Formalization of Gambler's Ruin Problem in Isabelle/HOL

Zibo Yang

zibo.yang@ip-paris.fr

Department of Informatics

Ecole Polytechnique/ Institut Polytechnique de Paris

July 30, 2021

Table of Contents

Introduction

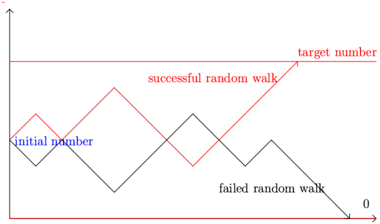
Karl Sigman describes the gambler's ruin problem as follows:

> *Let $N \geq 2$ be an integer and let $1 \leq i \leq N - 1$. Consider a gambler who starts with an initial fortune of \$i and then on each successive gamble either wins \$1 or loses \$1 independent of the past with probabilities p and $q = 1 - p$ respectively. Let X_n denote the total fortune after the n^{th} gamble. The gambler's objective is to reach a total fortune of \$N, without first getting ruined (running out of money). If the gambler succeeds, then the gambler is said to win the game. In any case, the gambler stops playing after winning or getting ruined, whichever happens first. Now we want to know the probability for the gambler to win(reaching the target \$N)*

Introduction

Tom Leighton and Ronitt Rubinfeld further generalize the problem into one-dimensional random walk model as below:

> *This problem is a classic example of a problem that involves a one-dimensional random walk. In such a random walk, there is some value, say the number of dollars we have, that can go up or down or stay the same at each step with some probabilities. In this example, we have a random walk in which the value can go up or down by 1 at each step.*



Model formalization of Gambler's ruin problem

locale infinite_coin_toss_space =
fixes p::real and M::"bool stream measure"
assumes p_gt_0: "$0 \leq p$"
and p_lt_1: "$p \leq 1$"
and bernoulli: "$M = $ bernoulli_stream p

measure =	space: Ω	+	power set: A = {all the measurable subsets}	+	measure function: μ: measurable set $\Rightarrow$ real + ∞
	$\Downarrow$		$\Downarrow$ (Instantiation)		$\Downarrow$
bernoulli measure =	{True,False}	+	{all the subsets of {True,False}}	+	$\mu(\{\text{True}\}) = p$ $\mu(\{\text{False}\}) = 1\text{-}p$
	$\Downarrow$		$\Downarrow$ (Infinite Cartesian Product)		$\Downarrow$
bernoulli stream =	$\{\text{True,False}\}^\infty$	+	{all the measurable subsets of $\{\text{True,False}\}^\infty$ produced by infinite cartesian product}	+	$\mu(x) = \inf_\infty \{\mu(A_i)\}$ where A_i is a sequence of simple measurable subset produced by infinite cartesian product

Model formalization of Gambler's ruin problem

locale gambler_model = infinite_coin_toss_space +
fixes geom_proc::"int ⇒ bool stream ⇒ enat ⇒ int"
assumes geometric_process:
"geom_proc init x step = gambler_rand_walk 1 (-1) init step x"

- init: the initial number at the beginning
- x: the infinite boolean sequence ([True, False,False,....])
- step: the position where we calculate

Formalization of probability calculation

Assume we start with $n > 0$ and randomly walk of plus 1 with the possibility of p or minus 1 with the possibility of $1 - p$, then we end this game once we get $m > n$ or 0. The possibility of reaching m from initially n without any stop is what we want to calculate and formalize in the final.

let P_n be the possibility of successfully reaching m with initial n, then it's fairly easy to have the equations:

$$P_n = pP_{n+1} + (1 - p)P_{n-1} \text{ if } 0 < n < m$$

$$P_0 = 0$$

$$P_m = 1$$

The final equation: $P_n = \frac{\left(\frac{1-p}{p}\right)^n - 1}{\left(\frac{1-p}{p}\right)^m - 1}$

Formalization of probability calculation

Key equation here:

$$P_n = pP_{n+1} + (1-p)P_{n-1} \textit{ if } 0 < n < m$$

The two lemmas below are the most important intermediate conclusions we need to prove:

```
lemma success_measurable:
fixes init target
assumes "0 < init < target"
shows "{x∈ space M . success init x target} ∈ sets M"
lemma semi_goal_true:
fixes init target
assumes "0 < init < target"
shows
"emeasure M {x∈ space M . success (init+1 ) (stl x ) target
∧ shd x} = emeasure M {x ∈ space M . shd x} × emeasure M
{x∈ space M . success (init+1 ) (x ) target}"
```

The probability theory of Archive of Formal Proofs(Isabelle/HOL)is established on measure theory

Calculate the probability = Measure the corresponding set

Probability of successful random walks = Measure value of the set of successful random walks

- Is the set of successful random walks measurable?
- How to calculate the measure value of a specific measurable set?

Is the set of successful random walks measurable?

```
lemma success_measurable:
fixes init target
assumes "0 < init < target"
shows "{x∈ space M . success init x target} ∈ sets M"
```

Biggest confusion : To measure the set, you need to prove it's measurable?
Biggest challenge : Proving if the set is measurable or not is theoretically normal but practically difficult and tedious

Is the set of successful random walks measurable?

- Lemma *finite_stake_measurable* states that for the function (λw. stake n w) listing the first n steps of random walk, the preimage of a finite set is measurable for measure M.
- Lemma *finite_image* states that sets filled with all bool lists of fixed length n is finite.
- Lemma *success_measurable*2 is the most important intermediate lemma prepared for lemma *success_measurable*. It clarifies that any list in the image of successful random walk over function stake will never contain another shorter list corresponding to another successful random walk, which sets up the bijection between successful random walks stopping at fixed-step and preimage of successful bool list with identical length.
- Lemma *success_measurable*1 demonstrates that a set of successful random walks is a countable union of sets of successful random walks stopping at some step.

The set of successful random walks is measurable?
({x. success init x target} ∈ sets M)

Lemma success_measurable1: countable union of measurable sets is measurable set.

The set of successful random walks stopping at step n is measurable?
({x. success init x target ∧ stop_at x = n} ∈ sets M)

Lemma success_measurable2: the preimage of measurable set through a measureable function is measurable

The function (λx. stake n x) is a measurable function?
(stake 2 [True,True,.....] = [True, True])

The image of function (λx. stake n x) is measurable?
(stake n {x. success init x target ∧ stop_at x = n} ∈ sets N)

Lemma finite_stake_measurable: the preimage of finite set through function (λx. stake n x) is measurable

The preimage of function (λx. stake n x) is finite?
(finite (stake n {x. success init x target ∧ stop_at x = n}))

Any subset of the finite set is finite

The set of boolean lists with specific length n is finite?
(finite {y ∈ boolean list. length y = n})

Lemma finite_image: the set of boolean lists with specific length is finite.

Yes, well done

How to calcuate the measure value of a specific measurable set?

```
lemma semi_goal_true:
fixes init target
assumes "0 < init < target"
shows
"emeasure M {x∈ space M . success init+1 (stl x) target
∧ shd x} = emeasure M {x ∈ space M . shd x} × emeasure M
{x∈ space M . success init+1 x target}"
```

How to calcuate the measure value of a specific measurable set?

Problem: how to measurable value of set of successful random walks?

measure =	space: Ω	+	power set: A = {all the measurable subsets}	+	measure function: μ: measurable set $\Rightarrow$ real + ∞
	$\Downarrow$		$\Downarrow$ (Instantiation)		$\Downarrow$
bernoulli measure	={True,False}	+	{all the subsets of {True,False}}	+	$\mu(\{\text{True}\}) = \text{p}$ $\mu(\{\text{False}\}) = \text{1-p}$
	$\Downarrow$		$\Downarrow$ (Infinite Cartesian Product)		$\Downarrow$
bernoulli stream	= $\{\text{True,False}\}^\infty$	+	{all the measurable subsets of $\{\text{True,False}\}^\infty$ produced by infinite cartesian product}	+	$\mu(x) = \inf_\infty\{\mu(A_i)\}$ where A_i is a sequence of simple measurable subset produced by infinite cartesian product

Challenge : Calculation of the set's measure value is theoretically feasible but practically almost "infeasible"?

Breakthough : Integral equation provided by Mnacho Echenim in the stochastic process library(2021)

```
lemma (in prob_space) emeasure_stream_space:
assumes "X ∈ sets (stream_space M)"
shows "emeasure (stream_space M) X =
(∫+t.  emeasure (stream_space M) {x∈space (stream_space
M). t ## x ∈ X } ∂M)"
```

- Lemma *semi_goal*1 states the equation between the measure value of set of successful random walks calculated from either first step or second step, which facilities our prepare for the function with first step as input and measure value as output. Preparation for the function
- Lemma *semi_goal*2*_final* declares the function we set up in lemma *semi_goal*1 can be calculated in integral prefectly. Integral for the function

Formalization of probability calculation

The final goal we will reach:

```
lemma final_goal:
fixes init target
assumes "0 < init < target" shows "emeasure M {x∈ space M
. success init x target} = emeasure M {x ∈ space M . shd
x} × emeasure M {x∈ space M . success (init+1) x target} +
emeasure M {x ∈ space M . ¬shd x} × emeasure M {x∈ space
M . success (init-1) x target}"
```

$$P_n = pP_{n+1} + (1-p)P_{n-1} \text{ } \mathit{if} \text{ } 0 < n < m$$

Summary

We aim to formalize gambler's ruin problem by Isabelle/HOL.

- Abstract interpretation: gambler's ruin model and recursive probability equation.

Summary

We aim to formalize gambler's ruin problem by Isabelle/HOL.

- Abstract interpretation: gambler's ruin model and recursive probability equation.
- The biggest challenges, Proving the measurability of specific sets and calculating its measure value, have been solved.

Summary

We aim to formalize gambler's ruin problem by Isabelle/HOL.

- Abstract interpretation: gambler's ruin model and recursive probability equation.
- The biggest challenges, Proving the measurability of specific sets and calculating its measure value, have been solved.
- Next step: formalization of security analysis of bitcoin network.