UC Berkeley
Department of Electrical Engineering and Computer Sciences

EECS 126: PROBABILITY AND RANDOM PROCESSES

Discussion 7
Fall 2023

1. **Sum of Rolls**

   You roll a fair 6-sided die 100 times, and you call the sum of the values of all your rolls X. Use the Central Limit Theorem to approximate the probability that $X > 400$. You may use a calculator and Gaussian lookup table.

   Solution: The value of an individual roll, distributed as Uniform([6]), has mean 3.5 and variance $\frac{6^2-1}{12} = \frac{35}{12}$. Call $\sigma = \sqrt{\frac{35}{12}}$. Then, by the Central Limit Theorem, $(X - 350)/(10\sigma)$ is approximately $\mathcal{N}(0,1)$ distributed, i.e. X is approximately $\mathcal{N}(350, 100\sigma^2)$. Therefore

$$\begin{aligned}\mathbb{P}(X > 400) &\approx \mathbb{P}(\mathcal{N}(350, 100\sigma^2) > 400) \\ &= \mathbb{P}(\mathcal{N}(0,1) > \tfrac{400-350}{10\sigma}) \\ &\approx 1 - \Phi(2.93) \\ &= 1 - 0.9983 = 0.0017.\end{aligned}$$

2. **Entropy of a Sum**

Let X_1, X_2 be i.i.d. Bernoulli($\frac{1}{2}$). Calculate $H(X_1+X_2)$ and show that $H(X_1+X_2) \geq H(X_1)$. Does this make intuitive sense?

Solution: $X_1 + X_2$ has the following distribution.

$$X_1 + X_2 = \begin{cases} 0 & \text{with probability } \frac{1}{4}, \\ 1 & \text{with probability } \frac{1}{2}, \\ 2 & \text{with probability } \frac{1}{4}, \end{cases}$$

Thus, the entropy of $X_1 + X_2$ is

$$H(X_1 + X_2) = -\frac{1}{4}\log_2 \frac{1}{4} - \frac{1}{2}\log_2 \frac{1}{2} - \frac{1}{4}\log_2 \frac{1}{4} = \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = \frac{3}{2},$$

which is greater than $H(X_1) = 1$. Intuitively, we might expect the sum of independent random variables to "have more randomness" than each individual random variable, so this makes sense because we think of entropy as a measure of randomness. In fact, it is generally true that adding independent random variables increases entropy.

One can also prove this using properties of entropy (which you may or may not see later in HW):

$$H(X_1, X_2) = H(X_1 + X_2, X_2) \leq H(X_1 + X_2) + H(X_2) \tag{1}$$

$$H(X_1, X_2) = H(X_1) + H(X_2) \quad \text{Using Independence.} \tag{2}$$

Note that the above result is not true in general if X_1 and X_2 are not independent.

3. **Mutual Information and Channel Coding**

The *mutual information* of X and Y is defined as

$$I(X;Y) := H(X) - H(X \mid Y),$$

where $H(X \mid Y)$ is the *conditional entropy* of X given Y,

$$\begin{aligned} H(X \mid Y) &= \sum_{y \in \mathcal{Y}} p_Y(y) \cdot H(X \mid Y = y) \\ &= \sum_{y \in \mathcal{Y}} p_Y(y) \sum_{x \in \mathcal{X}} p_{X|Y}(x \mid y) \log_2 \frac{1}{p_{X|Y}(x \mid y)}. \end{aligned}$$

Conditional entropy can be interpreted as the average amount of uncertainty remaining in the random variable X after observing Y. Then, mutual information is the amount of information about X gained by observing Y.

Now, the channel coding theorem says that the capacity of a channel with input X and output Y is the maximal possible amount of mutual information between them:

$$C = \max_{p_X} I(X;Y) = \max_{p_X} H(X) - H(X \mid Y).$$

a. Let X be the roll of a fair die and $Y = \mathbb{1}_{X \geq 5}$. What is $H(X \mid Y)$?

b. Suppose the channel is a noiseless binary channel, i.e. $X \in \{0, 1\}$ and $Y = X$. Use the theorem above to find its capacity C.

Solution:

a. $Y = 1$ with probability $\frac{1}{3}$, in which case X is equally likely to be 5 or 6, so $H(X \mid Y = 1) = \log_2(2) = 1$. In the other case, i.e. $Y = 0$ with probability $\frac{2}{3}$, X is equally likely to be 1 through 4, so $H(X \mid Y = 0) = \log_2(4) = 2$. Thus

$$H(X \mid Y) = \frac{1}{3} \cdot 1 + \frac{2}{3} \cdot 2 = \frac{5}{3}.$$

b. For a noiseless binary channel, $H(X \mid Y) = 0$: after observing Y, we know X certainly.

$$C = \max_{p_X} H(X) - H(X \mid Y) = \max_{p_X} H(X) - 0 = \log_2(2) = 1.$$

In other words, every bit we send over the channel also carries 1 bit of information.