A NOTE ON ASYMPTOTIC ENUMERATION OF CONTINGENCY TABLES WITH NON-UNIFORM MARGINS

DA WU

Abstract. In this short note, we compute the precise asymptotics for the number of contingency tables with non-uniform margins. More precisely, for parameter $n, \delta, B, C > 0$, we consider the set of matrices whose first $[n^\delta]$ rows and columns have sum $[BCn]$ and the rest $n$ rows and columns have sum $[Cn]$. We compute the precise asymptotics of the cardinality of this set when $B < B_c = 1 + \sqrt{1+1/C}$ using the maximal entropy methods in [1].

1. Introduction

Let $r = (r_1, \ldots, r_m)$ and $c = (c_1, \ldots, c_n)$ be two positive integer vectors such that

$$r_1 + \ldots + r_m = c_1 + \ldots + c_n = N$$

Let $M(r, c)$ be the set of $m \times n$ non-negative matrices with $i$th row sum $r_i$ and $j$th column sum $c_j$. (for all $1 \leq i \leq m$, $1 \leq j \leq n$) Suppose all the $r_i$ and $c_j$ depend on the dimension $m$ and $n$, one of the fundamental problems in Combinatorics is to provide the precise asymptotics of $\#M(r, c)$ as $m, n \to \infty$. Recently, we are interested in the case of non-uniform margin with two different values. More precisely, we consider the case when

$$\bar{r} = \bar{c} = ([BCn], \ldots, [BCn], [Cn], \ldots, [Cn]) \in \mathbb{N}^{[n^\delta]+n}$$

for parameters $B, C > 0$ and $0 \leq \delta < 1$. Let

$$M_{n, \delta}(B, C) := M(\bar{r}, \bar{c})$$

We are interested in the (precise) asymptotics of $\#M_{n, \delta}(B, C)$ when $n \to \infty$. It is shown in [2] that the typical table (defined in [1]) associated with $\bar{r}$ and $\bar{c}$ are uniform bounded in large $n$ limit when $B < B_c = 1 + \sqrt{1+1/C}$. In this case, we apply the maximal entropy method in [1] to compute the precise asymptotics of $M_{n, \delta}(B, C)$. When $B > B_c$, since entries in top left corner will blow up in large $n$ limit, the precise asymptotics is not known. However, loose estimate of $\log \#M_{n, \delta}(B, C)$ is known with error $O(n \log n + n^{2\delta})$, see the main theorem in [3].

2. Precise Asymptotics of $\#M_{n, \delta}(B, C)$ in sub-critical regime

In this section, we compute the precise asymptotic formula for $\#M_{n, \delta}(B, C)$ when $0 \leq \delta < 1$ and $B < B_c = 1 + \sqrt{1+1/C}$ (subcritical case). The computation is based on Theorem 1.3 in [1] and Lemma 5.1 in [2], which will be restated below.

2010 Mathematics Subject Classification. Primary: 05A05.
Key words and phrases. Random Contingency Tables, Maximum Entropy Principle.
2.1. Review of Literature. First, we recall the general asymptotic formula for \( \#M(r, c) \) when the all of the entries of typical table \( Z \) are of same order. More detailed description can be found in [1] Section 1. We say margins \( r = (r_1, \ldots, r_m) \) and \( c = (c_1, \ldots, c_n) \) are \( \delta'-smooth \) if they satisfy the following two conditions:

(i): \( m \geq \delta'n \) and \( n \geq \delta'm \). (dimensions of the matrix are of the same order asymptotically)

(ii): \( \delta' \tau \leq z_{ij} \leq \tau \) for some \( \tau \) such that \( \tau \geq \delta' \) and all \( 1 \leq i \leq m, 1 \leq j \leq n \). (entries of typical table are of the same order asymptotically)

Next, for typical table \( Z = (z_{ij}) \) associated with \( M(r, c) \), we define the quadratic form \( q : \mathbb{R}^{m+n} \to \mathbb{R} \) as the following:

\[
q(s, t) := \frac{1}{2} \sum_{1 \leq j \leq m, 1 \leq k \leq n} (z_{jk}^2 + z_{jk})(s_j + t_k)^2
\]

where \( s = (s_1, \ldots, s_m) \) and \( t = (t_1, \ldots, t_n) \). Notice that the null space is spanned by the vector \( \vec{u} = (1, 1, \ldots, 1, -1, \ldots, -1) \). Let \( H = u^\perp \subseteq \mathbb{R}^{m+n} \) and \( q|_H \) is a positive definite quadratic form and we can define its determinant \( \det (q|_H) \) to be the product of non-zero eigenvalues of \( q \).

We also define the polynomials \( f, h : \mathbb{R}^{m+n} \to \mathbb{R} \) by

\[
f(s, t) := \frac{1}{6} \sum_{1 \leq j \leq m, 1 \leq k \leq n} z_{jk}(z_{jk} + 1)(2z_{jk} + 1)(s_j + t_k)^3
\]

and

\[
h(s, t) := \frac{1}{24} \sum_{1 \leq j \leq m, 1 \leq k \leq n} z_{jk}(z_{jk} + 1)(6z_{jk}^2 + 6z_{jk} + 1)(s_j + t_k)^4
\]

where \( s = (s_1, \ldots, s_m) \) and \( t = (t_1, \ldots, t_n) \). We consider the Gaussian probability measure on \( H \) with density proportional to \( e^{-q} \). Define

\[
\mu := \mathbb{E}[f^2] \quad \text{and} \quad \nu := \mathbb{E}[h]
\]

Now, we can state the main theorem in [1].

**Theorem 2.1** ([1], Theorem 1.3). Fix \( 0 < \delta' < 1 \) and let \( r \) and \( c \) be \( \delta'-smooth \) margins and \( Z = (z_{ij}) \) be the associated typical table for \( M(r, c) \). Then

\[
\#M(r, c) \asymp \frac{e^{g(Z)}}{(4\pi)^{(m+n)/2} \sqrt{\det (q|_H)}} \exp \left( -\frac{\mu}{2} + \nu \right)
\]

as \( m, n \to +\infty \).

**Remark 2.2.** There exits some positive constants \( \gamma_1(\delta') \) and \( \gamma_2(\delta') \) such that

\[
\gamma_1(\delta') \leq \exp \left( -\frac{\mu}{2} + \nu \right) \leq \gamma_2(\delta')
\]

Therefore,

\[
\exp \left( -\frac{\mu}{2} + \nu \right) = O(1)
\]
Remark 2.3. Using the change of coordinate basis,
\begin{equation}
\det(q|_U) = (m + n) \cdot 2^{1-m-n} \det Q
\end{equation}
where $Q = (q_{ij})$ is the $(m + n - 1) \times (m + n - 1)$ symmetric matrix where
\begin{equation}
q_{ij} = r_j + \sum_{k=1}^{n} z_{jk}^2 = \sum_{k=1}^{n} (z_{jk} + z_{jk}^2) \quad \text{for } 1 \leq j \leq m
\end{equation}
\begin{equation}
q_{ij} = r_j + \sum_{k=1}^{n} z_{jk}^2 = \sum_{k=1}^{n} (z_{jk} + z_{jk}^2) \quad \text{for } 1 \leq j \leq m
\end{equation}
\begin{equation}
q_{ij} = r_j + \sum_{k=1}^{n} z_{jk}^2 = \sum_{k=1}^{n} (z_{jk} + z_{jk}^2) \quad \text{for } 1 \leq j \leq m
\end{equation}
Therefore, we can further simplify (2.3) to
\begin{equation}
\#M(r, c) \leq \frac{e^{\beta(Z)}}{(2\pi)^{(m+n-1)/2} \sqrt{\det Q}} \exp \left( -\frac{\mu}{2} + \nu \right)
\end{equation}
See [1] Section 1.4 for a more detailed explanation.

Next, we recall the key Lemma in [2] regarding the asymptotics of entries of $Z = (z_{ij})$ associated with $M_{n, \delta}(B, C)$.

Lemma 2.4 ([2], Lemma 5.1). Fix $0 \leq \delta < 1$ and let $Z = (z_{ij})_{1 \leq i, j \leq n+|n^3|}$ be the typical table of $M_{n, \delta}(B, C)$. Let $B_c = 1 + \sqrt{1 + 1/C}$ and we have the following,
(i): If $B < B_c$, then
\begin{equation}
z_{11} = \frac{B^2(C + 1)}{(B_c - B)(B_c + B - 2)} + O(n^{\delta-1}), \quad z_{1,n+1} = BC + O(n^{\delta-1})
\end{equation}
(ii): If $B > B_c$, then
\begin{equation}
z_{n+1,n+1} = C + O(n^{\delta-1}), \quad z_{1,n+1} = B_c C + O(n^{\delta-1}), \quad n^{\delta-1} z_{11} = C(B - B_c) + O(n^{\delta-1})
\end{equation}

Remark 2.5. The behaviour of $z_{n+1,n+1}$ is more predictable. It is shown in [1] that
\begin{equation}
|z_{n+1,n+1} - C| = n^{\delta-1} z_{1,n+1} \leq BC n^{\delta-1}
\end{equation}
Hence, it is trivial that
\begin{equation}
z_{n+1,n+1} = C + O(n^{\delta-1})
\end{equation}

2.2. Computation of $\#M_{n, \delta}(B, C)$. Now, we go back to our setting of $\#M_{n, \delta}(B, C)$. Recall $0 \leq \delta < 1$ and $B < B_c = 1 + \sqrt{1 + 1/C}$. First, notice that when $B < B_c$, all of entries of $Z = (z_{ij})$ have well-defined finite limits, and by symmetry
\begin{equation}
e^{\beta(Z)} = \prod_{1 \leq i, j \leq n+|n^3|} \frac{(z_{ij} + 1)z_{ij}+1}{z_{ij}}
= \left( \frac{z_{11}+1}{z_{11}} \right)^{[n^3]^2} \left( \frac{(z_{n+1,n+1}+1)z_{n+1,n+1}+1}{z_{n+1,n+1}^2} \right)^{n^2} \left( \frac{(z_{1,n+1}+1)z_{1,n+1}+1}{z_{1,n+1}^2} \right)^{2n[n^3]}
\end{equation}
Next, we compute the determinant of $Q$ in (2.7). By (2.6), $Q$ has entries

\begin{equation}
q_{jj} = [Bcn] + [n^\delta] \left( \frac{B^2(C + 1)}{(B_c - B)(B_c + B - 2)} + O(n^{\delta-1}) \right)^2 + n \left( BC + O(n^{\delta-1}) \right)^2
\end{equation}

when $1 \leq j \leq [n^\delta]$ and $[n^\delta] + n + 1 \leq j \leq 2[n^\delta] + n$

\begin{equation}
q_{jj} = [Cn] + [n^\delta] \left( BC + O(n^{\delta-1}) \right)^2 + n(C + O(n^{\delta-1}))^2
\end{equation}

when $[n^\delta] + 1 \leq j \leq [n^\delta] + n$ and $2[n^\delta] + n + 1 \leq j \leq 2([n^\delta] + n) - 1$.

\begin{equation}
q_{ij} = q_{ji} = \left( \frac{B^2(1 + C)}{(B_c - B)(B_c + B - 2)} + O(n^{\delta-1}) \right)^2 + \frac{B^2(1 + C)}{(B_c - B)(B_c + B - 2)} + O(n^{\delta-1})
\end{equation}

when $1 \leq i \leq [n^\delta]$ and $[n^\delta] + n + 1 \leq j \leq 2[n^\delta] + n$.

\begin{equation}
q_{ij} = q_{ji} = (BC + O(n^{\delta-1}))^2 + BC + O(n^{\delta-1})
\end{equation}

when $1 \leq i \leq [n^\delta]$, $2[n^\delta] + n + 1 \leq j \leq 2([n^\delta] + n) - 1$ and when $[n^\delta] + 1 \leq i \leq [n^\delta] + n$, $[n^\delta] + n + 1 \leq j \leq 2[n^\delta] + n$.

\begin{equation}
q_{ij} = q_{ij} = (C + O(n^{\delta-1}))^2 + C + O(n^{\delta-1})
\end{equation}

when $[n^\delta] + 1 \leq i \leq [n^\delta] + n$ and $2[n^\delta] + n + 1 \leq j \leq 2([n^\delta] + n) - 1$.

The rest of the entries are zero. Notice that all the off-diagonal entries have size $O(1)$ while all the entries on the diagonal has asymptotical order $n$. To compute the asymptotics of det $Q$, we write $Q = A + E$ where $A = \text{diag}(q_{11}, q_{22}, \ldots, q_{2([n^\delta]+n)-1,2([n^\delta]+n)-1})$ is the diagonal matrix. By diagonal expansion of the determinant,

\begin{equation}
\det(Q) = \det(A + E) = \det(A) + S_1 + S_2 + \ldots + S_{2([n^\delta]+n)-1} + \det(E)
\end{equation}

where

\[ S_k = \sum_{1 \leq i_1 < \ldots < i_k \leq 2([n^\delta]+n)-1} \left( \prod_{r=1}^{k} q_{i_r,i_r} \right) \det(E_{i_1,\ldots,i_k}) \]

$E_{i_1,\ldots,i_k}$ is the principle minor of order $2([n^\delta]+n) - 1 - k$ of $E$. Trivially, $\det(E) = 0$, and

\begin{equation}
\det(A) = \prod_{i=1}^{2([n^\delta]+n)-1} q_{ii}
\end{equation}
Furthermore, $S_{2([n^δ]+n)−3} = 0$, and

\[
S_{2([n^δ]+n)−4} = 0
\]

\[
S_{2([n^δ]+n)−5} = (n^δ)^2 n(n-1) \left[ (z_{11}^2 + z_{11})(z_{n+1,n+1}^2 + z_{n+1,n+1}) - (z_{1,n+1}^2 + z_{1,n+1})^2 \right]
\]

The above computation is based on the symmetry of typical table, i.e.

\[ z_{ij} = z_{i'j'} \quad \text{if} \quad r_i = r_{i'} \quad \text{and} \quad c_j = c_{j'} \]

Therefore,

\[
(2.20) \quad \det(Q) = \left( \prod_{i=1}^{2([n^δ]+n)-1} q_i \right) - (q_1 \cdots q_{[n^δ]+n-1}) (q_{[n^δ]+n+2} \cdots q_{2[n^δ]+2n-1}) \left[ (z_{1,n+1}^2 + z_{1,n+1})^2 \right]
\]

\[
+ (n^δ)^2 n(n-1) \left[ (z_{11}^2 + z_{11})(z_{n+1,n+1}^2 + z_{n+1,n+1}) - (z_{1,n+1}^2 + z_{1,n+1})^2 \right]
\]

where we write $q_i$ in place of $q_{ii}$. Finally, by (2.13), (2.14), (2.15), (2.16), (2.17), (2.7) and Lemma 2.4 we get the precise asymptotics of $\#M_{n,δ}(B,C)$.

3. LEFT HALF OF THE $M_{n,δ}(B,C)$

In this section, we compute the case when

\[
\mathbf{r}_1 = \left( \frac{Bc_n, \ldots, Bc_n}{[n^δ] \text{ entries}} \right) \left( \frac{Cn - Bc_n^δ, \ldots, Cn - Bc_n^δ}{n \text{ entries}} \right) \in \mathbb{Z}_{>0}^{n^δ}
\]

and

\[
\mathbf{c}_1 = (Cn, \ldots, Cn) \in \mathbb{Z}_{>0}^n
\]

By symmetry and margin conditions, $Z = (z_{ij})$ satisfies

\[
\begin{cases}
nz_{11} = Bc_n \\
n^δz_{11} + nz_{1,n+1} = Cn
\end{cases}
\]

which implies that

\[
\begin{cases}
z_{11} = BcC \\
z_{1,n+1} = C - z_{11}n^{δ-1} = C - BcCn^{δ-1}
\end{cases}
\]

Next, we compute the exact asymptotic formula of $\#M(\mathbf{r}_1, \mathbf{c}_1)$. Recall the formula,

\[
\frac{e^{g(Z)}}{(2π)^{(m+n-1)/2} \sqrt{\det Q}} \exp \left( -\frac{\nu}{2} + \nu \right)
\]
where $Q = (q_{ij}) \in \mathbb{R}_{\geq 0}^{(2n + n^\delta - 1) \times (2n + n^\delta - 1)}$ has entries

$$q_{ii} = \begin{cases} B_i C + n (B_i C)^2 & 1 \leq i \leq [n^\delta] \\ Cn - B_i C n^\delta + n (C - B_i C n^\delta - 1)^2 & [n^\delta] + 1 \leq i \leq [n^\delta] + n \\ Cn + n^\delta (B_i C)^2 + (n - n^\delta) (C - B_i C n^\delta - 1)^2 & [n^\delta] + n + 1 \leq i \leq 2n + [n^\delta] - 1 \end{cases}$$

and

$$q_{ij} = q_{ji} = \begin{cases} B_i C + (B_i C)^2 & 1 \leq i \leq [n^\delta], n + [n^\delta] + 1 \leq j \leq 2n + [n^\delta] - 1 \\ C - B_i C n^\delta - 1 + (C - B_i C n^\delta - 1)^2 & [n^\delta] + 1 \leq i \leq [n^\delta] + n, n + [n^\delta] + 1 \leq j \leq 2n + [n^\delta] - 1 \end{cases}$$

The rest of the entries are 0. We write $Q = A + E$ where

$$A = \text{diag} (q_{11}, \ldots, q_{2n + [n^\delta] - 1, 2n + [n^\delta] - 1})$$

By diagonal expansion of the determinants,

$$\det(Q) = \det(A + E) = \det(A) + S_1 + S_2 + \ldots + S_{2n + [n^\delta] - 2} + \det(E)$$

where

$$S_k = \sum_{1 \leq i_1 < \ldots < i_k \leq 2n + [n^\delta] - 2} \left( \prod_{r=1}^{k} q_{i_r, i_r} \right) \det(E_{i_1, \ldots, i_k})$$

$E_{i_1, \ldots, i_k}$ is the principle minor of order $2n + [n^\delta] - 2 - k$ of $E$. It is not hard to see that

$$S_{2n + [n^\delta] - 2} = 0$$
$$S_{2n + [n^\delta] - 3} = -[n^\delta] (n - 1) \left\{ B_i C + n (B_i C)^2 \right\}^{[n^\delta] - 1} \left\{ Cn - B_i C n^\delta + n (C - B_i C n^\delta - 1)^2 \right\}^n \times \left\{ Cn + n^\delta (B_i C)^2 + (n - n^\delta) (C - B_i C n^\delta - 1)^2 \right\}^{n - 2} (B_i C + (B_i C)^2)$$
$$- n (n - 1) \left\{ B_i C + n (B_i C)^2 \right\}^{[n^\delta]} \left\{ Cn - B_i C n^\delta + n (C - B_i C n^\delta - 1)^2 \right\}^{n - 1} \times \left\{ Cn + n^\delta (B_i C)^2 + (n - n^\delta) (C - B_i C n^\delta - 1)^2 \right\}^{n - 2} (C - B_i C n^\delta - 1 + (C - B_i C n^\delta - 1)^2)$$
$$S_{2n + [n^\delta] - 1} = 0$$
for all \( i \geq 4 \). Therefore,
\[
\det Q = \det A + S_{2n+[n^\delta]-3} = (q_{11})^{[n^\delta]} (q_{n+1,n+1})^n (q_{2n+1,2n+1})^{n-1} - [n^\delta](n-1)(q_{11})^{[n^\delta]-1} (q_{n+1,n+1})^n (q_{2n+1,2n+1})^{n-2} (B_c C + (B_c C)^2)^2 \\
- n(n-1)(q_{11})^{[n^\delta]} (q_{n+1,n+1})^{n-1} (q_{2n+1,2n+1})^{n-2} \left( C - B_c C n^{\delta-1} + (C - B_c C n^{\delta-1})^2 \right)^2 \\
= \left\{ B_c C n + n (B_c C)^2 \right\}^{[n^\delta]} \left\{ C n - B_c C n^{\delta} + n \left( C - B_c C n^{\delta-1} \right)^2 \right\}^{n-1} \times \\
\left\{ C n + n^\delta (B_c C)^2 + (n - n^\delta) \left( C - B_c C n^{\delta-1} \right)^2 \right\}^{n-1} - \left\{ B_c C n + n (B_c C)^2 \right\}^{[n^\delta]-1} \left\{ C n - B_c C n^{\delta} + n \left( C - B_c C n^{\delta-1} \right)^2 \right\}^{n-2} \times \\
\left\{ C n + n^\delta (B_c C)^2 + (n - n^\delta) \left( C - B_c C n^{\delta-1} \right)^2 \right\}^{n-2} (B_c C + (B_c C)^2)^2 \\
- n(n-1) \left\{ B_c C n + n (B_c C)^2 \right\}^{[n^\delta]} \left\{ C n - B_c C n^{\delta} + n \left( C - B_c C n^{\delta-1} \right)^2 \right\}^{n-1} \times \\
\left\{ C n + n^\delta (B_c C)^2 + (n - n^\delta) \left( C - B_c C n^{\delta-1} \right)^2 \right\}^{n-2} \left( C - B_c C n^{\delta-1} + (C - B_c C n^{\delta-1})^2 \right)^2 \\
\]