ON LIMITING DISTRIBUTION OF A CERTAIN CLASS OF RANDOM BINARY CONTINGENCY TABLES

DA WU

ABSTRACT. Motivated by recent work of Dittmer-Lyu-Pak and an old question posted by Barvinok, we study the limiting distribution of certain class of random binary contingency tables. More precisely, for parameters $n, \delta, B, C$, we consider $X = (x_{ij})$ the random binary contingency tables whose first $[n^\delta]$ rows and columns have margin $[BCn]$ and the rest columns and rows have margin $[Cn]$. We establish the limiting distribution of entries of $X$ under certain restrictions of parameters.

1. INTRODUCTION AND STATEMENT OF THE MAIN RESULT

1.1. Introduction. Let $r = (r_1, \ldots, r_m) \in \mathbb{N}^m$ and $c = (c_1, \ldots, c_n) \in \mathbb{N}^n$ be two positive integer vectors of length $m$ and $n$ with the same sum of entries (we call $r$ and $c$ margins). Namely,

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

Let $M^{(0,1)}(r,c)$ be the set of all $m \times n$ binary contingency tables with row sums $r_i$ and column sums $c_j$, i.e.

$$M^{(0,1)}(r,c) := \left\{ (d_{ij}) \in \{0,1\}^{mn} : \sum_{k=1}^{m} d_{ik} = r_i, \sum_{k=1}^{m} d_{kj} = c_j \text{ for all } 1 \leq i \leq m, 1 \leq j \leq n \right\}$$

For $B, C > 0, 0 \leq \delta \leq 1$. Let $\tilde{r} = \tilde{c} = ([BCn], \ldots, [BCn], [Cn], \ldots, [Cn]) \in \mathbb{N}^{[n^\delta]+n}$ be our margins. Define

$$M^{(0,1)}_{n,\delta}(B, C) := M^{(0,1)}(\tilde{r}, \tilde{c})$$

Let $X = (x_{ij})$ be sampled from $M^{(0,1)}_{n,\delta}(B, C)$ uniformly at random ($X$ is called the Random Binary Contingency Table). Our goal is to study the limiting distribution of each entry of $M^{(0,1)}_{n,\delta}(B, C)$ as $n \to \infty$.

2010 Mathematics Subject Classification. Primary: 60F05. Secondary: 60K35.

Key words and phrases. Random Contingency Table, Binary Contingency Table.
First, we obtain a trivial bound on $B$ and $C$ so that the set $M_{n,\delta}^{(0,1)}(B, C)$ is always non-empty as $n \to \infty$.

**Lemma 1.1.** As $n \to \infty$, we have the following natural bound on parameter $B$ and $C$,

$$\begin{cases} 0 < C \leq 1 \\ 0 < B \leq \frac{1}{C} \end{cases} \text{ if } 0 \leq \delta < 1 \quad \text{and} \quad \begin{cases} 0 < C \leq 2 \\ 0 < B \leq 2 \end{cases} \text{ if } \delta = 1$$

*Proof.* Since every entry of the matrix is restricted to $\{0, 1\}$,

$$\begin{align*}
    BCn &\leq \lfloor n\delta \rfloor + n \\
    Cn &\leq \lfloor n\delta \rfloor + n
\end{align*} \Rightarrow \begin{cases} BC \leq 1 + \frac{\lfloor n\delta \rfloor}{n} \\ C \leq 1 + \frac{\lfloor n\delta \rfloor}{n} \end{cases}$$

Taking the limit and the results follow. \qed

1.2. **Notation.** (1) For two random variables $X_1, X_2$ taking values on $\mathbb{N}$, the Total Variation Distance is defined as

$$\|X_1, X_2\|_{TV} := \sum_{k \geq 0} |\mathbb{P}(X_1 = k) - \mathbb{P}(X_2 = k)|$$

(2) If $\tilde{X} \sim \text{Ber}(q)$, then $\mathbb{P}(\tilde{X} = 0) = 1 - q$, $\mathbb{P}(\tilde{X} = 1) = q$.

1.3. **Main Result.** Our main result is the following.

**Theorem 1.2.** For $M_{n,\delta}^{(0,1)}(B, C)$ with parameter $n, \delta, B, C$. Let $X = (x_{ij})$ be sampled uniformly at random from $M_{n,\delta}^{(0,1)}(B, C)$. Fix $\varepsilon > 0$, we have

(i): When $0 \leq \delta < 1$, $0 < C \leq 1$ and $0 < B \leq \frac{1}{C}$,

$$\|X_{n+1, n+1}, \text{Ber}(C)\|_{TV} = O \left( n^{\delta-1} + n^{-\frac{1}{2}+\varepsilon} \right)$$

(ii): When $\frac{1}{2} < \delta < 1$, $0 < C < \frac{3}{4}$ and $B < \frac{1}{\sqrt{\frac{1}{2} - \frac{C^2}{2} + \varepsilon}}$,

$$\left\|X_{11}, \text{Ber} \left( \frac{B^2(1-C)}{B^2 - 2B + 1/C} \right) \right\|_{TV} = O \left( n^{\delta-1} + n^{\frac{1}{2} - \delta + \varepsilon} \right)$$

(iii): When $0 \leq \delta < 1$, $0 < C < \frac{3}{4}$ and $B < \frac{1}{\sqrt{\frac{1}{2} - \frac{C^2}{2} + \varepsilon}}$,

$$\|X_{1, n+1}, \text{Ber}(BC)\|_{TV} = \|X_{n+1, 1}, \text{Ber}(BC)\|_{TV} = O \left( n^{\delta-1} + n^{-\frac{1}{2}+\varepsilon} \right)$$
2. Analysis of Typical Table

A. Barvinok introduced the notion of Typical Table in order to answer the question What does a random contingency table look like? It turns out that as the dimension of matrix grows, the random contingency table is close in certain sense to the typical table (see, for example, [2], [3], [4], [5] for background and the precise statement) Here we only recall the construction by Barvinok and make several remarks.

Fix margins $r \in \mathbb{N}^m$ and $c \in \mathbb{N}^n$, we first define the binary transportation polytope to be

$$
\mathcal{P}^{[0,1]}(r, c) := \left\{ (x_{ij}) \in [0,1]^{mn} : \sum_{k=1}^{m} x_{ik} = r_i, \sum_{k=1}^{n} x_{kj} = c_j, \forall 1 \leq i \leq m, 1 \leq j \leq n \right\}
$$

**Definition 2.1 (Typical Table).** For all $X = (x_{ij}) \in (0,1)^{mn}$, define the function

$$
g(X) = \sum_{i,j} x_{ij} \ln \frac{1}{x_{ij}} + (1 - x_{ij}) \ln \frac{1}{1 - x_{ij}}
$$

For fixed margin $r$ and $c$, we define the typical table $Z = (z_{ij})$ to be the unique maximizer of $g$ on the interior of $\mathcal{P}^{[0,1]}(r, c)$.

**Remark 2.2.**

1: For fixed $i, j$,

$$
x_{ij} \ln \frac{1}{x_{ij}} + (1 - x_{ij}) \ln \frac{1}{1 - x_{ij}}
$$

is the (Boltzmann-Shannon) entropy of Bernoulli random variable with mean $x_{ij}$.

2: Since $g$ is strictly concave on the interior of $\mathcal{P}^{[0,1]}(r, c)$, $g$ attains the unique maximum in that region. Therefore the above definition is well-defined.

3: For fixed $i, j$,

$$
\frac{\partial}{\partial x_{ij}} g(X) = \ln \left( \frac{1 - x_{ij}}{x_{ij}} \right)
$$

For our typical table $Z = (z_{ij})$, we have the Lagrange multiplier condition (we are maximizing $g$ under the row sum and column sum constraints)

$$
\ln \left( \frac{1 - z_{ij}}{z_{ij}} \right) = \lambda_i + \mu_j, \quad i = 1, \ldots, m \quad j = 1, \ldots, n
$$

In our case of $\mathcal{M}^{[0,1]}_{n, \delta}(B, C)$, by symmetry and Lagrange multiplier, there exists some $\alpha, \beta$ (possibly depend on all the parameters) such
Let $P \implies$ implies that $z \in \mathbb{Z}$.

We also have the margin condition for $Z = (z_{ij})$,

\begin{equation}
\begin{cases}
([n^\delta]/n)z_{11} + z_{1,n+1} = BC \\
([n^\delta]/n)z_{1,n+1} + z_{n+1,n+1} = C
\end{cases}
\end{equation}

From (2.2), we can quickly get the following,

\begin{equation}
\begin{cases}
z_{n+1,n+1} \leq C, \\
z_{1,n+1} \leq BC,
\end{cases}
\end{equation}

and

\begin{equation}
|z_{n+1,n+1} - C| = n^{\delta - 1}z_{1,n+1} \leq BCn^{\delta - 1},
\end{equation}

\begin{equation}
\lim_{n \to \infty} z_{n+1,n+1} = C
\end{equation}

**Proposition 2.3.** When $0 \leq \delta < 1$ and $0 < C < 1$,

\[\limsup_{n \to \infty} z_{11} \leq \frac{B^2(1 - C)}{B^2 - 2B + 1/C} < \infty\]

**Proof.** Firstly,

\[z_{11}z_{n+1,n+1} = \frac{1}{(Q^2 + 1)(P^2 + 1)} = \frac{1}{P^2Q^2 + P^2 + Q^2 + 1}\]

\[z_{1,n+1}z_{n+1,1} = \frac{1}{(PQ + 1)^2} = \frac{1}{P^2Q^2 + 2PQ + 1}\]

implies that $z_{11}z_{n+1,n+1} \leq z_{1,n+1}z_{n+1,1}$. Next, we claim that $\frac{z_{1,n+1}}{z_{n+1,n+1}} \geq B$. Assume otherwise, i.e. $\frac{z_{1,n+1}}{z_{n+1,n+1}} < B$. Then

\[\frac{z_{11}}{z_{1,n+1}} \leq \frac{z_{n+1,1}}{z_{n+1,n+1}} = \frac{z_{1,n+1}}{z_{n+1,n+1}} < B\]

and

\[BC = \frac{[n^\delta]z_{11}}{n} + z_{1,n+1} < \frac{[n^\delta]Bz_{1,n+1}}{n} + Bz_{n+1,n+1} = BC\]

which is a contradiction. Now, notice that

\[\frac{(z_{n+1,n+1} - 1)z_{1,n+1}}{(z_{1,n+1} - 1)z_{n+1,n+1}} = \frac{z_{n+1,n+1}z_{1,n+1} - z_{1,n+1}}{z_{1,n+1}z_{n+1,n+1} - z_{n+1,n+1}} = \frac{Q^2}{PQ} = \frac{Q}{P}\]
To find the upper bound for $Q/P$, we solve the following optimization problem,

$$\text{maximize } \frac{(z_{n+1,n+1} - 1)z_{1,n+1}}{(z_{1,n+1} - 1)z_{n+1,n+1}}$$

subject to $z_{1,n+1} \geq Bz_{n+1,n+1}$, $\lim_{n \to \infty} z_{n+1,n+1} = C$

It is easy to see that the objective function is non-decreasing in $z_{n+1,n+1}$ and non-increasing in $z_{1,n+1}$. Hence,

$$\limsup_{n \to \infty} \frac{Q}{P} \leq \frac{BC - B}{BC - 1}$$

Since $z_{1,n+1} \leq BC \implies PQ \geq \frac{1}{BC} - 1 = \frac{1 - BC}{BC}$,

$$\liminf_{n \to \infty} P^2 = \liminf_{n \to \infty} \frac{PQ}{Q/P} \geq \frac{(1 - BC)/BC}{(BC - B)/(BC - 1)} = \frac{(BC - 1)^2}{B^2C(1 - C)}$$

This implies that

$$\limsup_{n \to \infty} z_{11} = \limsup_{n \to \infty} \frac{1}{P^2 + 1} \leq \frac{1}{(BC - 1)^2/B^2C(1 - C) + 1}$$

$$= \frac{B^2(1 - C)}{B^2 - 2B + 1/C}$$

$$= \frac{B^2(1 - C)}{(B - 1)^2 + (1/C - 1)} < \infty$$

□

**Lemma 2.4.** Let $Z = (z_{ij})$ be the typical table for $M_{n,\delta}^{(0,1)}(B, C)$ with $0 \leq \delta < 1$, $0 < C < \frac{3}{4}$ and $B < \frac{1}{\sqrt{\frac{C}{3} - \frac{C^2}{3} + \frac{1}{C}}}$

then we have

$$z_{11} = \frac{B^2(1 - C)}{B^2 - 2B + 1/C} + O(n^{\delta-1})$$

$$z_{1,n+1} = z_{n+1,1} = BC + O(n^{\delta-1})$$

**Proof.** Firstly, since $z_{11}$ is uniformly bounded in $n$,

$$(2.4) \quad |z_{1,n+1} - BC| \leq n^{\delta-1}z_{11} = O(n^{\delta-1})$$

This implies $\lim_{n \to \infty} z_{1,n+1} = BC$. Let $P = P(n), Q = Q(n)$ be as in (2.1), then

$$\lim_{n \to \infty} z_{1,n+1} = \lim_{n \to \infty} \frac{1}{PQ + 1} = BC$$

and

$$\lim_{n \to \infty} z_{n+1,n+1} = \lim_{n \to \infty} \frac{1}{Q^2 + 1} = C$$
which is equivalent to
\[ Q \to q^* := \sqrt{\frac{1}{C} - 1}, \quad PQ \to \frac{1}{BC} - 1 \]
Consequently,
\[ P \to p^* := \left( \frac{1}{BC} - 1 \right) / \sqrt{\frac{1}{C} - 1} \]
and
\[ z_{11} = \frac{1}{p^* + 1} - \frac{1}{(p^*)^2 + 1} = \frac{B^2(1 - C)}{B^2 - 2B + 1/C} \]
which is the correct limit. Now, we want to obtain the convergence rate for \( z_{11} \).
Let \( h(x) = \frac{1}{x^2 + 1} \) and \( h'(x) = \frac{-2x}{(x^2 + 1)^2} \). Since \( |h'(x)| \) is decreasing on \( (\sqrt{3}/3, \infty) \), when
\[ B < \frac{1}{\sqrt{\frac{C}{3} - \frac{C^2}{3} + \frac{1}{C}}} \]
we have \( p^* > \sqrt{3}/3 \). By Mean Value Theorem, for all \( p \) such that \( \sqrt{3}/3 < p < p^* \),
\[ |h(P) - h(p^*)| = |h(P(n)) - h(p^*)| \leq |h'(p)||P - p^*| \]
for sufficiently large \( n \). Next,
\[
|P - p^*| \leq \left| P - \frac{1/BC - 1}{Q} \right| + \left( \frac{1}{BC} - 1 \right) \left| \frac{1}{Q} - \frac{1}{q^*} \right|
\]
When \( C < 3/4, \ q^* > \sqrt{3}/3 \) and since \( z_{n+1,n+1} = h(Q), C = h(q^*) \), the Mean Value Theorem gives us
\[
BCn^{\delta-1} \geq |z_{n+1,n+1} - C| = |h(Q) - h(q^*)| \geq |h'(2q^*)| \cdot |Q - q^*|
\]
for sufficiently large \( n \). Hence, \( |Q - q^*| = O(n^{\delta-1}) \). Since \( Q \to q^* \), the second term in (2.5) is of order \( O(n^{\delta-1}) \). For the first term in (2.5),
\[
\left| P - \frac{1/BC - 1}{Q} \right| = \frac{(PQ + 1)/BC}{Q} \cdot \left| \frac{1}{PQ + 1} - BC \right|
\]
\[ = \frac{(PQ + 1)/BC}{Q} \cdot |z_{1,n+1} - BC| = O(n^{\delta-1}) \]
since both \( P \) and \( Q \) converge as \( n \to \infty \) and (2.4). Thus \( |P - p^*| = O(n^{\delta-1}) \). This completes the proof. \( \square \)
3. Bound on Total Variation Distance

First, we recall an important progress made by A. Barvinok.

**Theorem 3.1.** ([4]) Fix row margins \( r = (r_1, \ldots, r_m) \) and column margins \( c = (c_1, \ldots, c_n) \). Let \( Z = (z_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \) be the typical table for \( M^{[0,1]}(r, c) \). Let \( Y = (y_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \) be an matrix with independent Bernoulli random variables with \( y_{ij} \sim \text{Ber}(z_{ij}) \). Then we have the following conclusion:

(i): There exists an absolute constant \( \gamma \) such that

\[
(3.1) \quad (mn)^{-\gamma(m+n)} e^{g(Z)} \leq |M^{[0,1]}(r, c)| \leq e^{g(Z)}
\]

(ii): Conditioned on being in \( M^{[0,1]}(r, c) \), the matrix \( Y \) is uniform on \( M^{[0,1]}(r, c) \).

In other words, the probability mass function of \( Y \) is constant on the set \( M^{[0,1]}(r, c) \). More precisely, for any \( D \in M^{[0,1]}(r, c) \),

\[
(3.2) \quad P(Y = D) = e^{-g(Z)}
\]

(iii): For constant \( \gamma > 0 \),

\[
(3.3) \quad P \left( Y \in M^{[0,1]}(r, c) \right) = e^{-g(Z)}, \quad |M^{[0,1]}(r, c)| \geq (mn)^{-\gamma(m+n)}
\]

**Remark 3.2.**

(i): The matrix \( Y \) is sometimes called the Maximum Entropy Matrix.

(ii): For measurable set \( A \subseteq [0,1]^{mn} \),

\[
(3.4) \quad P(Y \in A) \geq P \left( Y \in A | Y \in M^{[0,1]}(r, c) \right) \cdot P \left( Y \in M^{[0,1]}(r, c) \right) \geq P(X \in A) \cdot (mn)^{-\gamma(m+n)}
\]

Next, we want to obtain an estimate on the total variation distance between entries of the uniform sampled matrix \( X \) and maximum entropy matrix \( Y \). Here we use the same large deviation type estimate method as in [1] and [6].

**Definition 3.3** (Blocks). Given the set \( M^{[0,1]}(r, c) \) of binary contingency tables, we call the subset of indices \( B \subseteq \{1, \ldots, m\} \times \{1, \ldots, n\} \) a block if

\[
r_i = r_{i'}, c_j = c_{j'} \quad \text{for all} \quad (i, j), (i', j') \in B
\]

**Remark 3.4.** Since we are picking \( X \) from \( M^{[0,1]}(r, c) \) uniformly at random, by symmetry, all the entries in the same block have the same distribution. Also, the entries of the typical matrix within the same block are the same.

**Lemma 3.5.** Let \( M^{[0,1]}(r, c) \) be the set of \( m \times n \) binary contingency tables. \( Z = (z_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \) is the typical table for \( M^{[0,1]}(r, c) \) and \( Y = (y_{ij})_{1 \leq i \leq m, 1 \leq j \leq n} \) is the matrix of independent Bernoulli random variables with \( y_{ij} \sim \text{Ber}(z_{ij}) \). \( X \) is sampled from \( M^{[0,1]}(r, c) \) uniformly at random.
Suppose $B_1, \ldots, B_k$ are $k$ not necessarily distinguished blocks in $\mathcal{M}^{[0,1]}(r, c)$ with $|B_1| \leq |B_2| \leq \cdots \leq |B_k|$, then there exists an absolute constant $\gamma' > 0$ such that for each $k$-tuple indices $I = (I_1, \ldots, I_k) \in B_1 \times \cdots \times B_k$ and $t > 0$,

$$
\left\| \prod_{r=1}^{k} X_{I_r} \prod_{r=1}^{k} Y_{I_r} \right\|_{TV} \leq t + (mn)^{\gamma'(m+n)} \exp \left(- \frac{[|B_1|]}{k} \frac{t^2}{2} \right)
$$

(3.5)

Proof. First, we can choose a subset $U \subseteq B_1 \times \cdots \times B_k$ such that every index $(i, j)$ appears only in one element of $U$ and $|U| \geq [B_1]/k$. We consider two extreme cases. First, let’s say all $B_i$ are the same. In this case, we can divide $B_1$ into at least $[B_1]/k$ boxes of size $k$. Let each sub-block to be an element of $U$ and we are done. Another case is when all of the $B_i$ are disjoint. In this case we have much more freedom of choosing elements. Again, divide the smallest block $B_1$ into at least $[B_1]/k$ blocks of size $k$. This surely can be done for larger blocks. Now, we only need to pick one index from each sub-block of $B_1$ so that together they are an element of $U$. Similar reasoning works for the cases in the middle.

Let $X = (x_{ij})$. For measurable subset $G \subseteq [0, 1]$ and each $I \in U \subseteq B_1 \times \cdots \times B_k$, let

$$
X^I = \prod_{r=1}^{k} X_{I_r} \quad \text{and} \quad S_X = \frac{1}{|U|} \sum_{I \in U} X^I \mathbb{1}_{X^I \in G}
$$

$X^I$s have the same distribution for all $I \in U$ and $Y^I$s have the same distribution for all $I \in U$ by the exchangeability within each blocks. This implies

$$
\mathbb{E}[S_X] = \mathbb{P}(X^I \in G) \quad \text{and} \quad Y^I \text{ independent of } Y^J \text{ for } I, J \in U
$$

By (3.4), we have

$$
\mathbb{P}(Y \in G^{mn}) \geq \mathbb{P}(X \in G^{mn}) \cdot (mn)^{-\gamma(m+n)}
$$

By Azuma-Hoeffding inequality, for every $I \in U$,

$$
\mathbb{P} \left( |S_Y - \mathbb{E}[Y^I \in G]| > t \right) \leq 2 \exp \left(- \frac{|U|t^2}{2} \right) \leq 2 \exp \left(- \frac{[B_1]}{k} \frac{t^2}{2} \right)
$$
ON LIMITING DISTRIBUTION OF A CERTAIN CLASS OF RANDOM BINARY CONTINGENCY TABLES

Now,
\[
\left| P(X^I \in G) - P(Y^I \in G) \right| = \left| E[S_X] - P(Y^I \in G) \right| \\
\leq E \left[ |S_X - P(Y^I \in G)| \right] \\
= E \left[ |S_X - P(Y^I \in G)|; |S_X - P(Y^I \in G)| \leq t \right] + E \left[ |S_X - P(Y^I \in G)|; |S_X - P(Y^I \in G)| > t \right] \\
\leq tP \left( |S_X - P(Y^I \in G)| \right) + 2P(|S_X - P(Y^I \in G)| > t)
\]

Since
\[
P \left( |S_Y - P(Y^I \in G)| > t \right) \geq P \left( |S_Y - P(Y^I \in G)| > t | Y \in M^{(0,1)}(r, c) \right) \\
\geq (mn)^{-\gamma(m+n)}P \left( |S_Y - P(Y^I \in G)| > t \right)
\]
we have
\[
2P \left( |S_X - P(Y^I \in G)| > t \right) \leq 4(mn)^{\gamma(m+n)} \exp \left( - \left[ \frac{[B_1]}{k} \right] \frac{t^2}{2} \right)
\]

Hence,
\[
\left| P(X^I \in G) - P(Y^I \in G) \right| \leq t + 4(mn)^{\gamma(m+n)} \exp \left( - \left[ \frac{[B_1]}{k} \right] \frac{t^2}{2} \right) \\
= t + (mn)^{\gamma(m+n)} \exp \left( - \left[ \frac{[B_1]}{k} \right] \frac{t^2}{2} \right)
\]

The set \( G \) is arbitrary, we are done. \( \square \)

Now, we are ready to prove the following key estimate.

**Theorem 3.6.** Let \( X = (X_{ij}) \) be sampled from \( M^{(0,1)}_{n,\delta}(B, C) \) uniformly. Let \( Z = (Z_{ij}) \) be the typical table for \( M^{(0,1)}_{n,\delta}(B, C) \) and let \( Y = (y_{ij}) \) be the matrix of independent Bernoulli random variables with \( y_{ij} \sim Ber(z_{ij}) \). For integers \( k \geq 1 \) and \( 1 \leq i_r, j_r \leq n + \lfloor n^{\delta} \rfloor \) with \( 1 \leq r \leq k \). Then, for every \( \varepsilon > 0 \), we have

\[
\left\| \prod_{r=1}^{\lfloor n^{\delta} \rfloor} X_{i_r, j_r} \prod_{r=1}^{k} Y_{i_r} \right\|_{\text{TV}} \leq \begin{cases} 
\lfloor n^{\delta} \rfloor + \varepsilon & 1 \leq i_r, j_r \leq n + \lfloor n^{\delta} \rfloor \quad \forall r = 1, \ldots, k \\
\lfloor n^{\delta} \rfloor + \varepsilon & 1 \leq i_r, j_r \leq \lfloor n^{\delta} \rfloor \quad \forall r = 1, \ldots, k \\
\lfloor n^{\delta} \rfloor + \varepsilon & 1 \leq i_r \leq \lfloor n^{\delta} \rfloor, 1 + \lfloor n^{\delta} \rfloor \leq j_r \leq n + \lfloor n^{\delta} \rfloor \quad \forall r = 1, \ldots, k \\
\lfloor n^{\delta} \rfloor + \varepsilon & 1 \leq j_r \leq \lfloor n^{\delta} \rfloor, 1 + \lfloor n^{\delta} \rfloor \leq i_r \leq n + \lfloor n^{\delta} \rfloor \quad \forall r = 1, \ldots, k
\end{cases}
\]
Proof. When $1 + |n^\delta| \leq i_r, j_r \leq n + |n^\delta|$ for $r = 1, \ldots, k$, let $t = \frac{1}{2} n^{-\frac{1}{2} + \epsilon}$ in (3.5), and we get
\[
\left\| \prod_{r=1}^{k} X_{i_r} \prod_{r=1}^{k} Y_{j_r} \right\|_{TV} \leq \frac{1}{2} n^{-\frac{1}{2} + \epsilon} + \left( n^2 + 2n^{1+\delta} + n^{2\delta} \right) 2^{\gamma(n+n^\delta)} \exp \left( - \left[ \frac{n^2}{k} \right] \frac{1}{8} n^{2\epsilon-1} \right) \\
\leq \frac{1}{2} n^{-\frac{1}{2} + \epsilon} + (4n^2)^{4\gamma/\epsilon} n \exp \left( - \left[ \frac{n^{2\delta}}{k} \right] \frac{1}{8} n^{2\epsilon-1} \right) \\
\leq \frac{1}{2} n^{-\frac{1}{2} + \epsilon} + C(n \log n) \exp \left( -C'n^{2\epsilon+1} \right) \\
\leq n^{-\frac{1}{2} + \epsilon}
\]

Similarly, when $1 \leq i_r, j_r \leq |n^\delta|$ for $r = 1, \ldots, k$, let $t = \frac{1}{2} n^{-\frac{1}{2} - \frac{1}{2} + \epsilon}$, then
\[
\left\| \prod_{r=1}^{k} X_{i_r} \prod_{r=1}^{k} Y_{j_r} \right\|_{TV} \leq \frac{1}{2} n^{-\frac{1}{2} + \epsilon} + (4n^2)^{4\gamma/\epsilon} n \exp \left( - \left[ \frac{n^{2\delta}}{k} \right] \frac{1}{8} n^{2\epsilon-(2\delta-1)} \right) \\
\leq \frac{1}{2} n^{-\frac{1}{2} + \epsilon} + C(n \log n) \exp \left( -C'n^{2\epsilon+1} \right) \\
\leq n^{-\frac{1}{2} + \epsilon}
\]

For the last two cases, we just let $t = \frac{1}{2} n^{-\frac{1}{2} + \epsilon}$ and the rest is the same. □

Next, we compute the total variation distance between two Bernoulli random variables with different mean.

Lemma 3.7. For $\lambda_1, \lambda_2 > 0$, $\lambda_1 \neq \lambda_2$,
\[
\|\text{Ber}(\lambda_1), \text{Ber}(\lambda_2)\|_{TV} = 2|\lambda_1 - \lambda_2|
\]

Proof. By definition,
\[
\|\text{Ber}(\lambda_1), \text{Ber}(\lambda_2)\|_{TV} = |(1 - \lambda_1) - (1 - \lambda_2)| + |\lambda_1 - \lambda_2| \\
= 2|\lambda_1 - \lambda_2|
\]

□

4. PROOF OF MAIN RESULTS

Proof of Theorem 1.2. First we prove (i). By Lemma 3.7,
\[
\|\text{Ber}(z_{n+1,n+1}), \text{Ber}(C)\|_{TV} = 2|z_{n+1,n+1}, C| = O(n^{\delta-1})
\]
By Theorem 3.6,
\[
\|X_{n+1,n+1}, \text{Ber} (z_{n+1,n+1})\|_{TV} = O \left( n^{-\frac{1}{2} + \epsilon} \right)
\]
Hence,
\[
\|X_{n+1,n+1}, \text{Ber}(C)\|_{TV} \leq \|X_{n+1,n+1}, \text{Ber}(z_{n+1,n+1})\|_{TV} + \|\text{Ber}(z_{n+1,n+1}), \text{Ber}(C)\|_{TV} = O \left( n^{\delta-1} + n^{-\frac{1}{2}+\varepsilon} \right)
\]

The rest of the theorem is proved by the same reasoning.

\[\Box\]

References

[1] S. Dittmer, H. Lyu and I. Pak: Phase Transition in Random Contingency Tables with non-uniform margins To appear in Transactions of the AMS.
[2] A. Barvinok: What does a random contingency table look like? Comb. Probab. Comp. 19 (2010), 517-539.
[3] A. Barvinok: On the number of matrices and a random matrix with prescribed row and column sums and 0-1 entries Advances in Mathematics 224 (2010) 316-339
[4] A. Barvinok: Matrices with prescribed row and column sums, Lin.Alg.Appl.436 (2012), 820-844
[5] A Barvinok and J. A. Hartigan, Maximum Entropy Gaussian approximations for the number of integer points and volumes of polytopes, Adv. Applied. Math. 45 (2010), 252-289
[6] S. Chatterjee, P. Diaconis and A.Sly, Properties of uniform doubly stochastic matrices, preprint (2010), 18pp; arXiv: 1010.6136.

Department of Mathematics, University of Pennsylvania, David Rittenhouse Laboratory, Philadelphia, PA, 19104-6395
E-mail address: dawu@sas.upenn.edu