ICAI 2023
12th International Conference on Applied Informatics
Eger, Hungary, 2023

Convergence rate for the longest T contaminated runs of heads

István Fazekas[a], Borbála Fazekas[b]Michael Suja[c]

[a]University of Debrecen
fazekas.istvan@inf.unideb.hu

[b]University of Debrecen
borbala.fazekas@science.unideb.hu

[c]University of Debrecen
michael.suja@science.unideb.hu

Abstract

The classical coin tossing experiment is studied. Limit theorems are obtained concerning the head - runs containing certain number T of tails called contaminated runs of heads.Asymptotic distributions are derived for the first hitting time of a T- contaminated run of heads and the length of the longest T- contaminated run of heads. These offer improved approximations for convergence of the random variables which out perform previous established results. A comparison is made through simulation of the two sets of estimates giving credence of excellent performance.

Problem 1 (First hitting time for the longest ***T*** contaminated run of heads)**.** *Consider the well known coin tossing experiment. Let p be the probability of heads and let $q = 1-p$ be the probability of tails. Here p is a fixed number with $0 < p < 1$. We toss a coin N independent times. We shall write 1 when the result is head and 0 when the result is tail. So we shall consider independent and identically distributed random variables $X_1, X_2, \ldots, X_n$ with $P(X_i = 1) = p$ and $P(X_i = 0) = q$, $i = 1, 2, \ldots, N$.*
Let $T > 0$ be a fixed non - negative integer. We shall study the T- interrupted runs of heads. It means that there are T zeros in an m length sequence of ones and zeros. So let m be a positive integer. Let $A_n = A_{n,m}$ denote the event that there are precisely T zeros in the sequence $X_n, X_{n+1}, \ldots X_{n+m-1}$.
Let τ_m be the first hitting time of the T- contaminated run of heads having length

m. We shall find the asymptotic distribution of τ_m as $m \to \infty$ for $T = 1$ and for $T = 2$.

For the proof, we apply the main lemma of Csáki, Földes, Komlós [1].

Theorem 2. *Let $T = 1$ or $T = 2$, $0 < p < 1$. Let τ_m be the first hitting time for the T contaminated run of heads having length m. Then, for $x > 0$,*

$$P(\tau_m \alpha P(A_1) > x) \sim e^{-x} \tag{1}$$

as $m \to \infty$. Here if $T = 1$, then $\alpha = q + \frac{2p^{m-1}-1}{m}$ and $P(A_1) = mp^{m-1}q$. When $T = 2$, then $\alpha = q - \frac{2}{m}$, $P(A_1) = \binom{m}{2}p^{m-2}q^2$.

Problem 3 (Length of the longest $\boldsymbol{T}$ contaminated run of heads)**.** *Erdős and Rényi [2] in their classical paper studied pure head runs for the case of a fair coin.*

Later on, [8] extended the study to the accuracy of the approximation to the distribution of the length of the longest head run in a Markov chain.

Almost sure limit results for the length of the longest runs containing at most T tails was detailed by [3].

Földes, in [5], presented asymptotic results for the distribution of the length of longest T contaminated head runs.

Móri in [7], obtained a so called almost sure limit theorem for the longest T contaminated head run.

Gordon, Schilling and Waterman [6] applied extreme value theory to obtain the asymptotic behaviour of the expectation and variance of the longest T contaminated head run.

Theorem 4. *Let $T = 1$ or $T = 2$, and let $0 < p < 1$ be fixed. Let $\mu(N)$ be the length of the longest T contaminated run of heads during N times of coin tossing. Let*

$$\begin{aligned} m(N) =& log(qN) + Tlog(log(qN)) + \\ &+ T^2 \frac{log(log(qN))}{clog(qN)} - \frac{T}{cq_0 log(qN)} - \frac{T^3}{2c}\left(\frac{log(log(qN))}{log(qN)}\right)^2 + \\ &+ T^2 \frac{log(log(qN))}{cq_0(log(qN))^2} + T^3 \frac{log(log(qN))}{(clog(qN))^2} + \\ &+ \left(Tlog(\frac{q}{p}) - log(T!)\right)\left(1 + \frac{T}{clog(qN)} - T^2 \frac{log(log(qN))}{c(log(qN))^2}\right), \end{aligned} \tag{2}$$

where $q_0 = \frac{2q}{2+Tq-q}$, log denotes the logarithm to base $1/p$ and $c = ln(1/p)$, ln denotes the natural logarithm to base e. Let $[m(N)]$ denote the integer part of $m(N)$ while $\{m(N)\}$ denotes the fractional part of $m(N)$, i.e $\{m(N)\} = m(N) - [m(N)]$. Then,

$$P(\mu(N) - [m(N)] < k) = e^{-p^{(k-\{m(N)\})\left(1 - \frac{T}{clog(qN)} + T^2 \frac{log(log(qN))}{c(log(qN))^2}\right)}} \left(1 + O\left(\frac{1}{(logN)^2}\right)\right) \tag{3}$$

for any integer k, where $f(N) = O(h(N))$ means that $f(N)/h(N)$ is bounded as $N \to \infty$.

Remark 5. In Proposition 3.3 from [4], the rate of convergence was $O(log(log(N))/log(N))$.

Using our method for $T = 1$ and $T = 2$ and for $m_0(N)$ (which is similar to $\mu_T(qN)$ in aforementioned paper), we obtain that the rate of convergence is $O\big(1/(log(N))^2\big)$. This is an improvement over this

$$P(\mu(N) - [m_0(N)] < k) = exp\Big(-p^{k-\{m_0(N)\}}\Big)(1 + O(log(log(N)/log(N)))).$$

So our Theorem considerably improves Theorem 1 of [6] in the cases of $T = 1$ and $T = 2$.

References

[1] E. Csáki, A. Földes, J. Komlós: *Limit theorems for Erdős-Rényi type problems*, Studia Sci. Math. Hungar 22 (1987), pp. 321–332.

[2] P. Erdős, A. Rény: *On a new law of large numbers*, Analyse Math. 23 (1970), pp. 103–111.

[3] P. Erdős, P. Révész: *On the length of the longest head-run*, Topics in information theory 16 (1975), pp. 219–228.

[4] I. Fazekas, S. Ochieng Michael: *Limit theorems for contaminated runs of heads*, Ann. Univ. Sci. Budapest, Sect. Comp 52 (2021), pp. 131–146.

[5] A. Földes: *The limit distribution of the length of the longest head-run*, Periodica Mathematica Hungarica 10.4 (1979), pp. 301–310.

[6] L. Gordon, M. F. Schilling, M. S. Waterman: *An extreme value theory for long head runs*, Probability Theory and Related Fields 72.2 (1986), pp. 279–287.

[7] T. F. Móri: *The a.s limit distribution of the longest head run*, Canadian journal of mathematics 45.6 (1993), pp. 1245–1262.

[8] S. Novak: *On the length of the longest head run*, Statistics & Probability Letters 130 (2017), pp. 111–114.