arXiv:2304.11423v1 [math.CO] 22 Apr 2023

给定阶数的随机诱导子图边数的二阶矩

Nicola Apollonio*

摘要

对于图 G 和正整数 c，令 $M_c(G)$ 表示由从 c 个顶点中随机抽取的子集所诱导的 G 的子图的边数。 $M_c(G)$ 的二阶矩揭示了 G 的部分结构。我们利用这一事实并结合经典矩不等式，证明图论结果，给出组合恒等式，分别从下界估计 c 最稠密子图的边数和从上界估计 c 最稀疏子图的边数，并为平凡子图的近似计数提供界。

关键词：诱导子图的大小、尾部不等式、平凡子图、最稠密和最稀疏子图、方差不等式。

1 引言

考虑以下三个看似毫不相关的图论结果：

定理 1（[1], [10], [25]）。设 G 是一个有 n 个顶点的图，且 c 是一个满足 $2 \le c \le n-2$ 的正整数。如果 G 的所有 c 阶诱导子图的边数都相同，那么 G 要么是完全图，要么是余完全图。

这里，余完全是指完全图的补图。

定理 2。设 G 是一个阶为 n 、边数为正的图。如果 c 是一个满足$2 \le c \le n-1$的整数，那么 G 有一个阶为 c 的子图，其边数至少为

$$\frac{c-2}{n-2}\left\{\frac{n-c}{n-3}\left(\sum_{v\in V(G)}\frac{d(v)(d(v)-1)}{m}\right)+\left(1-\frac{n-c}{n-3}\right)(m-1)\right\}+1.$$

此外，当且仅当以下情况成立时，该界可取到：

– 如果 $c=2$，则 G 是满足 $m \ge 1$ 的任意图；– 如果$c=3$，则 G 要么是完全二部图，要么 G 由 t 条不相交的边组成；– 如果 $4 \le c \le n-1$，则 G 是完全图，或 G 是星图，或 G 恰有一条边。

如果 $c=n$，那么该上界显然表明 G 的边数为 m。

这里， $d(v)$表示顶点 $v \in V(G)$的度。

定理3（[18], [23]）。每棵 n 阶的树至多有 $1+2^{n-1}$ 个独立集。此外，当且仅当该树是有 $n-1$ 片叶子的星图时，上界取等。

*计算 应用研究所，M. Picone，Via dei Taurini 19，00185 罗马，意大利 .
nicola.apollonio@cnr.it

这里将树的顶点的空子集计为一个独立集。

贯穿上述结果的共同主线，是随机变量 $M_c(G)$ 的二阶矩结构。该随机变量定义为：由随机抽取的 c 个顶点所诱导的 G 的子图的边数。请注意，顶点是以不放回抽样的方式抽取的。 $M_c(G)$ 的前两阶矩公式见 [4]，该公式还导出了一个用于线性时间计算这些矩的算法。提出这一公式，最初是为了评估：与适当零模型下的期望值相比，大型网络中由社区诱导的子图的观测边密度在统计上是否显著。本文将以不同方式呈现该公式。随机变量 $M_c(G)$ 编码了关于图 G 的多种信息。例如， $M_c(G)$ 的支撑集 $R_c(G)$，即 $M_c(G)$ 以正概率取到的值所构成的集合，就是 G 的一个 c 阶诱导子图可能具有的不同边数的集合。相关集合 $\cup_c R_c(G)$ 称为 [11]，中 G 的子图大小范围。Erd˝os 和 McKay 首先研究了这一集合（见 [2]），并提出了一个关于区间 $\{0, 1, \ldots, u\}$ 的最大宽度的猜想，该区间包含于 G 的子图大小范围之中。他们还给出了 u 的一个上界，该上界取决于 G 的诱导平凡子图的最大数量。回顾一下，若一个诱导子图是完全或余完全的，则称其为平凡子图。 [2]，显著改进了该上界，而对于随机图，这一猜想在 [11] 中得到了解决。基于上述讨论，定理 1 刻画了满足以下条件的图 G： $R_c(G)$ 是单元素集合 $\{u(c)\}$；也就是说，这些图 G 满足 $M_c(G)$ 几乎处处为常数。由于随机变量几乎处处为常数，当且仅当其方差为零，因此定理 1 实际上刻画了满足 $M_c(G)$ 的方差为零的图 G 。定理 1 在 [6] 中借助极值图论工具得到了大幅推广。定理 2 也可以看作关于 $R_c(G)$ 的一个陈述：实际上，它陈述的是 $R_c(G)$ 的最大值 $u(c)$。这样的最大值称为 c 最稠密子图 [17] 的边数。由于根据二阶 Fr´echet 不等式（见 [15] 和第 2 节），只要 $M_c(G)$ 不几乎处处为零，并以 $\mathbb{E}(\cdot)$ 表示期望，就有

$$\frac{\mathbb{E}\left(M_c(G)(M_c(G)-1)\right)}{\mathbb{E}\left(M_c(G)\right)} + 1 \leq u(c),$$

一旦我们计算出其中出现的期望——我们将在定理 5 中完成这项工作——利用这样的不等式便可相当直接地推出定理 2。实际上，当 $\ell(c)$ 中 $M_c(G)$ 的 $R_c(G)$ 最小值非零时，可以证明（见定理 4），Fr´echet 二阶不等式可由下列不等式加强：

$$\sigma_c^2(G) \leq \{u(c) - \mathbb{E}\left(M_c(G)\right)\}\{\mathbb{E}\left(M_c(G)\right) - \ell(c)\} \tag{1}$$

该不等式由 Bhatia 和 Davis [8] 针对离散有限支撑集的随机变量给出。这里， $\sigma_c^2(G)$ 表示 $M_c(G)$ 的方差。我们从 [3] 中了解到这一不等式；该文似乎也是这一不等式在图论社群中出现的最早出处。尽管 Bathia-Davis不等式加强二阶 Fr´echet 不等式这一断言确实需要证明（见定理 4），但除此之外，前者显然改进了后者，因为它利用了关于 $M_c(G)$ 支撑集的更多信息。注意， $\ell(c) = 0$ 当且仅当 $\alpha(G)$ 的独立数 G 至少为 c。一般而言，

$$c - \alpha(G) \leq \ell(c) \tag{2}$$

因为，如果 $c \geq \alpha(G)$，那么任意由 c 个顶点组成的集合 S 都会诱导出至少 $c - \alpha(G)$ 条边：如果由 S 诱导出的边数 h 少于 $c - \alpha(G)$，那么从 S 中删除至多 h 个顶点后，就会得到一个基数至少为 $\alpha(G)$ 的稳定集 $S' + 1$。因此，对于所有满足 $c \geq \alpha(G)$ 的 c，定理 2 给出的界都可以进一步加强。

除了 $M_c(G)$ 的支撑集 $R_c(G)$ 外，集合 $T(k) = \{r \in R_c(G) \mid r \geq k\}, k \in \mathbb{N} \cup \{0\}$ 也很有信息量。这些集合称为 $M_c(G)$ 的尾部概率。例如，如果 G 不是二分图，Pr $\left\{M_c(G) = c^2/2\right\}$ 是 G 中含有 c 个顶点的团所占的比例，而 $\binom{n}{c}$ Pr $\left\{M_c(G) = c^2/2\right\}$ 是

其数量。显然，$\Pr\left\{M_c(G) = c^{\underline{2}}/2\right\} = \Pr\left\{M_c \in T(c^{\underline{2}}/2)\right\}$。同样的说明也适用于给定树中阶为 c 的子树所占比例，甚至适用于二部图中阶为 c 的平衡完全二部图。同样，$\Pr\{M_c(G) = 0\}$ 是稳定集所占的比例，而 $\Pr\{M_c(G) = 0\} = 1 - \Pr\{M_c(G) \in T(1)\}$。由于尾概率 $\Pr\{M_c(G) \in T(k)\}$有若干已知上界，$k \in \mathbb{N}\cup\{0\}$，类似地，$\Pr\{M_c(G) \in T(1)\}$有若干已知下界，且这两类界都用 $M_c(G)$的二阶矩表示，因此，一旦我们知道 $M_c(G)$的二阶矩结构，就可以使用这些界中的任意一个，给出 G的平凡子图数量的上界。例如，著名的 Chung-Erd˝os 不等式的一个版本（见 [12],[15], [21],[22]）如下所示。

$$\Pr\{M_c(G) \geq 1\} \geq \frac{\left(\mathbb{E}\left(M_c(G)\right)\right)^2}{\mathbb{E}\left(M_c(G)^2\right)} \tag{3}$$

并且是 Petrov 不等式和 Paley-Zygmund 不等式的特例（另见 [21]）。由于随机变量的二阶矩总是不小于其一阶矩的平方，因此可得

$$\binom{n}{c}\left(1 - \frac{\left(\mathbb{E}\left(M_c(G)\right)\right)^2}{\mathbb{E}\left(M_c(G)^2\right)}\right)$$

始终是具有 c 个顶点的独立集数量的有效上界。

本文其余部分安排如下。第 2 节介绍本文将贯穿使用的矩不等式背景。第 3 节计算$M_c(G)$的二阶矩，并给出一些例子。第 4 节证明本文给出的结果，并进一步讨论 $M_c(G)$二阶矩结构的应用。

2 预备知识

对于正整数 t，符号 $[t]$ 表示前 t 个正整数构成的集合。对于非负整数 s，我们用 $t^{\underline{s}}$表示 t的 s 阶下降阶乘，即乘积 $t(t-1)\cdots(t-s+1)$。若 $s > t$，则 $t^{\underline{s}} = 0$。注意，$t^{\underline{0}} = 1$、$t^{\underline{t}} = t!$，且 $\binom{t}{s} = t^{\underline{s}}/s!$。我们讨论支撑集 R为非负整数区间 $\{\ell, \ell+1, \ldots, u\}$ 子集的离散随机变量。若 M 是任意此类随机变量且 $r \in R$，则用 p_r 表示概率 $\Pr\{M = r\}$。我们用 S_1表示 $\mathbb{E}(M)$，用 S_2 表示 $\frac{1}{2}\mathbb{E}(M(M-1))$。它们称为 M的前两个二项矩。我们用 μ_2表示 $\mathbb{E}\left(M^2\right)$，用 σ^2表示 $\mu_2 - S_1$（即 M的方差）。有 $2S_2 + S_1 = \mu_2$和 $\sigma^2 = 2S_2 - S_1(S_1 - 1)$。我们采用这种记法，是因为我们关注的尾部不等式通常以这种术语表述，这与 Bonferroni 不等式和容斥原理有关（见 [15], 和 [24]）。当 $M = M_c(G)$时，对于相应的矩，我们记作 $S_1(c)$、$S_2(c)$、$\mu_2(c)$和 $\sigma^2(c)$。

下面的引理是所谓的 Fr´echet 恒等式（见 [24]）的一个特例，我们将在下一节中用它来计算G的前两阶矩 M_c。

引理 1. 设 $\mathcal{A} = \{A_e \mid e \in E\}$ 是某个概率空间 Ψ中的有限个事件。若M 是一个随机变量，其定义为 $\mathcal{A}$中的事件中发生的事件数，即$M = \sum_e I(A_e)$，其中 $I(\cdot)$是指示函数，则

$$S_1 = \sum_{e\in E} \Pr\{A_e\} \;\; and \;\; S_2 = \sum_{\substack{\{e,f\}\subseteq E \\ e\neq f}} \Pr\{A_e \wedge A_f\}.$$

定理 4. 设 M是一个随机变量，使得 $R \subseteq \{\ell, \ell+1, \ldots, u\} \subseteq \mathbb{N}\cup\{0\}$ 成立，且可能有$\ell = u$。则

(i) Second order Fréchet inequality，$(u-1)S_1 \geq 2S_2$，当且仅当 $\#R = 1$ 或 $\#R = 2$ 且 $\ell = 0$ 时，等号成立。

(ii) Bathia-Davis inequality, $\sigma^2 \leq (u - S_1)(S_1 - \ell)$，当且仅当 $\#R \leq 2$ 时，等号成立。

(iii) Chung-Erdős inequality, $\Pr\{M \geq 1\} \leq \frac{S_1^2}{2S_2+S_1}$。

(iv) Petrov's inequality，对于任意满足 $1 \leq t \leq S_1$ 的 t，$\Pr\{M \geq t\} \leq \frac{(S_1-t+1\ ^2)}{2S_2+S_1}$。(v) 对于 $t \geq 1$, t 为整数, Second order factorial moment inequality, $\Pr\{M \geq t\} \leq \frac{2S_2}{t^2}$。

此外，当 $\ell = u$ 或 $\ell = 0$ 时，Bathia-Davis不等式与二阶 Fr´echet 不等式等价；而当 $0 < \ell < u$ 时，前者强于后者。

证明。上述陈述中的所有不等式（最后一个除外）均可见于 [15]，且几乎所有不等式都可通过线性规划对偶性得到（参见 [9]）。注意到对于正整数 t，有 $M \geq t \Longrightarrow M(M-1) \geq t(t-1)$，因此由马尔可夫不等式可直接推出 (v) 中的二阶阶乘不等式。Chung-Erd˝os 不等式最初写作

$$\Pr\left\{\bigvee_e A_e\right\} \geq \left(\sum_{e\in E} \Pr\{A_e\}\right)^2 \Bigg/ \sum_{(e,f)\in E^2} \Pr\left\{A_e \bigwedge A_f\right\}$$

对于概率空间 [12], 中的任意事件 A_e、$e \in E$，通过定义 M 为在 A_e's 中发生的事件数，它便具有引理 1 后 (iii) 中给出的形式。接下来需要证明 (i) 中不等式的紧性，以及该不等式与 (ii) 中不等式之间的关系。我们同时证明这两个事实。由于 $\sigma^2 = 2S_2 - S_1(S_1 - 1)$，如果 $\ell = 0$，则 Bathia-Davis不等式变为 $2S_2 \leq S_1(u-1)$，这恰好是关于 $u > 1$ 的二阶 Fr´echet 不等式。如果 $\ell = u$，则两个不等式都等价于 $S_2 = S_1(S_1 - 1)$：前者是因为 $0 = \sigma^2 = 2S_2 - S_1(S_1 - 1)$，后者是因为

$$2S_2 = \mathbb{E}\binom{M}{2} = u(u-1) = S_1(S_1-1)$$

因为 $p_u = 1$ 和 $\ell = S_1 = u$。现在设 $\ell > 0$。在这种情况下，Bathia-Davis不等式变为 $2S_2 \leq S_1(\ell + u - 1) - \ell u$。现在 S_1 为正，因为 $\ell > 0$。因此，我们可以将不等式两边除以 S_1，从而得到

$$\frac{2S_2}{S_1} + 1 + \ell\left(\frac{u}{S_1} - 1\right) \leq u$$

这意味着二阶 Fr´echet 不等式成立，因为 $\ell\left(\frac{u}{S_1} - 1\right)$ 为正。 □

3 $M_c(G)$ 的前两阶矩

设 G 是一个阶为 n、边数为 m 的图，$\mathcal{V}_c$ 是 $V(G)$ 中所有含 c 个元素的子集构成的集合。考虑概率空间 $\Psi_c(G)$，其样本空间为 $\mathcal{V}_c$，并配备均匀测度 p。因此，$p(U) = c!/n^{\underline{c}}$, $\forall U \in \mathcal{V}_c$，且所有 U 等可能。随机变量 $M_c(G)$ 定义在此概率空间上，其值为从集合 $U \in \mathcal{V}_c$ 中以概率 $p(U)$ 抽取的集合所诱导子图的边数。由于 p 是均匀测度，显然 $\Pr\{M_c(G) = k\} = i_G(c, k)c!/n^{\underline{c}}$，其中 $i_G(c,k)$ 表示 G 中阶为 c、边数为 k 的诱导子图数量。注意到

– 对任意图 G 和任意 $c \geq 2$，$i_G(c,0)$ 是 c 个顶点构成的独立集数量；

– 对于任意非二分图 G 和任意 $c \geq 2$，$i_G\left(c, \frac{c^2}{2}\right)$ 表示 c 个顶点的团的数量；而如果 G 是二分图，并且为简单起见，c 是偶数，那么 $i_G\left(c, \frac{c^2}{4}\right)$ 表示 G 的平衡完全二部图的数量；

– 若 T 是一棵树，则 $i_T(c, c-1)$ 是阶为 c 的子树数量。

$M_c(G)$ 的支撑集 $R_c(G)$ 包含于区间 $\{\ell(c)\ldots\ldots, u(c)\}$ 内。正如第 1 节所述，$\ell(c) \geq \max\{0, c-\alpha(G)\}$，其中 $\alpha(G)$ 是 G 的独立数。$u(c)$ 的一个平凡上界为 $\gamma(c)$，其中 γ_c 或为 $\lfloor c/2 \rfloor \lceil c/2 \rceil$，即以 c 个顶点构成的完全二部图的边数（若 G 是二分图）；或 γ_c 等于 $\binom{c}{2}$，即以 c 个顶点构成的完全图的边数。关于 $M_c(G)$ 最大值的一个更精细界来自 Motzkin-Straus 定理 [19]：该定理断言，若 A 是 G 的邻接矩阵，则二次型 $x'Ax/2$ 在 $\mathbb{R}^n$ 的标准单纯形上的最大值等于 $\frac{1}{2}(1-\frac{1}{\omega})$，其中 ω 是 G 的团数。因此，若 z 是 $V(G)$ 中一个含有 c 个顶点的集合的特征向量，则 $c^{-1}z$ 属于标准单纯形，由此可得 $u(c) \leq \frac{c^2}{2}\left(1-\frac{1}{\omega}\right)$。所以，$u(c)$ 的上界是 γ_c 与 Motzkin-Straus 界中较小者。注意，若该图是二分图，则这两个界相同。若 G 是非二分图，则当且仅当 $c \leq \omega - 1$ 时，γ_c 小于 Motzkin-Straus 界。下面来看几个例子。

例 1。

(a) 若 $G \cong K_n$，则对所有 $c \in [n]$，$R_c(K_n) = \{\gamma_c\}$。因此，$\Pr\{M_c(K_n) = \gamma_c\} = 1$；故该随机变量是退化的。(b) 若 $G \cong K_{1,n-1}$，则 $R_c(K_{1,n-1}) = \{0, c-1\}$，且 $M_c(K_{1,n-1})$ 以概率 $1-\frac{c}{n}$ 取值 0，而 $M_c(K_{1,n-1})$ 以概率 $\frac{c}{n}$ 取值 $c-1$。注意到 $1-\frac{c}{n} = \binom{n-1}{c}/\binom{n}{c}$ 和 $\frac{c}{n} = \binom{n-1}{c-1}/\binom{n}{c}$。(c) 若 $G \cong K_{d,d}$，$n = 2d$，并为简单起见假设 c 为偶数，则 $R_c(K_{d,d}) = \{0, c-1, 2(c-2)\ldots, c^2/2\}$。在这种情况下

$$\Pr\{M_c(K_{d,d}) = st\} = 2\frac{\binom{d}{s}\binom{d}{t}}{\binom{n}{s+t}},$$

也就是说，如果 $Z = \frac{c-\sqrt{c^2-4M_c(K_{d,d})}}{2}$，那么在事件 $(Z \leq c/2)$ 条件下 Z 的分布与在事件 $(Z \geq c/2)$ 条件下 Z 的分布相同，并且两者都服从参数为 c、d 和 $2d$ 的超几何分布。

(d) $R_c(K_2^d) = \{0, 1, \ldots, c\}$，其中 c 是一个偶数正整数，$c \leq d$ 且 K_2^d 是由 d 条不相交的边组成的图。对于满足 $0 \leq k \leq d$ 的 k，有

$$\Pr\left\{M_c(K_2^d) = k\right\} = \binom{2d}{c}^{-1}\binom{d}{k, c-2k}2^{c-2k}.$$

为说明这一点，注意到左边的概率可以解释为：在对顶点进行红色和蓝色的随机着色时，出现 k 条两端均为蓝色的边、$c-2k$ 条两端颜色不同的边，以及 $d-c+k$ 条两端均为红色的边的概率。上述公式定义一个概率分布这一事实归结为验证：

$$\binom{2d}{c} = \sum_{k=0}^{c/2}\binom{d}{k, c-2k}2^{c-2k},$$

一种通过观察以下事实即可得出的恒等式

$$\binom{2d}{c} = [X^c](1+X)^{2m} = [X^c](1+2X+X^2)^m = \sum_{k=0}^{c/2}\binom{d}{k, c-2k}2^{c-2k},$$

其中，在最后一个等式中，我们使用了多项式定理，而 $[X^c]Q(X)$ 表示多项式 $Q(X)$ 中 X^c 的系数。

下述定理以更清晰简洁的方式证明了 [4] 中给出的 $S_1(c)$ 和 $S_2(c)$ 的公式。

定理 5. 设 G 是一个阶为 n 、边数为 m 的图，并设 $d(v)$ 表示顶点$v \in V(G)$ 的度。对于满足 $2 \le c \le n$的整数 c， $M_c(G)$ 的一阶和二阶二项矩由下列公式给出

$$S_1 = \frac{c^{\underline{2}}}{n^{\underline{2}}} m \tag{4}$$

and

$$2S_2 = \frac{c^{\underline{3}}}{n^{\underline{3}}} \left\{ \frac{n-c}{n-3} \sum_{v \in V(G)} d(v)(d(v)-1) + \frac{c-3}{n-3} m(m-1) \right\}. \tag{5}$$

此外， $S_2(c)$ 可在线性时间内计算。

证明。我们使用引理 1。对于 $e \in E(G)$， $e = uv$，设 A_e 是 $\Psi_c(G)$ 中的事件，当 u 和 v 属于随机抽取的子集 $U \in \mathcal{V}_c$时，该事件发生。换言之， A_e 表示边 e 由 U诱导的事件。因此， $M_c(G)$ 是在所有 A_e 中发生的事件数；等价地， $M_c(G) = \sum_{e \in E(G)} I(A_e)$，其中 $I(\cdot)$ 是指示函数。根据引理 1，要计算 $S_1(c)$ 和 $S_2(c)$，我们只需计算所有 $e \in E(G)$ 的 $\Pr\{A_e\}$，以及 G中每一对不同的边 e 和 f 的 $\Pr\{A_e \wedge A_f\}$。第一个任务很容易完成： e 由 U 诱导的概率就是

$$\binom{n-2}{c-2} \Big/ \binom{n}{c} = \frac{n^{\underline{c-2}}}{(c-2)!} \frac{c!}{n^{\underline{c}}} = \frac{c^{\underline{2}}}{n^{\underline{2}}},$$

也就是说， U 包含两个指定顶点的概率。对$E(G)$求和后，我们得到（4）。下面计算$\{A_e \wedge A_f\}$ 中两条不同的边 e和 f的 Pr。我们分两种情况讨论： e 和 f 相邻，此时记作 $e \sim f$；或者它们不相邻，此时相应地记作 $e \not\sim f$。如果 $e \sim f$，那么 e 和 f 跨越三个顶点（我们不妨假设图是简单图）。因此，当且仅当这三个顶点属于随机抽取的子集 U时， $A_e \wedge A_f$才会发生。仿照之前的推理， U 包含三个指定顶点的概率为 $\frac{c^{\underline{3}}}{n^{\underline{3}}}$。如果$e \not\sim f$，那么它们跨越四个顶点；在这种情况下，Pr $\{A_e \wedge A_f\}$为 $\frac{c^{\underline{4}}}{n^{\underline{4}}}$，也就是 U 包含四个指定顶点的概率。由此可得

$$\begin{aligned} S_2(c) &= \sum_{e \sim f} \frac{c^{\underline{3}}}{n^{\underline{3}}} + \sum_{e \not\sim f} \frac{c^{\underline{4}}}{n^{\underline{4}}} \\ &= \frac{c^{\underline{3}}}{n^{\underline{3}}} \sum_{v \in V(G)} \binom{d(v)}{2} + \frac{c^{\underline{4}}}{n^{\underline{4}}} \left[\binom{m}{2} - \binom{d(v)}{2} \right] \end{aligned}$$

经过一些简单的运算后，我们得到（5）。根据（4）和（5），计算 M_c 的前两个二项矩是多项式时间可解的，因为这相当于计算 G 的度，而该任务显然可以在 $O(m)$时间内完成。 □

下面列出例 1 中 (a) 至 (d) 所示图的 $M_c(G)$ 的前两个二项矩。在下文中，当 G 已明确时，我们省略对 G 的显式引用。特别地，我们分别将 M_c、 R_c 写作 $M_c(G)$ 和 $R_c(G)$。

例 2（例 1 续）。为简化记号，在书写二项矩时，我们省略其对 c 的依赖关系。回忆一下，对于正整数 d， K_2^d 表示一个阶为 $2d$ 、边数为 d 且由 d 条不相交的边组成的图的同构类。

(a) $G \cong K_n$、 $n \ge 1$、 $0 \le c \le n$。有

$$S_1 = \frac{c^{\underline{2}}}{2}, \quad \textit{and} \quad 2S_2 = c^{\underline{3}} + \frac{c^{\underline{4}}}{4}.$$

注意， K_n 有 $\frac{n^3}{2}$ 个 P_3 副本，以及 $\frac{n^4}{8}$ 个 K_2^2副本。此外还要注意，

$$\frac{2S_2}{S_1}+1=\frac{c^2}{2},$$

是阶为 c 的诱导子图的最大值。

(b) $G\cong K_{1,\ n-1}$、 $n\geq 2$、 $2\leq c\leq n$。有】【。

$$S_1=\frac{c^2}{n},\quad and\quad 2S_2=\frac{c^3}{n}.$$

注意，$K_{1,\ n-1}$ 不包含 K_2^2 的副本。此外，还要注意的是

$$\frac{2S_2}{S_1}+1=c-1 \tag{6}$$

是阶为 c 的诱导子图的最大边数。

(c) $G\cong K_{d,\ d}$, $n=2d$ 并且，为简单起见， c 为偶数且 $2\leq c\leq 2d-1$。有以下结论。

$$S_1=\frac{c^2}{2}\frac{d}{2d-1},\quad and\quad 2S_2=\frac{c^3}{4}\frac{d}{2d-1}\left\{2+(c-3)\frac{d-1}{2d-3}\right\}.$$

注意到

$$\frac{2S_2}{S_1}+1=\frac{c^2}{4}+\left(\frac{(c-2)(c-3)}{4(2d-3)}\right).$$

因此，

$$c^2/4-\left\lceil\frac{2S_2}{S_1}+1\right\rceil$$

对所有满足 $2\leq c\leq 2d-1$ 的整数 c，该值为 $O(c)$；而如果 c 为 $\Theta(d)$，则该差值为 $o(1)$。

(d) $G\cong K_2^d$、 $n=2d$，并且为简单起见， c 为偶数，且满足 $2\leq c\leq 2d-1$。有

$$S_1=\frac{c^2}{2(2d-1)},\quad and\quad 2S_2=\frac{c^4}{4(2d-1)(2d-3)}.$$

请注意， G 有一个 P_3 的副本，并且

$$\frac{2S_2}{S_1}+1=\frac{c-2}{2}\left(\frac{c-3}{2d-3}\right)+1.$$

因此，当且仅当 c 是 $\Theta(d)$ 时，阶为 c 的诱导子图的最大边数 $\frac{c}{2}$ 与值 $\lceil\frac{2S_2}{S_1}+1\rceil$ 之间的差为 $o(c)$。

请注意，即使在这些引人注目的特殊情形下，计算前两个二项矩也可能并非易事。例如，考虑完全二部图的情形：直接计算 $2S_2$（甚至是 S_1）作为 $2\mathbb{E}\binom{M_c(G}{2})$，就需要计算以下求和式

$$\sum_{s\leq c/2}\frac{\binom{d}{s}\binom{d}{t}}{\binom{n}{c}}\left\{[s(c-s)]^2-s(c-s)\right\}.$$

因此，作为一个副产品，定理 5 为例 1 中 (c) 和 (d) 所定义的随机变量的前两阶矩给出了组合恒等式。需要注意的是，这些随机变量的支撑集并不是非负整数的区间。

4 应用

本节将证明第 1 节中的结果，并进一步讨论定理 5 的应用。我们先从定理 1 开始。

4.1 定理 1 的证明

为了证明该定理，我们需要以下广为人知的事实。我们将以一种稍不常见的方式给出证明（本质上避免使用柯西–施瓦茨不等式或詹森不等式）。

引理 2. 设 G 是一个阶为 n、边数为 m 的图。G 的度 $d(v)$、$v \in V(G)$ 满足以下不等式。

$$\sum_{v \in V(G)} d(v)(d(v)-1) \geq \frac{4m^2}{n} - 2m. \tag{7}$$

正则图的度数可使（7）式中的等号成立。

证明。为了证明（7），我们注意到 $\sum_{v \in V(G)} d(v)(d(v)-1)$ 是 $2na$，其中 a 是 G 的度分布的二项矩（二阶），将其视为一个离散随机变量。该变量的一阶二项矩 b 恰好是 G 的平均度 $\frac{2m}{n}$。由于任意随机变量的方差均非负（因为它是一个非负随机变量的期望），因此可得 $2a - b(b-1) \geq 0$，这正是我们要证明的不等式。当且仅当度分布的方差为零时，（7）中的等号成立。因此，当所有度都相等时，等号成立。

□

定理 1 的证明。设 $\mathcal{G}$ 为满足该命题的图的集合。我们证明 $\mathcal{G}$ 中的图都是正则图。记 s 为阶为 c 的诱导子图的共同边数。令 U 为 $V(G)$ 的任意一个含有 $c+1$ 个顶点的子集，令 H 为由 U 诱导的图。由于对于所有 $v \in V(H)$，$H - v$ 的边数为 s，因此 H 是正则的。于是，任意 $v \in V(G)$ 在 U' 的每个由 c 个顶点组成的子集中，都有相同数量 ν 的邻点，其中 $V(G)$。因此 $d_G(v)\binom{n-1}{c-1} = \nu\binom{n}{c}$，我们得出 $d_G(v)$ 与 v 无关。接下来，该假设等价于：对于每个 $G \in \mathcal{G}$，$M_c(G)$ 的方差 σ^2 为零，即 $2S_2 - S_1(S_1-1) = 0$ 于 $\mathcal{G}$ 上。根据引理 2，有 $\sum_{v \in V(G)} d_G(v)(d_G(v)-1) = 4\frac{m^2}{n} - 2m, \forall G \in \mathcal{G}$。将最后一个表达式代入由 (5) 定义的 S_2，得到关于 m 的一个具有有理系数的二次方程，因此至多有两个实根。其中一个根显然是 $m = 0$，它对应于余完全图。另一个根通过求解由 $\frac{2S_2}{S_1} + 1 = S_1$ 给出的关于 m 的一次方程得到，其中除以 S_1 的操作由 $m \neq 0$ 保证。由于 $m = n^2/2$ 满足该方程，我们得出另一个根对应于完全图。□

4.2 利用 Fr´echet 和 Bathia-Davis 不等式求最稠密与最稀疏诱导子图边数的界

定理 2 给出了所谓的 c-最稠密子图边数的下界。这个边数的近似问题似乎非常困难（见 [14, 16]）。相关问题——近似所谓的 c-最稀疏（或 c-最稀疏）子图的边数 $\ell(c)$——至少和前一个问题一样困难 [26]。因此，若能为这些问题求得易于计算的界，将十分理想。我们先证明定理 2。正如第 1 节所述，定理 2 可由以下形式的二阶 Fr´echet 不等式相当直接地推出

$$\frac{2S_2(c)}{S_1(c)} + 1 \leq u(c)$$

根据定理 5。接下来只需证明，只有完全图、星图以及由一条边和一个含有 $n-2$ 个顶点的独立集组成的图，才能达到该界。

由例 2 中的 (a) 可知，完全图能够达到该界，因此我们只需证明，星图和恰好诱导出一条边的图是仅有的其他能够达到该界的图。为此，我们需要以下两个极值结果。

定理 6（[6]）。设 G 是一个阶为 n 的图，且 c 是满足 $4 \leq c \leq n-4$的整数。若 $\#R_c = 2$，则 G要么是星图，要么是由一条边和 $n-2$ 个顶点构成的不相交并，要么是这类图的补图。

对于阶为 n 且边数为 m 的图 G，记 $\sum_{v \in V(G)} d(v)(d(v) - 1)$ 为 $D_2(G)$。另外，回顾一下，没有三角形的图称为三角形自由图。

引理 3。在阶为 n、$n \geq 4$且边数 m为正的图中，满足

$$\frac{2S_2(3)}{S_1(3)} + 1 = u(3) \tag{8}$$

的图要么是完全图，要么是完全二分图，要么是由 t 条不相交的边组成的图，$1 \leq t \leq \lfloor n/2 \rfloor$。

证明。设 G 是一个阶为 n 、边数为 m 且满足 (8) 的图。为简便起见，$u = u(3)$ 和$\ell = \ell(3)$。u 的可能取值为 0、1、2、3。由于 $m \geq 1$，$u \neq 0$。根据定理 4，对任意c，当且仅当 $\#R_c = 1$或 $\#R_c = 2$ 且 $\ell(c) = 0$时，M_c 达到二阶 Fr´echet 不等式的等号。若 $\#R_3 = 1$，则 G 是完全图：由于任意三个顶点都会诱导出一个三角形，因此任意两个顶点都相邻。另一方面，若 G 是完全图，则 $\#R_3 = 1$，结论成立。因此，我们可以假设 G 不是完全图且 $\#R_c = 2$。我们断言，当 $c = 3$时，$u \neq 3$ 必然成立。为此，注意到 $\ell = 0$ 意味着 G 有一个至少包含三个顶点的独立集 I，而 $u = 3$ 意味着 G 含有一个三角形 T。现在 $\#I \cap T \leq 1$。若 $n = 4$，则这不可能，因为 $\#T \cup I \geq 5$。因此 $n \geq 5$，且存在顶点 u、$v \in I \setminus T$。若这两个顶点之一的邻域（例如 u）漏掉了 T中的某个顶点w，则 u、w 以及 T 中的任意其他顶点会诱导出一条或两条边，这与 $R_3 = 2$矛盾。因此，存在一个顶点 $w \in T$，它同时与 u 和 v相邻。这三个顶点仍然诱导出两条边，与 $\#R_3 = 2$矛盾。由此可知，u 只能是 1 或 2，且 G 是三角形自由图。先设 $u = 1$。由于 m 如此，S_1 恒为正数，故由 (8) 得 $S_2 = 0$。根据定义，S_2 与 $D_2(G)$ 成正比，因此我们得到$d(v) \leq 1$。于是，G 由不相交的边组成，正如所述。现在设 $u = 2$。满足 $u = 2$ 的图必须是二分图。为此，注意到若 G 含有奇圈，则G 会含有一个长度至少为五的诱导奇圈 C。但这样一来，我们便能在 C 中找到三个恰好诱导出一条边的顶点（例如，一个顶点 u 和 $C - u$中任意两个相邻顶点），这与 $\#R_3 = 2$矛盾。此外，G 必须是完全二部图。事实上，设 A和 B 是 $V(G)$ 的任意两个部分，使得 $A \cup B = V(G)$，且 G 的所有边的一端都在 A 中，另一端都在 B 中。由于 $u = 2$，存在三个顶点 u、v 和 w，使得 uv、$uw \in E(G)$。不妨设 $u \in A$、v、$w \in B$。若 G 不是完全图，则存在一个顶点 $z \in A$ 和一个顶点 $z' \in B$，满足 z，且 z' 不相邻。因此，顶点 u、z' 和 v 将恰好诱导出一条边，这与 $\#R_3 = 2$矛盾。证毕。 □

定理 2 的证明。由定理 4 可知，当且仅当 $\#R_c = 1$ or $\#R_c = 2$ and $\ell(c) = 0$ 时，二阶 Fr´echet 不等式在 M_c 处取等号。由定理 6 和引理 3 可知，对于每个满足 $3 \leq c \leq n-4$ 的 c，定理 2 随即成立。还需对其他可行的 c 值证明该定理，即对 $c \in \{n-3, n-2, n-1\}$ 证明。我们断言，对于这些 c 值，如果 G 满足二阶 Fr´echet 不等式且取等号，那么 G 要么是一个有 $n-1$ 个叶子的星图，要么 G 恰有一条边。我们先证明 G 的独立数 $\alpha(G)$ 等于 $n-1$。设 I 是 $\alpha(G)$ 个顶点组成的独立集。假设 $\alpha(G) \leq n-2$。由于 $\ell(c) = 0$，因此 $\alpha(G) \geq c \geq n-3$。于是，存在两个顶点 $u, v \in V(G) \setminus I$。现在，$\{u\} \cup I$ 和 $\{v\} \cup I$ 都至少诱导出一条边。此外，u 和 v 必须都与 I 的每个顶点相邻：如果 u（例如）不具备这一性质，那么我们可以找到顶点 $w', w'' \in I$，使得 w' 与 u 相邻，但 w'' 不相邻；因此我们可以找到 $J \subseteq I \setminus \{w', w''\}$，使得 $\#J = c - 2$ and

$\{u, w'\} \cup J$和 $\{u, w''\} \cup J$诱导出的边数不同，这与$\#R_c \leq 2$矛盾。因此，u 和 v 必须与 I中的所有顶点相邻。然而，这是不可能的，因为对于 I的任意子集 I'和 I''，只要$\#I' = c-1$、$\#I'' = c-2$，集合 $\{u\} \cup I'$和 $\{u, v\} \cup I''$就会诱导出边数不同的子图，这再次与$\#R_c \leq 2$矛盾。这个矛盾说明 $\alpha(G) \geq n-1$，因此我们得出 $\alpha(G) = n-1$，因为 G 至少有一条边。设 I是 $n-1$ 顶点的一个独立集，并设 $u \notin I$。要么 u 恰好与I中的一个顶点相邻，此时结论成立；要么 u在 I 中有多个邻点。由上述相同理由， u 必须与 I 中的每个顶点相邻，因此我们得出 G 是一个有 $n-1$个叶子的星图。□

利用 Bathia-Davis不等式，我们可以得到更强的结果。

定理 7. 设 G 是一个阶为 n 、边数为 m的图。假设 G 既非完全图也非余完全图，并设 c为满足 $4 \leq c \leq n-4$的整数。若 $\ell_\star$和 $u^\star$是满足 $0 \leq \ell_\star \leq \ell(c) \leq u(c) \leq u^\star$的整数，则 G 有一个阶为 c 的子图，其边数至少为

$$\left\lceil \min\left\{ \frac{2S_2(c)}{S_1(c)} + 1,\, S_1(c) + \frac{\sigma^2(c)}{S_1(c) - \ell_\star} \right\} \right\rceil$$

以及一个阶为 c 的子图，其边数至多为

$$\left\lfloor S_1(c) - \frac{\sigma^2(c)}{u^\star - S_1(c)} \right\rfloor,$$

其中， $\sigma^2(c) = 2S_2(c) - S_1(c)(S_1(c) - 1)$、 $\ell(c)$和 $u(c)$分别是 R_c中 M_c 的最小值和最大值。当 $\ell_\star = \ell(c)$和 $u^\star = u(c)$时，下界和上界分别恰由定理 6中列出的图达到。对于 $u^\star$，我们总可以选取 γ_c；对于 $\ell_\star$，总可以选取 $c - \alpha^\star$，其中 $\alpha^\star$ 是 G的独立数的任一上界。此外，给定 $u^\star$ 和 $\ell_\star$后，这些界可以在与 G的边数成线性关系的时间内计算出来。

证明。该定理可直接通过反转 Bathia-Davis不等式得出，并利用定理 4 和定理 6 得到紧性。为作说明，考虑当 G 是具有 $n-1$ 个叶节点的星图且 $c \leq n-1$时， $\ell(c)$ 的上界。在这种情况下， $u(c) = c-1$。根据例 2 中的 (b)，有

$$S_1(c) = \frac{c^2}{n}, \quad \text{and} \quad 2S_2(c) = \frac{c^3}{n}.$$

因此

$$u(c) - S_1(c) = (c-1)\left(\frac{n-c}{n}\right)$$

and

$$\frac{\sigma^2(c)}{u(c) - S_1(c)} = \frac{c^2}{n} = S_1(c).$$

因此，在这种情况下，对于所有可行的 c 值，该界都正确地达到 0。 □

值得注意的是，这些思路导出了一种稍微不那么标准的 Mantel 定理表述：一个三角形自由图至多有 $\frac{n^2}{2}$ 条边，并且当颜色类基数相同的完全二分图取到此上界。

设 G 是一个三角形自由图。与上文一样，令 $D_2(G) = \sum_v d(v)(d(v) - 1)$，并为简便起见，用 D_2 代替 $D_2(G)$。因为 G 是三角形自由图，所以有 $u(3) \leq 2$。现在，$2S_2 = \frac{6}{n^3} D_2$ 而 $S_1 = \frac{6}{n^2} m$。因此

$$\frac{2S_2}{S_1} = \frac{D_2}{(n-2)m}.$$

因此

$$u(3) \leq 2 \iff \frac{2S_2}{S_1} \leq 1 \iff D_2 \leq (n-2)m.$$

上述等价关系最右侧通常通过将 D_2 写成 $\sum_{uv\in E(G)}\{d(u)+d(v)-2\}$，并注意到在三角形自由图中，对于任意一对不同的相邻顶点，$d(u)+d(v)\le n$ 来得到。这实际上是我们相较于 Mantel 定理的标准证明所能声称的唯一（值得商榷的）简化。证明的其余部分按标准方法进行：利用引理 2 给出 D_2 的下界；由此得到不等式 $4\frac{m}{n}-2\le n-2$，这恰好就是该定理的结论。

4.3 平凡子图及低密度/高密度诱导子图的计数

本节给出平凡子图数量的上界。回顾一下，这些子图要么是独立集，要么是团。我们先证明定理3。回顾一下，对于一个图，G, $D_2(G)$ 表示 $\sum_{v\in V(G)} d(v)(d(v)-1)$。下文中，我们将 $D_2(\cdot)$ 视为定义在给定阶的图上的函数，因此省略对 G 的引用。

定理3的证明。c 个顶点的独立集数量等于 $\binom{n}{c}\Pr\{M_c=0\}$。由 Chung-Erd˝os 不等式，$\Pr\{M_c=0\}=1-\frac{S_1^2}{2S_2+S_1}$，其中我们省略了对 c 的引用，因为这是不言而喻的。下面针对阶为 n 的树，给出 $\frac{S_1^2}{2S_2+S_1}$ 的下界。该界给出了 $\Pr\{M_c=0\}$ 的上界。为此，将此表达式写成

$$S_1\left(1+\frac{2S_2}{S_1}\right)^{-1},$$

其中除以 S_1 是由于 $S_1>0$ as $S_1=\frac{c^2}{n^2}m=\frac{c^2}{n}$。现在注意，对 $\frac{2S_2}{S_1}$ 的任何上界都会给出 $\Pr\{M_c=0\}$ 的一个上界。对于每棵 n 阶的树，其比值

$$\frac{2S_2}{S_1}=\frac{c-2}{n-2}\frac{n-c}{n-3}\left(\frac{D_2}{n-1}\right)+\frac{c-2}{n-2}\frac{c-3}{n-3}(n-2)$$

仅取决于 D_2。在树中，$D_2/(n-1)$ 在具有 $n-1$ 个叶的星图上达到绝对最大值 $\frac{(n-1)(n-2)}{n-1}=(n-2)$。因此，每棵具有 n 个顶点的树都满足 $\frac{2S_2}{S_1}\le c-2$，从而

$$S_1\left(1+\frac{2S_2}{S_1}\right)^{-1}\ge\frac{c}{n}$$

每棵树都满足该关系。此外，当且仅当该树为星图时，等号成立。由此我们得出，在任意 n 阶的树中，含有 c 个顶点的独立集数量 $i(c,0)$ 满足以下关系

$$i(c,0)=\binom{n}{c}\Pr\{M_c=0\}\le\binom{n}{c}\frac{c}{n}=\binom{n-1}{c-1}.$$

由于少于两个顶点的独立集有 $n+1$ 个（顶点的空集和单元素集合），且没有任何树具有包含多于 $n-1$ 个顶点的独立集，因此树中的独立集总数等于 $1+n+\sum_{c=2}^{n-1} i(c,k)$，从而有上界 $2^{n-1}+1=1+n+\sum_{c=2}^{n-1}\binom{n-1}{c-1}$。□

定理3依赖于Chung-Erd˝os不等式，这是一种形式为 $\Pr\{M_c\in T(1)\}$ 的尾部不等式。通过对Petrov不等式给出的尾部 $\Pr\{M_c\in T(\ell(c)+1)\}$ 的下界取补，可以用同样的方法从上界估计边数为 $\ell(c)$ 的所有诱导子图的数量。同样的思路也可用于从上界估计边数为$u(c)$ 的诱导子图数量。在这种情况下，$\Pr\{M_c=u(c)\}=\Pr\{M_c\in T(u(c))\}$，我们可以使用二阶阶乘矩不等式从上界估计 $i(c,u(c))$。我们称 G 中满足边数为 $\ell(c)$ 或 $u(c)$ 的子图为平凡子图。此外，对于一个图类 $\mathcal{G}$，如果某个诱导子图的同构类属于 $\mathcal{G}$，则称 G 包含一个 $\mathcal{G}$-副本。如果对于所有 c，G 中平凡子图的同构类都属于 $\mathcal{G}$，那么 $i(c,u(c))$ 就是 G 中 $\mathcal{G}$-副本的数量，我们可以使用尾部不等式从上界估计该数量。当 $\mathcal{G}$ 是

完全图类，且 $u(c) = \frac{c^2}{2}$ 时，平凡子图就是 G 的团。如果 G 是一棵树 T， $\mathcal{G}$ 是树类，且 $u(c) = c - 1$，那么平凡子图就是 T 的子树。当 G 是二部图， $\mathcal{G} = \{K_{\lfloor \frac{d}{2} \rfloor, \lceil \frac{d}{2} \rceil} \mid d \in \mathbb{N}\}$ 且 $u(c) = \lfloor \frac{c}{2} \rfloor \lceil \frac{c}{2} \rceil$时，平凡子图就是 G 的平衡完全二部图。通过补图运算，我们可以得到边数为 $\ell(c)$ 的平凡子图的类似概念。让我们用最简单的情形来说明这种方法，即给定一棵树的子树这一情形。

定理 8. 任意一棵阶为 n 的树至多有 2^{n-1} 个子树，且当且仅当 T 是星图时达到此上界。

证明。这个证明实际上与前一个几乎完全相同。如上所述，当 T 在阶为 n 的树中变化时， D_2 的绝对最大值为 $(n-1)(n-2)$，且在星图上取得。将这一事实代入 (5) 并作简单变形，可得 $S_2(c) \leq \frac{c^3}{n}$，等号当且仅当树为星图时成立。因此，由二阶阶乘矩不等式（定理 4 中的不等式 (v)），T 的阶为 c 的子树数 $i(c, c-1)$ 不超过 $\binom{n}{c} \frac{2S_2(c)}{c-1}() (c-2) = \binom{n}{c} \frac{c}{n}$。后一个表达式等于 $\binom{n-1}{c-1}$。注意到 T 的所有子树数为 $n + \sum_{c \geq 2} i(c, c-1)$，其中项 n 计数的是单顶点子树，即可得出结论。

□

平凡子图数量的上界可以用一种更一般的形式给出，这种形式也有助于网络科学中的应用。事实上，考虑以下常见情形：你拥有一个大型网络 G，观察到一组顶点，并想判断这些顶点是否构成一个社群。尽管社群的概念较为模糊，但判定一组 C 个 c 顶点构成社群时，最广为接受的标准之一是：这些顶点诱导出 G 的超稠密子图。这里的“超稠密”是指观测到的子图的（边）密度高于 G 的密度。由于 G 的密度为 $2m/n^2$，而 C 的密度为 $2m(C)/c^2$，其中 $m(C)$ 是由 C 诱导的边数，因此， C 为超稠密，当且仅当 $m(C) \geq S_1(c)$，也就是说，由 C 诱导的边数至少等于 $M_c(G)$ 的期望值。然而，仅知道 C 是一个超稠密集合，并没有太大信息量，因为我们对 $R_c(G)$ 一无所知。为克服这一判据缺乏显著性的问题，作者[4] 提出了一种基于 $M_c(G)$ 尾部概率的随机化检验：若 $\Pr\{M_c(G) \geq m(c)\} \leq \alpha$，则在显著性水平 α（$\alpha \in \in (0, 1)$）下，判定 C 显著超稠密；也就是说，随机情况下观测到比实际观测结果更稠密的子图的概率低于某个固定的小值 α。显然，即便只是近似计算这一概率，也是一项非常困难的任务。判断 0 或 $c^2/2$ 是否属于 $M_c(G)$ 的支撑集，分别相当于求解独立集和团的判定问题（而且这些问题甚至不属于 APX 类）。不过，可以利用尾部不等式给出 α 的上界。 [4] （并在 [5] 中得到令人满意的推进）采用了这一方法：作者将单侧切比雪夫不等式（又称切比雪夫–坎特利不等式）应用于随机变量 $Z_c(G) = (M_c(G) - S_1(c))/\sigma^2(c)$，从而得到以下界

$$\Pr\{M_c(G) \geq t\} \leq \min\left\{1, \frac{1}{\xi^2(t) + 1}\right\}$$

为 $\xi^2(t) = (S_1(c) - t)^2/\sigma^2(c)$。这里，我们提出使用二阶阶乘矩不等式和Petrov不等式来改进尾部界。利用这样一个事实： G 中的超高密度子图是 G 的补图中的低密度子图，可以得到以下界。对于 $t \geq S_1(c)$，该界至少与 [4],中给出的界一样好。

$$\Pr\{M_c(G) \geq t\} \leq \min\left\{1, \frac{S_2(c)}{t^2}, 1 - \frac{(S_1(c) - t + 1)^2}{2\overline{S}_2(c) + \overline{S}_1(c)}, \frac{1}{\xi^2(t) + 1}\right\}$$

是 $\overline{S}_i(c)$ 的 $i = 1,2$ 阶二项矩，其中 i 阶二项矩取自 $M_c(\overline{G})$，而 $\overline{G}$ 是 G 的补图。例如，通过选择 $t = c^2/2$，上述界是对具有 c 个顶点的团数量的上界。一般而言，通过适当特化对 G 和 t 的选择，上述界可给出 G 的平凡子图数量的上界。

5 总结

本文将对 $M_c(G)$ 的二阶二项矩结构的研究提升为一种工具，可广泛应用于理论和实际问题。我们在引理 1 中使用到二阶的 Fr´echet 恒等式，本质上等价于容斥原理（[24]）。相应地，本文始终使用的二阶矩不等式，可以视为所谓 Bonferroni 不等式 [9, 15, 24] 的概率版本。 [13], [20], 和 [7] 首创了使用这些不等式对组合对象进行渐近计数。就本文目的而言， r 阶 Bonferroni 型不等式是形如以下形式的不等式

$$L(S_1, S_2, \dots S_r, t) \leq \Pr\{M \geq t\} \leq U(S_1, S_2, \dots S_r, t)$$

其中， M 是一个非负随机变量， $L(\cdot)$和 $U(\cdot)$是（可能为非线性的）实值函数， t 是一个非负实数，而$(S_i)_i \geq 1$ 是M的二项矩序列。本文所介绍的所有应用都基于二阶 Bonferroni 不等式。我们认为，尤其是在计算下界时，了解完整的二项矩序列具有重要价值（无论是对于理论应用还是实际应用）。我们计划在未来的论文中给出这样一个序列。

参考文献

[1] J. Akiyama, G. Exoo, F.Harary, 《所有诱导子图均同构的图》, 《马来西亚数学学会公报》2(2) (1979) 43–44。[2] N. Alon, M. Krivelevich, B. Sudakov, 《指定阶数的诱导子图》, 《图论杂志》43 (2003) 239–251。[3] E. Angel, R. Campigotto, C. Laforest, 《关于图的独立数的一个新下界》, 《离散应用数学》161 (6) (2013) 847–852。[4] N. Apollonio, P.G. Franciosa, D. Santoni, 《通过二阶统计量评估和测量网络中的同配性的一种新方法》, 《科学报告》12, 9757 (2022)。[5]N. Apollonio, D. Blankenberg, F. Cumbo, P.G. Franciosa, D. Santoni, 《通过 HONTO（同配性网络工具）评估网络中的同配性：人类蛋白质相互作用网络中染色体相互作用的案例研究》, 《生物信息学》39 (1) (2023) btac763, https://doi.org/10.1093/bioinformatics/btac763[6] M. Axenovich, B. J´ozsef, 《诱导 k-子图中边数较少的图》, 《工业与应用数学学会离散数学杂志》21(1) (2007) 264–272。[7] E.A. Bender, 《计数中的渐近方法》, 《工业与应用数学学会评论》16(4) (1974) 485–515。[8] R. Bhatia, C. Davis, 《一个更好的方差界》, 《美国数学月刊》（美国数学学会）107(4) (2000) 353–357。[9] E. Boros, A. Pr´ekopa, 《恰有 r 和至少 r 个事件在 n 个事件中发生的概率的闭式双侧界》, 《数学运筹研究》14 (1989) 317–342[10] J. Bos´ak, 《具有相同阶数和边数的诱导子图》, 《斯洛伐克数学杂志》33 (1983) 105–119。[11] N. Calkin, A. Frieze, and B. McKay, 《随机图中的子图边数》, 《组合学、概率与计算》1 (1992) 123–134。[12] K.L. Chung, P. Erd˝os, 《Borel–Cantelli 引理的应用》, 《美国数学会汇刊》72(1) (1952) 179–186。

[13] P. Erd˝os, I. Kaplansky, 《拉丁矩形的渐近数量》, Amer. J. Math. 68 (1946) 230–236。[14] U. Feige, G. Kortsarz, D. Peleg, 《稠密 k-子图问题》, Algorithmica 29(3) (2001) 410–421。[15] F.M. Hoppe, E. Seneta, 《Gumbel 恒等式、二项矩与 Bonferroni 和》, Int. Stat. Rev. 80(2) (2012) 269–292。[16] S. Khot, 《排除图最小二分、稠密 k-子图和二分团问题的多项式时间近似方案》, SIAM J. Comput. 36(4) (2006) 1025-–107。[17]G. Kortsarz, D. Peleg, 《如何选择稠密子图》, 《1993 年 IEEE 第 34 届计算机科学基础年会论文集》, 美国加利福尼亚州帕洛阿尔托, 1993, 页 692-701, doi: 10.1109/SFCS.1993.366818。[18] S.-B. Lin and C. Lin, 《具有较大和较小独立指标的树与森林》, Chin. J. Math. 23(3) (1995), 199-210。[19] T.S. Motzkin, E.G. Straus, 《图的极大值及 Tur ´an 定理的新证明》, Canad. J. Math. 17 (1965) 533-–540。[20] P.E. O'Neil, 《具有行和与列和限制的渐近问题和随机矩阵》, Bull. Amer. Math. Soc. 75 (1964) 1276–1282。[21] V.V. Petrov, 《尾概率的下界》, J. Statist. Planning and Inference 137(8) (2007) 2703-2705。[22] V.V. Petrov, 《关于事件并的概率的 Chung-Erd ˝os 不等式的一种推广》, J. Math. Sci. 147 (2007) 6932-–6934。[23] H. Prodingerand R.F. Tichy, 《图的斐波那契数》, Fibonacci Quarterly 20(1) (1982), 16-21。[24] V.N. Sachov, 《组合分析中的概率方法》, 《数学及其应用百科全书》第 56 卷, 剑桥大学出版社, 1997。[25] J. ˇSir´aˇn, 《含有许多边数相同的子图的图》, 《斯洛伐克数学杂志》30 (1980) 267–268。[26] R. Watrigant, M. Bougeret, R. Giroudeau, 《弦图中最稀疏 k-子图的近似》, Theory Comput. Syst. 58(1) (2016) 111–132。