International Journal of Applied Mathematics

Volume 32 No. 3 2019, 537-547
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
doi: http://dx.doi.org/10.12732/ijam.v32i3.13

ON THE NUMBER OF EVENT APPEARANCES IN A MARKOV CHAIN

Natalia Mezhennaya

Applied Mathematics Department
Bauman Moscow State Technical University
ul. Baumanskaya 2-ya, 5/1
Moscow - 105005, RUSSIA

Abstract: The paper presents the estimate for the total variation distance between the distribution of the number of appearances of homogeneous disjoint events in a segment of strongly ergodic Markov chain on the finite state space and accompanying Poisson distribution (i.e., Poisson distribution with a parameter equal to the expectation of the random variable under consideration). For this purpose the Chen–Stein method was used. As a result Poisson and normal limit theorems for the number of events appearances are derived. The considered scheme describes the well-known number of runs on consecutive letters, the number of f-recurrent runs, etc., and can be used for describing the properties of distribution of the special form scan statistic.

AMS Subject Classification: 60B10, 60J10, 60F05, 62E20
Key Words: number of events, Markov chain, Chen–Stein method, limit theorem

1. Introduction

The problem of the number of runs of consecutive identical letters in a sequence of independent random variables [7] and its generalization to the case of a sequence forming a Markov chain (see, e.g., [4]) are well known. We consider the following more general scheme.

Received: April 5, 2019

Let $\{X_j, j = 1, \dots, n\}$ be a Markov chain, $s \geq 1$. We suppose (informally) that the random event A_j depends only on the random variables $X_j, \dots, X_{j+s}$, and that the set of events $\{A_j, j = 1, 2, \dots, n - s\}$ is homogeneous and has the property $\mathbf{P}\{A_i A_j\} = 0$ for $|i - j| \leq s$.

In this paper, we study the distribution of the number of appearances of events A_j in a segment of a Markov chain. For example, such a scheme describes the well-known run of consecutive letters. Thus, according to [7, p. 62], the letters $X_j, \dots, X_{j+s}$ form a run of consecutive a's of at least t length if $X_{j+1} \neq a, X_{j+2} = a, \dots, X_{j+t} = a$.

This definition could be expanded. Let $f : \mathcal{A}_N^l \to \mathcal{A}_N$ be a numerical function, $l \geq 1$. We define the event as follows:

$$A_j = \{X_{j+l} \neq f(X_j, \dots, X_{j+l-1}), X_{j+l+1} = f(X_{j+1}, \dots, X_{j+l}), \dots, \\ X_{j+l+t} = f(X_{j+t}, \dots, X_{j+l+t-1})\}.$$

In this case, the event A_j means that $X_j, \dots, X_{j+l+t}$ form an f-recurrent run of length at least t (see. [17]). The events A_i and A_j are incompatible for $|i - j| \leq s = t + l$.

It is easy to see that the concept of the f-recurrent run includes a run of consecutive a's ([7, p. 62]). Indeed, if $l = 1$ and the function $f \equiv a, a \in \mathcal{A}_N$, then

$$A_j = \{X_{j+1} \neq a, X_{j+2} = a, \dots, X_{j+t+1} = a\}$$

and the f-recurrent run coincides with the run of consecutive a's of at least t length.

The exact distributions of the numbers of runs in binary Markov chains were studied by Savelyev and Balakin [23, 24], Antzoulakos [1], Inoue [11], and their limit distributions in Markov chains with any number of states were obtained by Tikhomirova [25], Chryssaphinou et al. [5], and Fu et al. [9]. The distribution of the length of the longest run was considered by Erdos and Revesz [6], Fu [8], and Lou [13] for a sequence of independent random variables and by Chryssaphinou and Vaggelatou [4, 26] and Zhang [27] for a Markov chain.

The distribution of the number of f-recurrent runs in a sequence of independent random variables was studied by Mikhailov [16, 17]. Similar results for the number of usual and f-recurrent runs with possible omissions of letters were obtained by Mezhennaya in [14] and [15], respectively. There are other generalizations of the problem about the distribution of the number of runs. For example, Minakov [22] discussed the distribution of the number of monotone tuples and runs in a finite ergodic Markov chain.

The problem under consideration is closely related to the problem of the number of repetitions of tuples in a random sequence. The limiting distribution of the number of pairs of identical tuples in a Markov chain was studied by Mikhailov and Shoitov [18, 20, 21], and the number of tuples with the same structure by Mikhailov [19].

The considered problem is also related to the description of properties of widely known scan statistics. According to [10, p. 58], scan statistic is defined as follows:

$$T_{n,s}(u) = \max_{1 \le i \le n-k+1} \sum_{j=i}^{i+s-1} I_{\{X_j > u\}}$$

(in what follows I_A is an indicator of a random event A). The value of $T_{n,s}(u)$ is equal to the maximum number of exceedance occurrences by the random values $\{X_j, j = 1, \ldots, n\}$ of the threshold value u among all moving windows of length s. We denote

$$A_i = \left\{ \sum_{j=i}^{i+s-1} I_{\{X_j > u\}} = l - 1, \sum_{j=i+1}^{i+s} I_{\{X_j > u\}} = l \right\}.$$

Then

$$\{T_{n,s}(u) < l\} = \left\{ \sum_{i=1}^{n-s} I_{A_i} = 0 \right\}.$$

The distributions (both exact and asymptotic) of scan statistics for a sequence of independent random variables have been well studied [10], and their applications in various applied problems [10, 2] are also widely discussed.

In this paper, we estimate the total variation distance between the distribution of the number of appearances of events A_j in a segment of Markov chain and the accompanying Poisson distribution. As a result the Poisson and normal limit theorem for the random variable under consideration will be derived.

2. Main Results

Let $\{X_j, j = 1, \ldots, n\}$ be a strongly ergodic stationary Markov chain with the states set $\mathcal{A}_N = \{1, \ldots, N\}$, $N \ge 2$, with the transition probability matrix $P = ||p_{a,b}||_{a,b \in \mathcal{A}_N}$ and stationary distribution $\{\pi_a, a \in \mathcal{A}_N\}$. The elements of the matrix P^n are denoted by $p_{a,b}^{(n)}$, for $n = 1$ $p_{a,b}^{(1)} = p_{a,b}$.

According to [12, Cor. 4.1.5, p. 71] there are constants $C, \gamma > 0$ such as

$$|p_{a,b}^{(n)} - \pi_b| \le C\pi_b e^{-\gamma n},\ n \ge 1. \tag{1}$$

We assume that the random event A_j depends only on the random variables $X_j, \ldots, X_{j+s}$, $s \ge 1$, and that the set of events $\{A_j, j = 1, \ldots, n-s\}$ is homogeneous and has the property $\mathbf{P}\{A_i A_j\} = 0$ for $|i-j| \le s$.

Let $\Gamma = \{1, \ldots, n-s\}$, $\{\alpha_j = I_{A_j}, j \in \Gamma\}$ be the set of random indicators corresponding to the random events $\{A_j, j \in \Gamma\}$, and $Q_s = \mathbf{P}\{A_j\}$ be the probability of any event from the set $\{A_j, j \in \Gamma\}$.

We define the random variable $\xi = \sum_{j=1}^{n-s} \alpha_j$, which is equal to the number of appearances of events A_j in $\{X_j, j = 1, \ldots, n\}$, and its expectation

$$\lambda_s = \mathbf{E}\xi = (n-s)Q_s. \tag{2}$$

The following notation will be used: $\mathcal{L}(\xi)$ for the distribution of the random variable ξ, $\mathrm{Pois}(\lambda)$ for the Poisson distribution with parameter λ, $\mathcal{N}(0,1)$ for the standard normal distribution, and $\rho\left(\mathcal{L}(\xi), \mathcal{L}(\eta)\right)$ for the total variation distance between $\mathcal{L}(\xi)$ and $\mathcal{L}(\eta)$, respectively. For non-negative integer random variables ξ and η the total variation distance is given by the formula

$$\rho\big(\mathcal{L}(\xi), \mathcal{L}(\eta)\big) = \frac{1}{2}\sum_{u=0}^{\infty} |\mathbf{P}\{\xi = u\} - \mathbf{P}\{\eta = u\}|.$$

Theorem 1. *Let $s, m \ge 1$ and $\lambda_s \ge 1$. Then*

$$\rho\big(\mathcal{L}(\xi), \mathrm{Pois}(\lambda_s)\big) \le \left(2(s+2m) + 1 + \frac{2C}{e^{\gamma} - 1}\right)\frac{\lambda_s}{n-s} + Ce^{-\gamma(m+1)}\sqrt{\lambda_s}\left(2 + Ce^{-\gamma(m+1)} + e^{-\gamma(s+m+1)}\right),$$

where the constants C and γ are defined in (1).

Corollary 2. *Let $s, n \to \infty$, such that $s/n \to 0$, $Q_s \to 0$, $\lambda_s \to \lambda \in (0, \infty)$. Then $\mathcal{L}(\xi) \to \mathrm{Pois}(\lambda)$.*

Corollary 3. *Let $s, n \to \infty$, such that $\lambda_s \to \infty$, $s\lambda_s/n \to 0$, $\lambda_s = o(e^{\gamma s})$. Then $\mathcal{L}\big((\xi - \lambda_s)/\sqrt{\lambda_s}\big) \to \mathcal{N}(0,1)$.*

Corollaries 2 and 3 follow immediately from Theorem 1.

3. Proof of Theorem 1

We will use Theorem 1.A of [3, p. 9] and the proof scheme proposed by Mikhailov and Shoitov [21] and Minakov [22] to estimate the total variation distance between the distribution of the random variable ξ and the accompanying Poisson distribution (i.e., the Poisson distribution with parameter λ_s).

For each $i \in \Gamma$ we define the set of indices $\Gamma_i^s = \big\{j \in \Gamma\backslash\{i\}\big\}$ such that α_i and α_j are strongly dependent. The remaining indices are assigned to the set $\Gamma_i^w = \Gamma\backslash\big\{\Gamma_i^s \cup \{i\}\big\}$, which is called a set of weak dependence for a random indicator α_i. We present the formulation of Theorem 1.A from [3] for our case.

Theorem 4. *For each $i \in \Gamma$, let the set $\Gamma\backslash\{i\}$ be split into the disjoint subsets Γ_i^s and Γ_i^w. Then*

$$\rho\big(\mathcal{L}(\xi), \mathrm{Pois}(\lambda_s)\big) \le \min\left\{1, \frac{1}{\lambda_s}\right\}(S_1 + S_2) + \min\left\{1, \frac{1}{\sqrt{\lambda_s}}\right\} S_3, \tag{3}$$

where

$$S_1 = \sum_{i\in\Gamma} \sum_{j\in\Gamma_i^s\cup\{i\}} \mathbf{E}\alpha_i \mathbf{E}\alpha_j, \tag{4}$$

$$S_2 = \sum_{i\in\Gamma} \sum_{j\in\Gamma_i^s} \mathbf{E}\alpha_i\alpha_j, \tag{5}$$

$$S_3 = \sum_{i\in\Gamma} \mathbf{E}\big|\mathbf{E}\alpha_i - \mathbf{E}\left(\alpha_i |\, \{\alpha_j, j \in \Gamma_i^w\}\right)\big|. \tag{6}$$

Let $m \ge 1$. We put $\Gamma_i^s = \{j \in \Gamma : 1 \le |i-j| \le s+m\}$, $\Gamma_i^w = \Gamma\backslash\left(\{i\} \cup \Gamma_i^s\right)$, and estimate all summands in (3).

Let us begin with the sum S_1, which is given by (4). Since $|\Gamma_i^s \cup \{i\}| \le 2(s+m)+1$ and $|\Gamma| = n-s$, then considering the definition (2), we derive

$$S_1 \le (2(s+m)+1)\,(n-s)Q_s^2 = (2(s+m)+1)\,\lambda_s Q_s. \tag{7}$$

Next, we turn to S_2 (see (5)). The incompatibility of events A_i and A_j with $|i-j| \le s$ leads to

$$S_2 \le 2\sum_{i\in\Gamma} \sum_{j=i+s+1}^{i+s+m} \mathbf{E}\alpha_i\alpha_j.$$

Due to the Markov property we obtain

$$\mathbf{E}\alpha_i\alpha_j = \mathbf{P}\{A_iA_j\} = \mathbf{P}\{A_i\}\mathbf{P}\{A_j|A_i\} \le Q_s \max_{a\in\mathcal{A}_N} \mathbf{P}\{A_j|X_{i+s} = a\}. \tag{8}$$

Then

$$\mathbf{P}\{A_j|X_{i+s}=a\} \le \max_{b\in\mathcal{A}_N} p_{a,b}^{(j-i-s)}\frac{1}{\pi_b}Q_s$$
$$\le \max_{b\in\mathcal{A}_N} \pi_b(1+Ce^{-\gamma(j-i-s)})\frac{1}{\pi_b}Q_s = (1+Ce^{-\gamma(j-i-s)})Q_s.$$

Substituting this estimate into (8), we obtain

$$\mathbf{E}\alpha_i\alpha_j \le (1+Ce^{-\gamma(j-i-s)})Q_s^2.$$

Then, it follows from (5) that

$$S_2 \le 2(n-s)Q_s^2\sum_{k=1}^{m}(1+Ce^{-\gamma k}) = 2(n-s)Q_s^2\left(m+C\frac{1-e^{-\gamma m}}{e^\gamma-1}\right).$$

Applying definition (2) to the last estimate produces

$$S_2 \le 2\lambda_s Q_s\left(m+\frac{C}{e^\gamma-1}\right). \tag{9}$$

To estimate the sum S_3 we use the scheme proposed by Mikhailov and Shoitov [21] and Minakov [22]. We start with a separate summand of S_3, which we denote as $s_{3,i}$ as follows:

$$s_{3,i} = \mathbf{E}\left|Q_s - \mathbf{E}\left(\alpha_i|\{\alpha_j, j\in\Gamma_i^w\}\right)\right|.$$

According to the Markov property there are three possible cases:

a) for $m+2 \le i \le n-s-m-1$

$$s_{3,i} = \mathbf{E}\left|Q_s - \mathbf{E}\left(\alpha_i|\{X_{i-m-1}, X_{i+s+m+1}\}\right)\right|;$$

b) for $1 \le i \le m+1$

$$s_{3,i} = \mathbf{E}\left|Q_s - \mathbf{E}\left(\alpha_i|X_{i+s+m+1}\right)\right|;$$

c) for $n-s-m \le i \le n-s$

$$s_{3,i} = \mathbf{E}\left|Q_s - \mathbf{E}\left(\alpha_i|X_{i-m-1}\right)\right|.$$

First, for case a) we have

$$s_{3,i} = \sum_{a,b\in\mathcal{A}_N}\left|Q_s - \mathbf{E}\left(\alpha_i|X_{i-m-1}=a, X_{i+s+m+1}=b\right)\right| \times$$
$$\times\mathbf{P}\{X_{i-m-1}=a, X_{i+s+m+1}=b\} =$$

$$= \sum_{a,b\in\mathcal{A}_N} |Q_s \mathbf{P}\{X_{i-m-1} = a, X_{i+s+m+1} = b\} -$$
$$-\mathbf{P}\{A_i, X_{i-m-1} = a, X_{i+s+m+1} = b\}| . \qquad (10)$$

Since $\mathbf{P}\{X_{i-m-1} = a, X_{i+s+m+1} = b\} = \pi_a p_{a,b}^{s+2(m+1)}$, then the use of (1) leads to the following estimate

$$|p_{a,b}^{s+2(m+1)} - \pi_b| \le C\pi_b e^{-\gamma(s+2(m+1))}.$$

Now we can find the upper and lower bounds for the second probability in (10):

$$\mathbf{P}\{A_i, X_{i-m-1} = a, X_{i+s+m+1} = b\} \le \pi_a \max_{u,v\in\mathcal{A}_N} p_{a,u}^{(m+1)} \frac{1}{\pi_u} Q_s p_{v,b}^{(m+1)}$$
$$\le \pi_a Q_s \max_{u,v\in\mathcal{A}_N} \pi_u (1 + Ce^{-\gamma(m+1)}) \frac{1}{\pi_u} \pi_b (1 + Ce^{-\gamma(m+1)})$$
$$= \pi_a \pi_b Q_s (1 + Ce^{-\gamma(m+1)})^2. \qquad (11)$$

Analogously,

$$\mathbf{P}\{A_i, X_{i-m-1} = a, X_{i+s+m+1} = b\} \ge \pi_a \min_{u,v\in\mathcal{A}_N} p_{a,u}^{(m+1)} \frac{1}{\pi_u} Q_s p_{v,b}^{(m+1)} q$$
$$\ge \pi_a Q_s \min_{u,v\in\mathcal{A}_N} \pi_u (1 - Ce^{-\gamma(m+1)}) \frac{1}{\pi_u} \pi_b (1 - Ce^{-\gamma(m+1)})$$
$$= \pi_a \pi_b Q_s (1 - Ce^{-\gamma(m+1)})^2. \qquad (12)$$

Substituting (11) and (12) into (10), we obtain

$$s_{3,i} \le \sum_{a,b\in\mathcal{A}_N} \pi_a \pi_b Q_s \left[(1 + Ce^{-\gamma(m+1)})^2 - (1 - Ce^{-\gamma(s+2(m+1))}) \right]$$
$$\le Ce^{-\gamma(m+1)} Q_s \left[2 + Ce^{-\gamma(m+1)} + e^{-\gamma(s+m+1)} \right]. \qquad (13)$$

The number of such summands in the sum S_3 is $n - s - 2m - 2$.

We turn to case b). Similarly to case a) we derive

$$s_{3,i} = \sum_{a\in\mathcal{A}_N} |Q_s - \mathbf{E}\,(\alpha_i | X_{i+s+m+1} = a)| \,\mathbf{P}\{X_{i+s+m+1} = a\}$$
$$= \sum_{a\in\mathcal{A}_N} |Q_s \mathbf{P}\{X_{i+s+m+1} = a\} - \mathbf{P}\{A_i, X_{i+s+m+1} = a\}|$$

$$= \sum_{a \in \mathcal{A}_N} \left| Q_s \pi_a - \mathbf{P}\{A_i, X_{i+s+m+1} = a\} \right|. \tag{14}$$

Now we write the upper and lower bounds for the second probability in (14):

$$\mathbf{P}\{A_i, X_{i+s+m+1} = a\} \le \max_{v \in \mathcal{A}_N} Q_s p_{v,a}^{(m+1)}$$
$$\le Q_s \max_{u,v \in \mathcal{A}_N} \pi_a (1 + Ce^{-\gamma(m+1)}) = \pi_a Q_s (1 + Ce^{-\gamma(m+1)}),$$
$$\mathbf{P}\{A_i, X_{i+s+m+1} = a\} \ge \min_{v \in \mathcal{A}_N} Q_s p_{v,a}^{(m+1)} \ge \pi_a Q_s (1 - Ce^{-\gamma(m+1)}).$$

From the last two formulas we derive

$$s_{3,i} \le \sum_{a \in \mathcal{A}_N} \pi_a Q_s \left[(1 + Ce^{-\gamma(m+1)}) - 1 \right] = Ce^{-\gamma(m+1)} Q_s. \tag{15}$$

Similar calculations in case c) lead to the same estimate. The total number of summands in cases b) and c) is $2m + 2$. Substituting the estimates of the terms (13) and (15) into (6), we obtain the following estimator:

$$S_3 \le Ce^{-\gamma(m+1)} Q_s (n - s - 2m - 2) \left(2 + Ce^{-\gamma(m+1)} + e^{-\gamma(s+m+1)} \right)$$
$$+ Ce^{-\gamma(m+1)} Q_s (2m + 2)$$
$$\le Ce^{-\gamma(m+1)} \lambda_s \left(2 + Ce^{-\gamma(m+1)} + e^{-\gamma(s+m+1)} \right). \tag{16}$$

The statement of the theorem arises from the formulas (7), (9) and (16). The proof of Theorem 1 is complete.

Acknowledgements

The author is grateful to V.G. Mikhailov for the helpful suggestions and remarks.

References

[1] D.L. Antzoulakos, S. Chadjiconstantinidis, Distributions of numbers of success runs of fixed length in Markov dependent trials, *Ann. Inst. Statist. Math.*, **53**, No 3 (2001), 599–619.

[2] N. Balakrishnan, M.V. Koutras, *Runs and Scans with Applications*, Whiley, New York (2002).

[3] A.D. Barbour, L. Holst, S. Janson, *Poisson Approximation*, Oxford Univ. Press, Oxford (1992).

[4] O. Chryssaphinou, S. Papastavridis, E. Vaggelatou, Poisson approximation for the number of non-overlapping appearances of several words in Markov chain, *Combinatorics Probab.*, **10** (2001), 293–308.

[5] O. Chryssaphinou, E. Vaggelatou, Compound Poisson approximation for multiple runs in a Markov chain, *Ann. Inst. Statist. Math.*, **54**, No 2 (2002), 411–424.

[6] P. Erdos, P. Revesz, On the length of the longest head-run, In: *Topics in Information Theory. Colloquia Math. Soc. J. Bolyai* **16**, Keszthely, Hungary (1975), 219–228.

[7] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I, 3rd Ed., Wiley, New York (1968).

[8] J.C. Fu, Distribution theorem of runs and patterns associated with a sequence of multi-state trials, *Statist. Sinica*, **6** (1996), 957–974.

[9] J.C. Fu, W.Y.W. Lou, Z.-D. Bai, G. Li, The exact and limiting distributions for the number of successes in success runs within a sequence of Markov-dependent two-state trials, *Ann. Inst. Statist. Math.*, **54**, No 4 (2002), 719–730.

[10] J. Glaz, V. Pozdnyakov, S. Wallenstein (Eds.), *Scan Statistics. Methods and Applications*, Birkhäuser, Boston-Basel-Berlin (2009).

[11] K. Inoue, S. Aki, Joint distributions of numbers of runs of specified length in a sequence of Markov dependent multistate trials, *Ann. Inst. Statist. Math.*, **59**, No 3 (2007), 577–595.

[12] J.G. Kemeny, J.L. Snell, *Finite Markov Chains*, Springer-Verlag, New York (1976).

[13] W.Y.W. Lou, On runs and longest runs tests: a method of finite Markov chain imbedding, *J. Amer. Statist. Assoc.*, **91** (1996), 1595–1601.

[14] N.M. Mezhennaya, Limit theorems for the number of dense series in a random sequence, *Discrete Mathematics and Applications*, **19**, No 2 (2009), 215–228.

[15] N.M. Mezhennaya, Limit theorems for dense F-recurrent series and chains numbers in sequence of independent random variables, *Herald of the Bauman Moscow State Technical University, Series Natural Sciences*, **3** (2014), 11–25 (in Russian).

[16] V.G. Mikhailov, On the limit theorem of B.A. Sevastiyanov for sums of dependent random indicators, *Review of Applied and Industrial Mathematics*, **10**, No 3 (2003), 571–578 (in Russian).

[17] V.G. Mikhailov, On asymptotic properties of the number of runs of events, *Tr. Diskr. Mat.*, **9** (2006), 152–163 (in Russian).

[18] V.G. Mikhailov, Estimates of accuracy of the Poisson approximation for the distribution of number of runs of long string repetitions in a Markov chain, *Discrete Math. Appl.*, **26**, No 2 (2016), 105–113.

[19] V.G. Mikhailov, On the probability of existence of substrings with the same structure in a random sequence *Discrete Math. Appl.*, **27**, No 6 (2017), 377–386.

[20] V.G. Mikhailov, A.M. Shoitov, On multiple repetitions of long tuples in a Markov chain, *Mat. Vopr. Kriptogr.*, **6**, No 3 (2015), 117–133 (in Russian).

[21] V.G. Mikhailov, A.M. Shoitov, On repetitions of long tuples in a Markov chain, *Discrete Math. Appl.*, **25**, No 5 (2015), 295–303.

[22] A.A. Minakov, Poisson approximation for the number of non-decreasing runs in Markov chains, *Mat. Vopr. Kriptogr.*, **9**, No 2 (2018), 103–116.

[23] L.Ya. Savelyev, S.V. Balakin, B.V. Khromov, Covering runs in binary Markov sequence, *Discrete Math. Appl.*, **13**, No 2 (2003), 111–138.

[24] L.Ya. Savelyev, S.V. Balakin, Some applications of the stochastic theory of runs, *Sib. Zh. Ind. Mat.*, **15**, No 3 (2012), 111–123 (in Russian).

[25] M.I. Tikhomirova, Limit distributions of the number of absent chains of identical outcomes, *Discrete Math. Appl.*, **18**, No 3 (2008), 293–300.

[26] E. Vaggelatou, On the length of the longest run in a multi-state Markov chain, *Statist. Probab. Letters*, **62** (2003), 211–221.

[27] Y.Z. Zhang, X.Y. Wu, Some results associated with the longest run in a strongly ergodic Markov chain, *Acta Mathematica Sinica*, **29**, No 10 (2013), 1939–1948.