ON THE NUMBER OF CONTINGENCY TABLES AND THE INDEPENDENCE HEURISTIC

HANBAEK LYU* AND IGOR PAK*

Abstract. We obtain sharp asymptotic estimates on the number of $n \times n$ contingency tables with two linear margins Cn and BCn. The results imply a second order phase transition on the number of such contingency tables, with a critical value at $B_c := 1 + \sqrt{1 + 1/C}$. As a consequence, for $B > B_c$, we prove that the classical *independence heuristic* leads to a large undercounting.

1. Introduction

Sometimes a conjecture is more than a straightforward claim to be proved or disproved. A conjecture can also represent an invitation to understand a certain phenomenon, a challenge to be confirmed or refuted in every particular instance. Regardless of whether such a conjecture is true or false, the advances toward resolution can often reveal the underlying nature of the objects.

This paper concerns with the *independence heuristic* for approximating the number of contingency tables, introduced by I. J. Good as far back as in 1950. The independence heuristic has been both proved and disproved in several extreme cases. This paper investigates an intermediate case of the margins when the asymptotics are very subtle. Unreachable until now with the existing techniques, the results are quite surprising, providing a new piece of the puzzle.

Let $\boldsymbol{a} = (a_1, \ldots, a_m) \in \mathbb{N}^m$ and $\boldsymbol{b} = (b_1, \ldots, b_n) \in \mathbb{N}^n$, such that $a_1 + \ldots + a_m = b_1 + \ldots + b_n = N$. A *contingency table* with *margins* $(\boldsymbol{a}, \boldsymbol{b})$ is a $m \times n$ matrix $X = \big(x_{ij}\big)$, s.t. $x_{ij} \in \mathbb{N}$, and

$$\sum_{j=1}^{n} x_{ij} \, = \, a_i \, , \quad \sum_{i=1}^{m} x_{ij} \, = \, b_j \quad \text{for all} \quad 1 \le i \le n, \;\; 1 \le j \le m. \tag{1.1}$$

Denote by $\mathcal{T}(\boldsymbol{a}, \boldsymbol{b})$ the set of such tables, and let $\mathrm{T}(\boldsymbol{a}, \boldsymbol{b}) = \big|\mathcal{T}(\boldsymbol{a}, \boldsymbol{b})\big|$.

Computing and approximating $\mathrm{T}(\boldsymbol{a}, \boldsymbol{b})$ is a fundamental and classical problem in Statistics and Combinatorics, with many connections and applications to other fields, see e.g. [DG] (see also [BLP] for recent references). While there are a number of algorithmic approaches and asymptotic results for small margins, the lower and upper bounds for large margins remain far apart, see [BLP]. In fact, there is a dearth of asymptotic tools in the latter case, and very little hope to get a tight asymptotic bound in full generality.

The *independence heuristic* is a classical approximation formula:

$$\mathrm{T}(\boldsymbol{a}, \boldsymbol{b}) \, \approx \, \mathrm{G}(\boldsymbol{a}, \boldsymbol{b}), \tag{1.2}$$

where

$$\mathrm{G}(\boldsymbol{a}, \boldsymbol{b}) \, := \, \binom{N + mn - 1}{mn - 1}^{-1} \prod_{i=1}^{m} \binom{a_i + n - 1}{n - 1} \prod_{j=1}^{n} \binom{b_j + m - 1}{m - 1}. \tag{1.3}$$

The idea behind the independence heuristic is the asymptotic independence of rows and columns of random continency tables $X \in \mathcal{T}(\boldsymbol{a}, \boldsymbol{b})$, see §2.2. We postpone the history of (1.2) and numerical examples until §6.1.

For the uniform margins, the independence heuristic was studied by Canfield and McKay [CM]. In particular, for $m = n$, $a_i = b_i = Bn$,[1] they prove that

1991 *Mathematics Subject Classification.* 05A16, 62H17, 52C07.

Key words and phrases. Contingency tables, asymptotic enumeration, independent heuristic.

August 30, 2021.

*Department of Mathematics, UCLA, Los Angeles, CA, 90095; Email: {hlyu,pak}@math.ucla.edu.

[1] To simplify the presentation, throughout the introduction we drop the floor/ceiling notation, and use Bn, εn, n^{δ}, etc., to mean the nearest integer to these values.

$$\mathrm{T}(\boldsymbol{a},\boldsymbol{b}) \sim \sqrt{e}\cdot \mathrm{G}(\boldsymbol{a},\boldsymbol{b}) \quad \text{as} \quad n\to\infty, \tag{1.4}$$

where $B>0$ is a fixed constant. The same asymptotics (1.4) was proved by Greenhill and McKay [GM] for small margins: $\max\{a_i\}\cdot\max\{b_j\} = o\big(N^{2/3}\big)$.

In the opposite direction, Barvinok proved that the independence heuristic (1.2) fails for nonuniform "cloned margins". In a notable special case, for $m=n$, $\boldsymbol{a}=\boldsymbol{b}=(Bn,\ldots,Bn,n,\ldots,n)$, with εn and $(1-\varepsilon)n$ rows/columns of each sum, $B>1$, he proves:

$$\lim_{n\to\infty}\frac{1}{n^2}\log \mathrm{T}(\boldsymbol{a},\boldsymbol{b}) > \lim_{n\to\infty}\frac{1}{n^2}\log \mathrm{G}(\boldsymbol{a},\boldsymbol{b}).$$

In other words, the independence heuristic greatly undercounts the number of contingency tables for constant fraction of each sum.

In this paper we consider an intermediate case $m=n$, $\boldsymbol{a}=\boldsymbol{b}=(Bn,\ldots,Bn,n,\ldots,n)$, where n^δ rows/columns have larger sums, and fixed $B>1$, $0\le\delta<1$. In this case it is known and easy to see that

$$\lim_{n\to\infty}\frac{1}{n^2}\log \mathrm{T}(\boldsymbol{a},\boldsymbol{b}) = \lim_{n\to\infty}\frac{1}{n^2}\log \mathrm{G}(\boldsymbol{a},\boldsymbol{b}) = 2\log 2.$$

We show that the independence heuristic works fairly well in this case as the second terms of the asymptotics have the same order:

$$\log \mathrm{T}(\boldsymbol{a},\boldsymbol{b}) = (2\log 2)\, n^2 + \Theta\big(n^{1+\delta}\big),$$
$$\log \mathrm{G}(\boldsymbol{a},\boldsymbol{b}) = (2\log 2)\, n^2 + \Theta\big(n^{1+\delta}\big).$$

But the similarities stop when we compute the exact constant implied by the $\Theta(\cdot)$ notation. We present the detailed results in the next section, but here is the qualitative version of (the special case of) the main theorem.

Corollary 1.1. *Fix $0<\delta<1$, $B>1$, and denote $B_c := 1+\sqrt{2}$. Let $m=n$, $\boldsymbol{a}=\boldsymbol{b}=(Bn,\ldots,Bn,n,\ldots,n)$ with n^δ sums Bn. Then:*

$$\lim_{n\to\infty}\frac{1}{n^{1+\delta}}\log\frac{\mathrm{T}(\boldsymbol{a},\boldsymbol{b})}{\mathrm{G}(\boldsymbol{a},\boldsymbol{b})} \quad \begin{cases} = 0 & \text{if } 1<B\le B_c \\ > 0 & \text{if } B>B_c \end{cases}$$

This is quite surprising since the independence heuristic does not "notice" the phase transition at B_c and changes smoothly with B. The corollary then implies a second order phase transition for the number of contingency tables, see the discussion below.

The significance of the critical value $B_c = 1+\sqrt{2}$ for the distribution of random contingency tables has already been predicted in [B3] and proved in [DLP], but until now they never appeared in the context of counting contingency tables.

To summarize the idea of the proof, we combined Barvinok's classical bounds and our previous results on the distribution of entries in contingency tables. We then use *self-reduction* to derive the asymptotics for the number of contingency tables. Put succinctly, the difference in these distributions before and after the phase transition then amplifies the undercounting by the independence heuristic.

2. Main results

2.1. **Barvinok margins.** Fix parameters $0\le\delta\le1$, $B\ge1$ and $C>0$. As in the introduction, define *Barvinok margins*

$$\boldsymbol{a} = \boldsymbol{b} := \big(\lfloor BCn\rfloor,\ldots,\lfloor BCn\rfloor,\lfloor Cn\rfloor,\ldots,\lfloor Cn\rfloor\big) \in \mathbb{N}^{\lfloor n^\delta\rfloor+n}. \tag{2.1}$$

To simplify the notation, for the Barvinok margins we write $\mathcal{T}_{n,\delta}(B,C)$, $\mathrm{T}_{n,\delta}(B,C)$ and $\mathrm{G}_{n,\delta}(B,C)$. Formally, $\mathcal{T}_{n,\delta}(B,C)$ is the set of contingency tables whose first $\lfloor n^\delta\rfloor$ rows and columns have sums

$\lfloor BCn \rfloor$, and the other n rows and columns have sums $\lfloor Cn \rfloor$. Similarly, $\mathrm{T}_{n,\delta}(B,C) = |\mathcal{T}_{n,\delta}(B,C)|$, and $\mathrm{G}_{n,\delta}(B,C)$ is the corresponding independence heuristic approximation (1.3).

The main result of this paper is a sharp asymptotics for the number $\mathrm{T}_{n,\delta}(B,C)$ of contingency tables for Barvinok's margins. The result establishes a phase transition at a critical value $B_c = 1+\sqrt{1+1/C}$, where the second order term in $\log \mathrm{T}_{n,\delta}(B,C)$ grows in B for $B < B_c$, but remains constant for $B > B_c$.

Theorem 2.1 (Main theorem)**.** *Fix $0<\delta<1$, $B,C>0$, and $n \geq 1$. Let $B_c = 1+\sqrt{1+1/C}$ and denote $f(x) := (x+1)\log(x+1) - x\log x$.*
(i) For $B \leq B_c$, we have:

$$\log \mathrm{T}_{n,\delta}(B,C) \, = \, f(C)n^2 \, + \, \left[2f(BC) \, - \, BC\log\left(1+\frac{1}{C}\right)\right] n^{1+\delta} \, + \, Dn^{2\delta} \, + \, O\big(n^{3\delta-1} + n\log n\big),$$

where

$$D \, := \, f(E) + E\log\frac{(1+C)(BC)^2}{C(BC+1)^2} \, - \, \frac{B^2C}{2(C+1)} \quad \text{and} \quad E \, := \, \frac{B^2C(C+1)}{(B_c-B)(B_c+B-2)}\,.$$

(ii) For $B > B_c$, we have:

$$\log \mathrm{T}_{n,\delta}(B,C) \, = \, f(C)n^2 \, + \, \left[2f(B_cC) \, - \, B_cC\log\left(1+\frac{1}{C}\right)\right] n^{1+\delta} \, + \, O\big(n^{2\delta} + n\log n\big).$$

We prove the theorem in the next section. Note that for $\frac{1}{2} < \delta < 1$, we obtain three terms in the asymptotics in the first case, and for $0<\delta<1$, two terms in the second case. It is important to note that the coefficient of the second order term $n^{1+\delta}$ as a function of B attains global maximum at $B = B_c$ and decreases in the interval (B_c, ∞).

In fact, the first two terms of $\log \mathrm{T}_{n,\delta}(B,C)$ agree with those of $\log \mathrm{G}_{n,\delta}(B,C)$, see Lemma 4.1. Hence, the independence heuristic predicts that the number $\mathrm{T}_{n,\delta}(B,C)$ of contingency tables to *decrease* in B when $B > B_c$. However, Theorem 2.1 proves that $\mathrm{T}_{n,\delta}(B,C)$ remains constant for $B > B_c$ (up to the second order) due to the phase transition at $B = B_c$, which was 'invisible' to the independence heuristic. See §6.4 for further discussion.

2.2. **Asymptotic independence.** To understand the main theorem, consider the *correlation ratio* in contingency tables, defined as

$$\rho(\boldsymbol{b},\boldsymbol{a}) \, := \, \frac{\mathrm{T}(\boldsymbol{a},\boldsymbol{b})}{\mathrm{G}(\boldsymbol{a},\boldsymbol{b})}\,.$$

Let us show how $\rho(\boldsymbol{b},\boldsymbol{a})$ can be interpreted as the asymptotic independence of rows and columns. First, recall the following is the argument essentially in [G3] (see also [B1]).

Let $\mathcal{S}_N$ be the set of all $m \times n$ tables with total sum $N = a_1 + \ldots + a_m = b_1 + \ldots + b_n$. Let X be a uniformly chosen contingency table from the set $\mathcal{S}_N$, and consider the following events that X satisfies the row and column margins:

$$\mathcal{R}_n(\mathbf{a}) \, = \, \big\{X \text{ has row margins } \mathbf{a}\big\} \quad \text{and} \quad \mathcal{C}_m(\mathbf{b}) \, = \, \big\{X \text{ has column margins } \mathbf{b}\big\}. \tag{2.2}$$

By definition

$$\mathbb{P}\big(\mathcal{R}_n(\boldsymbol{a}) \cap \mathcal{C}_m(\boldsymbol{b})\big) \, = \, \frac{\mathrm{T}(\boldsymbol{a},\boldsymbol{b})}{|\mathcal{S}_N|}, \qquad \mathbb{P}\big(\mathcal{R}_n(\boldsymbol{a})\big) \, = \, \frac{|\mathcal{R}_n(\boldsymbol{a})|}{|\mathcal{S}_N|}, \qquad \mathbb{P}\big(\mathcal{C}_n(\boldsymbol{b})\big) \, = \, \frac{|\mathcal{C}_n(\boldsymbol{b})|}{|\mathcal{S}_N|}\,.$$

Since

$$|\mathcal{S}_N| \, = \, \binom{N+mn-1}{mn-1}, \qquad |\mathcal{R}_n(\mathbf{a})| \, = \, \prod_{i=1}^{m}\binom{a_i+n-1}{n-1}, \qquad |\mathcal{C}_m(\mathbf{b})| \, = \, \prod_{j=1}^{n}\binom{b_j+m-1}{m-1},$$

using the definition of independence heuristic (1.3) we conclude that

$$\frac{\mathbb{P}\big(\mathcal{R}_n(\mathbf{a}) \cap \mathcal{C}_m(\mathbf{b})\big)}{\mathbb{P}\big(\mathcal{R}_n(\mathbf{a})\big)\,\mathbb{P}\big(\mathcal{C}_m(\mathbf{b})\big)} \;=\; \frac{\mathrm{T}(\boldsymbol{a},\boldsymbol{b})}{\mathrm{G}(\boldsymbol{a},\boldsymbol{b})} \;=\; \rho(\boldsymbol{b},\boldsymbol{a}).$$

In other words, $\log \rho(\mathbf{a},\mathbf{b})$ measures the independence of row and column margins in random contingency tables.

2.3. **Critical correlation exponent.** We now compute the asymptotics of the correlation ratio for the Barvinok margins (2.1). As stated in the introduction, we prove that the ratio also exhibits a second order phase transition in parameter B.

Theorem 2.2. *Fix $0 < \delta < 1$, $B, C > 0$, and $n \geq 1$. Let $B_c = 1 + \sqrt{1 + 1/C}$ and denote $f(x) := (x+1)\log(x+1) - x\log x$. Then:*

(2.3)
$$\lim_{n\to\infty} \frac{1}{n^{1+\delta}} \log \frac{\mathrm{T}_{n,\delta}(B,C)}{\mathrm{G}_{n,\delta}(B,C)} = \begin{cases} 0 & \textit{if} \quad B \leq B_c \\ C(B - B_c)\log\left(1 + \frac{1}{C}\right) - 2\big(f(BC) - f(B_cC)\big) \;>\; 0 & \textit{if} \quad B > B_c. \end{cases}$$

Corollary 1.1 follows immediately from the theorem for $C = 1$. Note that the *critical factor* $n^{1+\delta}$ can also be found in the asymptotics of the total sum N in the $(n + \lfloor n^{\delta} \rfloor) \times (n + \lfloor n^{\delta} \rfloor)$ contingency tables with Barvinok margins:

$$N \;=\; \lfloor Cn \rfloor \cdot n \,+\, 2\lfloor BCn \rfloor \cdot \lfloor n^{\delta} \rfloor \;=\; Cn^2 \,+\, 2BCn^{1+\delta} \,+\, O(n).$$

We call the left hand side of (2.3) the *critical correlation exponent* for the contingency tables $\mathcal{T}(\mathbf{a},\mathbf{b})$. Theorem 2.2 implies that the row and column margin events $\mathcal{R}_n$ and $\mathcal{C}_n$ are asymptotically independent for $B < B_c$ and asymptotically positively correlated for $B > B_c$. Moreover, it is easy to check that the right hand side of (2.3) as well as its first derivative in B is continuous for $B > 0$, but its second derivative is discontinuous at $B = B_c$ (see Figure 1). Hence we are uncovering a second-order phase transition in the correlation structure in contingency tables.

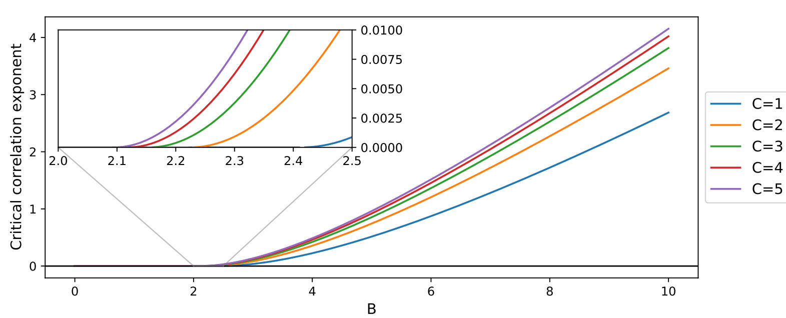

Figure 1. Plot of the critical correlation coefficient in uniform contingency tables with Barvinok's margins with parameters n, δ, B and C. For each $0 < \delta < 1$ and $C > 0$, there exists a second-order phase transition of the critical correlation coefficient in B at critical value $B_c = 1 + \sqrt{1 + 1/C}$. Below B_c the rows and columns are asymptotically independent, but above B_c, they are asymptotically positively correlated.

3. Proof of Theorem 2.1

The proof below relies on the notion of *typical table* introduced in [B3] (Definition 3.1), and a result in [DLP] that was used prove a probabilistic phase transition for the uniformly sampled contingency table for the Barvinok margins (Lemma 3.3).

Let $\mathcal{P}(\boldsymbol{a},\boldsymbol{b}) \subseteq \mathbb{R}_+^{mn}$ be the *transportation polytope* of *real* nonnegative contingency tables with margins $\boldsymbol{a}$ and $\boldsymbol{b}$, i.e. defined by (1.1) over $\mathbb{R}_+$. Clearly, $\mathcal{T}(\boldsymbol{a},\boldsymbol{b}) = \mathcal{P}(\boldsymbol{a},\boldsymbol{b}) \cap \mathbb{Z}^{mn}$. Next, we define the *typical table* introduced by Barvinok [B3].

Definition 3.1 (Typical table)**.** Fix margins $\boldsymbol{a} \in \mathbb{N}^m$ and $\boldsymbol{b} \in \mathbb{N}^n$. Let $\mathcal{P}(\boldsymbol{a},\boldsymbol{b}) \subseteq \mathbb{R}_+^{mn}$ denote the transportation polytope. For each $X = (x_{ij}) \in \mathcal{P}(\boldsymbol{a},\boldsymbol{b})$, define

$$g(X) \, = \sum_{1\le i,j\le n} f(x_{ij}), \tag{3.1}$$

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

$$f(x) \, = \, (x+1)\log(x+1) - x\log x. \tag{3.2}$$

The *typical table* $Z \in \mathcal{P}(\boldsymbol{a},\boldsymbol{b})$ for $\mathcal{T}(\boldsymbol{a},\boldsymbol{b})$ is defined by

$$Z \, = \, \operatorname*{arg\,max}_{X\in\mathcal{P}(\boldsymbol{a},\boldsymbol{b})} g(X).$$

Since the function g defined at (3.1) is strictly concave, it attains a unique maximizer on the transportation polytope $\mathcal{P}(\boldsymbol{a},\boldsymbol{b})$ and thus the typical table is well-defined. The function $f(x)$ in (3.2) equals the Shannon-Bolzmann entropy of Geometric distribution with mean x.

In [B1, Thm 1.1], Barvinok gave the following upper and lower bound on the number of contingency tables in terms of the typical table (cf. 6.2).

Theorem 3.2 ([B1, BH])**.** *Fix margins $\boldsymbol{a} \in \mathbb{N}^m$ and $\boldsymbol{b} \in \mathbb{N}^n$. Let $Z = (z_{ij})$ be the typical table for $\mathcal{T}(\boldsymbol{a},\boldsymbol{b})$. Then there exists some absolute constant $\gamma > 0$, such that*

$$g(Z) \, - \, \gamma(m+n)\log N \, \le \, \log \mathrm{T}(\boldsymbol{a},\boldsymbol{b}) \, \le \, g(Z),$$

where $N = a_1 + \ldots + a_m = b_1 + \ldots + b_n$ is the total sum of the entries.

Let $Z = (z_{ij})$ denote the typical table for $\mathcal{T}_{n,\delta}(B,C)$. By the symmetry of margins (2.1), matrix Z is also symmetric and can be viewed as a 2×2 block matrix with constant entries within each block. To simplify the notation, we use the following entries of Z as representatives of each block:

- z_{11} for the top-left block of size $\lfloor n^\delta \rfloor \times \lfloor n^\delta \rfloor$,
- $z_{1,n+1}$ for both the top-right and bottom-left blocks of size $\lfloor n^\delta \rfloor \times n$, and
- $z_{n+1,n+1}$ for the bottom-right block of size $n \times n$.

A slight modification of the argument in [DLP] shows the following asymptotic expression of the entries of Z:

Lemma 3.3 ([DLP, Lem 5.1])**.** *Let $Z = (z_{ij})$ be the typical table for $\mathcal{T}_{n,\delta}(B,C)$, where $0 \le \delta < 1$. Let $B_c = 1 + \sqrt{1 + 1/C}$. Then there exists a constant $\alpha = \alpha(C)$ independent of B, such that:*

(i) *If $B \le B_c$, then:*

$$\begin{aligned} \left| z_{n+1,n+1} - C \right| \, &\le \, BCn^{\delta-1} \\ \left| z_{1,n+1} - BC \right| \, &\le \, \frac{\alpha}{B_c - B} n^{\delta-1} \\ \left| z_{11} - \frac{B^2C(C+1)}{(B_c-B)(B_c+B-2)} \right| \, &\le \, \frac{\alpha}{B_c - B} n^{\delta-1}. \end{aligned}$$

(ii) *If $B > B_c$, then:*

$$\begin{aligned} \big|z_{n+1,n+1} - C\big| \;&\le\; B_c C n^{\delta-1} \\ \big|z_{1,n+1} - B_c C\big| \;&\le\; \frac{\alpha}{B-B_c}\, n^{\delta-1} \\ \Big|n^{\delta-1} z_{11} - C(B-B_c)\Big| \;&\le\; \frac{\alpha}{B-B_c}\, n^{\delta-1}. \end{aligned}$$

We use the lemma to prove the following result.

Proposition 3.4. *Let $Z = (z_{ij})$ be the typical table for $\mathcal{T}(\boldsymbol{a},\boldsymbol{b})$, where $0 \le \delta < 1$. Let $B_c = 1 + \sqrt{1+1/C}$. Let f, g be the functions defined at (3.2) and (3.1). Then the following hold:*

(i) *If $B < B_c$, then for all $n \ge 1$,*

$$\begin{aligned} g(Z) \;=\; f(C)n^2 \;&+\; \left[2f(BC) \;-\; BC\log\left(1+\frac{1}{C}\right)\right] n^{1+\delta} \\ &+\; \left[f(z_{11}) \;+\; z^*_{11}\log\left(\frac{(1+C)(BC)^2}{C(BC+1)^2}\right) \;-\; \frac{B^2C}{2(C+1)}\right] n^{2\delta} + O(n^{3\delta-1}) + O(n), \end{aligned}$$

where $z^*_{11} = B^2C(C+1)/(B_c-B)(B_c+B-2)$.

(ii) *If $B > B_c$, then for all $n \ge 1$,*

$$g(Z) \;=\; f(C)n^2 \;+\; \left[2f(B_cC) \;-\; B_cC\log\left(1+\frac{1}{C}\right)\right] n^{1+\delta} \;+\; O\big(n^{2\delta}\big) + O(n).$$

Proof. Throughout the proof, the constants B and C are held fixed, and $n \to \infty$. The implied constant in $O(\cdot)$ may depend on B and C but not on n.

First recall that due to the symmetry, the entries z_{11}, $z_{1,n+1}$ and $z_{n+1,n+1}$ of the typical table Z satisfy the following margin condition:

$$\begin{cases} (\lfloor n^\delta \rfloor/n)\, z_{11} + z_{1,n+1} \;=\; (\lfloor BCn \rfloor/n) = BC + O(n^{-1}), \\ (\lfloor n^\delta \rfloor/n)\, z_{1,n+1} + z_{n+1,n+1} \;=\; (\lfloor Cn \rfloor/n) = C + O(n^{-1}). \end{cases}$$

In accordance to Lemma 3.3, define a block table $Z^* = (z^*_{ij})$ by $z^*_{n+1,n+1} = C$, $z^*_{1,n+1} = BC$, and $z^*_{11} = B^2C(C+1)/(B_c-B)(B_c+B-2)$. Combining with Lemma 3.3, for $B < B_c$, we obtain

$$\begin{aligned} C - z_{n+1,n+1} &= z_{1,n+1}(\lfloor n^\delta \rfloor/n) + O(n^{-1}) \\ &= BC(\lfloor n^\delta \rfloor/n) - (BC - z_{1,n+1})(\lfloor n^\delta \rfloor/n) + O(n^{-1}) \\ &= BC(\lfloor n^\delta \rfloor/n) - z^*_{11}(\lfloor n^\delta \rfloor/n)^2 - (z_{11} - z^*_{11})(\lfloor n^\delta \rfloor/n)^2 + O(n^{-1}) \\ &= BC(\lfloor n^\delta \rfloor/n) - z^*_{11}(\lfloor n^\delta \rfloor/n)^2 + O(n^{3\delta-3}) + O(n^{-1}), \end{aligned}$$

and also

$$\begin{aligned} BC - z_{1,n+1} &= z^*_{11}(\lfloor n^\delta \rfloor/n) + (z_{11} - z^*_{11})(\lfloor n^\delta \rfloor/n) + O(n^{-1}) \\ &= z^*_{11}(\lfloor n^\delta \rfloor/n) + O(n^{2\delta-2}) + O(n^{-1}). \end{aligned}$$

Similarly, for $B > B_c$,

$$\begin{aligned} C - z_{n+1,n+1} &= B_cC(\lfloor n^\delta \rfloor/n) + (z_{1,n+1} - B_cC)(\lfloor n^\delta \rfloor/n) + O(n^{-1}) \\ &= B_cC(\lfloor n^\delta \rfloor/n) + \Big[C(B-B_c) - (\lfloor n^\delta \rfloor/n)z_{11}\Big](\lfloor n^\delta \rfloor/n) + O(n^{-1}) \\ &= B_cC(\lfloor n^\delta \rfloor/n) + O(n^{2\delta-2}) + O(n^{-1}). \end{aligned}$$

Now suppose $B \le B_c$. We use the following Taylor expansion of f:

$$f(x) \;=\; f(y) + (x-y)\log(1+x^{-1}) + \frac{(x-y)^2}{2y(y+1)} + O(|x-y|^3),$$

where the constant in $O(\cdot)$ above is bounded when $x, y > 0$ remain bounded. Then observe that

$$\begin{aligned}
f(z_{n+1,n+1}) &= f(C) + (z_{n+1,n+1} - C)\log\left(1+\frac{1}{C}\right) \;+\; \frac{(z_{n+1,n+1}-C)^2}{2C(C+1)} \;+\; O(n^{3\delta-3}) \;+\; O(n^{-1}),\\
&= f(C) - BC(\lfloor n^\delta\rfloor/n)\log\left(1+\frac{1}{C}\right) + z^*_{11}(\lfloor n^\delta\rfloor/n)^2\log\left(1+\frac{1}{C}\right) \;+\; O(n^{3\delta-3}) \;+\; O(n^{-1}),
\end{aligned}$$

$$\begin{aligned}
f(z_{1,n+1}) &= f(BC) + (z_{1,n+1} - BC)\log(1+1/BC) + O(n^{2\delta-2}) + O(n^{-1})\\
&= f(BC) - z^*_{11}(\lfloor n^\delta\rfloor/n)\log(1+1/BC) \;+\; O(n^{2\delta-2}) \;+\; O(n^{-1}),\\
f(z_{11}) &= f(z^*_{11}) + O(n^{\delta-1}) + O(n^{-1}).
\end{aligned}$$

Noting that $g(Z) = n^2 f(z_{n+1,n+1}) + 2n\lfloor n^\delta\rfloor f(z_{1,n+1}) + \lfloor n^\delta\rfloor^2 f(z_{11})$, a straightforward computation shows **(i)**. Next, suppose $B > B_c$. By a similar argument, we have

$$\begin{aligned}
&f(z_{n+1,n+1}) = f(C) - B_c C n^{\delta-1}\log\left(1+\frac{1}{C}\right) + O(n^{2\delta-2}) + O(n^{-1}),\\
&f(z_{1,n+1}) = f(B_cC) + O(n^{\delta-1}), \qquad \text{and} \qquad f(z_{11}) = f(z^*_{11}) + O(n^{\delta-1}) + O(n^{-1}).
\end{aligned}$$

Then **(ii)** follows from here. □

Proof of Theorem 2.1. Suppose $0 < \delta < 1$. Note that $N = Cn^2 + BCn^{1+\delta} + O(n)$, where N denotes the total sum of entries in a contingency table in $\mathcal{T}_{n,\delta}(B,C)$. Combining with Theorem 3.2, we have:

$$\big|\log \mathrm{T}(\mathbf{a},\mathbf{b}) - g(Z)\big| \;\le\; \gamma' n\log n\,,$$

for all $n \ge 1$, for some absolute constant $\gamma' > 0$. Now the theorem follows from Proposition (3.4). □

4. Proof of Theorem 2.2

We start with the following lemma.

Lemma 4.1. *We have:*

$$\begin{aligned}
\log \mathrm{G}_{n,\delta}(B,C) \;=\; f(C)n^2 &+ \left[f(BC) - BC\log\left(1+\frac{1}{C}\right)\right] n^{1+\delta}\\
&+ \left[2\log(BC+1) - \log(C+1) + \frac{(2-4B+B^2)C}{2(1+C)}\right] n^{2\delta} + O(n^{3\delta-1} + n\log n).
\end{aligned}$$

The proof of Lemma 4.1 involves a straightforward computation of expanding the right hand side of (1.3) under Barvinok's margins (2.1). Details are given in the next Section (5).

Proof of Theorem 2.2. By Lemma 4.1,

$$\log \mathrm{G}_{n,\delta}(B,C) = f(C)n^2 + \left[2f(BC) - BC\log\left(1+\frac{1}{C}\right)\right] n^{1+\delta} + O(n^{2\delta} + n\log n). \tag{4.1}$$

Suppose $B < B_c$. Recall that by Theorem 2.1 (i),

$$\log \mathrm{T}_{n,\delta}(B,C) \,=\, f(C)n^2 \,+\, \left[2f(BC) - BC\log\left(1+\frac{1}{C}\right)\right] n^{1+\delta} \,+\, O(n^{2\delta} + n\log n),$$

Hence for $0 \le \delta < 1$, we have $1+\delta > 2\delta$, so we obtain:

$$\lim_{n\to\infty} \frac{1}{n^{1+\delta}} \log \frac{\mathrm{T}_{n,\delta}(B,C)}{\mathrm{G}_{n,\delta}(B,C)} \,=\, 0.$$

On the other hand, suppose $B > B_c$. Then by Theorem 2.1 (ii),

$$\log \mathrm{T}_{n,\delta}(B,C) \,=\, f(C)n^2 + \left[2f(B_cC) - B_cC\log\left(1+\frac{1}{C}\right)\right] n^{1+\delta} + O(n^{2\delta} + n\log n).$$

Hence by (4.1), for $0 < \delta < 1$, we have:

$$\lim_{n\to\infty} \frac{1}{n^{1+\delta}} \log \frac{\mathrm{T}_{n,\delta}(B,C)}{\mathrm{G}_{n,\delta}(B,C)} \,=\, 2f(B_cC) - 2f(BC) + C(B-B_c)\log\left(1+\frac{1}{C}\right).$$

Let us denote the right hand side of the above equation as $\lambda(B)$. Then

$$\frac{\partial \lambda}{\partial B} \,=\, C\log\left(1+\frac{1}{C}\right) \,-\, 2C\log\left(1+\frac{1}{BC}\right),$$

and it is easy to see $\frac{\partial\lambda}{\partial B} > 0$ if and only if $B > B_c$ and the derivative at $B = B_c$ equals zero. Hence for each fixed C, the function λ is a strictly increasing on $[B_c, \infty)$, and has minimum at $B = B_c$. Note also that $\lambda(B_c) = 0$. This shows $\lambda(B) > 0$ for all $B > B_c$. □

We remark that the second derivative in B of the limiting expression in (2.3) for $B > B_c$ is $2C/(B(BC+1)) > 0$ which is also strictly positive at $B = B_c$. Hence the phase transition given in Theorem 2.2 is indeed of second order.

5. Proof of Lemma 4.1

We first compute $\log \mathrm{G}(\mathbf{a},\mathbf{b})$ for the general $m \times n$ tables with total sum N. By Stirling's approximation,

$$\log \binom{a+b}{a} \,=\, (a+b)\log(a+b) - a\log a - b\log b + O(\log(a+b)).$$

Then we have:

$$\begin{aligned} \log \mathrm{G}(\mathbf{a},\mathbf{b}) \,=\, & \sum_{i=1}^{m} (r_i+n)\log(r_i+n) \,+\, \sum_{j=1}^{n} (c_j+m)\log(c_j+m) \,-\, \sum_{i=1}^{m} r_i \log r_i \\ & - \sum_{j=1}^{n} c_j \log c_j \,-\, (N+m\cdot n)\log(N+m\cdot n) + N\log N + O\big((n+m)\log(N+m\cdot n)\big). \end{aligned} \tag{5.1}$$

Now assume Barvinok margins (2.1). Denote $\phi(x) := x\log x$. We use (5.1) with $m \leftarrow n+n^\delta$, $n \leftarrow n+n^\delta$, and $N \leftarrow Cn^2 + BCn^{1+\delta}$. We have:

$$\begin{aligned} \log \mathrm{G}_{n,\delta}(B,C) \,=\, & 2n\left[\phi\left((C+1)n+n^\delta\right) - \phi(Cn)\right] + 2n^\delta\left[\phi\left((BC+1)n+n^\delta\right) - \phi\left(BCn\right)\right] \\ & - \phi\left(N+mn\right) + \phi\left(N\right) + O(n\log n). \end{aligned}$$

Using Taylor expansion, we have:

$$\log\left((C+1)n+n^\delta\right) \,=\, \log n + \log(C+1) \,+\, \frac{n^{\delta-1}}{C+1} \,-\, \frac{n^{2\delta-2}}{2(C+1)^2} \,+\, O(n^{3\delta-3}),$$

$$\log\left(Cn^2+BCn^{1+\delta}\right) \;=\; \log Cn^2 + Bn^{\delta-1} - \frac{B^2n^{2\delta-2}}{2} + O(n^{3\delta-3}),$$
$$\log\left(N+mn\right) \;=\; \log(C+1)n^2 + \frac{BC+2}{C+1}n^{\delta-1} + \frac{(-1+C-2BC-B^2C^2)n^{2\delta-2}}{(1+C)^2} + O(n^{3\delta-3}).$$

Hence we get

$$\begin{aligned}
\log \mathrm{G}_{n,\delta}(B,C) \;=\; & 2n\left((C+1)n+n^{\delta}\right)\left[\log(C+1)n+\frac{n^{\delta-1}}{C+1}-\frac{n^{2\delta-2}}{2(C+1)^2}\right]-2Cn^2\log Cn \\
& + 2n^{\delta}\left((BC+1)n+n^{\delta}\right)\left[\log(BC+1)n+\frac{n^{\delta-1}}{BC+1}-\frac{n^{2\delta-2}}{2(BC+1)^2}\right]-2n^{1+\delta}BC\log BCn \\
& - \left((C+1)n^2+(BC+2)n^{1+\delta}+n^{2\delta}\right)\times \\
& \qquad \times\left[\log(C+1)n^2+\frac{BC+2}{C+1}n^{\delta-1}+\left[-1+C-2BC-B^2C^2\right]\frac{n^{2\delta-2}}{(1+C)^2}\right] \\
& + \left(Cn^2+BCn^{1+\delta}\right)\left[\log Cn^2+Bn^{\delta-1}-\frac{B^2n^{2\delta-2}}{2}\right]+O(n^{3\delta-1}+n\log n) \\
= \; & f(C)n^2+\left[f(BC)-BC\log\left(1+\frac{1}{C}\right)\right]n^{1+\delta} \\
& + \left[2\log(BC+1)-\log(C+1)+\frac{(2-4B+B^2)C}{2(1+C)}\right]n^{2\delta}+O(n^{3\delta-1}+n\log n).
\end{aligned}$$

This completes the proof. $\square$

6. Final remarks

6.1. The story behind the *independence heuristic* (1.2) and (1.3) is rather interesting. This approximation was given implicitly by Good in [G1, p. 100], and later stated formally in [G2, G3]. Good writes that "the conjecture appears to be confirmed" by his calculations [G3, p. 1166], but later admits he is "leaving aside finer points of rigor" (ibid, p. 1184).

Because of the small constant similar to the $\sqrt{e}$ in (1.4), there was an effort to "improve" upon (1.2). Unfortunately, from the asymptotic point of view, many such heuristics behave poorly. Notably, Diaconis and Efron [DE, (3.14)], see also [DG, (7.2)], propose another heuristic estimate. For linear margins $a_i, b_j = \Theta(n)$ and $m = n$, our calculation shows that this formula gives $\log \mathrm{T}(\mathbf{a},\mathbf{b}) \approx \Theta(n^2\log n)$, thus implying the wrong leading term of the asymptotics.

A number of papers tested the independence heuristic numerically, see e.g. [DG, GC]. The results are nothing short of remarkable, showing that (1.2) holds up to a small constant. For example, for the 4×4 case with $N = 592$ introduced in [DE], we have $\mathrm{T}(\boldsymbol{a},\boldsymbol{b}) = 1.226\times 10^{15}$ [DG], while $\mathrm{G}(\boldsymbol{a},\boldsymbol{b}) = 1.211\times 10^{15}$ [B1]. The asymptotic analysis shows that this level of agreement is purely coincidental. In fact, for a much larger 4×4 case with $N = 65159458$ computed in [D1, D2], we have $\mathrm{T}(\boldsymbol{a},\boldsymbol{b}) = 4.3\times 10^{61}$ vs. $\mathrm{G}(\boldsymbol{a},\boldsymbol{b}) = 3.7\times 10^{61}$, so $\mathrm{T}(\boldsymbol{a},\boldsymbol{b})/\mathrm{G}(\boldsymbol{a},\boldsymbol{b}) \approx \sqrt{2.5}$ in this case. This and other numerical estimates are collected in [BLP, §11.4].

6.2. There are several better lower and upper bounds known for the number $\mathrm{T}(\mathbf{a},\mathbf{b})$ of contingency tables, see [BLP] for a recent overview. Notably, the upper bound in [Sha] and the most recent lower bound in [BLP] give improvement over Theorem 3.2 in the lower order terms. For the linear margins these improvement are of the order n^{cn}. Thus, they give an improvement in the second order term in the Main Theorem 2.1 (i) when $\delta < \frac{1}{2}$. It would be interesting to further explore these bounds.

6.3. Consider $m \times n$ contingency tables with general margins $\boldsymbol{a}$ and $\boldsymbol{b}$ and total sum N. Let $W = (w_{ij})$ be as $w_{ij} = r_i c_j / N$ for $1 \le i \le m$, $1 \le j \le n$. Let g be defined in (3.1). In [B1, §2], Barvinok showed that $g(W) - \log \mathrm{G}(\mathbf{a}, \mathbf{b}) \ge 0$. This implies that

$$\log \frac{\mathrm{T}(\boldsymbol{a}, \boldsymbol{b})}{\mathrm{G}(\boldsymbol{a}, \boldsymbol{b})} \ge -\gamma\,(m+n) \log N. \tag{6.1}$$

This shows that the row and column margin events $\mathcal{R}$ and $\mathcal{C}$ defined by (2.2), have asymptotically *nonnegative correlation* at a scale where the right hand side of (6.1) vanishes. For instance, for $m = n$ and linear margins $a_i, b_j = \Theta(n)$, we have $\gamma(m+n)\log N = O(n \log n)$. Then:

$$\limsup_{n\to\infty} \frac{1}{n^{\delta+1}} \log \frac{\mathrm{T}(\boldsymbol{a}, \boldsymbol{b})}{\mathrm{G}(\boldsymbol{a}, \boldsymbol{b})} \ge 0,$$

for every $\delta > 0$.

6.4. To understand the discussion of the Main Theorem 2.1 at the end of §2.1, consider an extreme case of margins $\mathbf{a} = \mathbf{b} = (Bn^2, n, \ldots, n) \in \mathbb{R}^{n+1}$. For every fixed n, when $B > 1$ is large enough, the number $\mathrm{T}(\mathbf{a}, \mathbf{b})$ stabilizes as the corner entry x_{11} is forced to absorb bulk of the total sum $N = (B+1)n^2$. Meanwhile, the independence heuristic approximation $\mathrm{G}(\mathbf{a}, \mathbf{b})$ maximizes at a certain constant B and then decreases exponentially, eventually becoming < 1. In other words, our Main Theorem 2.1 implies a lower order version of the same phenomenon.

Note also a different but related phenomenon of the lower bound in Barvinok's Theorem 3.2, which works well asymptotically for large n, but for small n and large marginals gives lower bonds which are < 1, see [BLP, §10].

6.5. In [B2], Barvinok gave the analogue of Theorem 3.2 for *binary* (0-1) *contingency tables*. Most recently, Wu [Wu] investigated the limiting distribution of the entries in binary contingency tables, and proved the analogue of our Lemma 3.3. It would be interesting to find the analogue of our Main Theorem 2.1 in this setting.

Acknowledgements. We are grateful to Sasha Barvinok, Sam Dittmer, Jonathan Leake and Greta Panova for numerous helpful discussions, and to Robin Pemantle for telling us about paper [Wu]. Both authors were partially supported by the NSF.

References

[B1] A. Barvinok, Asymptotic estimates for the number of contingency tables, integer flows, and volumes of transportation polytopes, *Int. Math. Res. Notices* **2009** (2009), 348–385.

[B2] A. Barvinok, On the number of matrices and a random matrix with prescribed row and column sums and 0-1 entries, *Adv. Math.* **224** (2010), 316–339.

[B3] A. Barvinok, What does a random contingency table look like?, *Comb. Probab. Comp.* **19** (2010), 517–539.

[BH] A. Barvinok and J. A. Hartigan, An asymptotic formula for the number of non-negative integer matrices with prescribed row and column sums, *Trans. AMS* **364** (2012), 4323–4368.

[BBK] A. Békéssy, P. Békéssy and J. Komlós, Asymptotic enumeration of regular matrices, *Studia Sci. Math. Hungar.* **7** (1972), 343–353.

[BLP] P. Brändén, J. Leake and I. Pak, Lower bounds for contingency tables via Lorentzian polynomials, to appear in *Israel J. Math.*; `arXiv:2008.05907`.

[CM] E. R. Canfield and B. D. McKay, Asymptotic enumeration of integer matrices with large equal row and column sums, *Combinatorica* **30** (2010), 655–680.

[D1] J. A. De Loera, Counting and Estimating Lattice Points: Tools from Algebra, Analysis, Convexity, and Probability, *Optima* **81** (2009), 1–9; available at `http://www.mathopt.org/Optima-Issues/optima81.pdf`

[D2] J. A. De Loera, Details on experiments (c1ounting and estimating lattice points), an appendix to [D1], ibid., 17-22; available at `http://www.mathopt.org/Optima-Issues/optima81-app.pdf`

[DE] P. Diaconis and B. Efron, Testing for independence in a two-way table: new interpretations of the chi-square statistic, *Ann. Stat.* **13** (1985), 845–913.

[DG] P. Diaconis and A. Gangolli, Rectangular arrays with fixed margins, *Disc. Prob. Alg.* **72** (1995), 15–41.

[DLP] S. Dittmer, H. Lyu and I. Pak, Phase transition in random contingency tables with non-uniform margins, *Trans. AMS* **373** (2020), 8313–8338.

[G1] I. J. Good, *Probability and the Weighing of Evidence*, Hafner, New York, 1950, 119 pp.

[G2] I. J. Good, On the application of symmetric Dirichlet distributions and their mixtures to contingency tables, *Annals Math. Stat.* **34** (1963), 911–934.

[G3] I. J. Good, Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables, *Annals. Stat.* **4** (1976), 1159–1189.

[GC] I. J. Good and J. F. Crook, The enumeration of arrays and a generalization related to contingency tables, *Discrete Math.* **19** (1977), 23–45.

[GM] C. Greenhill and B. D. McKay, Asymptotic enumeration of sparse nonnegative integer matrices with specified row and column sums, *Adv. Appl. Math.* **41** (2008), 459–481.

[LW] A. Liebenau and N. Wormald, Asymptotic enumeration of digraphs and bipartite graphs by degree sequence, preprint (2020), 25 pp.; `arXiv:2006.15797`.

[Sha] A. Shapiro, Bounds on the number of integer points in a polytope via concentration estimates, preprint (2010), 23 pp.; `arXiv:1011.6252`.

[Wor] N. Wormald, Asymptotic enumeration of graphs with given degree sequence, in *Proc. ICM Rio de Janeiro*, Vol. 3, 2018, 3229–3248.

[Wu] Da Wu, On limiting distribution of a certain class of random binary contingency tables, preprint (2020), 11 pp.; `arXiv:2002.12559`.