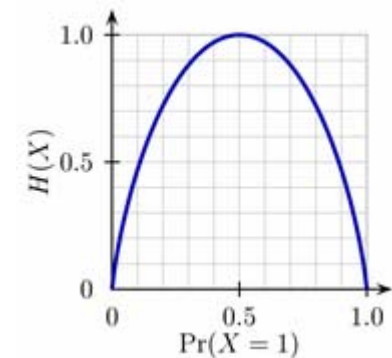


# Information und Kommunikation

Hartmut Klauck  
Universität Frankfurt  
SS 07  
20.4.



# Gemeinsame Entropie

- Gegeben seien zwei Zufallsvariablen, die nicht unabhängig sein müssen
- $X$  auf  $1, \dots, n$  und  $Y$  auf  $1, \dots, m$  mit einer gemeinsamen Verteilung  $p(i, j)$
- $p_X(i) = \sum_{j=1, \dots, m} p(i, j)$  bzw.  
 $p_Y(j) = \sum_{i=1, \dots, n} p(i, j)$  seien die Grenzverteilungen
- $H(X) = -\sum_{i=1, \dots, n} p_X(i) \log p_X(i)$
- $H(XY) = -\sum_{i=1, \dots, n; j=1, \dots, m} p(i, j) \log p(i, j)$

# Bedingte Entropie

- Wieder betrachten wir zwei Zufallsvariablen. Was ist die Unsicherheit über  $X$ , wenn  $Y$  bekannt ist?
- Die bedingte Wahrscheinlichkeit  
$$p_X(i|j) = p(i,j) / p_Y(j)$$
- $$H(X|Y) = \sum_{j=1, \dots, m} p_Y(j) H(X|Y=j)$$
$$= \sum_{j=1, \dots, m} p_Y(j) \sum_{i=1, \dots, n} p_X(i|j) \log(1/p_X(i|j))$$
$$= -\sum_{i,j} p(i,j) \log(p_X(i|j))$$
- Klar:  $0 \leq H(X|Y) \leq \log n$

# Kettenregel

- **Theorem 2.1**

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

- **Beweis:**

$$\begin{aligned} H(XY) &= -\sum_{i=1, \dots, n; j=1, \dots, m} p(i, j) \log p(i, j) \\ &= -\sum_{i=1, \dots, n; j=1, \dots, m} p(i, j) \log (p_X(i) p_Y(j|i)) \\ &= -\sum_{i=1, \dots, n; j=1, \dots, m} p(i, j) \log p_X(i) \\ &\quad - \sum_{i=1, \dots, n; j=1, \dots, m} p(i, j) \log (p_Y(j|i)) \\ &= H(X) + H(Y|X) \end{aligned}$$

- Weiterhin gilt:  $H(XY) \leq H(X) + H(Y)$

# Beispiele

- Zufallsvariablen  $X, Y$  mit  $X=Y$  (d.h.  $p(i, j)=0$  für alle  $i \neq j$ ).
  - $H(XY)=H(X)+H(Y|X)=H(X)$
  - $H(XX)=H(X)$
- Unabhängige Zufallsvariablen
  - $H(XY)=H(X)+H(Y|X)=H(X)+H(Y)$
- Zufallsvariablen  $X_1, \dots, X_k$ :
  - $H(X_1, \dots, X_k)=\sum_{i=1 \dots k} H(X_i | X_1, \dots, X_{i-1})$

# Eigenschaften

- $H(X|Y) \geq 0$
- $H(X|Y) = H(XY) - H(Y)$ 
  - leicht nachzurechnen
- $H(X|Y) = H(XY) - H(Y) \leq H(X) + H(Y) - H(Y)$   
 $= H(X)$
- $H(X|Y) \neq H(Y|X)$  im allgemeinen

# Information

- Information ist ein Maß der Korrelation zwischen Zufallsvariablen
- „System 1 hat Information über System 2“
- Seien wieder  $X, Y$  Zufallsvariablen
- $I(X:Y) = H(X) - H(X|Y)$ 
  - „Reduktion in der Unsicherheit über  $X$ , wenn man  $Y$  erfährt.“

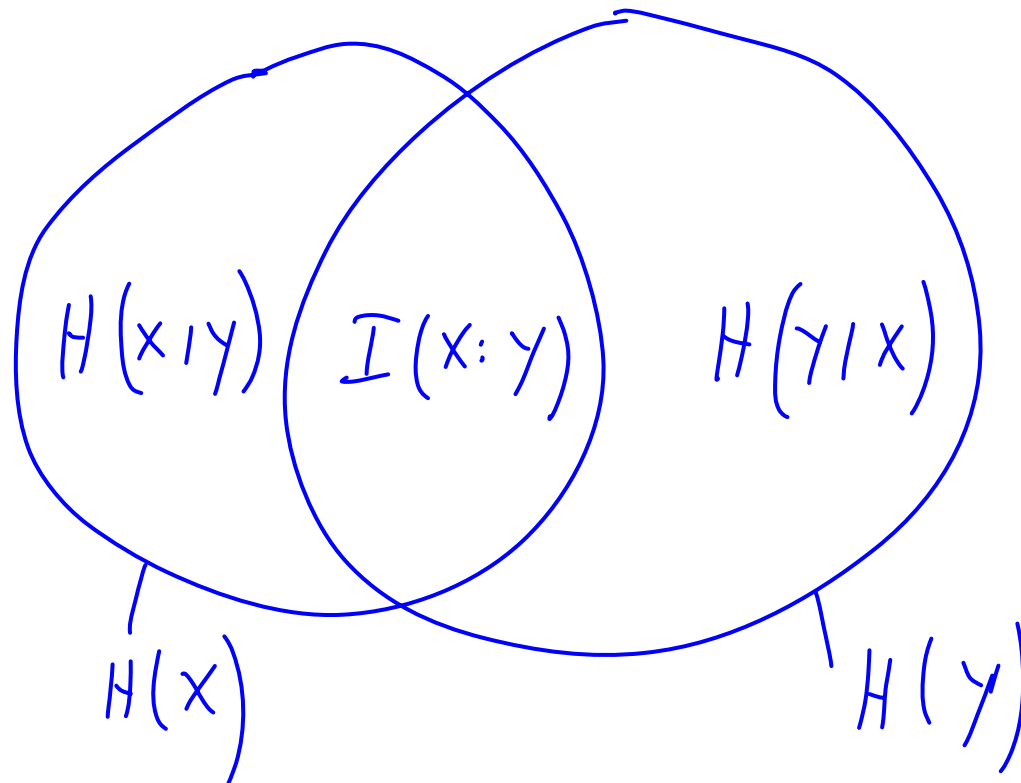
# Information

- **Theorem 2.2**

- $I(X:Y) = H(X) + H(Y) - H(XY)$ ,  
denn  $H(XY) = H(Y) + H(X|Y)$
- $I(X:Y) \leq H(X)$
- $I(X:Y) \leq H(Y)$
- $I(X:Y) = I(Y:X)$
- $I(X:X) = H(X)$

# Information

- Venn Diagramm



# Bedingte Information

- $I(X:Y|Z)$ : Information zwischen  $X, Y$ , wenn  $Z$  gegeben ist.
- Verteilung  $p(i, j, k)$  auf  $\{1, \dots, n\} \times \{1, \dots, m\} \times \{1, \dots, l\}$
- $I(X:Y|Z) = \sum_{k \in \{1, \dots, l\}} p_Z(k) I(X:Y|Z=k)$
- Dabei sei  $I(X:Y|Z=k)$  die Information zwischen  $X$  und  $Y$ , unter der Verteilung  $p_{X,Y|Z=k}$  auf  $\{1, \dots, n\} \times \{1, \dots, m\}$ , die sich durch Konditionierung von  $p$  auf  $Z=k$  ergibt
- Man kann leicht sehen, dass  
$$I(X:Y|Z) = H(XZ) + H(YZ) - H(Z) - H(XYZ)$$
- Weitere Eigenschaften?
  - Kann Information negativ sein?
  - Kettenregel?

# Kettenregel

- **Theorem 2.3**

$$I(X, Y: Z) = I(X: Z) + I(Y: Z | X)$$

- **Induktiv:**

$$I(X_1, \dots, X_k: Z) = \sum_{i=1, \dots, k} I(X_i: Z | X_1, \dots, X_{i-1})$$

- **Beweis:**

$$\begin{aligned} I(X, Y: Z) &= H(XY) - H(XY | Z) \\ &= H(X) + H(Y | X) - H(X | Z) - H(Y | XZ) \\ &= H(X) - H(X | Z) + H(Y | X) - H(Y | XZ) \\ &= I(X: Z) + I(Y: Z | X) \end{aligned}$$

# Relative Entropie

- Gegeben seien zwei Zufallsvariablen  $X$ ,  $Y$  mit Verteilungen  $p, q$  auf  $\{1, \dots, n\}$
- Die relative Entropie ist
  - $D(X||Y) = D(p||q) = \sum_{i=1, \dots, n} p(i) \log(p(i)/q(i))$
- Auch Kullback Leibler Distanz genannt
  - Die relative Entropie misst, wie weit Verteilungen voneinander entfernt sind
  - ist keine Metrik (insbesondere nicht symmetrisch)

# Eigenschaften

- $D(p||q)$  kann unendlich sein: wenn  $q(i)=0$ , aber  $p(i)>0$  (unter den Konventionen  $p \log(p/0)=\infty$ ,  $0 \log 0/q=0$ )
  - Wir verwenden daher  $D(p||q)$  gewöhnlich, wenn  $p(i)>0 \Rightarrow q(i)>0$  für alle  $i$
- $D(p||p)=0$
- $D(p||q)=0 \Rightarrow p=q$
- Im allgemeinen gilt  $D(p||q) \neq D(q||p)$