Überblick

- Beschreibende Statistik:
  Auswertung von Experimenten und Stichproben
- Wahrscheinlichkeitsrechnung:
  Schlüsse aus gegebenen Wahrscheinlichkeiten, Hilfsmittel: Kombinatorik
- Beurteilende Statistik:
  Schlüsse aus Experimenten, Beurteilung von exp. Ergebnissen (machen wir nicht)

Linguistische Anwendungen:

- Spracherkennung
- Textretrival
- probabilistische Grammatiken: z.B. Disambiguierung

Problem: woher Daten?

Beschreibende Statistik

Statistische Erhebung: Bestimmung der *Ausprägung* eines *Merkmals* bei allen Individuen einer Grundgesamtheit.

qualitative vs. quantitative Merkmale,

diskrete vs. stetige qualitative Merkmale...

Beispiele:

- Geschlecht/ Gewicht aller Neugeborenen an einem Tag in einem Krankenhaus
- Anzahl der Wörter in jedem Artikel einer Ausgabe einer Tageszeitung
- Anzahl des Vorkommens von bestimmten Wörtern in einem Text-Korpus

Begriffe

- *absolute Häufigkeit*:

  Anzahl des Vorkommens einer Ausprägung.

- *relative Häufigkeit*:

  absolute H. / Anzahl der Individuen

- *Häufigkeitsverteilung*:

  Funktion von allen Ausprägungen eines Merkmals auf Häufigkeiten.

- *Zentralwert*:

  Bedingung: Ausprägungen geordnet. Der Zentralwert ist diejenige Ausprägung, für die gilt: es liegen nicht mehr als die Hälfte der Erhebungswerte darunter oder darüber.

- *arithmetisches Mittel* $\overline{x}$ von Erhebungswerten $x_1, .., x_n$:

  Bedingung: quantitatives Merkmal.

  $\overline{x} = \frac{1}{n}(x_1 + .. + x_n) = \frac{1}{n} \Sigma_{i=1}^{n} x_i$

- *Varianz, Streuung* $\overline{s}^2$ (mittlere quadratische Abweichung):

  $\overline{s}^2 = \frac{1}{n}((x_1 - \overline{x})^2 + .. + (x_n - \overline{x})^2) = \frac{1}{n} \Sigma_{i=1}^{n} (x_i - \overline{x})^2$

- *Standardabweichung*: Quadratwurzel aus Varianz

Zufallsexperimente

Zufallsvariable X : unsicherer Ausgang eines Zufallsexperiments mit endlicher Zahl möglicher Ausgänge $E_1, ..E_k$, Ausgangsmenge oder Ereignisraum $V(X)$.

Bsp: Werfen einer Münze, Ziehung der Lottozahlen, Alter des nächsten Passanten.

Jede Teilmenge von $V(X)$ heisst *Ereignis*, die einzelnen Elemente auch Elementarereignisse.

Das Komplement eines Ereignisses A heisst Gegenereignis $\overline{A}$.

relative Häufigkeit eines Ausgangs, $h(E)$:

$\#$ Eintreten von E $/$ $\#$ Versuche.

Bemerkung zum Übergang zur Wahrscheinlichkeitsrechnung: es gelten gleiche Gesetzmässigkeiten, aber W'keitsrechnung lässt sich nicht statistisch begründen. Daher axiomatische Einführung mit gegebenen *Wahrscheinlichkeiten* der Elementarereignisse.

Axiome der Wahrscheinlichkeit (Kolmogoroff)

Wahrscheinlichkeit:

Sei $\{E_1, ..., E_k\}$ ein Ereignisraum mit den Elementar-ereignissen E_i.

Wahrscheinlichkeitsverteilung:

Funktion $P : \{E_1, ..., E_k\} \to [0, 1]$ (1)

mit $P(E_1) + P(E_2) + ... + P(E_k) = 1,$ (2)

$P(E_i)$ heisst *Wahrscheinlichkeit* von E_i.

Sei A ein Ereignis mit Ereignisraum wie oben.

Wahrscheinlichkeit von A:

$P(A) = P(E_1) + .. + P(E_i)$, falls $A = E_1 \cup .. \cup E_i$;

$P(A) = 0$, falls $A = \emptyset$ (3)

Folgerungen daraus: für alle Ereignisse A, B gilt:

$0 \leq P(A) \leq 1$

$P(\overline{A}) = 1 - P(A)$

$A \subset B \Rightarrow P(A) < P(B)$

$A \cap B = \emptyset \Rightarrow P(A \cup B) = P(A) + P(B)$

(A und B heissen *unvereinbar*)

Gleichverteilung

Gleichverteilung: W'keitsverteilung, bei der alle Elementarereignisse die gleiche W'keit haben.

Zufallsexperimente mit Gleichverteilung heissen *Laplace-Experimente*.

Für Laplace-Experimente gilt für Ereignis A :

$$P(A) = \frac{\text{Anzahl der günstigen Ausgänge}}{\text{Anzahl der möglichen Ausgänge}}$$

Beispiele:

- X_1: Augenzahl bei Wurf eines fairen (idealen) Würfels.

  $V(X_1) = \{1, 2, 3, 4, 5, 6\}, P(E_i) = 1/6$
- X_2: Augenzahl bei Wurf von zwei fairen Würfeln gleichzeitig.

  $V(X_2) = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\}$

  $P(2) = 2/36, P(3) = 4/36, \ldots$

Kombinatorik

Produktregel zur Bestimmung möglicher Kombinationen:

Sei folgendes Lexikon gegeben:

$\{$ die, keine, schönen, grünen, schnellen, Hunde, Katzen, Mäuse $\}$.

Wieviele NPs lassen sich unter Verwendung der Regel NP $\rightarrow$Det Adj N daraus bilden?

NP = # Det · # Adj · # N.

- Geordnete Stichprobe mit Zurücklegen, k-mal eines von n Elementen ziehen:

  n^k Möglichkeiten.
- Geordnete Stichprobe ohne Zurücklegen:

  $n \cdot (n-1)... \cdot (n-k+1) = \frac{n!}{(n-k)!}$
- Geordnete Vollerhebung (Permutation): $n!$
- Ungeordnete Stichprobe ohne Zurücklegen (Bsp: Lottozahlen):

  Binomialkoeffizient, ”n über k”: $\begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k!(n-k)!}$

Bedingte Wahrscheinlichkeiten

$P(A \cup B) = P(A) + P(B)$ gilt nur wenn $A \cap B = \emptyset$.

Sonst: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Wie kann man $P(A \cap B)$ bestimmen?

Bedingte relative Häufigkeit:

n Durchführungen, $n_1 > 0$ mal Ereignis A, davon k-mal auch B, dann ist k/n_1 die *relative Häufigkeit von* B *bezüglich* A, $h_A(B)$, auch $h(B \mid A)$.

$$h(A \cap B) = h(A) \cdot h_A(B)$$

Bedingte Wahrscheinlichkeit:

Gegeben: Ereignisse A und B, $P(A) \neq 0$, dann heisst

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

die *durch* A *bedingte Wahrscheinlichkeit von* B oder Wahrscheinlichkeit von B bezgl. A.

Allgemeiner Multiplikationssatz:

$P(A) \neq 0,$ dann $P(A \cap B) = P(A) \cdot P_A(B)$

Beispiel

Wenn sich jemand noch genau erinnert, dass eines der beiden Kinder seiner Cousine ein Junge ist, wie gross ist dann die Wahrscheinlichkeit, dass beides Jungen sind? (P(Junge) = P(Mädchen) = 0.5).

Gesucht: $P_{1\ \text{Junge}}(2\ \text{Jungen})$

Lösung

Ohne Information:

P(2 Jungen) = 1/4

P(2 Mädchen)= 1/4

P(Junge/Mädchen) = 1/2

P(A) = P(1 Junge) = 1/4 + 1/2 = 3/4

P(B) = P(2 Jungen), $P(A \cap B) = P(B)$

Mit der Information, dass ein Kind ein Junge ist:

$P_{1\ \text{Junge}}(2\ \text{Jungen}) = P(2\ \text{Jungen}) / P(1\ \text{Junge})$

= 1/4 / 3/4 = 1/3

Weiter: bedingte W'keiten

Seien A und B Ereignisse mit $P(A) \neq 0, P(B) \neq 0$.

Dann gilt:

$$P_B(A) = P_A(B) \cdot P(A)/P(B).$$

verallgemeinert, Satz von Bayes:

Seien $A_1, ...A_n$ Ereignisse, die den Ereignisraum zerlegen, d.h. $P(A_1) + .. + P(A_n) = 1$, und $P(A_j) \geq 0$ für $1 \leq j \leq n$. Sei $P(B) > 0$.

Dann gilt für A_i mit $1 \leq i \leq n$:

$$P_B(A_i) = \frac{P(A_i) \cdot P_{A_i}(B)}{P(A_1) \cdot P_{A_1}(B) + .. + P(A_n) \cdot P_{A_n}(B)}$$

Zwei Ereignisse heissen *unabhängig*, wenn gilt: $P_A(B) = P(B)$ (und $P_B(A) = P(A)$).

Spezieller Multiplikationssatz:

Sind A und B unabhängig, dann gilt:

$$P(A \cup B) = P(A) \cdot P(B)$$

bedingte W'keiten, linguistisch

Wortfolgen:

$$P(w_{1,n}) = P(w_1)P(w_2 \mid w_1)P(w_3 \mid w_1, w_2)...P(w_n \mid w_{1,n-1})$$

$$\text{der} \left\{ \begin{array}{c} \text{kleine} \\ \text{Schweine} \end{array} \right\} \text{Hund}$$

Sei P(kleine $\mid$ der) = P(Schweine $\mid$ der).

P$_1$ = P(der kleine Hund) =
P(der) P(kleine $\mid$ der) P(Hund $\mid$ der kleine)

P$_2$ = P(der Schweine Hund) =
P(der) P(Schweine $\mid$ der) P(Hund $\mid$ der Schweine)

P$_1$ $>$ P$_2$,
falls P(Hund $\mid$ der kleine) $>$ P(Hund $\mid$ der Schweine)