

Paper of the Month, April 2019: Shuo-Yen Robert Li (1980) [A Martingale Approach to the Study of Occurrence of Sequence Patterns in Repeated Experiments](#), The Annals of Probability, Vol. 8, No. 6 (Dec., 1980), pp. 1171-1176

If a monkey types only capital letters, and is on every occasion equally likely to type any of the 26, how long on average will it take the monkey to produce the sequence: 'ABRACADABRA'? This seemingly childish question has a lot of connections to important real-world problems. For instance, the distribution of scan statistics can be related to this type of waiting time. The study of occurrence of patterns is a classical problem in probability theory. In the first volume of his famous book Feller has several sections devoted to this topic. The distribution of the waiting time can be derived in many different ways. But my personal favorite is the martingale approach developed in Li (1980). Li's key observation is that information on the occurrence times of patterns can be obtained from the values assumed by a specially constructed auxiliary martingale at a certain well-chosen time. Li's method is a textbook example of a clever application of fundamental Doob's Optional Stopping Theorem. This elegant mathematical trick did not change the world, but I like it a lot. And by the way, the expected time is $26^{11} + 26^4 + 26$.