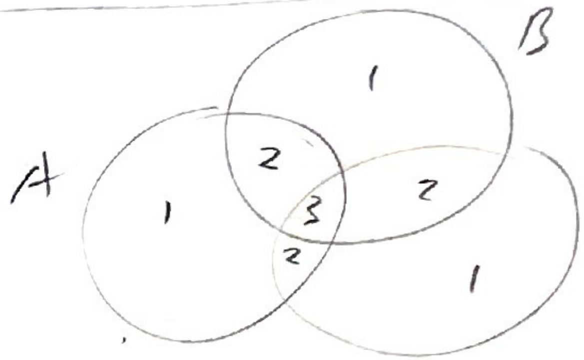
Lecture 3a Information Theory : Relative Entropy.

Review threeplusone.com/info Homework.

Entropy $S(A) = -\sum_a P_A(a) \log P_A(a)$ discrete $0 \geq S \geq \ln N$

(Alan Turing)

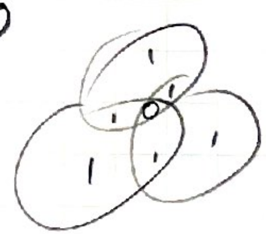
Base 2 bit "binary digit"
e nat "natural digit"
10 ban/digits.

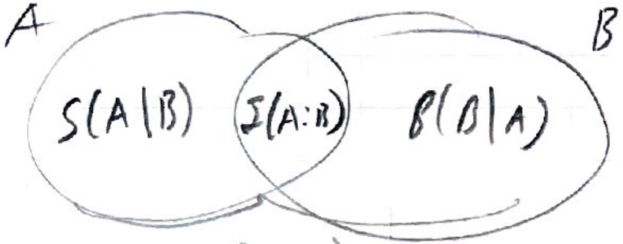


$S(A|B) = -\sum P(A,B) \log P(A|B)$

$I(A:B) = \sum P(A,B) \log \dfrac{P(A,B)}{P(A)P(B)} \geq 0$

$S(A,B,C) = S(A) + S(B) + S(C)$
$- I(A:B) - I(A:C) - I(B:C)$
$+ I(A:B:C)$
positive or negative.

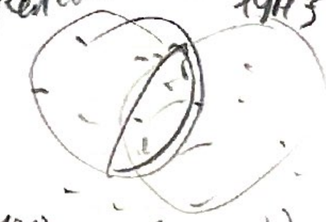




$S(A,B)$

, $A \cup B$
: $A \cap B$
| $A \setminus B$

Color = Red lines type 3 UVs



$\| (\text{Sorgeln}) \leftarrow$

No Red Means
Not Venn Diagram

P(a and b)

Relative Entropy

(entropy narrow sense Shannon Entropy → wide sense, can we measure it in bits or some generalization)

Relative Entropy

Kullback-Leibler divergence (1951) D_{KL} "Dee-Kay-ell"

$$D(A \| B) = \sum_x P_A(x) \ln \frac{P_A(x)}{P_B(x)}$$

$$D(p \| q) = \sum p(x) \ln \frac{p(x)}{q(x)}$$

(Cover & Thomas)

≥ 0

0 only if $P_A(x) = P_B(x)$

$$= S(A;B) - S(A)$$

$$= -\sum P_A(x) \log P_B(x) - S(A)$$

Not symmetric $D(A \| B) \neq D(B \| A)$

"How different A is from B"

Expectation

Jensen

3

$$\langle f(a) \rangle_A = \sum_{x} P_A(a)\, f(x)$$

Expectation

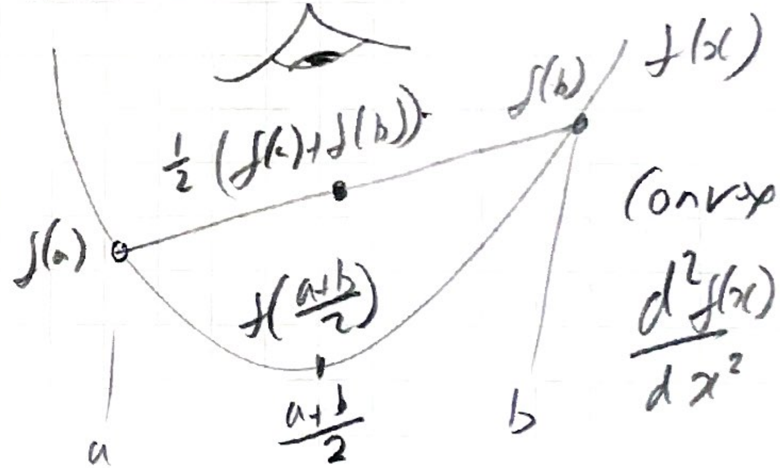
$$\mathbb{E}_A\big(f(a)\big) \equiv \langle f(x) \rangle_A$$

Not expected.

$$S(A) = \langle -\ln P_A(a) \rangle_A$$

Jensen's inequality (1906)

x^2 2



Convex function

$$\frac{d^2 f(x)}{dx^2} \geq 0$$

$$\langle f(x) \rangle \geq f(\langle x \rangle)$$

convex

$$\langle e^{x} \rangle \geq e^{\langle x \rangle}$$

Gibbs-Bogoliubov-Feynman bound

Gibbs Inequality & others

Gibbs inequality (Information inequality)

$$D(A\|B) \geq 0$$

$$= \left\langle -\log \frac{P_B(x)}{P_A(x)} \right\rangle$$

$$\geq -\log \left\langle \frac{P_B(x)}{P_A(x)} \right\rangle$$

$$-\log \underbrace{\sum \cancel{P_A(x)} \frac{P_B(x)}{\cancel{P_A(x)}}}_{=1}$$

$$\geq 0$$

$$D(A\|\underset{\text{uniform}}{U}) = \sum_a P_A(x) \log \frac{P_A(x)}{1/N}$$

$$-S(A) + \underbrace{\sum_a P_A(x) \log N}_{\log N} \geq 0$$

$$S(A) \leq \log N$$

Relative entropy

Encoding inefficiency

331323 $\frac{15}{6} = 2\frac{1}{2}$

	~~expected~~	~~alt tea~~	al tea / actual
␣	$\frac{1}{2}$	0	$\frac{1}{6}$
e	$\frac{1}{4}$	10	$\frac{1}{6}$
t	$\frac{1}{8}$	110	$\frac{1}{3}$
a	$\frac{1}{8}$	111	$\frac{1}{3}$
bits	S $1\frac{1}{2}$		1.92

b

$-\frac{1}{2^n} \log 2^{-n} = -n$

$-\frac{n}{2^n}$

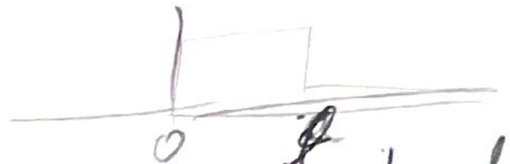
$\frac{1}{6} \log \frac{1/6}{1/2}$ $\quad$ $\frac{1}{6} \log \frac{1/6}{1/4}$

$\frac{1}{3} \log \frac{1/3}{1/8}$ $\quad$ $\frac{1}{3} \log \frac{1/3}{1/8}$

0.58 inefficiency

Differential Entropy

$$S = -\int P(x) \log P(x) \, dx$$



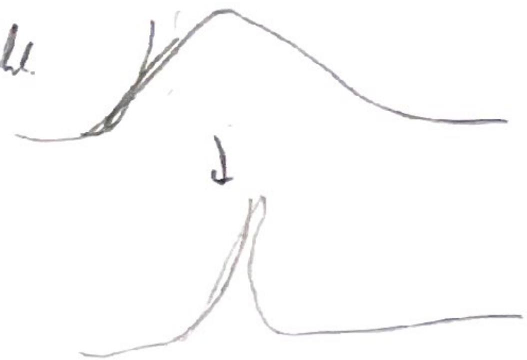
6

$S = \log l$

(Gaussian)

~~infinite~~ $-\infty \leq S \leq \infty$!

Not invariant to change of variable



$\Rightarrow$ use Relative entropy.

Classical Thermodynamics entropies are Quantum

~~$D(A|B \,\|\, A'|B')$~~ ~~Change of Variables~~

$$S(aX) = S(X) + \log|a|$$

$y = f(x)$

$$S(Y) \leq S(X) + \int p(x) \log \left|\frac{\partial f}{\partial x}\right| dx$$

Jacobian

Free energy

$$F =$$

"capacity to do work.

$$S = \left\langle \ln \frac{e^{-\beta E}}{Z} \right\rangle$$

$$S = \ln Z + \beta \langle E \rangle$$

$$-\ln Z = \beta \langle E \rangle - S$$

$$F = \boxed{}$$

$$p^{0}(x) = \frac{e^{-\beta E(x)}}{Z} \qquad Z = \sum e^{-\beta E(x)}$$

Z ← partition function

"Boltzmann distribution"

(Heart of statistical mechanics)

$$\beta F \overset{eq}{=} -\ln Z, \quad p(x) = e^{\beta F - \beta E(x)}$$

<u>Kullback-Leibler</u> $D\left(p \,\|\, p^{eq}\right) \geq 0$

75b

$$D(p\|p^{eq}) = -S(A) - \sum p_a e^{\beta F - \beta E}$$

$$- \beta F^{a} + \langle E \rangle_A$$

$$= \beta F^{neq} - \beta F^{can}$$

(will defer discussion of why this is a good definition of nonequilibrium free energy
→ (equate to do work)