

Project – TMA4265 Stochastic Processes Autumn 2009

In the dice game of *Yahtzee* (<http://en.wikipedia.org/wiki/Yahtzee>) five dice are used. One of the tasks is to obtain five-of-a-kind (five equal faces) in three rolls. After the first and second roll the player can save any dice that are wanted to complete the task and re-roll the other dice.

We will investigate statistical aspects of obtaining i equal faces after n rolls using five dice by modelling such a sequence of rolls by a Markov chain.

You may use any computing tools you wish to simulate this Markov chain. You may also use computing tools to answer the analytical questions (*i.e.*, to solve systems of equations or to invert a matrix).

This assignment is related to Problem 1 of the TMA4265 exam of 2008 Autumn – see <http://www.math.ntnu.no/~bakke/TMA4265/2008H/e.pdf>.

1. Argue that the Markov chain of the 2008 exam can be simplified by omitting state 0 and instead starting with $X_0 = 1$. Specify the states and, assuming that an optimal strategy to obtain five-of-a-kind is applied, the transition probabilities of the resulting five-state Markov chain. Assume that state 5 is absorbing. Demonstrate how you arrive at P_{12} , P_{13} , P_{22} and P_{23} .

2. Simulate 10000 sequences of 100 rolls, *i.e.* 10000 realizations of the Markov chain.

You may choose whether to save all of the simulations, or to accumulate only quantities that are needed for the following questions. Also, you may choose whether to simulate the dice rolls and the strategy directly, or, more indirectly, to simulate according to the transition probabilities.

You may instead choose to simulate according to the modified process needed for problem 5 below – you should still be able to answer the other problems on the basis of the modified simulations.

Enclose all computer code used to answer the questions.

Hint: Use a smaller number of simulations, *i.e.* 1000, when testing your code.

In (3–7) below, explain how you used the simulations to obtain estimates, and compare with the analytical results.

3. Find the expected number of rolls that are needed to get five-of-a-kind for the first time. Also find the expected number of rolls to get three-of-a-kind (or more than three equal faces) for the first time, and the expected number of additional rolls needed to get five-of-a-kind if you have already saved three-of-a-kind (and the two other dice differ).

Then use the simulations to estimate these expected numbers, and also to estimate the standard deviation of the number and to estimate the standard deviation (*i.e.*, the uncertainty) of the estimate of the number (previous standard deviation divided by the square root of the number of simulations).

4. Find the probability that three-of-a-kind appears (and the two other dice differ) in the sequence of rolls. Verify by the simulations.

5. Assume that instead of remaining in state 5 when five-of-a-kind is obtained for the first time, all five dice are re-rolled. Modify the transition probabilities accordingly, and find the limiting probability distribution. Verify by estimating the probabilities from the simulations.

6. Find the probability that five-of-a-kind is obtained in three or fewer rolls (one of the tasks in *Yahtzee*). If you try this task, and obtain two-of-a-kind in the first roll, should you be happy or sad? Verify by simulations.

7. Let T be the number of rolls when five-of-a-kind occurs for the first time. Make a graph illustrating $P(T \leq n)$. In the same figure, show three graphs of estimates of $P(T \leq n)$ based on the 100 first, the 1000 first, and all 10000 simulations.