

These notes are from <http://math.cmu.edu/~cocox/teaching/discrete20/rec10.pdf>

In a nearby prison, a bored guard decides to play a game with the prisoners. The guard has with her hats of k different colors. She tells the prisoners that she'll place a hat on each of their heads: if the prisoner guesses their hat's color correctly, they'll be freed, otherwise she'll prolong their sentence. The prisoners will not be allowed to look at their own hat, nor communicate in any way with the other prisoners once the hats are placed; however, each prisoner can see the hats of every other prisoner.

The guard tells the prisoners that they can meet beforehand to discuss a strategy; however, the guard may be eavesdropping and will try to foil any plan devised by the prisoners. Furthermore, the prisoners' strategy must be deterministic. We will formalize what is meant by a "deterministic strategy" in a moment, but informally this means that no form of randomness can be used in the strategy. The prisoners decide to act selflessly and try to maximize the total number of correct guesses. If there are n prisoners and k hat colors, how many correct guesses can be guaranteed?

Since there are k different hat colors, we may label them $0, 1, \dots, k-1$. In other words, the hat placement is a function $h: P \rightarrow \mathbb{Z}/k\mathbb{Z}$. Partition the prisoners into $P = P_0 \sqcup \dots \sqcup P_{k-1}$ so that $|P_i| \in \{\lfloor n/k \rfloor, \lceil n/k \rceil\}$ for each $i \in \mathbb{Z}/k\mathbb{Z}$. Those prisoners in P_i , will assume that the sum of all the hats is i modulo k and guess accordingly. That is, a prisoner $p \in P_i$ will guess color

$$i - \sum_{q \neq p} h(q) \pmod k.$$

Since $\sum_{p \in P} h(p)$ is congruent to some value modulo k and $h(p) = (\sum_{p \in P} h(p)) - (\sum_{q \neq p} h(q)) \pmod k$, all of the prisoners in some part are correct; hence there are at least $\min_i |P_i| = \lfloor n/k \rfloor$ correct guesses, regardless of how the guard actually placed the hats.

Question. With n prisoners and k hat colors, is there a deterministic strategy which can guarantee to save strictly more than $\lfloor n/k \rfloor$ of the prisoners (no matter how the guard places the hats)?

It turns out that the answer is "no". In order to show this, we need to formalize what is meant by a "deterministic strategy" and use randomness to our advantage.

Let H be the set of hat colors and P be the set of prisoners so a hat placement is just a function $h: P \rightarrow H$; in other words H^P is the set of all hat placements. A *deterministic strategy* is a tuple of functions $(f_p)_{p \in P}$ where $f_p: H^{P \setminus \{p\}} \rightarrow H$. That is, each prisoner $p \in P$ has a function f_p whose input is the hat placement on all prisoners other than themselves and whose output is a hat color: given a hat placement $h \in H^P$, prisoner $p \in P$ guesses color $f_p((h(q))_{q \neq p})$. In particular, each prisoner's guess is determined fully by the hats on the other prisoners' heads.

Let $(f_p)_{p \in P}$ be a deterministic strategy and let $h \in H^P$ be a hat placement. Observe that prisoner p guesses correctly if and only if

$$f_p((h(q))_{q \neq p}) = h(p).$$

Lastly, recall that the guard can eavesdrop on the prisoners' planning meeting. In other words, the guard knows the strategy $(f_p)_{p \in P}$ and can select the hat placement based on this.

Claim 1. Suppose that $|P| = n$ and $|H| = k$. For any deterministic strategy $(f_p)_{p \in P}$, there is a hat placement in which at most $\lfloor n/k \rfloor$ prisoners guess correctly.

Proof. Consider choosing a hat placement uniformly at random from H^P and let X be the random variable which counts the total number of correct guesses. Formally, $X : H^P \rightarrow \mathbb{R}$ where

$$X(h) = \sum_{p \in P} \mathbf{1}[f_p((h(q))_{q \neq p}) = h(p)].$$

Using linearity of expectation, we see that

$$\mathbb{E} X = \sum_{p \in P} \mathbb{E} \mathbf{1}[f_p((h(q))_{q \neq p}) = h(p)] = \sum_{p \in P} \Pr[f_p((h(q))_{q \neq p}) = h(p)].$$

For a $p \in P$, we can partition the probability space based on the value of $(h(q))_{q \neq p}$, so we can use the law of total probability to compute

$$\begin{aligned} \Pr[f_p((h(q))_{q \neq p}) = h(p)] &= \sum_{x \in H^{P \setminus \{p\}}} \Pr[f_p((h(q))_{q \neq p}) = h(p) \mid (h(q))_{q \neq p} = x] \Pr[(h(q))_{q \neq p} = x] \\ &= \sum_{x \in H^{P \setminus \{p\}}} \Pr[h(p) = f_p(x)] \Pr[(h(q))_{q \neq p} = x] \\ &= \sum_{x \in H^{P \setminus \{p\}}} \frac{1}{k} \Pr[(h(q))_{q \neq p} = x] = \frac{1}{k}, \end{aligned}$$

since $\Pr[h(p) = c] = 1/k$ for any $c \in H$ and $\sum_{x \in H^{P \setminus \{p\}}} \Pr[(h(q))_{q \neq p} = x] = 1$. Therefore,

$$\mathbb{E} X = \sum_{p \in P} \Pr[f_p((h(q))_{q \neq p}) = h(p)] = \sum_{p \in P} \frac{1}{k} = \frac{n}{k}.$$

Finally, there must exist some $h \in H^P$ for which $X(h) \leq \mathbb{E} X = n/k$. Since $X(h)$ is the number of correct guesses, which is an integer, there actually must be some $h \in H^P$ which which $X(h) \leq \lfloor n/k \rfloor$. Hence, there is some hat placement in which at most $\lfloor n/k \rfloor$ prisoners guess correctly. \square

Exercise: Suppose there are three prisoners and two hat colors. However, the guard didn't plan ahead and has only two hats of each color (so all three prisoners cannot be given the same hat color). Find a deterministic strategy which guarantees to save two prisoners.¹

Moreover, fix any deterministic strategy for the three prisoners. Show that if there's a way for the guard to place hats which ensures three correct guesses, then there is also a way for the guard to place hats that ensures at most one correct guess. In other words, *any* deterministic strategy in this setting which has the "potential" to save all three prisoners cannot guarantee to always save two prisoners (this may be a good fact to keep in mind while trying to find a strategy in the problem above). Hint: imitate the proof of Claim 1 and carefully analyze the application of the law of total probability.

¹Interestingly, it can be shown that, up to relabeling the prisoners, there is a *unique* strategy which accomplishes this task.