EE Qualifying Exam 2015 – John Duchi

Question 1: Let $\delta \in (0, 1)$. I have two unfair coins in my pocket. The first coin, when flipped, lands heads with probability 1. The second coin, when flipped, lands heads with probability $1 - \delta$ and tails with probability δ. I give you one of the coins but do not tell you which one. How many flips do you need to guarantee that, no matter which coin I gave you, you can identify which coin it is with probability at least 3/4?

Answer: Let P_1 be the distribution of heads under the first coin, P_2 the second. The probability of seeing no heads after n flips for coin 1 is always 1; the probability of seeing 0 heads after n flips for coin 2 is

$$P_2(\text{no heads after } n \text{ flips}) = (1 - \delta)^n.$$

If our procedure is "guess coin 1 if there are no heads and guess coin 2 otherwise," then we never make a mistake with coin 1, while with coin 2 we have $P_2(\text{mistake}) = (1 - \delta)^n$. To get $(1 - \delta)^n \le 1/4$, it is sufficient that $n \log(1 - \delta) \le \log \frac{1}{4}$, and using the approximation $\log(1 - \delta) \approx -\delta$, we see that $n \ge \frac{\log 4}{\delta}$ coin flips are sufficient. $\square$

Question 2: Let $\delta \in (0, 1)$. I have two unfair coins in my pocket. The first coin, when flipped, lands heads with probability $\frac{1+\delta}{2}$ and tails with probability $\frac{1-\delta}{2}$. The second coin, when flipped, lands heads with probability $\frac{1-\delta}{2}$ and tails with probability $\frac{1+\delta}{2}$. I give you one of the coins but do not tell you which one. How many flips do you need to guarantee that, no matter which coin I gave you, you can identify which coin it is with probability at least 3/4?

Answer: Let P_1 and P_2 be the distributions of the coins as in Question 2. We flip the coin given n times, and if there are more heads report coin 1, more tails report coin 2. Now we must approximate the probability of success under this scheme.

There are several ways to answer this question. Let X_i be 1 if the ith flip is heads, 0 otherwise. The simplest is to recall the central limit theorem. We know that $\sigma^2 = \mathrm{Var}(X_i) = \frac{1-\delta^2}{4}$ for either distribution, and the CLT implies that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} (X_i - \mathbb{E}[X_i]) \xrightarrow{d} \mathsf{N}(0, \sigma^2).$$

Thus, if $\Phi(t) = \mathbb{P}(Z \le t)$, where Z is a standard normal random variable, we (approximately) have that

$$\mathbb{P}\left(\frac{1}{\sqrt{n \,\mathrm{Var}(X)}} \sum_{i=1}^{n} (X_i - \mathbb{E}[X_i]) \le t \right) \approx \Phi(t)$$

for $t \in \mathbb{R}$. Now, let us focus on coin 1 (the cases are symmetric). Then we have that

$$P_1(\text{wrong after } n \text{ flips}) = P_1\left(\sum_{i=1}^{n} X_i \le \frac{n}{2} \right) = P_1\left(\sum_{i=1}^{n} \left(X_i - \frac{1+\delta}{2} \right) \le \frac{-\delta n}{2} \right),$$

and by the CLT, we have (noting that $\mathrm{Var}(X) = \frac{1-\delta^2}{4}$)

$$P_1(\text{wrong after } n \text{ flips}) = P_1\left(\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left(X_i - \frac{1+\delta}{2}\right) \le \frac{-\delta\sqrt{n}}{2}\right)$$
$$\approx \Phi\left(\frac{-\delta\sqrt{n}}{2\sqrt{\mathrm{Var}(X)}}\right) = \Phi\left(\frac{-\delta\sqrt{n}}{\sqrt{1-\delta^2}}\right).$$

For this probability to be small—i.e. for us to be unlikely to make a mistake—we see that we need $\sqrt{n}\delta$ to be large, or $n \approx \frac{1}{\delta^2}$.

A second way to see this result is to recall Hoeffding's bound (or the Chernoff bound), which says that if a sequence of independent random variables $X_i \in [0,1]$, then

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \ge t\right) \le \exp\left(-2nt^2\right) \quad \text{for all } t \ge 0.$$

(And an identical inequality holds for the event $\le -t$.) Focusing on the first coin, we see that if $X_i = 1$ if the ith flip is heads, $X_i = 0$ if the ith flip is tails, then $\mathbb{E}[X_i] = (1+\delta)/2$, and

$$P_1(\text{wrong}) = P_1\left(\sum_{i=1}^{n} X_i \le \frac{n}{2}\right) = P_1\left(\frac{1}{n}\sum_{i=1}^{n} X_i \le \frac{1}{2}\right)$$
$$= P_1\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i]) \le -\frac{\delta}{2}\right) \le \exp\left(-\frac{n\delta^2}{2}\right),$$

by Hoeffding's bound. To guarantee that this bound is less than 1/4, we require

$$n \ge \frac{2\log 4}{\delta^2}.$$

This is much worse than $1/\delta$. $\square$

Question 3: Can you design a procedure that does better than the one you identified in question 1? Can you design a procedure that does better than the one you identified in question 2? Can you give a lower bound on the number of flips needed to identify the coin in problems 1 and 2?

Answer: The procedures given above are both optimal, as they are likelihood ratio tests for the simple problem of deciding whether we have coin 1 or coin 2, so they are not improvable. (This is the Neyman-Pearson lemma.)

The bounds are also—in terms of δ—sharp. For the first question, this is an immediate consquence of the Neyman-Pearson lemma. For the second question, the normal approximation shows that asymptotically (as $n \to \infty$, or as $\delta \to 0$), the likelihood ratio test looks normal, so that we require $n \approx 1/\delta^2$ coin flips. It is possible to make this more rigorous.

I did not expect anyone to be able to answer for the second part, and you can look up a solution on your own. But for some intuition, note that there is a single bit of information in our system: whether we have coin 1 or coin 2. Each flip of the coin gives a number of bits roughly equal to $D_{\rm kl}\left(P_1 \| P_2\right)$, and n flips gives $D_{\rm kl}\left(P_1^n \| P_2^n\right) = n D_{\rm kl}\left(P_1 \| P_2\right)$ In this case, of course,

$$D_{\rm kl}\left(P_1 \| P_2\right) = \frac{1+\delta}{2} \log \frac{1+\delta}{1-\delta} + \frac{1-\delta}{2} \log \frac{1-\delta}{1+\delta} = \delta \log \frac{1+\delta}{1-\delta},$$

and for $\delta \in [0, \frac{1}{2}]$, we have $D_{\rm kl}\left(P_1 \| P_2\right) \le 3\delta^2$. Thus we observe at most $3n\delta^2$ bits of information after n flips, so that $n \gtrsim 1/\delta^2$ flips are necessary. $\square$