How long does a tennis game last?

M. Ferrante*, G. Fonseca** and S. Pontarollo*

*Dip. di Matematica "Tullio Levi-Civita", Università di Padova, Via Trieste 63, 35121-Padova, Italy
email: ferrante@math.unipd.it, spontaro@math.unipd.it
** Dip. di Sc. Econom. e Stat., Univ. di Udine, via Tomadini, 30/A, 33100-Udine, Italy
email: giovanni.fonseca@uniud.it

Abstract

In this paper we present a generalisation of previously considered Markovian models for Tennis that overcome the assumption that the points played are i.i.d and includes the time into the model. Firstly we postulate that in any game there are two different situations: the first 6 points and the, possible, additional points after the first deuce, with different winning probabilities. Then we assume that the duration of any point is distributed with an exponential random time. We are able to compute the law of the (random) duration of a game in this more general setting.

1 Introduction

Markovian framework is particularly suitable to describe the evolution of a tennis match. The usual assumption is that the probability that a player wins one point is independent of the previous points and constant during the match. Under these hypotheses the score of a game, set and match can be described by a set of nested homogeneous Markov chains. Hence, theoretical results concerning winning probabilities and mean duration of a game, set and match can be easily obtained. A complete account on this approach can be found in Klaassen and Magnus (2014).

Anyway, some authors criticise the assumption that the point winning probability is constant along the match and independent of the previous points played, see e.g. Klaassen and Magnus (2001). In particular, it is pointed out that playing decisive points, i.e. points after a deuce score, modifies players attitude and this reflects heavily on the probability to win these points.

In Carrari *et al.* (2017), we propose a modification of the model at the game's level. Indeed, we assume that during any game there are two different situations: the first points and the, possible, additional points played after the (30,30) score that in our model coincide with the "Deuce". Under this hypothesis, following the approach used in Ferrante and Fonseca (2014), we computed the winning probabilities and the expected number of points played in a game.

In the present work we include in the model the time needed to play a single point. The aim of such a modification is to obtain the computation of the expected length of a match in terms of actual time and not just as number of points. Indeed there is a concern about the length of tennis matches and several modification to the game rules are nowadays proposed in order to fix the length of a match or at least to avoid too long matches. In Section 2. we present the model, in Section 3. we compute the game winning probabilities and in Section 4. we obtain the expected length of a game.

2 The continuous time model

In this paper we model the tennis game as a continuous-time Markov chain (see Norris (1998) for a complete account on this topic). We define the state space S of the chain, which collects all the possible scores in the game, and the generator matrix Q on S. In order to determine the matrix Q, we define independently the transition matrix of the associated Jump chain, which is a discrete-time Markov chain, and the exponential holding times.

The transition matrix of the associated Jump chain follows the model defined in Carrari *et al.* (2017) for a discrete-time Markov chain of the tennis. The classical assumptions previously considered in the literature (see e.g Newton and Keller (2005)) were that the probability to win any point by the player on service was independent of the previous points and constant during the game. In Carrari *et al.* (2017) we assume that p, the probability to win a point, does not remain the same during the game. As empirical data on the matches confirm, the estimated winning probability of the first 6 points of a game is different from that of the additional played points from the "Deuce" on. For this reason, we consider a second parameter $\bar{p}$, that describe this part of the game and, to avoid trivial cases, we assume that both p and $\bar{p}$ belong to $(0,1)$.

Regarding the holding times, it is not easy to find in the literature data sets to test different scenarios that better describe the true course of a tennis game (see Morante and Brotherhooc (2007) for some statistics on the duration). For this reason in this paper we assume that all the holding times have the same distribution, i.e. share the same rate λ of their Exponential Laws.

Let us now define precisely our model: the state space is the set $S = \{1,2,\ldots,17\}$ which describes the score of a game as defined in Table 1.

Score	(0,0)	(15,0)	(0,15)	(30,0)	(15,15)	(0,30)	(40,0)	(30,15)	(15,30)
State	1	2	3	4	5	6	7	8	9
Score	(0,40)	(40,15)	(15,40)	*Deuce*	Adv_A	Adv_B	Win_A	Win_B	
State	10	11	12	13	14	15	16	17	

Table 1: Scores and corresponding states used in equations

Note that in the present model the scores $(30,30)$ and *Deuce* are represented by the single state 13, since they share the same mathematical properties, as it happens to the pairs $(40,30)$–Adv_A and $(30,40)$–Adv_B. The graph representing the transition probabilities of the Jump process is presented in Fig. 1, where $q = 1-p$ and $\bar{q} = 1-\bar{p}$.

By our construction, we define the transition rates in the generator matrix Q by $\lambda_1 = \lambda p$ and $\lambda_2 = \lambda q$ (see Norris (1998)) and in Fig. 2 we present the graph of the continuous time Markov chain describing a tennis game. Note that the expected length of a point is equal to $1/\lambda$ and it does not depend on who is the winner of the point.

From the graph it is immediate to write down the generator matrix $Q = (q_{ij})_{i,j\in S}$ and to prove that the states

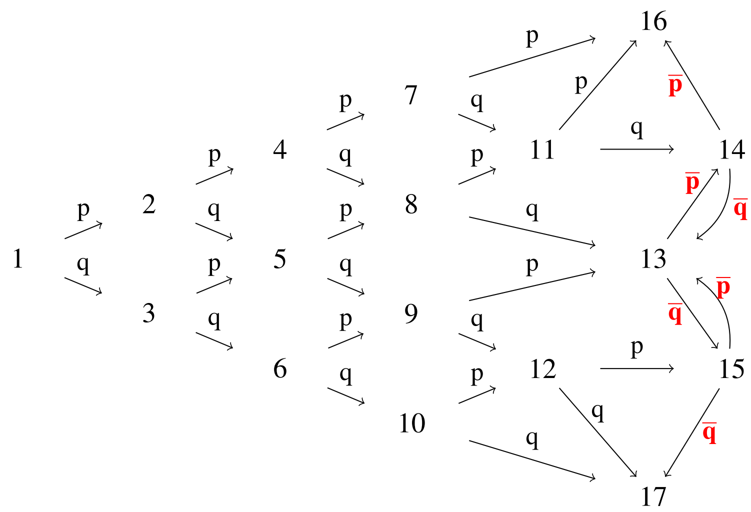

Figure 1: Graph of the Jump Markov chain

16 and 17 are absorbing, while all the other states are transient.

$$Q = \begin{bmatrix}
-\lambda & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\lambda & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -\lambda & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & \lambda_1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & 0 & \lambda_2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_2 & \lambda_1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & \lambda_1 & 0 & 0 & 0 & 0 & \lambda_2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_2 & 0 & \lambda_1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & 0 & \lambda_2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\bar{\lambda} & \bar{\lambda}_1 & \bar{\lambda}_2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{\lambda}_2 & -\bar{\lambda} & 0 & \bar{\lambda}_1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{\lambda}_1 & 0 & -\bar{\lambda} & 0 & \bar{\lambda}_2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}$$

In order to compute the winning probabilities, we need to determine the absorption probabilities in states 16 and 17 for the transition matrix of the Jump process, while to investigate the distribution of the expected length of a game, we need to evaluate the exponential matrix of Q, which is in general not very simple.

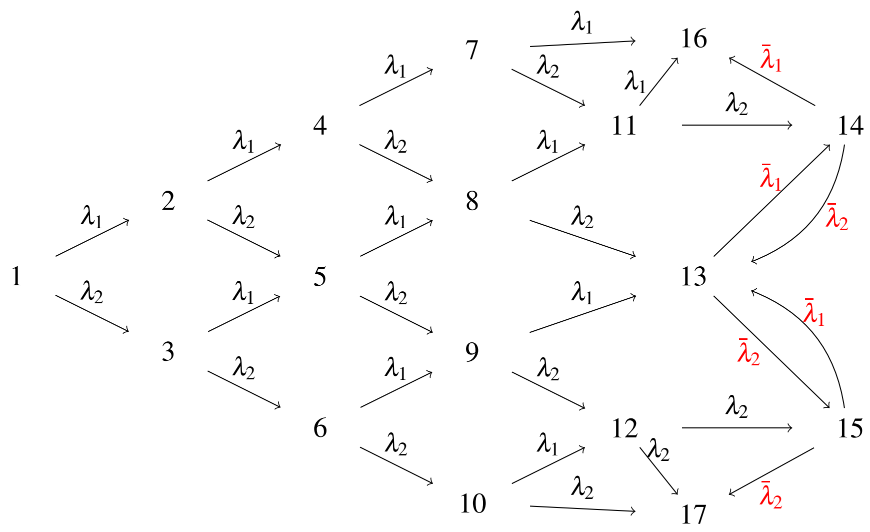

Figure 2: Graph of the continuous time Markov chain describing a tennis game, with its transition rates.

3 Winning probabilities

In this section we recall some of the result proved in Carrari *et al.* (2017). The transition matrix of the Jump process is

$$P = \begin{bmatrix}
0 & p & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & p & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & p & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & p & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & p & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p & q & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q & 0 & 0 & 0 & 0 & p & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p & 0 & q & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q & p & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p & 0 & 0 & 0 & 0 & q \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & q & 0 & p & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & p & 0 & q \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{\mathbf{p}} & \bar{\mathbf{q}} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{\mathbf{q}} & 0 & 0 & \bar{\mathbf{p}} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{\mathbf{p}} & 0 & 0 & 0 & \bar{\mathbf{q}} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}$$

The winning probability of the game for the player A on service, denoted by h_1, coincides with the absorption probability in the state 16 of the previous Markov chain starting from 1, which can be obtained (see e.g.

Norris (1998)) as the minimal, non negative solution of the linear system

$$h_i = \sum_{j\in S} p_{ij} h_j \qquad \text{for } 1 \le i \le 15\ ,\ h_{16} = 1\ ,\ h_{17} = 0\ .$$

The solution can be easily calculated and we obtain that

$$h_1 = p^2 \left[5p^2 - 4p^3 + 4(p-1)^2 p\bar{p} - \frac{2(p-1)^2 \bar{p}^2 (p(4\bar{p}-2) - 2\bar{p} - 3)}{2\bar{p}^2 - 2\bar{p} + 1} \right]\ .$$

Denoting by $G(p,\bar{p}) = h_1$, by A and B the two players, and by P^G_{xY} the probability that the player Y wins a game when X serves, thanks to the symmetry of the model we obtain that:

$$\begin{aligned} P^G_{aA} &= G(p_A, \bar{p}_A) \\ P^G_{aB} &= G(1-p_A, 1-\bar{p}_A) \\ P^G_{bB} &= G(p_B, \bar{p}_B) \\ P^G_{bA} &= G(1-p_B, 1-\bar{p}_B) \end{aligned} \tag{1}$$

Note that, since $P^G_{xX} + P^G_{xY} = 1$, $G(1-p_X, 1-\bar{p}_X) = 1 - G(p_X, \bar{p}_X)$ and that for $p_X = \bar{p}_X$, the previous probabilities coincides with those well known in the literature (see e.g. Newton and Keller (2005)). In Table 2 we report the values of G for increasing p and $\bar{p}$.

	$\bar{p}$									
p	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
0.1	0.001	0.004	0.011	0.021	0.034	0.049	0.062	0.071	0.078	0.081
0.2	0.011	0.022	0.043	0.076	0.119	0.165	0.204	0.233	0.252	0.263
0.3	0.040	0.061	0.099	0.158	0.234	0.312	0.378	0.425	0.455	0.472
0.4	0.102	0.132	0.185	0.264	0.363	0.464	0.549	0.607	0.643	0.663
0.5	0.206	0.242	0.302	0.391	0.500	0.609	0.697	0.758	0.794	0.812
0.6	0.357	0.392	0.451	0.535	0.636	0.736	0.815	0.868	0.898	0.913
0.7	0.545	0.575	0.622	0.688	0.766	0.842	0.901	0.939	0.960	0.969
0.8	0.748	0.767	0.795	0.835	0.881	0.924	0.957	0.978	0.989	0.993
0.9	0.922	0.929	0.938	0.951	0.965	0.979	0.989	0.995	0.998	0.999
1.0	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 2: Winning probabilities of a game

4 Duration of a game

In this section we evaluate the duration of a game and of a game given that the player on serve wins the game. By the Markovian structure of the present model, the problem we face is equivalent to evaluate the distribution of the absorption time to one of the states 16 and 17. These distributions are called in the literature Phase-type distributions (see e.g. Neuts (1981)) and we have explicit formulas for their densities and moments. The only drawback is that the computation of the density passes through the evaluation of the matrix exponential of a 16×16 matrix, which is usually not feasible. On the contrary, for the moments we only need to be able to evaluate the inverse of the same matrix and its powers.

4.1 The unconditioned case

Due to the difficulties described above, in this section we consider only the case where $\overline{p} = p$. Starting from Q, the generator matrix defined above, we can compute the distribution of the duration time of a game. Indeed, let T be equal to the matrix Q where the last two rows and two columns have been erased, that is

$$T = \begin{bmatrix}
-\lambda & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\lambda & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -\lambda & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & \lambda_2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 & 0 & \lambda_2 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_2 & \lambda_1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & \lambda_1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_2 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & 0 & 0 & \lambda_1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\lambda & \lambda_1 & \lambda_2 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & -\lambda & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 & 0 & -\lambda
\end{bmatrix}$$

and let $\mathbf{T}^0$ be a vector of length 15 where each entries is equal to the jumping time needed for each state (excluding Win_A and Win_B) to reach one of the two absorbing states, that is

$$\mathbf{T}^0 = (0,0,0,0,0,0,\lambda_1,0,0,\lambda_2,\lambda_1,\lambda_2,0,\lambda_1,\lambda_2)^\top .$$

Then, if $\boldsymbol{\alpha} = (1,0,\ldots,0)$, the density function of the duration time, denoted by $f_{p,\lambda}$, can be computed as

$$f_{p,\lambda} = \boldsymbol{\alpha} e^{Tt} \mathbf{T}^0$$

(see Neuts (1981) for the simple proof). Therefore, if we set

$$A_{p,\lambda}(t) = -3\lambda^2 p^4 t^2(\lambda t - 4) + 6\lambda^2 p^3 t^2(\lambda t - 4) - p^2\left(4\lambda^3 t^3 - 21\lambda^2 t^2 + 30\right) + p\left(\lambda^3 t^3 - 9\lambda^2 t^2 + 30\right) - 15 + \lambda^2 t^2$$

and

$$B_{p,\lambda}(t) = 15\sqrt{2}\left(2p^2 - 2p + 1\right)\left(e^{2\lambda t\sqrt{2(1-p)p}} - 1\right) ,$$

the time distribution is given by

$$f_{p,\lambda}(t) = \frac{\lambda}{24\sqrt{(1-p)p}} e^{-\lambda t\left(1+\sqrt{2(1-p)p}\right)} \left(B_{p,\lambda}(t) + \left(4\lambda t\sqrt{(1-p)p}\, e^{\lambda t\sqrt{2(1-p)p}}\right) A_{p,\lambda}(t)\right) .$$

In Figure 3. we plot the density $f_{p,\lambda}$ in the cases $p = 0.5, 0.65$ and 0.9, while $\lambda = 2$.

We are also able to evaluate the moments of these random times. Indeed, the expected time μ needed to win a game can be computed as $\boldsymbol{\alpha}(-T)^{-1}\mathbf{e}$, where $\mathbf{e} = (1,\ldots,1)^T$. Therefore, the average time is given by

$$\mu = \frac{4\left(-6p^6 + 18p^5 - 18p^4 + 6p^3 + p^2 - p + 1\right)}{\lambda\left(2p^2 - 2p + 1\right)} .$$

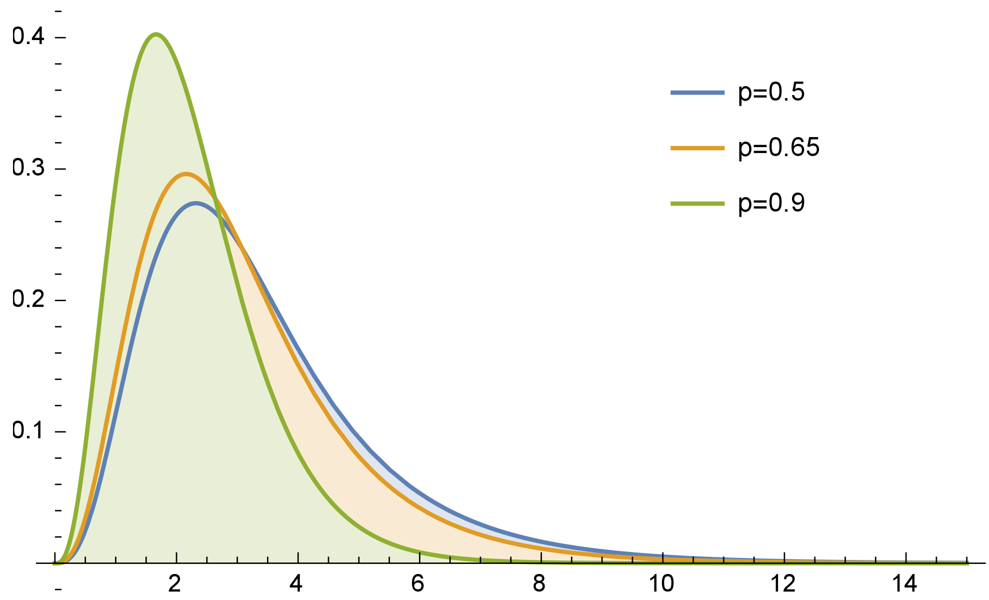

Figure 3: Graph of the distribution of a game duration time for three different values of p, and for $\lambda = 2$.

Finally the moment of order two, computed as $\mu^2 = 2\alpha(-T)^{-2}\mathbf{e}$, allows us to obtain the variance

$$\sigma^2 = \frac{4}{\lambda^2 (2p^2 - 2p + 1)^2} \left(-144p^{12} + 864p^{11} - 2160p^{10} + 2880p^9 - 2232p^8 + 1152p^7 - 618p^6 + 414p^5 - 197p^4 + 40p^3 + 3p^2 - 2p + 1\right) .$$

4.2 The conditioned case

Let us now compute the expected duration of a game given that the player on serve wins the game. In this case it is easy to prove that the Jump chain of the conditioned chain is the matrix P' on the state space $\{1, \dots, 16\}$ given by:

$$p'_{ij} = p_{ij} \frac{h_j}{h_i} \quad \text{with } i, j \in \{1, \dots, 16\},$$

where the h_i are the absorption probabilities in 16. Moreover, the conditional generator matrix is the matrix Q' obtained from P' and the original exponential holding times of parameter λ. The mean and variance computed above are now

$$\mu = \frac{4\left(20p^5 - 84p^4 + 148p^3 - 143p^2 + 79p - 21\right)}{\lambda (2p^2 - 2p + 1)(8p^3 - 28p^2 + 34p - 15)} ,$$

$$\sigma^2 = \frac{4}{\lambda^2 (2p^2 - 2p + 1)^2 (-8p^3 + 28p^2 - 34p + 15)^2} \left(320p^{10} - 2880p^9 + 11616p^8 - 27744p^7 + 43608p^6 - 47460p^5 + 36746p^4 - 20540p^3 + 8287p^2 - 2288p + 336\right) .$$

In Table 3 we show the $Mean \pm StandardDeviation$ of the actual time of a game when the point winning probabilities for the serving player are p and $\bar{p} = p$ (the rate λ is set arbitrarily equal to 2 for ease of exposition). For different values of p (recall that we have set $\overline{p} = p$) the first row shows the ($mean \pm StandardDeviation$) time distribution for finishing a game when the player with probability of winning a point equal to p is serving. The second row shows ($mean \pm StandardDeviation$) when the distribution is conditioned to the winning of the player that is serving. Note that, due to symmetry of the problem, when $p < 0.5$ and the model is conditioned as above, the results represent both the average time for the player (characterized by $p < 0.5$) to win a game on his turn of serving, and the average time to win a game for the player with same p when the other player is serving. The same empirical quantities are derived in Morante and Brotherhooc (2007). This is one of the few quantitative studies about the duration of points and games in Tennis, although they are more interested in relating playing time with performance indicators. They compute the average duration using a sample of Grand Slam matches for both male and female professional players.

$\mathbf{p = 0.9}$	$\mathbf{p = 0.8}$	$\mathbf{p = 0.7}$	$\mathbf{p = 0.6}$	$\mathbf{p = 0.5}$	$\mathbf{p = 0.4}$	$\mathbf{p = 0.3}$	$\mathbf{p = 0.2}$	$\mathbf{p = 0.1}$
2.23±1.13	2.54±1.34	2.92±1.60	3.24±1.82	3.37±1.90	3.24±1.82	2.92±1.60	2.54±1.34	2.23±1.13
2.23±1.13	2.53±1.33	2.87±1.58	3.18±1.80	3.37±1.90	3.40±1.86	3.30±1.71	3.13±1.52	2.96±1.35

Table 3: ($mean \pm StandardDeviation$) time for finishing a game (first row) and for winning a game while serving.

5 Conclusions

In this paper we present a model for Tennis including non i.i.d. point winning probabilities and the time of play. Indeed, we allow winning point probabilities to change depending on the score of the game. Moreover, we are interested in describing the time of play since there is some concerns about the excessive length of matches, especially in male Grand Slam competitions. In particular, in the present work, we obtain the distribution of the actual time of a game.

References

[1] Carrari, A., Ferrante M. and Fonseca G. (2017) *A new Markovian model for tennis matches.* Electronic Journal of Applied Statistical Analysis, to appear.

[2] Ferrante, M. and Fonseca G. (2014) *On the winning probabilities and mean durations of volleyball.* Journal of Quantitative Analysis in Sports **10**, 91-98.

[3] Klaassen, F. and Magnus, J.R. (2001) *Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model.* Journal of the American Statistical Association **96**, 500-509.

[4] Klaassen, F. and Magnus, J.R. (2014) *Analyzing Wimbledon*. Oxford University Press, Oxford.

[5] Morante, S. and Brotherhooc, J. (2007) *Match Characteristics of Professional Singles Tennis*. `http://www.cptennis.com.au/pdf/CooperParkTennisPDF_MatchCharacteristics.pdf`.

[6] Neuts, M.F. (1981) *Matrix-Geometric Solutions in Stochastic Models: An algorithmic approach.* Johns Hopkins University Press, Baltimore.

[7] Newton, P.K. and Keller, J.B. (2005) *Probability of winning at tennis (I). Theory and data.* Studies in Applied Mathematics **114**, 241-269.

[8] Norris, J.R. (1998) *Markov chains.* Cambridge University Press, Cambridge.