

Übung 1

Aufgabe 1

Formulieren Sie folgende Ausdrücke mathematisch:

- die Menge der Student*innen, die jünger als 23 Jahre sind
- das durchschnittliche Alter der Student*innen
- das Alter der/des ältesten Studierenden
- der/die älteste Studierende

Dabei sei:

S : die Menge der Student*innen

$|S|$: die Größe der Menge S (= Zahl der Student*innen)

$alter(x)$: das Alter der Person x

$\{x|cond(x)\}$: Menge aller x , welche die Bedingung $cond(x)$ erfüllen

$\sum_{x \in X} f(x)$: Summe der Werte der Funktion f für alle x aus der Menge X

$\max_{x \in X} f(x)$: der maximale Wert der Funktion f

$\arg \max_{x \in X} f(x)$: dasjenige x mit dem größten Funktionswert $f(x)$

Aufgabe 2

Angenommen Sie finden in einem großen englischen Korpus das Wort *New* 1712 Mal, das Wort *York* 911 Mal und das Wortpaar *New York* 870 Mal.

Schätzen Sie (1) die bedingte Wahrscheinlichkeit von *York* nach *New* und (2) die Wahrscheinlichkeit von *New* vor *York*.

Aufgabe 3

Zeigen Sie, dass die folgenden Gleichungen gelten:

- $p(x|y) = \frac{p(y|x)p(x)}{p(y)}$ (aus $p(x|y) = \frac{p(x,y)}{p(y)}$)
- $Var(X) = E((X - E(X))^2) = E(X^2) - E(X)^2$
- $H(X, Y) = H(X) + H(Y|X)$
- $H(Y) - H(Y|X) = I(X; Y)$

- $H(X, Y) = H(X) + H(Y) - I(X; Y)$
- $I(X; Y) = D(p(x, y) \parallel p(x)p(y))$

Hierbei ist $E(X)$ der Erwartungswert der Zufallsvariablen X , $Var(X)$ ihre Varianz und $H(X)$ ihre Entropie. $I(X, Y)$ ist die Mutual Information zwischen den Zufallsvariablen X und Y . $D(p, q)$ ist der Kullback-Leibler-Abstand der Wahrscheinlichkeitsverteilungen p und q .