ist die Wahrscheinlichkeit, dass $A$ und $B$ eintreten?

Es gibt $|A| \times |B|$ Möglichkeiten dafür und gesamt gibt es $|\Omega_1| \times |\Omega_2|$ Kombinationen. Daher

$$P(A \, \text{und} \, B) = \frac{|A| \times |B|}{|\Omega_1| \times |\Omega_2|} = \frac{|A|}{|\Omega_1|} \times \frac{|B|}{|\Omega_2|}.$$

$\Longrightarrow$ wir multiplizieren also die Wahrscheinlichkeiten.

# Unabhängigkeit von Ereignissen

Unabhängigkeit lässt sich auch auf mehr als 2 Ereignisse verallgemeinern.

Unabhängigkeit: Wir nennen $A_1, \ldots, A_n$ unabhängig, falls

$$(*) \quad P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i) \quad \text{für alle } \emptyset \neq I \subset \{1, \ldots, n\}.$$

Das läßt sich für unendliche Folgen verallgemeinern. Dann muss $(*)$ für alle endlichen Teilmengen $I$ der Indexmenge halten.

# Unabhängigkeit von Ereignissen

**Lemma:** Ist $A$ von $B$ unabhängig, dann ist auch $A$ von $B^C$ unabhängig.

## Beispiel

Wir spielen Roulette. Wir setzen in 3 Runden jeweils 5 €. Im ersten Spiel setzen wir auf rot, im zweiten setzen wir auf $1 - 12$ und im letzten Spiel setzen wir auf die Zahl 23. Wie groß ist die Wahrscheinlichkeit, dass wir mindestens 10 € gewinnen?

# Unabhängigkeit von Zufallsvariablen

Das Konzept der Unabhängigkeit für Ereignisse lässt sich auf Unabhängigkeit von Zufallsvariablen erweitern.

Zwei Zufallsvariablen $X$ and $Y$ heißen unabhängig, falls

$$P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$$

für alle infrage kommenden Mengen $A$ und $B$ (z.B. Intervalle).

Man kann zeigen, dass dies genau dann der Fall ist, wenn

$$\begin{aligned} F^{XY}(x, y) &= P(X \leq x, Y \leq y) \\ &= P(X \leq x)P(Y \leq y) = F^X(x)F^Y(y). \end{aligned}$$

## Unabhängigkeit von Zufallsvariablen

Beachte: üblicherweise versuchen wir nicht herauszufinden ob zwei ZV unabhängig sind. Unabhängigkeit ist eine nützliche und oft realistische Modellannahme. Dies ist durch das Laplace Beispiel motiviert, welches unsere intuitive Auffassung von Unabhängigkeit untermauert. Wir können das Konzept auch auf mehrere ZV erweitern:

### Definition (Unabhängigkeit von Zufallsvariablen)

Wir bezeichnen $X_1, \ldots, X_n$ als unabhängig falls

$$F^{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \prod_{i=1}^{n} F^{X_i}(x_i) \quad \text{für alle } x_i \in \mathbb{R} \text{ und alle } n \geq 1.$$

Haben alle $X_i$ dieselbe Verteilung, nennen wir $(X_i)$ unabhängig und identisch verteilt (i.i.d.).

### Beispiel

Ein Empfänger bekommt Signale von 3 Sendestationen. Die Signalstärke $X_j$ der $j$-ten Station folgt einer Verteilung $F_j$, also $P(X_j \leq x) = F_j(x) = 1 - e^{-\lambda_j x}$, $x \geq 0$. Dabei ist $\lambda_1 = 1.3$, $\lambda_2 = 2.1$ und $\lambda_3 = 0.8$. Man kann annehmen, dass die Signalstärken der Stationen voneinander unabhängig sind. Um das Signal korrekt zu verarbeiten, braucht der Empfänger von mindestens einem Sender eine Signalstärke $> 0.5$. Wie groß ist die Wahrscheinlichkeit, dass das Signal korrekt verarbeitet wird?

# Unabhängigkeit von Zufallsvariablen

## Beispiel

An einem Ort ist die maximale 3-stündige Niederschlagsmenge innerhalb eines Jahres Weibull-verteilt, d.h. die CDF ist gleich

$$F(x) = 1 - e^{-(\lambda(x-\mu))^k}, \quad x \geq \mu.$$

Angenommen wir wüßten, dass $\lambda = 1/10$, $k = 2$ und $\mu = 15$. Weiters sind diese extremen Niederschlagsmengen voneinander unabhängig. Wir suchen einen Wert $z$ derart, dass die Wahrscheinlichkeit innerhalb der nächsten 100 Jahren $z$ zu überschreiten, kleiner als 5% ist.

# Unabhängigkeit von Zufallsvariablen

## Proposition

*Zufallsvariablen $X_1, \ldots, X_n$ sind unabhängig genau dann, wenn gilt"*

- $$f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f^{X_i}(x_i),$$

  *falls eine Dichte existiert.*

- $$p_{k_1,\ldots,k_n} = \prod_{i=1}^{n} p_{k_i}^{X_i},$$

  *falls eine PMF existiert.*

## Unabhängigkeit von Zufallsvariablen

### Proposition

*Sind die Zufallsvariablen $X_n$ unabhängig, so sind auch Funktionen $g_n(X_n)$ unabhängig.*

# Unabhängigkeit von Zufallsvariablen

Mit dem Konzept der Unabhängigkeit können wir klassische Modelle auf eine andere, oft leichtere Art erhalten.

## Beispiel (Binomialmodell)

Seien $X_1, \ldots, X_n$ unabhängige ZVen mit $P(X_i = 1) = p$ und $P(X_i = 0) = 1 - p$.
Sei $S_n = X_1 + \cdots + X_n$. Was ist dann $P(S_n = k)$?

### Beispiel

Seien $X_1, X_2, \ldots$ iid ZVen, die die Werte 0 und 1 annehmen. Angenommen $P(X_i = 1) = p$ und $P(X_i = 0) = 1-p$. Sei $T$ der erste Zeitpunkt in welchem die Zahl 1 vorkommt. Bestimme $P(T = k)$.

## Unabhängigkeit und Momente

### Proposition

*Wenn $X$ und $Y$ unabhängig sind, und jeweils in $L^1$ (also Erwartungswerte existieren), dann gilt, dass auch $XY$ in $L^1$ und dann ist*

$$E(XY) = E(X)E(Y).$$

*Das lässt sich auch auf mehr als 2 ZVen verallgemeintern.*

### Beweis.

Nehmen wir an der Vektor $(X, Y)$ hat eine Dichte $f(x, y)$.

$$\begin{aligned} E(XY) &= \int \int xyf(x,y)dxdy \\ &= \int \int xyf^X(x)f^Y(y)dxdy \\ &= \int xf^X(x)dx \int yf^Y(y)dy. \end{aligned}$$

$\square$

## Unabhängigkeit und Momente

Ein wichtiges Anwendungsbeispiel sind momenterzeugende Funktionen.

Sei $X$ eine ZV. Wir betrachten die Funktion

$$M^X(s) = Ee^{sX},$$

sofern diese existiert. (Der Erwartungswert muss nicht für alle $s$ endlich sein.)

Dann heißt $M^X(s)$ momenterzeugende Funktion von $X$.

Man kann zeigen: wenn $M^X(s)$ für alle $s$ in einer Umgebung von 0 existiert, ist die Verteilung

$$F^X(s) = EI\{X \leq s\} = P(X \leq s), \quad s \in \mathbb{R}$$

eindeutig festlegt.

## Unabhängigkeit und Momente

Insbesondere gilt für eine Summe von unabhängigen ZVen $X_1, \ldots, X_n$, dass

$$\begin{aligned} M^{X_1+\cdots+X_n}(s) &= E(e^{s(X_1+\cdots+X_n)}) \\ &= E(e^{sX_1}\cdots e^{sX_n}) \\ &= E(e^{sX_1})\cdots E(e^{sX_n}) \\ &= M^{X_1}(s)\cdots M^{X_n}(s). \end{aligned}$$

Das werden wir später (beim zentralen Grenzwertsatz) sehr gut verwenden können.

## Unabhängigkeit und Momente

Zum Begriff: Beachte, dass

$$e^{sX} = 1 + sX + \frac{s^2X^2}{2!} + \frac{s^3X^3}{3!} + \cdots$$

Also (wegen der Linearität des Erwartungswertes)

$$M^X(s) = 1 + sE(X) + \frac{s^2}{2}E(X^2) + \frac{s^3}{3!}E(X^3) + \cdots$$

Wenn man nun $k$ Mal nach $s$ ableitet und diese Funktion in 0 bestimmt, ergibt sich:

$$\frac{d^k}{ds^k}M^X(s)|_{s=0} = E(X^k).$$

Erinnerung: $E(X^k)$ ist das $k$-te Moment von $X$.