1.2. Informationsmaße

Shannon-Information

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

Es gilt stets $\underline{I}(P) \leq 0$

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

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

$$\underline{I}(P) \overset{!}{=} \text{minimal}$$

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

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

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

$$\ln p_i + 1 + \lambda = 0 \qquad \forall i$$

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

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

$$\Rightarrow \boxed{p_i = \frac{1}{N}} \qquad \text{Gleichverteilung}$$

(iii) Verallgemeinerte Informationsmaße (Renyi)

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

($q \to 1$ Shannon-Information)

Informationsgewinn:

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

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

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

notwendige Bitzahl um p_i' in p_i zu verwandeln durch eine Nachricht

mittlere Bitzahl (mit der korrigierten W.-Verteilung p_i gewichtet)

$$\boxed{K(p, p') = \sum_{i=1}^{N} p_i \ln \frac{p_i}{p_i'}}$$

Informationsgewinn

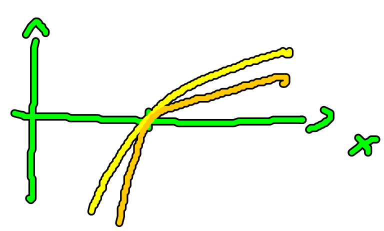
(Kullback-Information)

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

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

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

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



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

(iv) Für $p_i' = \frac{1}{N}$ (Gleichverteilung)

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

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

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

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

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

$$\Rightarrow \quad P_i \sim P_i'$$

Normierung $\Rightarrow \quad P_i = P_i' \quad \Rightarrow \quad K = 0$

(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 $I(P) = K(P, \tfrac{1}{N}) - \ln N$ konvex.

__Kontinuierliche Ereignismenge__ $(x \in \mathbb{R}^d, \rho(x))$

$$P_i = \rho(x^i)\, \Delta^d x$$

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

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

Bem. Interpretation von $-k\dot{K}(\rho,\rho')$ in der Thermodynamik als Entropieproduktion, von $kT\cdot K(\rho,\rho')$ 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 Zufallsvariable. Rückschluss von $M(x)$ auf die Wahrsch.-verteilung $\rho(x)$?

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

$\longrightarrow$ Verallgemeinerung des Laplace'schen Prinzips vom unzureichenden Grund

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

Jetzt Zusätzlich zur Normierung von P_i sind Mittelwerte von m Zufallsvariablen M_i^ν $(\nu = 1, \dots, m)$ gegeben: $\langle M^\nu \rangle = \sum_{i=1}^{N} P_i M_i^\nu$

Annahme: Jedes Elementarereignis A_i hat gleiche a-priori-Wahrscheinlichkeit, d. h. ohne zusätzl. Kenntnisse $\langle M^\nu \rangle$ gilt Gleichverteilung über den A_i.

Informationstheoret. Prinzip (Jaynes)

Suche die Wahrscheinlichkeitsverteilung, die unter Einhaltung aller bekannten Angaben als Nebenbedingungen die minimale Information enthält.

Also $I(P) = \sum_{i=1}^{N} P_i \ln P_i \overset{!}{=}$ Minimum

Nebenbed. $(m+1)$

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

$$\langle M^\nu \rangle = \sum_{i=1}^{N} P_i M_i^\nu \qquad (\nu = 1, \dots, m)$$

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

$$\sum_i \delta P_i = 0 \qquad \text{Lagrange-Multiplikator } \lambda = -(\psi + 1)$$

$$\sum_i M_i^\nu \delta P_i = 0 \qquad \text{Lagrange-Mult. } \lambda_\nu$$

Wähle ψ, λ_ν so, dass die Koeff von $(m+1)$ δP_i's verschwinden, die übrigen $N-(m+1)$ sind dann frei variierbar.

$$\Rightarrow \quad \sum_i \Big(\ln P_i - \psi + \underbrace{\lambda_\nu M_i^\nu}_{\text{Summationskonvention über } \nu} \Big) \delta P_i \overset{!}{=} 0$$

$$\Rightarrow \quad \boxed{P_i = \exp\left(\psi - \lambda_\nu M_i^\nu \right)}$$

<u>verallgemeinerte kanonische Verteilung</u>

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