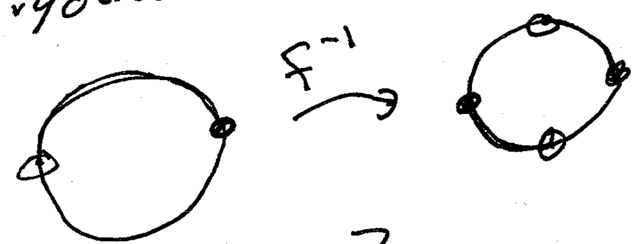
Entropy, Cont.

Example: Define $f: S^1 \to S^1$ via $\theta \mapsto 2\theta \bmod 1$ or treating $S^1 \subseteq \mathbb{C}$ as $\{z : |z| = 1\}$, $z \mapsto z^2$ with Lebesgue measure μ and $\mathcal{B}$ = the Borels. We showed previously that f is ergodic



Let $\rho = \{[0, 1/2), [1/2, 1)\}$

$f^{-1}(\rho) = \{[0, 1/4), [1/4, 1/2), [1/2, 3/4), [3/4, 1)\}$

$f^{-n}(\rho) = \{[0, \frac{1}{2^{n+1}}), \ldots [1 - \frac{1}{2^{n+1}}, 1)\}$

So every interval in S^1 is the countable union of sets from $\bigvee_{i=0}^{\infty} f^{-i}(\rho)$ and so

$$\bigvee_{i=0}^{\infty} f^{-i}(A(\rho)) \doteq \mathcal{B}$$

Thus using Kolmogorov-Sinai for non-invertible transformations $h(f) = h(f, \rho)$. Notice that $\bigvee_{i=0}^{n-1} f^{-i}(\rho) = f^{-(n-1)}(\rho)$

$$H\left(\bigvee_{i=0}^{n-1} f^{-i}(\rho)\right) = H(f^{-(n-1)}\rho)$$

$$= -2^n \left[\frac{1}{2^n} \log\left(\frac{1}{2^n}\right)\right]$$

$$= n \log 2$$

So $h(f) = h(f, \rho) = \lim_{n \to \infty} \frac{1}{n} n \log 2 = \log 2.$

<u>Alternative method</u>: We showed previously that $(f, \mathcal{B}, \mu)$ is isomorphic to $(\Sigma_2, \mathcal{B}(\Sigma_2^+), \nu)$ where ν is the $(1/2, 1/2)$-Bernoulli measure on Σ_2^+.

It follows from a problem in HW that

$$h_\nu(\sigma) = -\left[\tfrac{1}{2}\log \tfrac{1}{2} + \tfrac{1}{2}\log \tfrac{1}{2}\right] = \log 2$$

MODELING and entropy

Bernoulli and Markov measures are common tools in modeling processes or sequences.

This is a vast subject which will just be touched on here.

We assume the data or output takes the form of a sequence of integers in $\{0, \ldots, k-1\}$. or an element of Σ_k

<u>Bernoulli measure</u>: If the next entry doesn't depend on the previous, perhaps just its probability of occurrence is a good model

The standard example is a coin toss with a crooked coin with 70% heads and 30% tails.

Then $\vec{p} = [.7, .3]$ yields a Bernoulli measure

$$\mu([s_0 \ldots s_n]) = p_{s_0} \cdots p_{s_n}$$

So $\mu([HHT]) = (.7)(.7)(.3)$

Another example is the base 10 digits of a number $x \in [0,1]$, we expect these to be distributed according to $\vec{p} = (1/10, \dots, 1/10)$ [more on this later]

<u>Markov measure</u>: If the next step depends on the previous we want to assign measures to $[s_0 s_1]$ which take this into account. So there is a matrix involved which computes $\mu[s_0 s_1] = p_{s_0} A_{s_0 s_1}$ where $\vec{p} A = \vec{p}$

Notice that $\vec{p} > 0$ is determined by A when A is primitive via Peron-Fröbenius.

For example, in a disease with states uninfected, infected, recovered, died from disease, the transition probabilities depend on the previous state

Another example is the sequence of A, C, G, T's in DNA.

What we have described is a one-step Markov process or measure. The next stage could depend on several previous so one wants to specify the measure or probability of longer cylinder sets $[s_0 \ldots s_n]$.

In both Bernoulli and Markov models the entropy indicates the complexity or information of the process or structure

NOTE: (a) See Probability and Statistics books for more information and other invariant measures.

(b) There a lots of invariant measures for (Σ_K, T) that are not Bernoulli or Markov. Recall Krieger's theorem implies that if $f : (X, \mathcal{B}, \mu)$ is ergodic with $h(f) < \infty$, it is isomorphic to some $(\Sigma_K, \mathcal{B}(\Sigma_K), \nu)$, so shifts contain all finite entropy ergodic invariant measures.

Normal numbers

Fix an integer $b>1$. A point $x \in [0,1]$ has base b expansion $.s_0 s_1 s_2 \ldots$ if

$$x = \sum_{n=0}^{\infty} \frac{s_n}{b^{n+1}}$$

Base b expansion is not unique so let

$$X_b = \{x \in [0,1] : x \text{ has unique base } b \text{ expansion}\}$$
$$= [0,1] - \left\{\frac{\ell}{b^j} : 0 \le \ell \le b^j,\ j = 1,2,\ldots\right\}$$

So X_b is $[0,1]$ minus a countable set, so if μ is Lebesque measure, $\mu(X_b) = 1$.

Since one expects every finite sequence or block $a_0 \ldots a_n$ to be equally probable, for each $k > n+1$ define for $B = a_0 \ldots a_{n-1}$

$$y(x, B, k) = \#\text{ of occurrences of } B \text{ in } s_0 s_1 \ldots s_k$$

and so x is normal in base b

$$\lim_{k \to \infty} \frac{y(x, B, k)}{k} = \frac{1}{b^n} \quad \text{for all blocks } B$$

So, on average, x in base b expansion has the correct number of occurrences of the block B.

DEF: $x \in [0,1]$ is normal if it is normal in all bases $b > 1$.

<u>Theorem</u>: The collection of normal numbers in $[0,1]$ has full Lebesgue measure.

<u>Proof</u>: Let Z_b be all $\underline{s} \in \Sigma_b^+$ that do <u>not</u> end in 0^∞ or $(b-1)^\infty$. Like in the case of $b=2$ we did

$$\varphi(\underline{s}) = \sum_{i=0}^{\infty} \frac{s_i}{b^{i+1}} \qquad \varphi : Z_b \to X_b$$

gives an isomorphism between the $(\frac{1}{b}, \ldots, \frac{1}{b})$ Bernoulli measure ν on Σ_b^+ and Lebesgue measure μ on $[0,1]$. By definition $\varphi^{-1}(x)$ is the base b expansion of x.

Fix a block $B = a_0 a_1 \ldots a_{n-1}$, using the isomorph.

$y(x, B, k) = $ # of occurrences of B in $s_0 s_1 \ldots s_k$

= # of occurrences of B is $\varphi^{-1}(x)$ up to k

= # of times $\sigma^i(\varphi^{-1}(x))$ is in the cylinder set ${}_0[a_0 a_1 \ldots a_{n-1}]$ for $0 \le i \le k-n$

$$= \sum_{i=0}^{k-n} \chi_{{}_0[a_0 \ldots a_{n-1}]}(\sigma^i(\varphi^{-1}(x)))$$

Now recall $\sigma : (\Sigma_b^+, \mathcal{B}(\Sigma_b), \nu)$?

is ergodic. Let $Z_b(B)$ be the full measure set for which

$$\lim_{n\to\infty} \frac{1}{\cancel{K}} \sum_{i=0}^{\cancel{K}} \chi_{{}_0[B]}\left(\sigma^i\left(\varphi^{-1}(x)\right)\right)$$

$$= \int \chi_{{}_0[B]}\, d\nu = \frac{1}{b^n} = \nu\left({}_0[B]\right)$$

note that $\dfrac{y(x,B,k)}{\cancel{K}} = \lim_{K\to\infty} \frac{1}{K} \sum_{i=0}^{K-n} \chi_{{}_0[B]}\, \sigma^i\left(\varphi^{-1}(x)\right)$

$$= \lim_{K\to\infty} \frac{1}{K} \sum_{i=0}^{k} \chi_{{}_0[B]}\, \sigma^i\left(\varphi^{-1}(x)\right) = 1/b^n$$

when $\varphi^{-1}(x) \in Z_b(B)$

let $Z_b = \bigcap_{\text{blocks } B} Z_b(B)$, then $\nu(Z_b) = 1$

and $\varphi^{-1}(x) \in Z_b$ implies x is normal for b

or $x \in \varphi(Z_b)$ implies x is normal for b

By the isomorphism $\mu(\varphi(Z_b)) = 1$

and $x \in \bigcap_{b=2}^{\infty} \varphi(Z_b)$ implies x is normal

and $\mu\left(\bigcap_{b=2}^{\infty} \varphi(Z_b)\right) = 1$. $\blacksquare$

<u>Remark</u>: This can also be done directly using $f:[0,1) \circlearrowleft$ $f(x) = bx \bmod 1$ is ergodic and for $x \in X_b$, its base b expansion is the itinerary of x under f with the partition $A_i = \left[\frac{i}{b}, \frac{i+1}{b}\right)$.

$(i(x))_j = i \Leftrightarrow f^j(x) \in A_i \Leftrightarrow$ the j^{th} entry of the base b expansion of x is i.