Regularities Unseen, Randomness Observed: Addenda

James P. Crutchfield[1]

[1] *Complexity Sciences Center & Physics Department, One Shields Avenue, University of California, Davis, California 95656*
Electronic Address: chaos@cse.ucdavis.edu

(Dated: March 8, 2013)

We make points in [1] that used to be simple simple again and correct a sign error. The first addendum concerns the mutual entropy form of excess entropy, showing additional steps to go from Eq. (A23) to Eq. (A24). The second clarifies and corrects a nonstandard use of information gain in Prop. 1. The third corrects a typographical error in the published version.

PACS numbers: 02.50.Ey 05.45.-a 05.45.Tp 89.75.Kd

Contents

I. PROPOSITION 8 OF [1]

Proposition 8: $\mathbf{E} = I[\overrightarrow{S};\overleftarrow{S}]$.

Proof: We rewrite the definition so that we can use the finite-L forms of various entropies:

$$I[\overrightarrow{S};\overleftarrow{S}] \equiv \lim_{L\to\infty} I[\overrightarrow{S}^L;\overleftarrow{S}^L] \,. \tag{1}$$

We begin with the definition of mutual information, Eq. (8) of [1], which expresses I as the difference between two entropies:

$$I[\overrightarrow{S}^L;\overleftarrow{S}^L] = H[\overrightarrow{S}^L] - H[\overrightarrow{S}^L \mid \overleftarrow{S}^L] \,. \tag{2}$$

Recall that $H[\overrightarrow{S}^L] \equiv H(L)$.

Using the conditional entropy chain rule [2] we have

$$\begin{aligned} H[\overrightarrow{S}^L \mid \overleftarrow{S}^L] & \\ &= H[S_0, S_1, \ldots, S_{L-1} | S_{-L}, \ldots, S_{-1}] && (3) \\ &= \sum_{i=0}^{L-1} H[S_i | S_{-L} S_{-L+1} \cdots S_{i-1}] \,. && (4) \end{aligned}$$

Putting these together we have

$$\begin{aligned} I[\overrightarrow{S};\overleftarrow{S}] &= \lim_{L\to\infty} \left[H(L) - \sum_{i=0}^{L-1} H[S_i | S_{-L} S_{-L+1} \cdots S_{i-1}] \right] \\ &= \lim_{L\to\infty} \left[H(L) - \sum_{i=0}^{L-1} H(S_L | S^{L+i}) \right] \,. \end{aligned} \tag{5}$$

Add and subtract $L \times h_\mu$ and rearrange:

$$\begin{aligned} I[\overrightarrow{S};\overleftarrow{S}] &= \lim_{L\to\infty} \left[H(L) - L h_\mu + L h_\mu - \sum_{i=0}^{L-1} H(S_L | S^{L+i}) \right] \\ &= \lim_{L\to\infty} \left[H(L) - L h_\mu \right] \\ &\quad + \lim_{L\to\infty} \left[L h_\mu - \sum_{i=0}^{L-1} H(S_L | S^{L+i}) \right] \end{aligned} \tag{6}$$

Focusing on the second limit, we see that it vanishes:

$$\begin{aligned} \lim_{L\to\infty} & \sum_{i=0}^{L-1} \left[h_\mu - H(S_L | S^{L+i}) \right] \\ &= \sum_{i=0}^{\infty} \lim_{L\to\infty} \left[h_\mu - H(S_L | S^{L+i}) \right] \,, \end{aligned} \tag{7}$$

since, by definition, $\lim_{L\to\infty} H(S_L | S^{L+i}) = h_\mu$ for any fixed i.

In short,

$$I[\overrightarrow{S};\overleftarrow{S}] = \lim_{L\to\infty} \left[H(L) - L h_\mu \right] \,. \tag{8}$$

The LHS of which is $\mathbf{E}$ by Prop. 7 of [1]. $\square$

II. PROPOSITION 1 OF [1]

Proposition 1 introduces a nonstandard form of information gain that compares, in effect, a joint distribution to one of its marginals. Unlike the usual information gain that is non-negative, this form is non-positive.

We have the following corrections (adding minus signs) to make.

1. Eq. (24) of that proposition should read:

$$\Delta H(L) = -\mathcal{D}\left(\Pr(s^L)||\Pr(s^{L-1})\right) \tag{9}$$

This change should also be made to the restatement of Prop. 1 in Appendix A1.

2. Following Prop. 1 in the main text replace the two sentences:

> When this is the case, ... We then sum ... sequences.

with

> We appeal to the consistency conditions for word distributions: $\Pr(s^{L-1}) = \sum_{s_{L-1}} \Pr(s^L)$. This means, in particular, that (i) $\Pr(s^L) > 0 \Rightarrow \Pr(s^{L-1}) > 0$ and (ii) $\Pr(s^{L-1}) = 0 \Rightarrow \Pr(s^L) = 0$. This addresses the boundary case for the information gain when the second distribution assigns zero probability to events of positive probability under the first distribution. (The information gain diverges.) That is, this cannot occur. This and realizing that $\Pr(s^{L-1}) \leq \sum_{s_{L-1}} \Pr(s^L)$ means that this form of the information gain is nonpositive.

3. The first sentence of the immediately following paragraph should read:

> Note that, due to this, it follows from Prop. 1 that $\Delta H(L) \equiv H(L) - H(L-1) \geq 0$, as remarked earlier.

4. Eqs (A3) and (A4) have the incorrect sign. For example, Eq. (A4) should read $= -H(L) + H(L-1)$.

III. EXCESS ENTROPY AS MUTUAL INFORMATION OF [1]

Eq. (54) of [1] is missing an ellipsis and should read:

$$\mathbf{E} = \lim_{L\to\infty} I[S_0 S_1 \ldots S_{L-1}; S_L S_{L+1} \ldots S_{2L-1}] \tag{10}$$

Acknowledgments

The author thanks many for the questions that led to these clarifications, including Benny Brown, Chris Ellison, Dave Feldman, Spencer Mathews, Richard Watson, and Karoline Wiesner.

[1] J. P. Crutchfield and D. P. Feldman. Regularities unseen, randomness observed: Levels of entropy convergence. *CHAOS*, 13(1):25–54, 2003.

[2] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, second edition, 2006.