13.04.2016 Mathematical focus

continue proof:

Proof. We now show that axioms 1-4 imply (A),(B^*),(C) for U_N.

(A) follows from axiom 1

(C) follows from axiom 4

$$H\left(\underbrace{\frac{1}{NM},\dots,\frac{1}{NM}}_{NM}\right) = H\left(\frac{1}{N}U_M^{(1)},\dots,\frac{1}{N}U_M^{(N)}\right)$$
$$= H(U_N) + N\frac{1}{N}H(U_M)$$

Now show (B^*):

$$H(U_N) \overset{\text{Claim } \textbf{??}}{=} H\left(\frac{1}{N},\underbrace{\frac{1}{N},\dots,\frac{1}{N}}_{\frac{N-1}{N}}\right)$$
$$= \underbrace{H\left(\frac{1}{N},\frac{N-1}{N}\right)}_{\delta_N} + \underbrace{\frac{N-1}{N}}_{=1-\frac{1}{N}}H(U_{N-1})$$

$$d_N = H(U_N) - H(U_{N-1}) = \delta_N - \frac{1}{N}H(U_{N-1})$$

Axiom 2 implies that

$$\delta_N \to H(0,1) = 0$$

$$\delta_N = d_N + \frac{1}{N}\left(H(U_{N-1}) - H(U_{N-2}) + H(U_{N-2}) \pm \dots - H(U_2) + H(U_2) - \underbrace{H(U_1)}_{=0}\right)$$
$$= d_N + \frac{1}{N}(d_2 + d_3 + \dots + d_{N-1})$$

Then

$$\sum_{n=2}^{N} n\delta_n = \sum_{n=2}^{N} n\left(d_n + \underbrace{\frac{1}{n}\left(\sum_{k=2}^{n-1} d_k\right)}_{\frac{1}{N}\sum_{k=2}^{N} d_k - \frac{1}{N}d_n}\right) = \dots = N\sum_{k=2}^{N} d_k$$

$N \leftrightarrow N-1$

$$\frac{1}{N}(d_2+\cdots+d_{N-1}) = \frac{1}{N(N-1)}\sum_{n=2}^{N-1} n\delta_n \overset{\text{why?}}{\to} 0 \frac{1}{N}(d_2+\cdots+d_{N-1}) \quad = \frac{1}{N(N-1)}\sum_{n=1}^{N-1} n\delta_n$$

$$= \frac{1}{2}\left(\frac{2}{N(N-1)}\sum_{n=1}^{N-1} n\delta_n\right)$$

$$= \frac{1}{2}\sum_{n=1}^{N-1}\frac{2n}{N(N-1)}\delta_n$$

$$\underbrace{\sum_{n=1}^{N_\varepsilon}\frac{2n}{N(N-1)}\delta_n}_{\to 0} + \text{Rest}$$

where $|\text{Rest}| < \varepsilon$. For this convergence we use known arguments form analysis. Sow (B^*) holds, $H(U_N) = \log_2(N)$.

Conclude:

$$\log_2(N) = H\left(\underbrace{\frac{1}{N},\ldots}_{K}\underbrace{\ldots,\frac{1}{N}}_{N-K}\right)$$

$$= H\left(\frac{K}{N},\frac{N-K}{N}\right) + \frac{K}{N}\underbrace{H(U_K)}_{\log_2(K)} + \frac{N-K}{N}\underbrace{H(U_{N-k})}_{\log_2(N-K)}$$

$$H\left(\frac{K}{N},\frac{N-K}{N}\right) = -\frac{K}{N}\left(\log_2(N)-\log_2(K)\right) + \frac{N-K}{N}\left(\log_2(N)-\log_2(N-K)\right)$$

$$= -\frac{K}{N}\log_2\left(\frac{K}{N}\right) - \frac{N-K}{N}\log_2\left(\frac{N-K}{N}\right)$$

Axiom 2 implies

$$H(p_1, 1-p_1) = -p_1\log_2 p_1 - (1-p_1)\log_2(1-p_1)$$

for all p_1. Now use axiom 4 and induction. □

Proof. of Lemma (2016-03-15, 1.2) fix $q \in \mathbb{N}$, $q \geq 2$

$$f(N) = H(U_N)g(N) \qquad = f(N) - \frac{f(q)\log_2(N)}{\log_2(q)}$$

$$\varepsilon_N = g(N+1) - g(N) = \underbrace{f(n+1) - f(N)}_{\overset{(B^*)}{\to} 0} - \frac{f(q)}{\log_2(q)}\underbrace{\left(\log_2(N+1) - \log_2(N)\right)}_{\to 0} \to 0$$

$$g(q) = 0$$

$$g(q^k N) \overset{\text{(C)}}{=} g(q) + q(N) = g(N)$$

$$N' = \lfloor \frac{N}{q} \rfloor$$
$$N = N'q + r \ \ 0 \le r \le q-1$$

$$g(N) - g(N') = g(N) - g(qN') = \underbrace{\sum_{j=qN'}^{N-1} \varepsilon_j}_{\text{at most } q-1}$$

$$\begin{aligned} N^{(k+1)} &= (N^{(k)})' \\ N^{(0)} &= N \\ N^{(k)} &\le \frac{N}{q^k} \\ k_N = \lfloor \log_q(N) \rfloor : N^{(k_n+1)} &= 0 \\ N^{(k+1)} &= \lfloor \frac{N^{(k)}}{q} \rfloor \end{aligned}$$

$$\begin{aligned} g(N) &= g(N) - g(N^{(1)} + g(N^{(1)}) - g(N^{(2)} + \cdots + g(N^{(k_N)}) - 0 \\ &= \sum_{\text{some } j} \varepsilon_j \end{aligned}$$

How many: S_N at most $(q-1)\log_q(N)$.

$$\begin{aligned} \frac{1}{S_N} \sum_{\text{these } j} \varepsilon_j &\to 0 \\ \frac{1}{(q-1)\log_q(N)} \sum_{\text{these } j} \varepsilon_j &\to 0 \\ \implies \frac{1}{(q-1)(k_N+1)} \sum \varepsilon_j &\to 0 \end{aligned}$$

This tells us that

$$\frac{g(N)}{k_N} \to 0 \implies \frac{g(N)}{\log_2(N)} \to 0$$

Thus,

$$\frac{f(N)}{\log_2(N)} = \frac{g(N)}{\log_2(N)} + \frac{f(q)}{\log_2(q)} \overset{N\to\infty}{\to} \frac{f(q)}{\log_2(q)} = \text{ constant in } q = 1$$

$\square$

Mixed session

Recall:
$$H(Y|X) = \sum_{x\in\mathcal{X}} p_X(x) \underbrace{H(Y|X=x)}_{p_{(Y|X=0)}(y)=\mathbb{P}[Y=y|X=x]=\frac{p_{X,Y}(x,y)}{p_X(x)}}$$
$$H(Y|X) = H(X,Y) - H(X)$$

Remark (Exercise)**.** *Write down the meaning of $H(X,Y|Z)$ and $H(Y|X,Z)$ and show that*
$$H(X,Y|Z) = H(X|Z) + H(Y|X,Z).$$
Hint: Use $H(X,Y) = H(X) + H(Y|X)$.

Try to understand where this comes from.

0.1 Relative entropy and mutual information

Definition. *Let $p(\cdot)$ and $q(\cdot)$ be two prob distr on $\mathcal{X}$. The relative entropy of Kullback-Leibler distance/divergence of p with respect to q is*
$$D(p||q) = \sum_x p(x)\log_2 \frac{p(x)}{q(x)}$$
If X is a RV with $p_X = p$, then
$$\mathbb{E}\left(\log_2 \frac{p(X)}{q(X)}\right)$$

Convention:
$$0\log_2 \frac{0}{b} := 0 \ \ \forall b \geq 0$$
$$a\log_2 \frac{a}{0} := +\infty \ \ \forall a > 0$$

in general: $D(p||q) \neq D(q||p)$

If $p = q$, then $D(p||p) = 0$.

We will see that $D(p||q) \geq 0$. And we will see that $D(p||q) = 0 \implies p = q$.