LOOKING FOR A GOOD TIME TO BET

LAURENT SERLET

Abstract. Suppose that the cards of a well shuffled deck of cards are turned up one after another. At any time-but once only- you may bet that the next card to be turned up will be black. If it is correct you win 1$, if not you lose 1$. If you never bet you lose 1$. Is there a good strategy to play that game and make it profitable ? Discussing this will be a good opportunity to see an example of a random walk and introduce the notion of a martingale.

1. Let's play a game !

To play this game one needs an ordinary deck of cards, for instance with $N = 52$ cards. In fact we will only look at the colours of the cards, black for half of them and red for the other half. First the cards are properly shuffled. Then the cards are turned up one after another in front of you. At any time-but once only- you may bet that the next card to be turned up will be black. If it is correct you win 1$, if not you lose 1$. If you never bet you lose 1$. Naturally, the question is :

is there a way to make money playing this game ?

Of course the gain is random so "making money" must be understood as choosing a strategy which ensures that the mean gain is strictly positive, that is : probability to win 1$ minus probability to lose 1$ is strictly positive. If it is the case, the most famous theorem of probability theory, called Kolmogorov's law of large numbers, implies that if you play the game a large number of times, the average gain per play will be close to this mean, hence strictly positive.

Let us examine two naive strategies. First betting at a fixed time, for instance at the very beginning, before the first card is revealed–that will at least save time. In that case black or red have the same probability of occuring so the average gain is 0. This is also true for betting at any other fixed time in particular on the last card. Note that in that case the colour of this last card is known for sure but it happens to be black only with probability 1/2 ! However you also have the right to bet at a *random* time depending of the colours of the cards previously revealed. And there are random times such that the next card revealed has more chances to be black than red which implies a strictly positive (conditional) mean gain. Typically it happens if there is a deficit of black cards among the cards

Date: January 24, 2014.

revealed up to that time. More precisely, suppose that, among the k first cards turned up ($1 \leq k \leq N-1$), there are b black ones and r red ones ($k = b + r$). Then, betting at this time, your mean gain will be :

$$\frac{r-b}{N-k}. \tag{1}$$

Indeed, in the specified situation, $(N/2) - b$ black cards and $(N/2) - r$ red ones are left to be drawn so the probability of a black card to be drawn is

$$\frac{(N/2) - b}{(N/2) - b + (N/2) - r} = \frac{1}{2}\left(1 + \frac{r-b}{N-k}\right). \tag{2}$$

Similarly, the probability of a red one is

$$\frac{1}{2}\left(1 - \frac{r-b}{N-k}\right) \tag{3}$$

and (1) follows.

We can see from this formula that the mean gain is proportional to the deficit of black cards and inversely proportional to the number of cards still to be drawn, so one could be tempted to wait for a large deficit of black cards or a low number of remaining cards. But the longer you wait the less probability there is that this situation will occur and remember that in the case of non-betting there is a 1$ penalty.

To go further we need to formalize the problem a little more. In view of formula (1), let us introduce, for $0 \leq k \leq N$, the random variable S_k defined as the number of red cards minus the number of black cards, among the k first cards drawn. In probability theory, $(S_k)_{0 \leq k \leq N}$ is called an unbiased random walk on $\mathbb{Z}$ starting from 0 and conditioned on $S_N = 0$. Many explicit computations can be performed on this object by combinatorial techniques as we will see in the next section.

For the moment let us give a general formula for the mean gain generated by a strategy. Let us say that a strategy is associated with a random time τ if it consists in betting after τ cards have been revealed. Thus τ is a random variable taking its values in $\{0, 1, \ldots, N-1\}$ and we add the possibility that τ takes the value $+\infty$ to mean that the player never bets. We restrict our analysis to the strategies that we call admissible, in the sense that, for every $k \in \{0, 1, \ldots, N-1\}$, the event $\tau = k$ only depends on the values $S_1, \ldots, S_k$. In other words, we assume that the player has no psychic power to anticipate the future. For instance the strategy of waiting for a deficit of q black cards drawn ($q \geq 1$) is denoted by τ_q given by $\tau_q = \min\{k \in \{1, \ldots, N-1\},\ S_k = q\}$ with the convention $\tau_q = +\infty$ if $S_k < q$ for all $k \in \{0, 1, \ldots, N\}$. In probabilistic terms τ_q is the hitting time of q for the random walk $(S_k)_{0 \leq k \leq N}$. Here hitting time means first hitting time but one could possibly consider the strategy associated to the second, third,… hitting time.

With the notation above, it follows from (1) that the mean gain of the admissible strategy associated to τ is :

$$\mathrm{MG}(\tau) = \mathbb{E}\left[\sum_{k=0}^{N-1} \frac{S_k}{N-k}\, \mathbf{1}_{\{\tau=k\}}\right] - \mathbb{P}(\tau = +\infty). \tag{4}$$

In this formula, we denote by $\mathbf{1}_{\{\tau=k\}}$ the variable equal to 1 if the event $\tau = k$ is realized and 0 if not ; $\mathbb{P}$ is our notation for probability and $\mathbb{E}$ is the associated expectation (mean).

The problem we want to solve is to find the admissible strategies τ which maximize $\mathrm{MG}(\tau)$ given by (4) and determine if the maximum is strictly positive. Probability theory has the ideal tool to solve the problem completely as we will explain in Section 3. However this argument could seem rather abstract so we first perform the explicit computation with a natural candidate to be a good strategy. This strategy is simply τ_1 which consists in betting as soon as red cards have been turned up more often than black ones. This will be an opportunity to see the techniques needed to do such a computation and the difficulty to conclude this way. By contrast, it will hopefully convince the reader of the power of the abstract argument developed in Section 3.

2. Is the "natural strategy" profitable ?

We want to compute $\mathrm{MG}(\tau_1)$ using (4). We first remark that on the event $\tau_1 = k$, we have $S_k = 1$ so the formula simplifies to

$$\mathrm{MG}(\tau_1) = \left(\sum_{k=0}^{N-1} \frac{1}{N-k}\, \mathbb{P}(\tau_1 = k)\right) - \mathbb{P}(\tau_1 = +\infty). \tag{5}$$

We are left with computing the probabilities appearing in this formula which amounts to computing the law of τ_1. To state the results we need to introduce the usual binomial coefficients :

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!}$$

when n is a non-negative integer and $k \in \{0, \ldots, n\}$; moreover, when these requirements are not satisfied, the binomial coefficient above is set equal to 0. We claim that

$$\mathbb{P}(\tau_1 = +\infty) = 1 - \frac{\binom{N}{\frac{N}{2}+1}}{\binom{N}{N/2}} \tag{6}$$

and, for every $k \in \{1, \ldots, N-1\}$,

$$\mathbb{P}(\tau_1 = k) = \frac{1}{k} \frac{\binom{k}{\frac{k+1}{2}} \binom{N-k}{\frac{N-k-1}{2}}}{\binom{N}{N/2}}. \tag{7}$$

The proof is given in Section 5. Substituting these values in (8) and putting $N = 2n$, we get :

$$\mathrm{MG}(\tau_1) = \frac{1}{2} \sum_{i=0}^{n-1} \frac{1}{(2n-2i-1)\,(i+1)} \frac{\binom{2i}{i}\,\binom{2n-2i}{n-i}}{\binom{2n}{n}} - \frac{1}{n+1}. \tag{8}$$

Remember that we want to determine if this expression is strictly positive. A priori it is not clear but computing the value for $n = 1, 2, 3$, we always find 0. In fact the value is 0 for every $n \geq 1$. This will be a consequence of the arguments given later in this article, but very similar identities exist in the literature, see for instance [Go] Formula (1.34) attributed to Bruckman.

This result is disappointing because the strategy associated to τ_1 which seemed rather sensible is not better than the strategy of betting on the first card. Both strategies are not profitable. However both are fair is the sense that the mean gain is 0. One could ask about the strategy associated to τ_q for $q > 1$. The computations are rather similar but it is even harder to see the sign of $\mathrm{MG}(\tau_q)$.

Because of the rather negative result of this "experimental" approach, we feel the need of a more abstract probabilistic tool to settle the problem. The next section is going to present the ad hoc concept.

3. Let's use a martingale !

In the gambler vocabulary a martingale is a technique which enables one to win. In probability theory a martingale is simply a sequence of random variables which satisfy a certain property concerning the (conditional) means. See [MS] for the historical aspects and, for instance, [Wi] or [Ne] for an account of the full theory. The paradox is that a martingale of probability theory can often be used to show that a "gambler's martingale" does not exist. In our setting a martingale $(M_k)_{0 \leq k \leq N}$ is defined as a sequence of random variables such that, for every $i \in \{0, \ldots N-1\}$ and for every event B depending only on the values of $S_1, \ldots, S_i$, we have

$$\mathbb{E}\left[M_i \, \mathbf{1}_B\right] = \mathbb{E}\left[M_{i+1} \, \mathbf{1}_B\right]. \tag{9}$$

In our problem, it happens that the sequence

$$(M_k)_{0 \leq k \leq N} = \left(\frac{S_k}{N-k} \right)_{0 \leq k \leq N} \tag{10}$$

is a martingale. Let us justify this fact. Equation (9) amounts to showing that, for all integers $s_1, \ldots, s_i$, the mean of $M_{i+1} = S_{i+1}/(N-i-1)$ conditionally on $S_1 = s_1, \ldots, S_i = s_i$ is $s_i/(N-i)$. But as seen in (2) and (3), conditionally on $S_1 = s_1, \ldots, S_i = s_i$, we have $S_{i+1} = s_i - 1$ with probability $\frac{1}{2}\left(1 + \frac{s_i}{N-i}\right)$ and $S_{i+1} = s_i + 1$ with probability $\frac{1}{2}\left(1 - \frac{s_i}{N-i}\right)$ so the conditional mean of S_{i+1} is

$$(s_i - 1)\,\frac{1}{2}\left(1 + \frac{s_i}{N-i}\right) + (s_i + 1)\,\frac{1}{2}\left(1 - \frac{s_i}{N-i}\right) = \frac{N-(i+1)}{N-i}\, s_i.$$

Simply dividing by $N-(i+1)$, we obtain the martingale property. From this property we deduce the following result.

Theorem 1. *The mean gain associated to an admissible strategy τ is*

$$\mathrm{MG}(\tau) = -2\,\mathbb{P}(S_{N-1}=1;\,\tau=+\infty). \tag{11}$$

Proof. Let $k \in \{1,\dots,N-1\}$. If we iterate (9) for the martingale given in (10), letting i take the values $k, k+1, \dots, N-2$ and $B=\{\tau=k\}$, we obtain that

$$\mathbb{E}\left[\frac{S_k}{N-k}\ \mathbf{1}_{\{\tau=k\}}\right] = \mathbb{E}\left[S_{N-1}\ \mathbf{1}_{\{\tau=k\}}\right].$$

Substituting this expression in (4), we get that the mean gain is

$$\mathrm{MG}(\tau) = \mathbb{E}\left[S_{N-1}\ \left(\sum_{k=0}^{N-1}\mathbf{1}_{\{\tau=k\}}\right)\right] - \mathbb{P}(\tau=+\infty). \tag{12}$$

But

$$\sum_{k=0}^{N-1}\mathbf{1}_{\{\tau=k\}} = 1-\mathbf{1}_{\{\tau=+\infty\}}$$

and it is clear by symmetry that $\mathbb{E}(S_{N-1})=0$. Substituting in (12), we obtain that

$$\mathrm{MG}(\tau) = -\mathbb{E}\left[(1+S_{N-1})\ \mathbf{1}_{\{\tau=+\infty\}}\right] \tag{13}$$

and (11) is just another formulation. The proof is complete.

Let us now discuss the consequences of this theorem. The most important is that the mean gain cannot be strictly positive. There is no way to make money with this game. Conversely any strategy associated with τ such that $\mathbb{P}(\tau=+\infty)=0$ –i.e. such that the player always bets at some time– is fair. It also confirms that the strategy associated with τ_1 is fair because, for this strategy, the event $\tau_1=+\infty$ (no bet) automatically implies $S_{N-1}=-1$. So this theorem shows no strategy can do better than the strategy of betting on the first card drawn. More sophisticated strategies like those associated to τ_q for $q>1$ can in fact be even worse since, as we will see in (14), the mean gain is strictly negative. Conversely it is astonishing that certain strategies which sound stupid are in fact fair and by consequence among the best ones. For instance, suppose that you bet at the first time when black cards have turned up q times more than red ones ($q \geq 1$) or on the last card if this never happens. It does not seem to make sense but in fact this strategy is as clever as the best ones. In fact (11) shows that losing money is caused by non-betting even when you know for sure that the last card drawn will be black, which is effectively a stupid behaviour !

4. Further questions and remarks

Some new questions arise if the problem is changed or extended. In particular what happens if the penalty for non-betting is cancelled ? Of course profitable strategies arise. Starting again from (12), we obtain that the mean gain of an admissible strategy associated to τ is now

$$\begin{aligned}\widetilde{\mathrm{MG}}(\tau) &= \mathbb{E}\left[S_{N-1}\,\mathbf{1}_{\{\tau<+\infty\}}\right] \\ &= \mathbb{P}(S_{N-1}=1,\ \tau<+\infty)-\mathbb{P}(S_{N-1}=-1,\ \tau<+\infty).\end{aligned}$$

The strategy which maximizes this quantity is $\tilde{\tau}$ defined by $\tilde{\tau} = N-1$ if $S_{N-1} = 1$ and $+\infty$ otherwise. This is the "last minute" strategy which consists in waiting for the last draw and in that case $\widetilde{\mathrm{MG}}(\tilde{\tau}) = 1/2$. In that case too the best strategy is not a clever one.

One can also imagine several generalizations of this game with bets on the suit, but note that the modelling will then require to consider a random walk in dimension bigger than 1, which certainly complicates the arguments.

We end this article by giving a few additional results that the reader can treat as application exercises.

(1) Show that

$$\mathrm{MG}(\tau_q) = -2\frac{\binom{N-1}{N/2}-\binom{N-1}{\frac{N}{2}+q-1}}{\binom{N}{N/2}}. \tag{14}$$

and, using the well known monotonicity properties of binomial coefficients, show that, for $q > 1$, the mean gain $\mathrm{MG}(\tau_q)$ is strictly negative.

(2) Combine (14) with (4) to deduce a –not so obvious– combinatorial identity : for N even integer and $q \geq 1$,

$$\sum_{k=q}^{N-q}\frac{1}{k\,(N-k)}\binom{k}{\frac{k+q}{2}}\binom{N-k}{\frac{N-k-q}{2}} = \frac{2}{q\,N}\binom{N}{\frac{N}{2}+q}$$

which is a generalization of [Go] Formula (1.34).

(3) With the notation $N = 2n$, show that

$$\widetilde{\mathrm{MG}}(\tau_q) = \frac{q}{n}\frac{(n!)^2}{(n+q)!\ \ (n-q)!}.$$

Check that this quantity reaches a maximum for q_0 being the least integer q such that $2\,q\,(q+1) > n$. In particular when n is large, $q_0 \approx \sqrt{n/2}$ and $\widetilde{\mathrm{MG}}(\tau_{q_0})$ is, up to multiplicative constants, of order $n^{-1/2}$ which is much lower than the last minute strategy.

5. Appendix : computation of a law

This section is devoted to the computation of the law of τ_1 that we announced in (6) and (7). Since it requires no more effort, we will compute the law of τ_q for $q \geq 1$. It uses well known combinatorial techniques on the set of trajectories of random walks, the main argument being the so-called "reflection principle". Chapter III of Feller's book ([Fe]) is a nice and classical reference on the techniques we use in the present section. Note that τ_q takes in fact its values in $\{q, \dots, N-q\} \cup \{+\infty\}$ and we will moreover see that τ_q may only take integer values having the same parity as q.

Let us denote by $\mathcal{S}(N, a, b)$ the set of all finite sequences of integers $(s_k)_{0 \le k \le N}$ of length $N+1$ with increments $+1$ or -1, starting from $s_0 = a$ and arriving at $s_N = b$. Such a sequence is easily represented by a graph in the plane. A sequence belonging to $\mathcal{S}(N, a, b)$ is entirely determined by specifying which increments are $+1$ and the number of those increments equal to $+1$ is exactly $(N+b-a)/2$ in order to meet the requirements on s_0 and s_N. Hence the cardinal (number of elements) of $\mathcal{S}(N, a, b)$ is

$$\#[\mathcal{S}(N, a, b)] = \binom{N}{\frac{N+b-a}{2}}. \tag{15}$$

The random sequence $(S_k)_{0 \le k \le N}$ takes its values in $\mathcal{S}(N, 0, 0)$ and moreover, the distribution of this random sequence is the uniform probability on $\mathcal{S}(N, 0, 0)$. As a consequence computing the desired probabilities will consist essentially in determining the cardinal of the specified events. For instance let us start with $\mathbb{P}(\tau_q = +\infty)$ and for convenience, we prefer to compute the probability $\mathbb{P}(\tau_q < +\infty)$ of the complementary event. It is the cardinal of the subset of $\mathcal{S}(N, 0, 0)$ formed by the sequences that reach q, divided by the cardinal of $\mathcal{S}(N, 0, 0)$. But there is a one-to-one correspondence between sequences in $\mathcal{S}(N, 0, 0)$ that reach q and sequences in $\mathcal{S}(N, 0, 2q)$. The correspondence consists in applying a reflection with respect to level q from the hitting time of q up to time N. Therefore this argument is called the "reflection principle". By (15), it follows that

$$\mathbb{P}(\tau_q < +\infty) = \frac{\#[\mathcal{S}(N, 0, 2q)]}{\#[\mathcal{S}(N, 0, 0)]} = \frac{\binom{N}{\frac{N}{2}-q}}{\binom{N}{N/2}} \tag{16}$$

hence (6). Now we pass to $\mathbb{P}(\tau_q = k)$. Decomposing into the part before time k and the part after time k, we get

$$\mathbb{P}(\tau_q = k) = \frac{\#[\mathcal{S}(k, 0, q; < q)] \times \#[\mathcal{S}(N-k, q, 0)]}{\#[\mathcal{S}(N, 0, 0)]} \tag{17}$$

where $\mathcal{S}(k, 0, q; < q)$ is the subset of $\mathcal{S}(k, 0, q)$ consisting of sequences that reach level q for the first time at time k. This set has same cardinal as the set of sequences in $\mathcal{S}(k-1, 0, q-1)$ which do not reach level q. Another application of the reflection principle shows that the latter set has same cardinal as

$$\#[\mathcal{S}(k-1, 0, q-1)] - \#[\mathcal{S}(k-1, 0, q+1)] = \frac{q}{k}\binom{k}{\frac{k+q}{2}}.$$

Substituting in (17) and expressing the other terms using (15), we get the final expression of the law of τ_q :

$$\mathbb{P}(\tau_q = k) = \frac{q}{k} \frac{\binom{k}{\frac{k+q}{2}}\binom{N-k}{\frac{N-k-q}{2}}}{\binom{N}{N/2}} \tag{18}$$

and (7) is a particular case.

REFERENCES

Fe: Feller W. *An introduction to probability theory and its applications*, volume 1. Third edition. Wiley.

Go: Gould H.W. Tables of combinatorial identities, unpublished manuscripts collected by J. Quaintance.
Available on the web page : www.math.edu/~gould/

MS: Mazliak L., Shafer G. (editors). *The splendors and miseries of martingales.* Electronic Journal for History of Probability and Statistics 5 (1), june 2009.

Ne: Neveu J. *Discrete-parameter martingales.* Vol. 10 of North-Holland math. library. Elsevier (1975)

Wi: Williams D. *Probability with martingales.* Cambridge University Press. (1991)

Laurent Serlet. After studying at the École Normale Supérieure de Cachan, I obtained my Ph.D. at the University Paris 6 in 1993. Since 2005, I am a professor at the University Blaise Pascal in Clermont-Ferrand. My research interests have been first measure-valued and path-valued processes; now I am mostly interested in perturbed random walks and reinforced processes.

Clermont Université, Université Blaise Pascal, Laboratoire de Mathématiques (CNRS umr 6620), Complexe scientifique des Cézeaux, BP 80026, 63171 AUBIÈRE, FRANCE

E-mail address: `Laurent.Serlet@math.univ-bpclermont.fr`