Jaynes (1957): stat. phys. as an inference problem using limited data (such as $\bar{E}$)

Maximize $S = -k_B \sum_i p_i \log p_i$

(p_i – prob. to be in cell i of phase space)

with constraints:

(1) $\sum_i p_i = 1$ normal'n

(2) $\langle E \rangle \equiv \sum_i p_i E_i = \bar{E}$

E_i – energy of a system of particles

$$\delta \left[\frac{S}{k_B} - \beta \left(\sum_i p_i E_i - \bar{E} \right) - \alpha \left(\sum_i p_i - 1 \right) \right] = 0$$

δ wrt $\{p_i\}$ & α, β

$$p_i^* = \frac{e^{-\beta E_i}}{\underbrace{\sum_j e^{-\beta E_j}}_{\equiv Q}}$$ $\equiv Q$, part'n f'n

$$S_{max} = -k_B \sum_i p_i^* \log p_i^* = -k_B \Big[\sum_i p_i^* (-\beta E_i) -$$

$$- \sum_i p_i^* \underbrace{\log (\sum_j e^{-\beta E_j})}_{Q} \Big] =$$

$$= + k_B \log Q + \underbrace{k_B \beta}_{1/T} \langle E \rangle = \frac{\langle E \rangle - F}{T},$$

where

$F = -k_B T \log Q$ is the free en.

Shore & Johnson:

(1980, 1981)

Max of entropy $\Rightarrow$ self-consistent way to draw inferences from prob. distrib's.

Max of H with constraints should be:

1) unique
2) inv wrt coord. transform'n
3) relative prob. of two subsets of outcomes should not depend on other subsets if data are provided indep. ~~...~~ for each subset
4) Prob. for 2 indep. systems should be a product of marginal probs.

Consider

$$H(\{p_i, q_i\}) - \lambda(\sum_i a_i p_i - \bar{a})$$

probs. prior probs.

a - some property of interest

$p_i \to q_i$ when there're no constraints (i.e., no data)

q_i " priors

Constraints/data lead to $p_i \neq q_i$

discard $\lambda\bar{a}$

Start with $(\partial_{p_i} - \partial_{p_k})[H - \lambda \sum_i a_i p_i]$

E.g. p_j increased, p_k decreased $\Rightarrow$

$\Rightarrow$ other cells $l \neq j, k$ unaffected

(axiom 3):

$$p_\ell (\partial p_j - \partial p_k)[H - \lambda \sum_i a_i p_i] = 0$$

$\ell \neq k, j$

$\Downarrow$

$(\partial p_j - \partial p_k) H$ must depend only on j, k

$\Downarrow$

$\partial p_m H$ must depend only on m

$\Downarrow$

$$H = \sum_i \underbrace{f(p_i, q_i)}_{\text{"cell-level" quantities}}$$

Continuous repres'n:

$$H = \int \mathcal{D}x \, f(p(x), q(x))$$

Coord. inv. of the max (axiom 2):

$x \to y \Rightarrow \mathcal{D}y = \mathcal{D}x \cdot J$ ← Jacobian

$$\int \mathcal{D}x \, p(x) = 1 \Rightarrow \int \underbrace{\mathcal{D}y}_{\mathcal{D}x \cdot J} \underbrace{p(y)}_{\frac{p(x)}{J}} = 1$$

$$\int \mathcal{D}x \, p(x) a(x) = \int \underbrace{\mathcal{D}y}_{\mathcal{D}x \cdot J} \underbrace{p(y)}_{\frac{p(x)}{J}} a(y) = \bar{a}$$

$\Downarrow$

$$a(x) = a(y)$$

So, $H' = \int \mathcal{D}x \, J f\left(\frac{p(x)}{J}, \frac{q(x)}{J}\right)$

$$\frac{\delta}{\delta p(x)}\left[H - \lambda \int \mathcal{D}x\, p(x)\, d(x)\right] = 0 \quad \text{gives}$$

$$\underbrace{\frac{\partial f[p(x), q(x)]}{\partial p(x)}}_{\equiv\, g[p(x),\, q(x)]} - \lambda d(x) = 0 .$$

$$\frac{\delta}{\delta p(y)}\left[H' - \lambda' \int \mathcal{D}y\, p(y)\, a(y)\right] =$$

$$= g\left[\frac{p(x)}{J}, \frac{q(x)}{J}\right] - \lambda' \underbrace{a(y)}_{=\, a(x)} = 0$$

So, $$\underbrace{g\left[\frac{p(x)}{J}, \frac{q(x)}{J}\right]}_{\substack{\text{since } J \text{ is} \\ \text{arbitrary, it has to} \\ \text{vanish from LHS:}}} = \underbrace{(\lambda' - \lambda)}_{\text{const}} a(x) + g[p(x), q(x)]$$

$$g\left[\frac{p(x)}{J}, \frac{q(x)}{J}\right] = g\left[\frac{p(x)}{q(x)}\right]$$

$$\swarrow$$

$$f[p(x), q(x)] = p(x)\, h\left[\frac{p(x)}{q(x)}\right] + \upsilon[q(x)]$$

Indeed, $$\frac{\partial f[p(x), q(x)]}{\partial p(x)} = \underbrace{h\left[\frac{p(x)}{q(x)}\right] + h'\left[\frac{p(x)}{q(x)}\right]\frac{p(x)}{q(x)}}_{g\left[\frac{p(x)}{q(x)}\right]}$$

$h\left[\frac{p(x)}{q(x)}\right]$ is arbitrary thus far

Finally, axiom 4 (system indep.):

Consider two indep. systems described by x_1 & x_2: (2 constraints)

$$\int \mathcal{D}[x]\, a_k(x_k)\, p(x_1, x_2) = A_k \qquad k = 1, 2$$

Define $r(x) = \frac{p(x)}{q(x)}$, $x \equiv \{x_1, x_2\}$

Now, $H = \int \mathcal{D}[x]\, p(x)\, h(r(x))$ gives

$$\frac{\delta}{\delta p(x)}\Big[H - \lambda_1 \int \mathcal{D}[x]\, p(x_1, x_2)\, a_1(x_1) - \lambda_2 \int \mathcal{D}[x]\, p(x_1, x_2)\, a_2(x_2) \Big] =$$

$$= h(r(x)) + r(x)\, h'(r(x)) - \lambda_1 a_1(x_1) - \lambda_2 a_2(x_2) = h[r_1 r_2] + r_1 r_2 h'[r_1 r_2] - \lambda_1 a_1 - \lambda_2 a_2 = 0.$$

system indep.

$\frac{\partial^2}{\partial x_1 \partial x_2}$: $\frac{\partial}{\partial x_1}\Big\{ h'[r_1 r_2](r_2' r_1) + r_1 r_2 h''[r_1 r_2](r_1 r_2') + r_1 r_2' h'[r_1 r_2] - \lambda_2 a_2' \Big\} =$

$$= h''[r_1 r_2](r_1' r_2)(r_2' r_1) + h'[r_1 r_2](r_2' r_1') + h'''[r_1 r_2](r_1' r_2)(r_1^2 r_2 r_2') + h''[r_1 r_2](2 r_1 r_1' r_2 r_2') + r_1' r_2' h'[r_1 r_2] + r_1 r_2' h''[r_1 r_2](r_1' r_2) =$$

$$= r_1' r_2' \{ 4 r_1 r_2 h'' + 2h' + r_1^2 r_2^2 h''' \} = 0,$$ or

$$4\underbrace{r}_{{}^{\shortparallel}r_1 r_2}h'' + 2h' + r^2 h''' = 0$$

(system indep.)

$$h(r) = -\underset{\uparrow}{K}\log(r) + \underset{\uparrow}{B} + C/r$$

scale factor, >0

some const

(just a const in H)

So, $H = \int \mathcal{D}[x]\, p(x)\left[-K\log(r_{(x)}) + \cancel{\beta} + C\frac{q_0(x)}{p(x)}\right] =$

$$= -K\int \mathcal{D}[x]\, p(x)\log\left[\frac{p(x)}{q_0(x)}\right] +$$

$$+ C\int \mathcal{D}[x]\, q_0(x) = -K\int \mathcal{D}[x]\, p(x)\log\left[\frac{p(x)}{q_0(x)}\right]$$

also an irrelev. const

Note that

$$H = -K\int \mathcal{D}[x]\, p_1 p_2\left[\log\frac{p_1}{q_1} + \log\frac{p_2}{q_2}\right] =$$

$$= -K\left[\int \mathcal{D}x_1\, p_1\log\frac{p_1}{q_1} + \int \mathcal{D}x_2\, p_2\log\frac{p_2}{q_2}\right] =$$

$$= \underline{\underline{H_1 + H_2}}, \text{ as expected}$$

Note that H is a KL distance between $p(x)$ & $q_0(x)$

Maximize $H \to$ minimize distance between $p(x)$ & $q(x)$ given the constraints.

Types of constraints

higher moments, combinations of moments?

SJ use $\int \mathcal{D}x\, p(x)\, a(x)$

Nonlinear in $p(x)$ constraints?

Axiom 1: unique max of H under constraints → no constraints that would mess up the convexity of $\int \mathcal{D}x\, p(x) \log p(x)$

What should $a(x)$ be? Could be higher moments, but not always useful...

Thermodynamics: a system is in contact with a large bath

system + bath = closed

Then $F = -\sum_{i,a} p_{ia} \log p_{ia} + \nu\left(\sum_{i,a} p_{i,a} - 1\right) + \lambda \sum_{i,a} p_{ia}\left(1 - \delta_{E_i + E_a, \underbrace{E_{tot}}_{const}}\right)$

i ↑ system, a ↑ bath

↑ $p_{ia} = 0$ if $E_i + E_a \neq E_{tot}$

Note that $p_i = \sum_a p_{ia}$,

$$p(a|i) = \frac{p_{ia}}{p_i} \leftarrow p_{ia} = p(a|i)\, p_i$$

[indep. not assumed]

$$\sum_i p_i = 1, \quad \sum_a p(a|i) = 1$$

Then $-\sum_{i,a} p_{ia} \log p_{ia} = -\sum_{i,a} p(a|i)\, p_i \left[\log p(a|i) + \log p_i\right] = -\sum_i p_i \log p_i - \sum_{i,a} p(a|i)\, p_i \log p(a|i)$.

So, $F_{\text{new}} = -\sum_i p_i \log p_i - \sum_{i,a} p(a|i)\, p_i \log p(a|i) +$

$$+ \sum_i \nu_i \Big[\sum_a p(a|i) - 1 \Big] \delta_{E_i+E_a,\, E_{tot}} + \alpha\Big(\sum_i p_i - 1\Big) +$$

$$+ \lambda \sum_{i,a} p(a|i)\, p_i \,[1 - \delta_{E_i+E_a,\, E_{tot}}] .$$

δF_{new} wrt $p(a|i)$, ν_i, λ :

$\underbrace{(p_i \,\&\, \alpha}_{\text{open system vars}}$ given)

$$\begin{cases} \sum_{i,a} p(a|i)\, p_i \,[1 - \delta_{E_i+E_a,\, E_{tot}}] = 0, \\ \sum_a p(a|i) = 1, \quad \text{if } E_i + E_a = E_{tot} \\ -p_i \log p(a|i) - p_i + \nu_i + \lambda p_i \,[1 - \delta_{E_i+E_a,\, E_{tot}}] = 0 \end{cases}$$

$\downarrow$

$$\log p(a|i) = \begin{cases} \frac{\nu_i}{p_i} - 1, & E_i + E_a = E_{tot} \\ \frac{\nu_i}{p_i} + \lambda - 1, & E_i + E_a \neq E_{tot} \end{cases}$$

$\sum_{i,a} p(a|i) p_i = 0$, or $p(a|i) = 0$, $\forall a$

$E_i + E_a \neq E_{tot}$:

consistent with $\lambda = -\infty$

$\cancel{\#\text{ microstates at } E_a \neq E_{tot} - E_i = \emptyset}$

$$\sum_{i,a} e^{\frac{\nu_i}{p_i} + \lambda - 1} p_i = 0 \quad \cancel{\Rightarrow \Big(\sum_a 1\Big)\Big(\sum_i e^{\frac{\nu_i}{p_i}} p_i\Big) = 0}$$

$$\cancel{\sum_a e^{\frac{\nu_i}{p_i} + \lambda - 1} = 1 \Rightarrow \underbrace{\Big(\sum_a 1\Big)}_{=\emptyset} = e^{1-\lambda-\frac{\nu_i}{p_i}} \Rightarrow \lambda = +\infty}$$

\# bath microstates w/energy $E_a = E_{tot} - E_i$

$E_i + E_a = E_{tot}$:

$$\sum_a 1 = \Omega(E_{tot} - E_i)$$

$$\sum_a \underbrace{e^{\frac{\nu_i}{p_i} - 1}}_{\text{indep. of } a} = 1 \quad \Rightarrow \quad p(a|i) = \frac{1}{\Omega(E_{tot} - E_i)}$$

So, $$p(a|i) = \frac{\delta_{E_a,\, E_{tot} - E_i}}{\Omega(E_{tot} - E_i)}$$

Now,

$$F_{new} = -\sum_i p_i \log p_i + \alpha\left(\sum_i p_i - 1\right) -$$

$$-\sum_{i,a} p_i \underbrace{\frac{\delta_{E_a, E_{tot}-E_i}}{\Omega(E_{tot}-E_i)} \log \frac{\delta_{E_a, E_{tot}-E_i}}{\Omega(E_{tot}-E_i)}}_{\text{"0 unless } E_a + E_i = E_{tot}} =$$

$$\sum_a \frac{1}{\Omega(E_{tot}-E_i)} = 1$$

$$= -\sum_i p_i \log p_i + \alpha\left(\sum_i p_i - 1\right) + \sum_i p_i \log \Omega(E_{tot}-E_i)$$

Large bath: $E_{tot} \gg E_i$,

$$\log \Omega(E_{tot} - E_i) \simeq \log \Omega(E_{tot}) - \beta E_i$$

So the constraint becomes

$-\beta \sum_i p_i E_i$, just as before

Higher-order moments drop out when the bath is large → 1st order moments arise naturally.

Indeed, as $N \to \infty$ in any system,

$$Q = \sum_E g(E)\, e^{-\beta E}$$

energy level degeneracy

fluct's $\langle E^2 \rangle - \langle E \rangle^2$ become small as $N \uparrow$, (as $\frac{1}{N}$)

only $\langle E \rangle$ remains finite

Higher-order moments also vanish...

↳ In "nanothermodynamics", higher-order moments may play a role, but the MAXENT framework still holds.