A mathematical model for "Who wants to be a millionaire?"

Robert C. Dalang[1] and Violetta Bernyk[1]

Abstract

We propose a mathematical model for the TV game show "Who wants to be a millionaire?" Using stochastic optimization methods, we obtain the optimal strategy maximizing the player's expected payoff. The model provides answers to questions such as "What can the player expect to win?" and "What are the chances of winning a million dollars?" This optimal stategy is presented in simple form in an appendix.

1 Introduction

The TV game show "Who wants to be a millionaire?" (rules described below) is shown in many countries. The objective of this paper is to provide a mathematical model for this game and to compute the player's optimal strategy. Modeling this game is interesting for several reasons: the rules do not uniquely determine the mathematical model; the game involves chance and decisions by the player under uncertainty, but the randomness is not produced by a physical mechanism and important facts about how the randomness is produced are not available; it is necessary to model the influence on the player's strategy of external sources of information; much statistical data that one would like to have is not available; and selecting an optimization criterion for the player is part of the modeling problem.

In the face of these difficulties, the objective of a model is to provide a rational strategy that helps the player decide which risks are worth taking. Since the game has a combinatorial flavor, no advanced probability theory is needed and this problem and the resulting model can be discussed in a first or second course in probability theory or stochastic processes. The strategy that we propose is presented in a self-contained appendix (see Section 7) and can easily be used even by non-mathematicians.

Rules of the game. The player is successively confronted with up to fifteen questions. Question number n has a face value f_n, shown in Table 1 for the USA version of the game. For each question, four answers a, b, c and d are proposed to the player.

[1]Institut de Mathématiques, Ecole Polytechnique Fédérale, 1015 Lausanne, Switzerland. Emails: robert.dalang@epfl.ch violetta.bernyk@epfl.ch The first author's research is partly supported by the Swiss National Foundation for Scientific Research.

MSC 2002 Subject classifications. Primary 60G40; Secondary 60J20, 91A35, 60-01.

Keywords: Stochastic optimization, probabilistic modelling.

n	f_n	w_n
1	100	0
2	200	0
3	300	0
4	500	0
5	1000	0

n	f_n	w_n
6	2000	1000
7	4000	1000
8	8000	1000
9	16000	1000
10	32000	1000

n	f_n	w_n
11	64000	32000
12	125000	32000
13	250000	32000
14	500000	32000
15	1000000	32000

Table 1: The face values f_n and the payoffs w_n for an incorrect answer (USA version).

After seeing question n and the four possible answers, the player may choose to answer, to use one of three *"lifelines"* (described below) or to quit without replying and to receive the payoff f_{n-1} (f_0 is set to 0). There is no time limit for answering.

If the player answers correctly, he moves to question $n+1$. If the player answers incorrectly, the game is over and the player receives the amount w_n, also shown in Table 1. The values of w_n increase strictly at $n = 5$ and $n = 10$, which are called *"milestones."* If the player is not sure of the correct answer, he may use one of the three lifelines. He may use more than one lifeline for the same question, but each lifeline can be used only once during the entire game.

The three lifelines. One lifeline is entitled *"50:50."* When the player asks to use this lifeline, a computer eliminates two incorrect answers, leaving two answers to choose from. A second lifeline is called *"Phone-a-friend"*: the host telephones a friend selected by the player and the player has 30 seconds to communicate the question, the four answers, and to obtain the reply from his friend. He is then free to use the information supplied by his friend as he sees fit. The third lifeline is *"Ask-the-audience"*: the host asks all members of the audience to enter their answer into the computer. The computer tabulates the audience's answers in a histogram, which is shown to the player, and the player is free to use this information as he sees fit.

The need for a mathematical model. Stochastic optimization techniques can be applied to this game, and there is a well-developed theory for this, expounded in books such as [1], [2], [4] and [6]. But none of these are directly applicable, since a mathematical model for the game is needed, whereas the theory begins by assuming that such a model is given! The main objective of this paper is to create a suitable model, which turns out to be a nice example of a stochastic control problem for processes indexed by a (partially) ordered set (see [1]). Though we do choose various numerical values in order to propose an actual strategy, the choices can be refined.

2 The generic model

We let $n = 1, \ldots, 15$ denote the number of the question that the player is contemplating, and $n = 16$ after the player has correctly answered all fifteen questions.

The player's states of knowledge. The central parameters in the game are the likelihoods that the player will be able to answer each question n. Let $\pi^n = (\pi_a^n, \pi_b^n, \pi_c^n, \pi_d^n)$ be the player's prior probability vector for question n, where the components are non-negative and sum to 1, and π_i^n, $i = a, b, c, d$, is the probability that the player assigns to the event "the correct answer is i."

Since a player will not be able to state his prior probabilities very precisely, we shall consider a finite set $\mathcal{K}$ of possible states of knowledge for any given question. Each state of knowledge corresponds to a particular prior probability vector.

When the player gives an incorrect answer, he has lost, and since his payoff depends on how many of the milestones he has passed, we need three "lost" states, which we denote L_1, L_2 and L_3. We shall denote by K_n the set of possible states of the player after answering $n-1$ questions. Taking into account the "lost states" and a state W (= win) for the (fortunate) player who has correctly answered all fifteen questions, these are $K_1 = \mathcal{K}$,

$$\begin{array}{rcll} K_n & = & \mathcal{K} \cup \{L_1\}, & n = 2, \ldots, 6, \\ K_n & = & \mathcal{K} \cup \{L_1, L_2\}, & n = 7, \ldots, 11, \\ K_n & = & \mathcal{K} \cup \{L_1, L_2, L_3\}, & n = 12, \ldots, 15, \end{array}$$

and $K_{16} = \{W, L_1, L_2, L_3\}$. We note that if the player gives an incorrect answer to question n, then he moves to state $(n+1, L_m)$, with $(n, m) \in \Lambda$, where

$$\Lambda = (\{1, \ldots, 5\} \times \{1\}) \cup (\{6, \ldots, 10\} \times \{2\}) \cup (\{11, \ldots, 15\} \times \{3\}).$$

Stages in the game. In order to describe the player's progression, we must keep track of the number n and of the lifelines that remain available. Let Lifeline 1 be the 50:50 lifeline, Lifeline 2 be "Phone-a-friend," and Lifeline 3 be "Ask-the-audience."

Since each lifeline can be used only once, we set $S = \{0, 1\}^3$, and for $s = (s_1, s_2, s_3) \in S$, s_m will denote the number of times lifeline m has been used, $m = 1, 2, 3$. The three lifelines are symbolized by ℓ_1, ℓ_2 and ℓ_3, where

$$\ell_1 = (1, 0, 0), \qquad \ell_2 = (0, 1, 0), \qquad \ell_3 = (0, 0, 1).$$

Each $s \in S$ represents the lifelines still available: for instance, $s = (1, 1, 0)$ if only Lifeline 3 is still available. For $s = (s_1, s_2, s_3)$ and $t = (t_1, t_2, t_3)$ in S, define

$$s \vee t = (\max(s_1, t_1), \max(s_2, t_2), \max(s_3, t_3)), \qquad s - t = (s_1 - t_1, s_2 - t_2, s_3 - t_3),$$

and set $D(s) = \{s \vee \ell_1, s \vee \ell_2, s \vee \ell_3\} \setminus \{s\}$. For $\sigma \in D(s)$, $\sigma - s \in \{\ell_1, \ell_2, \ell_3\}$ represents the lifeline that was used to pass from s to σ. For $s \neq (1, 1, 1)$, the set $D(s)$ represents the possible lifeline situation, after using one of the remaining available lifelines.

At any stage in the game, the progression of the player in the game is fully described by the couple $(n, s) \in \mathbb{P} \stackrel{\text{def}}{=} \{1, \ldots, 16\} \times S$. These are the *stages* in the game.

Markovian nature of the game. At the stage (n, s), the status of the player is a random variable $X_{n,s}$ with values in K_n. The event $\{X_{n,s} = k\}$ occurs if the player's state of knowledge concerning question n is k, and $\{X_{n,s} = L_m\}$ occurs if the player previously gave an incorrect answer. Before question n is displayed, we do not know what the player's status will be at that or future stages of the game.

At a given stage (n, s) in the game, once we know the status $X_{n,s}$, the previous statuses, such as the degree of knowledge in previous questions, are no longer relevant. The player can either answer and move to a new stage $(n + 1, s)$ with a new status $X_{n+1,s}$, or he can use a lifeline and move to a new stage (n, σ), with $\sigma \in D(s)$, and a new status $X_{n,\sigma}$. The lifeline he just used was $\sigma - s$. In order to describe the evolution of the player's status, we need the transition probabilites

$$p_n(j, k) = P\{X_{n+1,s} = k \mid X_{n,s} = j\}, \qquad j \in K_n,\ k \in K_{n+1},$$

and, for $s \in S$ such that $s \vee \ell_m \in D(s)$,

$$p_{n,\ell_m}(j, k) = P\{X_{n,s\vee\ell_m} = k \mid X_{n,s} = j\}, \qquad j, k \in K_n.$$

We assume for the moment that we know these transition probabilities. In Sections 3 and 4, we shall explain how we determine them.

Optimization. For $(n, s) \in \mathbb{P}$ and $k \in K_n$, the payoff to the player if he quits the game at stage (n, s) does not depend on s, so we denote it $f(n, k)$. According to the rules of the game, we set $f(n, k) = f_{n-1}$, $k = 0, \ldots, 4$. For $n = 2, \ldots, 16$, $f(n, L_1) = w_1$, for $n = 7, \ldots, 16$, $f(n, L_2) = w_6$, and for $n = 12, \ldots, 16$, $f(n, L_3) = w_{11}$, where w_1, w_6 and w_{11} are as in Table 1. Finally, $f(16, W) = f_{15} = 1000000$.

We shall assume that the player seeks to maximize the expected utility of the payoff, for a given *utility function* $u : \mathbb{R}_+ \to \mathbb{R}$ (for a reasonable choice of this function, see Section 5), and we let $g(n, k) = u(f(n, k))$.

For $(n, s) \in \mathbb{P}$ and $k \in K_n$, let $u^*(n, s, k)$ be the expected utility of a player currently at stage (n, s) and in status k, who proceeds *optimally* from this stage on. Bellman's equation of dynamic programming (see e.g. [1, Sections 3.6 and 9.2]) for $u^*(\cdot, \cdot, \cdot)$ is

$$u^*(n, s, j) = \max\left(g(n, j), \mathcal{C}(n, s, j), \max_{\sigma\in D(s)} \mathcal{L}(n, s, \sigma, j)\right), \tag{1}$$

where

$$\mathcal{C}(n, s, j) = \sum_{k\in K_{n+1}} p_n(j, k)u^*(n + 1, s, k), \qquad \mathcal{L}(n, s, \sigma, j) = \sum_{k=0}^{4} p_{n,\sigma-s}(j, k)u^*(n, \sigma, k).$$

The values of u^* are now computed by backward induction, since $u^*(16, s, W) = g(16, W)$ $(= u(10^6))$, for all $s \in S$, and for $m = 1, 2, 3$, if $L_m \in K_n$, then $u^*(n, s, L_m) = g(n, L_m)$, for $n = 2, \dots, 16$, $s \in S$.

With these values of $u^*(\cdot, \cdot, \cdot)$, we use (1) successively for $n = 15, 14, \dots, 2, 1$, to compute $u^*(n, s, j)$, first for $s = (1, 1, 1)$ $(D(s) = \emptyset)$, then for $s = (0, 1, 1)$, $s = (1, 0, 1)$ and $s = (1, 1, 0)$ $(D(s) = \{(1, 1, 1)\})$, then for $s = (0, 0, 1)$, $s = (0, 1, 0)$ and $s = (1, 0, 0)$ (two lifelines available), and finally for $s = (0, 0, 0)$ (all lifelines available).

Description of the optimal strategy. Once the $u^*(n, s, k)$ have been computed, the optimal strategy for the player is as follows: when at stage (n, s) with knowledge status j, determine which of the three quantities on the right-hand side of (1) is equal to $u^*(n, s, j)$. If this is $g(n, j)$, then quit and receive the payoff $f(n, j)$. If this is $\mathcal{C}(n, s, j)$, then answer the question. If this is $\max_{\sigma \in D(s)} \mathcal{L}(n, s, \sigma, j)$, then pick m such that $\sigma = s \vee \ell_m$ achieves this maximum, and use lifeline m. In the last two cases, the player moves respectively to a new stage $(n + 1, s)$ or (n, σ), and repeats the procedure.

3 Adjusting the model to the game

We now describe how to determine the transition probabilities $p_n(j, k)$ from more basic quantities. Let $q_n(k)$ be the likelihood that the player will have a particular state $k \in \mathcal{K}$ of knowledge for question n, and let r_k be the probability that the player's answer is correct if his state of knowledge is k. Then $p_n(j, k)$ can be easily expressed from the $q_n(k)$ and r_k, if we assume that giving a correct answer to question n is independent of the player's state of knowledge for question $n + 1$: for $j \in \mathcal{K}$,

$$p_n(j, k) = r_j \, q_{n+1}(k), \qquad k \in \mathcal{K}, \; n = 1, \dots, 14,$$

$$p_n(j, L_m) = \begin{cases} 1 - r_j & \text{if } (n, m) \in \Lambda, \\ 0 & \text{otherwise,} \end{cases} \qquad p_{15}(j, W) = r_j. \tag{2}$$

Note that if the player gives an incorrect answer when at stage (n, s), then he moves to state $L_m \in K_{n+1}$ at stage $(n + 1, s)$, where $(n, m) \in \Lambda$. For completeness, we set $p_n(L_m, L_m) = 1$ if $L_m \in K_n$, $n = 2, \dots, 15$.

Specifying the states of knowledge and numerical choices of parameters. In order to propose a specific strategy, we need to specify the set $\mathcal{K}$ and to select values for the parameters $q_n(k)$ and r_j. This requires prior (statistical) information about the game. There are, however, two distinct situations.

Imagine that a player learns with only a few days' advance notice that he will be on the show (this is generally the case). At that point, he might like to create a rational strategy, but he will not have time to gather much additional data by watching further shows, and he will have to use only the information that he already has. He can go out and buy the game box version of the show, and use the 100 or so questions for each face value to create some statistical data to include in the model.

We shall consider this situation. We decided to limit ourselves to a few states in $\mathcal{K}$, and after examining the questions from the game box version, found that we were almost always in one of the following *five basic states of knowledge:*

State 0: the player definitely knows the answer (zero uncertainty).

State 1: the player is confident, but not certain, that he knows the correct answer.

State 2: the player hesitates between 2 answers (the other two are unlikely).

State 3: Two answers were just eliminated by the computer by using the 50:50 lifeline, but the player still hesitates between the two remaining answers.

State 4: the player has no idea of the answer (hesitates between all 4 answers).

In terms of prior probability vectors, we found that it was consistent with our observations from the game box version to associate to state k a permutation of the prior probability vectors shown in Table 2. Counting the distinct permutations of these vectors, $\mathcal{K}$ should contain 21 elements, but since the permutations of a given vector all play the same role, we simply set $\mathcal{K} = \{0, 1, 2, 3, 4\}$.

k	Probability vector
0	(1, 0, 0, 0)
1	(4/5, 1/15, 1/15, 1/15)
2	(2/5, 2/5, 1/10, 1/10)
3	(1/2, 1/2, 0, 0)
4	(1/4, 1/4, 1/4, 1/4)

Table 2: For each state k of knowledge, the player's prior probabilities are a permutation of those that are shown.

Our selection of the $q_n(k)$ must reflect the fact that the questions become more and more difficult. Since there are two milestones, we divide the fifteen questions into three groups and set, for $k \in \{0, \dots, 4\}$, $q_n(k) = \rho_m(k)$, where $(n, m) \in \Lambda$.

The numerical values for the $\rho_m(k)$ resulting from our study of the game box version are shown in Table 3. Since state $k = 3$ can only be reached by using the 50:50 lifeline, we have set $\rho_m(3) = 0$, $m = 1, 2, 3$.

According to our choice of states, we should set $r_0 = 1$, $r_1 = 4/5$, $r_2 = 2/5$, $r_3 = 1/2$, $r_4 = 1/4$. The numerical values of the matrix $(p_n(j, k),\ j = 0, \dots, 4,\ k = 0, \dots, 4)$ are now easily computed using the formulas in (2).

Another possible statistical situation. A distinct statistical situation would be that of a player who prepares for the game for many months. This player can gather lots of data, which should lead to more refined estimates of the various parameters. However, since the randomness in the game is not produced by a physical mechanism,

	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$
$\rho_1(k)$	0.5	0.4	0.05	0	0.05
$\rho_2(k)$	0.2	0.3	0.2	0	0.3
$\rho_3(k)$	0.2	0.2	0.1	0	0.5

Table 3: Our choice for the probabilities that the player's state of knowledge for a question is k, for the three groups of questions.

there is no guarantee that the data will be consistent over time: past data may not be so useful for future games. While much data about the game may be recorded somewhere, the (apparently unavailable) data we would need would be a record of which state $k \in \mathcal{K}$ the player was in for each question, and how well he did when in each state. We decided not to seek such data, with the idea that any particular player can tailor the parameters of our model to his particular situation and abilities.

4 Modeling the lifelines

We now turn to the transition probabilites $p_{n,\ell_m}(j,k)$. Each lifeline has its own specific characteristics. We begin with the 50:50 lifeline.

4.1 The 50:50 lifeline

The game is designed so that the probability that the computer will eliminate any one of the three possible pairs of incorrect answers is 1/3. When in state k, the player has assigned prior probabilites to each event C_a = "the correct answer is a," so the total probability formula gives us the probability of events $E_{b,c}$ = "the computer eliminates answers b and c." However, our model does not yet determine the new state to which the player moves after two particular incorrect answers have been eliminated. For this, we use Bayes' formula to determine the posterior probability that each remaining answer is correct, and then we select the player's possible new state(s) in such a way that this posterior probability becomes the player's probability of giving a correct answer.

For instance, if the player is in state 2, assume that the player has assigned probabilities $\frac{2}{5}$, $\frac{2}{5}$, $\frac{1}{10}$, $\frac{1}{10}$ to the respective events "a, b, c, d, is the correct answer."

There are three distinct possibilities: $E_{a,b}$ occurs, $E_{c,d}$ occurs, or one of $\{a, b\}$ and one of $\{c, d\}$ are eliminated. These events have respective probabilities $\frac{1}{15}$, $\frac{4}{15}$, and $\frac{2}{3}$; indeed, the first event has probability

$$\begin{aligned} P(E_{a,b}) &= P(E_{a,b} \mid C_c)\, P(C_c) + P(E_{a,b} \mid C_d)\, P(C_d) \\ &= \frac{1}{3} \cdot \pi_c^n + \frac{1}{3} \cdot \pi_d^n = \frac{1}{3} \cdot \frac{1}{10} + \frac{1}{3} \cdot \frac{1}{10} = \frac{1}{15}, \end{aligned}$$

and the two others are computed in a similar way. Using Bayes' formula, we can compute the probability that either of the remaining answers is correct. For instance,

$$P(C_a \mid E_{b,d}) = \frac{P(E_{b,d} \mid C_a) \cdot P(C_a)}{P(E_{b,d})} = \frac{\frac{1}{3} \cdot \frac{2}{5}}{\frac{1}{6}} = \frac{4}{5},$$

and so the player's state is best represented by state 1 in this case. If $E_{a,b}$ or $E_{c,d}$ occurs, then the player's state is best represented by state 3. Therefore, $p_{n,\ell_1}(2,1) = 4P(E_{b,d}) = 2/3$, and $p_{n,\ell_1}(2,3) = P(E_{a,b} \cup E_{c,d}) = 1/3$.

If the player is in state 1, assume that the player has assigned probability $\frac{4}{5}$ to answer a and $\frac{1}{15}$ to each of answers b, c and d. There are two distinct possibilities: either a is eliminated, or it is not. The first has probability $\frac{2}{15}$ and the second $\frac{13}{15}$.

If a is eliminated, we shall assume that the player's new state of knowledge is 3. If a is not eliminated, then we note that

$$P(C_a \mid E_{b,c}) = \frac{P(E_{b,c} \mid C_a) \cdot P(C_a)}{P(E_{b,c})} = \frac{\frac{1}{3} \cdot \frac{4}{5}}{\frac{13}{45}} = \frac{12}{13} \simeq 0.92,$$

which is substantially higher than the probability $\frac{4}{5}$ assigned to state 1. Being somewhat conservative, we consider that the player's new state of knowledge can be 0 or 1, with respective probabilities $\frac{6}{15}$ and $\frac{7}{15}$.

The transition probabilites associated with the 50:50 lifeline are summarized in Table 4. Notice that they do not depend on n.

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$p_{n,\ell_1}(0,k)$	1	0	0	0	0
$p_{n,\ell_1}(1,k)$	6/15	7/15	0	2/15	0
$p_{n,\ell_1}(2,k)$	0	2/3	0	1/3	0
$p_{n,\ell_1}(3,k)$	0	0	0	1	0
$p_{n,\ell_1}(4,k)$	0	0	0	1	0

Table 4: Transition probabilites when the 50:50 lifeline is used.

4.2 The lifeline "Phone-a-friend"

For this lifeline, there are two clearly distinct issues: the friend's state of knowledge of the answer, and how his state of knowledge influences our own state of knowledge.

We address these two issues in turn. First of all, we assume that the possible states of knowledge for the friend are the same as for the player, and we let $F_n(k)$ be the probability that the friend's state of knowledge for question n is k.

Consider now the issue of how the friend's state of knowledge affects our own state of knowledge. Let $I(j, i, k)$ denote the probability, given the player's state j and the friend's state i, that after using the lifeline "Phone-a-Friend", the player's new state is k. The $p_{n,\ell_2}(j,k)$, $j,k = 0,\dots,4$, are now easily computed:

$$p_{n,\ell_2}(j,k) = \sum_{i=0}^{4} F_n(i)\, I(j,i,k). \tag{3}$$

We now examine how to determine the values of the $I(i,j,k)$. Clearly, the player's new state of knowlege could just as well be the friend's new state of knowledge, so we should have $I(j,i,k) = I(i,j,k)$, and therefore we only consider the case $j \geq i$.

If the player's state is $j = 4$, and the friend's state is i, then the player's new state k of knowledge is certainly $k = i$, so $I(4,i,i) = 1$, $i = 0,\dots,4$.

If the player's state of knowledge is $j = 3$, meaning he has just used the 50:50 lifeline, and the friend's state is also $i = 3$, then the player remains in state 3, so $I(3,3,3) = 1$ and $I(3,3,k) = 0$ for $k \neq 3$. If the friend's state is $i = 2,1$ or 0, then we assume that the interaction between friend and computer is the same as between player and computer when the 50:50 lifeline was used, so we set

$$I(3,i,k) = p_{n,\ell_1}(i,k), \qquad i = 2,1,0, \quad k = 0,\dots,4.$$

Dependence between player's and the friend's selections. Bringing together two knowledgeable people should enhance each one's state of knowledge. One can check that this property does *not* hold under the assumption that given their states of knowledge, the answers they select are independent. Since player and friend derive their knowledge from the same sources (school, books, etc.), independence of selections is in fact not a natural assumption. We shall assume that the friend's and player's states of knowledge *are* independent, but given their states of knowledge, the answers they think correct are *not* independent.

To formalize this, when player and friend are in states 1 or 2, we shall assume that the friend's preferred answer(s) are selected with the same probabilities as drawing (one or two) tickets labelled a, b, c, d, without replacement from an urn, using the prior probabilities $\pi_a^n,\dots,\pi_d^n$ assigned by the player.

If player and friend are both in state 1, and if they have the same preferred answer, then the player's confidence in the correct answer increases, and decreases otherwise. According to the above rule, the former occurs with probability r_1: in this case, we assume that the player moves to state 0, and otherwise, that he moves to state 2. Therefore, $I(1,1,0) = r_1$ and $I(1,1,2) = 1 - r_1$.

If the player's state is $j = 2$, and if the friend is in state 1, then either the friend's selection is contained in the player's two selections (with probability 4/5), or not. In the first case, we assume that the player will go with his friend's suggestion and move to state 1. In the second case, we assume that this brings some confusion to the player and so his new state is 4. Therefore, $I(2,1,1) = \frac{4}{5}$ and $I(2,1,4) = \frac{1}{5}$.

If both player and friend are in state 2, and their preferred pairs are identical, then we assume that the player's state does not change. If their preferred pairs

	$i=0$	$i=1$	$i=2$	$i=3$	$i=4$
$j=0$	0 (1)	0 (1)	0 (1)		0 (1)
$j=1$	0 (1)	0 (4/5) ; 2 (1/5)	1 (4/5); 4 (1/5)		1 (1)
$j=2$	0 (1)	1 (4/5) ; 4 (1/5)	1 (20/45); 2 (24/45); 4 (1/45)		2 (1)
$j=3$	0 (1)	0 (6/15); 1 (7/15); 3 (2/15)	1 (2/3); 3 (1/3)	3 (1)	3 (1)
$j=4$	0 (1)	1 (1)	2 (1)		4 (1)

Table 5: The box at row j, column i, contains the values of k for which $I(j,i,k) > 0$, and next to each k, the value of $I(j,i,k)$ in parentheses.

are disjoint, then we assume that the player moves to state 4, and if they have one answer in common, then we assume that the player will give his preference to the one answer both he and his friend consider likely, and so his new state of knowledge is 1. Computing the respective probabilities of these events leads to

$$I(2,2,1) = 20/45, \qquad I(2,2,2) = 24/45, \qquad I(2,2,4) = 1/45.$$

Finally, for all the triples (j,i,k) that have not been explicitely discussed above, we set $I(j,i,k) = 0$.

Numerical values. The friend has been chosen because he is knowledgeable, and so we could consider that he is more likely to know the correct answer than the player. On the other hand, only 30 seconds are allowed for communication. Rather often, this time limit is significant. In the end, we have chosen simply to set $F_n(k) = q_n(k)$, that is, friend and player have the same probabilites of being in each state.

The numerical values of $I(j,i,k)$, computed according to the formulas described above, are summarized in Table 5. The numerical values of $(p_{n,\ell_2}(j,k),\ j = 0,\dots,4,\ k = 0,\dots,4)$ are now easily computed using (3).

4.3 The lifeline "Ask-the-audience"

Typical audience responses are shown in Figure 1. For instance, in Case 1, 70% of the audience says that a is the correct answer.

How does the player use the histogram of audience responses? We assume that the audience aims to help the player (though this seems *not* always to be the case in at least one country, Russia). We also assume that the player determines from the histogram that the audience's state of knowledge is one of the states $0,\dots,4$. This should generally be easy to do, though some borderline cases may arise.

With this assumption, we can treat this lifeline in the same way as the "Phone-a-Friend" lifeline. Let $A_n(k)$ be the probability that the audience's state of knowledge for question n is k. The transition probabilities $p_{n,\ell_3}(j,k)$ are now computed as in (3), with $F_n(i)$ replaced by $A_n(i)$.

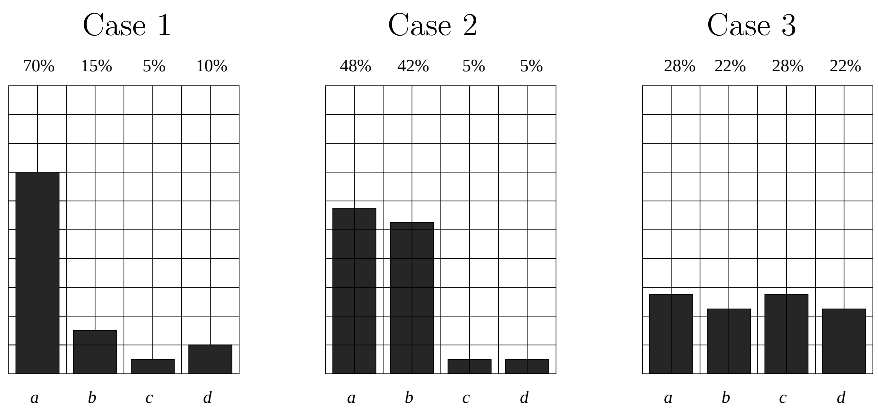

Figure 1: Three typical histograms of audience responses.

Numerical values. Since members of the audience can talk to their neighbors, and since they are not under the same stress as the player, they should be more likely to know the correct answer, at least during the early stages of the game. It seems reasonable to set $A_n(k) = \alpha_m(k)$ if $(n,m) \in \Lambda$, with our choice for the $\alpha_m(k)$ shown in Table 6.

	$k=0$	$k=1$	$k=2$	$k=3$	$k=4$
$\alpha_1(k)$	0.6	0.3	0.05	0	0.05
$\alpha_2(k)$	0.4	0.3	0.2	0	0.1
$\alpha_3(k)$	0.2	0.2	0.1	0	0.5

Table 6: Our choice for the probabilities that the audience's state of knowledge for a question is k, for the three groups of questions.

Using the numerical values of $I(j,i,k)$ shown in Table 5, we easily obtain the numerical values of the matrix $(p_{n,\ell_3}(j,k),\ j=0,\dots,4,\ k=0,\dots,4)$.

5 Modeling the player's risk tolerance

Recall, as in [5], that a "utility function" is any function $u : I\!R_+ \to I\!R$ that is continuous, non-decreasing and *concave*. A standard class of utility functions are the power functions $u(x) = x^p$, where $p < 1$ ($1-p$ is known as the Arrow-Pratt risk-aversion index, as in [5, p.20]). A value of $1-p$ near 0 is to be used for a player who has a high risk tolerance, while a large value of $1-p$ should be used for a cautious player.

What is a reasonable choice for p? The most extreme risks that confront the player occur at questions 10 and 15: if the player quits, he receives the payoff $x = f_9$ (respectively $x = f_{14}$). If he answers incorrectly, his payoff is w_6 (respectively w_{11}), which is approximately $x/15$. If he answers correctly, then his payoff will be at least w_{11} (respectively f_{15}), which equals $2x$. In short, if he is uncertain about the answer, he can quit the game with the payoff x, or answer and receive either $x/15$ or $2x$.

Consider the following alternative: (A) you are given x =\$500000; or (B) a fair coin is tossed and you receive \$32000 if it falls on heads, and \$1000000 if it falls on tails. Which of these alternatives would you prefer? Even though the expected reward is slightly higher in the second case, most people prefer (A). If the fixed amount is lowered to \$400000, most people still prefer (A). When this amount is set to \$300000, then most people feel that (B) is more attractive. We can set the breakeven point, at which both alternatives are equally attractive, at $\$350000 = \frac{7}{10}x$.

Based on these considerations, we seek a utility function $u(x) = x^p$ such that

$$(7x/10)^p = \tfrac{1}{2}\,(x/15)^p + \tfrac{1}{2}(2x)^p.$$

Dividing both sides of this equation by x^p, one easily checks that this equality is satisfied to the third decimal place for $p = \frac{1}{2}$, so that the utility function $u(x) = \sqrt{x}$ does a satisfactory job of capturing a typical level of risk tolerance. We shall use this particular utility function to compute our optimal strategy.

6 The optimal stragegy

We now have in hand all the numerical quantities needed to compute the optimal strategy according to the procedure described in Section 2. Since there are 15 questions, 5 states of knowledge and 8 configurations for the availability of lifelines, there are 600 values of $u^*(n, s, j)$ to compute, which could, at least in principle, be computed by hand (though we used Mathematica and Excel, which gave identical results).

The optimal strategy that results is shown in the Appendix (Section 7). Some comments on this strategy are in order.

"Essential" uniqueness. The optimal strategy is not unique. For instance, if we are at stage $(15, s)$, with $s \neq (1,1,1)$ and $X_{15,s} = 0$, then we may either answer immediately, or first use one or more lifelines. Additional non-uniqueness comes from the fact that the lifelines ℓ_2 and ℓ_3 are equivalent for questions 11 to 15 (though priority should be given to ℓ_3 in this case: see [3, Chapter 9]). At stage $(15, (0,1,0), 4)$, it turns out that it is optimal to use either lifeline ℓ_1 or ℓ_3, and at stage $(15, (0,0,1), 4)$, it is optimal to use either lifeline ℓ_1 or ℓ_2. Except for these cases, the strategy always prescribes exactly one optimal action, which is indicated in Figure 3.

Relative values of the lifelines. The "most powerful" lifeline is Ask-the-audience, as can be seen from the inequality $u^*(1,(1,1,0),j) > u^*(1,s,j)$, for $s \in \{(1,0,1),\ (0,1,1)\}$ and all $j \in \{0,\dots,4\}$, which comes out of the computations and means that if initially we are allowed only one lifeline, then we should select Ask-the-audience. For

$n = 1, \ldots, 5$, if ℓ_2 and ℓ_3 are available, then our strategy prefers to use the weaker lifeline Phone-a-friend, and to save Ask-the-audience for later.

For $n = 6, 7, 8$, if all three lifelines are available and we are in state 2, then it is preferable to use 50:50 and save the other two lifelines for later, but if we are in state 4, then 50:50 is not powerful enough and we should use Ask-the-audience.

In state 1, it is worth answering without using a lifeline if $n \leq 12$. In states 2 or 4, it is optimal to use a lifeline if one is available, unless $n = 6$ and $s = (1, 0, 1)$. In state 4, if no lifeline is available, then we should quit unless we have just passed a milestone ($n = 1, 6, 7$ and 11). In state 2, if no lifeline is available, we can risk answering even if the milestones are farther away ($n =$ 1–4, 6–8 and 11–12).

Estimates of the expected reward. If we do use the strategy described above, what will be our payoff R, on average? The expected utility V of our payoff is

$$V = \sum_{k=0}^{4} q_1(k) u^*(1, (0, 0, 0), k),$$

the numerical value of which is $V = 63.34$, so an estimate of our expected payoff is $V^2 \simeq 4012$ dollars. By Jensen's inequality $E(Y) \geq (E(\sqrt{Y}))^2$, valid for all non-negative random variables Y, this is a (rather severe) underestimate.

We can get an upper bound on R by setting $p = 1$, which means that we set $u(x) = x$, and again go through the calculation of the $u^*(n, s, j)$. This leads to the strategy that maximizes the expected payoff. The value of the expected payoff under this new strategy is \$19252.92. The optimal strategy for this second optimization criterion is not so different from the previous one, so this upper bound is probably relatively close to the expected payoff of the strategy presented in Figure 3.

When the player has correctly answered $n - 1$ questions and has not yet seen question n, we can use the quantity

$$\sum_{k=0}^{4} q_n(k)\, u^*(n, s, k) \tag{4}$$

(computed with $p = 1$) to estimate the expected payoff if we proceed optimally from that stage on, and the availability of lifelines is described by $s \in S$. These quantities (rounded to the nearest integer) are shown in Table 7 for $n = 1, 6, 11$. It is interesting to note that even after correctly answering the first ten questions, the expected payoff is still far below a million dollars (\$230627 if all three lifelines are still available, and \$84585 if all lifelines have been used).

The chances of winning a million. We can estimate this probability by seeking to "maximize the probability that we reach state $(16, W)$." For this, set $f(16, W) = 1$ and $f(n, k) = 0$ for all other values of n and k, and then again calculate the $u^*(n, s, j)$. With these new values, formula (4) gives the probability, given that the player has

n	$	I (0,0,0)	II (1,0,0)	III (0,1,0)	IV (0,0,1)	V (0,1,1)	VI (1,0,1)	VII (1,1,0)	VIII (1,1,1)
1	100	19252	12347	11609	10319	5537	6319	7139	3257
6	2000	37927	24653	24232	21063	12072	13041	15120	7233
11	64000	230627	174533	172597	172597	121571	124859	124859	84585

Table 7: Upper bounds on the expected payoff under the optimal strategy. In row n, for each value of s, the entry is an estimate of the player's expected payoff given that he has correctly answered $n-1$ questions and has not yet seen question n.

correctly answered $n-1$ questions and has not yet seen question n, that with this new strategy, he will reach state $(16, W)$. This is an upper bound for the probability that he will win a million dollars if he uses the strategy presented in Figure 3.

For $n = 1, 6, 11$, these probabilities are shown in Table 8. In particular, a player who starts with three lifelines has a chance of about $1/100$ of winning a million dollars.

n	$	I (0,0,0)	II (1,0,0)	III (0,1,0)	IV (0,0,1)	V (0,1,1)	VI (1,0,1)	VII (1,1,0)	VIII (1,1,1)
1	100	0.0111	0.0065	0.0060	0.0053	0.0026	0.0029	0.0033	0.0013
6	2000	0.0227	0.0134	0.0128	0.0112	0.0056	0.0063	0.0072	0.0030
11	64000	0.1771	0.1180	0.1197	0.1197	0.0718	0.0735	0.0735	0.0399

Table 8: Upper bounds on the probability of winning a million. In row n, for each s, the entry is an estimate of the conditional probability that the player will win a million dollars, given that he has correctly answered $n-1$ questions and has not yet seen question n.

Acknowledgement. The authors thank the referee for a careful reading of this paper.

References

[1] Cairoli, R. & Dalang, R.C. *Sequential Stochastic Optimization.* John Wiley & Sons, New York, 1996.

[2] Dubins, L.E. & Savage, L.J. *Inequalities for Stochastic Processes: How To Gamble If You Must.* Dover Publ., New York, 1976.

[3] Haigh, J. *Taking Chances* (2nd ed.). Oxford University Press, New York, 2003.

[4] Maitra, A.P. & Sudderth, W.D. *Discrete Gambling and Stochastic Systems.* Springer Verlag, Berlin, 1996.

[5] Merton, R.C. *Continuous-Time Finance.* Blackwell Publ., Cambridge, Massachusetts, 1992.

[6] Puterman, M.L. *Markov Decision Processes.* John Wiley & Sons, New York 1994.

7 Appendix. "Who wants to be a millionaire?": the optimal strategy

The eight tables in Figure 3, numbered from I to VIII, indicate the optimal manner of using the lifelines and the best time to quit the game.

The symbols. The symbols R, Q, and the three lifeline symbols shown in Figure 2 respectively correspond to the actions "Answer the question," "Quit the game," "use the 50:50 lifeline," "use the lifeline Phone-a-Friend," and "use the lifeline Ask-the-Audience."

Figure 2: The symbols for the three lifelines

The numbers 0, 1, 2, 3, and 4 on the second line of each of the eight tables represent the player's degree of uncertainty about the correct answer. The number 0 is for the case where the player knows the correct answer (zero uncertainty), the number 1 for the case where the player is quite confident, but not certain, that he knows the correct answer, the number 2 for the case where the player hesitates between two of the answers and considers the other two as unlikely, the number 3 is for the case where the player has just used the 50:50 lifeline and still does not know which of the two remaining answers is correct, and the number 4 is for the case where the player has no idea of the correct answer.

How to use the tables. At each stage in the game, the player selects the table that corresponds to the lifelines that are still available (the lifelines whose symbols are crossed out are those that are no longer available). He then selects the column in that table which corresponds to his degree of uncertainty about the answer. Finally, he selects the row in that table labelled with the number (and dollar value) of the question. Then he should carry out the action indicated in the table at the intersection of that column and row.

An example. If the player has not yet used any of the lifelines, is currently at question 6 and is hesitating between two answers, then table I tells him to use the 50:50 lifeline. Once he has done this and this lifeline is no longer available, he moves to table II. If he is now confident that he knows the answer, then column 1 of table II tells him to give the answer. On the other hand, if he is still hesitating between the two remaining answers, then column 3 of table II tells him to use the lifeline Phone-a-friend. In this case, his only remaining lifeline is Ask-the-audience, so he moves to table VII, and so on.

A comment. The most important factor in the game is the player's knowledge: a well-informed player will generally do better than one who is less-informed, and the strategy for using the lifelines is at a second level of importance. The player should consider that the tables are designed to help him decide when to use each lifeline and when to quit. Since they have been computed for an "average player," it may be reasonable in some cases to act differently from what the tables indicate.

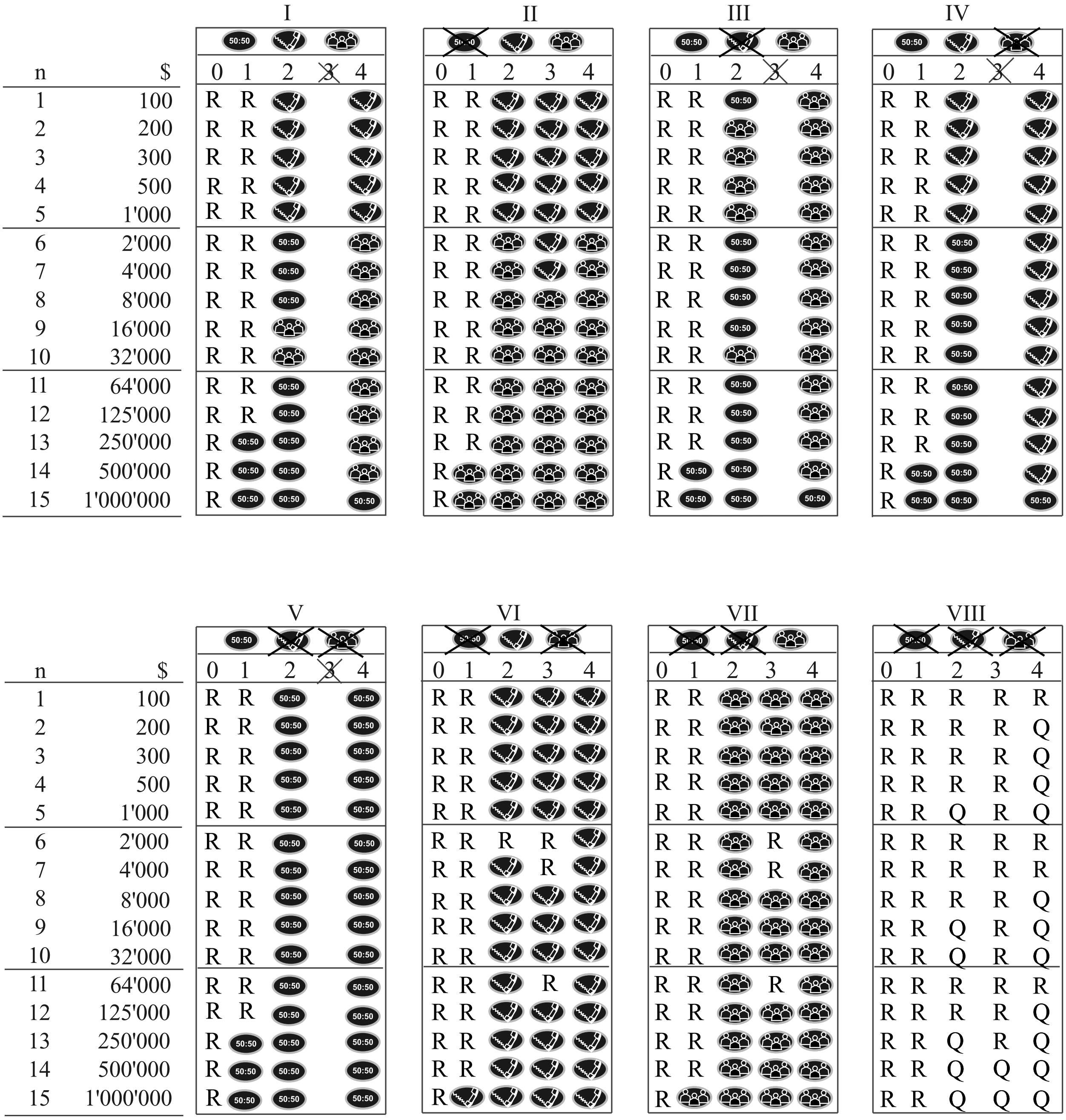

Figure 3: The optimal strategy. The symbols are explained at the beginning of Section 7.