Theory of Stochastic Processes
Vol. 12 (28), no. 1–2, 2006, pp. 106–115

UDC 519.21

VOLODYMYR I. MASOL AND SVITLANA V. POPERESHNYAK

POISSON ESTIMATES OF THE DISTRIBUTION OF THE RANK OF A RANDOM MATRIX OVER THE FIELD $GF(2)$

The estimates of the rate of convergence of the distribution of the rank of a random matrix over the field $GF(2)$ and the Poisson distribution with the parameter depending on both the matrix sizes and distributions of its elements are found.

Introducton

A matrix A over the field $GF(2)$ consisting of two elements is called a sparse Boolean matrix if the probability of the occurrence of 1 in its positions (i, j) equals $\frac{1}{n}(\ln n + x_{ij})$, where $|x_{ij}| \leq c$, $c = const$, $i = \overline{1, T}$, $j = \overline{1, n}$, $T/n/$ is the number of rows/columns/ of the matrix A.

In works [1], [2], using different methods, the limit Poisson distribution of the rank $r(A)$ of a random sparse Boolean matrix A was established when $T = T(n)$ and $n \to \infty$. For finite values of the parameters T and n, the distribution of the rank $r(A)$ can be presented in terms of the factorial moments of the random variable $r(A)$, as stated in [3]. The asymptotics of these moments under certain conditions, in particular, if $x_{ij} = x_i$, $i = \overline{1, T}$, $j = \overline{1, n}$ as $n \to \infty$, is given in ([4], Theorem 4).

The question on the rate of convergence of the distribution of the rank of a matrix A to the Poisson distribution with a properly chosen parameter has not been examined yet. The suggested paper is devoted to the investigation of the mentioned issue.

The proof of our main result (Theorem 1) is based on the theorem on the rate of convergence in a Poisson scheme given in [5, p.67]. It should also be noted that, in contrast with works [1], [2], the present paper considers a random matrix A, the distributions of elements in which can depend on their (elements) positions.

The Main Result

Let the elements of a $T \times n$ matrix $A = \|a_{ij}\|$, $i = \overline{1, T}$, $j = \overline{1, n}$ be the independent random variables that take values in the field $GF(2)$ and have the distribution

$$P\{a_{ij} = 1\} = 1 - P\{a_{ij} = 0\} = \frac{\ln n + x_{ij}}{n}, \tag{1}$$

where

$$|x_{ij}| \leq c, \quad c = const, \quad i = \overline{1, T}, \quad j = \overline{1, n}. \tag{2}$$

Let us assume that the matrix A has at least n_0 columns so that it is correct to determine a distribution by formula (1) for $n \geq n_0$. We denote, by $r(A)$, the rank of the

2000 *AMS Mathematics Subject Classification.* Primary 60C05.
Key words and phrases. Random matrix over the field $GF(2)$, rank of a matrix, Poisson distribution, rate of convergence.

matrix A and put λ equal

$$\lambda = \frac{1}{n}\sum_{i=1}^{T}\exp\left\{-\frac{1}{n}\sum_{j=1}^{n}x_{ij}\right\}. \tag{3}$$

Theorem 1. *Let conditions (1) and (2) be satisfied, and*

$$\frac{T}{n} \le 1 - \frac{\log_2 \ln n}{(\ln n)^q}, q = const \quad , \quad 0 < q < 1, \tag{4}$$

$$\underline{\lim}_{n\to\infty} \frac{T}{n} > 0. \tag{5}$$

Then, for k, k=0, 1, 2 . . . ,

$$\left|P\left\{r\left(A\right) = T - k\right\} - \frac{e^{-\lambda}\cdot\lambda^k}{k!}\right| \le 2\left(1+\delta\right)c(n,\, k)\frac{\ln^4 n}{n\left(\ln\ln n\right)^2},$$

where $1 < \underline{\lim}_{n\to\infty}\, c\left(n,\, k\right) \le \overline{\lim}_{n\to\infty}\, c\left(n,\, k\right) \le e^{e^c}$, $\delta > 0$, $\delta = const.$

Remark. *The explicit form of the coefficient $c(n, k)$ is given by equality (30).*

Auxiliary Statements

By $\xi_{n,T}$, we denote the number of zero rows of the matrix A.

Lemma 1. *When conditions (1) and (2) hold, the distribution of the random variable $\xi_{n,T}$ satisfies the inequality*

$$\left|P\left\{\xi_{n,T} = k\right\} - \frac{e^{-\lambda}\cdot\lambda^k}{k!}\right| \le Q(n,\, k), \tag{6}$$

where $Q(n,\, k) = \frac{\ln^2 n}{n}c_1(n,\, k)$ *and* $0 < \underline{\lim}_{n\to\infty}\, c_1\left(n,\, k\right) \le \overline{\lim}_{n\to\infty}\, c_1\left(n,\, k\right) \le e^{-1}$ *if* $k = 0$, $0 < \underline{\lim}_{n\to\infty}\, c_1\left(n,\, k\right) \le \overline{\lim}_{n\to\infty}\, c_1\left(n,\, k\right) \le \frac{e^{ck}}{2k!}\max\left(k,\, e^c\right)$ *if* $k \ge 1$.

Proof. The probability $p_n^{(i)}$ that the row i of the matrix A consists of zeros exclusively is, obviously, equal to

$$p_n^{(i)} = \prod_{j=1}^{n}\left(1 - \frac{\ln n + x_{ij}}{n}\right), \quad i = \overline{1,\, T}.$$

Let us put $a = p_n^{(1)} + p_n^{(2)} + \ \ldots + p_n^{(n)}$.

According to the Poisson theorem [5, p. 67], we get the inequality for arbitrary $k = 0, 1, 2, \ldots$

$$\left|P\left\{\xi_{n,T} = k\right\} - \frac{e^{-a}\cdot a^k}{k!}\right| \le \sum_{i=1}^{T}\left(p_n^{(i)}\right)^2. \tag{7}$$

Applying the inequality $p_n^{(i)} \le \frac{1}{n}\exp\left\{-\frac{1}{n}\sum_{j=1}^{n}x_{ij}\right\}$ to the right-hand side of (7), we obtain

$$\left|P\left\{\xi_{n,T} = k\right\} - \frac{e^{-a}\cdot a^k}{k!}\right| \le \frac{1}{n^2}\sum_{i=1}^{T}\exp\left\{-\frac{2}{n}\sum_{j=1}^{n}p_{ij}\right\}. \tag{8}$$

Since

$$\lambda\left(1-\gamma_n\right) \le a \le \lambda,$$

where $\gamma_n = \left(\frac{\ln^2 n}{2n}\right) \frac{\left(1-c(\ln n)^{-1}\right)^2}{1-(\ln n - c)n^{-1}}$, we get

$$-\lambda^{k+1} e^{-\lambda} \gamma_n \left(1 + \frac{\lambda \gamma_n}{2} e^{\lambda \gamma_n}\right) \le \lambda^k e^{-\lambda} - a^k e^{-a} \le \lambda^k e^{-\lambda} k \gamma_n.$$

Hence,

$$\left|\lambda^k e^{-\lambda} - a^k e^{-a}\right| \le \lambda^k e^{-\lambda} \gamma_n \max\left\{k,\, \lambda\left(1 + \frac{\lambda\gamma_n}{2} e^{\lambda\gamma_n}\right)\right\}. \tag{9}$$

Combining ratios (8) and (9) leads to

$$\left| P\left\{\xi_{n,T} = k\right\} - \frac{e^{-\lambda} \cdot \lambda^k}{k!} \right| \le Q(n,\, k),$$

where

$$Q\left(n,\, k\right) = \frac{\ln^2 n}{n} c_1(n,\, k), \tag{10}$$

$$c_1\left(n,\, k\right) = \frac{1}{n \ln^2 n} \sum_{i=1}^{T} \exp\left\{-\frac{2}{n} \sum_{j=1}^{n} p_{ij}\right\} +$$

$$+ \frac{\lambda^k e^{-\lambda}}{2\, k!} \frac{\left(1 - c\left(\ln n\right)^{-1}\right)^2}{1 - \left(\ln n - c\right) n^{-1}} \max\left\{k,\, \lambda\left(1 + \frac{\lambda\gamma_n}{2} e^{\lambda\gamma_n}\right)\right\},$$

$0 < \underline{\lim}_{n\to\infty} c_1\left(n,\, k\right) \le \overline{\lim}_{n\to\infty} c_1\left(n,\, k\right) \le e^{-1}$ *for k=0,* $0 < \underline{\lim}_{n\to\infty} c_1\left(n,\, k\right) \le \overline{\lim}_{n\to\infty} c_1\left(n,\, k\right) \le \frac{e^{ck}}{2k!} \max\left(k,\, e^c\right)$ for $k \ge 1$.

Lemma 1 is proved.

By $S_1\left(A\right)$, we denote the maximum number of independent critical sets of the matrix A (see [2, p. 147]), each containing at least one non-zero row.

Lemma 2. *If condition (1) is satisfied, the expectation of the random variable* $S_1\left(A\right)$ *equals*

$$MS_1\left(A\right) = \sum_{k=0}^{T} \sum_{1 \le i_1 < \cdots < i_k \le T} \frac{1}{2^n} \prod_{j=1}^{n} \left(1 + \prod_{s=1}^{k} \left(1 - \frac{2\left(\ln n + x_{i_s j}\right)}{n}\right)\right) -$$

$$- \sum_{k=0}^{T} \sum_{1 \le i_1 < \cdots < i_k \le T} \prod_{s=1}^{k} \prod_{j=1}^{n} \left(1 - \frac{\ln n + x_{i_s j}}{n}\right).$$

Proof. Let us take into account that the probability that the number of successes in a series of k independent trails, with the probability of a success p_i, $i = \overline{1,\, k}$ being an even number, equals

$$\frac{1}{2} + \frac{1}{2} \prod_{i=1}^{k} \left(1 - 2p_i\right).$$

(It is not difficult to verify it using induction on the parameter $k \ge 1$).

Hence, the probability that k rows constitute a critical set can be presented by the expression

$$\frac{1}{2^n} \prod_{j=1}^{n} \left(1 + \prod_{i=1}^{k} \left(1 - \frac{2\left(\ln n + x_{ij}\right)}{n}\right)\right).$$

Note that the probability that there is no single 1 in k rows equals

$$\prod_{i=1}^{k}\prod_{j=1}^{n}\left(1-\frac{\ln n+x_{ij}}{n}\right).$$

Now it is not difficult to complete the proof of Lemma 2.

Let us put

$$f(n)=\left[(\ln\ln n)^{-1}(1+\delta_1)\ln n\right],\quad \delta_1=const,\quad \delta_1>0,$$

$\mu(n)=\sum_{k=0}^{f(n)}\sum_{1\le i_1<\cdots<i_k\le T}\exp\left\{-\sum_{s=1}^{k}\sum_{j=1}^{n}\frac{\ln n+x_{i_s j}}{n}\right\}.$

Lemma 3. *If conditions (1), (2), and (4) are satisfied, and*

$$T\le n, \tag{11}$$

then

$$1<\varliminf_{n\to\infty}\mu(n)\le\varlimsup_{n\to\infty}\mu(n)\le e^{e^{c}}. \tag{12}$$

Proof. Let us estimate $\mu(n)$ from above. The sum $\mu(n)$ can be presented as

$$\mu(n)=\theta_1(n)-\theta_2(n), \tag{13}$$

where $\theta_1(n)=\sum_{k=0}^{T}\sum_{1\le i_1<\cdots<i_k\le T}\exp\left\{-\sum_{s=1}^{k}\sum_{j=1}^{n}\frac{\ln n+x_{i_s j}}{n}\right\}$,

$\theta_2(n)=\sum_{k=f(n)+1}^{T}\sum_{1\le i_1<\cdots<i_k\le T}\exp\left\{-\sum_{s=1}^{k}\sum_{j=1}^{n}\frac{\ln n+x_{i_s j}}{n}\right\}.$

Then $\theta_1(n)$ can be estimated in such a way:

$$\theta_1(n)\le e^{\lambda}. \tag{14}$$

Since $\theta_2(n)\ge 0$, relations (13) and (14) imply

$$\mu(n)\le e^{\lambda}. \tag{15}$$

Let us estimate $\mu(n)$ from below. Since

$$\theta_1(n)\ge e^{\lambda}\exp\left\{-\frac{1}{2n^2}\sum_{i=1}^{T}\exp\left(-\frac{2}{n}\sum_{j=1}^{n}x_{ij}\right)\right\},$$

we get

$$\mu(n)\ge e^{\lambda}\exp\left\{-\frac{1}{2n^2}\sum_{i=1}^{T}\exp\left(-\frac{2}{n}\sum_{j=1}^{n}x_{ij}\right)\right\}-\theta_2(n).$$

From the last inequality and (15), it follows that

$$e^{\lambda}\exp\left\{-\frac{1}{2n^2}\sum_{i=1}^{T}\exp\left(-\frac{2}{n}\sum_{j=1}^{n}x_{ij}\right)\right\}-\theta_2(n)\le\mu(n)\le e^{\lambda}. \tag{16}$$

To analyze $\theta_2(n)$, it is necessary to take into account conditions (2) and (11), that is

$$\theta_2(n)\to 0\quad as\ \ n\to\infty. \tag{17}$$

Letting $n\to\infty$ in inequality (16) and taking into account (17) and conditions (2), (5), and (11), we obtain (12).

Lemma 3 is proved.

Let us set the following $\Gamma(n)=u\left(1+\frac{u(1+\delta_1)}{2}e^{u(1+\delta_1)}\right)$, $u=\frac{\ln^4 n}{n(\ln\ln n)^2}$.

Lemma 4. *If conditions (1) and (2) are satisfied, then we have*

$$\sum_{k=0}^{f(n)} \sum_{1\le i_1<\dots.<i_k\le T} \frac{1}{2^n}\prod_{j=1}^{n}\left(1+\prod_{s=1}^{k}\left(1-\frac{2\left(\ln n+x_{i_s j}\right)}{n}\right)\right)\le \tag{18}$$
$$\le \mu\left(n\right)+\mu\left(n\right)\left(1+\delta\right)\Gamma\left(n\right).$$

Proof. The left-hand side of relation (18) can be estimated as

$$\sum_{k=0}^{f(n)} \sum_{1\le i_1<\cdots<i_k\le T} \frac{1}{2^n}\prod_{j=1}^{n}\left(1+\prod_{s=1}^{k}\left(1-\frac{2\left(\ln n+x_{i_s j}\right)}{n}\right)\right)\le$$
$$\le \sum_{k=0}^{f(n)} \sum_{1\le i_1<\cdots<i_k\le T} \prod_{j=1}^{n}\left(1-\sum_{s=1}^{k}\frac{\left(\ln n+x_{i_s j}\right)}{n}+\left(\sum_{s=1}^{k}\frac{\left(\ln n+x_{i_s j}\right)}{n}\right)^2\right).$$

By (2), the last inequality leads to

$$\sum_{k=0}^{f(n)} \sum_{1\le i_1<\cdots<i_k\le T} \frac{1}{2^n}\prod_{j=1}^{n}\left(1+\prod_{s=1}^{k}\left(1-\frac{2\left(\ln n+x_{i_s j}\right)}{n}\right)\right)\le \tag{19}$$
$$\le \mu(n)\cdot\exp\left\{\frac{f^2\left(n\right)\left(\ln n+c\right)^2}{n}\right\}.$$

Further, we obtain

$$\exp\left\{\frac{f^2\left(n\right)\left(\ln n+c\right)^2}{n}\right\}\le 1+u\left(1+\delta\right)\left(1+\frac{u}{2}\left(1+\delta\right)\exp\left\{u\left(1+\delta\right)\right\}\right). \tag{20}$$

Relations (19) and (20) give, obviously, (18). Lemma 4 is proved.

Lemma 5. *If conditions (1), (2), (5), and (11) are satisfied, then*

$$\sum_{k=0}^{f(n)} \sum_{1\le i_1<\cdots<i_k\le T} \prod_{s=1}^{k}\prod_{j=1}^{n}\left(1-\frac{\ln n+x_{i_s j}}{n}\right)\ge\mu\left(n\right)-c_2\left(n\right)\frac{\ln^3 n}{n\cdot\ln\ln n},$$

where $\frac{(1+\delta_1)}{2}<\underline{\lim}_{n\to\infty}c_2\left(n\right)\le\overline{\lim}_{n\to\infty}c_2\left(n\right)\le\frac{e^{e^c}}{2}\left(1+\delta_1\right).$

Proof. With the help of relation (2), we find the following inequalities:

$$\sum_{k=0}^{f(n)} \sum_{1\le i_1<\cdots<i_k\le T} \prod_{s=1}^{k}\prod_{j=1}^{n}\left(1-\frac{\ln n+x_{i_s j}}{n}\right)\ge$$
$$\ge \mu(n)\cdot\exp\left\{-\frac{f(n)}{2n}\left(\ln n+c\right)^2\cdot\frac{1}{1-\frac{\ln n+c}{n}}\right\}.$$

Therefore, using the notations introduced earlier, we come to

$$\sum_{k=0}^{f(n)} \sum_{1\le i_1<\cdots<i_k\le T} \prod_{s=1}^{k}\prod_{j=1}^{n}\left(1-\frac{\ln n+x_{i_s j}}{n}\right)\ge\mu(n)-c_2(n)\cdot\frac{\ln^3 n}{n\cdot\ln\ln n},$$

where

$$c_2(n)=\frac{\mu(n)}{2}\left(1+\delta_1\right)\left(1+\frac{c}{\ln n}\right)^2\frac{1}{1-\frac{\ln n+c}{n}}.$$

Taking into account (12), we get $\frac{(1+\delta_1)}{2}<\underline{\lim}_{n\to\infty}c_2(n)\le\overline{\lim}_{n\to\infty}c_2(n)\le\frac{1}{2}e^{e^c}\left(1+\delta_1\right)$ as $n\to\infty$.

Lemma 5 is proved.

Lemma 6. *If conditions (1), (2), (5), and (11) are satisfied, then*

$$\sum_{\substack{k:\frac{n}{2}\left(1-\frac{1}{\ln n}\right)<k\le\frac{n}{2}\left(1+\frac{1}{\ln n}\right)\\ k-integer}} \binom{n}{k}\frac{1}{2^n}\sum_{l=f(n)}^{T}\binom{T}{l}\left(\left(1-\frac{2(\ln n-c)}{n}\right)^k\right)^l \le$$

$$\le \left(\frac{c_3(n)}{f(n)}\right)^{f(n)}\cdot\frac{1}{\sqrt{f(n)}}\cdot c_4(n),$$

where $0<\underline{\lim}_{n\to\infty} c_3(n)\le\overline{\lim}_{n\to\infty} c_3(n)\le e^{2+c}$, $\lim_{n\to\infty} c_4(n)=(2\pi)^{-1/2}$.

Proof. For $(n/2)\left(1-1/\ln n\right)<k\le(n/2)\left(1+1/\ln n\right)$, we find

$$\left(1-\frac{2(\ln n-c)}{n}\right)^k\le e^{-2k(\ln n-c)n^{-1}}\le\frac{1}{n}e^{c+1-\frac{c}{\ln n}}.$$

Using the obtained estimation, we establish that

$$\sum_{\substack{k:\frac{n}{2}\left(1-\frac{1}{\ln n}\right)<k\le\frac{n}{2}\left(1+\frac{1}{\ln n}\right)\\ k-integer}} \binom{n}{k}\frac{1}{2^n}\sum_{l=f(n)}^{T}\binom{T}{l}\left(\left(1-\frac{2(\ln n-c)}{n}\right)^k\right)^l \le$$

$$\sum_{\substack{k:\frac{n}{2}\left(1-\frac{1}{\ln n}\right)<k\le\frac{n}{2}\left(1+\frac{1}{\ln n}\right)\\ k-integer}} \binom{n}{k}\frac{1}{2^n}\sum_{l=f(n)}^{T}\binom{T}{l}\left(\frac{1}{n}e^{c+1-\frac{c}{\ln n}}\right)^l\le\sum_{l=f(n)}^{T}\left(\frac{T}{n}\right)^l\frac{\left(e^{c+1-\frac{c}{\ln n}}\right)^l}{l!}.$$

Applying the Stirling formula, we get

$$\sum_{l=f(n)}^{T}\left(\frac{T}{n}\right)^l\frac{\left(e^{c+1-\frac{c}{\ln n}}\right)^l}{l!}\le\left(\frac{c_3(n)}{f(n)}\right)^{f(n)}\cdot\frac{1}{\sqrt{f(n)}}\cdot c_4(n),$$

where $c_3(n)=\frac{T}{n}\cdot e^{2+c-\frac{c}{\ln n}}$, $c_4(n)=\frac{1}{\sqrt{2\pi}}\cdot\left(1-\frac{c_3(n)}{f(n)}\right)^{-1}$; taking into account (5) and (11), we find the following expressions: $0<\underline{\lim}_{n\to\infty} c_3(n)\le\overline{\lim}_{n\to\infty} c_3(n)\le e^{2+c}$, $\lim_{n\to\infty} c_4(n)=(2\pi)^{-1/2}$

Lemma 6 is proved.

Lemma 7. *If conditions (1), (2), (4), and (5) are satisfied, then*

$$\sum_{k=f(n)+1}^{T}\sum_{1\le i_1<\dots<i_k\le T}\frac{1}{2^n}\prod_{j=1}^{n}\left(1+\prod_{s=1}^{k}\left(1-\frac{2(\ln n+x_{i_sj})}{n}\right)\right)\le c_5(n)\cdot(\ln n)^{-\frac{n\cdot c_6(n)}{(\ln n)^q}}+$$

$$+\left(\frac{c_3(n)}{f(n)}\right)^{f(n)}\cdot\frac{1}{\sqrt{f(n)}}\cdot c_4(n)+c_7(n)\cdot\exp\left\{-\frac{n}{2\ln^2 n}c_8(n)\right\},$$

where $\lim_{n\to\infty} c_5(n)=(2\pi)^{-1/2}$, $\lim_{n\to\infty} c_6(n)=1-q$, $\lim_{n\to\infty} c_8(n)=1$, $\sqrt{\frac{2}{\pi}}<\underline{\lim}_{n\to\infty} c_7(n)\le\overline{\lim}_{n\to\infty} c_7(n)\le\frac{e^{c+1}+e^{c-1}}{\sqrt{2\pi}}$.

Proof. Denote $S=\sum_{k=f(n)}^{T}\sum_{+1\ \ 1\le i_1<\dots<i_k\le T}\frac{1}{2^n}\prod_{j=1}^{n}\left(1+\prod_{s=1}^{k}\left(1-\frac{2(\ln n+x_{i_sj})}{n}\right)\right)$.

With the help of condition (2), we find that, for $n\ge n_0$,

$$0\le S\le\sigma,$$

where $\sigma=\sum_{k=f(n)}^{T}\binom{T}{k}\frac{1}{2^n}\left(1+\left(1-\frac{2(\ln n-c)}{n}\right)^k\right)^n$.

Using elementary transformations, we can obtain the following presentation for the sum σ:

$$\sigma=\sum_{k=0}^{n}A_k\left(f\left(n\right)\right),$$

where $A_k\left(f(n)\right)=\binom{n}{k}\frac{1}{2^n}\sum_{l=f(n)}^{T}\binom{T}{l}\left(1-\frac{2(\ln n-c)}{n}\right)^{kl}$.

Using this presentation, we obtain the inequality

$$\sigma\le S_1+S_2+S_3+S_4,$$

where $S_1=\sum_{0\le k\le k_1}A_k(0)$, $S_2=\sum_{k_1<k\le k_2}A_k(0)$, $S_3=\sum_{k_2<k\le k_3}A_k(f(n))$, $S_4=$ $=\sum_{k_3<k\le n}A_k(0)$,

$$k_1=\left[\frac{n}{(\ln n)^q}\right],k_2=\left[\frac{n}{2}\left(1-\frac{1}{\ln n}\right)\right],k_3=\left[\frac{n}{2}\left(1+\frac{1}{\ln n}\right)\right].$$

Let us proceed to the estimations of the sum S_1. We have

$$S_1\le(1+k_1)\cdot 2^{T-n}\cdot\frac{n^{k_1}}{k_1!}.$$

With the help of the Stirling formula and (4), we come to the relation

$$S_1\le c_5(n)\cdot(\ln n)^{-\frac{n\cdot c_6(n)}{(\ln n)^q}}, \tag{21}$$

where $c_6(n)=1-q-\frac{1}{\ln\ln n}\left(1-\ln\left(1-\frac{(\ln n)^q}{n}\right)+\frac{1}{2n}(\ln n)^{1+q}-\frac{q(\ln n)^q\ln\ln n}{2n}\right)$, $c_5(n)=\frac{1}{\sqrt{2\pi}}\cdot\left(1+k_1^{-1}\right)$ and $\lim_{n\to\infty}c_5(n)=(2\pi)^{-1/2}$, $\lim_{n\to\infty}c_6(n)=1-q$.

Let us proceed to the estimation of the sum S_2. Taking into consideration that $A_k(0)$ is a non-decreasing function with the parameter k to be in the scope $k_1<k\le k_2$, we get

$$S_2\le\frac{n}{2}A_{k_2}(0)=\frac{n}{2}\binom{n}{k_2}\frac{1}{2^n}\theta_3(n),$$

where $\theta_3(n)=\left(1+\left(1-\frac{2(\ln n-c)}{n}\right)^{k_2}\right)^T$.

It is obvious that $\theta_3(n)$ satisfies the inequality

$$\theta_3(n)\le\left(1+\exp\left\{-\frac{2k_2\left(\ln n-c\right)}{n}\right\}\right)^T\le\left(1+\frac{1}{n}e^{c+1-\frac{c}{\ln n}+2\frac{\ln n-c}{n}}\right)^T\le c_9(n), \tag{22}$$

where $c_9(n)=\exp\left\{\frac{T}{n}e^{c+1-\frac{c}{\ln n}+2\frac{\ln n-c}{n}}\right\}$.

Then

$$S_2\le\frac{1}{\sqrt{2\pi\left(1-\frac{1}{\ln^2 n}\right)}}\cdot\exp\left\{-\frac{n}{2\ln^2 n}\left(1-\frac{1}{2\ln n}-\frac{1}{2\ln^2 n}-\frac{\ln^3 n}{n}\right)\right\}\cdot c_9(n). \tag{23}$$

Using condition (5) and relation (11) [it follows directly from (4)], we have

$$1<\varliminf_{n\to\infty}c_9(n)\le\varlimsup_{n\to\infty}c_9(n)\le e^{1+c}.$$

To estimate the sum S_4, we use the following expression:

$$S_4\le\frac{n}{2}A_{k_3}(0)=\frac{n}{2}\binom{n}{k_3}\frac{1}{2^n}\left(1+\left(1-\frac{2\left(\ln n-c\right)}{n}\right)^{k_3}\right)^T.$$

Hence,

$$S_4 \le \frac{1}{\sqrt{2\pi\left(1-\frac{1}{\ln^2 n}\right)}} \cdot \exp\left\{-\frac{n}{2\ln^2 n}\left(1-\frac{1}{2\ln n}-\frac{1}{2\ln^2 n}-\frac{\ln^3 n}{n}\right)\right\} \cdot c_{10}(n), \tag{24}$$

where $c_{10}(n) = \exp\left\{\frac{T}{n}e^{c-1+\frac{c}{\ln n}}\right\}$. From condition (5) and relation (11), it follows that

$$1 < \varliminf_{n\to\infty} c_{10}(n) \le \varlimsup_{n\to\infty} c_{10}(n) \le e^{c-1}.$$

Inequalities (23) and (24) result in

$$\begin{aligned} S_2+S_4 \le \frac{1}{\sqrt{2\pi\left(1-\frac{1}{\ln^2 n}\right)}} \cdot \exp\left\{-\frac{n}{2\ln^2 n}\left(1-\frac{1}{2\ln n}-\frac{1}{2\ln^2 n}-\frac{\ln^3 n}{n}\right)\right\} \times \\ \times\left(c_9(n)+c_{10}(n)\right) \quad \le c_7(n)\cdot\exp\left\{-\frac{n}{2\ln^2 n}c_8(n)\right\}, \end{aligned} \tag{25}$$

where $c_8(n) = \left(1-\frac{1}{2\ln n}-\frac{1}{2\ln^2 n}-\frac{\ln^3 n}{n}\right)$, $c_7(n) = \frac{1}{\sqrt{2\pi\left(1-\frac{1}{\ln^2 n}\right)}}\cdot\left(c_9(n)+c_{10}(n),\right)$ $\sqrt{\frac{2}{\pi}} < \varliminf_{n\to\infty} c_7(n) \le \varlimsup_{n\to\infty} c_7(n) \le \frac{e^{c+1}+e^{c-1}}{\sqrt{2\pi}}$, *and* $\lim_{n\to\infty} c_8(n) = 1$.

Combining estimates (21), (25), and the estimate for the sum S_3 found in Lemma 6, we obtain Lemma 7.

We now denote the maximum number of independent critical sets of the matrix A by $S(A)$.

Lemma 8. *If conditions (1) and (2) are satisfied, then, for k, $k = 0, 1, 2, \dots$,*

$$\left|P\{S(A)=k\}-\frac{e^{-\lambda}\cdot\lambda^k}{k!}\right| \le 2MS_1(A)+Q(n,k),$$

where $MS_1(A)$is found in Lemma 2, while $Q(n,k)$ in Lemma 1.

Proof. Taking into account Lemma 1, we get the following relation:

$$\begin{aligned}\left|P\{S(A)=k\}-\frac{e^{-\lambda}\cdot\lambda^k}{k!}\right| &\le \left|P\{S(A)=k\}-P\{\xi_{n,T}(A)=k\}\right|+Q(n,k) = \\ &= \left|P\{S_1(A)+\xi_{n,T}(A)=k\}-P\{\xi_{n,T}(A)=k\}\right|+Q(n,k).\end{aligned}$$

Applying the relations

$$P\{S_1(A)+\xi_{n,T}=k\} = \sum_{l=0}^{k} P\{S_1(A)=l,\, \xi_{n,T}=k-l\},$$

$$\left|P\{S_1(A)=0,\, \xi_{n,T}=k\} - P\{\xi_{n,T}=k\}\right| \le P\{S_1(A)\ge 1\}$$

together with Chebyshev inequality, we find

$$\left|P\{S(A)=k\}-\frac{e^{-\lambda}\cdot\lambda^k}{k!}\right| \le 2P\{S_1(A)\ge 1\}+Q(n,k) \le 2MS_1(A)+Q(n,k).$$

Lemma 8 is proved.

The proof of Theorem 1

To estimate $MS_1(A)$, we introduce the notation

$$\Delta_1(k) = \sum_{1\le i_1<\dots<i_k\le T} \frac{1}{2^n}\prod_{j=1}^{n}\left(1+\prod_{s=1}^{k}\left(1-\frac{2\left(\ln n + x_{i_s j}\right)}{n}\right)\right),$$

$$\Delta_2(k) = \sum_{1\le i_1<\cdots<i_k\le T} \prod_{s=1}^{k}\prod_{j=1}^{n}\left(1-\frac{\ln n + x_{i_s j}}{n}\right)$$

for $k \ge 0$. With the help of Lemma 2, $MS_1(A)$ can be presented as

$$MS_1(A) = \sum_{k=0}^{f(n)} \Delta_1(k) + \sum_{k=f(n)+1}^{T} \Delta_1(k) - \left(\sum_{k=0}^{f(n)} \Delta_2(k) + \sum_{k=f(n)+1}^{T} \Delta_2(k)\right).$$

This leads to

$$MS_1(A) \le \sum_{k=0}^{f(n)} \Delta_1(k) + \sum_{k=f(n)+1}^{T} \Delta_1(k) - \sum_{k=0}^{f(n)} \Delta_2(k).$$

From Lemmas 4 and 5, it follows that

$$\sum_{k=0}^{f(n)} \Delta_1(k) - \sum_{k=0}^{f(n)} \Delta_2(k) \le$$

$$\le \mu(n) + \mu(n)(1+\delta)\Gamma(n) - \left(\mu(n) - c_2(n)\frac{\ln^3 n}{n\cdot\ln\ln n}\right) = \tag{26}$$

$$= \mu(n)(1+\delta)\Gamma(n) + c_2(n)\frac{\ln^3 n}{n\cdot\ln\ln n}.$$

Combining (26) with the estimate for the sum $\sum_{k=f(n)+1}^{T} \Delta_1(k)$ obtained in Lemma 7, we establish the inequality

$$MS_1(A) \le \mu(n)(1+\delta)\Gamma(n) + F(n), \tag{27}$$

where

$$F(n) = c_2(n)\frac{\ln^3 n}{n\cdot\ln\ln n} + c_5(n)\cdot(\ln n)^{-\frac{n\cdot c_6(n)}{(\ln n)^q}} +$$

$$+\left(\frac{c_3(n)}{f(n)}\right)^{f(n)}\frac{1}{\sqrt{f(n)}}c_4(n) + c_7(n)\exp\left\{-\frac{n}{2\ln^2 n}c_8(n)\right\}. \tag{28}$$

According to (27) and Lemma 8,

$$\left|P\{S(A)=k\} - \frac{e^{-\lambda}\cdot\lambda^k}{k!}\right| \le 2(\mu(n)(1+\delta)\Gamma(n) + F(n)) + Q(n,k).$$

Using the explicit expression for $\Gamma(n)$, we can write

$$\left|P\{S(A)=k\} - \frac{e^{-\lambda}\cdot\lambda^k}{k!}\right| \le 2(1+\delta)c_{11}(n)u + 2F(n) + Q(n,k),$$

where

$$c_{11}(n) = \mu(n) + \mu(n)\frac{u}{2}(1+\delta)\exp\{u(1+\delta)\}, \tag{29}$$

The last inequality and the relation $r(A) + s(A) = T$ (see [2, p. 148]) imply

$$\left|P\{r(A)=T-k\} - \frac{e^{-\lambda}\cdot\lambda^k}{k!}\right| \le 2(1+\delta)c(n,k)\frac{\ln^4 n}{n(\ln\ln n)^2},$$

where

$$c(n,k) = c_{11}(n) + \frac{F(n)}{(1+\delta)u} + \frac{Q(n,k)}{2u(1+\delta)}, \tag{30}$$

$c_{11}(n)$, $F(n)$, and $Q(n,k)$ are determined in equalities (29), (28), and (10).

Limited distribution of the rank of a sparse random Boolean matrix

Theorem 2. *I. Let the conditions of Theorem 1 hold and*

$$\lambda \to \lambda_0 \quad if \quad n \to \infty. \tag{31}$$

Then $0 < \lambda_0 < \infty$*, and, for a fixed* k*,* $k = 0, 1, 2, \dots$ *,*

$$P\left\{r(A) = T - k\right\} \to e^{-\lambda_0} \frac{\lambda_0^k}{k!}, \quad n \to \infty. \tag{32}$$

II. If conditions (1), (2), (4), and (31) are satisfied, and $\lambda_0 > 0$*, then* $\lambda_0 < \infty$*, and relation (32) holds true.*

Theorem 1 makes it not difficult to prove Theorem 2.

Corollary. *If condition (1) holds for* $x_{ij} = x$*,* $i = \overline{1, T}$*,* $j = \overline{1, n}$*, where* x *is a fixed number, and* $\frac{T}{n} \to \alpha$ *as* $n \to \infty$*, where* $0 < \alpha < 1$*, then relation (32) holds for* $\lambda_0 = \alpha e^{-x}$*.*

To prove the corollary, it is enough to note that the conditions of the corollary imply immediately the conditions of the first (or the second) part of Theorem 2.

Remark 2. *The result of the above-mentioned corollary was obtained, in particular, in ([1], Theorem 1) and in ([2], Theorem 3.3.1).*

Conclusions

In the paper, the rate of convergence of the distribution of the rank of a sparse Boolean matrix to the Poisson distribution (Theorem 1) has been found. Theorem 1 implies Theorem 2 that generalizes the known result on the limit distribution of the rank of the mentioned matrix in case where the distributions of its elements depend on their positions, and the ratio of the number of rows to the number of columns does not exceed $1 - \gamma(n)$, where $\gamma(n)$ tends slowly to zero and takes only positive values.

Bibliography

1. G.V. Balakin, *Distribution of the rank of random matrices over a finite field* **XIII** (1968), no. 4, Probability theory and its application, 631-641. (in Russian)
2. V.F. Kolchin, *Random Graphs*, Fizmatlit, Moscow, 2004, pp. 256. (in Russian)
3. V.I. Masol, *Moments of the number of solutions of a system of random Boolean equations*, Random Oper. and Stoch. Equations. **1** (1993), no. 2, 171-179.
4. V.I. Masol, *Theorems of invariance for systems of random Boolean equations* (1993), Sixth Intern. Vilnius Conf. of Probability Theory and Math. Statist.: Abstr. of Communic., Vilnius, 19-20.
5. B.A. Sevastyanov, *Course of the Theory of Probability and Mathematical Statistics*, Nauka, Moscow, 1982, pp. 256. (in Russian)

Department of Probability Theory and Mathematical Statistics, Taras Shevchenko Kyiv National University, 6 Academician Glushkov Str., Kyiv 03127, Ukraine.
E-mail: vimasol@ukr.net

Department of Probability Theory and Mathematical Statistics, Taras Shevchenko Kyiv National University, 6 Academician Glushkov Str., Kyiv 03127, Ukraine.
E-mail: Popereshnyak_sv@mail.ru