Chapter III: Games of Pure Chance Generated by Gambler's Ruin with Unlimited Credit: The Fuss-Catalan case

In Chapter II, we only allowed positive steps. Now we will also allow negative steps, and treat games that may be viewed as a "gambler's ruin with infinite credit", with *arbitrary* 'die'. In the next chapter we will treat the case of a *general* die, while in this chapter we only consider two-faced dice, where one of the faces is marked 1 and the other markded $-k$, and the more difficult case where one of the faces is marked -1 and the other marked k. We will start with their intersection, the very classical case of $\{-1, 1\}$, treated in Feller's classic [F]. However, even in this case we will be able to go beyond Feller, since he did not use a computer.

The general set-up, to be considered in full generality in Chapter IV is as follows.

On the discrete line, you start at the origin $x = 0$, and there is a fixed allowed set of steps consisting of both positive and negative integers and a probability distribution on them, let's call it $\mathcal{P}$. You are allowed to go as far left as possible (i.e. you can owe as much as necessary). At each round, you roll the $\mathcal{P}$ die, and move accordingly. You win as soon as you reach a location ≥ 1, or more generally when you reach a location $\geq n$. In other words, your goal is to exit the casino with at least one dollar (or more generally, at least n dollars). In the two-player (or multi-player) version, the players take turns rolling the $\mathcal{P}$ die, and whoever achieves the goal first is declared the winner. As before, we will first discuss the solitaire game, where the goal is to reach it as soon as possible.

Classical Gambling: Winning a dollar or losing a dollar

Let's start with the simplest, most classical case, of simple random walk, where you start with 0 dollars, and at each round you win a dollar with probability p and lose a dollar with probability $1 - p$. The expected gain at each individual round is $p \cdot 1 + (1 - p) \cdot (-1) = 2p - 1$, so if $p > \frac{1}{2}$, then sooner or later you will reach your goal. If $p < \frac{1}{2}$, then you may never make it, sliding down to infinite debt. In the border-line case of a **fair** coin, $p = \frac{1}{2}$, as we will soon see, you are also guaranteed to 'eventually' be in possession of 1 dollar (and more generally, n dollars for each $n > 0$, as big as you wish). Alas, as we will also soon see, the expected time until that happens is infinite, and since life is finite, there is a good chance that when you will pass away, your heirs will have a huge debt.

Analyzing Gambling histories

For typographical clarity, let's denote -1 by $\overline{1}$.

Our **alphabet** is $\{-1, 1\}$=$\{\overline{1}, 1\}$. A 'gambling history' consists of a **word** that ends in 1, whose sum is 1, and whose proper partial sums are all non-positive. Obviously the length of such a game is odd.

If you are really lucky, you exit after one step, since you won a dollar right away.

If you lost a dollar at the first round, you can recover at the second round, and then win a dollar at the third round. Etc.

For the sake of clarity and concreteness, let's list the first few 'histories'.

Length 1: $\{1\}$. Probability $= p$.

Length 3: $\{\bar{1}\,1\,1\}$. Probability $= p^2\,(1-p)$.

Length 5: $\{\bar{1}\,1\,\bar{1}\,1\,1 \quad , \quad \bar{1}\,\bar{1}\,1\,1\,1\}$. Probability $2 \cdot p^3\,(1-p)^2$.

Length 7:

$$\{\bar{1}\,1\,\bar{1}\,1\,\bar{1}\,1\,1 \quad , \quad \bar{1}\,1\,\bar{1}\,\bar{1}\,1\,1\,1 \quad , \quad \bar{1}\,\bar{1}\,1\,1\,\bar{1}\,1\,1 \quad , \quad \bar{1}\,\bar{1}\,1\,\bar{1}\,1\,1\,1 \quad , \quad \bar{1}\,\bar{1}\,\bar{1}\,1\,1\,1\,1\} \quad ,$$

with probability $5 \cdot p^3\,(1-p)^2$.

It is useful, for humans, to visualize such a history as a lattice path in the discrete plane starting at $(0,0)$ where $\bar{1}$ corresponds to a step $(1,-1)$ and 1 corresponds to a step $(1,1)$. For example, the word (gambling history)

$$\bar{1}\,\bar{1}\,1\,1\,\bar{1}\,1\,1 \quad ,$$

corresponds to the walk

$$(0,0) \to (1,-1) \to (2,-2) \to (3,-1) \to (4,0) \to (5,-1) \to (6,0) \to (7,1) \quad .$$

Let's study the *anatomy* of such histories, or equivalently, paths . Obviously they are all of odd length, and they all end with 1. So we can write, for *any* history W

$$W = U\,1 \quad ,$$

where U is a word that sums to 0, all whose partial sums are non-positive. Such words are called *Dyck words.*

Let's analyze such a Dyck word U or rather its corresponding path from $(0,0)$ to $(2n,0)$, say. Of course, it may be the empty word, but if it is not, let $(2r,0)$ $0 < r \leq n$ be the **first** time that it hits the x-axis. Then we can write

$$U = U_1\,U_2 \quad ,$$

where U_2 is another word of that kind (of length $2n-2r$), but U_1, consisting of the first $2r$ letters of U, has the special property that all its partial sums (except the 0-th and the last) are **strictly negative**, or in terms of its path, except for its starting and ending points, they lie **strictly** below the x-axis. Such a word must **necessarily** start with a $\bar{1}$ and end with a 1, and may be written as $\bar{1}U_3 1$, where U_3 is an arbitrary Dyck word. Conversely, for any Dyck word U_3, $\bar{1}\,U_3\,1$ corresponds to such a 'strictly below the x-axis' path. So we have the (context-free) *grammar*

$$U = EmptyWord \vee \bar{1}\,U\,1U \quad , \qquad (DyckGrammar)$$

where now U stands for 'an arbitrary Dyck word'.

let x_1 and x_{-1} be **commuting** variables.

For any word $u = u_1 \dots u_m$, let the *weight* of u be $x_{u_1} \cdots x_{u_m}$. For example,

$$weight(\overline{1}\,\overline{1}\,\overline{1}\,1\,1\,1\,\overline{1}\,1\,1) = x_{-1}x_{-1}x_{-1}x_1x_1x_1x_{-1}x_1x_1 \;=\; x_{-1}^4 x_1^5 \quad .$$

Let $F(x_{-1}, x_1)$ be the **weight enumerator** of the set of Dyck words, i.e. the *sum* of all the weights of all these words, a certain **formal power series** in x_{-1}, x_1.

Obviously the weight of the empty word is 1 (the empty product), hence applying *weight* to ($DyckGrammar$), we get the **quadratic equation**

$$F = 1 \,+\, x_{-1}\, F\, x_1\, F \quad .$$

Abbreviating $X = x_{-1}\, x_1$, we get

$$F = 1 \,+\, X\, F^2 \quad .$$

Recalling what we learned in seventh grade (or what the Babylonians knew more than 3000 years ago), we can express F **explicitly**

$$F = \frac{1 - \sqrt{1-4X}}{2X} \quad .$$

Recalling what we learned in 12-th grade (or what Isaac Newton knew more that 300 years ago) we can write

$$F = \sum_{m=0}^{\infty} \frac{(2\,m)!}{m!\,(m+1)!} X^m \quad ,$$

implying the fact that the number of Dyck paths of length $2m$ is the super-famous **Catalan** number $C_m = \frac{(2m)!}{m!\,(m+1)!}$, that is the subject of Richard Stanley's modern classic [St], and the most popular sequence, `A108`, in the great OEIS [Sl].

The above is the standard, very boring proof of that famous fact. We know at least a dozen proofs, some of them are given in [St]. Here is one of our favorite proofs due to Aryeh Dvoretzky and Theodore Motzkin [DM].

The fact that the number of Dyck paths of length $2m$ equals the Catalan number C_m is equivalent the fact that the number of words in $\{1, -1\}$ of length $2m+1$ whose sum is 1 and all whose proper-partial sums are non-positive is C_m. Every word of length $2m+1$ in $\{-1, 1\}$ that adds up to 1 has $m+1$ '1' and m '$\overline{1}$'. There are $\binom{2\,m+1}{m}$ such words. The $2\,m\,+\,1$ *cyclic shifts* of each such word are all **different** (why?), and exactly one of them has the property that its partial sums are all non-positive (why?). Hence the number of gambling histories that we are interested in is $\frac{1}{2m+1} \cdot \binom{2m+1}{m} = C_m$.

Enter Probability

So far what we did was *enumerative combinatorics*. We found out that the weight-enumerator of the set of Dyck words is

$$\frac{1-\sqrt{1-4x_{-1}x_1}}{2\,x_{-1}x_1} \quad ,$$

and hence the weight enumerator of words in $\{-1,1\}$ that add-up to 1, and such that all their proper partial sums are ≤ 0, is x_1 times that, i.e.

$$\frac{1-\sqrt{1-4x_{-1}x_1}}{2\,x_{-1}} \quad .$$

Assume that each round in the gambling game is **independent** of the other ones, and for each of them the probability of winning a dollar is p, and hence of losing a dollar is $1-p$. Plugging-in $x_{-1}=(1-p)\,t$, $x_1=p\,t$, in the above explicit enumerating generating function, we get the following human-generated, well-known (see [F]) proposition.

Proposition 5: The **probability generating function** of the random variable 'numer of rounds it takes until the first time you have one dollar', if you start with 0 dollars and at each round you win a dollar with probability p and lose a dollar with probability $1-p$, let's call it $g(t)$ is

$$g(t)=\frac{1-\sqrt{1-4\,(1-p)\,pt^2}}{2\,(1-p)\,t} \quad .$$

So far all our power series were *formal*, but it is easy to see that if $p \geq \frac{1}{2}$ then plugging-in $t=1$ leads to a convergent series, that sums-up to 1, in agreement with the obvious fact that if $p>\frac{1}{2}$ sooner or later you will succeed, and the slightly less obvious fact that it is still true when $p=\frac{1}{2}$. If $p<\frac{1}{2}$, then we must take the other sign of the square-root, leading to the classical and well-known fact that the probability of one day having one dollar in your possession is $\frac{p}{1-p}$.

More generally, suppose that your goal in life is not just to exit the casino with one dollar, but you want to make n dollars. Since each additional dollar is yet another 1-dollar game, we immediately get.

Proposition 5': The **probability generating function** of the random variable 'numer of rounds it takes until the first time you have n dollars', if you start with 0 dollars and at each round you win a dollar with probability p and lose a dollar with probability $1-p$, is given by

$$\left(\frac{1-\sqrt{1-4\,(1-p)\,pt^2}}{2\,(1-p)\,t}\right)^n \quad .$$

From now let's assume that $p \geq \frac{1}{2}$. To get the **expected duration** we can sill do it by hand, find $(g(t)^n)' = ng(t)^{n-1}g'(t)$, then compute $g'(t)$, plug-in $t=1$ and simplify, getting that the expectation is $\frac{n}{2p-1}$.

For the k-th moment, we compute $(t\frac{d}{dt})^k(g(t)^n)$, plug-in $t=1$, and simplify, expressing all higher derivatives of $g(t)$ in terms of $g(t)$ and t, followed by substituting $t=1$.

An even better way, that would be the only way later on when we do the general gambling caes, is to use *implicit differentiation*, using the relation

$$f(t) \,=\, 1 + p\,(1-p)\,t^2 f(t)^2 \quad ,$$

and its implied relation for $g(t) = p\,t\,f(t)$.

It turns out that if you use the explicit expression $g(t) = \frac{1-\sqrt{1-4\,(1-p)\,pt^2}}{2\,(1-p)\,t}$ all the radicals disappear, and if you use implicit differentiation, and then plug-in $t=1$, you never have to divide 0 by 0, so either way you would get that all the moments are **polynomials** in n and rational functions in p. In particular, if p is a rational number, then they are all also rational numbers. The expectation, is $\frac{n}{2p-1}$.

For higher moments, We get the following computer-generated proposition.

Proposition 6: Let $X_{n,p}$ be the random variable "Number of rounds until reaching n dollars for the first time" in a gambling game where the probability of winning a dollar is p and of losing a dollar is $1-p$. Assume that $p > \frac{1}{2}$. We have

$$E[X_{n,p}] \,=\, \frac{n}{2p-1} \quad .$$

$$Var[X_{n,p}] \,=\, \frac{4\,n\,p\,(1-p)}{(2p-1)^3} \quad .$$

The *skewness* (aka scaled third moment about the mean) is

$$\alpha_3[X_{n,p}] \,=\, \left(-2\,p^2+2\,p+1\right)(-1+2\,p)^{-2}\,\frac{1}{\sqrt{-\frac{np(-1+p)}{(-1+2\,p)^3}}} \quad .$$

The *kurtosis* (aka scaled fourth moment about the mean) is

$$\alpha_4[X_{n,p}] \,=\, \frac{-4\,p^4+(6\,n+8)\,p^3+(-9\,n+6)\,p^2+(3\,n-10)\,p-1}{np\,(-1+p)\,(-1+2\,p)} \quad .$$

For the 5-th through 10-th scaled moments, see the output file

`http://sites.math.rutgers.edu/~zeilberg/tokhniot/oGenPileGames1.txt` .

The Two Player version for the $(1,-1)$ case

Using Lagrange inversion (see [Z4] for a lucid statement and proof) or otherwise, it is easy to see that the probability of reaching m dollars for the first time after exactly n rounds, in a solitaire game where the probability of winning a dollar is p and the probability of losing a dollar is $1-p$, let's call it $b_{n,m}$ is

$$b_{n,m} = \frac{m\,(2\,n+m-1)!\,p^{n+m}\,(1-p)^n}{n!\,(n+m)!} \quad .$$

Suppose that two players take turns and whoever reaches m dollars first is declared the winner. As before, the probability of winning the game for the player whose turn is to move is $a(m) = (1+f(m))/2$, where

$$f(m) = \sum_{n=1}^{\infty} b_{n,m}^2 \quad .$$

Using the Zeilberger algorithm once again we have the next computer-generated proposition.

Proposition 7: In the two player version game with a **fair** coin, i.e. the probability of winning a dollar and losing a dollar are both $\frac{1}{2}$, the winning probability of the player whose turn is to move is $(1+f(m))/2$ where $f(m)$ satisfies the second-order recurrence

$$\left(2\,m^2+5\,m+2\right) f\,(m+2) + \left(-12\,m^2-24\,m-10\right) f\,(m+1) + \left(2\,m^2+3\,m\right) f\,(m) = -\frac{8}{\pi} \quad ,$$

subject to the initial conditions

$$f\,(1) = -\frac{-4+\pi}{\pi} \quad , \quad f\,(2) = -\frac{-16+5\,\pi}{\pi} \quad .$$

For the loaded case, where $p > \frac{1}{2}$, we have the next proposition.

Proposition 8: In the two player version game with the probability of winning a dollar is p and losing a dollar is $1-p$, provided $\frac{1}{2} < p < 1$, the winning probability of the player whose turn is to move is $(1+f(m))/2$ where $f(m)$ satisfies the fourth-order recurrence

$$m\,(-1+p)^4\,(m-3)\,f\,(m) - (-1+p)^2\left(2\,m^2-7\,m+4\right) f\,(m-1)$$
$$+\left(-2\,m^2p^4+4\,m^2p^3+8\,mp^4-2\,m^2p^2-16\,mp^3-4\,p^4+8\,mp^2+8\,p^3+m^2-4\,p^2-4\,m+4\right) f\,(m-2)$$
$$-p^2\left(2\,m^2-9\,m+8\right) f\,(m-3) + p^4\,(m-1)\,(m-4)\,f\,(m-4) = 0 \quad ,$$

subject to the appropriate initial conditions.

Winning a dollar or losing k dollars

Now let's generalize to the gambling game where, as before, you start with a capital of 0 dollars, but now at each round you **win** a dollar with probability p or **lose** k dollars with probability $1-p$, and the game ends as soon as you owe 1 dollar. Very soon we will treat the more general case where the goal is to exit with m dollars, but for now let's consider the case of $m=1$.

In order to guarantee that the game ends, the expected gain of a single round, $p\cdot 1-(1-p)\cdot k = (k+1)\,p - k$ should be positive. So we will assume that $p > \frac{k}{k+1}$. In the border-line case $p = \frac{k}{k+1}$ the game still ends with probability 1, but its expected duration is infinite.

Now the **alphabet** is $\{1,-k\}$, and we will try to adapt the above argument that workded for the classical case. Let's abbreviate $\bar{k} := -k$. Now the steps are $(1,1)$ and $(1,-k)$.

Let's study the *anatomy* of such words (histories) or, equivalently, paths. Obviously all these words are of length $n(k+1)+1$, for some non-negative integer n, and they all end with 1. So we can write, for *any* history W,

$$W = U\,1 \quad ,$$

where U is a word that sums to 0, all whose partial sums are non-positive. we will call such words $(1,-k)$-*Dyck words.*

Let's analyze such a $(1,-k)$-Dyck word U or rather its corresponding path from $(0,0)$ to $((k+1)n,0)$, say. Of course, it may be the empty word, but if it is not, let $(r(k+1),0)$ $0<r\leq n$ be the **first** time that it hits the x-axis. Then we can write

$$U = U_1\,U_2 \quad ,$$

where U_2 is another arbitrary $(1,-k)$-Dyck word, but U_1 has the special property that all its partial sums (except the 0-th and the last) are **strictly negative**, or in terms of its path, except for its starting and ending points, they lie **strictly** below the x-axis. Such a word must **necessarily** start with a $\bar{k}$ and end with a 1, but to recover the 'debt' of k, must regain these lost k dollars, one dollar at a time, so it may be written as $\bar{k}\,(U_3 1)^k$, where U_3 is an arbitrary $(1,-k)$-Dyck word. Conversely, for any such word U_3, $\bar{k}\ (U_3\,1)^k$ is such a strictly below the x-axis word. So we have the (context-free) *grammar*

$$U = EmptyWord \vee \bar{k}\ (U\,1)^k U \quad , \qquad ((1,-k)-DyckGrammar)$$

where now U stands for 'an arbitrary $(1,-k)$-Dyck word'.

Let $F(x_{-k},x_1)$ be the *weight-enumerator* for all such words. Applying the *weight* operation, we get that $F=F(x_{-k},x_1)$ satisfies

$$F = 1 + (x_{-k}x_1^k)\,F^{k+1} \quad .$$

Abbreviating $X := x_{-k}\,x_1^k$, this can be written

$$F = 1 + X\,F^{k+1} \quad .$$

When $k=2$ and $k=3$, we can solve these equations 'explicitly' using 'radicals', thanks to Cardano and Ferrari, but thanks to Abel, Ruffini, and Galois we know that we can **not** do it for $k\geq 4$. Even the 'explicit' solutions for $k=2$ and $k=3$ are not very useful. On the other hand, thanks to *Lagrange inversion* (see,e.g. [Z4]) we can find the Maclaurin expansion explicitly.

$$F(X) = \sum_{m=0}^{\infty} \frac{((k+1)\,m)!}{m!\,(km+1)!} X^m \quad ,$$

featuring the **Fuss-Catalan** numbers $C_{k,m} = \frac{((k+1)\,m)!}{m!\,(km+1)!}$.

It follows that the weight-enumerator of words in $\{-k, 1\}$ that add-up to 1, and such that the proper-partial sums are all non-positive is $F(x_{-k}x_1^k)\,x_1$, since the last letter must be 1.

Equivalently (and that's is our actual object of interest) the number of words with m '$-k$' and $mk+1$ '1' whose proper-partial sums are all non-positive equals the Fuss-Catalan number $C_{k,m}$. This can be also proved by adapting the [DM] proof. There are $\binom{mk+1+m}{m}$ words altogether, and for each of these its $mk+1+m$ cyclic shifts are all different, and exactly one of them is a 'good' word, hence there are $\frac{1}{mk+1+m}\binom{mk+1+m}{m} = C_{k,m}$ such words.

Since, in order to exit with n dollars , we must gain one dollar, n times, the weight-enumerator of words that reach n for the first time is $(F(x_{-k}x_1^k)\,x_1)^n$.

So far we did enumerative combinatorics. To convert it to probability, we plug-in the above $x_1 = p\,t$ and $x_{-k} = (1-p)\,t$. Using implicit differentiation, we can compute the expectation, variance, and higher moments. Since in this case we do not encounter $0/0$, all the moments are **rational** functions of p. In particular, if the number p is rational, all the quantities are rational numbers.

Using implicit differentiation, for **symbolic** k and **symbolic** p and **symbolic** n, our beloved computer generated the next proposition.

Proposition 9: Suppose that at each round, you win a dollar with probability p and lose k dollars with probability $1-p$, and you quit as soon as you reach n dollars. If $p > k/(k+1)$, then, of course, sooner or later you will reach your goal. How long should it take? Denote by $X_{n,k,p}$ the random variable, 'number of moves until reaching n dollars'. We have the following facts.

Let $g(t)$ be the formal power series, in t, satisfying the algebraic equation

$$g\,(t) - 1 - p^k\,(1-p)\,t^{k+1}\,g(t)^{k+1} \;=\; 0 \quad .$$

The probability generating function of $X_{n,k,p}$ is

$$(\,p\,t\,g(t)\,)^n \quad .$$

By implicit differentiation, followed by substituting $t=1$, we can compute any desired derivative, and hence the expectation, variance, and higher moments. We have

$$E[X_{n,k,p}] \;=\; \frac{n}{(p-1)k+p} \quad ,$$

[as *expected* (npi), since the expected gain in one move is $(p-1)k+p$]. The variance is given by

$$Var[X_{n,k,p}] \;=\; -\frac{np\,(k+1)^2\,(p-1)}{((p-1)\,k+p)^3} \quad .$$

The *skewness* (aka 'third scaled-moment about the mean') is

$$\alpha_3[X_{n,k,p}] \;=\; -\,(k+1)\left(kp^2+p^2-k-2\,p\right)(kp-k+p)^{-2}\,\frac{1}{\sqrt{-\frac{np(k+1)^2(p-1)}{((p-1)k+p)^3}}} \quad .$$

The *kurtosis* (aka 'fourth scaled-moment about the mean') is

$$\alpha_3[X_{n,k,p}] =$$

$$\frac{-\left(k+1\right)^2 p^4 - 2\,\left(k+1\right)\left(k-\frac{3}{2}\,n-3\right)p^3 + \left(6\,k^2+\left(-6\,n+6\right)k-3\,n-6\right)p^2 - 2\,k\left(k-\frac{3}{2}\,n+4\right)p - k^2}{np\left(p-1\right)\left(p\left(k+1\right)-k\right)} \quad .$$

For the scaled fifth and sixth moments, see the output file

`http://sites.math.rutgers.edu/ zeilberg/tokhniot/oGenPileGames2.txt` .

The Two Player version for the $(1,-k)$ case

Since the probability mess function is explicit, given in terms of the Fuss-Catalan numbers, we can use the Zeilberger algorithm to compute recurrences for the probability of the first player winning, for **symbolic** n, and **symbolic** p (assuming that it is larger than $\frac{k}{k+1}$). Alas, we can **not** do it for symbolic k, since the Fuss-Catalan numbers are not bi-holonomic in both n and k.

For the case $k=2$ we have the next proposition.

Proposition 10: In the two player version game, if the probability of winning a dollar is p and of losing two dollars is $1-p$, provided $\frac{2}{3}<p<1$, the probability of the player whose turn is to move of winning the game is $(1+f(m))/2$ where $f(m)$ satisfies the sixth-order linear recurrence

$$m\left(p-1\right)^4\left(m-5\right)f\left(m\right)$$

$$-2\,\left(p-1\right)^2\left(m^2-6\,m+6\right)f\left(m-2\right)-p^2\left(p-1\right)^2\left(2\,m^2-13\,m+12\right)f\left(m-3\right)$$

$$+\left(m-3\right)\left(m-4\right)f\left(m-4\right)-p^2\left(2\,m^2-15\,m+24\right)f\left(m-5\right)+p^4\left(m-2\right)\left(m-6\right)f\left(m-6\right)=0 \quad ,$$

subject to the appropriate initial conditions.

For the case $k=3$ we have the next proposition.

Proposition 11: In the two player version game, if the probability of winning a dollar is p and of losing three dollars is $1-p$, provided $\frac{3}{4}<p<1$, the probability of the player whose turn is to move of winning the game is $(1+f(m))/2$ where $f(m)$ satisfies the eighth-order linear recurrence

$$m\left(p-1\right)^4\left(m-7\right)f\left(m\right)-\left(p-1\right)^2\left(2\,m^2-17\,m+24\right)f\left(m-3\right)$$

$$-2\,p^2\left(p-1\right)^2\left(m^2-9\,m+12\right)f\left(m-4\right)+\left(m-4\right)\left(m-6\right)f\left(m-6\right)$$

$$-p^2\left(2\,m^2-21\,m+48\right)f\left(m-7\right)+p^4\left(m-3\right)\left(m-8\right)f\left(m-8\right)=0 \quad ,$$

subject to the appropriate initial conditions.

For the case $k=4$ we have the next proposition.

Proposition 12: In the two player version game, if the probability of winning a dollar is p and of losing four dollars is $1-p$, provided $\frac{4}{5} < p < 1$, the probability of the player whose turn is to move of winning the game is $(1+f(m))/2$ where $f(m)$ satisfies the tenth-order linear recurrence

$$m\,(p-1)^4\,(m-9)\,f\,(m) - 2\,(p-1)^2\,\left(m^2-11\,m+20\right)f\,(m-4)$$

$$-p^2\,(p-1)^2\,\left(2\,m^2-23\,m+40\right)f\,(m-5)+(m-5)\,(m-8)\,f\,(m-8)$$

$$-p^2\left(2\,m^2-27\,m+80\right)f\,(m-9)+p^4\,(m-4)\,(m-10)\,f\,(m-10)=0 \quad ,$$

subject to the appropriate initial conditions.

For the case $k=5$ we have the next proposition.

Proposition 13: In the two player version game, if the probability of winning a dollar is p and of losing five dollars is $1-p$, provided $\frac{5}{6} < p < 1$, the probability of the player whose turn is to move of winning the game is $(1+f(m))/2$ where $f(m)$ satisfies the 12^{th}-order linear recurrence

$$m\,(p-1)^4\,(m-11)\,f\,(m) - (p-1)^2\,\left(2\,m^2-27\,m+60\right)f\,(m-5)$$

$$-2\,p^2\,(p-1)^2\,\left(m^2-14\,m+30\right)f\,(m-6)+(m-6)\,(m-10)\,f\,(m-10)$$

$$-p^2\left(2\,m^2-33\,m+120\right)f\,(m-11)+p^4\,(m-5)\,(m-12)\,f\,(m-12)=0 \quad ,$$

subject to the appropriate initial conditions.

Winning k dollars or losing one dollar

This case is more complicated than the previous one, and we will have to treat one k at a time even for the expectation. Also, we only consider the case of reaching at least one dollar for the first time, rather than the more general case of reaching n dollars for the first time.

Now our **alphabet** is $\{\,k\,,\,-1\,\}$ and, in terms of lattice paths, the atomic steps are $(1,k)$ and $(1,-1)$.

Since the last step of such a path must be $(1,k)$ it can terminate at $y=k$, or $y=k-1,\ldots,y=1$, so we are forced to consider, in addition to $U_{0,0}$ the set of paths that start at $y=0$ and end at $y=0$ and never go above the x-axis, also $U_{0,-1}$ the set of paths that start at $y=0$ and end at $y=-1$ and never go above the x-axis, all the way to $U_{0,-(k-1)}$, the set of paths that start at $y=0$ and end at $y=-(k-1)$ and never go above the x-axis.

Such a word looks like

$$U_{0,0}\,k \vee U_{0,-1}\,k \vee \quad \ldots \quad \vee U_{0,-(k-1)}\,k \quad .$$

Let $U := U_{0,0}$. Then the weight-enumerator of U is $F(x_k x_{-1}^k)$ where $F(X)$ is as above, the solution of

$$F(X) \,=\, 1 \,+\, XF(X)^{k+1} \quad .$$

It can be seen that $U_{0,-r} = (\overline{1}\, U_{0,0})^r$, hence its weight-enumerator is $(x_{-1}\, F(X))^r$.

Substituting for $x_{-1} \, = \, p\,t$ and $x_k \, = \, (1-p)\,t$, we get the following human-generated proposition.

Proposition 14: Suppose that at each round, you lose one dollar with probability p and win k dollars with probability $1-p$, and you quit as soon as you reach at least 1 dollar. If $0 < p < \frac{k}{k+1}$ then, of course, sooner or later, you will reach your goal. Let $g(t)$, be the formal power series, in t, satisfying the algebraic equation

$$g(t) \, - \, 1 \, - \, p^k\,(1-p)\,t^{k+1}\,g(t)^{k+1} \, = \, 0 \quad .$$

The probability generating function, let's call it $f(t)$, for the number of rounds until having a positive capital is

$$f(t) \, = \, (1-p)\,t\,g(t)\,\sum_{i=0}^{k-1}(p\,t\,g(t))^i \quad .$$

If you will apply implicit differentiation to the defining equation of $g(t)$, and then express $f'(t)$ in terms of $g(t)$ and $g'(t)$ and then plug-in $t=1$, you will get $0/0$. It turns out that the expressions for the expectation, variance, and higher moments are no longer rational functions of p, but are roots of **algebraic** equations. The reason is that when $t=1$, 1 is a double (or higher-order) root of the defining equation for the probability itself $f(1)=1$.

Since Maple knows how to differenitate, both explicitly and implicitly, our beloved computer can handle it all automatically, and get explicit algebraic equation for symbolic p, or specific algebraic numbers for specific $p < \frac{k}{k+1}$, alas only for **one k at a time**.

We have the following computer-generated proposition for the case $k\,=\,2$, i.e. for the gambling options $\{-1,2\}$, with $Pr(-1)=p$ and $Pr(2)=1-p$.

Proposition 15: Let X be the random variable 'number of rounds until you reach positive capital' if you start at 0, and at each round, you lose 1 dollar with probability p and win 2 dollars with probability $1-p$. Assume that $p < \frac{2}{3}$.

The expectation is given by

$$E[X] \, = \, \frac{3\,p + \sqrt{-\,(3\,p+1)\,(-1+p)} - 1}{2\,p\,(2\,-\,3\,p)}$$

For the variance, and third through the sixth moment, see

`http://sites.math.rutgers.edu/~zeilberg/tokhniot/oGenPileGames3.txt` .

Note that for the most interesting case, $p=\frac{1}{2}$, the expectation is the **beautiful number** $1+\sqrt{5}$ (twice the golden ratio). This is so nice that we will single it out.

Beautiful Corollary: If a one-dimensional random walker starts at 0 and moves **one step back** with probability $\frac{1}{2}$ and **two steps forward** with probability $\frac{1}{2}$ and keeps going until he is at a location ≥ 1 for the first time, the expected number of steps that he takes is twice the Golden Ratio, i.e. $1+\sqrt{5}$.

For $k \geq 3$ and *symbolic* p, things get too complicated to reproduce here, so let's just mention the expectations for a few cases for the most interesting case, $p = \frac{1}{2}$.

$k = 3$: The expected duration of a random walk with $Pr(-1) = Pr(3) = \frac{1}{2}$ until reaching a location ≥ 1 for the first time is the positive root of

$$x^3 - 4\,x - 4 = 0 \quad ,$$

that equals $2.382975767906237494\ldots$.

$k = 4$: The expected duration of a random walk with $Pr(-1) = Pr(4) = \frac{1}{2}$ until reaching a location ≥ 1 for the first time is the positive root of

$$3\,x^4 + 4\,x^3 - 8\,x^2 - 24\,x - 16 \;=\; 0 \quad ,$$

that equals $2.1561901553356811691\ldots$.

$k = 5$: The expected duration of a random walk with $Pr(-1) = Pr(5) = \frac{1}{2}$ until reaching a location ≥ 1 for the first time is the positive root of

$$2\,x^5 + 5\,x^4 - 20\,x^2 - 32\,x - 16 \;=\; 0 \quad ,$$

that equals $2.07050432323944926\ldots$.

$k = 6$: The expected duration of a random walk with $Pr(-1) = Pr(6) = \frac{1}{2}$ until reaching a location ≥ 1 for the first time is the positive root of

$$5\,x^6 + 18\,x^5 + 20\,x^4 - 40\,x^3 - 144\,x^2 - 160\,x - 64 \;=\; 0 \quad ,$$

that equals $2.0333823565252879532\ldots$.

$k = 7$: The expected duration of a random walk with $Pr(-1) = Pr(7) = \frac{1}{2}$ until reaching a location ≥ 1 for the first time is the positive root of

$$3\,x^7 + 14\,x^6 + 28\,x^5 - 112\,x^3 - 224\,x^2 - 192\,x - 64 \;=\; 0 \quad ,$$

that equals $2.0162018012796575781\ldots$.

$k = 8$: The expected duration of a random walk with $Pr(-1) = Pr(8) = \frac{1}{2}$ until reaching a location ≥ 1 for the first time is the positive root of

$$7\,x^8 + 40\,x^7 + 112\,x^6 + 112\,x^5 - 224\,x^4 - 896\,x^3 - 1280\,x^2 - 896\,x - 256 \;=\; 0 \quad ,$$

that equals $2.00796926912597191\ldots$.

References

[DM] A. Dvoretzky and Th. Motzkin, *A problem of arrangements*, Duke Math. J. **14** (1947), 305-313.

[F] William Feller, "*An Introduction to Probability Theory and Its Application*", volume 1, three editions. John Wiley and sons. First edition: 1950. Second edition: 1957. Third edition: 1968.

[Sl] Neil J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `http://www.oeis.org` .

[St] Richard Stanley, "*Catalan Numbers*", Cambridge University Press, 2015.

[Z4] Doron Zeilberger, "*Lagrange Inversion Without Tears (Analysis) (based on Henrici)*, The Personal Journal of Shalosh B. Ekhad and Doron Zeilberger
`http://sites.math.rutgers.edu/~zeilberg/mamarim/mamarimhtml/lag.html` .

Written: Sept. 10, 2019.