Some Classic Problems in Probability

Ankita Sarkar

October 7, 2016

These problems are well-known staples in undergraduate study and beyond, but are restated below in a way that assumes a Class 12 level knowledge of probability. Young readers might want to tackle them from a puzzle-solving and intuitive point of view before attempting a formal mathematical approach.

1. *(Gambler's Ruin*[1]*)* Two gamblers, A and B, bet on an event with two possible outcomes – say, the flip of a coin – and they have starting money i and $(M - i)$ rupees respectively. If the coin comes up heads, A gets one rupee from B, and if the coin comes up tails, A gives one rupee to B. The game ends when one gambler loses completely i.e. goes broke, while the other, obviously, possesses the entire starting money M. Given that the coin turns up heads with probability p, **what is the probability that A wins?**
   To think about this situation further:

   - Is it possible that the game goes on forever?
   - What about the case with N gamblers instead of two, where each pair play independent games and we consider the final result?
   - What if the money gained/lost by the players are unequal instead of being one rupee each?
   - What if one of the players had infinite amount of money?

   For more mathematical ideas originating in gambling, look up *martingales* and *gambler's fallacy.*

[1] This is the oldest formalized mathematical idea to which the term 'gambler's ruin' was applied, but is certainly not the first gambling problem to be called that.

2. *(Pólya Urn Model)*[2] This model involves an urn being filled with different coloured balls by a 'rich gets richer'-type rule, the simplest formulation of which is as follows: an urn contains a black ball and a white ball. A ball is drawn at random. Its colour is noted, and then it is replaced into the urn along with one new ball of the same colour as the one drawn. In such a model:

   - What is the probability that there will be i black balls after the n^{th} stage of draw and replacement?
   - How is the above problem modified if c extra balls are added at each stage instead of one, and we begin with b black balls and w white ones instead of one each?[3]
   - Encode the process as a sequence of random variables $X_0, X_1...$, such that $X_j = 1$ if the j^{th} ball drawn is black, and zero if it is white. Try to find $P(S_n = s|S_1 = s_1, ..., S_{n-1} = s_{n-1})$. Is S_n equal to the number of black balls accumulated in the urn at stage n?
   - Consider the above with χ colours instead of two: what modifications must you make? Is the encoding hinted in the third problem still useful?
   - *(Optional)* How would the model change if the number of balls added depended on the colour drawn?

3. *(The Monty Hall Problem)* In a game show[4], the contestant is shown three doors – numbered 1,2,3 – and told that behind one of the doors, there is a car. Behind the other two, there are goats. The contestant has to play a game via which he chooses a door, and receives whatever is behind that door. We assume the car is the favourable (or 'winning') choice (ie. the goats could just as well be nothing). The game proceeds as follows: the contestant picks a door at first and declares the choice. The host then opens one of the two unchosen doors that conceal a goat, without revealing anything about the other two. The contestant is then given a chance to either retain their earlier choice, or switch to the other remaining door. This choice is considered the final one. **Is it advisable for the contestant to switch?** Although with these

[2]Named after Hungarian Mathematician George Pólya.

[3]*Hint:* Which values of i become impossible to attain?

[4]The Monty Hall Problem is loosely based on the game show *Let's Make A Deal*, the original host of which was Monty Hall. It was first published by American biostatistician Steve Selvin in 1975, in the magazine *American Statistician.*

smaller numbers one can write out the cases, attempt a mathematical solution – and hence arrive at the answer for a general n number of doors where, as before, only one is favourable.[5]

4. *(The Banach Match Problem)* [6] A pipe-smoking mathematician keeps two matchboxes with him at all times – one in his left pocket and one in his right pocket. When he wants a match, he is equally likely to take one from either box. Consider the first instance when he reaches for a match and finds the box empty. Assuming that both matchboxes had N matches each initially, what is the probability that the other matchbox has exactly k matches left in it?

5. *(The Drunkard's Walk[7])* A drunkard exits a tavern and decides to go home, but he is so drunk that he cannot walk properly. His city is an infinite grid on integer axes, and the vertices are the various locations, with the origin being the tavern, and one of the other locations being his home. At each step, the drunkard moves either forwards, backwards, left or right with equal probability ($= 1/4$). More mathematically, at $t = 0$, the drunkard is at $(0, 0)$ and the locations in his city are all points in the set $\mathbb{Z} \times \mathbb{Z}$. If at $t = i$ he is at (x_i, y_i), then at $t = i+1$, he is at (or, (x_{i+1}, y_{i+1}) equals) (x_i+1, y_i+1), (x_i+1, y_i-1), $(x_i - 1, y_i + 1)$ or $(x_i - 1, y_i - 1)$ with equal probability, $= 1/4$. **Now, what is the probability that he reaches home (say, (h, k))?** You will find that computing this is not as easy as it sounds (even for the simpler one-dimensional case of locations on a straight path i.e. just $\mathbb{Z}$ – try). Hence, we shall ponder on some simpler questions:

   - Let us define the drunkard's positions as a sequence of random variables $\{A_i\} = A_0, A_1, .., A_i, ...$ where A_i is the position at time $t = i$. Let us also define $(\{X_i\}, \{Y_i\}) := \{A_i\}$ ie. two new sequences of the separate co-ordinates such that $\forall i \geq 0$, $A_i = (X_i, Y_i)$. Are $\{X_i\}$ and $\{Y_i\}$ independent?

[5] This puzzle if famous for misleading humans. Pigeons, however, figured out the winning strategy after a few repeated tries. Other such statistical puzzles, of varying difficulties, include the *Three Prisoners Problem* (logically identical to Monty Hall), the *Boy-Girl Paradox* and the *Sleeping Beauty Problem*.

[6] After Polish mathematician Stefan Banach. Whether he formulated the problem or someone else named it after him is debated. Either way, it refers to Banach's smoking habit; which brings us to this: *Smoking is injurious to health. Tobacco causes cancer.*

[7] This is nothing but an example of a *Random Walk*. Do look it up, but you might want to make a first attempt at this problem without that formalisation.

- A_i clearly depends on A_{i-1}. Does it also depend on $A_0, .., A_{i-2}$? What about the corresponding dependencies in $\{X_i\}$ and $\{Y_i\}$?[8]
- Let us define a new sequence of random variables $\{Z_i\}$ to be the man's 'displacement' ie. $\forall i \geq 0,\ Z_i = \sqrt{{X_i}^2 + {Y_i}^2}$. Repeat the above problem for this sequence.

Those interested in more such problems and further reading are encouraged to look up *A First Course in Probability* by Sheldon Ross. Students familiar with the related material in Ross could read from more advanced books, like *Stochastic Processes with Applications* by R.N. Bhattacharya and E.C. Waymire.

[8] *Hint:* Compare $P(A_i = a_i | A_{i-1} = a_{i-1}$ and $P(A_i = a_i | A_{i-1} = a_{i-1}, A_{i-2} = a_{i-2}, .., A_0 = a_0)$. When you are done, look up *Markov Chains*.