Three Variants of a Simple Card Game

iMathTutor® • 20 Jan 2025

Introduction

The problem considered in this post and the problems in the two posts immediately prior to this one, here and here, were among a number of problems posed to me last semester by a client who was taking an upper division probability course at Penn State. I found these three problems particularly interesting. I hope you do too.

Variants of the Game

A deck of cards numbered 1 through n is shuffled such that all possible $n!$ orderings are equally like. Cards are removed from the deck sequentially until no more cards remain. At each step you make a guess as to the the value of the card removed. Three variants of the game are considered.

Variant 1:
You are not given any information about the cards removed at previous step before you make a guess at the current step.

Variant 2:
At each step, you are shown the card that is drawn.

Variant 3:
You are told at each step if your guess is correct or not, but not the value of the card removed.

The Problem

Under each of the three variants find a guessing strategy that maximizes the expected number of correct guesses and find the the maximum.

Solutions

Two different methods of solution are used, one for both variants 1 and 2, and a second for variant 3.

Set $\mathbf{s} = (s_1, s_2, \ldots, s_n) \in [n]^n$ [1] where s_i is your guess for the card at step $i = 1, 2, \ldots, n$. $\mathbf{s}$ represents your strategy. Denote the number of correct guesses under strategy $\mathbf{s}$ by $N(\mathbf{s})$.

The first method is the standard approach used to find the expected number of success in a sequence of Bernoulli trials. In this approach a sequence of Bernoulli random variables $X_1(\mathbf{s}), X_2(\mathbf{s}), \ldots . X_n(\mathbf{s})$ by

$$X_i(\mathbf{s}) := \begin{cases} 1 & \text{if the guess at step } i \text{ under } \mathbf{s} \text{ is correct} \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$

$i = 1, 2 \ldots, n$ is introduced. Clearly,

$$N(\mathbf{s}) = X_1(\mathbf{s}) + X_2(\mathbf{s}) + \cdots + X_n(\mathbf{s}). \tag{2}$$

The linearity of expectation, and the expected value of Bernoulli random variable give

$$\begin{aligned} \mathbf{E}\left[N(\mathbf{s})\right] =& \mathbf{E}\left[X_1(\mathbf{s})\right] + \mathbf{E}\left[X_2(\mathbf{s})\right] + \cdots + \mathbf{E}\left[X_n(\mathbf{s})\right] \\ =& \mathbf{P}\left[X_1(\mathbf{s}) = 1\right] + \mathbf{P}\left[X_2(\mathbf{s}) = 1\right] + \cdots + \mathbf{P}\left[X_n(\mathbf{s}) = 1\right]. \end{aligned} \tag{3}$$

This reduces the problem to find $\mathbf{P}\left[X_i(\mathbf{s}) = 1\right], i = 1, 2 \ldots, n.$

The second approach starts with a well-known result on the expectation of nonnegative discrete random variables:

$$\mathbf{E}\left[N(\mathbf{s})\right] = \sum_{r=1}^{n} \mathbf{P}\left[N(\mathbf{s}) \geq r\right]. \tag{4}$$

Let $Y_r(\mathbf{s})$ denote the number of trials needed to observe $r = 1.2. \ldots, n$ correct guesses under strategy $\mathbf{s}$. One has

$$N(\mathbf{s}) \geq r \Longleftrightarrow Y_r(\mathbf{s}) \leq n. \tag{5}$$

Consequently

$$\mathbf{E}\left[N(\mathbf{s})\right] = \sum_{r=1}^{n} \mathbf{P}\left[Y_r(\mathbf{s}) \leq n\right]. \tag{6}$$

At this point the problem is reduced to finding the marginal distributions of $Y_r(\mathbf{s}), r = 1, 2, \ldots, n.$

The reason that two methods are employed is that in the first two variants, the information about the cards removed prior to a given step does not depend on the the outcomes of the prior steps, while in the third variant that information does depend on the outcomes of the prior steps. So if the first method were used to solve the third variant one would have to condition on the prior steps. One way to handle this would be to introduce the Y sequence which would be conditioned on. This would necessitate finding the marginal distributions of the random variables in this sequence, it makes sense to simply use them directly to find the expectation.

Variant 1

Let $G_i(\mathbf{s})$ denote the event that your guess at step $i = 1, 2, \ldots, n$ is a card that remains in the deck under strategy $\mathbf{s}$. Conservation of total probability gives

$$\begin{aligned}\mathbf{P}\left[X_i(\mathbf{s}) = 1\right] =& \mathbf{P}\left[X_i(\mathbf{s}) = 1 | G_i(\mathbf{s})\right]\mathbf{P}\left[G_i(\mathbf{s})\right] + \mathbf{P}\left[X_i(\mathbf{s}) = 1 | G_i(\mathbf{s})^c\right]\mathbf{P}\left[G_i(\mathbf{s})^c\right] \\ =& \mathbf{P}\left[X_i(\mathbf{s}) = 1 | G_i(\mathbf{s})\right]\mathbf{P}\left[G_i(\mathbf{s})\right]\end{aligned} \tag{7}$$

In this variant, there is no information about the cards removed in prior steps. Hence for any strategy

$$\mathbf{P}\left[G_i(\mathbf{s})\right] = \frac{\binom{n-i+1}{1}}{\binom{n}{1}} = \frac{n-i+1}{n}. \tag{8}$$

Further,

$$\mathbf{P}\left[X_i(\mathbf{s}) = 1 | G_i(\mathbf{s})\right] = \frac{1}{\binom{n-i+1}{1}} = \frac{1}{n-i+1}. \tag{9}$$

Thus

$$\mathbf{P}\left[X_i(\mathbf{s}) = 1\right] = \frac{1}{n} \tag{10}$$

$i = 1, 2, \ldots, n$. This is substituted into (3) to obtain

$$\mathbf{E}\left[N(\mathbf{s})\right] = 1, \tag{11}$$

for all $\mathbf{s}$.

Variant 2

In this variant, at each step you know all of the cards that remain in the deck. Thus for any strategy $\mathbf{s}'$ which selects a card that remains in the deck at each step

$$\mathbf{P}\left[G_i(\mathbf{s}')\right] = 1, \tag{12}$$

$i = 1, 2, \ldots, n$, In this case, (7) gives

$$\mathbf{P}\left[X_i(\mathbf{s}') = 1\right] = \mathbf{P}\left[X_i(\mathbf{s}') = 1 | G_i(\mathbf{s})\right] = \frac{1}{n-i+1}. \tag{13}$$

This is substituted in (2) to yield

$$\mathbf{E}\left[N(\mathbf{s}')\right] = \sum_{i=1}^{n} \frac{1}{n-i+1} = \sum_{k=1}^{n} \frac{1}{n} \tag{14}$$

for any strategy $\mathbf{s}'$ which selects a card at each step that remains in the deck. For any strategy $\mathbf{s}''$ that selects a card not in the deck as some step j, $\mathbf{P}\left[G_j(\mathbf{s}'')\right] < 1$. Hence $\mathbf{E}\left[N(\mathbf{s}'')\right] < \mathbf{E}\left[N(\mathbf{s}')\right]$. Therefore equation (14) gives the maximum.

Variant 3

In this variant, at the first step you know your guess is a card in the deck, at each step afterward you only know which of the cards were removed form the deck at previous steps where your guess was correct, but you do not know what cards were removed from the deck at previous steps at which your guess was incorrect. As will be seen, the optimal strategy for this variant is to stick with a guess until it is correct, and once it is correct change your guess to a guess you have not yet made. Let τ denote such a strategy.

As discussed above the equation

$$\mathbf{E}\left[N(\mathbf{s})\right] = \sum_{r=1}^{n} \mathbf{P}\left[Y_r(\mathbf{s}) \leq n\right]. \tag{15}$$

where Y_r is the number of steps until r correct guess are made,$r = 1, 2, \ldots, n$ is used here. This necessitates finding the distributions of these "negative binomial"-like random variables.

Since the first guess is always a card that remains in the deck,

$$\mathbf{P}\left[Y_1(\tau) = 1\right] = \frac{1}{n} \tag{16}$$

For $1 < i \leq n$, $\mathbf{P}\left[Y_1(\tau) = i\right]$ is found by a straightforward conditioning argument. The first correct guess occurs at a given step if and only if at all previous steps the guess was incorrect. Thus

$$\begin{aligned}\mathbf{P}\left[Y_1(\tau)=i\right]=&\mathbf{P}\left[Y_1(\tau)=1, X_{i-1}=0, X_{i-2}=0, \ldots, X_1=0\right]\\ =&\mathbf{P}\left[Y_1(\tau)=1 | X_{i-1}=0, X_{i-2}=0, \ldots, X_1=0\right]\\ &\times \mathbf{P}\left[X_{i-1}=0 | X_{i-2}=0, \ldots, X_1=0\right] \times \cdots \times \mathbf{P}\left[X_1=0\right]\\ =&\mathbf{P}\left[Y_1(\tau)=1 \middle| \bigcap_{k=1}^{i-1} X_{i-k}=0\right]\\ &\times \prod_{k=1}^{i-2} \mathbf{P}\left[X_{i-k}=0 \middle| \bigcap_{k=1}^{i-2} X_{i-k-1}=0\right] \mathbf{P}\left[X_1=0\right],\end{aligned} \tag{17}$$

where the dependence of the Xs on the strategy has been suppressed to uncluttered the notation; This should not lead to any confusion. Under the strategy τ the guess at step i and all previous steps is the same as the initial guess whenever no correct guesses have occurred in the previous steps, further the initial guess remains a card in the deck at all steps up to i. Consequently

$$\mathbf{P}\left[Y_1(\tau)=1 \middle| \bigcap_{k=1}^{i-1} X_{i-k}=0\right] =\frac{1}{n-1+1}, \tag{18}$$

$$\mathbf{P}\left[X_{i-k}=0 \middle| \bigcap_{k=1}^{i-2} X_{i-k-1}=0\right] =\frac{n-i+k}{n-i+k+1}, \qquad \text{and} \tag{19}$$

$$\mathbf{P}\left[X_1=0\right] =\frac{n-1}{n} \tag{20}$$

Therefore for all $1 \leq i \leq n$,

$$\mathbf{P}\left[Y_1(\tau)=i\right]=\frac{1}{n}. \tag{21}$$

Conservation of total probability yields the following set of coupled equations for the distributions of the other Ys.

$$\begin{aligned}\mathbf{P}\left[Y_{r+1}(\tau)=i\right]=\sum_{k=r}^{i-1}&\left(\mathbf{P}\left[Y_{r+1}(\tau)=i | G_{k+1}(\tau), Y_r(\tau)=k\right]\right.\\ &\left.\times \mathbf{P}\left[G_{k+1}(\tau) | Y_r(\tau)=k\right] \mathbf{P}\left[Y_r(\tau)=k\right]\right)\end{aligned} \tag{22}$$

for $i>r$, and it is zero for $i \leq r$; $1 \leq r<n$. Let $Y_1'(\tau)$ denote the number of trial needed to observe the first correct guess in a deck with $n-k$ cards. Then for $r \leq k<i$

$$\mathbf{P}\left[Y_{r+1}(\tau)=i | G_{k+1}(\tau), Y_r(\tau)=k\right]=\mathbf{P}\left[Y_1'(\tau)=i-1+k\right]=\frac{1}{n-k}. \tag{23}$$

On the other hand, if the r^{th} correct guess occurs at step k, then under strategy τ your guess at step $k+1$ will not be one of the r cards you know have been removed from the deck. Thus you guess will be selected from the $n-r$ cards which may remain in the deck. Since $n-k$ cards remain in the deck following step k

$$\mathbf{P}\left[G_{k+1}(\tau)|Y_r(\tau)=k\right]=\frac{n-k}{n-r} \tag{24}$$

for $r \leq k < i$. Equations (23) and (24) are substituted into (22) to obtain the coupled equations

$$\mathbf{P}\left[Y_{r+1}(\tau)=i\right]=\frac{1}{n-r}\mathbf{P}\left[Y_r(\tau)\leq i-1\right] \tag{25}$$

$i>r, 1\leq r<n$.

Induction on r is used to prove that for $r \leq m \leq n, 1 \leq r \leq n$ that

$$\mathbf{P}\left[Y_r(\tau)=m\right]=\frac{\binom{m-1}{r-1}}{\binom{n}{r}r!}. \tag{26}$$

The reader is encouraged to work out a few cases to convince themselves that this ansatz is reasonable.

The base case is $r=1$. One has from (21) that

$$\mathbf{P}\left[Y_1(\tau)=m\right]=\frac{1}{n}=\frac{\binom{m-1}{1-1}}{\binom{n}{1}1!} \tag{27}$$

Hence the base case holds. The induction hypothesis is given in (26). It is used long with (22) to obtain

$$\mathbf{P}\left[Y_{r+1}(\tau)=m\right]=\sum_{k=r}^{m-1}\frac{1}{n-r}\frac{\binom{k-1}{r-1}}{\binom{n}{r}r!} \tag{28}$$

The hockey-stick combinatorical identity gives

$$\sum_{k=r}^{m-1}\binom{k-1}{r-1}=\sum_{k=r-1}^{m-2}\binom{k}{r-1}=\binom{m-1}{r}. \tag{29}$$

On the other hand, it is straightforward to verify algebraically that

$$(n-r)\binom{n}{r}=\binom{n}{r+1}(r+1), \tag{30}$$

Therefore

$$\mathbf{P}\left[Y_{r+1}(\tau)=m\right]=\frac{\binom{m-1}{r}}{\binom{n}{r+1}(r+1)!}=\frac{\binom{m-1}{(r+1)-1}}{\binom{n}{r+1}(r+1)!}. \tag{31}$$

This completes the proof of (26).

The hockey-stick identity is used again to obtain

$$\mathbf{P}\left[Y_r(\tau) \leq n\right] = \sum_{m=r}^{n} \mathbf{P}\left[Y_r(\tau) = m\right] = \frac{1}{r!}. \tag{32}$$

It follows that

$$\mathbf{E}\left[N(\tau)\right] = \sum_{r=1}^{n} \mathbf{P}\left[Y_r(\tau) \leq n\right] = \sum_{r=1}^{n} \frac{1}{r!}. \tag{33}$$

To prove that (33) is the maximum expected number of correct guess in variant 3, return to (24)

$$\mathbf{P}\left[G_{k+1}(\tau)|Y_r(\tau) = k\right] = \frac{n-k}{n-r}$$

The strategy τ is at every step never guess a card that you know has been removed from the deck, and only change guess after a correct guess. Let $\mathbf{s}'$ denote a strategy such that after a correct guess at step k your guess at step $k+1$ is a card that you know has already been removed from the deck at a previous step. Then

$$\mathbf{P}\left[G_{k+1}(\mathbf{s}')|Y_r(\tau) = k\right] = 0$$

Thus under $\mathbf{s}'$ $\mathbf{E}\left[N(\mathbf{s}')\right] < \mathbf{E}\left[N(\tau)\right]$.

On the other hand, let $\mathbf{s}''$ be a strategy in which a guess is changed after an incorrect at step j that comes after the the r correct guess at step $k < j$ and before the next correct guess at step $i > j$. With this strategy, the appropriate probability to consider is

$$\mathbf{P}\left[G_{j+1}(\mathbf{s}'')|Y_r(\mathbf{s}'') = k\right] = \frac{n-j}{n-r} < \frac{n-k}{n-r} = \mathbf{P}\left[G_{k+1}(\tau)|Y_r(\tau) = k\right].$$

Hence $\mathbf{E}\left[N(\mathbf{s}'')\right] < \mathbf{E}\left[N(\tau)\right]$.

It follows that τ maximizes the expected number of correct guesses.

Conclusion

The solution to variant 3 presented here is a brute force calculation. I would like to see a subtle probabalistic solution for this variant. If you have one, please post it in comments.

Update, November 3, 2025

Over the summer, I came up with a probabilitic arguement for Variant 3. I decided to make the variant a problem in my Fall 2025 Featured Problem Series that is currently running on my website. It was last week's featured problem.

The solution to it posted today. You may see it at The Fall 2025 Featured Problem Series. Starting next Monday the solution will be moved to the archive, which is also accessible at the link.

---

1. $[n] := \{1, 2, \ldots, n\}$.↩