Asymptotic Properties of Random Contingency Tables with Uniform Margin

Da Wu

Abstract
Let $C \geq 2$ be a positive integer. Consider the set of $n \times n$ non-negative integer matrices whose row sums and column sums are all equal to $Cn$ and let $X = (X_{ij})_{1 \leq i,j \leq n}$ be uniformly distributed on this set. This $X$ is called the random contingency table with uniform margin. In this paper, we study various asymptotic properties of $X = (X_{ij})_{1 \leq i,j \leq n}$ as $n \to \infty$.

Keywords Random contingency tables · Maximum entropy principle · Concentration inequality · Asymptotic statistics

Mathematics Subject Classification (2020) 60F05 · 60C05

1 Introduction

Contingency tables model the dependence structure in large data sets. Mathematically, it is the set of matrices with fixed row and column sum. Let $r = (r_1, \ldots, r_m)$ and $c = (c_1, \ldots, c_n)$ be two positive integer vectors with same total sum of entries $N$, i.e.,

$$\sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j = N.$$ 

The $r$ and $c$ are called row margin and column margin, respectively. Let $\mathcal{M}(r, c)$ be the set of $m \times n$ non-negative integer matrices with $i$th row sum $r_i$ and $j$th column sum $c_j$ for $1 \leq i \leq m$ and $1 \leq j \leq n$, i.e.,
\[ \mathcal{M}(r, c) = \left\{ (x_{ij}) \in \mathbb{Z}_{\geq 0}^{mn} : \sum_{j=1}^{n} x_{kj} = r_k, \sum_{i=1}^{m} x_{ir} = c_r, \forall 1 \leq k \leq m, 1 \leq r \leq n \right\}. \]

The Random Contingency Table \( X = (X_{ij}) \) is defined as the uniform sample from \( \mathcal{M}(r, c) \) and we are interested in various asymptotic statistics of \( X \) when the dimensions grow to infinity.

### 1.1 Notations

1. We use \( f(n) = O(g(n)) \) or \( f(n) \ll g(n) \) to denote the estimate \( |f(n)| \leq Cg(n) \) for some \( C \) independent of \( n \) and for all \( n \geq C \). If this \( C \) depends on some parameter \( k \), then we will write \( f(n) = O_k(g(n)) \) or \( f(n) \ll_k g(n) \).

2. Let \( \text{Geom}(C) \) denote the geometric distribution with mean \( C \). That is, if \( X' \sim \text{Geom}(C) \), then

\[ P(X' = x) = \left( \frac{1}{1+C} \right) \left( \frac{C}{1+C} \right)^x \]

for \( x \geq 0 \).

3. For every two probability distributions \( \mu_1 \) and \( \mu_2 \) on the countable sample space, the total variation distance metric is defined as

\[ \|\mu_1, \mu_2\|_{TV} := \frac{1}{2} \sum_{x \in \Omega} |\mu_1(x) - \mu_2(x)|. \]

4. For any measurable set \( A \), let \( \mathbb{E}[X; A] \) denote the expectation of \( X \) over \( A \), i.e.,

\[ \mathbb{E}[X; A] = \mathbb{E}[X 1_A] = \int_A X \, d\mathbb{P}. \]

### 1.2 Setup and Statements of the Main Results

In this paper, we consider the discrete random contingency table with uniform margin. Namely, all the row sums and column sums are equal. Precisely, let \( C \geq 2 \) be a positive integer and let

\[ \tilde{r} = \tilde{c} = (Cn, \ldots, Cn) \in \mathbb{Z}_{\geq 0}^n, \]

where \( n \) is the dimension of contingency tables. Denote the set of \( n \times n \) non-negative integer matrices whose row sums and column sums are all equal to \( Cn \) by \( \mathcal{M}(Cn, n) \). Throughout this paper, the matrix-valued random variable \( X = (X_{ij})_{1 \leq i, j \leq n} \) is uniformly distributed on \( \mathcal{M}(Cn, n) \), i.e.,

\[ P(X = D) = \frac{1}{\#\mathcal{M}(Cn, n)}, \quad \forall D \in \mathcal{M}(Cn, n). \]
We are interested in various asymptotic statistics of $X$ as $n \to \infty$. By symmetry, all the $X_{ij}$, $1 \leq i, j \leq n$, have the same marginal distribution. It was proved in [1, Theorem 2.1] that the marginal distribution of a single entry of $X$ will converge to $\text{Geom}(C)$ in total variation distance and this can be viewed as the discrete counterpart of [2, Theorem 1].

**Theorem 1.1** (Marginal Distribution, [1, Theorem 2.1]) Let $X = (X_{ij})_{1 \leq i, j \leq n}$ be uniformly distributed on $\mathcal{M}(Cn, n)$, then

$$X_{11} \xrightarrow{\text{as } n \to \infty} \text{Geom}(C),$$

where convergence is in total variation distance. Moreover, for any $\varepsilon > 0$,

$$\|X_{11}, \text{Geom}(C)\|_{TV} = O(n^{-1/2+\varepsilon}).$$

The proof of this theorem depends heavily on the Maximum Entropy Principle, which was introduced by I. J. Good in [3]. More recently, Alexander Barvinok managed to apply this principle to the context of random contingency tables and answer the question *What does a random contingency table looks like?* He found that as dimensions of the matrix grow, the random contingency table behaves much like the matrix of independent geometric random variables. Precise meaning of *much like* was given in his sequence of papers [4–6] and we will review these material carefully in the later section.

Going one step further, it is shown in [1, Theorem A.3] that if we take a $k \times k$ sub-matrix of $X$, where $k = o\left(\frac{n^{1/2}}{(\log n)^{1/2}}\right)$, then the joint distribution of this sub-matrix converges in total variation distance to the $k \times k$ matrix of independent geometric random variables with mean $C$. This can be viewed as the discrete counterpart of [2, Theorem 4].

**Theorem 1.2** (Joint Distribution, [1, Theorem A.3]) Let $W_k$ denote the projection of a uniform sample $X$ onto the $k \times k$ sub-matrix of its first $k$ rows and columns and let $Y_k$ be the $k \times k$ matrix of independent geometric random variables with mean $C$. When $k = o\left(\frac{n^{1/2}}{(\log n)^{1/2}}\right)$, we have

$$\|W_k, Y_k\|_{TV} = o(1).$$

**Remark 1.3** Notice that the above theorem holds in a greater generality in the sense that as long as we are considering $k^2 = o\left(\frac{n}{\log n}\right)$ entries in $X$, they are asymptotically independent geometric random variables with mean $C$. In other words, the shape does not matter.

For the convenience of readers, we will provide self-contained proofs of both Theorem 1.1 and Theorem 1.2 in Sect. 2. Next, we show that the moments of entries of $X$ converge to the moments of the i.i.d. $\text{Geom}(C)$ variables.
Theorem 1.4 (Moment Convergence) Let \((i_1, j_1), \ldots, (i_L, j_L)\) be a fixed sequence of indices and let \(\alpha_1, \ldots, \alpha_L\) be \(L\) fixed positive integers. Let \(X = (X_{ij})_{1 \leq i, j \leq n}\) be uniformly distributed on \(\mathcal{M}(Cn, n)\), then

\[
\mathbb{E} \left[ \prod_{k=1}^{L} X_{i_k, j_k}^{\alpha_k} \right] \to \mathbb{E} \left[ \prod_{k=1}^{L} Y_{i_k}^{\alpha_k} \right],
\]

where \(Y_1, \ldots, Y_k\) are i.i.d. Geom\((C)\).

Furthermore, we prove that the maximum entries of \(X\) are of the same order as that of \(n^2\) i.i.d. Geom\((C)\) variables.

Theorem 1.5 (Maximum Entry) For any fixed \(\varepsilon > 0\),

\[
\lim_{n \to \infty} \mathbb{P} \left( \max_{1 \leq i, j \leq n} X_{ij} > \frac{1}{\log \left( \frac{C+1}{C} \right)} \log \left( \frac{C}{1 + C \cdot n^{2+\varepsilon}} \right) \right) = 0.
\]

By the above theorem, the distribution of each individual \(X_{ij}\) is compactly supported on \(\left[ 0, \frac{(2+\varepsilon) \log n}{\log \left( \frac{C+1}{C} \right)} \right]\) asymptotically. Thus, by the concentration of measure for large Wishart Matrix (see [7]) and the Marčenko–Pastur distribution (see [8]), we can obtain the following result on the limiting empirical singular value distribution of \(X\). The proof of Theorem 1.6 will be presented in Sect. 5.

Theorem 1.6 (Limiting Empirical Singular Value Distribution) Let \(X\) be uniformly distributed on \(\mathcal{M}(Cn, n)\) and let \(\tilde{\Upsilon} = \frac{1}{\sqrt{n}} (X - \mathbb{E}[X])\). Let \(0 \leq \sigma_1(\tilde{\Upsilon}) \leq \sigma_2(\tilde{\Upsilon}) \leq \ldots \leq \sigma_n(\tilde{\Upsilon})\) be singular values of \(\tilde{\Upsilon}\) and let

\[
\mu_n^s(\tilde{\Upsilon}) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\sigma_i(\tilde{\Upsilon})}
\]

be the empirical singular value distribution of \(\tilde{\Upsilon}\). Then

\[
\mu_n^s(\tilde{\Upsilon}) \to \frac{\sqrt{4C(1+C) - y^2}}{\pi C(1+C)} 1_{[0,2\sqrt{C(1+C)}]}(y) dy
\]

weakly in probability.

1.3 Comparison with Known Literature and Open Problems

Our results can be viewed as discrete counterparts of the work of Chatterjee, Diaconis and Sly [2], where they studied the doubly stochastic matrix, which is the set of non-negative real-valued matrices whose row and column sums are all equal to 1. It is also worth mentioning that in [1], authors studied the non-uniform margin case and obtained
the phase transition regarding some asymptotic statistics. The case of non-uniform binary contingency tables was studied in [9]. From a combinatorial perspective, recent works of [10] and [11] compared sharp asymptotics of $\#M(r, c)$ in non-uniform case with the so-called independence heuristic introduced in [12]. It was shown in [10] that in binary non-uniform case, the independence heuristic overestimates the $\#M(r, c)$ whereas in [11], it was proved that the independence heuristic underestimates the $\#M(r, c)$. All of the above results are based on the Maximum entropy principle, which we will set up carefully in Sect. 2.

In non-uniform case of two different margins, due to the loose estimate of $\#M(r, c)$, it is still unknown how to obtain the moment convergence (similar to Theorem 1.4). This is a prerequisite to prove the central limit theorem for certain rows and columns (see [1] for detailed discussions). It is even more interesting to understand what would be the limiting empirical singular value distribution in this case.

## 2 Proof of Theorem 1.1 and 1.2

In this section, we provide self-contained proofs of both Theorem 1.1 and 1.2. First recall the notion of Typical Table introduced by Barvinok in [4].

**Definition 2.1** (Typical Table, [4, Definition 1.2]) Fix row margin $r = (r_1, \ldots, r_m) \in \mathbb{Z}_{>0}^m$ and column margin $c = (c_1, \ldots, c_n) \in \mathbb{Z}_{>0}^n$ and let

$$\mathcal{P}(r, c) = \left\{ (x_{ij}) \in \mathbb{R}^{mn}_{\geq 0} : \sum_{j=1}^{n} x_{kj} = r_k, \sum_{i=1}^{m} x_{ir} = c_r, \forall 1 \leq k \leq m, 1 \leq r \leq n \right\}.$$ 

For each $M = (m_{ij}) \in \mathcal{P}(r, c)$, define

$$g(M) = \sum_{1 \leq i \leq m, 1 \leq j \leq n} f(m_{ij}),$$

where $f : [0, \infty) \to [0, \infty)$ is defined by

$$f(x) = (x + 1) \log(x + 1) - x \log x.$$ 

The typical table $Z \in \mathcal{P}(r, c)$ is defined as

$$Z := \arg \max_{M \in \mathcal{P}(r, c)} g(M).$$

**Remark 2.2** 1. $f(x) = (x + 1) \log(x + 1) - x \log x$ is the Shannon–Gibbs–Boltzmann entropy of $\text{Geom}(x)$.

2. The function $g$ in the above definition is strictly concave; hence, it achieves the unique maximum on $\mathcal{P}(r, c)$. Therefore, the typical table is well-defined.

**Lemma 2.3** The typical table for $M(Cn, n)$ is $C \cdot I_n$, where $I_n$ is the $n \times n$ matrix with all entries equal to 1.
**Proof** Notice that

\[ \nabla g = (\log(1 + 1/m_{ij}))_{1 \leq i,j \leq n}, \]

and the transportation polytope for \( M(Cn, n) \) is defined by the intersections of hyperplanes in \( \mathbb{R}^{n^2}_{\geq 0} \) given by

\[
\begin{align*}
    h_{i\bullet} &:= \left( \sum_{k=1}^{m} m_{ik} \right) - C = 0 \quad \text{for all } 1 \leq i \leq n, \\
    h_{\bullet j} &:= \left( \sum_{k=1}^{m} m_{kj} \right) - C = 0 \quad \text{for all } 1 \leq j \leq n.
\end{align*}
\]

The gradient \( \nabla h_{i\bullet} \) is the \( n \times n \) matrix \( E_{i\bullet} \), which has 1’s in the \( i \)-th row and 0’s elsewhere. Similarly, the gradient \( \nabla h_{\bullet j} \) is the \( n \times n \) matrix \( E_{\bullet j} \), which has 1’s in the \( j \)-th column and 0’s elsewhere.

Hence, by multivariate Lagrange’s method, when evaluated at typical table \( Z = (z_{ij}) \), there exist some non-negative constants \( \lambda_1, \ldots, \lambda_n \) and \( \mu_1, \ldots, \mu_n \) such that

\[ \log(1 + 1/z_{ij}) = \lambda_i + \mu_j \quad \text{for all } 1 \leq i,j \leq n. \]

Since all the row sums and column sums are equal to \( Cn \), by symmetry,

\[ \log(1 + 1/z_{ij}) = \lambda + \mu \quad \text{for all } 1 \leq i,j \leq n. \]

Therefore, all the \( z_{ij} \)’s are equal to \( C \). \( \square \)

**Theorem 2.4** ([4, Theorem 1.7, 2.1]) Fix margins \( \mathbf{r} = (r_1 \ldots, r_m) \in \mathbb{Z}_{>0}^m \) and \( \mathbf{c} = (c_1, \ldots, c_n) \in \mathbb{Z}_{>0}^n \) with \( \sum_{i=1}^{m} r_i = \sum_{j=1}^{n} c_j = N \). Let \( Z = (z_{ij}) \) be the typical table for \( M(\mathbf{r}, \mathbf{c}) \) and \( Y = (Y_{ij}) \) be the \( m \times n \) matrix of independent geometric random variables with \( Y_{ij} \sim \text{Geom}(z_{ij}) \). Then, we have the following:

1. There exists some absolute constant \( \gamma > 0 \) such that

\[ N^{-\gamma(m+n)}e^{g(Z)} \leq \#M(\mathbf{r}, \mathbf{c}) \leq e^{g(Z)}. \]

2. Conditioned on being in \( M(\mathbf{r}, \mathbf{c}) \), the matrix \( Y \) is uniformly distributed on \( M(\mathbf{r}, \mathbf{c}) \), i.e.,

\[ P(Y = D) = e^{-g(Z)} \quad \text{for all } D \in M(\mathbf{r}, \mathbf{c}). \]

3. Combining 1 and 2, we have that

\[ P(Y \in M(\mathbf{r}, \mathbf{c})) = e^{-g(Z)} \cdot \#M(\mathbf{r}, \mathbf{c}) \geq N^{-\gamma(m+n)}. \]
Remark 2.5 By 2 and 3 in the above theorem, for fixed measurable set $C \subseteq \mathbb{R}_+^{m \times n}$, we have the following transformation of mass inequality:

$$
\mathbb{P}(Y \in C) \geq \mathbb{P}(Y \in C \mid Y \in M(r, c)) \cdot \mathbb{P}(Y \in M(r, c))
= \mathbb{P}(X \in C) \cdot \mathbb{P}(Y \in M(r, c))
\geq \mathbb{P}(X \in C) \cdot N^{-\gamma(m+n)}.
$$

(2.1)

Now, we are ready to prove the Theorem 1.1.

Proof of Theorem 1.1 Let $Y = (Y_{ij})$ with $Y_{ij} \sim$ i.i.d. Geom$(C)$. By Theorem 2.4, we have

$$
\mathbb{P}(Y \in M(C, n)) \geq (Cn^2)^{-\gamma'}
$$

for some absolute constant $\gamma'$. Fix a measurable set $A \subseteq [0, \infty)$, $\mathbb{1}_{[Y_{ij} \in A]}$ is a sequence of i.i.d. random variables. Let $\mathcal{F} = \{(i, j), 1 \leq i, j \leq n\}$ with $\#\mathcal{F} = n^2$. Fix $\varepsilon > 0$, and by Hoeffding inequality on $\frac{1}{\#\mathcal{F}} \sum_{(i, j) \in \mathcal{F}} \mathbb{1}_{[Y_{ij} \in A]}$,

$$
\mathbb{P}\left(\left|\frac{1}{\#\mathcal{F}} \sum_{(i, j) \in \mathcal{F}} \mathbb{1}_{[Y_{ij} \in A]} - \mathbb{P}(Y_{11} \in A)\right| > \frac{1}{2} n^{-\frac{1}{2}+\varepsilon}\right) \leq \exp\left(-\frac{2 \cdot \frac{1}{2} n^{-1+2\varepsilon}}{\#\mathcal{F} \cdot \left(\frac{1}{\#\mathcal{F}}\right)^2}\right)
= \exp\left(-\frac{1}{2} n^{-1+2\varepsilon} \cdot \#\mathcal{F}\right).
$$

Equivalently,

$$
\mathbb{P}\left(\left|\frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{1}_{[Y_{ij} \in A]} - \mathbb{P}(Y_{11} \in A)\right| > \frac{1}{2} n^{-\frac{1}{2}+\varepsilon}\right) \leq \exp\left(-\frac{1}{2} n^{1+2\varepsilon}\right).
$$

Next,

$$
|\mathbb{P}(X_{11} \in A) - \mathbb{P}(Y_{11} \in A)|
\leq \mathbb{E}\left[\frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{1}_{[X_{ij} \in A]} - \mathbb{P}(Y_{11} \in A)\right]
\leq \mathbb{E}\left[\frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{1}_{[X_{ij} \in A]} - \mathbb{P}(Y_{11} \in A)\right]
\leq \frac{1}{2} n^{-\frac{1}{2}+\varepsilon} \mathbb{P}\left(\left|\frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{1}_{[X_{ij} \in A]} - \mathbb{P}(Y_{11} \in A)\right| \leq \frac{1}{2} n^{-\frac{1}{2}+\varepsilon}\right)
+ \mathbb{P}\left(\left|\frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{1}_{[X_{ij} \in A]} - \mathbb{P}(Y_{11} \in A)\right| > \frac{1}{2} n^{-\frac{1}{2}+\varepsilon}\right).
$$
\[
\leq \frac{1}{2} n^{-\frac{1}{2} + \varepsilon} + \mathbb{P}\left( \left| \frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{I}_{\{X_{ij} \in A\}} - \mathbb{P}(Y_{11} \in A) \right| > \frac{1}{2} n^{-\frac{1}{2} + \varepsilon} \right).
\]

Notice that

\[
\mathbb{P}\left( \left| \frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{I}_{\{Y_{ij} \in A\}} - \mathbb{P}(Y_{11} \in A) \right| > \frac{1}{2} n^{-\frac{1}{2} + \varepsilon} \right) \geq \mathbb{P}\left( \left| \frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{I}_{\{X_{ij} \in A\}} - \mathbb{P}(Y_{11} \in A) \right| > \frac{1}{2} n^{-\frac{1}{2} + \varepsilon} \right) \cdot \mathbb{P}(Y \in \mathcal{M}(Cn, n)) \cdot (Cn^2)^{-\gamma n}.
\]

The above estimate follows from the transformation of mass inequality (2.1). The key observation is that if \(Y\) is conditioned on being in \(\mathcal{M}(Cn, n)\), then \(Y\) is uniformly distributed and has the same law as \(X\). Hence,

\[
|\mathbb{P}(X_{11} \in A) - \mathbb{P}(Y_{11} \in A)| \leq \frac{1}{2} n^{-\frac{1}{2} + \varepsilon} + (Cn^2)^{\gamma n} \cdot \mathbb{P}\left( \left| \frac{1}{n^2} \sum_{1 \leq i, j \leq n} \mathbb{I}_{\{Y_{ij} \in A\}} - \mathbb{P}(Y_{11} \in A) \right| > \frac{1}{2} n^{-\frac{1}{2} + \varepsilon} \right)
\]

\[
\leq \frac{1}{2} n^{-\frac{1}{2} + \varepsilon} + (Cn^2)^{\gamma n} \exp\left( -\frac{1}{2} n^{1+2\varepsilon} \right).
\]

Therefore, for any fixed \(\varepsilon\),

\[
\|X_{11}, Y_{11}\|_{TV} = O(n^{-\frac{1}{2} + \varepsilon}),
\]

which is equivalent to

\[
\|X_{11}, \text{Geom}(C)\|_{TV} = O(n^{-\frac{1}{2} + \varepsilon}).
\]

This completes the proof of Theorem 1.1. \(\square\)
Furthermore, we can also study the joint distribution of the sub-matrix of $X$. Again, let $Y = (Y_{ij})$ be the matrix of i.i.d. $\text{Geom}(C)$. Fix some $k = k(n)$ and let

$$W_{\ell_1\ell_2} = \begin{pmatrix} Y_{\ell_1k,\ell_2k} & Y_{\ell_1k,\ell_2k+1} & \cdots & Y_{\ell_1k,\ell_2k+(k-1)} \\ Y_{\ell_1k+1,\ell_2k} & Y_{\ell_1k+1,\ell_2k+1} & \cdots & Y_{\ell_1k+1,\ell_2k+(k-1)} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{\ell_1k+(k-1),\ell_2k} & Y_{\ell_1k+(k-1),\ell_2k+1} & \cdots & Y_{\ell_1k+(k-1),\ell_2k+(k-1)} \end{pmatrix}.$$ 

Also, let

$$V_{\ell_1\ell_2} = \begin{pmatrix} X_{\ell_1k,\ell_2k} & X_{\ell_1k,\ell_2k+1} & \cdots & X_{\ell_1k,\ell_2k+(k-1)} \\ X_{\ell_1k+1,\ell_2k} & X_{\ell_1k+1,\ell_2k+1} & \cdots & X_{\ell_1k+1,\ell_2k+(k-1)} \\ \vdots & \vdots & \ddots & \vdots \\ X_{\ell_1k+(k-1),\ell_2k} & X_{\ell_1k+(k-1),\ell_2k+1} & \cdots & X_{\ell_1k+(k-1),\ell_2k+(k-1)} \end{pmatrix}.$$ 

**Proof of Theorem 1.2** Let $A \subseteq \mathbb{R}^{k^2}$ be a measurable set. By Hoeffding inequality,

$$\mathbb{P} \left( \frac{1}{\left[ \frac{n-1}{k} \right]^2} \sum_{\ell_1=1}^{\left[ \frac{n-1}{k} \right]} \sum_{\ell_2=1}^{\left[ \frac{n-1}{k} \right]} \mathbb{I}_{\{W_{\ell_1\ell_2} \in A\}} - \mathbb{P}\left(W_{11} \in A\right) \right) > \frac{1}{2} \varepsilon$$

$$\leq \exp \left( -\frac{\varepsilon^2}{\left[ \frac{n-1}{k} \right]^2} \right) = \exp \left( -\frac{\varepsilon^2}{\left[ \frac{n-1}{k} \right]^2} \right).$$

Again, by (2.1),

$$\mathbb{P} \left( \frac{1}{\left[ \frac{n-1}{k} \right]^2} \sum_{\ell_1=1}^{\left[ \frac{n-1}{k} \right]} \sum_{\ell_1=1}^{\left[ \frac{n-1}{k} \right]} \mathbb{I}_{\{V_{\ell_1\ell_2} \in A\}} - \mathbb{P}\left(W_{11} \in A\right) \right) > \frac{1}{2} \varepsilon$$

$$\leq (Cn^2)^{\nu n} \cdot \exp \left( -\frac{1}{2} \varepsilon^2 \cdot \left[ \frac{n-1}{k} \right]^2 \right).$$

By exchangeability of entries of $X$,

$$|\mathbb{P}(V_{11} \in A) - \mathbb{P}(W_{11} \in A)| = \left| \mathbb{E} \left( \frac{1}{\left[ \frac{n-1}{k} \right]^2} \sum_{\ell_1=1}^{\left[ \frac{n-1}{k} \right]} \sum_{\ell_2=1}^{\left[ \frac{n-1}{k} \right]} \mathbb{I}_{\{V_{\ell_1\ell_2} \in A\}} - \mathbb{P}\left(W_{11} \in A\right) \right) \right|$$
\[
\begin{align*}
\leq & \mathbb{E} \left[ \frac{1}{\frac{n-1}{k}} \sum_{\ell_1 = 1}^{n-1} \sum_{\ell_2 = 1}^{n-1} \mathbb{1}_{\{V_{\ell_1, \ell_2} \in A\}} - \mathbb{P}(W_{11} \in A) \right] \\
\leq & \frac{1}{2} \varepsilon + (Cn^2)^{\gamma'} n \cdot \exp \left( - \frac{1}{2} \varepsilon^2 \cdot \left[ \frac{n-1}{k} \right]^2 \right).
\end{align*}
\]

Hence, when \( k = o \left( \frac{n^2}{(\log n)^2} \right) \),

\[
\left| \mathbb{P}(V_{11} \in A) - \mathbb{P}(W_{11} \in A) \right| \leq \frac{1}{2} \varepsilon + o(1).
\]

Taking \( \varepsilon \downarrow 0 \) and this completes the proof. \( \square \)

### 3 Convergence of Moments

In this section, we show that finite moments of entries of \( X \) converge to those of independent Geom(\( C \)) variables. We first show that the uniform margin maximizes the number of matrices among all the possible margins.

**Lemma 3.1** Let \( a = (a_1, \ldots, a_m) \) be a \( m \)-dimensional positive integer vector and let \( N = a_1 + \cdots + a_m \). Let \( b_r = (r, N-r) \) be a 2-dimensional positive integer vector. We have that

\[
\begin{cases}
\# M(a, b_r) \leq \# M(a, b_{N/2}) & \text{if } N \equiv 0 \mod 2, \\
\# M(a, b_r) \leq \# M(a, b_{(N+1)/2}) = \# M(a, b_{(N-1)/2}) & \text{if } N \equiv 1 \mod 2.
\end{cases}
\]

**Proof** Observe that

\[
\# M(a, b_r) = N_m(r; a_1, \ldots, a_m),
\]

where \( N_m(r; a_1, \ldots, a_m) \) denotes the number of ways of partitioning \( r \) into \( m \) pieces, each of which has size less or equal to \( a_i \). Next, consider the generating function of \( N_m(r; a_1, \ldots, a_m) \),

\[
\prod_{\ell=1}^{m} (1 + q + \cdots + q^{a_{\ell}}) = \sum_{r=0}^{a_1 + \cdots + a_m} N_m(r; a_1, \ldots, a_m) q^r = \sum_{r=0}^{N} N_m (r; a_1, \ldots, a_m) q^r.
\]
It suffices to show that
\[
N_m(r; a_1, \ldots, a_m) \leq N_m(N/2; a_1, \ldots, a_m) \quad \text{if } N \text{ even,}
\]
\[
N_m(r; a_1, \ldots, a_m) \leq N_m((N + 1)/2; a_1, \ldots, a_m) = N_m((N - 1)/2; a_1, \ldots, a_m) \quad \text{if } N \text{ odd.}
\]

Without loss of generality, assume \(a_1 \leq a_2 \leq \cdots \leq a_m\). We prove (3.1) by induction. Write
\[
\prod_{\ell=1}^{m'} (1 + q + q^2 + \cdots + q^{a_\ell}) = c_0^{m'} + c_1^{m'} q + \cdots + c_i^{m'} q^i + \cdots + c_{a_1+\cdots+a_m'} q^{a_1+\cdots+a_m'}
\]
for \(1 \leq m' \leq m\). Our induction hypotheses are that

1. If \(\sum_{n=1}^{m'} a_n\) is even, then
   \[
   1 = c_0^{m'} \leq c_1^{m'} \leq \cdots \leq c_{a_1+\cdots+a_{m'+1}}^{m'} \geq \cdots \geq 1 = c_{a_1+\cdots+a_{m'}}^{m'}.
   \]

2. If \(\sum_{n=1}^{m'} a_n\) is odd, then
   \[
   1 = c_0^{m'} \leq c_1^{m'} \leq \cdots \leq c_{a_1+\cdots+a_{m'+1}}^{m'} = c_{a_1+\cdots+a_{m'-1}}^{m'} \geq \cdots \geq 1 = c_{a_1+\cdots+a_{m'}}^{m'}.
   \]

3. For all \(0 \leq j \leq a_1 + \cdots + a_{m'}\),
   \[
   c_j^{m'} = c_{a_1+\cdots+a_{m'-j}}^{m'}.
   \]

When \(m' = 1\), all the coefficients are equal to 1. It satisfies the hypotheses. Suppose the case of \(m'\) is true and we consider the case of \(m' + 1\). By simply expanding the product, we have that
\[
\prod_{\ell=1}^{m'+1} (1 + q + \cdots + q^{a_\ell})
= \left(1 + c_1^{m'} q + \cdots + c_i^{m'} q^i + \cdots + c_{a_1+\cdots+a_m'} q^{a_1+\cdots+a_m'}\right) (1 + q + \cdots + q^{a_{m'+1}})
= 1 + c_1^{m'} q + \cdots + c_i^{m'} q^i + \cdots + c_{a_1+\cdots+a_m'} q^{a_1+\cdots+a_m'}
+ q + c_1^{m'} q^2 + \cdots + c_i^{m'} q^{i+1} + \cdots + c_{a_1+\cdots+a_m'} q^{a_1+\cdots+a_m'+1}
+ \cdots
+ q^{a_{m'+1}} + c_1^{m'} q^{a_{m'+1}+1} + \cdots + c_i^{m'} q^{a_{m'+1}+i} + \cdots + c_{a_1+\cdots+a_m'} q^{a_1+\cdots+a_m'+a_{m'+1}}.
\]

After summing up coefficients and applying the induction hypothesis, we are done. □
Lemma 3.2 Let \( a = (a_1, \ldots, a_m) \) and \( b = (b_1, \ldots, b_n) \) be two positive integer vectors such that
\[
\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j = N.
\]

Suppose \( b' \) is a vector such that
\[
\begin{cases}
  b' := \left( \frac{b_1+b_2}{2}, \frac{b_1+b_2}{2}, b_3, \ldots, b_n \right) & \text{if } b_1 + b_2 \text{ is even,} \\
  b' := \left( \frac{b_1+b_2-1}{2}, \frac{b_1+b_2+1}{2}, b_3, \ldots, b_n \right) & \text{if } b_1 + b_2 \text{ is odd,}
\end{cases}
\]
then we have that
\[
\#M(a, b) \leq \#M(a, b').
\]

Proof Let
\[
\mathcal{A} = \left\{ (\hat{a}_i)_{i=1,\ldots,m} : \sum_{i=1}^{m} \hat{a}_i = N - b_1 - b_2, 0 \leq \hat{a}_i \leq a_i \right\}
\]
be the set of possible truncated row margins (except the first two columns) and we have
\[
\#M(a, b) = \sum_{(\hat{a}_i)_{1\leq i \leq m} \in \mathcal{A}} \#M((\hat{a}_i)_{1\leq i \leq m}, (b_i)_{i\geq 3}) \cdot \#M((a_i - \hat{a}_i)_{1\leq i \leq m}, (b_i)_{i=1,2}).
\]

Similarly,
\[
\#M(a, b') = \sum_{(\hat{a}_i)_{1\leq i \leq m} \in \mathcal{A}} \#M((\hat{a}_i)_{1\leq i \leq m}, (b_i)_{i\geq 3}) \cdot \#M((a_i - \hat{a}_i)_{1\leq i \leq m}, (b'_i)_{i=1,2}),
\]
where
\[
\begin{cases}
  b'_1 = b'_2 = \frac{b_1+b_2}{2} & \text{if } b_1 + b_2 \text{ is even,} \\
  b'_1 = \frac{b_1+b_2-1}{2}, b'_2 = \frac{b_1+b_2+1}{2} & \text{if } b_1 + b_2 \text{ is odd.}
\end{cases}
\]

By Lemma 3.1,
\[
\#M((a_i - \hat{a}_i)_{1\leq i \leq m}, (b_i)_{i=1,2}) \leq \#M((a_i - \hat{a}_i)_{1\leq i \leq m}, (b'_i)_{i=1,2}),
\]
which implies that
\[
\#M(a, b) \leq \#M(a, b').
\]
This completes the proof. \( \square \)
Corollary 3.3  Fix positive integers $m$, $n$ and $N$. Let $\ell_1$ and $\ell_2$ be such that

$$N \equiv \ell_1 \mod m \quad \text{and} \quad N \equiv \ell_2 \mod n.$$ 

Let

$$\mathcal{P} = \left\{ (a, b) \in \mathbb{Z}_{>0}^m \times \mathbb{Z}_{>0}^n : a = (a_1, \ldots, a_m), b = (b_1, \ldots, b_n), \sum_{i=1}^{m} a_j = \sum_{j=1}^{n} b_j = N \right\}$$

be the set of all possible margins and let

$$a^* = \begin{cases} N - \ell_1 + 1, & \ell_1 \text{ entries} \\ \frac{m}{m - \ell_1} & m - \ell_1 \text{ entries} \end{cases},$$

$$b^* = \begin{cases} N - \ell_2 + 1, & \ell_2 \text{ entries} \\ \frac{n}{n - \ell_2} & n - \ell_2 \text{ entries} \end{cases}.$$ 

Then, we have

$$#\mathcal{M}(a^*, b^*) = \max_{(a, b) \in \mathcal{P}} #\mathcal{M}(a, b).$$

Proof  Starting with any pair of margins $(a, b) \in \mathcal{P}$, applying the Lemma 3.2 repeatedly on $(a, b)$ proves the above result. \qed

Next, consider the set $\mathcal{R}(Cn, r), 1 \leq r < n$, which consists of all the $r \times n$ non-negative integer matrices with row sums $Cn$. There are no restrictions on column sums. Consider the following two probability measures on $\mathcal{R}(Cn, r)$:

1. $\mu_r$, a uniform measure on $\mathcal{R}(Cn, r)$.
2. $\nu_r$, a measure induced by the uniform measure on $\mathcal{M}(Cn, n)$.

Our next task is to understand the Radon–Nikodym derivative $\frac{d\nu_r}{d\mu_r}$. Precisely, our goal is to find a uniform finite upper bound on $\frac{d\nu_r}{d\mu_r}$ for all $n$.

Lemma 3.4  Suppose $0 < \gamma < 1$ is fixed and $m = (\gamma + o(1))n$. Then,

$$\binom{n}{m} = \exp \left[ (h(\gamma) + o(1))n \right]$$

$$= \exp \left[ h(\gamma)n + \frac{1}{2} \log \frac{1}{\gamma} - \frac{1}{2} \log(2\pi(1 - \gamma)n) \right],$$

where

$$h(\gamma) = \log \frac{1}{\gamma} + (1 - \gamma) \log \frac{1}{1 - \gamma}.$$
This is the Exercise 1.2.1 in Tao’s book *Topics in Random Matrix Theory* [13]. The proof is trivial by the Stirling Formula \( n! = \sqrt{2\pi n} \left( \frac{n}{e} \right)^n (1 + o(1)) \). Here, we write down the error term \( o(1) \) in (3.2) explicitly as in (3.3) in order to do a careful analysis later.

Next, recall an important theorem by Canfield and McKay [14] on precise asymptotic enumeration of the number of non-negative integer matrices with equal row and column sums.

**Theorem 3.5** ([14, Theorem 1]) Let \( \mathbf{r} = (r, \ldots, r) \in \mathbb{Z}^m_0 \) and \( \mathbf{c} = (c, \ldots, c) \in \mathbb{Z}^n_0 \) with \( mr = nc \). Let \( \lambda = \frac{r}{n} = \frac{c}{m} \) and let \( a, b > 0 \) be constants such that \( a + b < \frac{1}{2} \). Suppose \( m, n \to \infty \) in such a way that

\[
\frac{(1 + 2\lambda)^2}{4\lambda(1 + \lambda)} \left( 1 + \frac{5m}{6n} + \frac{5n}{6m} \right) \leq a \log n,
\]

then

\[
\# \mathcal{M}(\mathbf{r}, \mathbf{c}) = \frac{(n+r-1)^m (m+c-1)^n}{(mn+\lambda mn-1)^c \lambda^{mn}} \exp \left( \frac{1}{2} + O(n^{-b}) \right).
\]

where \( A = \frac{\lambda(1+\lambda)}{2} \).

**Proposition 3.6** Fix a positive integer \( r \) and for \( n > r \),

\[
\frac{dv_r}{d\mu_r} \leq (1 + o(1)) e^\frac{r}{2}.
\]

**Proof** For each \( (X_{ij})_{1 \leq i \leq r, 1 \leq j \leq n} \), \( v_r \) gives weights to \( (X_{ij})_{1 \leq i \leq r, 1 \leq j \leq n} \) proportional to \( \# \mathcal{M}((Cn)_{n-r}, (Cn-\sum_{j=1}^{r} X_{ij})_{j=1, \ldots, n}) \). Here, \( (Cn)_{n-r} \) just means the \( n-r \) dimensional vector with all entries \( Cn \). Therefore,

\[
\frac{dv_r}{d\mu_r} ((X_{ij})_{1 \leq i \leq r, 1 \leq j \leq n}) \propto \# \mathcal{M}((Cn)_{n-r}, (Cn-\sum_{i=1}^{r} X_{ij})_{j=1, \ldots, n}).
\]

Next, we determine the constant in the above proportionality. Notice that

\[
\# \mathcal{M}(Cn, n) = \sum_{(X_{ij})_{1 \leq i \leq r, 1 \leq j \leq n} \in \mathcal{R}(Cn, r)} \# \mathcal{M}((Cn)_{n-r}, (Cn-\sum_{i=1}^{r} X_{ij})_{j=1, \ldots, n}).
\]

\[
= \# \mathcal{R}(Cn, r) \int_{\mathcal{R}(Cn, r)} \# \mathcal{M}((Cn)_{n-r}, (Cn-\sum_{i=1}^{r} X_{ij})_{j=1, \ldots, n}) d\mu_r ((X_{ij})_{1 \leq i \leq r, 1 \leq j \leq n}).
\]
\[
1 = v_r(\mathcal{R}(C_n, r)) = \frac{\#\mathcal{R}(C_n, r)}{\#\mathcal{M}(C_n, n)} \int_{\mathcal{R}(C_n, r)} \#\mathcal{M}((C_n)_{n-r}, (C_n - \sum_{i=1}^{r} X_{ij})_{j=1,...,n}) \times d\mu_r((X_{ij})_{1 \leq i \leq r, 1 \leq j \leq n}),
\]

and

\[
\frac{dv_r}{d\mu_r}((X_{ij})_{1 \leq i \leq r, 1 \leq j \leq n}) = \frac{\#\mathcal{R}(C_n, r)}{\#\mathcal{M}(C_n, n)} \cdot \#\mathcal{M}((C_n)_{n-r}, (C_n - \sum_{i=1}^{r} X_{ij})_{j=1,...,n}) \leq \frac{\#\mathcal{R}(C_n, r)}{\#\mathcal{M}(C_n, n)} \cdot \#\mathcal{M}((C_n)_{n-r}, (C(n - r))_{n}).
\]

The last inequality above follows from Corollary 3.3. By Theorem 3.5,

\[
\#\mathcal{M}((C_n)_{n-r}, (C(n - r))_{n}) = \frac{(n-r+C(n-r)-1)_{n} (n+Cn-1)_{n-r}}{(n-r+Cn(n-r)-1)_{n-r}} \exp\left(\frac{1}{2} + o(1)\right)
\]

and

\[
\#\mathcal{M}(C_n, n) = \frac{(C_{n+1})_{n-1} (C_{n+1})_{n+1}^{n}}{(n^2+Cn^2-1)_{n-1}} \exp\left(\frac{1}{2} + o(1)\right).
\]

Trivially,

\[
\#\mathcal{R}(C_n, r) = \left(\frac{C_{n} + n - 1}{n-1}\right)^{r} = \exp\left(\frac{\log(1+C)}{1+C} + \frac{C}{1+C} \log \frac{1+C}{C} + o(1)\right) n^r.
\]

By plugging them into (3.5) and Lemma 3.4, we have

\[
\frac{\#\mathcal{R}(C_n, r)}{\#\mathcal{M}(C_n, n)} \cdot \#\mathcal{M}((C(n - r))_{n}, (C(n - r))_{n-r}) = \frac{(n-r+C(n-r)-1)_{n} (n+Cn-1)_{n-r}}{(n-r+Cn(n-r)-1)_{n-r}} \cdot \frac{(n^2+Cn^2-1)_{n-1}}{(n-1)_{n-1} (C_{n+1})_{n+1}^{n}} (1 + o(1))
\]

\[
= \exp\left[ -\frac{r}{2} \log\left(2\pi \cdot \frac{C}{1+C} (n-r+C(n-r)-1)\right) - \frac{1}{2} \log(2\pi \cdot \frac{C}{1+C} (Cn^2 + n^2 - 1)) \right]
\]

\[
\times (1 + o(1))
\]

\[\square\ Springer\]
This completes the proof. \(\square\)

**Remark 3.7** From the above proof, we can easily see that

\[
(2\pi \cdot \frac{C}{1+C} (n-r+C(n-r)-1))^{-\frac{1}{2}} \leq \left(2\pi \cdot \frac{C}{1+C} (Cn^2+n^2-1)\right)^{-\frac{1}{2}} \times (1+o(1))
\]

\[
= \frac{(Cn(n-r)+n(n-r)-1)^{\frac{1}{2}}}{(Cn^2+n^2-1)^{\frac{1}{2}}} \frac{(Cn+1-n)^{\frac{1}{2}}}{(C(n-r)+n-r-1)^{\frac{1}{2}}}(1+o(1))
\]

\[
\leq \left(1 + \frac{Cr + r}{Cn + n - Cr - r - 1}\right)^{\frac{n}{2}} (1+o(1))
\]

\[
\leq e^{\frac{\xi}{2}} (1+o(1)).
\]

This completes the proof.

**Theorem 3.8** (Same as Theorem 1.4) Let \((i_1, j_1), \ldots, (i_L, j_L)\) be a fixed sequence of indices and \(\alpha_1, \ldots, \alpha_L\) be \(L\) fixed positive integers. Let \(X = (X_{ij})_{1 \leq i, j \leq n}\) be uniformly distributed on \(\mathcal{M}(Cn, n)\), then

\[
\mathbb{E}\left[ \prod_{k=1}^{L} X_{i_k, j_k}^{\alpha_k} \right] \to \mathbb{E}\left[ \prod_{k=1}^{L} Y_{k}^{\alpha_k} \right],
\]

where \(Y_1, \ldots, Y_k\) are i.i.d. \(\text{Geom}(C)\).

**Proof** By Theorem 1.2, the joint distribution of \((X_{i_k, j_k})_{1 \leq k \leq L}\) converges to i.i.d. \(\text{Geom}(C)\) variables in total variation distance. Hence,

\[
\mathbb{E}\left[ \prod_{k=1}^{L} X_{i_k, j_k}^{\alpha_k}; \max_{1 \leq k \leq L} X_{i_k, j_k} < M \right] \to \mathbb{E}\left[ \prod_{k=1}^{L} Y_{k}^{\alpha_k}; \max_{1 \leq k \leq L} Y_k < M \right].
\]

Therefore, it suffices to show that

\[
\lim_{M \to \infty} \sup_n \mathbb{E}\left[ \prod_{k=1}^{L} X_{i_k, j_k}^{\alpha_k}; \max_{1 \leq k \leq L} X_{i_k, j_k} \geq M \right] \to 0.
\]

Without loss of generality, by symmetry, assume that \(1 \leq i_k, j_k \leq L\) for all \(1 \leq k \leq L\). Let \(\tilde{Y} = (\tilde{Y}_{ij})\) be uniformly distributed on \(\mathcal{R}(Cn, L)\). Recall that \(\mathcal{R}(Cn, L)\) is the set
of \( L \times n \) non-negative integer matrices with row sums \( Cn \). By Proposition 3.6,
\[
\mathbb{E} \left[ \prod_{k=1}^{L} X_{i_k,j_k}^{\alpha_k} ; \max_{1 \leq k \leq L} X_{i_k,j_k} \geq M \right] \leq e^{L} (1 + o(1)) \mathbb{E} \left[ \prod_{k=1}^{L} \tilde{Y}_{i_k,j_k}^{\alpha_k} ; \max_{1 \leq k \leq L} \tilde{Y}_{i_k,j_k} \geq M \right].
\] (3.7)

For each fixed \( k \), \( \tilde{Y}_{i_k,j_k} \) has the law of \( \mu_1 \) on \( \mathcal{R}(Cn, 1) \). Hence,
\[
\mathbb{P}(\tilde{Y}_{i_k,j_k} = x) = \frac{(Cn-x+n-2)}{(n-1)} \left( \frac{Cn}{n} \cdot \frac{Cn-1}{n-1} \cdot \ldots \cdot \frac{Cn-x+1}{n-x-1} \right).
\]
It is easy to see that
\[
\lim_{n \to \infty} \mathbb{P}(\tilde{Y}_{i_k,j_k} = x) = \left( \frac{1}{1+C} \right) \left( \frac{C}{1+C} \right)^x,
\]
and for any fixed \( \ell < \infty \) and \((i_k, j_k)\),
\[
\mathbb{E} \left[ \tilde{Y}_{i_k,j_k}^{\ell} \right] = \sum_{k=0}^{\infty} \mathbb{P}(\tilde{Y}_{i_k,j_k} = k) k^{\ell} \xrightarrow{n \to \infty} \sum_{k=0}^{\infty} \left( \frac{1}{1+C} \right) \left( \frac{C}{1+C} \right)^k k^{\ell} < \infty.
\]
Therefore,
\[
\sup_n \mathbb{E} \left[ Y_{i_k,j_k}^{\ell} \right] < \infty \quad (3.8)
\]
for any finite \( \ell < \infty \). By (3.8),
\[
\sup_n \mathbb{E} \left[ \prod_{k=1}^{L} \tilde{Y}_{i_k,j_k}^{\alpha_k} ; \max_{1 \leq k \leq L} \tilde{Y}_{i_k,j_k} \geq M \right] \leq \sup_n \mathbb{E} \left[ \prod_{k=1}^{L} \left( \frac{\sum_{k=1}^{L} \alpha_k}{\sum_{k=1}^{L} \alpha_k} \right)^{\sum_{k=1}^{L} \alpha_k} ; \max_{1 \leq k \leq L} \tilde{Y}_{i_k,j_k} \geq M \right]
\]
\[
\leq \sup_n \mathbb{E} \left[ \sum_{k=1}^{L} \frac{\alpha_k}{\sum_{k=1}^{L} \alpha_k} \left( \sum_{k=1}^{L} \alpha_k \right)^{\sum_{k=1}^{L} \alpha_k} ; \max_{1 \leq k \leq L} \tilde{Y}_{i_k,j_k} \geq M \right]
\]
\[
\leq \sup_n \mathbb{E} \left[ \sum_{k=1}^{L} \frac{\alpha_k}{\sum_{k=1}^{L} \alpha_k} \left( \max_{1 \leq k \leq L} \tilde{Y}_{i_k,j_k} \right)^{\sum_{k=1}^{L} \alpha_k + 1} ; \max_{1 \leq k \leq L} \tilde{Y}_{i_k,j_k} \geq M \right]
\]
\[
\leq \frac{1}{M} \sup_n \mathbb{E} \left[ \sum_{k=1}^{L} \frac{\alpha_k}{\sum_{k=1}^{L} \alpha_k} \left( \max_{1 \leq k \leq L} \tilde{Y}_{i_k,j_k} \right)^{\sum_{k=1}^{L} \alpha_k + 1} \right].
\]
\[
\leq \frac{1}{M} \sum_{k=1}^{L} \frac{\alpha_k}{\sum_{k'=1}^{L} \alpha_{k'}} \sup_n \mathbb{E} \left[ \sum_{k'=1}^{L} \sum_{k=1}^{L} \alpha_{k+1} \right] \xrightarrow{M \to \infty} 0
\]

as \( M \to \infty \). In the above estimate, the first inequality follows from the power mean inequality

\[
\prod_{i=1}^{n} x_i^{\omega_i} \leq \sum_{i=1}^{n} \omega_i x_i,
\]

where \( x_i > 0 \) and \( \sum_{i=1}^{n} \omega_i = 1 \). The third inequality follows from the following classical estimate:

\[
\mathbb{E} \left[ |Y|^{\alpha}; Y > M \right] \leq \mathbb{E} \left[ |Y|^{\alpha+1}; |Y| > M \right] \leq \mathbb{E} \left[ \frac{|Y|^{\alpha+1}}{M}; |Y| > M \right].
\]

Together with (3.7), we obtain the desired result. \( \square \)

**Corollary 3.9** Let \( X = (X_{ij})_{1 \leq i, j \leq n} \) be uniformly distributed on \( \mathcal{M}(Cn, n) \), then

\[
\frac{1}{\sqrt{n}} \left( \sum_{j=2}^{n} X_{1j} - C(n-1) \right) \to 0
\]

almost surely.

**Proof** Let \( \bar{X}_{ij} = X_{ij} - C \) and

\[
\bar{S}_{n-1} = \sum_{j=2}^{n} \bar{X}_{1j} = \sum_{j=2}^{n} X_{1j} - C(n-1) = Cn - X_{11} - C(n-1) = C - X_{11}.
\]

For any fixed \( \varepsilon > 0 \),

\[
\mathbb{P} \left( \frac{|\bar{S}_{n-1}|}{\sqrt{n}} > \varepsilon \right) = \mathbb{P} \left( |\bar{S}_{n-1}| > \sqrt{n} \varepsilon \right) \leq \frac{\mathbb{E} \left[ |\bar{S}_{n-1}|^4 \right]}{\varepsilon^4 n^2} = \frac{\mathbb{E} \left[ |C - X_{11}|^4 \right]}{\varepsilon^4 n^2} = O \left( \frac{1}{n^2} \right).
\]

By Borel–Cantelli Lemma,

\[
\mathbb{P} \left( \frac{|\bar{S}_{n-1}|}{\sqrt{n}} > \varepsilon \text{ i.o.} \right) = 0,
\]

and

\[
\frac{1}{\sqrt{n}} \left( \sum_{j=2}^{n} X_{1j} - C(n-1) \right) \to 0
\]

almost surely. \( \square \)
4 Maximum Entry

By the maximum entropy principle, we believe that the uniform distributed $X = (X_{ij})$ behaves in many sense like the matrix $Y$ of i.i.d. $\text{Geom}(C)$ variables. In this section, we show that the maximum entries of $X$ and $Y$ are of the same order. Let’s first take a look at the behaviors of the maximum of i.i.d. $\text{Geom}(C)$. This question has been well studied. See [15] for a more detailed treatment. By the exact same approach there, we have the following Lemma:

Lemma 4.1 Let $Y_i, 1 \leq i \leq n^2$ be sequence of i.i.d. $\text{Geom}(C)$, then

1. For any $\varepsilon > 0$,

$$
\lim_{n \to \infty} \mathbb{P} \left( \max_{1 \leq i \leq n^2} Y_i \leq \frac{1}{\log \left( \frac{C+1}{C} \right)} \log \left( \frac{C}{1+C} \cdot n^{2+\varepsilon} \right) \right) = 1.
$$

2. We have that

$$
\frac{1}{\log \left( \frac{1+C}{1+C} \right)} H_{n^2} - 1 < \mathbb{E} \left[ \max_{1 \leq i \leq n^2} Y_i \right] < \frac{1}{\log \left( \frac{1+C}{1+C} \right)} H_{n^2},
$$

where $H_{n^2} = \sum_{k=1}^{n^2} \frac{1}{k} = \log (n^2) + \gamma + O(1/n^2)$.

Proof First, we prove 1. Since $Y_1 \sim \text{Geom}(C)$, the probability density function $\mathbb{P}(Y_1 = x) = \left( \frac{1}{1+C} \right) \left( \frac{C}{1+C} \right)^x$ and the cumulative density function takes the following form:

$$
\mathbb{P}(Y_1 \leq x) = \sum_{k=0}^{x} \mathbb{P}(Y_1 = x) = \sum_{k=0}^{x} \left( \frac{1}{1+C} \right) \left( \frac{C}{1+C} \right)^x = 1 - \left( \frac{C}{1+C} \right)^{x+1}.
$$

Hence,

$$
\mathbb{P} \left( \max_{1 \leq i \leq n^2} Y_i \leq x \right) = \left[ 1 - \left( \frac{C}{1+C} \right)^{x+1} \right]^{n^2},
$$

and it is easy to verify that

$$
\lim_{n \to \infty} \mathbb{P} \left( \max_{1 \leq i \leq n^2} Y_i \leq \frac{1}{\log \left( \frac{C+1}{C} \right)} \log \left( \frac{C}{1+C} \cdot n^{2+\varepsilon} \right) \right) = 1.
$$

Next, we prove 2. To start, we have

$$
\mathbb{E} \left[ \max_{1 \leq i \leq n^2} Y_i \right] = \sum_{x=0}^{\infty} \mathbb{P} \left( \max_{1 \leq i \leq n^2} Y_i > x \right) = \sum_{x=0}^{\infty} 1 - \left[ 1 - \left( \frac{C}{1+C} \right)^{x+1} \right]^{n^2}.
$$
Notice that
\[
\sum_{x=0}^{\infty} 1 - \left[ 1 - \left( \frac{C}{1 + C} \right)^{x+1} \right] n^2 = \sum_{x=0}^{\infty} 1 - \left[ 1 - \left( \frac{C}{1 + C} \right)^{x} \right] n^2 - 1.
\]
Let \( e^{-\lambda} = \frac{C}{1 + C} \), i.e., \( \lambda = \log \left( \frac{1+C}{C} \right) \), then
\[
\int_0^\infty 1 - (1 - e^{-\lambda x}) n^2 \, dx < \sum_{x=0}^{\infty} 1 - \left[ 1 - \left( \frac{C}{1 + C} \right)^{x} \right] n^2 \leq 1 + \int_0^\infty 1 - (1 - e^{-\lambda x}) n^2 \, dx.
\]
Equivalently,
\[
\frac{1}{\lambda} H_{n^2} < \sum_{x=0}^{\infty} 1 - \left[ 1 - \left( \frac{C}{1 + C} \right)^{x} \right] n^2 < 1 + \frac{1}{\lambda} H_{n^2}.
\]
This completes the proof of 2. \( \square \)

It turns out that we have a similar result for \( X \).

**Theorem 4.2** (Same as Theorem 1.5) Fix any fixed \( \epsilon > 0 \),
\[
\lim_{n \to \infty} \mathbb{P} \left( \max_{1 \leq i, j \leq n} X_{ij} > \frac{1}{\log \left( \frac{C+1}{C} \right)} \log \left( \frac{C}{1 + C} \cdot n^{2+\epsilon} \right) \right) = 0.
\]

**Proof** By Proposition 3.6,
\[
\mathbb{P} \left( X_{11} > \frac{1}{\log \left( \frac{C+1}{C} \right)} \log \left( \frac{C}{1 + C} \cdot n^{2+\epsilon} \right) \right) \leq (e^{1/2} + o(1)) \mathbb{P} \left( \tilde{Y}_{11} > \frac{1}{\log \left( \frac{C+1}{C} \right)} \log \left( \frac{C}{1 + C} \cdot n^{2+\epsilon} \right) \right). \tag{4.1}
\]
Recall that \( \tilde{Y}_{11} \) is uniformly distributed on \( \mathcal{R}(Cn, 1) \) and has the following probability density function:
\[
\mathbb{P} (\tilde{Y}_{11} = x) = \frac{Cn-x+n-2}{(Cn+n-1)} \cdot \frac{Cn-1}{Cn+n-2} \cdot \frac{Cn-1}{Cn+n-3} \cdots \frac{Cn-x+1}{Cn+n-x-1}.
\]
Notice that

\[
\mathbb{P}\left(\tilde{Y}_{11} > \frac{1}{\log\left(\frac{C+1}{C}\right)} \log\left(\frac{C}{1+C} \cdot n^{2+\varepsilon}\right)\right) = \sum_{x=\left\lfloor \frac{1}{\log\left(\frac{C+1}{C}\right)} \log\left(\frac{C}{1+C} \cdot n^{2+\varepsilon}\right) \right\rfloor + 1}^{Cn} \mathbb{P}(\tilde{Y}_{11} = x) = \left(\sum_{\frac{\log(n^{2+\varepsilon})}{\log\left(\frac{C+1}{C}\right)} \leq x \leq Cn} \mathbb{P}(\tilde{Y}_{11} = x)\right) + o(1).
\]

Therefore,

\[
\mathbb{P}\left(\tilde{Y}_{11} > \frac{1}{\log\left(\frac{C+1}{C}\right)} \log\left(\frac{C}{1+C} \cdot n^{2+\varepsilon}\right)\right) = \sum_{C_n,\varepsilon \leq x \leq Cn} \frac{n-1}{Cn+n-1} \frac{Cn}{Cn+n-2} \frac{Cn-1}{Cn+n-3} \cdots \frac{Cn-x+1}{Cn+n-x-1} + o(1)
\]

\[
= \sum_{C_n,\varepsilon \leq x \leq Cn} \left(\frac{1}{1+C + \frac{C}{n-1}}\right) \left(\frac{C}{1+C - \frac{2}{n}} \frac{C - \frac{1}{n}}{1+C - \frac{3}{n}} \cdots \frac{C - \frac{x-1}{n}}{1+C - \frac{x+1}{n}}\right) + o(1)
\]

\[
\ll \sum_{C_n,\varepsilon \leq x \leq Cn} \left(\frac{1}{1+C}\right)^x \left(\frac{C}{1+C}\right)^{C_n-C_n,\varepsilon} + o(1)
\]

\[
\ll \left(\frac{C}{1+C}\right)^{C_n,\varepsilon} \left[1 - \left(\frac{C}{1+C}\right)^{C_n-C_n,\varepsilon}\right] + o(1)
\]

\[
\ll n^{-2-\varepsilon},
\]

where \(C_{n,\varepsilon} = \frac{\log(n^{2+\varepsilon})}{\log\left(\frac{C+1}{C}\right)}\). Hence, \(\mathbb{P}\left(\tilde{Y}_{11} > \frac{1}{\log\left(\frac{C+1}{C}\right)} \log\left(\frac{C}{1+C} \cdot n^{2+\varepsilon}\right)\right)\) decays to 0 of order \(n^{-2-\varepsilon}\) and by (4.1),

\[
\mathbb{P}\left(X_{11} > \frac{1}{\log\left(\frac{C+1}{C}\right)} \log\left(\frac{C}{1+C} \cdot n^{2+\varepsilon}\right)\right) = O(n^{-2-\varepsilon}).
\]

Finally, by symmetry of \(X_{ij}\) and the union bound, we are done. \(\square\)
5 Empirical Singular Value Distribution

In this section, we study the distribution of singular values of \( \frac{1}{\sqrt{n}}(X - \mathbb{E}[X]) \), where \( X \) is uniformly distributed on \( M(Cn, n) \). To begin with, let \( 0 \leq \sigma_1(n) \leq \sigma_2(n) \leq \cdots \leq \sigma_n(n) \) denote the singular values of the matrix \( \frac{1}{\sqrt{n}}(X - \mathbb{E}[X]) \). These are positive square roots of eigenvalues of \( \frac{1}{n}(X - \mathbb{E}[X])(X - \mathbb{E}[X])^* \). For any \( n \times n \) Hermitian matrix \( H \), let

\[
\mu_n(H) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(H)}
\]

denote the empirical eigenvalue distribution of \( H \), where \( \lambda_i \) denotes the \( i \)th eigenvalue of \( H \). Let \( \tilde{M}(Cn, n) \) be the subset of \( M(Cn, n) \) with maximum entries not exceeding \( \frac{10 \log n}{\log(C + 1)} \), i.e.,

\[
\tilde{M}(Cn, n) = \left\{ (\tilde{M}_{ij}) : \sum_{k=1}^{n} \tilde{M}_{kj} = Cn, \sum_{k=1}^{n} \tilde{M}_{ik} = Cn, \max_{1 \leq i, j \leq n} \tilde{M}_{ij} \leq \frac{10 \log n}{\log(C + 1)} \right\}.
\]

Let \( \tilde{Y} = (\tilde{Y}_{ij})_{1 \leq i, j \leq n} \) be the matrix of i.i.d. \text{Geom}(C) variables conditioned on not exceeding \( \frac{10 \log n}{\log(C + 1)} \), i.e.,

\[
P(\tilde{Y}_{ij} = x) = \begin{cases} 
\frac{Z_n}{Z_n} \left( \frac{1}{1+C} \right) \left( \frac{C}{1+C} \right)^x & 0 \leq x \leq \left[ \frac{10 \log n}{\log(C + 1)} \right], \\
0 & \text{otherwise}
\end{cases}
\]

where

\[
Z_n = \sum_{x=0}^{\left[ \frac{10 \log n}{\log(C + 1)} \right]} \left( \frac{1}{1+C} \right) \left( \frac{C}{1+C} \right)^x.
\]

Lemma 5.1 We have the following:

1. Conditioning on \( \tilde{Y} \in \tilde{M}(Cn, n) \), the \( \tilde{Y} \) is uniformly distributed on \( \tilde{M}(Cn, n) \).
2. There exists some absolute constant \( \gamma'' > 0 \) such that \( \mathbb{P}(\tilde{Y} \in \tilde{M}(Cn, n)) \geq n^{-\gamma'' n} \).

Proof First, we prove 1. Fix any matrix \( D = (D_{ij}) \in \tilde{M}(Cn, n) \),

\[
\mathbb{P}(\tilde{Y} = D) = \prod_{1 \leq i, j \leq n} \mathbb{P}(\tilde{Y}_{ij} = D_{ij})
\]
\[
\prod_{1 \leq i, j \leq n} \frac{1}{Z_n} \left( \frac{1}{1 + C} \right) ^{D_{ij}} = \left( \frac{1}{Z_n} \right) ^{n^2} \left( \frac{1}{1 + C} \right) ^{n^2} \sum_{1 \leq i, j \leq n} D_{ij} = \left( \frac{1}{Z_n} \right) ^{n^2} \left( \frac{1}{1 + C} \right) ^{n^2} \left( C n^2 \right).
\]

Hence, the probability density function of \( \tilde{Y} \) is constant on \( \mathcal{M}(\tilde{C}n, n) \). This proves 1.

Next, we prove 2. Observe that

\[
Z_n = \sum_{x=0}^{10 \log \left( \frac{C+1}{C} \right) \log n} \left( \frac{1}{1 + C} \right) ^{x} \left( \frac{C}{1 + C} \right) ^{x} = \frac{1}{1 + C} \cdot \left( \frac{C}{1 + C} \right) ^{10 \log \left( \frac{C+1}{C} \right) \log n} + 1
\]

\[
= 1 - \left( \frac{C}{1 + C} \right) ^{10 \log \left( \frac{C+1}{C} \right) \log n} + 1
\]

\[
= 1 - \left( \frac{C}{1 + C} \right) ^{10 \log \left( \frac{C+1}{C} \right) \log n} \cdot \left( \frac{C}{1 + C} \right) ^{10 \log \left( \frac{C+1}{C} \right) \log n} + 1 = 1 - \left( \frac{C}{1 + C} \right) ^{10 \log \left( \frac{C+1}{C} \right) \log n}
\]

where \( C' = \left( \frac{C}{1 + C} \right) ^{10 \log \left( \frac{C+1}{C} \right) \log n} + 1 - \frac{10 \log \left( \frac{C+1}{C} \right) \log n}{\log \left( \frac{C+1}{C} \right) \log n} \) and it is easy to see that \( \frac{C}{1 + C} \leq C' \leq 1 \). Next,

\[
\begin{align*}
\mathbb{P}(\tilde{Y} \in \mathcal{M}(\tilde{C}n, n)) &= \# \mathcal{M}(\tilde{C}n, n) \cdot \left( \frac{1}{Z_n} \right) ^{n^2} \left( \frac{1}{1 + C} \right) ^{n^2} \left( \frac{C}{1 + C} \right) ^{Cn^2} \\
&= \# \mathcal{M}(\tilde{C}n, n) \cdot (1 - C' \cdot n^{-10})^{-n^2} \left( \frac{1}{1 + C} \right) ^{n^2} \left( \frac{C}{1 + C} \right) ^{Cn^2} \\
&= (1 + o(1)) \cdot \# \mathcal{M}(\tilde{C}n, n) \cdot \left( \frac{1}{1 + C} \right) ^{n^2} \left( \frac{C}{1 + C} \right) ^{Cn^2}.
\end{align*}
\]
By Theorem 1.5,
\[
\frac{\#\mathcal{M}(Cn,n)}{\#\mathcal{M}(Cn,n)} = P\left( \max_{1 \leq i,j \leq n} X_{ij} \leq \frac{10}{\log \left( \frac{C+1}{C} \right)} \log n \right) = 1 - o(1). \quad (5.2)
\]
Combining (5.1), (5.2), and Theorem 3.5 by Canfield and McKay, we have that
\[
P\left( \tilde{Y} \in \mathcal{M}(Cn,n) \right)
= (1 + o(1)) \cdot \left( \frac{1}{1 + C} \right)^{n^2} \cdot \frac{\left( C-C(1+C)^{1+C} \right)^{n^2}}{(2\pi C(1+C))^{\frac{2n-1}{2} n^{n-1}}} \cdot \exp(O(1))
= (1 + o(1)) \cdot \exp(O(1)) \cdot \frac{1}{(2\pi C(1+C))^{n-\frac{1}{2}n^{n-1}}}
\geq n^{-\gamma'' n},
\]
where \( \gamma'' > 0 \) is some absolute constant. \( \square \)

Let \( \Upsilon = \frac{1}{n} (\tilde{Y} - \mathbb{E}[\tilde{Y}]) (\tilde{Y} - \mathbb{E}[\tilde{Y}])^* \). By results of Guionnet and Zeitouni [7, Corollary 1.8] on the concentration of measure for empirical spectrum of Wishart Matrix,
\[
P( W_1(\mu_n(\Upsilon), \mathbb{E}[\mu_n(\Upsilon)]) > \varepsilon ) \leq \exp(-C'(\varepsilon) \cdot n^2 \cdot (\log n)^{-2})
\quad (5.3)
\]
for some constant \( C'(\varepsilon) > 1 \) and large \( n \). Here, \( W_1 \) is the Wasserstein distance defined by
\[
W_1(\mu, \nu) := \inf_{\gamma \in \Gamma(\mu, \nu)} \left( \int_{\mathbb{R} \times \mathbb{R}} |x - y|^\gamma(dx, dy) \right),
\]
where \( \Gamma(\mu, \nu) \) is the set of coupling of \( \mu \) and \( \nu \) on \( \mathbb{R} \times \mathbb{R} \). In addition, by the work of Marchenko–Pastur [8],
\[
\mu_n(\Upsilon) \rightarrow \mu'
\]
weakly in probability, where
\[
\mu' = \frac{\sqrt{[4C(1+C)-x]} x}{2\pi C(1+C)} \mathbb{1}_{[0,4C(1+C)]} dx.
\]
Consequently, \( \mathbb{E}[\mu_n(\Upsilon)] \rightarrow \mu' \) weakly. Finally, we are ready to prove the following Theorem 1.6:

**Theorem 5.2** (Same as Theorem 1.6) Let \( X \) be uniformly distributed on \( \mathcal{M}(Cn,n) \) and let \( \tilde{Y} = \frac{1}{\sqrt{n}} (X - \mathbb{E}[X]) \). Let \( 0 \leq \sigma_1(\tilde{Y}) \leq \sigma_2(\tilde{Y}) \leq \ldots \leq \sigma_n(\tilde{Y}) \) be singular
values of $\tilde{\Upsilon}$ and let
\[
\mu_n^s(\tilde{\Upsilon}) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\sigma_i(\tilde{\Upsilon})}
\]
be the empirical singular value distribution of $\tilde{\Upsilon}$. Then
\[
\mu_n^s(\tilde{\Upsilon}) \to \frac{\sqrt{4C(1+C) - y^2}}{\pi C(1+C)} - \frac{1}{\sqrt{C(1+C)}} d\gamma
\]
weakly in probability.

**Proof** By Lemma 5.1,
\[
P(W_1(\mu_n(\Upsilon), E(\mu_n(\Upsilon))) > \epsilon) \geq P(W_1(\mu_n(\Upsilon), E(\mu_n(\Upsilon))) > \epsilon | \tilde{\Upsilon} \in M(Cn, n)) \cdot P(\tilde{\Upsilon} \in M(Cn, n))
\]
\[
= P\left(W_1\left(\mu_n\left(\frac{1}{n} (\tilde{X} - E[\tilde{X}]) (\tilde{X} - E[\tilde{X}])^*\right), E[\mu_n(\Upsilon)]\right) > \epsilon \right) P(\tilde{\Upsilon} \in M(Cn, n)),
\]
where $\tilde{X}$ is uniformly distributed on $M(Cn, n)$. Therefore, by (5.3),
\[
P\left(W_1\left(\mu_n\left(\frac{1}{n} (\tilde{X} - E[\tilde{X}]) (\tilde{X} - E[\tilde{X}])^*\right), E[\mu_n(\Upsilon)]\right) > \epsilon \right) \leq \frac{P(W_1(\mu_n(\Upsilon), E(\mu_n(\Upsilon))) > \epsilon)}{P(\tilde{Y} \in M(Cn, n))}
\]
\[
\leq n^{-n} \cdot \exp\left(-C'(\epsilon) \cdot n \cdot (\log n)^{-2}\right)
\]
\[
= o(1).
\]
Let $E_{C,n}$ be the event that maximum entries of $X$ do not exceed $\frac{10 \log n}{\log(C+1)}$, i.e.,
\[
E_{C,n} = \left\{ X = (X_{ij}) : \max_{1 \leq i, j \leq n} X_{ij} \leq \frac{10 \log n}{\log(C+1)} \right\}.
\]
Since $\tilde{X} \sim (X|E_{C,n})$, we have
\[
P\left(W_1\left(\mu_n\left(\frac{1}{n} (X - E[X]) (X - E[X])^*\right), E[\mu_n(\Upsilon)]\right) > \epsilon \left| E_{C,n}\right.\right) = o(1).
\]
Now, by Theorem 1.5, $P(E_{C,n}) \to 1$ as $n \to \infty$, which implies that
\[
P\left(W_1\left(\mu_n\left(\frac{1}{n} (X - E[X]) (X - E[X])^*\right), E[\mu_n(\Upsilon)]\right) > \epsilon \right) = o(1).
\]
Since $E[\mu_n(\Upsilon)] \to \mu'$, we have

$$
\mu_n\left(\frac{1}{n}(X - E[X])(X - E[X])^*\right) \to \mu' = \frac{\sqrt{[4C(1+C) - x]x}}{2\pi C(1+C)} I_{[0,4C(1+C)]} \, dx
$$

weakly in probability as $n \to \infty$. By simple change of variables $x = y^2$,

$$
\int_a^b \frac{\sqrt{[4C(C+1) - x]x}}{2\pi C(C+1)x} \, dx = \int_{\sqrt{a}}^{\sqrt{b}} \frac{\sqrt{4C(C+1) - y^2}}{\pi C(1+C)} \, dy.
$$

Hence,

$$
\mu_n^x(X) \to \frac{\sqrt{4C(1+C) - y^2}}{\pi C(1+C)} I_{[0,2\sqrt{C(C+1)}]} \, dy
$$

weakly in probability. \hfill \Box

**Acknowledgements** The author would like to thank Robin Pemantle for many helpful discussions.

**Data Availability** Data sharing is not applicable to this article as no datasets were generated or analyzed during the current study.

**Declarations**

**Conflict of interest** The author declares that there is no conflict of interest.

**References**

1. Dittmer, S., Lyu, H., Pak, I.: Phase transition in random contingency tables with non-uniform Margins (2019). arXiv:1903.08743
2. Chatterjee, S., Diaconis, P., Sly, A.: Properties of uniform doubly stochastic matrices (2010). arXiv:1010.6136
3. Good, I.J.: Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables. Annu. Math. Stat. 34(3), 911–934 (1963)
4. Barvinok, A.: What does a random contingency table look like? Comb. Probab. Comput. 19(4), 517–539 (2010)
5. Barvinok, A.: Matrices with prescribed row and column sums. Linear Algebr. Appl. 436, 820–844 (2012)
6. Barvinok, A.: On the number of matrices and a random matrix with prescribed row and column sums and 0-1 entries. Adv. Math. 224, 316–339 (2000)
7. Guionnet, A., Zeitouni, O.: Concentration of the spectral measure for large matrices. Electron. Commun. Probab. 5, 119–136 (2000)
8. Marčenko, V.A., Pastur, L.A.: Distribution of eigenvalues in certain sets of random matrices. Mat. sb. (N.S.) 72, 507–536 (1967)
9. Wu, D.: On properties of random binary contingency tables with non-uniform margin (2022). arxiv:2002.12559
10. Wu, D.: Asymptotic enumeration of binary contingency tables and comparison with independence heuristic (2022). arXiv:2010.12969
11. Lyu, H., Pak, I.: On the number of contingency tables and the independence heuristic (2020). arXiv:2009.10810
12. Good, I.J., Crook, J.F.: The enumeration of arrays and a generalization related to contingency tables. Discret. Math. 19, 23–45 (1977)
13. Terence, T.: Graduate studies in mathematics. In: Topics in Random Matrix Theory, vol. 132. American Mathematical Society, ISBN-10: 0-8218-7430-6 (2012)
14. Rodney Canfield, E., McKay, B.D.: Asymptotic enumeration of integer matrices with large equal row and column sums. Combinatorica 30, 655 (2010)
15. Eisenberg, B.: On the expectation of the maximum of IID geometric random variables. Stat. Probab. Lett 78(2), 135–143 (2008)

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.