1 Entropy Scale

Consider the following scenario: you have 9 coins, and you know that all of the coins have the same weight, *except* one of the coins (which is heavier than the rest). You have a scale at your disposal which will tell you if one set of coins is heavier, lighter, or the same weight as another set of coins. Your goal is to find the heavier coin with the fewest number of weighings.

(a) At the beginning, there are 9 possibilities for the identity of the heavier coin, and each possibility is equally likely. Calculate the entropy using the base-3 logarithm.

(b) Now suppose that you pick up two sets of three coins each and weight the two sets against each other. What is the entropy after the weighing?

(c) From the previous part, you should have been able to narrow down the possibilities to just three coins. Now, you pick two of those three coins and weigh them against each other. What is the entropy after the weighing?

(d) Give an intuitive argument for why no scheme can find the heavier coin in fewer weighings.

2 Relative Entropy of the Gaussian

Recall that we defined the relative entropy of distribution q from distribution p via

$$D_{\mathrm{KL}}(\mathsf{p} \,\|\, \mathsf{q}) = \sum_{x \in \mathscr{X}} \mathsf{p}(x) \log_2 \frac{\mathsf{p}(x)}{\mathsf{q}(x)}.$$

By analogy, we define the continuous relative entropy of the density function g from the density function f by

$$D_{\mathrm{KL}}(f \,\|\, g) = \int_{-\infty}^{\infty} f(x) \ln \frac{f(x)}{g(x)} \, \mathrm{d}x$$

(we choose to use the natural logarithm for this problem out of convenience). Let $X \sim \mathcal{N}(\mu_X, 1)$ and $Y \sim \mathcal{N}(\mu_Y, 1)$. Compute $D_{\mathrm{KL}}(f_X \,\|\, f_Y)$.

3 Polling Your Constituency

Suppose the whole population of California has 60% Democrats, 40% Republicans, and no other parties. You choose 1000 people independently and uniformly at random from the Californian population, and for each person, you record whether s/he is a Democrat or a Republican. Let S denote the number of Democrats among the 1000 people. What is the probability that you chose more than 625 Democrats?

(a) Use Markov's inequality.

(b) Use Chebyshev's inequality.

(c) Use the Chernoff bound.

(d) Use Hoeffding's inequality, which is stated below:

If $X_1, \ldots, X_n$ are i.i.d. with mean μ and $|X_i| \leq 1$, then for $\lambda > 0$,

$$\mathbb{P}\left(\frac{X_1 + \cdots + X_n}{n} - \mu \geq \lambda\right) \leq \exp(-2\lambda^2 n).$$

[You will prove Hoeffding's inequality in the homework.]

(e) Approximate by the CLT.

4 Entropy

[*This problem is optional and may not be covered during discussion.*]

Let X_i, $1 \leq i \leq n$, be a sequence of i.i.d. Bernoulli random variables with parameter p, i.e. $\mathbb{P}(X_i = 1) = p$ and $\mathbb{P}(X_i = 0) = 1 - p$.

(a) Express $\mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$ in terms of p and n_0, where $x_i \in \{0, 1\}$ for all $1 \leq i \leq n$ and n_0 depends on $(x_1 \ldots x_n)$ and represents the number of 1's in $(x_1, \ldots, x_n)$. Call this result $f(x_1, x_2, \ldots, x_n)$.

(b) Define the random variable Y_n as $Y_n = -n^{-1} \log f(X_1, X_2, \ldots, X_n)$. What does Y_n tend to as n grows large?