



## **Deciding when to Quit the Gambler's Ruin Game with Unknown Probabilities**

Filipo Studzinski Perotto, Imen Trabelsi, Stéphanie Combettes, Valérie Camps,  
Elsy Kaddoum, Nicolas Verstaevel

### **► To cite this version:**

Filipo Studzinski Perotto, Imen Trabelsi, Stéphanie Combettes, Valérie Camps, Elsy Kaddoum, et al.. Deciding when to Quit the Gambler's Ruin Game with Unknown Probabilities. *International Journal of Approximate Reasoning*, 2021, 137, pp.16-33. <10.1016/j.ijar.2021.06.013>. <hal-03284848>

**HAL Id: hal-03284848**

**<https://hal.science/hal-03284848v1>**

Submitted on 12 Jul 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# Deciding when to Quit the Gambler’s Ruin Game with Unknown Probabilities<sup>\*</sup>

Filipo Studzinski Perotto<sup>a,1,\*</sup>, Imen Trabelsi<sup>b</sup>, Stéphanie Combettes<sup>a</sup>,  
Valérie Camps<sup>a</sup>, Elsy Kaddoum<sup>a</sup>, Nicolas Verstaevel<sup>a</sup>

<sup>a</sup>Toulouse University / Université Paul Sabatier / IRIT, Toulouse, France

<sup>b</sup>Normandy University / INSA de Rouen / LITIS, Rouen, France

---

## Abstract

In the standard definition of the classical *gambler’s ruin* game, a persistent player enters in a stochastic process with an initial budget  $b_0$ , which is, round after round, either increased by 1 with probability  $p$ , or decreased by 1 with probability  $1 - p$ . The player *wins* the game if the budget reaches a given objective value  $g$ , and *loses* the game if the budget drops to zero (the gambler is ruined). This article introduces the *decisional gambling process*, where the parameter  $p$  is hidden, and the player has the possibility to stop the game at any round keeping earnings. In this case, the best a player can do is to maintain an estimate of  $p$  based on the observed outcomes, and use it to decide whether is better to stay or quit the game. The main contribution of this article is to bring the question of finding the *optimal stopping time* to the *gambler’s ruin* game. Different heuristics are analyzed and evaluated according to their performance in maximizing the gambler’s expected final budget.

**Keywords:** Gambler’s Ruin, Stochastic Process, Stopping Time, Bayesian Estimation

---

---

<sup>\*</sup>This work was supported by the Regional Council of Occitanie, France.

<sup>\*</sup>Corresponding author

Email address: [filipo.perotto@irit.fr](mailto:filipo.perotto@irit.fr) (Filipo Studzinski Perotto)

<sup>1</sup>IRIT - Université Paul Sabatier, 118 Route de Narbonne, F-31062, Toulouse, France.

## 1. Introduction

The *gambler's ruin* game is a fundamental problem in the domain of stochastic processes [1, 2, 3, 4, 5]. The earliest known mention of it dates back to a 1656 letter from Blaise Pascal to Pierre de Fermat, followed by the books of Christiaan Huygens (1657), Jacob and Nicolaus Bernoulli (1713), Pierre de Montmort (1708), and Abraham de Moivre (1718), inaugurating the mathematical study of probabilities [6, 7, 8, 9, 10].

Despite such a long history in the literature, the *gambler's ruin* is still an actively studied problem [11, 12, 13, 14, 15, 16, 17, 18, 19, 20]. It is also used as a model for diverse applied problems, such as business risk [21, 22, 23], quantum mechanics [24, 25], anomaly detection [26], material properties [27, 28, 29], or yet other physical [30, 31], biological [32, 33], and even social phenomena [34]. The *gambler's ruin* has recently been associated with a survival version of *multiarmed bandits* [35, 36, 37], which in turn is an essential model for studying sequential decision and learning problems [38].

In the standard setting, the player enters the game with an initial budget  $b_0$ , which must be positive and integer, and the goal is to increase the budget to a given objective value  $g$ . At each round  $t \in \mathbb{N}^+$ , the player either earns 1 unit with probability  $p$ , increasing the current budget, or conversely loses 1 unit with probability  $q = 1 - p$ , decreasing the current budget. The process stops the first time the budget reaches 0 or  $g$ : the gambler either lost everything and is ruined, or earned as much as wanted and won the game. The probability of success  $p$  is stationary, i.e. identical at every round.

Different variations of the problem can be found in the literature. A common alternative is to consider the objective budget  $g$  to be infinite [1, 39, 40]. In such a scenario, depending on the initial budget  $b_0$  and on the probability of success  $p$ , the process will either stop, with the gambler getting ruined, or, conversely, will run indefinitely, with the gambler becoming arbitrarily rich. Other modified versions of the game include variable or asymmetric rewards for success and failure [41, 42, 43, 44, 45], non-stationary probabilities [46], multiple players or

dimensions [47, 48, 49, 50, 51, 52], possibility of ties [1, 53], catastrophes and windfalls [54], intermediate absorption [55, 56], delayed rewards [57], continuous vectors [58], etc.

This article addresses an original variation of the *gambler’s ruin* game, called  
 35 a *decisional gambling process*, where the parameter that regulates the distribution function from which the results are drawn is hidden, i.e. the probability of success  $p$  is unknown to the player. In addition, at each round, the gambler must decide between two options: either to continue playing the game in the next round, or to quit it and stop gambling, in this case keeping the current budget. An  
 40 agent facing this alternative setting seeks to maximize its expected final budget, either by playing until winning if the context is favorable, or by avoiding ruin, quitting the game as soon as possible, if the context seems adverse. The agent can estimate the underlying parameter  $p$  from its previous observations, and, round by round, based on that, decide whether it is preferable to continue or to  
 45 quit the game.

Terminology and notation can vary in different references. In this article, we use “gambler” and “player” interchangeably, and generally to refer to someone who plays the *gambler’s ruin* game (GRG). We adopt the term “agent” to refer to an automatic method which plays a *decisional gambling process* (DGP), since  
 50 it must actively participate in the game by choosing an action at each round. We prefer to use the term *round* instead of “time”, “cycle”, or “step”, to refer to the discrete temporal evolution of the process. Some authors employ the terms “fortune”, “wealth”, and “bankroll” to refer to the *budget*. Others make reference to an “adversary”, while we prefer to consider that the game is played against  
 55 an impersonal process. The objective budget  $g$  is sometimes called the “total fortune inside the game”. Also, some authors refer to the *stopping rule* (i.e. the conditions for winning and losing) as the “stopping strategy”. Because in our work the agent can decide to quit the game before that, we reserve the term *strategy* to the policy of actions (in the case of a DGP, the criterion adopted by  
 60 the agent for deciding, at each round, whether to play or to quit the game).

In contrast with the classical *gambler’s ruin*, which is a pure stochastic game

where the flow of events does not depend on any player’s decision, the *decisional gambling process* constitutes a sequential decision problem in which the agent must choose an action (either to continue or to quit the game) at each round,  
65 based on the previous observations, and guided by an objective optimization function: to maximize its expected final budget. Similar to [59], the agent searches for an optimal stopping time, but here without knowing  $p$ .

The problem involves a kind of exploration-exploitation dilemma: on the one hand, when the agent chooses to abandon the game, it knows the amount  
70 of its current budget, compared to the estimated future earnings that could be obtained if it had kept playing; on the other hand, that estimate comes from the observations, and playing another round brings more information about the underlying parameter, which allows to improve the estimation precision, but implies the risk of having additional losses.

75 In addition to introducing the *decisional gambling process* as an original problem, another contribution of this article is the definition and evaluation of a set of heuristics for deciding when to stop playing the game. The expected utilities of three straightforward heuristics (*bold*, *timid*, and *naive*) are analytically demonstrated, and those methods are redefined in terms of tolerable loss. The  
80 theoretical optimal behavior (*oracle*), which indicates the best decision to take if the parameter  $p$  could be known, is also defined, serving as a reference to calculate the regret of other strategies. From a broader point of view, the scenario studied in this article brings together two classical problems: the *gambler’s ruin* game and the search for the *optimal stopping time* in stochastic processes [60, 61].

85 In the rest of the article, Section 2 formally states the definition of a *decisional gambling process*, explains the similarities and differences in relation to the classical *gambler’s ruin* game, and presents some of its properties. Section 3 introduces a set of heuristic methods supported by theoretical analysis. Section 4 compares the methods and discusses the results, and Section 5 concludes the  
90 article, pointing out the limits of this research, and indicating possible future work.

## 2. Decisional Gambling Process

The first contribution of this article is defining a modified version of the classical *gambler's ruin* game (GRG) in which the parameter  $p$  is not known,  
 95 and the gambler can decide to stop the game at any time. We call this variation of the problem a *decisional gambling process* (DGP), into which the gambler becomes an *agent*.

DGPs and GRGs share the same parameters. Both models can be seen as discrete stochastic machines. At each successive round, the agent's budget can  
 100 be either increased or decreased by 1 with respective stationary probabilities  $p$  and  $1 - p$ . The parameter  $b_0$  defines the initial budget and  $g$  defines the goal or objective budget.

**Definition 1.** *A DGP is constituted of a Bernoulli process into which  $p$  is the constant probability of success in a trial. An initial natural budget  $b_0$  evolves by  
 105 steps of size  $+1$  or  $-1$  depending on the trial result at each round. The process stops when the budget attains either 0 or  $g$ , or when a stopping decision is made by the policy  $\pi$ .*

Formally, let  $b_0$ ,  $g$ , and  $p$  be the parameters of a DGP, such that:

$$\begin{cases} b_0 \in \mathbb{N}^+ & \text{is the initial budget,} \\ g \in \mathbb{N} \mid g > b_0 & \text{is the goal or objective budget,} \\ p \in [0, 1] & \text{is the probability of } \textit{success} \text{ in a single round.} \end{cases} \quad (1)$$

From the parameter  $p$  we can state two other derived parameters:

$$\begin{cases} q := 1 - p & \text{is the probability of } \textit{failure} \text{ in a single round,} \\ r := \frac{1-p}{p} & \text{is the } \textit{failure-success} \text{ ratio.} \end{cases} \quad (2)$$

At each round  $t$ , the gambler can observe either a *success*, represented by  $X_t = 1$ , which returns a constant reward  $R_t = +1$ , or a *failure*, represented by  $X_t = 0$ , which returns a constant reward  $R_t = -1$ . It means that the immediate reward received by the agent only depends on the result of the corresponding

trial:

$$\forall t \in \mathbb{N}^+ : \begin{cases} X_t \in \{0, 1\} & \text{is the trial result at round } t, \\ R_t := 2X_t - 1 & \text{is the reward received at round } t. \end{cases} \quad (3)$$

In this way, executing a DGP (as well as a GRG) corresponds to running a series of rewarded *Bernoulli trials* where each observation comes from the *realization* of the *random variable*  $X_t$  at round  $t$ , and where each trial is independent and identically distributed, drawn from a stationary Bernoulli distribution:

$$\forall t \in \mathbb{N}^+ : X_t \sim \mathbf{Bern}(p) = \begin{cases} 1 & \text{with probability } p, \\ 0 & \text{with probability } q = 1 - p. \end{cases} \quad (4)$$

As a model, the DGP extends the GRG, which in turn extends the *Bernoulli process* [62], defined as a *stochastic process*  $\{X_t\}_{t \in \mathbb{N}^+}$  over the *sample space*  $\Omega = \{0, 1\}^{\mathbb{N}^+}$ .

Let  $K_t$ ,  $Z_t$ ,  $S_t$  and  $B_t$  be the state of the game after  $t$  rounds, such that:

$$\forall t \in \mathbb{N} : \begin{cases} K_t := \sum_{i=1}^t X_i & \text{is the number of observed } \textit{successes}, \\ Z_t := t - K_t & \text{is the number of observed } \textit{failures}, \\ S_t := \sum_{i=1}^t R_i = K_t - Z_t & \text{is the cumulated reward,} \\ B_t := b_0 + S_t & \text{is the agent's budget at round } t, \end{cases} \quad (5)$$

with  $K_0 = Z_0 = S_0 = 0$ , and  $B_0 = b_0$ .

Let the stopping time  $\tau \in \mathbb{N}$  be a random variable indicating when the considered DGP stops (i.e. the last played round). The difference between DGP and GRG lies in their *stopping rules*. A GRG stops in the earliest round at which the budget reaches either 0 or  $g$ . A DGP, in addition, can stop when the agent decides to quit the game. For this reason, in a DGP, the action chosen by the agent at each round,  $D_t \in \{\top, \perp\}$ , must be taken into account:

$$\forall t \in \mathbb{N} : D_t = \begin{cases} \perp & \text{if the agent decides to quit the game after round } t, \\ \top & \text{if the agent decides to play the round } t + 1. \end{cases} \quad (6)$$

Since the value of  $\tau$  depends on the realization of  $B_t$  and  $D_t$ , the stopping time is also a *random variable*:

$$\tau := \inf_{t \in \mathbb{N}} \left[ (B_t = 0) \vee (B_t = g) \vee (D_t = \perp) \right]. \quad (7)$$

By analogy with GRG [62, 63, 64], because  $g$  is finite, and since the DGP stopping rule is more restrictive than the GRG stopping rule, it can be demonstrated that the probability for a game to run indefinitely tends to zero as time approaches infinity:

$$\lim_{t \rightarrow \infty} \mathbb{P}(\tau \geq t) = 0, \quad (8)$$

which means that a DGP will *almost always* stop (i.e. eventually the agent wins the game, quits it, or is ruined), and the finiteness of  $\tau$  holds *almost surely*, independently of the adopted policy.

115 Let  $\pi(b_0, g, t, K_t, Z_t, B_t) \rightarrow \{\top, \perp\}$  be the *decision function*, or *policy of actions*, that defines the agent's behavior to produce the sequence  $\{D_t\}_{t \in \mathbb{N}}$ . Note that  $\pi$  can depend on the initial budget  $b_0$ , on the number of played rounds  $t$ , on the number of successes  $K_t$  and failures  $Z_t$  observed until the current round, on the current budget  $B_t$ , and on the objective budget  $g$ , but does not have access  
120 to  $p$ . The decision is made by the agent without knowing this parameter. The only information related to  $p$  comes, indirectly, from  $K_t$  and  $Z_t$ . Also note that the agent makes its first decision at round  $t = 0$ , i.e. before the effective start of the game at round  $t = 1$ .

Let  $\mathbb{E}[B_\tau]$  be the expected final budget, which is kept when the game eventually stops, and which the agent would like to maximize. Thanks to Eq. (8), it can be defined as:

$$\mathbb{E}[B_\tau] = \sum_{t \in \mathbb{N}} \left[ \mathbb{P}(\tau = t) \cdot \mathbb{E}[B_t \mid \tau = t] \right]. \quad (9)$$

Let  $\Pi$  be the set of all possible policies. An optimal strategy  $\pi^*$  is a policy of actions that maximizes the agent's expected final budget for any game configuration:

$$\pi^* := \arg \max_{\pi \in \Pi} \mathbb{E}[B_\tau \mid \pi]. \quad (10)$$



Nevertheless, even if the value of  $p$  is hidden, two important characteristics are assumed to be true: first, the parameter is *stationary* (i.e.  $p$  does not change over time):

$$\forall x \in \{0, 1\}, \forall t \in \mathbb{N}^+ : \mathbb{P}(X_t = x) = \mathbb{P}(X_1 = x), \quad (11)$$

and second, it has a *uniform prior* (i.e. there is no assumed predefined tendency, and  $p$  has *a priori* equal chances to assume any real value between 0 and 1, inclusive). It is equivalent to say that  $p$  is the realization of a random variable  $\theta$ , drawn from a uniform distribution when the game starts [65]:

$$\theta \sim \mathbb{U}(0, 1) = \mathbf{Beta}(1, 1). \quad (12)$$

### 3. Analyzing Theoretically Grounded Strategies

Another contribution of this article is formally analyzing the expected performance of different intuitive heuristic methods: *bold* (which always plays), *timid* (which always quits), *radical* (a combined version of the previous two), and *naive* (which estimates  $p$ ). Their performance is compared to the theoretical performance of an *oracle* policy, i.e. an optimal method that can see the hidden parameter  $p$ . The difference between the *oracle*'s expected final budget, and the expected final budget of a given strategy corresponds to its expected *regret*:

$$\lambda_\pi := \mathbb{E}[B_\tau \mid \pi_{\text{oracle}}] - \mathbb{E}[B_\tau \mid \pi]. \quad (13)$$

125 These four mentioned strategies are described in the next subsections. A comparison between them is shown in Figures 7 and 8. We also show that these strategies can be generalized in the form of a single parametric  $\ell$ -strategy, where  $\ell$  represents the maximal tolerable loss.

Let  $\pi$  be the strategy adopted by the agent,  $B_t$ , its current budget,  $g$ , the objective budget, and  $p$ , the underlying DGP parameter. The probability that the process will stop with the agent quitting the game is defined as:

$$\{\mathbb{P}_\pi^{\text{quit}} \mid B_t, g, p\} := \mathbb{P}(\exists h \in \mathbb{N} : [(\tau = h) \wedge (D_h = \perp)] \mid B_t, g, p, \pi); \quad (14)$$

The probability that the agent will win the game is define as:

$$\{\mathbb{P}_\pi^{win} \mid B_t, g, p\} := \mathbb{P}(\exists h \in \mathbb{N} : [(\tau = h) \wedge (B_h = g)] \mid B_t, g, p, \pi); \quad (15)$$

And the probability that the agent will lose the game is defined as:

$$\{\mathbb{P}_\pi^{lose} \mid B_t, g, p\} := \mathbb{P}(\exists h \in \mathbb{N} : [(\tau = h) \wedge (B_h = 0)] \mid B_t, g, p, \pi). \quad (16)$$

### 3.1. Bold Strategy

A DGP can be reduced to a GRG by the adoption of a *bold* policy, indicated in Eq. (17), which never quits the game. This “everything or nothing” strategy makes the agent play until either winning or getting ruined, which is equivalent to the classical *gambler’s ruin* routine:

$$\pi_{bold} := \top \implies \forall b_0, g, t, K_t, Z_t, B_t \in \mathbb{N} : \{D_t \mid \pi_{bold}\} = \top. \quad (17)$$

130 While a general DGP can be described in the form of an MDP, as shown in Figure 1, the *bold* strategy can be represented as a *finite terminating one-dimensional Markov chain* [66] on the state space  $\{0, \dots, g\}$ , as shown in Figure 2, where each node corresponds to a possible budget. The two *barrier* nodes, 0 and  $g$ , are *absorbing*, and all the other nodes, between 0 and  $g$ , are *transient*.

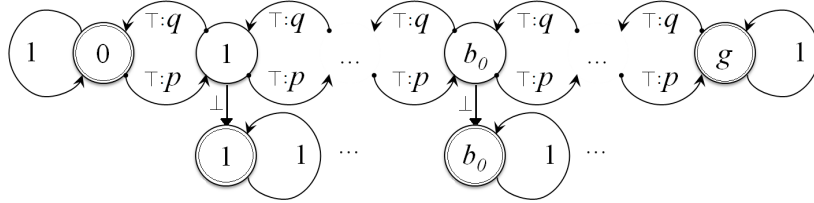


Figure 1: A *decisional gambling process* can be modeled as a *Markovian decision process*. Each node in the top of the graph corresponds to a possible budget while the game is running. The nodes 0 and  $g$  are absorbing (self-loop with probability 1), corresponding, respectively, to the *losing* and *winning* situations. The nodes in the bottom of the graph are also absorbing, corresponding to stopping the game with an intermediate budget, following the agent’s decision to quit the game. The initial state is  $b_0$ .

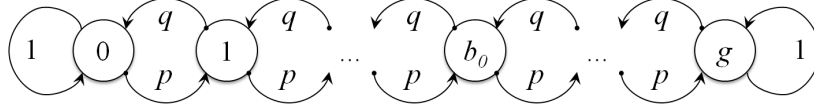


Figure 2: The *gambler's ruin* game can be modeled as a *Markov chain*, and the course of the game as a *random walk*. It also corresponds to the *decisional gambling process* when the player never wants to quit the game (*bold strategy*). Each state corresponds to a possible budget. The process starts on  $b_0$  and, from each transient node, probabilities  $p$  and  $q$  define, respectively, the chance to step into the right, and into the left neighbor. The probability of winning or losing can be deduced by looking at the limiting distribution as time goes to infinity (steady state probabilities) of the chain, since the probability of being in a transient node tends to 0.

Since the *bold* strategy never chooses to quit the game, observing a quitting decision at any round is impossible when using that strategy:

$$\mathbb{P}_{bold}^{quit} = 0. \quad (18)$$

Then, if the process stops (which eventually happens, *almost surely*), the final budget cannot be other than 0 or  $g$ , independent of any other parameter:

$$\forall h \in \mathbb{N} : \left( \tau = h \longrightarrow \{B_h \mid \pi_{bold}\} \in \{0, g\} \right). \quad (19)$$

Thanks to the reducibility of the *bold* strategy to a GRG, it can also be demonstrated [1, 10] that a *bold* gambler with current budget  $B_t$  in a game where the objective budget is  $g$  has the following probability to win:

$$\{\mathbb{P}_{bold}^{win} \mid B_t, g, p\} = \begin{cases} 0 & \text{if } p = 0, \\ 1 & \text{if } p = 1, \\ \frac{1 - \left(\frac{1-p}{p}\right)^{B_t}}{1 - \left(\frac{1-p}{p}\right)^g} & \text{if } p \in (0, 1) \mid p \neq \frac{1}{2}, \\ \frac{B_t}{g} & \text{if } p = \frac{1}{2}. \end{cases} \quad (20)$$

Since the probability that the game will run forever (Eq. (8)) tends to 0, and because the *bold* strategy never asks to quit it (Eq. (18)), then the probability of losing the game, being ruined at the end, is complementary to the probability of winning it:

$$\{\mathbb{P}_{bold}^{lose} \mid B_t, g, p\} = 1 - \{\mathbb{P}_{bold}^{win} \mid B_t, g, p\}. \quad (21)$$

Figure 3 illustrates the probabilities of winning a game in different scenarios using the *bold* strategy.

Since a *bold* gambler always plays until winning ( $B_\tau = g$ ) or losing the game ( $B_\tau = 0$ ), and because these probabilities are known for any given value of  $p$ , the gambler's expected final budget corresponds to:

$$\begin{aligned}\mathbb{E}[B_\tau \mid B_t, g, p, \pi_{bold}] &= g \cdot \{\mathbb{P}_{bold}^{win} \mid B_t, g, p\} + 0 \cdot \{\mathbb{P}_{bold}^{lose} \mid B_t, g, p\} \\ &= g \cdot \{\mathbb{P}_{bold}^{win} \mid B_t, g, p\}.\end{aligned}\tag{22}$$

It is possible to determine a general expectation for  $B_\tau$ , independent of  $p$ , based on the fact that the *a priori* probability distribution of  $p$  is uniform. Thanks to the *lotus* theorem, the general expectation, denoted by  $\xi$ , corresponds to the definite integral between 0 and 1 of Eq. (22) [67, p.2]. That value indicates the expected final budget of a *bold* gambler playing a DGP, without knowing  $p$ , which has been drawn from a uniform distribution. We call this expectation the *utility* of the method:

$$\begin{aligned}\xi_{bold}(b_0, g) &:= \int_0^1 \mathbb{E}[B_\tau \mid b_0, g, p, \pi_{bold}] \, dp \\ &= \int_0^1 g \cdot \{\mathbb{P}_{bold}^{win} \mid b_0, g, p\} \, dp \\ &= g \int_0^1 \{\mathbb{P}_{bold}^{win} \mid b_0, g, p\} \, dp \\ &= g \int_0^1 \frac{1 - \left(\frac{1-p}{p}\right)^{b_0}}{1 - \left(\frac{1-p}{p}\right)^g} \, dp.\end{aligned}\tag{23}$$

The expected duration of the game for a *bold* gambler can be obtained by *first-passage* distribution analysis [64, 42]:

$$\mathbb{E}[\tau \mid b_0, g, p, \pi_{bold}] = \begin{cases} \frac{b_0}{1-2p} - \frac{g}{1-2p} \cdot \frac{1 - \left(\frac{1-p}{p}\right)^{b_0}}{1 - \left(\frac{1-p}{p}\right)^g} & \text{if } p \neq \frac{1}{2}, \\ b_0(g - b_0) & \text{if } p = \frac{1}{2}. \end{cases}\tag{24}$$

### 3.2. Timid and Radical Strategies

In contrast to the *bold* strategy, the *timid* strategy, indicated in Eq. (25), always chooses to quit the game:

$$\pi_{timid} := \perp \implies \forall b_0, g, t, K_t, Z_t, B_t \in \mathbb{N} : \{D_t \mid \pi_{timid}\} = \perp.\tag{25}$$

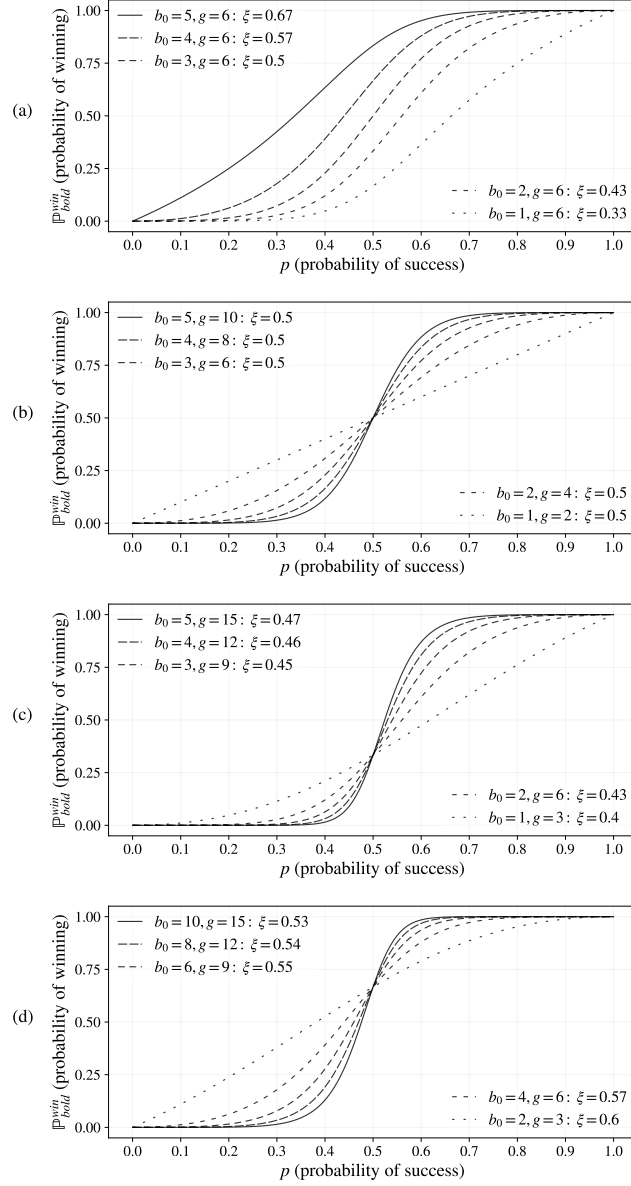


Figure 3: The probability of winning the game (vertical axis) adopting the *bold* strategy (which always chooses playing) depends on the parameter  $p$  (horizontal axis), on the objective budget ( $g$ ), and on the initial budget ( $b_0$ ). The graphic (a) illustrates these probabilities for a constant  $g = 6$ , varying  $b_0$  from 1 to 5. The other graphics show the probabilities for different constant  $b_0/g$  ratios: (b)  $b_0/g = 1/2$ , (c)  $b_0/g = 1/3$ , and (d)  $b_0/g = 2/3$ . The utility ( $\xi$ ) of the *bold* strategy for each case, given a uniform prior for  $p$ , corresponds to the definite integral of the function (i.e. the area under each curve).

The properties of *timid* are trivial to deduce. When the agent uses this strategy, the process always terminates before the first effective round, at  $t =$   
140 0 (consequently, with  $\tau = 0$ ), and the gambler keeps an unchanged budget equivalent to  $b_0$ :

$$\mathbb{E}[\tau \mid \pi_{timid}] = 0, \quad (26)$$

$$\mathbb{P}_{timid}^{quit} = 1, \quad (27)$$

$$\mathbb{P}_{timid}^{win} = \mathbb{P}_{timid}^{lose} = 0, \quad (28)$$

$$\mathbb{E}[B_\tau \mid b_0, \pi_{timid}] = \xi_{timid}(b_0) = b_0. \quad (29)$$

Both *timid* and *bold* are blind strategies: they do not take into account any information about the game for improving their decision making. In fact, by looking at the  $b_0/g$  ratio, it is easy to see that when  $b_0$  is closer to  $g$  than to 0 (i.e.  $b_0/g > 0.5$ ) the gambler has more to lose than to gain. Therefore, a simple improvement over those two heuristics, called *radical* strategy, corresponds to becoming *timid* when  $b_0 > g/2$  and *bold* otherwise:

$$\pi_{radical}(b_0, g) := \begin{cases} \perp & \text{if } b_0 > g/2, \\ \top & \text{if } b_0 \leq g/2. \end{cases} \quad (30)$$

### 3.3. Oracle Policy

When  $p = 1/2$ , the gambler neither earns nor loses money on average. In that case, the progression of the budget over time constitutes a *martingale*: the expected budget at round  $t + 1$  is equivalent to the budget at round  $t$  [68]. When  $p < 1/2$ , the gambler loses money on average, and the game is said to be a *supermartingale*: the expected budget at round  $t + 1$  is smaller than the budget at round  $t$  [69]. When  $p > 1/2$ , the gambler earns money on average,

then the process constitutes a *submartingale*: the expected budget at round  $t+1$  is greater than the budget at round  $t$  [69]. Formally:

$$\forall t \in \mathbb{N} : \begin{cases} \mathbb{E}[B_{t+1} \mid \pi_{bold}] = B_t & \text{if } p = 0.5, \\ \mathbb{E}[B_{t+1} \mid \pi_{bold}] < B_t & \text{if } p < 0.5, \\ \mathbb{E}[B_{t+1} \mid \pi_{bold}] > B_t & \text{if } p > 0.5. \end{cases} \quad (31)$$

Considering a GRG, or a DGP with a *bold* strategy (i.e.  $D_t = \top$ ,  $\forall t \in \mathbb{N}$ ), and based on Eq. 31, it can be demonstrated that:

$$\forall t \in \mathbb{N} : \begin{cases} \mathbb{E}[B_\tau \mid \pi_{bold}] = b_0 & \text{if } p = 0.5, \\ \mathbb{E}[B_\tau \mid \pi_{bold}] = 0 & \text{if } p < 0.5, \\ \mathbb{E}[B_\tau \mid \pi_{bold}] = g & \text{if } p > 0.5. \end{cases} \quad (32)$$

Based on these facts (Eqs. (31) and (32)), and considering the objective of maximizing the gambler's expected final budget, if the parameter  $p$  could be known in advance, the decision would be trivial. In a martingale, this result does not depend on the gambler's decision, and any expected future budget is equivalent to the initial budget. In a supermartingale, the budget is expected to decrease round after round, so the best decision is to quit the game immediately. In a submartingale, the budget is expected to increase round after round, and the best decision is to continue playing until the game finishes. The *oracle* policy can be used for calculating the regret (Eq. (13)) of other strategies. The resulting behavior of that policy corresponds to being *bold* when the probability of success is favorable ( $p \geq 0.5$ ), and being *timid* otherwise:

$$\{\pi_{oracle} \mid p\} := \begin{cases} \perp & \text{if } p < 0.5, \\ \top & \text{if } p \geq 0.5. \end{cases} \quad (33)$$

Since the *oracle* always quits the game at the beginning if  $p < 0.5$ , keeping the initial budget, and always plays until winning or losing the game when  $p \geq 0.5$ , the gambler's expected final budget corresponds to:

$$\mathbb{E}[B_\tau \mid B_t, g, p, \pi_{oracle}] = \begin{cases} g \cdot \{\mathbb{P}_{bold}^{win} \mid B_t, g, p\} & \text{if } p \geq 0.5, \\ b_0 & \text{if } p < 0.5. \end{cases} \quad (34)$$

and its general utility, always considering a uniform prior for  $p$ , is:

$$\begin{aligned}
\xi_{oracle}(b_0, g) &= \int_0^1 \mathbb{E}[B_\tau \mid b_0, g, p, \pi_{oracle}] \, dp \\
&= \int_0^{0.5} b_0 \, dp + \int_{0.5}^1 g \cdot \{\mathbb{P}_{bold}^{win} \mid B_t, g, p\} \, dp \\
&= b_0 + g \cdot \int_{0.5}^1 \{\mathbb{P}_{bold}^{win} \mid B_t, g, p\} \, dp \\
&= b_0 + g \cdot \int_{0.5}^1 \frac{1 - \left(\frac{1-p}{p}\right)^{b_0}}{1 - \left(\frac{1-p}{p}\right)^g} \, dp.
\end{aligned} \tag{35}$$

### 3.4. Naive Strategy

In a DGP, the parameter  $p$  is not known in advance, and the agent has the power to stop the game before the end, to try to maximize its expected final budget. For doing so, the agent can estimate  $p$  while playing, from observations, which corresponds to estimating the parameter that regulates a binomial distribution [70, 71]. The binomial distribution gives the probability of having a certain number of successes on a sequence of Bernoulli trials. The probability of having  $K_t = k$  successes (and  $Z_t = z = t - k$  failures) in  $t$  trials, given  $p$  (and  $q = 1 - p$ ) is:

$$\mathbb{P}(K_t = k \mid t, p) = \mathbf{Bin}(k; t, p) = \binom{t}{k} p^k q^z. \tag{36}$$

In a Bayesian approach [72], the beta distribution is conjugate to the binomial likelihood. Then, assuming a uniform prior (i.e. considering that  $p$  can be *a priori* placed uniformly between 0 and 1, following Eq. (12), which corresponds to a **Beta**(1, 1) distribution), the posterior probability density function  $\varphi$  of  $p$  takes the form of a beta distribution:

$$\varphi(p \mid k, t) = \mathbf{Beta}(p; k + 1, z + 1) = (t + 1) \binom{t}{k} p^k q^z, \tag{37}$$

and the Bayesian estimate of  $p$ , denoted by  $\hat{p}$ , corresponds to the mean of the posterior distribution [73]:

$$\hat{p} = \mathbb{E}[p \mid k, t] = \int_0^1 p \cdot \varphi(p \mid k, t) \, dp = \frac{k + 1}{t + 2}. \tag{38}$$



A naive approach for approximating the probability of winning a game based on the observations is estimating  $p$  using Eq. (38), and then using it in Eq. (20). The strategy corresponds to continuing playing the game while the estimate of  $p$  suggests a positive mean reward. It is a naive approach since only the estimation of  $p$  is taken into account for taking a decision, independent of the current budget and ignoring the confidence on the estimation. The *naive* strategy is similar to the *oracle* policy, but using the estimation of  $p$  at the decision time:

$$\pi_{naive}(t, \hat{p}) = \begin{cases} \perp & \text{if } \hat{p} < 0.5, \\ \top & \text{if } \hat{p} \geq 0.5. \end{cases} \quad (39)$$

We would like to demonstrate that the naive strategy, described in Eq. (39) based on the estimate of  $p$ , is equivalent to the strategy that quits the game when the budget drops to  $b_0 - 1$ , as proposed by Theorem (1):

**Theorem 1.** *In a DGP, the Bayesian estimate of  $p$  is lower than  $1/2$ , considering any symmetric beta prior (i.e.  $\alpha = \beta$ ), if and only if the current budget  $B_t$  is lower than the initial budget  $b_0$ .*

*Proof.* Since, following Eq. (5), the budget at round  $t$  ( $B_t$ ) must correspond to the initial budget  $b_0$  plus the number of successes  $K_t$  (which return reward +1) minus the number of failures  $Z_t = t - K_t$  (which return reward -1), it is possible to deduce the number of observed successes based on the number of played rounds, the initial and current budget:

$$B_t = b_0 + K_t - Z_t = b_0 + 2K_t - t \implies K_t = \frac{B_t - b_0 + t}{2}. \quad (40)$$

Then, considering Eq. (40), it is possible to demonstrate that the conditions presented in Eqs. (38) and (39), which guide the *naive* strategy, supposing a

uniform prior, can be transformed as follows:

$$\begin{aligned}
\frac{K_t + 1v}{t + 2v} < \frac{1}{2} &\iff K_t + 1v < \frac{t + 2v}{2} \\
&\iff K_t < \frac{t + 2v}{2} - 1v \\
&\iff K_t < \frac{t}{2} \\
&\iff \frac{B_t - b_0 + t}{2} < \frac{t}{2} \\
&\iff B_t - b_0 + t < t \\
&\iff B_t - b_0 < 0 \iff B_t < b_0.
\end{aligned} \tag{41}$$

150

□

The transformation presented in Eq. (41) demonstrates that, in fact, the *naive* strategy stops the game at the first round in which the current budget falls under the initial budget ( $B_t < b_0$ ). Because the budget is integer, it means stopping when  $B_t = b_0 - 1$ .

The proof can be specialized to demonstrate that using the uniform distribution **Beta**(1, 1) results in the same behavior: the agent quits the game the first time the budget falls below  $b_0$ , as indicated in Eq. (42):

$$\hat{p} < 0.5 \iff \frac{K_t + 1}{t + 2} < \frac{1}{2} \iff \frac{B_t - b_0 + t}{2} < \frac{t}{2} \iff B_t < b_0. \tag{42}$$

Note that using the frequentist estimator ( $v = 0$ ) instead of the Bayesian estimator of  $p$  results in the same policy, as evidenced by Eq. (43), and illustrated in Figure 4:

$$\bar{X}_t < 0.5 \iff \frac{K_t}{t} < \frac{1}{2} \iff \frac{B_t - b_0 + t}{2} < \frac{t}{2} \iff B_t < b_0. \tag{43}$$

155

A *naive* gambler either wins the game or quits it with budget  $B_\tau = b_0 - 1$ . In general (i.e. when  $b_0 > 1$ ), quitting the game means deciding to stop the process, but in the particular case in which  $b_0 = 1$ , *naive* behaves like *bold* (i.e. the gambler either wins or is ruined, since  $B_\tau = b_0 - 1 = 0$  in that case).

The probability of winning the game using the *naive* strategy can be deduced by reinterpreting Eq. (20). That formula indicates the probability of winning

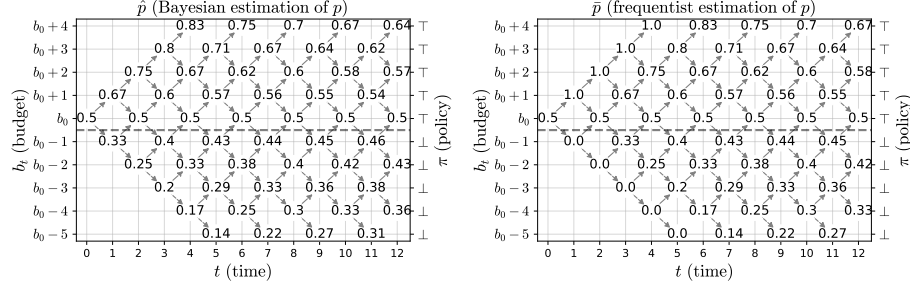


Figure 4: The naive policy estimates  $p$  using the Bayesian average. When  $\hat{p} < 0.5$  the gambler stops the game. It corresponds to the situation in which  $b_t < b_0$ . For that reason, the equations used for calculating *lose* probabilities ( $B_\tau = 0$ ) with the *bold* strategy can be translated into *quit* probabilities ( $B_\tau = b_0 - 1$ ) with the *naive* strategy. A similar alternative is using the frequentist estimation of  $p$ , but the resulting policy corresponds exactly to *naive*. In fact, roughly speaking, in this case, conducting a Bayesian estimation corresponds to starting a frequentist estimation after have seeing one success and one failure (the prior). Note that, because both the confidence in the estimation and the budget magnitude are not taken into account in that strategy, a failure at the very first round makes the gambler quit the game.

when playing boldly. The same scheme of probabilities can be used to describe an agent which plays either until winning, or until observing a budget equivalent to  $b_0 - 1$ . The insight for this translation is based on the fact that the structure of paths leading from  $b_0$  to  $b_0 - 1$  or to  $g$  is isomorphic to the structure of paths leading from 1 to 0 or to  $g - b_0 + 1$ , as shown in Figure 5. In this way, the probability of winning the game, using the *naive* strategy, before the budget drops to  $b_0 - 1$  corresponds to:

$$\{\mathbb{P}_{naive}^{win} \mid B_t, b_0, g, p\} = \{\mathbb{P}_{bold}^{win} \mid B'_t, g', p\}, \quad (44)$$

where  $B'_t := 1$ , and  $g' := g - b_0 + 1$ .

The (complementary) probability of stopping the game using the *naive* strategy, since the budget  $b_0 - 1$  is observed, is:

$$\{\mathbb{P}_{naive}^{quit} \mid B_t, b_0, g, p\} = \begin{cases} 1 - \{\mathbb{P}_{naive}^{win} \mid B_t, b_0, g, p\} & \text{if } b_0 > 1, \\ 0 & \text{if } b_0 = 1. \end{cases} \quad (45)$$

The only case in which the gambler can be ruined using the *naive* strategy is

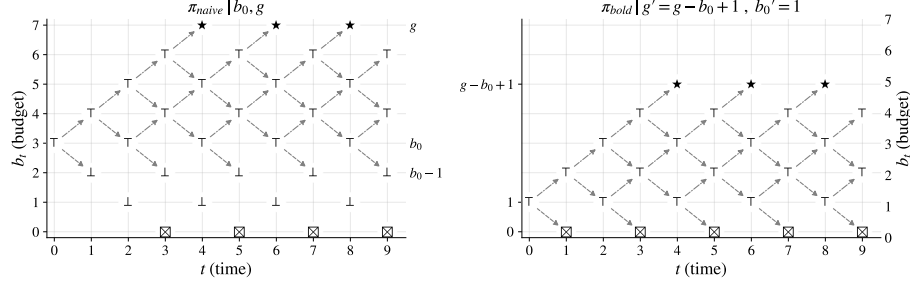


Figure 5: The probability of reaching  $g$  (winning the game) before  $b_0 - 1$  is the same as reaching  $g - b_0 + 1$  before being ruined when starting with budget 1. Note the isomorphic structure of both graphs if they are taken as Markov chains where up arrows have probability  $p$  and down arrows have probability  $q$ .

when  $b_0 = 1$ :

$$\{\mathbb{P}_{naive}^{lose} \mid B_t, b_0, g, p\} = \begin{cases} 1 - \{\mathbb{P}_{naive}^{win} \mid B_t, b_0, g, p\} & \text{if } b_0 = 1, \\ 0 & \text{if } b_0 > 1. \end{cases} \quad (46)$$

Thanks to Equations (44) and (45), it is possible to define the expected final budget of a gambler using the *naive* strategy:

$$\mathbb{E}[B_\tau \mid B_t, b_0, g, p, \pi_{naive}] = g \cdot \{\mathbb{P}_{naive}^{win} \mid B_t, b_0, g, p\} + (b_0 - 1) \cdot \{\mathbb{P}_{naive}^{quit} \mid B_t, b_0, g, p\}, \quad (47)$$

and its general utility:

$$\xi_{naive}(b_0, g) = \int_0^1 \mathbb{E}[B_\tau \mid b_0, g, p, \pi_{naive}] \, dp. \quad (48)$$

### 160 3.5. Acceptable Loss Strategy

A general formulation, called *loss* strategy, and connecting *timid*, *bold* and *naive*, can be done based on a maximal tolerable loss parameter  $\ell \in \mathbb{N}$ , into which  $\ell = b_0$  for *bold*,  $\ell = 0$  for *timid*, and  $\ell = 1$  for *naive*:

$$\pi_\ell(K_t, Z_t) = \begin{cases} \perp & \text{if } Z_t - K_t \leq \ell \text{ (i.e. when } B_t \leq b_0 - \ell), \\ \top & \text{if } Z_t - K_t > \ell \text{ (i.e. when } B_t > b_0 - \ell), \end{cases} \quad (49)$$

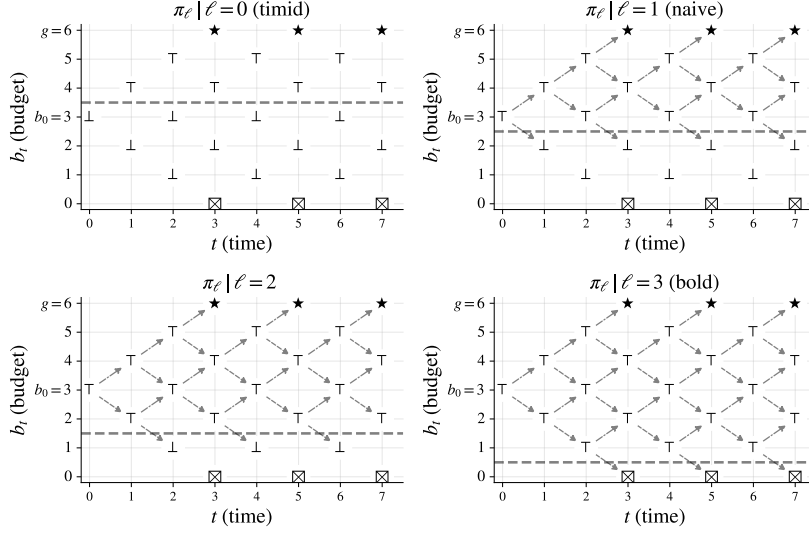


Figure 6: Illustration of the *maximal loss* strategy for different values of  $\ell$  when  $b_0 = 3$  and  $g = 6$ . Note that when  $b_t$  drops to  $b_0 - \ell$  the agent decides to stop gambling, except if it is already ruined (case  $\ell = b_0$ ).

where  $K_t$  and  $Z_t$  are respectively the number of observed successes and failures.

Based on a reasoning similar to the one used for deriving Equations (44), (45) and (46), illustrated in Figure 5, we can use the well-known GRG winning and losing probability formulas to understand the  $\ell$ -strategy. All we need to do is translating the reference points on the path structure, using  $\ell$  in the place of  $b_0$ , and  $g - b_0 + 1$  instead of  $g$ , then obtaining:

$$\{\mathbb{P}_\ell^{win} \mid B_t, b_0, g, p\} = \{\mathbb{P}_{bold}^{win} \mid B_t'', g'', p\}, \quad (50)$$

with  $B_t'' = \ell$  and  $g'' = g - b_0 + \ell$ , and:

$$\{\mathbb{P}_\ell^{quit} \mid B_t, b_0, g, p\} = \begin{cases} 1 - \{\mathbb{P}_\ell^{win} \mid B_t, b_0, g, p\} & \text{if } b_0 > \ell, \\ 0 & \text{if } b_0 \leq \ell, \end{cases} \quad (51)$$

$$\{\mathbb{P}_\ell^{lose} \mid B_t, b_0, g, p\} = \begin{cases} 1 - \{\mathbb{P}_\ell^{win} \mid B_t, b_0, g, p\} & \text{if } b_0 \leq \ell, \\ 0 & \text{if } b_0 > \ell. \end{cases} \quad (52)$$

The expected final budget of a gambler using the *loss* strategy is:

$$\begin{aligned} \mathbb{E}[B_\tau \mid B_t, b_0, g, p, \pi_\ell] = \\ g \cdot \{\mathbb{P}_\ell^{win} \mid B_t, b_0, g, p\} + (b_0 - \ell) \cdot \{\mathbb{P}_\ell^{quit} \mid B_t, b_0, g, p\}, \end{aligned} \quad (53)$$

and its general utility:

$$\xi_\ell(b_0, g) = \int_0^1 \mathbb{E}[B_\tau \mid b_0, g, p, \pi_\ell] \, dp. \quad (54)$$

The intuition behind this generalization can be understood by looking at Figure 6 and observing the displacement of the decision frontier (the horizontal line that divides the plane, indicating where the decision changes from *playing* to *quitting*).

#### 4. Discussion

Figures 7 and 8 present the performance of each of the analyzed strategies given the game conditions ( $b_0$  and  $g$ ). Not surprisingly, the *oracle* strategy, which has access to the hidden parameter  $p$ , always performs better than the other ones, copying the *timid* strategy when  $p < 0.5$  and the *bold* strategy when  $p \geq 0.5$ . As explained in Section 3.3, that is the theoretically best possible expected performance.

In terms of preferability, we say that a strategy  $\pi'$  is *preferred* over strategy  $\pi$  if its utility is greater, given  $b_0$  and  $g$ :

$$\pi' \succ \pi \mid b_0, g \iff \xi_{\pi'}(b_0, g) > \xi_\pi(b_0, g), \quad (55)$$

and we say that a strategy  $\pi'$  *dominates* a strategy  $\pi$  if it is never worse, and is strictly preferable at least in one configuration:

$$\pi' \succcurlyeq \pi \iff \begin{cases} \forall b_0, g \in \mathbb{N}^+ : \pi' \succeq \pi \mid b_0, g, \\ \exists b'_0, g' \in \mathbb{N}^+ : \pi' \succ \pi \mid b'_0, g'. \end{cases} \quad (56)$$

Figure 8 suggests that the *bold* strategy outperforms the *timid* strategy in expectation when the initial budget is low, closer to 0 than  $g$ . However, *timid*

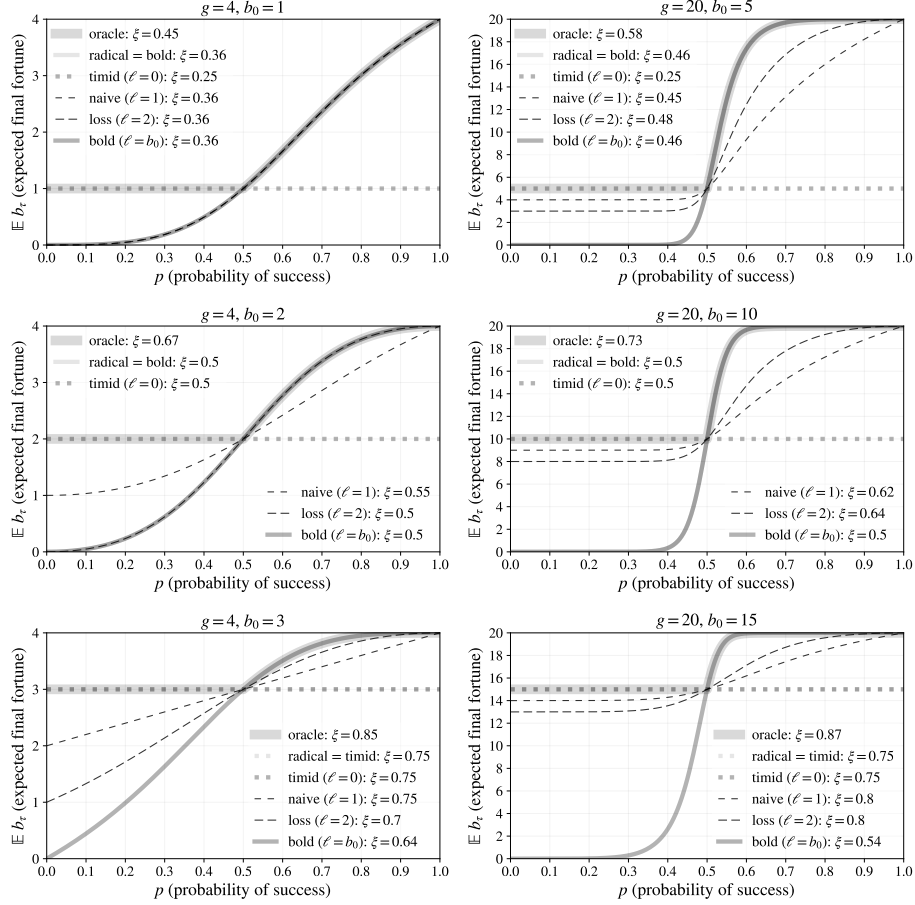


Figure 7: Expected final budget for *bold*, *timid*, *radical*, *naive*, *maximal loss* with  $\ell = 2$ , and *oracle* strategies. The graphics in left column present a case in which the objective budget is low ( $g = 4$ ), in contrast to the right column, where it is higher ( $g = 20$ ). The graphics on the top present a case in which the initial budget is low, compared to the objective budget ( $b_0 = 1/4g$ ). In the middle row, the initial budget is half of the objective ( $b_0 = 1/2g$ ). In the bottom row, the initial budget is close to the objective ( $b_0 = 3/4g$ ). The utility ( $\xi$ ) of each strategy (the area under the corresponding curve) is also indicated.

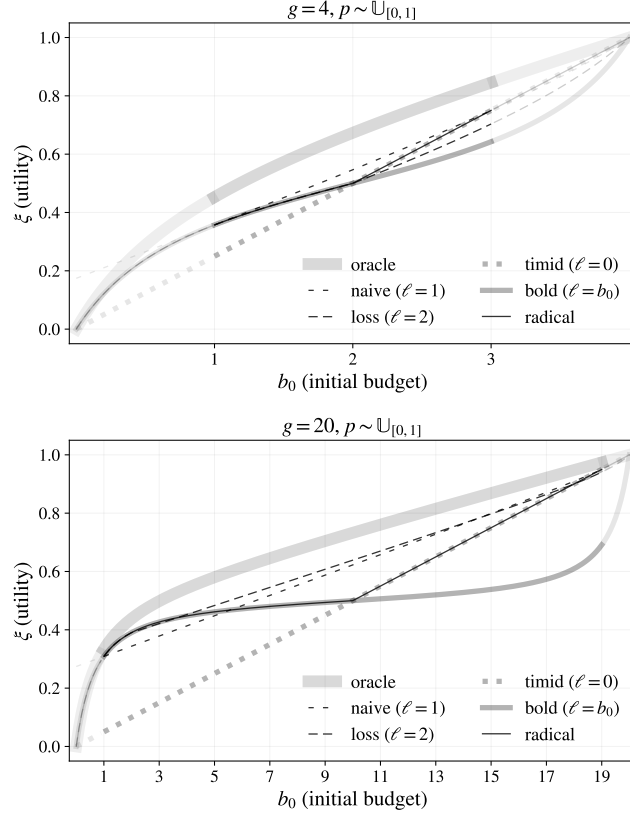


Figure 8: Expected final budget for the analyzed strategies (*bold*, *timid*, *naive*, *radical*, *loss* with  $\ell = 2$ , and *oracle*), considering a uniform prior distribution for  $p$ . The curves correspond to the definite integral with respect to  $p$  between 0 and 1 of the expected utility of each strategy, depending on  $g$ , and on the ratio  $b_0/g$ . The graphic on the top presents a case in which the objective budget is low ( $g = 4$ ), and the graphic on the bottom presents a case in which it is higher ( $g = 20$ ). Note the superior performance of the *bold* strategy in relation to *naive* when the objective budget  $g$  is high, and the ratio  $b_0/g$  is low.



175 outperforms *bold* when the initial budget is high, closer to  $g$  than 0. Both strategies are equivalent when  $b_0 = g/2$ . Again, this is not surprising: when the initial budget is low, the gambler has more to gain than to lose, then it is preferable to take the risk of playing, and vice versa. It leads to the hypothesis described in Theorem 2:

**Theorem 2.** *In a DGP, the relation of preferability between bold and timid strategies is determined by the ratio of  $b_0$  to  $g$ , such that:*

$$\forall b_0, g \in \mathbb{N}^+ : \begin{cases} \pi_{bold} \succ \pi_{timid} & \iff g > 2b_0, \\ \pi_{bold} \sim \pi_{timid} & \iff g = 2b_0, \\ \pi_{bold} \prec \pi_{timid} & \iff g < 2b_0, \end{cases} \quad (57)$$

which means, according to Eqs. (55) and (57):

$$\forall b_0, g \in \mathbb{N}^+ : \begin{cases} g > 2b_0 & \iff \xi_{bold}(b_0, g) > \xi_{timid}(b_0), \\ g = 2b_0 & \iff \xi_{bold}(b_0, g) = \xi_{timid}(b_0), \\ g < 2b_0 & \iff \xi_{bold}(b_0, g) < \xi_{timid}(b_0). \end{cases} \quad (58)$$

For proving Theorem 2, and consequently the proposition made in Eqs. (57) and (58), we must recall the utility formulas concerning *bold* and *timid* strategies, defined by Eqs. (23) and (29):

$$\xi_{bold}(b_0, g) = \int_0^1 g \cdot \frac{1 - \left(\frac{1-p}{p}\right)^{b_0}}{1 - \left(\frac{1-p}{p}\right)^g} dp \quad ; \quad \xi_{timid}(b_0) = b_0. \quad (59)$$

180 The solution for that definite integral cannot be expressed with a simple closed formula, but a proof for the inequalities presented in Eq. (58) can be done based on Lemmas 2.1 and 2.2, which follow well-known definite integral properties [74, 75, 76]:

**Lemma 2.1.** *The definite integral of a given integrable function  $f$  between 0 and 1*  
 185 *is equivalent to the definite integral of  $f'$  between 0 and  $1/2$  if  $f'(p) = f(p) + f(1-p)$  for any value of  $p$  within that interval.*

*Proof of Lemma 2.1.*

$$\begin{aligned}
\int_0^1 f(p) \, dp &= \int_0^{\frac{1}{2}} f(p) \, dp + \int_{\frac{1}{2}}^1 f(p) \, dp \\
&= \int_0^{\frac{1}{2}} f(p) \, dp - \int_1^{\frac{1}{2}} f(p) \, dp \\
&= \int_0^{\frac{1}{2}} f(p) \, dp + \int_0^{\frac{1}{2}} f(1-p) \, dp \\
&= \int_0^{\frac{1}{2}} f(p) + f(1-p) \, dp.
\end{aligned} \tag{60}$$

□

**Lemma 2.2.** *Let  $a$  and  $b$  be real values such that  $a < b$ , and  $f$  and  $g$  be functions, both integrable on  $[a, b]$ . If  $f(x) > g(x)$  for any  $x$  between  $a$  and  $b$ , then  $\int_a^b f(x) \, dx > \int_a^b g(x) \, dx$ .*

190

*Proof of Lemma 2.2.*

$$\begin{aligned}
&\forall a, b, x \in \mathbb{R} \mid a \leq x \leq b \quad : \quad f(x) > g(x) \\
\Rightarrow &f(x) - g(x) > 0 \\
\Rightarrow &\int_a^b f(x) - g(x) \, dx > 0 \\
\Rightarrow &\int_a^b f(x) \, dx - \int_a^b g(x) \, dx > 0 \\
\Rightarrow &\int_a^b f(x) \, dx > \int_a^b g(x) \, dx.
\end{aligned} \tag{61}$$

□

**Corollary 2.2.1.** *Let  $f'(p) = f(p) + f(1-p)$  be a function, integrable between 0 and 1/2, and  $b_0$  be a constant. If  $f'(p) > 2b_0$  for any  $p$  between 0 and 1/2, then  $\int_0^{1/2} f'(p) \, dp > b_0$ .*

*Proof of Corollary 2.2.1.* Replacing the terms in Lemma 2.2, we obtain:

$$\begin{aligned}
& \forall p \in [0, 1/2] : f(p) + f(1-p) > 2b_0 \\
\Rightarrow & \int_0^{\frac{1}{2}} f(p) + f(1-p) \, dp > \int_0^{\frac{1}{2}} 2b_0 \, dp \\
\Rightarrow & \int_0^{\frac{1}{2}} f(p) + f(1-p) \, dp > b_0 \\
\Rightarrow & \int_0^1 f(p) \, dp > b_0.
\end{aligned} \tag{62}$$

195

□

*Proof of Theorem 2.* Let  $f$  be the expected final budget using the bold strategy (Eq. (22)):

$$f(p) := \mathbb{E}[B_\tau \mid b_0, g, p, \pi_{bold}] = g \frac{1 - r^{b_0}}{1 - r^g}, \tag{63}$$

using the ratio  $r = \frac{1-p}{p}$  for readability (Eq. (2)). The sum of  $f(p)$  with  $f(1-p)$  can be re-written as follows:

$$\begin{aligned}
f(p) + f(1-p) &= g \left( \frac{1 - r^{b_0}}{1 - r^g} \right) + g \left( \frac{1 - \left(\frac{1}{r}\right)^{b_0}}{1 - \left(\frac{1}{r}\right)^g} \right) \\
&= g \left( \frac{1 - r^{b_0}}{1 - r^g} + \frac{1 - r^{-b_0}}{1 - r^{-g}} \right) \\
&= g \left( \frac{1 - r^{b_0}}{1 - r^g} + r^{g-b_0} \frac{1 - r^{b_0}}{1 - r^g} \right) \\
&= g (1 + r^{g-b_0}) \frac{1 - r^{b_0}}{1 - r^g} \\
&= g \frac{(1 + r^{g-b_0})(1 - r^{b_0})}{1 - r^g} \\
&= g \frac{-r^g + r^{g-b_0} - r^{b_0} + 1}{1 - r^g}.
\end{aligned} \tag{64}$$

Observe that:

$$0 < p < 1/2 \iff 1 < r < \infty, \tag{65}$$

and:

$$\forall r \in \mathbb{R} \mid r > 1 : \begin{cases} r^{g-b_0} > r^{b_0} & \iff g - b_0 > b_0 & \iff g > 2b_0, \\ r^{g-b_0} = r^{b_0} & \iff g - b_0 = b_0 & \iff g = 2b_0, \\ r^{g-b_0} < r^{b_0} & \iff g - b_0 < b_0 & \iff g < 2b_0. \end{cases} \tag{66}$$

Manipulating the first inequality of the first row presented in Eq. (66), we obtain:

$$\begin{aligned}
\forall r \in \mathbb{R} \mid r > 1 : \quad & r^{g-b_0} > r^{b_0} \\
\iff & r^{g-b_0} - r^{b_0} > 0 \\
\iff & 1 - r^g + r^{g-b_0} - r^{b_0} > 1 - r^g \\
\iff & \frac{1 - r^g + r^{g-b_0} - r^{b_0}}{1 - r^g} > \frac{1 - r^g}{1 - r^g} \quad (67) \\
\iff & \frac{1 - r^g + r^{g-b_0} - r^{b_0}}{1 - r^g} > 1 \\
\iff & g \frac{-r^g + r^{g-b_0} - r^{b_0} + 1}{1 - r^g} > g.
\end{aligned}$$

Using similar algebraic reasoning on the other two rows, we can obtain:

$$\forall r \in \mathbb{R} \mid r > 1 : r^{g-b_0} = r^{b_0} \iff g \frac{-r^g + r^{g-b_0} - r^{b_0} + 1}{1 - r^g} = g, \quad (68)$$

and:

$$\forall r \in \mathbb{R} \mid r > 1 : r^{g-b_0} < r^{b_0} \iff g \frac{-r^g + r^{g-b_0} - r^{b_0} + 1}{1 - r^g} < g. \quad (69)$$

Therefore, combining Eqs. (66)–(69), we verify:

$$\forall p \in [0, 1/2] : \begin{cases} f(p) + f(1-p) > g \iff g > 2b_0, \\ f(p) + f(1-p) = g \iff g = 2b_0, \\ f(p) + f(1-p) < g \iff g < 2b_0, \end{cases} \quad (70)$$

allowing to infer that:

$$\forall p \in [0, 1/2], r = \frac{1-p}{p} : \begin{cases} g \frac{1-r^g+r^{g-b_0}-r^{b_0}}{1-r^g} > 2b_0 \implies \int_0^1 g \frac{1-r^{b_0}}{1-r^g} dp > b_0, \\ g \frac{1-r^g+r^{g-b_0}-r^{b_0}}{1-r^g} = 2b_0 \implies \int_0^1 g \frac{1-r^{b_0}}{1-r^g} dp = b_0, \\ g \frac{1-r^g+r^{g-b_0}-r^{b_0}}{1-r^g} < 2b_0 \implies \int_0^1 g \frac{1-r^{b_0}}{1-r^g} dp < b_0. \end{cases} \quad (71)$$

□

The *radical* strategy appears as a combination of *timid* and *bold*, taking the best of both strategies given the  $b_0/g$  ratio (Eq. (58)). For this reason, *radical* should dominate both *timid* ( $\pi_{\text{radical}} \succ \pi_{\text{timid}}$ ) and *bold* ( $\pi_{\text{radical}} \succ \pi_{\text{bold}}$ ).

200 The *naive* strategy does not take into account the  $b_0/g$  ratio, but estimates  $p$  by observing the past results. A *naive* gambler quits the game immediately if the estimated parameter  $\hat{p}$  falls under  $1/2$  (i.e. at any round  $t$  when  $B_t < b_0$ ). Figure 8 suggests that *naive* presents a better (or at least equal) utility than *timid* in any scenario (but a formal proof of dominance is still to be done).  
205 It also appears that *naive* is superior to *bold* in most cases, except when the objective budget  $g$  is relatively high (i.e. when the magnitude of the goal budget is significantly bigger than the single unit wagered each round), and the gambler's initial budget  $b_0$  is relatively low (i.e. closer to 0 than  $g$ ). In this case, the agent has more to gain than to lose, and the *naive* strategy (even if better than  
210 *timid*) still induces a too conservative behavior. The exact point of intersection, from where *naive* becomes preferable than *bold*, depends on  $g$ , as can be seen in Figure 8.

The *loss* strategy generalizes the *naive* strategy allowing to modify  $\ell$ , the “tolerable loss”. It can be observed in Figure 8 that each different value of  $\ell$  changes the point where *loss* “takes off” from *bold*, like a “tensioned rope”. There is an optimal parameter  $\ell^*$  that maximizes the utility  $\xi$  of that strategy, depending on  $b_0$  and  $g$ :

$$\ell^* = \arg \max_{i=0}^{b_0} [\xi_{\ell^*}(b_0, g) \geq \xi_{\ell=i}(b_0, g) \mid b_0, g \in \mathbb{N}^+]. \quad (72)$$

Table 1 presents the optimal value of  $\ell$  for all possible scenarios in which  $g \leq 30$ . The regularity of the values within that table suggests, empirically, that  $\ell^*$  could be obtained by the following rule:

$$\ell^* \mid g, b_0 = \max \left( 1, \min \left( b_0, \left\lfloor \frac{-1 + \sqrt{2g - 2b_0 - 1}}{2} \right\rfloor + 1 \right) \right). \quad (73)$$

Among the heuristic methods analyzed in this article, the *loss* strategy with  $\ell^*$  is the one that presents the best performance for the analyzed scenarios, which  
215 allows to hypothesize that it dominates the other strategies. However, formally proving the dominance of the  $\ell^*$  strategy over the other ones is still to be done.

		$\{\ell^* \mid b_0, g\}$																												
$g \backslash b_0$		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
2		1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
3		1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
4		1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
5		1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
6		1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
7		1	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
8		1	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
9		1	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
10		1	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
11		1	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
12		1	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
13		1	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
14		1	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
15		1	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
16		1	2	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
17		1	2	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-
18		1	2	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-	-
19		1	2	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-	-
20		1	2	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-	-
21		1	2	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-	-
22		1	2	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-	-
23		1	2	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-	-
24		1	2	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-	-
25		1	2	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-	-
26		1	2	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-	-
27		1	2	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-	-
28		1	2	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-	-
29		1	2	3	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1	-
30		1	2	3	4	3	3	3	3	3	3	3	3	3	3	3	3	2	2	2	2	2	2	2	2	1	1	1	1	1

Table 1: The optimal maximal tolerable loss parameter  $\ell^*$  given  $b_0$  and  $g$ , considering all possible scenarios with  $g \leq 30$ . To construct this table, the utility of the *loss* strategy has been calculated exhaustively for each scenario and for all possible values of  $\ell$ , from 0 (*timid*), then 1 (*naive*) to  $b_0$  (*bold*). The values of  $\ell$  having highest utility are presented. The regularity of the values inside this triangular matrix suggests that the optimal maximal tolerable loss parameter  $\ell^*$  can be calculated analytically.

## 5. Conclusion

This article contributes to the state-of-the-art by presenting an original version of the classical gambler’s ruin game, called decisional gambling process, in which the gambler can decide to stop the game, but does not know the parameter  $p$  in advance. An *oracle* policy is defined, allowing to indicate the theoretically best possible performance of a gambler if the value of the underlying parameter could be known.

Some heuristic strategies are proposed and formally evaluated. The *naive* strategy (which simply bases its decision on the estimation of  $p$ ) is often better than *timid*, *bold*, and *radical*. However, *naive* can be surpassed by *bold* and *radical* when the ratio  $b_0/g$  is low (i.e. when the gambler does not have much to lose, compared with the potential earnings). The exact point from where those strategies behave better than *naive* depends on  $g$ .

We finally propose a generalized method, the *loss* strategy, which is able to reproduce the behavior of the previously analyzed strategies, depending on how a parameter  $\ell$  is tuned. That parameter specifies how much the agent is willing to lose in relation to its initial budget (similar to the “stop-loss” order in the stock market [77]). When  $\ell = 0$ , the player does not accept any risk of decreasing the budget, quitting the game immediately, which corresponds to the *timid* strategy. When  $\ell = 1$ , the gambler quits the game if the budget falls to  $b_0 - 1$ , which corresponds to the *naive* strategy, as demonstrated. Other integer values of  $\ell$  can be defined between 0 and  $b_0$ . When  $\ell = b_0$ , the agent accepts to play until being ruined, which corresponds to the *bold* strategy.

For each combination of  $b_0$  and  $g$ , an optimal parameter  $\ell^*$  can be deduced in order to maximize the utility of the strategy. The use of the *loss* strategy tuned with  $\ell^*$  proved to be the best strategy among the described ones in the analyzed scenarios. In this article, however, we do not demonstrate that this method is optimal. In fact, finding an optimal method for DGPs remains an open problem. In particular, it could be interesting to investigate whether considering the confidence on the estimation (which depends on the number of observations)

during the decision making can lead to some utility improvement.

Future work includes considering the case in which the objective budget can be infinite ( $g \rightarrow \infty$ ), and finding a policy which can be formally proved to be optimal. Another possibility consists in extending the analysis made in this article to the corresponding *multiarmed bandit* problem, where the decision is not related to stay or quit, but to the choice of playing, at each round  $t$ , a specific game  $i$  among  $n$  different games.

## References

- [1] W. Feller, An Introduction to Probability Theory and Its Applications, Wiley, New York, NY, 1966.
- [2] E. Shoesmith, Huygens' solution to the gambler's ruin problem, *Historia Math.* 13 (2) (1986) 157–164.
- [3] A. W. F. Edwards, Pascal's problem: The 'gambler's ruin', *International Statistical Review* 51 (1) (1983) 73–79.
- [4] M. Kraitchik, *Mathematical Recreations*, W. W. Norton, New York, 1942, Ch. The Gambler's Ruin., pp. 140–140.
- [5] E. W. Weisstein, Gambler's ruin (2002).
- [6] G. Shafer, V. Vovk, *Game-Theoretic Foundations for Probability and Finance*, Probability and Statistics, Wiley, 2019.
- [7] D. Stirzaker, *Elementary Probability*, 2nd Edition, Cambridge University Press, 2003.
- [8] A. Hald, *A History of Probability and Statistics and their Applications before 1750*, John Wiley, New York, 1990.
- [9] S. Song, J. Song, A note on the history of the gambler's ruin problem, *Communications for Statistical Applications and Methods* 20 (2) (2013) 157–168.



- [10] P. Gorroochurn, *Huygens and the Gambler's Ruin (1657)*, Wiley, 2012, Ch. 5, pp. 39–48.
- 275 [11] G. Orosi, Linear algebra-based solution of the gambler's ruin problem, *International Journal of Mathematical Education in Science and Technology* 48 (1) (2017) 107–111.
- [12] G. J. Morrow, Laws relating runs and steps in gambler's ruin, *Stochastic Processes and their Applications* 125 (5) (2015) 2010–2025.
- 280 [13] G. J. Morrow, Laws relating runs, long runs, and steps in gambler's ruin, with persistence in two strata, in: G. E. Andrews, C. Krattenthaler, A. Krinik (Eds.), *Lattice Path Combinatorics and Applications*, Springer, Cham, 2019, pp. 343–381.
- [14] D. S. Grebenkov, D. Holcman, R. Metzler, Preface: new trends in first-  
285 passage methods and applications in the life sciences and engineering, *Journal of Physics A: Mathematical and Theoretical* 53 (19) (2020) 190301.
- [15] M. El-Shehawey, On the gambler's ruin problem for a finite markov chain, *Statistics & Probability Letters* 79 (14) (2009) 1590 – 1595.
- [16] S. N. Ethier, D. Khoshnevisan, Bounds on gambler's ruin probabilities in  
290 terms of moments, *Methodology And Computing In Applied Probability* 4 (1) (2002) 55–68.
- [17] W. Hürlimann, Improved analytical bounds for gambler's ruin probabilities, *Methodol. Comput. Appl. Probab.* 7 (2005) 79–95.
- [18] G. Katriel, Gambler's ruin probability – a general formula, *Statistics &*  
295 *Probability Letters* 83 (10) (2013) 2205 – 2210.
- [19] R. M. Canjar, Gambler's ruin revisited: The effects of skew and large jackpots, in: S. N. Ethier, W. R. Eadington (Eds.), *Optimal Play: Mathematical Studies of Games and Gambling*, University of Nevada, 2007, pp. 439–470.

- [20] J. Anděl, Šárka Hudecová, Variance of the game duration in the gambler’s ruin problem, *Statistics & Probability Letters* 82 (9) (2012) 1750 – 1754.
- [21] J. W. Wilcox, The gambler’s ruin approach to business risk, *Sloan Management Review* 18 (1) (1976) 33–46.
- [22] A. Coad, J. Frankish, R. G. Roberts, D. J. Storey, Growth paths and survival chances: An application of gambler’s ruin theory, *Journal of Business Venturing* 28 (5) (2013) 615–632.
- [23] R. Jayasekera, Prediction of company failure: Past, present and promising directions for the future, *International Review of Financial Analysis* 55 (2017) 196–208.
- [24] C. F. Lardizabal, R. R. Souza, Open quantum random walks: Ergodicity, hitting times, gambler’s ruin and potential theory, *Journal of Statistical Physics* 164 (2016) 1122–1156.
- [25] F. Debbasch, The gambler’s ruin problem and quantum measurement, in: G. Di Molfetta, V. Kendon, Y. Shikano (Eds.), *Proceedings 9th International Conference on Quantum Simulation and Quantum Walks*, Marseille, France, 20-24/01/2020, Vol. 315 of *Electronic Proceedings in Theoretical Computer Science*, Open Publishing Association, 2020, pp. 100–111.
- [26] P. Lorek, M. Słowik, F. Zagórski, Statistical testing of prng: Generalized gambler’s ruin problem, in: J. Blömer, I. S. Kotsireas, T. Kutsia, D. E. Simos (Eds.), *Mathematical Aspects of Computer and Information Sciences*, Springer, Cham, 2017, pp. 425–437.
- [27] C. M. Guttman, E. A. DiMarzio, J. D. Hoffman, Modelling the amorphous phase and the fold surface of a semicrystalline polymer—the gambler’s ruin method, *Polymer* 22 (11) (1981) 1466–1479.
- [28] C. M. Guttman, E. A. DiMarzio, Rotational isomeric modeling of a polyethylene-like polymer between two plates: connection to “gambler’s ruin” problem, *Macromolecules* 15 (2) (1982) 525–531.

- [29] X. Z. Cheng, M. B. A. Jalil, H. K. Lee, Analytical solution to transport in brownian ratchets via the gambler’s ruin model, *Phys. Rev. Lett.* 99 (2007) 070601.
- 330 [30] P. Grassberger, Partons, the gambler’s ruin, and pomerons, *Nuclear Physics B* 125 (1) (1977) 83 – 107.
- [31] D. J. Bicut, T. W. Burkhardt, Absorption of a randomly accelerated particle: gambler’s ruin in a different game, *Journal of Physics A: Mathematical and General* 33 (39) (2000) 6835–6841.
- 335 [32] E. H. Bulte, G. van Kooten, Harvesting and conserving a species when numbers are low: population viability and gambler’s ruin in bioeconomic models, *Ecological Economics* 37 (1) (2001) 87 – 100.
- [33] S. Petteruti, Fast simulation of random walks on the endoplasmic reticulum by reduction to the gambler’s ruin problem, Ph.D. thesis, Harvard University,  
340 Cambridge, MA (2017).
- [34] A. T. Coram, Social class and luck: Some lessons from gambler’s ruin and branching processes, *Political Studies* 45 (1) (1997) 66–77.
- [35] N. Akbarzadeh, C. Tekin, Gambler’s ruin bandit problem, in: 54th Annual Allerton Conference on Communication, Control, and Computing, Allerton 2016, Monticello, IL, USA, September 27-30, 2016, 2016, pp. 1236–1243.  
345
- [36] F. S. Perotto, M. Bourgain, B. C. Silva, L. Vercouter, Open problem: Risk of ruin in multiarmed bandits, in: *Conference on Learning Theory, COLT 2019*, 25-28 June 2019, Phoenix, AZ, USA, 2019, pp. 3194–3197.
- [37] F. S. Perotto, S. Vakili, P. Gajane, Y. Faghan, M. Bourgain, Gambler bandits and the regret of being ruined, in: 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), AAMAS ’21, International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS), Richland, SC, 2021, p. 1664–1667.  
350

- 355 [38] T. Lattimore, C. Szepesvári, *Bandit Algorithms*, Cambridge University Press, 2020.
- [39] A. G. Munford, S. M. Lewis, A note on gambler’s ruin against an infinitely rich adversary, *International Journal of Mathematical Education in Science and Technology* 12 (2) (1981) 165–168.
- 360 [40] S. N. Ethier, Gambler’s ruin, in: *The Doctrine of Chances, Probability and its Applications*, Springer, 2010, pp. 241–274.
- [41] J. Bak, The anxious gambler’s ruin, *Mathematics Magazine* 74 (3) (2001) 182–193.
- [42] G. Katriel, Gambler’s ruin: The duration of play, *Stochastic Models* 30 (3) (2014) 251–271.
- 365 [43] T. Tran Cong, M. Sato, One-dimensional random walk with unequal step lengths restricted by an absorbing barrier, *Discrete Mathematics* 40 (2) (1982) 153 – 162.
- [44] R. Pelessoni, P. Vicig, M. Zaffalon, Inference and risk measurement with the pari-mutuel model, *International Journal of Approximate Reasoning* 51 (9) (2010) 1145 – 1158, imprecise probability in statistical inference and decision making.
- 370 [45] J. D. Harper, K. A. Ross, Stopping strategies and gambler’s ruin, *Mathematics Magazine* 78 (4) (2005) 255–268. doi:10.1080/0025570X.2005.11953340.
- 375 [46] P. Lorek, P. Markowski, Conditional gambler’s ruin problem with arbitrary winning and losing probabilities with applications, CoRR abs/1812.00687. arXiv:1812.00687.
- [47] B. Hajek, Gambler’s ruin: A random walk on the simplex, in: T. M. Cover, B. Gopinath (Eds.), *Open Problems in Communication and Computation*, Springer, New York, 1987, pp. 204–207.
- 380

- [48] Y. C. Swan, F. T. Bruss, A matrix-analytic approach to the  $n$ -player ruin problem, *Journal of Applied Probability* 43 (3) (2006) 755–766.
- [49] A. Kmet, M. Petkovšek, Gambler’s ruin problem in several dimensions,  
385 *Advances in Applied Mathematics* 28 (2) (2002) 107 – 118.
- [50] A. L. Rocha, F. Stern, The gambler’s ruin problem with  $n$ -players and asymmetric play, *Statistics & Probability Letters* 44 (1) (1999) 87 – 95.
- [51] A. Rocha, F. Stern, The asymmetric  $n$ -player gambler’s ruin problem with equal initial fortunes, *Advances in Applied Math.* 33 (3) (2004) 512 – 530.
- 390 [52] P. Lorek, Generalized gambler’s ruin problem: Explicit formulas via siegmund duality, *Methodology and Computing in Applied Probability* 19 (2016) 603–613.
- [53] T. Lengyel, The conditional gambler’s ruin problem with ties allowed, *Applied Mathematics Letters* 22 (3) (2009) 351 – 355.
- 395 [54] B. Hunter, A. Krinik, C. Nguyen, J. Switkes, H. von Bremen, Gambler’s ruin with catastrophes and windfalls, *J. Stat. Theory Pract.* 2 (2008) 199–219.
- [55] T. van Uem, Maximum and minimum of modified gambler’s ruin problem, *CoRR* abs/1301.2702. [arXiv:1301.2702](#).
- [56] T. van Uem, Stopping strategies and gambler’s ruin problem, *CoRR*  
400 abs/1302.7247. [arXiv:1302.7247](#).
- [57] T. Imai, T. Ohira, Delayed gambler’s ruin, in: Q. Gao, H. R. Karimi (Eds.), *Stability, Control and Application of Time-delay Systems*, Butterworth-Heinemann, 2019, Ch. 16, pp. 333 – 347.
- [58] M. L. Mansfield, A continuum gambler’s ruin model, *Macromolecules* 21 (1)  
405 (1988) 126–130.
- [59] R. K. Ankomah, R. Oduro, E. K. Amoah, Gambler’s risk of ruin and optimal bet, *Communications in Mathematical Finance* 9 (2020) 13–27.

- [60] A. N. Shiryaev, Optimal stopping rules, in: International Encyclopedia of Statistical Science, Springer, 2011, pp. 1032–1034.
- 410 [61] G. Edgar, L. Sucheston, Stopping Times and Directed Processes, Vol. 42 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1992.
- [62] O. C. Ibe, Chapter 12 - special random processes, in: O. C. Ibe (Ed.), Fundamentals of Applied Probability and Random Processes (Second Edition), second edition Edition, Academic Press, Boston, 2014, pp. 369 – 425.
- 415 [63] K. Sigman, Gambler’s Ruin Problem, Vol. 1, Wiley, 2009, Ch. 2, pp. 2:1–2:6.
- URL <http://www.columbia.edu/~ks20/stochastic-I/stochastic-I-GRP.pdf>
- 420 [64] J. M. Steele, Stochastic Calculus and Financial Applications, Springer-Verlag, 2001.
- [65] K. P. Murphy, Machine Learning: A Probabilistic Perspective, The MIT Press, 2012.
- [66] P. Gagniuc, Markov Chains: From Theory to Implementation and Experimentation, Wiley, 2017.
- 425 [67] S. Ross, Applied Probability Models with Optimization Applications, Holden-Day, 1970.
- [68] J. Jacod, P. Protter, Martingales, in: Probability Essentials, Universitext, Springer, Berlin, Heidelberg, 2000, Ch. 24, pp. 207–213.
- 430 [69] J. Jacod, P. Protter, Supermartingales and submartingales, in: Probability Essentials, Universitext, Springer, Berlin, Heidelberg, 2000, Ch. 25, pp. 215–218.
- [70] L. M. Łudański, Binomial distribution, in: Between Certainty and Uncertainty: Statistics and Probability in Five Units with Notes on Historical

- 435      Origins and Illustrative Numerical Examples, Springer, Berlin, Heidelberg,  
2013, pp. 87–127.
- [71] T. Yaacoub, G. V. Moustakides, Y. Mei, Optimal stopping for interval  
estimation in bernoulli trials, *IEEE Transactions on Information Theory*  
65 (5) (2019) 3022–3033.
- 440 [72] H. Harney, Bayesian Inference: Data Evaluation and Decisions, Springer,  
2016.
- [73] H. Liu, L. Wasserman, Bayesian Inference, Carnegie Mellon, 2014, Ch. 12,  
pp. 299–351.
- [74] T. M. Apostol, The concepts of integral calculus, in: *Calculus: One-Variable*  
445 *Calculus with an Introduction to Linear Algebra*, 2nd Edition, Vol. 1, Wiley,  
1967, Ch. 1.
- [75] C. T. Moss, C. Plumpton, Properties of definite integrals, Macmillan,  
London, 1983, Ch. 9, pp. 58–62. doi:10.1007/978-1-349-06045-0\_9.
- [76] D. Hughes-Hallett, A. M. Gleason, W. G. McCallum, Theorems about  
450 definite integrals, in: *Calculus: single variable*, 8th Edition, Wiley, 2020,  
Ch. 5.4, pp. 319–332.
- [77] T. onn Talpsepp, T. Vaarmets, The disposition effect, performance, stop  
loss orders and education, *Journal of Behavioral and Experimental Finance*  
24 (2019) 100240.