

# Reward Variance in Markov Chains: A Computational Approach

Tom Verhoeff \*

May 2004

## Abstract

We consider the variance of the reward until absorption in a Markov chain. This variance is usually calculated from the second moment (expectation of the square). We present a direct system of equations for the variance, involving the first moment (expectation) but not the second moment. This method is numerically superior to the calculation from the second moment.

## 1 Introduction

Consider the following problem. A spider is located on the ceiling of a cubic room. Each day it travels from one face to a neighboring face, crossing a single edge. The spider randomly chooses among the four neighboring faces, with uniform probability. Successive choices are independent. What is the expected number of days for the spider to reach the floor? And, in particular, what is the corresponding variance?

This problem can be modeled as a Markov chain, where each transition contributes a fixed ‘reward’ of one day. The ceiling is the initial state, the walls are other transient states, and the floor is an absorbing state. The problem can then be reformulated as finding the expected reward until absorption, and its variance.

We present, what we believe to be, a new method for calculating the variance in the reward until absorption. We came to this method when analyzing strategies for playing a solitaire version of the dice game Yahtzee.

Markov chains are not only useful for analyzing puzzles and games of chance, but also play a prominent role in economics and engineering. Attention is often focused on the expected reward (or cost). However, in practice, the variance is also important, because it relates to risk and buffer capacity needed for handling the swings around the expected value. For instance, it can be more economical to aim for a suboptimal expected value in favor of a lower variance, because this improves the predictability of a budget. The Dutch government only recently decided [1] that its goal in addressing the traffic jam problem would no longer be a reduction of the expected travel time, but rather a reduction in the variability of travel time to improve predictability even if that means longer travel times.

---

\*Technische Universiteit Eindhoven, Fac. of Math. & CS, P.O. Box 513, 5600 MB EINDHOVEN, The Netherlands

In the remainder of this section, we explain some general concepts and results from probability theory and our notations. In Section 2, we do the same for Markov chains. Section 3 contains our new result. The spider in the cubic room is a running example. Finally, Section 4 concludes the article.

## 1.1 Probability Theory

For an introduction to probability theory see for instance [2]. We summarize the concepts needed for understanding this article.

A sample space is a set  $\Omega$  of (mutually exclusive) sample points or outcomes. To model a stochastic experiment, each outcome  $s \in \Omega$  is assigned a probability  $P.s$ , measuring the likelihood that  $s$  occurs. We have  $0 \leq P.s \leq 1$  and  $\sum_{s \in \Omega} P.s = 1$ . The pair  $(\Omega, P)$  is also called a probability space.

For example, the set of six faces (ceiling, four walls, and floor) of the cubic room forms a sample space. Assuming the spider is on the ceiling, let  $P.s$  be the probability that the spider will be on face  $s$  the next day. We then have  $P.s = \frac{1}{4}$  if  $s$  is one of the four walls, and  $P.s = 0$  if  $s$  is the ceiling or the floor.

Given a probability space  $(\Omega, P)$ , a random variable  $X$  is a function  $\Omega \rightarrow \mathbb{R}$ , where  $X.s$  is the value of  $s \in \Omega$ . The expectation (first moment)  $\mathcal{E}[X]$  of this random variable  $X$  is given by

$$\mathcal{E}[X] = \sum_{s \in \Omega} P.s * X.s . \quad (1)$$

It is also called the mean, and it captures the central tendency of the random variable. More generally,  $\mathcal{E}[X^k]$  is the  $k$ -th moment of  $X$ . We will write  $\mathcal{E}_\Omega[X]$ ,  $\mathcal{E}_{\Omega, P}[X]$ ,  $\mathcal{E}_s[X.s]$ , or  $\mathcal{E}_{s \in \Omega}[X.s]$  to make the relevant probability space and random variable more explicit.

For example, a roll of a fair die can be modeled by

- the sample space  $\Omega$  consisting of the six faces of a cube,
- the probability  $P.s = \frac{1}{6}$  for each face  $s$  to appear on top, and
- a function  $D$  associating a unique value in the range 1 through 6 with each face.

For the expectation of  $D$  we have  $\mathcal{E}[D] = \sum_{s \in \Omega} \frac{1}{6} * D.s = \sum_{k=1}^6 \frac{1}{6} * k = 3.5$ .

An important property of the expectation is its linearity. For constants  $c$  and  $d$  (whose value does not depend on the sample space) and random variables  $X$  and  $Y$  on the same probability space,  $X + cY + d$  is also a random variable on that probability space, having expectation

$$\mathcal{E}[X + cY + d] = \mathcal{E}[X] + c\mathcal{E}[Y] + d . \quad (2)$$

Note that, in general,  $\mathcal{E}[X * Y] = \mathcal{E}[X] * \mathcal{E}[Y]$  does *not* hold, but it does hold if  $X$  and  $Y$  are *independent*. In particular,  $\mathcal{E}[X^2]$  and  $\mathcal{E}[X]^2$  are not necessarily equal. For instance, for the fair die we have  $\mathcal{E}[D^2] = \sum_{k=1}^6 \frac{1}{6} * k^2 = 15\frac{1}{6}$ , whereas  $\mathcal{E}[D]^2 = 3.5^2 = 12\frac{1}{4}$ .

The amount of variability of a random variable  $X$  around its mean can be measured by its variance  $\mathcal{V}[X]$ , defined by

$$\mathcal{V}[X] = \mathcal{E}[(X - \mathcal{E}[X])^2] . \quad (3)$$

The variance can be expressed in terms of the first and second moment, as the following calculation<sup>1</sup> shows:

$$\begin{aligned}
& \mathcal{V}[X] \\
&= \quad \{ \text{definition of } \mathcal{V} \} \\
& \quad \mathcal{E} [(X - \mathcal{E}[X])^2] \\
&= \quad \{ \text{algebra} \} \\
& \quad \mathcal{E} [X^2 - 2X\mathcal{E}[X] + \mathcal{E}[X]^2] \\
&= \quad \{ \text{linearity of } \mathcal{E}, \text{ observing that } \mathcal{E}[X] \text{ is a constant} \} \\
& \quad \mathcal{E} [X^2] - 2\mathcal{E}[X]\mathcal{E}[X] + \mathcal{E}[X]^2 \\
&= \quad \{ \text{algebra} \} \\
& \quad \mathcal{E} [X^2] - \mathcal{E}[X]^2
\end{aligned} \tag{4}$$

Numerically, however, the expression  $\mathcal{E} [X^2] - \mathcal{E}[X]^2$  is inferior to (3), because of the risk of cancelation as illustrated by the following example. Consider a random variable  $X$  which takes on one of two values  $a = 999$  and  $b = 1001$  with probabilities  $p = 0.1$  and  $q = 0.9$  respectively. We then have  $\mathcal{V}[X] = pq(a - b)^2 = 0.36$ . When evaluating (3) and (4) using the IEEE-754 single format for floating-point numbers, the following dramatic results are obtained

$$\begin{aligned}
\mathcal{E} [(X - \mathcal{E}[X])^2] &= 0.36000\dots \\
\mathcal{E} [X^2] - \mathcal{E}[X]^2 &= 0.50660\dots
\end{aligned}$$

The latter value is 40% too high!

Note that  $\mathcal{V}[X]$  is expressed in the square of the units of  $X$ . If  $X$  is expressed in m/s, then  $\mathcal{V}[X]$  is expressed in m<sup>2</sup>/s<sup>2</sup>. The standard deviation  $\sigma$  measures the variability in the same units as the random variable by taking the square root of the variance:

$$\sigma[X] = \sqrt{\mathcal{V}[X]}. \tag{5}$$

In general, the variance does not satisfy a linearity property like (2), but we do have:

$$\begin{aligned}
& \mathcal{V}[cX + d] \\
&= \quad \{ \text{definition of } \mathcal{V} \} \\
& \quad \mathcal{E} [(cX + d - \mathcal{E}[cX + d])^2] \\
&= \quad \{ \text{linearity of } \mathcal{E} \} \\
& \quad \mathcal{E} [(cX + d - (c\mathcal{E}[X] + d))^2] \\
&= \quad \{ \text{algebra} \} \\
& \quad \mathcal{E} [c^2(X - \mathcal{E}[X])^2] \\
&= \quad \{ \text{linearity of } \mathcal{E}, \text{ definition of } \mathcal{V} \} \\
& \quad c^2\mathcal{V}[X]
\end{aligned} \tag{6}$$

---

<sup>1</sup>This way of recording calculations is due to Feijen, see [3]. The hint in braces explains why the relationship shown on the left holds between the expressions above and below it.

## 2 Markov Chains

For an introduction to Markov chains see for instance [4, 5, 6]. We summarize the concepts needed for understanding this article.

A Markov chain models a stochastic process, where an experiment with outcomes in a sample space  $\Omega$  is repeated and where the probability distribution for the outcome of each experiment can depend on the outcome of the preceding experiment.

It is often more convenient to view the sample space  $\Omega$  of a Markov chain as a state space. At each time step, the system is in a state  $s \in \Omega$ . The transition from state  $s$  to state  $t$  in the next time step occurs with probability  $p.s.t$ . For all  $s \in \Omega$ , the transition probabilities  $p.s.t$  satisfy

$$\sum_{t \in \Omega} p.s.t = 1. \quad (7)$$

Let us define a Markov chain for the spider in the cubic room. The state space consists of the six faces where the spider can be located. We abbreviate them as C (Ceiling),  $W_i$  (Wall,  $0 \leq i < 4$ ), and F (Floor). The initial state is C. The transition probabilities are given by

$p.s.t$	C	$W_0$	$W_1$	$W_2$	$W_3$	F
C	0	1/4	1/4	1/4	1/4	0
$W_0$	1/4	0	1/4	0	1/4	1/4
$W_1$	1/4	1/4	0	1/4	0	1/4
$W_2$	1/4	0	1/4	0	1/4	1/4
$W_3$	1/4	1/4	0	1/4	0	1/4
F	0	1/4	1/4	1/4	1/4	0

(8)

Note that the matrix of transition probabilities is symmetric. However, this is not generally the case.

We are interested only in the spider's behavior until it reaches the floor. Therefore, the transition probabilities from the floor, given in the last row of (8), are irrelevant. We might as well make the spider stay on the floor:  $p.F.F = 1$  and  $p.F.s = 0$  for  $s \neq F$ .

Furthermore, the four states  $W_i$  are *equivalent* in view of what they offer for the future. We can collapse them into a single state  $W$ . This yields the following simplified Markov chain, which is also pictured in Fig. 1.

$p.s.t$	C	W	F
C	0	1	0
W	1/4	1/2	1/4
F	0	0	1

(9)

### 2.1 Walks

We now turn to sequences of successive state transitions. A nonempty sequence  $s_0 s_1 \cdots s_n$  of  $n+1$  states  $s_i \in \Omega$  is called a walk of length  $n$  from  $s_0$  to  $s_n$ . This sequence represents  $n$  successive state transitions  $s_{i-1} \rightarrow s_i$ . The length of a walk is the number of state transitions it involves. Note that for a walk  $w$  of

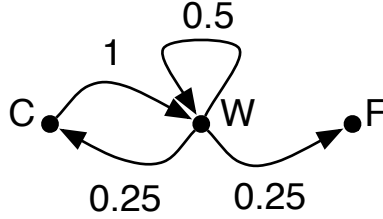


Figure 1: Simplified Markov chain for spider in cubic room (transitions with probability 0 not drawn)

length zero we have  $w = s_0$ . We denote catenation of sequences by juxtaposition. Variables  $s$  and  $t$  range over states, whereas variables  $v$  and  $w$  range over sequences of states.

The successive experiments in a Markov chain are independent and, hence, the transition probabilities can be multiplied. Thus, the probability  $P.sw$  of a walk  $sw$ , given that it starts in  $s$ , satisfies the recurrence:

$$P.s = 1, \quad (10)$$

$$P.stv = p.s.t * P.tv. \quad (11)$$

We are interested in the analysis of walks until their first arrival in some nonempty subset  $A \subseteq \Omega$ . In the example of the spider, our observation of the walk ends when the spider arrives at the floor for the first time, that is,  $A = \{F\}$ . In this article, we will use ‘first arrival in  $A$ ’ and ‘absorption in  $A$ ’ interchangeably.<sup>2</sup>

For nonempty  $A \subseteq \Omega$ , let  $W_A.s$  be the set of walks starting in  $s \in \Omega$  until first arrival in  $A$ . We will leave out the subscript  $A$ , because  $A$  will not vary. If  $s \in A$ , then the walk ends immediately. If  $s \notin A$ , then such walks involve at least one transition to some state  $t$ , from where the walk proceeds until first arrival in  $A$ . Formally:

$$W.s = \{s\} \quad \text{if } s \in A, \quad (12)$$

$$W.s = \bigcup_{t \in \Omega} \{stv \mid tv \in W.t\} \quad \text{if } s \notin A. \quad (13)$$

For this article, we make one important assumption about  $A$ : starting in state  $s$ , the probability that a walk eventually ends in  $A$  is 1. That is, for all  $s$ , we have

$$\sum_{sw \in W.s} P.sw = 1. \quad (14)$$

Hence,  $W.s$  is a sample space, and  $P.sw$  for  $w \in W.s$  is a probability function on it.

## 2.2 Rewards

In the example of the spider, we are interested in the expected *length* of a walk until absorption. More generally, we associate with each transition  $s \rightarrow t$  a

<sup>2</sup>Strictly speaking, the transition probabilities should be redefined to make states  $s \in A$  truly absorbing:  $p.s.s = 1$  and  $p.s.t = 0$  for  $t \neq s$ .

reward  $r.s.t$ . If we are only interested in the walk length, then we take  $r.s.t = 1$ . The (total) reward  $R.sw$  of walk  $sw$  is defined inductively by

$$R.s = 0 , \quad (15)$$

$$R.stv = r.s.t + R.tv . \quad (16)$$

That is, successive rewards are independent and, hence, are simply added.

The reward  $R.sw$  obtained when starting in state  $s$  and walking until first arrival in  $A$  is a random variable on the probability space  $(W.s, P.sw)$ . In the remainder of this section, we deal with the expectation  $\mathcal{E}_{W.s}[R.sw]$ . Even though this is a well-known result, we have included it here in detail for two reasons:

- We have not seen it treated in this way elsewhere.
- The treatment of the variance follows the same pattern.

First, however, we derive a pair of convenient properties for the expectation  $\mathcal{E}_{W.s}[X]$  of an arbitrary random variable  $X$  on walks from  $s$  until first arrival in  $A$ .

**Property** For  $s \in A$ :

$$\begin{aligned} & \mathcal{E}_{W.s}[X] \\ = & \quad \{ \text{definition of } \mathcal{E} \} \\ & \sum_{w \in W.s} P.w * X.w \\ = & \quad \{ W.s = \{ s \}, \text{ because } s \in A \} \\ & P.s * X.s \\ = & \quad \{ \text{by definition } P.s = 1 \} \\ & X.s \end{aligned}$$

**Property** (conditioning on the first step toward absorption) For  $s \notin A$ :

$$\begin{aligned} & \mathcal{E}_{W.s}[X] \\ = & \quad \{ \text{definition of } \mathcal{E} \} \\ & \sum_{w \in W.s} P.w * X.w \\ = & \quad \{ \text{write } w = sv \text{ for } v \in W.t, \text{ because } s \notin A, \text{ cf. (13)} \} \\ & \sum_{t \in \Omega} \sum_{v \in W.t} P.sv * X.sv \\ = & \quad \{ \text{recurrence for walk probability: } P.sv = p.s.t * P.v \text{ for } v \in W.t \} \\ & \sum_{t \in \Omega} \sum_{v \in W.t} p.s.t * P.v * X.sv \\ = & \quad \{ \text{distribute } p.s.t * \text{ outside } \sum_v, \text{ using that } p.s.t \text{ does not depend on } v \} \\ & \sum_{t \in \Omega} p.s.t * \sum_{v \in W.t} P.v * X.sv \\ = & \quad \{ \text{definition of } \mathcal{E} \} \\ & \sum_{t \in \Omega} p.s.t * \mathcal{E}_{v \in W.t}[X.sv] \\ = & \quad \{ \text{definition of } \mathcal{E} \} \\ & \mathcal{E}_{t \in \Omega} [ \mathcal{E}_{v \in W.t}[X.sv] ] \end{aligned}$$

Concerning the expected reward  $\mathcal{E}_{W.s}[R]$  on a walk from state  $s$  until absorption in  $A$ , we can now calculate the following (well-known) result.

For  $s \in A$ :

$$\begin{aligned} & \mathcal{E}_{W.s}[R] \\ = & \quad \{ \text{property above, using } s \in A \} \\ & R.s \\ = & \quad \{ \text{definition of } R \} \\ & 0 \end{aligned}$$

For  $s \notin A$ :

$$\begin{aligned} & \mathcal{E}_{W.s}[R] \\ = & \quad \{ \text{conditioning on first state } t \text{ after state } s, \text{ using } s \notin A \} \\ & \mathcal{E}_{t \in \Omega} [\mathcal{E}_{v \in W.t}[R.sv]] \\ = & \quad \{ \text{recurrence for walk reward: } R.sv = r.s.t + R.v \text{ for } v \in W.t \} \\ & \mathcal{E}_{t \in \Omega} [\mathcal{E}_{v \in W.t}[r.s.t + R.v]] \\ = & \quad \{ \text{linearity of expectation, using that } r.s.t \text{ is independent of } v \} \\ & \mathcal{E}_{t \in \Omega} [r.s.t + \mathcal{E}_{v \in W.t}[R.v]] \\ = & \quad \{ \text{simplify notation} \} \\ & \mathcal{E}_{t \in \Omega} [r.s.t + \mathcal{E}_{W.t}[R]] \end{aligned}$$

This gives us a system of linear equations with as unknowns  $\mu_s = \mathcal{E}_{W.s}[R]$  for each  $s \in \Omega$ :

$$\mu_s = \sum_{t \in \Omega} p.s.t * (r.s.t + \mu_t) . \quad (17)$$

If  $r.s.t = 1$  (measuring the length of a walk), then this can be simplified to

$$\mu_s = 1 + \sum_{t \in \Omega} p.s.t * \mu_t . \quad (18)$$

Consider the three-state Markov chain (9) for the spider in the cubic room. We take  $A = \{ \mathbf{F} \}$ . According to (18), the system of equations for the expected walk lengths  $\mu_s = \mathcal{E}_{W.s}[R]$  from face  $s$  to absorption on the floor is:

$$\begin{aligned} \mu_{\mathbf{C}} &= 1 + \mu_{\mathbf{W}} \\ \mu_{\mathbf{W}} &= 1 + \left( \frac{1}{4} \mu_{\mathbf{C}} + \frac{1}{2} \mu_{\mathbf{W}} + \frac{1}{4} \mu_{\mathbf{F}} \right) \\ \mu_{\mathbf{F}} &= 0 \end{aligned}$$

This has as solution:

$$\begin{aligned} \mu_{\mathbf{C}} &= 6 \\ \mu_{\mathbf{W}} &= 5 \\ \mu_{\mathbf{F}} &= 0 \end{aligned}$$

Thus, when starting on the ceiling, the expected duration for the spider to hit the floor is exactly 6 days.

### 3 Reward Variance

We now turn to the variance  $\mathcal{V}_{W.s}[R]$  in the reward on a walk from state  $s$  to absorption in  $A$ . For  $s \in A$ , we calculate

$$\begin{aligned}
& \mathcal{V}_{W.s}[R] \\
= & \quad \{ \text{definition of } \mathcal{V} \} \\
& \mathcal{E}_{W.s} \left[ (R - \mathcal{E}_{W.s}[R])^2 \right] \\
= & \quad \{ \mathcal{E}_{W.s}[R] = 0, \text{ because } s \in A \text{ and } R.s = 0 \} \\
& \mathcal{E}_{W.s} [R^2] \\
= & \quad \{ \text{property of } \mathcal{E}_{W.s}, \text{ using } s \in A \text{ and } R.s = 0 \} \\
& 0
\end{aligned}$$

Before tackling  $s \notin A$ , we observe that for constant  $c$  and random variable  $X$ :

$$\mathcal{V}[X] = \mathcal{V}[c + X] = \mathcal{E}[(c + X)^2] - \mathcal{E}[c + X]^2 = \mathcal{E}[(c + X)^2] - (c + \mathcal{E}[X])^2 .$$

And, hence,

$$\mathcal{E}[(c + X)^2] = (c + \mathcal{E}[X])^2 + \mathcal{V}[X] . \quad (19)$$

Finally, for  $s \notin A$ , we calculate

$$\begin{aligned}
& \mathcal{V}_{W.s}[R] \\
= & \quad \{ \text{definition of } \mathcal{V} \} \\
& \mathcal{E}_{W.s} \left[ (R - \mathcal{E}_{W.s}[R])^2 \right] \\
= & \quad \{ \text{conditioning on first state } t \text{ after state } s, \text{ using } s \notin A \} \\
& \mathcal{E}_{t \in \Omega} \left[ \mathcal{E}_{v \in W.t} \left[ (R.s.v - \mathcal{E}_{W.s}[R])^2 \right] \right] \\
= & \quad \{ \text{recurrence for reward: } R.s.v = r.s.t + R.v \text{ for } v \in W.t \} \\
& \mathcal{E}_{t \in \Omega} \left[ \mathcal{E}_{v \in W.t} \left[ (r.s.t + R.v - \mathcal{E}_{W.s}[R])^2 \right] \right] \\
= & \quad \{ (19), \text{ using that } r.s.t - \mathcal{E}_{W.s}[R] \text{ does not depend on } v \} \\
& \mathcal{E}_{t \in \Omega} \left[ (r.s.t + \mathcal{E}_{W.t}[R] - \mathcal{E}_{W.s}[R])^2 + \mathcal{V}_{W.t}[R] \right]
\end{aligned}$$

This yields a system of linear equations with as unknowns  $\sigma_s^2 = \mathcal{V}_{W.s}[R]$  for each  $s \in \Omega$ , involving  $\mu_s = \mathcal{E}_{W.s}[R]$  as parameters:

$$\sigma_s^2 = \sum_{t \in \Omega} p.s.t * ((r.s.t + \mu_t - \mu_s)^2 + \sigma_t^2) . \quad (20)$$

My earlier derivations of this result were quite messy. The derivation presented here is kept simple by using (19).

When applying (20) to the example of the spider, we obtain as system of equations for the variance in walk length  $\sigma_s^2 = \mathcal{V}_{W.s}[R]$  from face  $s$  to absorption on the floor:

$$\begin{aligned}
\sigma_C^2 &= (1 + \mu_W - \mu_C)^2 + \sigma_W^2 \\
\sigma_W^2 &= \frac{1}{4} \left( (1 + \mu_C - \mu_W)^2 + \sigma_C^2 \right) + \frac{1}{2} \left( (1 + \mu_W - \mu_W)^2 + \sigma_W^2 \right) + \frac{1}{4} \left( (1 + \mu_F - \mu_W)^2 + \sigma_F^2 \right) \\
\sigma_F^2 &= 0
\end{aligned}$$



The first equation yields  $\sigma_{\text{C}}^2 = \sigma_{\text{W}}^2$ , because  $\mu_{\text{C}} = 1 + \mu_{\text{W}}$ . This is understandable, since the probability for a transition from Ceiling to Wall equals 1 and, hence, there is no variability on this part of the walk.

The solution to the equation system is:

$$\begin{aligned}\sigma_{\text{C}}^2 &= 22 \\ \sigma_{\text{W}}^2 &= 22 \\ \sigma_{\text{F}}^2 &= 0\end{aligned}$$

Hence, the standard deviation in the walk length from the ceiling to the floor is  $\sqrt{22} \approx 4.69$ . This is considerable compared to the expectation  $\mu_{\text{C}} = 6$ . It means<sup>3</sup> that almost two million simulation runs are needed to estimate the expectation with an accuracy of 0.01 and a confidence level within  $3\sigma$ .

The equations (20) can be generalized for covariance. Given two random variables  $X$  and  $Y$  on the same probability space, their *covariance*  $\text{Cov}.X.Y$  is defined by

$$\text{Cov}.X.Y = \mathcal{E}[(X - \mathcal{E}[X])(Y - \mathcal{E}[Y])] . \quad (21)$$

Note that  $\text{Cov}.X.X = \mathcal{V}[X]$ . Similar to (4), one can derive

$$\text{Cov}.X.Y = \mathcal{E}[XY] - \mathcal{E}[X]\mathcal{E}[Y] \quad (22)$$

Hence, we have (compare this to (2))

$$\mathcal{V}[X + cY + d] = \mathcal{V}[X] + c^2\mathcal{V}[Y] + 2c\text{Cov}.X.Y \quad (23)$$

Now consider two reward functions  $r$  and  $q$  on the same Markov chain. These induce the reward functions  $R$  and  $Q$  on walks. The covariances  $z_s$  between  $R$  and  $Q$  on walks starting in state  $s$  until absorption in  $A$  satisfy

$$z_s = \sum_{t \in \Omega} p.s.t * ((r.s.t + \mu_t - \mu_s)(q.s.t + \nu_t - \nu_s) + z_t) . \quad (24)$$

where  $\mu_s = \mathcal{E}_{W.s}[R]$  and  $\nu_s = \mathcal{E}_{W.s}[Q]$ .

## 4 Conclusion

We have presented a new system of equations (20) for determining the variance of the reward until absorption in a Markov chain. Compared to the standard approach using the second moment, these equations have a lower risk of cancellation when solved numerically.

We applied this technique in our analysis of the dice game Yahtzee [7, 8]. The Markov chain for solitaire Yahtzee involves close to  $10^9$  states. Because it has no cycles, the resulting equations for expectation and variance are recurrence equations. These can be solved simply by backward substitution and dynamic programming. The analysis yields the optimal expected final score and its variance, also broken down by the individual scoring categories and their covariances. Cremers [9] extended the analysis to the beating of high scores.

I would like to acknowledge the helpful comments from Onno Boxma and Ivo Adan on an earlier version of this article.

<sup>3</sup>The variance in the average taken over  $n$  simulation runs equals the variance in a single run divided by  $n$ , and hence, the standard deviation in the average taken over  $n$  simulation runs equals  $\sigma/\sqrt{n}$ . The number of runs needs to be at least  $(3\sqrt{22}/0.01)^2$ .

## References

- [1] Karla Peijs. Nota Mobiliteit. Ministerie van Verkeer en Waterstaat, 2004. On-line: <http://www.notamobiliteit.nl/>.
- [2] W. Feller. *Introduction to Probability Theory and Its Applications*, volume 1. Wiley, London, third edition, 1970.
- [3] E. W. Dijkstra and C. S. Scholten. *Predicate Calculus and Program Semantics*. Springer, Berlin, 1990.
- [4] John G. Kemeny and J. Laurie Snell. *Finite Markov Chains*. Springer, Berlin, 1976.
- [5] Edward P. C. Kao. *An Introduction to Stochastic Processes*. Duxbury Press, Bonn, 1997.
- [6] Sheldon M. Ross. *Introduction to Probability Models*. Academic Press, San Diego, eighth edition, 2003.
- [7] T. Verhoeff. Optimal solitaire Yahtzee advisor and Yahtzee proficiency test, 1999. On-line since July 1999: <http://www.win.tue.nl/~wstomv/misc/yahtzee/>.
- [8] H. van Maanen. Ook Yahtzee bezwijkt (Eng.: “Also Yahtzee succumbs”). *Het Parool*, 1999. Dutch daily newspaper, 9 oktober 1999.
- [9] C. Cremers. How best to beat high scores in Yahtzee: A caching structure for evaluating large recurrent functions. Master’s thesis, Fac. of Math. and CS, Technische Universiteit Eindhoven, The Netherlands, 2002.