A simple countable infinite-entropy distribution

(notes by Michael Baer)

It is well known that probability distributions exist with infinite entropy. A common example of such a distribution is the uniform distribution on $[0,1]$. There are also instances of events with a countable number of outcomes having infinite entropy. Kato, Han, and Nagaoka ("Huffman Coding with an Infinite Alphabet," *IEEE Trans. Inf. Theory*, vol. IT-42, No. 3, 977–984, May 1996) construct an abstract one in their paper, but, curiously, simple, concrete examples seem lacking. Here we give one.

First, let us use the convention that ln denotes the natural logarithm $\log_e$, while lg denotes $\log_2$. Consider the probability mass function over the positive integers for event X characterized by

$$p(k) = \Pr[X = k] = \frac{1}{\lg(k+1)} - \frac{1}{\lg(k+2)}.$$

Clearly this is positive over k and sums to 1. In order to prove that this distribution has infinite entropy, we need to first derive a few basic properties of related functions.

First, the monotonicity and convexity of the related function $1/\ln x$ is easily seen by noting that its first derivative is $-(x\ln^2 x)^{-1}$ (negative for $x > 1$) and its second derivative is $(\ln x + 2)/(x^2 \ln^3 x)$ (positive for $x > 1$). This will allow us to bound $p(k)$ from below with $\tilde{p}(k) = (\ln 2)/(k \ln^2 k)$. Even easier to show via basic calculus is that $p \ln p$ is monotonically decreasing with p for $p < e^{-1}$. Also, $\tilde{p}(k)$ is monotonically decreasing, so $\tilde{p}(k) \ln \tilde{p}(k)$ is montonically decreasing for $k \geq 2$ since $\tilde{p}(k) < e^{-1}$ for these k. Also $\ln^2 m$ is greater than $\ln 2$ for $m \geq 3$. We can thus lower-bound $H(p)$ in nats by an expression equal to infinity, thereby showing that $H(p)$ is infinite:

$$\begin{aligned}
H(p) &= -p(1)\ln p(1) - p(2)\ln p(2) - \sum_{m=5}^{\infty} \left(\frac{1}{\lg(m-1)} - \frac{1}{\lg m}\right) \ln\left(\frac{1}{\lg(m-1)} - \frac{1}{\lg m}\right) \\
&> (\ln 2) \sum_{m=5}^{\infty} \frac{1}{m \ln^2 m} \left(\ln m + \ln\ln^2 m - \ln\ln 2\right) \\
&> (\ln 2) \sum_{m=5}^{\infty} \frac{1}{m \ln m} > (\ln 2) \int_5^{\infty} \frac{dm}{m \ln m} \\
&> \lim_{n\to\infty} (\ln 2)(\ln\ln n - \ln\ln 5) = \infty
\end{aligned}$$

and we thus have our countable infinite-entropy distribution. A Shannon code for this distribution must satisfy the Kraft inequality and has a redundancy of no more than 1, using the definition of Klimesh ("Redundancy and Optimality of Codes for Infinite-Entropy Sources," *Proc., International Symposium on Information Theory*, 1949–1953, July 2008). One good question might be whether a well-known asymptotically optimal universal prefix code — such as Elias' ω code ("Universal Codeword Sets and Representations of the Integers," *IEEE Trans. Inf. Theory*, vol. IT-21, No. 2, 194–203, March 1975) — has bounded redundancy. This is not guaranteed by the definintion of asymptotic optimality, in which the ratio, not the difference, between redundancy and entropy must go to zero as entropy goes to infinity.