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:

$S = \{ \dots \}$

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 $f(x)$

- $\{x \in S \mid alter(x) < 23\}$
- $Avg. = \frac{total}{num} = \frac{\sum_{x \in S} alter(x)}{|S|}$
- $\max_{x \in S} alter(x)$

  $\max(alter(s_1), alter(s_2), \dots, alter(s_n))$

  $\arg\max(alter(s_1), alter(s_2), \dots, alter(s_n))$

- $\arg\max_{x \in S} alter(x)$

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*.

(1) $$P(\text{York} \mid \text{New}) = \frac{P(\text{New}, \text{York})}{P(\text{New})} = \frac{870}{1712}$$

(2) $$P(\text{New} \mid \text{York}) = \frac{P(\text{New}, \text{York})}{P(\text{York})} = \frac{870}{911}$$

Aufgabe 3

Zeigen Sie, dass die folgenden Gleichungen gelten:

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

Hierbei is $E(X)$ 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.

- $P(x|y) = \frac{P(y|x) \cdot P(x)}{P(y)}$

$P(x|y) = \frac{P(x,y)}{P(y)} \Rightarrow P(x|y) \cdot P(y) = P(x,y)$ (1)

$P(y|x) = \frac{P(y,x)}{P(x)} \Rightarrow P(y|x) \cdot P(x) = P(y,x)$ (2)

$\Rightarrow P(x|y) \cdot P(y) = P(y|x) \cdot P(x)$

$\Rightarrow P(x|y) = \frac{P(y|x) \cdot P(x)}{P(y)}$

- $Var(X) = E((X - E(X))^2) = E(X^2) - (E(X))^2$

$(a-b)^2 = a^2 - 2ab + b^2$

$E((X - E(X))^2)$

$= E(X^2 - 2 \cdot E(X) \cdot X + E(X)^2)$

$= E(X^2) - E(2E(X) \cdot X) + E(X)^2$

a : constant

X : Variable (random)

$$= E(X^2) - 2E(X)\cdot E(X) + E(X)^2$$

$$= E(X^2) - 2E(X)^2 + E(X)^2$$

$$= E(X^2) - E(X)^2$$

$E(a) = a$

$E(X+a) = E(X) + E(a) = E(X)+a$

$E(aX) = a\cdot E(X)$

- $H(X,Y) = H(X) + H(Y|X)$

$$-\sum_{x\in\Omega_x}\sum_{y\in\Omega_y} p(x,y)\log_2 p(x,y)$$

$$-\sum_{x\in\Omega_x}\sum_{y\in\Omega_y} p(x,y)\log_2 p(y|x)$$

$$-\sum_{x\in X} p(y)\log p(y)$$

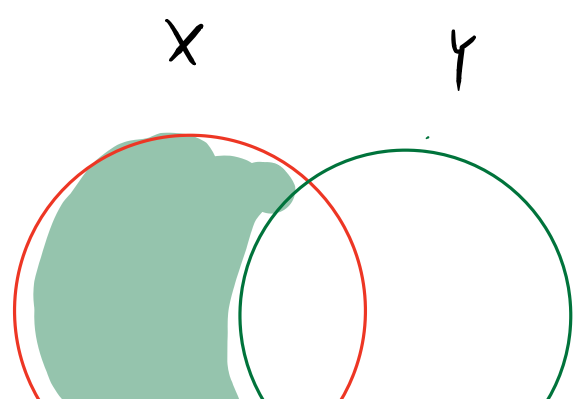
$$I(X;Y) = \sum_{x\in\Omega_x}\sum_{y\in\Omega_y} p(x,y)\log_2\frac{p(x,y)}{p(x)p(y)}$$

$$= \underbrace{\sum_{x\in X}\sum_{y\in Y} p(x,y)\log_2\frac{p(x,y)}{p(x)}}_{p(x|y)} - \underbrace{\sum_{x\in X,\, y\in Y} p(x,y)\log_2 p(y)}_{H(Y)}$$

$$= \sum_{x\in X}\sum_{y\in Y} p(x,y)\log_2\frac{p(x)}{p(x,y)}$$

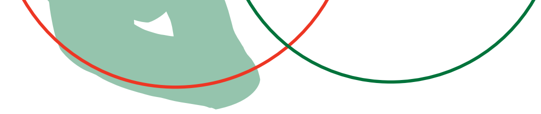
$H(X|Y)$

$$= -H(X|Y) + H(Y)$$



$H(X,Y)$

$H(X,Y)$



$H(X)$ $H(Y)$

$I(X;Y)$

$$I(X;Y) = D(p(x,y) \,\|\, p(x)p(y))$$

$$\sum p(x,y) \log_2 \frac{p(x,y)}{p(x)p(y)}$$