

1.2. Informationsmaße

Shannon - Information

$$\boxed{I(P) = \sum_{i=1}^n P_i \ln P_i}$$

Es gilt stets $I(P) \leq 0$

Maximum: $I(P) = 0$ für $P_i = \delta_{ij}$
(scharfe Verteilung mit sicheren Ereignis A_j)

Minimum Variation der P_i um δP_i unter der Nebenbedingung $\sum_i \delta P_i = 0$
(wegen Normierung $\sum_{i=1}^n P_i = 1$)

$I(P)$ ist minimal

$$\begin{aligned} \delta I &= \sum_i \left(\ln P_i + P_i \cdot \frac{1}{P_i} \right) \delta P_i = 0 \\ &= \sum_i (\ln P_i + 1) \delta P_i = 0 \end{aligned}$$

Addition der Nebenbedingung $\sum \delta P_i = 0$ mit Lagrange-Multiplikator λ :

$$\sum_i (\ln P_i + 1 + \lambda) \delta P_i = 0$$

unabhängige Variation δP_i

$$\ln P_i + 1 + \lambda = 0 \quad \forall_i$$

$$\ln P_i = -(1 + \lambda) = \text{const}$$

Normierung $\sum_{i=1}^N P_i = N P_i = 1$

$$\Rightarrow \boxed{P_i = \frac{1}{N}} \quad \text{Gleichverteilung}$$

(iii) Verallgemeinerte Informationsmaße (Rényi)

$$I_q = - \frac{1}{1-q} \ln \left(\sum_i P_i^q \right), \quad q = 1, 2, \dots$$

($q \rightarrow 1$ Shannon-Information)

Informationsgewinn:

Maß für die Zusatzinformation einer Wahrscheinlichkeitsverteilung $\{P_i\}$ im Vergleich zu einer Referenzverteilung $\{P_i'\}$ über derselben Ereignismenge

Bitzahl $\boxed{b(P_i) = -\ln P_i}$

$$b(P_i') - b(P_i) = \ln \frac{P_i}{P_i'}$$

erforderliche Bitzahl um P_i' in P_i zu wandeln durch eine Nachricht

mittlere Bitzahl (mit der Zufallszahlen w.-Verteilung P_i gerichtet)

$$\boxed{K(P, P') = \sum_{i=1}^n P_i \ln \frac{P_i}{P_i'}}$$

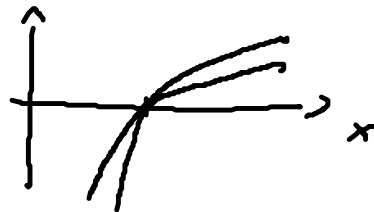
Informationsgewinn
(Kullback - Information)

Bem: (i) Asymmetrisch bzgl. $P \leftrightarrow P'$

(ii) Es gilt $K(P, P') \geq 0$, da

$$\sum_i P_i \ln \frac{P_i}{P_i'} \geq \sum_i P_i \left(1 - \frac{P_i'}{P_i}\right) = \sum_i P_i - \sum_i P_i' = 1 - 1 = 0$$

$$\left(\ln x \geq 1 - \frac{1}{x} \quad \text{für } x > 0\right)$$



(iii) $P_i' \neq 0$, damit $K < \infty$

(iv) Für $P_i' = \frac{1}{n}$ (Gleichverteilung)

$$K(P, P') = \underbrace{\sum_{i=1}^N P_i \ln P_i}_{I(P)} + \underbrace{\sum_{i=1}^N P_i \ln N}_1 = I(P) - I\left(\frac{1}{N}\right)$$

(v) Minimum von K : Variation der P_i um δP_i unter Nebenbedingung $\sum_i \delta P_i = 0$

$$\delta K = \sum_i \left(\ln \frac{P_i}{P_i} + 1 \right) \delta P_i$$

$$\sum_i \left(\ln \frac{P_i}{P_i} + 1 + \lambda \right) \delta P_i = 0$$

$$\Rightarrow \ln \frac{P_i}{P_i} = -(1 + \lambda) = \text{const}$$

$$\Rightarrow P_i \sim P_i'$$

$$\text{Normalisierung} \Rightarrow P_i = P_i' \Rightarrow K = 0$$

(vi) $K(P, P')$ ist konvexe Funktion von P da

$$\frac{\partial^2 K}{\partial P_i \partial P_j} = \frac{\partial}{\partial P_j} \left(\ln \frac{P_i}{P_i} + 1 \right) = \frac{1}{P_i} \delta_{ij} \geq 0$$

somit ist auch $I(P) = K(P, \frac{1}{N}) - \ln N$ konvex.

Kontinuierliche Ereignismenge ($x \in \mathbb{R}^d, \rho(x)$)

$$P_i = \rho(x_i) \Delta^d x$$

$$K(P, P') = \sum_i \Delta^d x \rho(x_i) \ln \frac{\rho(x_i)}{\rho'(x_i)}$$

$$\Delta^d x \rightarrow 0 : \boxed{K(\rho, \rho') = S d^d x \rho \ln \frac{\rho}{\rho'}}$$

Bem Interpretation von $-2\dot{K}(\rho, \rho')$ in der Thermodynamik als Entropieproduktion, von $2T \cdot K(\rho, \rho')$ als Exergie (availability)

1.3. Verallgemeinerte kanonische Verteilung

Motivation: Makroskop. thermodyn. Zustand ist gegeben durch Mittelwerte $\langle M(x) \rangle$ von Mikroobservablen $m(x)$, interpretiert als z -falls variable.
Rückschlüsse von $M(x)$ auf die Wahrsch.-verteilung $\rho(x)$?

Methode: vorurteilsfreie Schätzung (Jaynes, 1957)
(unbiased guess, Prinzip des maximalen Nichtwissens)

→ Verallgemeinerung des Laplace'schen Prinzips vom unzureichenden Grund

(Maximum der Shannon-Info $I(\rho)$
= Maximum des Nichtwissens $S(\rho)$
liefert Gleichverteilung)

Fakt Zusätzlich zur Normierung von P_i sind Mittelwerte von n Zufallsvariablen M_i^0 ($0=1, \dots, n$) gegeben: $\langle M^0 \rangle = \sum_{i=1}^n P_i M_i^0$

Annahme: Jedes Elementarergebnis A_i hat gleiche a-priori-Wahrscheinlichkeit, d. h. dem zusätzl. Kenntnisse $\langle M^0 \rangle$ gibt Gleichverteilung über A_i .

Informationstheoret. Prinzip (Jaynes)

Suche die Wahrscheinlichkeitsverteilung, die unter Festlegung aller bekannten Angaben als Nebenbedingung die minimale Information enthält.

$$\text{Also } I(P) = \sum_{i=1}^n P_i \ln P_i \stackrel{!}{=} \text{Minimum}$$

Nebenbed. $\sum_{i=1}^n P_i = 1$ Normierung
($n+1$)

$$\langle M^0 \rangle = \sum_{i=1}^n P_i M_i^0 \quad (0=1, \dots, n)$$

Variation: $\delta I = \sum_{i=1}^n (\ln P_i + 1) \delta P_i$

$$\sum_i \delta P_i = 0$$

Lagrange-Multiplikator $\lambda = -(\psi + 1)$

$$\sum_i M_i^0 \delta P_i = 0$$

Lagrange-Mult. λ_0

Wähle η, λ_0 so, dass die Koeff. von $(m+1) \delta P_i$'s verschwinden,
die übrigen $N - (m+1)$ sind dann frei variierbar.

$$\Rightarrow \sum_i (\ln P_i - \eta + \lambda_0 M_i^0) \delta P_i \stackrel{!}{=} 0$$

Summationskonvention über i

$$\Rightarrow \boxed{P_i = \exp(\eta - \lambda_0 M_i^0)}$$

verallgemeinerte kanonische Verteilung

Die Lagrange-Multiplikatoren η, λ_0 sind durch die
 $m+1$ Nebenbedingungen eindeutig bestimmt.