UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: Probability and Random Processes

Discussion 5
Fall 2019

---

1. **Entropy Warmup** Entropy seems like a very weird concept. We'll walk through a few examples together to build up your intuition. Let's say that we have a random variable X that can take on values from $\{\text{lecture}, \text{midterm}, \text{pop quiz}\}$. (Don't worry, we don't actually have pop quizzes in this class). Each day you go to class, you observe a random value of X, which is determined according to the distribution P_X. If $P(\text{lecture}) = 0.85$, $P(\text{midterm}) = 0.1$, and $P(\text{pop quiz}) = 0.05$, we'd mainly see lectures, occasionally have a midterm, and have a pop quiz very rarely. We can describe how "interesting" it is to see a particular $X = x$ with the notion of the "surprise," which is a function $S(x) = \log_2 \frac{1}{P_X(x)}$. This function is large for low-probability events and small for high-probability events.

   (a) For the probabilities above, calculate $S(\text{lecture})$, $S(\text{midterm})$, and $S(\text{pop quiz})$.

   (b) If $P(\text{lecture}) = \frac{1}{3}$, $P(\text{midterm}) = \frac{1}{3}$, and $P(\text{pop quiz}) = \frac{1}{3}$, calculate the surprises again. Given that $\log_2 \frac{1}{0.85} = 0.234$, $\log_2 \frac{1}{1/3} = 1.58$, $\log_2 \frac{1}{0.1} = 3.32$, and $\log_2 \frac{1}{0.05} = 4.32$, do the relative magnitudes of the values in (a) and (b) make sense intuitively?

   (c) The entropy is the *expected surprise*. Formally,

   $$H(X) = \sum_x P_X(x) S(x) = \sum_x P_X(x) \log_2 \frac{1}{P_X(x)}$$

   We will follow the convention that, if for a particular x, $P_X(x) = 0$, then $P_X(x) \log_2 \frac{1}{P_X(x)} = 0$. Calculate the entropy for the original probability values $(0.85, 0.1, 0.05)$, the entropy of the uniform distribution $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and the entropy of the deterministic RV with distribution $(1, 0, 0)$.

   (d) Do the entropies in part (c) make sense to you?

2. **Breaking a Stick**

   I break a stick n times, where n is a positive integer, in the following manner: the ith time I break the stick, I keep a fraction X_i of the remaining stick where X_i is uniform on the interval $[0, 1]$ and $X_1, X_2, \dots, X_n$ are i.i.d. Let $P_n = \prod_{i=1}^n X_i$ be the fraction of the original stick that I end up with.

   (a) Show that $P_n^{1/n}$ converges almost surely to some constant function.

   (b) Compute $\mathbb{E}[P_n]^{1/n}$.

3. Bounds on Entropy

There's actually a limit to how much "randomness" there is in a random variable X that takes on $|\mathcal{X}|$ distinct values. Prove that for any distribution p_X, $H(X) \leq \log |\mathcal{X}|$.

Hint: We can write $H(X)$ as $\mathbb{E}[\log \frac{1}{P_X(X)}]$. By Jensen's inequality, if we have a random variable Z and a concave function f, $f(\mathbb{E}[Z]) \geq \mathbb{E}[f(Z)]$.