1 Expanded Proof: $H(X_n|h_n) \leq H_F(\pi)$

This is an expaned proof to that presented as appendix A in *A Refined Limit on the Predictability of Human Movement* by Gavin Smith, Romain Wieser, James Goulding and Duncan Barrack. It makes explict the claim in the Supporting Online Material for *Limits of Predictability in Human Mobility* by Song, Qu, Blumm and Barabasi, that $H(X_n|h_n) \leq H_F(\pi)$ "represents an appropriate rewritting of Fano's inequality".

Let X_n be a discrete random variable within a stochastic process at the n^{th} time point for which the possible outcomes are a set of spatial areas corresponding to a finite quantisation of a spatial region. Let the set of possible spatial areas be denoted by the set Ω and h_{n-1} be the outcome of the preceding random variables within the random process.

Consider the distribution $P(X_n|h_{n-1})$. In context this distribution denotes the *next step* probabilities over all possible spatial regions in the spatial area under consideration. h_{n-1} is a specific history of points of length $n-1$.

Let the probability of the most probable location, $x_{ML} \in \Omega$, equal π, given the history h_{n-1}.

Let E be a binary random variable.

Let:

$$P(e|h_{n-1}) = \sum_{x \in \Omega, x \neq x_{ML}} P(x|h_{n-1}) \tag{1}$$

Then:

$$1 - P(e|h_{n-1}) = P(x_{ML}|h_{n-1}) \tag{2}$$

The corresponding entropy, H, of the binary variable E is then:

$$H(E|h_{n-1}) = -P(e|h_{n-1}) \log_2 P(e|h_{n-1}) - (1 - P(e|h_{n-1})) \log_2(1 - P(e|h_{n1-})) \tag{3}$$

Now define:

$$P_C(x|h_{n-1}) = \begin{cases} \frac{P(x|h_{n-1})}{P(e|h_{n-1})} & \text{if } x \neq x_{ML} \\ \text{undefined} & \text{otherwise} \end{cases} \tag{4}$$

Note that by definition:

$$\begin{aligned} \sum_{x \in \Omega, x \neq x_{ML}} P_C(x|h_{n-1}) &= \sum_{x \in \Omega, x \neq x_{ML}} \frac{P(x|h_{n-1})}{P(e|h_{n-1})} && (5) \\ &= \frac{1}{P(e|h_{n-1})} \sum_{x \in \Omega, x \neq x_{ML}} P(x|h_{n-1}) && (6) \\ &= \frac{\sum_{x \in \Omega, x \neq x_{ML}} P(x|h_{n-1})}{\sum_{x \in \Omega, x \neq x_{ML}} P(x|h_{n-1})} && \text{[via equation 1]} \quad (7) \\ &= 1 && (8) \end{aligned}$$

Consider the entropy $H(X_n|h_{n-1})$:

$$H(X_n|h_{n-1}) = -\sum_{x\in X_n} P(x|h_{n-1})\log_2(P(x|h_{n-1})) \quad \text{[definition of entropy]} \tag{9}$$

$$= -P(x_{ML}|h_{n-1})\log_2(P(x_{ML}|h_{n-1})) - \sum_{x\in\Omega, x\neq x_{ML}} P(x|h_{n-1})\log_2(P(x|h_{n-1})) \tag{10}$$

$$= -(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) - \sum_{x\in\Omega, x\neq x_{ML}} P(x|h_{n-1})\log_2(P(x|h_{n-1})) \quad \text{[via equation 2]} \tag{11}$$

$$= -P(e|h_{n-1})\log_2 P(e|h_{n-1}) + P(e|h_{n-1})\log_2 P(e|h_{n-1}) - (1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) - \sum_{x\in\Omega, x\neq x_{ML}} P(x|h_{n-1})\log_2(P(x|h_{n-1})) \quad \text{[zero change, +/- of identical terms]} \tag{12}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] + P(e|h_{n-1})\log_2 P(e|h_{n-1}) - \sum_{x\in\Omega, x\neq x_{ML}} P(x|h_{n-1})\log_2(P(x|h_{n-1})) \quad \text{[regrouping of terms]} \tag{13}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] + P(e|h_{n-1})\log_2 P(e|h_{n-1}) - P(e|h_{n-1})\sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})\log_2(P(x|h_{n-1})) \quad \text{[substitution via equation 4]} \tag{14}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] - P(e|h_{n-1})\left[-\log_2 P(e|h_{n-1}) + \sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})\log_2(P(x|h_{n-1}))\right] \quad \text{[regrouping of terms]} \tag{15}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] - P(e|h_{n-1})\left[-\log_2 P(e|h_{n-1})\left(\sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})\right) + \sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})\log_2(P(x|h_{n-1}))\right] \quad \text{[zero change via multiplication by 1, via equation 8]} \tag{16}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] - P(e|h_{n-1})\left[\sum_{x\in\Omega, x\neq x_{ML}} -P_C(x|h_{n-1})\log_2 P(e|h_{n-1}) + \sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})\log_2(P(x|h_{n-1}))\right] \tag{17}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] - P(e|h_{n-1})\left[\sum_{x\in\Omega, x\neq x_{ML}} -P_C(x|h_{n-1})\log_2 P(e|h_{n-1}) + P_C(x|h_{n-1})\log_2(P(x|h_{n-1}))\right] \tag{18}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] - P(e|h_{n-1})\left[\sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})[\log_2(P(x|h_{n-1})) - \log_2 P(e|h_{n-1})]\right] \tag{19}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] - P(e|h_{n-1})\left[\sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})\log_2\left(\frac{P(x|h_{n-1})}{P(e|h_{n-1})}\right)\right] \quad \text{[log laws]} \tag{20}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] - P(e|h_{n-1})\left[\sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})\log_2(P_C(x|h_{n-1}))\right] \quad \text{[definition from equation 4]} \tag{21}$$

$$= -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] + P(e|h_{n-1})H(X_{x\neq x_{ML}}|h_{n-1}) \tag{22}$$

where:

$$H(X_{x\neq x_{ML}}|h_{n-1}) = \left[\sum_{x\in\Omega, x\neq x_{ML}} P_C(x|h_{n-1})\log_2 P_C(x|h_{n-1})\right] \tag{23}$$

is the entropy of an ensemble of $N-1$ elements whose value cannot exceed $\log_2(N-1)$. Thus,

$$H(X_n|h_{n-1}) \leq -[(1-P(e|h_{n-1}))\log_2(1-P(e|h_{n-1})) + P(e|h_{n-1})\log_2 P(e|h_{n-1})] + P(e|h_{n-1})\log_2(N-1) \tag{24}$$

Now recall that $P(e|h_{n-1})$ is the probability of $X \neq x_{ML}$ (equation 2). Recall that $P(X = x_{ML}|h_{n-1}) = \pi$.
Therefore:

$$P(e|h_{n-1}) = 1-\pi \tag{25}$$

Substituting 25 into 24:

$$H(X_n|h_{n-1}) \leq -[\pi\log_2\pi + (1-\pi))\log_2(1-\pi)] + (1-\pi)\log_2(N-1) \tag{26}$$

$$\leq S_F(\pi) \tag{27}$$

Where $S_F(\pi)$ is defined in the Supporting Online Material for *Limits of Predictability in Human Mobility* by Song, Qu, Blumm and Barabasi, with $\pi = \pi(h_{n-1})$, which completes the proof.