Generalized birthday problem for October 12

Sumit Mukherjee,
Department of Statistics, Columbia

Joint work with B.B. Bhattacharya (U Penn) and S. Mukherjee (NUS)

Outline

The classical birthday problem

- Suppose there are n students in a class.

The classical birthday problem

- Suppose there are n students in a class. What is the chance that there are two students with the same birthday?

The classical birthday problem

- Suppose there are n students in a class. What is the chance that there are two students with the same birthday? Assume that a year has c days.

The classical birthday problem

- Suppose there are n students in a class. What is the chance that there are two students with the same birthday? Assume that a year has c days.
- The answer to this question is very well known.

The classical birthday problem

- Suppose there are n students in a class. What is the chance that there are two students with the same birthday? Assume that a year has c days.
- The answer to this question is very well known.
- In this case the number of student pairs with a common birthday is approximate Poisson with mean $\lambda = \frac{\binom{n}{2}}{c}$.

The classical birthday problem

- Suppose there are n students in a class. What is the chance that there are two students with the same birthday? Assume that a year has c days.
- The answer to this question is very well known.
- In this case the number of student pairs with a common birthday is approximate Poisson with mean $\lambda = \frac{\binom{n}{2}}{c}$.
- Consequently the chance of at least one common birthday is approximately $1 - e^{-\lambda}$.

First variant: Unfriendly class

- Suppose in a class of size n, not everybody knows each other.

First variant: Unfriendly class

- Suppose in a class of size n, not everybody knows each other.
- There is a friendship graph G between the n nodes, which determines which pair of students are friends.

First variant: Unfriendly class

- Suppose in a class of size n, not everybody knows each other.
- There is a friendship graph G between the n nodes, which determines which pair of students are friends.
- What is the chance that there are two people who know each other, and have the same birthday?

First variant: Unfriendly class

- Suppose in a class of size n, not everybody knows each other.
- There is a friendship graph G between the n nodes, which determines which pair of students are friends.
- What is the chance that there are two people who know each other, and have the same birthday?
- It was shown in Bhattacharya-Diaconis-M., AAP-2017 that the number of friend pairs with a common birthday is approximate Poisson with mean $\lambda' = \frac{|E(G)|}{c}$.

First variant: Unfriendly class

- Suppose in a class of size n, not everybody knows each other.
- There is a friendship graph G between the n nodes, which determines which pair of students are friends.
- What is the chance that there are two people who know each other, and have the same birthday?
- It was shown in Bhattacharya-Diaconis-M., AAP-2017 that the number of friend pairs with a common birthday is approximate Poisson with mean $\lambda' = \frac{|E(G)|}{c}$. Consequently, the answer is $1 - e^{-\lambda'}$.

First variant: Unfriendly class

- Suppose in a class of size n, not everybody knows each other.
- There is a friendship graph G between the n nodes, which determines which pair of students are friends.
- What is the chance that there are two people who know each other, and have the same birthday?
- It was shown in Bhattacharya-Diaconis-M., AAP-2017 that the number of friend pairs with a common birthday is approximate Poisson with mean $\lambda' = \frac{|E(G)|}{c}$. Consequently, the answer is $1 - e^{-\lambda'}$.
- Here $|E(G)|$ is the number of edges in G, i.e. the total number of friendship pairs in the class.

First variant: Unfriendly class

- Suppose in a class of size n, not everybody knows each other.
- There is a friendship graph G between the n nodes, which determines which pair of students are friends.
- What is the chance that there are two people who know each other, and have the same birthday?
- It was shown in Bhattacharya-Diaconis-M., AAP-2017 that the number of friend pairs with a common birthday is approximate Poisson with mean $\lambda' = \frac{|E(G)|}{c}$. Consequently, the answer is $1 - e^{-\lambda'}$.
- Here $|E(G)|$ is the number of edges in G, i.e. the total number of friendship pairs in the class.
- The classical birthday problem is a special case with $G = K_n$, where everyone knows everyone else.

Second variant: Fix the date

- Suppose as before there is a graph G which determines the friendship structure.

Second variant: Fix the date

- Suppose as before there is a graph G which determines the friendship structure.
- But in addition, we now require both of them to be born on October 12.

Second variant: Fix the date

- Suppose as before there is a graph G which determines the friendship structure.
- But in addition, we now require both of them to be born on October 12.
- Question: What is the chance that there are two people who know each other, and have birthday October 12?

Second variant: Fix the date

- Suppose as before there is a graph G which determines the friendship structure.
- But in addition, we now require both of them to be born on October 12.
- Question: What is the chance that there are two people who know each other, and have birthday October 12?
- In this case things are more interesting, and the answer depends on the graph in a more delicate way (than just the number of edges).

Formal set up

- Suppose G_n is a large (non random) labelled graph on the vertex set $V(G_n) := \{1, 2, \cdots, n\}$ with edge set $E(G_n)$.

Formal set up

- Suppose G_n is a large (non random) labelled graph on the vertex set $V(G_n) := \{1, 2, \cdots, n\}$ with edge set $E(G_n)$.
- Color each vertex of G_n with one of c_n colors chosen uniformly at random, independent of other vertices.

Formal set up

- Suppose G_n is a large (non random) labelled graph on the vertex set $V(G_n) := \{1, 2, \cdots, n\}$ with edge set $E(G_n)$.
- Color each vertex of G_n with one of c_n colors chosen uniformly at random, independent of other vertices.
- What is the limiting distribution of T_n, the number of monochromatic edges of color 1?

Formal set up

- Suppose G_n is a large (non random) labelled graph on the vertex set $V(G_n) := \{1, 2, \cdots, n\}$ with edge set $E(G_n)$.

- Color each vertex of G_n with one of c_n colors chosen uniformly at random, independent of other vertices.

- What is the limiting distribution of T_n, the number of monochromatic edges of color 1?

- Note that

$$T_n = \sum_{i<j} G_n(i,j) X_i X_j,$$

where $(X_1, \ldots, X_n)$ are i.i.d. Bern(p_n), with $p_n = \frac{1}{c_n}$.

Formal set up

- Suppose G_n is a large (non random) labelled graph on the vertex set $V(G_n) := \{1, 2, \cdots, n\}$ with edge set $E(G_n)$.
- Color each vertex of G_n with one of c_n colors chosen uniformly at random, independent of other vertices.
- What is the limiting distribution of T_n, the number of monochromatic edges of color 1?
- Note that

$$T_n = \sum_{i<j} G_n(i,j) X_i X_j,$$

where $(X_1, \ldots, X_n)$ are i.i.d. Bern(p_n), with $p_n = \frac{1}{c_n}$.

- We will assume that the sequence $\{G_n, p_n\}_{n \geq 1}$ are chosen such that $\mathbb{E}T_n = \frac{|E(G_n)|}{c_n^2} = O(1)$.

Example 0: Complete graph

- Suppose $G_n = K_n$ is the complete graph, and $p_n = \frac{1}{n}$.

Example 0: Complete graph

- Suppose $G_n = K_n$ is the complete graph, and $p_n = \frac{1}{n}$.
- In this case we have

$$T_n = \sum_{i<j} X_i X_j$$

Example 0: Complete graph

- Suppose $G_n = K_n$ is the complete graph, and $p_n = \frac{1}{n}$.
- In this case we have

$$T_n = \sum_{i<j} X_i X_j = \frac{1}{2}\left[\Big(\sum_{i=1}^{n} X_i\Big)^2 - \sum_{i=1}^{n} X_i\right].$$

Example 0: Complete graph

- Suppose $G_n = K_n$ is the complete graph, and $p_n = \frac{1}{n}$.
- In this case we have

$$T_n = \sum_{i<j} X_i X_j = \frac{1}{2}\left[\Big(\sum_{i=1}^n X_i\Big)^2 - \sum_{i=1}^n X_i\right].$$

- Now

$$S_n := \sum_{i=1}^n X_i \sim Bin(n, 1/n)$$

Example 0: Complete graph

- Suppose $G_n = K_n$ is the complete graph, and $p_n = \frac{1}{n}$.
- In this case we have

$$T_n = \sum_{i<j} X_i X_j = \frac{1}{2}\left[\Big(\sum_{i=1}^{n} X_i\Big)^2 - \sum_{i=1}^{n} X_i\right].$$

- Now

$$S_n := \sum_{i=1}^{n} X_i \sim Bin(n, 1/n) \xrightarrow{D} S \sim Pois(1).$$

Example 0: Complete graph

- Suppose $G_n = K_n$ is the complete graph, and $p_n = \frac{1}{n}$.

- In this case we have

$$T_n = \sum_{i<j} X_i X_j = \frac{1}{2}\left[\Big(\sum_{i=1}^{n} X_i\Big)^2 - \sum_{i=1}^{n} X_i\right].$$

- Now

$$S_n := \sum_{i=1}^{n} X_i \sim Bin(n, 1/n) \xrightarrow{D} S \sim Pois(1).$$

- Thus $T_n \xrightarrow{D} \binom{S}{2}$, where $S \sim Pois(1)$.

Example I: Erdős-Rényi graph

- Suppose G_n is an Erdős-Rényi random graph with parameter q fixed.

Example I: Erdős-Rényi graph

- Suppose G_n is an Erdős-Rényi random graph with parameter q fixed.

- Then the total number of vertices of color 1 is

$$S_n = \sum_{i=1}^{n} X_i \sim Bin(n, 1/n)$$

Example I: Erdős-Rényi graph

- Suppose G_n is an Erdős-Rényi random graph with parameter q fixed.

- Then the total number of vertices of color 1 is

$$S_n = \sum_{i=1}^{n} X_i \sim Bin(n, 1/n)$$

- Throwing away all vertices of other colors, we are left with a (smaller) Erdős-Rényi random graph on S_n vertices with parameter q.

Example I: Erdős-Rényi graph

- Suppose G_n is an Erdős-Rényi random graph with parameter q fixed.

- Then the total number of vertices of color 1 is

$$S_n = \sum_{i=1}^{n} X_i \sim Bin(n, 1/n)$$

- Throwing away all vertices of other colors, we are left with a (smaller) Erdős-Rényi random graph on S_n vertices with parameter q.

- Total number of edges in this smaller Erdős-Rényi random graph is $Bin(\binom{S_n}{2}, q)$, and so

$$T_n \stackrel{D}{=} Bin\Big(\binom{S_n}{2}, q\Big)$$

Example I: Erdős-Rényi graph

- Suppose G_n is an Erdős-Rényi random graph with parameter q fixed.

- Then the total number of vertices of color 1 is

$$S_n = \sum_{i=1}^{n} X_i \sim Bin(n, 1/n)$$

- Throwing away all vertices of other colors, we are left with a (smaller) Erdős-Rényi random graph on S_n vertices with parameter q.

- Total number of edges in this smaller Erdős-Rényi random graph is $Bin(\binom{S_n}{2}, q)$, and so

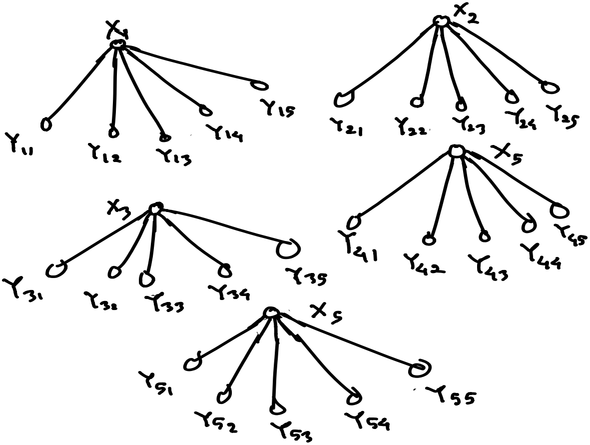
$$T_n \stackrel{D}{=} Bin\left(\binom{S_n}{2}, q\right) \stackrel{D}{\rightarrow} Bin\left(\binom{S}{2}, q\right), \text{ where } S \sim Pois(1).$$

Example II: Disjoint union of stars

- Suppose G_n is a disjoint union of $\sqrt{n}$ many $K_{1,\sqrt{n}}$ graphs, and $p_n = \frac{1}{\sqrt{n}}$.

Example II: Disjoint union of stars

- Suppose G_n is a disjoint union of $\sqrt{n}$ many $K_{1,\sqrt{n}}$ graphs, and $p_n = \frac{1}{\sqrt{n}}$.



EG : n=25, 5 COPIES OF $K_{1,5}$

Example II: Disjoint union of stars

- From the last figure,

$$T_n = \sum_{i=1}^{\sqrt{n}} X_i \sum_{j=1}^{\sqrt{n}} Y_{ij},$$

where $(X_1, \ldots, X_{\sqrt{n}})$ and $(Y_{ij})_{1 \le i,j \le \sqrt{n}}$ are mutually independent $\text{Bern}(p_n)$.

Example II: Disjoint union of stars

- From the last figure,

$$T_n = \sum_{i=1}^{\sqrt{n}} X_i \sum_{j=1}^{\sqrt{n}} Y_{ij},$$

where $(X_1, \ldots, X_{\sqrt{n}})$ and $(Y_{ij})_{1 \leq i,j \leq \sqrt{n}}$ are mutually independent Bern(p_n).

- Note that

$$\sum_{j=1}^{\sqrt{n}} Y_{ij} \sim Bin(\sqrt{n}, \frac{1}{\sqrt{n}})$$

Example II: Disjoint union of stars

- From the last figure,

$$T_n = \sum_{i=1}^{\sqrt{n}} X_i \sum_{j=1}^{\sqrt{n}} Y_{ij},$$

where $(X_1, \ldots, X_{\sqrt{n}})$ and $(Y_{ij})_{1 \leq i,j \leq \sqrt{n}}$ are mutually independent Bern(p_n).

- Note that

$$\sum_{j=1}^{\sqrt{n}} Y_{ij} \sim Bin(\sqrt{n}, \frac{1}{\sqrt{n}}) \xrightarrow{D} Y_i \sim Pois(1).$$

Example II: Disjoint union of stars

- From the last figure,

$$T_n = \sum_{i=1}^{\sqrt{n}} X_i \sum_{j=1}^{\sqrt{n}} Y_{ij},$$

where $(X_1, \ldots, X_{\sqrt{n}})$ and $(Y_{ij})_{1 \le i,j \le \sqrt{n}}$ are mutually independent Bern(p_n).

- Note that

$$\sum_{j=1}^{\sqrt{n}} Y_{ij} \sim Bin(\sqrt{n}, \frac{1}{\sqrt{n}}) \xrightarrow{D} Y_i \sim Pois(1).$$

- Thus $T_n \overset{D}{\approx} \sum_{i=1}^{\sqrt{n}} X_i Y_i$.

Example II: Disjoint union of stars

- From the last figure,

$$T_n = \sum_{i=1}^{\sqrt{n}} X_i \sum_{j=1}^{\sqrt{n}} Y_{ij},$$

where $(X_1, \ldots, X_{\sqrt{n}})$ and $(Y_{ij})_{1 \le i,j \le \sqrt{n}}$ are mutually independent Bern(p_n).

- Note that

$$\sum_{j=1}^{\sqrt{n}} Y_{ij} \sim Bin(\sqrt{n}, \frac{1}{\sqrt{n}}) \xrightarrow{D} Y_i \sim Pois(1).$$

- Thus $T_n \overset{D}{\approx} \sum_{i=1}^{\sqrt{n}} X_i Y_i$. Conditioning on $(X_1, \ldots, X_{\sqrt{n}})$, this has a $Pois\Big(\sum_{i=1}^{\sqrt{n}} X_i\Big)$ distribution.

Example II: Disjoint union of stars

- From the last figure,

$$T_n = \sum_{i=1}^{\sqrt{n}} X_i \sum_{j=1}^{\sqrt{n}} Y_{ij},$$

where $(X_1, \ldots, X_{\sqrt{n}})$ and $(Y_{ij})_{1 \le i,j \le \sqrt{n}}$ are mutually independent Bern(p_n).

- Note that

$$\sum_{j=1}^{\sqrt{n}} Y_{ij} \sim Bin(\sqrt{n}, \frac{1}{\sqrt{n}}) \xrightarrow{D} Y_i \sim Pois(1).$$

- Thus $T_n \overset{D}{\approx} \sum_{i=1}^{\sqrt{n}} X_i Y_i$. Conditioning on $(X_1, \ldots, X_{\sqrt{n}})$, this has a $Pois\Big(\sum_{i=1}^{\sqrt{n}} X_i\Big)$ distribution.

- This gives $T_n \xrightarrow{D} Pois(S)$, where $S \sim Pois(1)$.

Question: Study T_n

- Can we characterize the class of all possible limit distributions of T_n?

Question: Study T_n

- Can we characterize the class of all possible limit distributions of T_n?
- As already seen, this class contains mixtures of Poissons, and Binomials of quadratic functions of Poissons.

Question: Study T_n

- Can we characterize the class of all possible limit distributions of T_n?
- As already seen, this class contains mixtures of Poissons, and Binomials of quadratic functions of Poissons.
- Also, can we characterize when is this limit exactly a Poisson?

Outline

Towards a general result

- The Erdős-Rényi example captures the contribution of the "dense" part of the graph, i.e. edges between high degree vertices.

Towards a general result

- The Erdős-Rényi example captures the contribution of the "dense" part of the graph, i.e. edges between high degree vertices.

- The disjoint star example captures the contribution of edges between high degree vertices and low degree vertices.

Towards a general result

- The Erdős-Rényi example captures the contribution of the "dense" part of the graph, i.e. edges between high degree vertices.
- The disjoint star example captures the contribution of edges between high degree vertices and low degree vertices.
- Finally, edges between low degree vertices gives rise to a Poisson limit.

Towards a general result

- The Erdős-Rényi example captures the contribution of the "dense" part of the graph, i.e. edges between high degree vertices.
- The disjoint star example captures the contribution of edges between high degree vertices and low degree vertices.
- Finally, edges between low degree vertices gives rise to a Poisson limit.
- Using this philosophy, we partition the edge set into 3 types,

$$\text{High} \leftrightarrow \text{High}, \quad \text{High} \leftrightarrow \text{Low}, \quad \text{Low} \leftrightarrow \text{Low}.$$

Adjacency matrix $\mapsto$ function on positive reals

- Define a function $W_{G_n}(.,.) : [0,\infty)^2 \mapsto [0,1]$ by setting

$$\begin{aligned} W_{G_n}(x,y) =& 1 \text{ if } (\lceil xc_n \rceil, \lceil yc_n \rceil) \in E(G_n) \\ =& 0 \text{ otherwise} \end{aligned}$$

Adjacency matrix $\mapsto$ function on positive reals

- Define a function $W_{G_n}(.,.) : [0,\infty)^2 \mapsto [0,1]$ by setting

$$\begin{aligned} W_{G_n}(x,y) =& 1 \text{ if } (\lceil xc_n \rceil, \lceil yc_n \rceil) \in E(G_n) \\ =& 0 \text{ otherwise} \end{aligned}$$

- By this definition, $W_{G_n}(x,y) = 0$ outside $[0, np_n]^2$.

Adjacency matrix $\mapsto$ function on positive reals

- Define a function $W_{G_n}(.,,) : [0,\infty)^2 \mapsto [0,1]$ by setting

$$\begin{aligned} W_{G_n}(x,y) =& 1 \text{ if } (\lceil xc_n \rceil, \lceil yc_n \rceil) \in E(G_n) \\ =& 0 \text{ otherwise} \end{aligned}$$

- By this definition, $W_{G_n}(x,y) = 0$ outside $[0, np_n]^2$.

- Also,

$$\int_{[0,\infty)^2} W_{G_n}(x,y) \mathrm{d}x \mathrm{d}y = \frac{2|E(G_n)|}{c_n^2} = O(1).$$

Adjacency matrix $\mapsto$ function on positive reals

- Define a function $W_{G_n}(.,.) : [0,\infty)^2 \mapsto [0,1]$ by setting

$$\begin{aligned} W_{G_n}(x,y) =& 1 \text{ if } (\lceil xc_n \rceil, \lceil yc_n \rceil) \in E(G_n) \\ =& 0 \text{ otherwise} \end{aligned}$$

- By this definition, $W_{G_n}(x,y) = 0$ outside $[0, np_n]^2$.

- Also,

$$\int_{[0,\infty)^2} W_{G_n}(x,y) \mathrm{d}x \mathrm{d}y = \frac{2|E(G_n)|}{c_n^2} = O(1).$$

- Let $d_{G_n}(x) := \int_0^\infty W_{G_n}(x,y) \mathrm{d}x$ be the degree function.

First assumption (A1)

- Given two bounded measurable functions f, g from $[0, K]^2 \mapsto [0, 1]$, define the strong cut distance between f and g by

$$\sup_{A,B \subset [0,1]} \left| \int_{A \times B} f(x,y) \mathrm{d}x \mathrm{d}y - \int_{A \times B} g(x,y) \mathrm{d}x \mathrm{d}y \right|.$$

First assumption (A1)

- Given two bounded measurable functions f, g from $[0,K]^2 \mapsto [0,1]$, define the strong cut distance between f and g by

$$\sup_{A,B\subset[0,1]} \left| \int_{A\times B} f(x,y)\mathrm{d}x\mathrm{d}y - \int_{A\times B} g(x,y)\mathrm{d}x\mathrm{d}y \right|.$$

- Assumption (A1):

  There is a function $W : [0,\infty)^2 \mapsto [0,1]$, such that for every K fixed, the function W_{G_n} converges in strong cut distance to the function W on $[0,K]^2$.

First assumption (A1)

- Given two bounded measurable functions f, g from $[0,K]^2 \mapsto [0,1]$, define the strong cut distance between f and g by

$$\sup_{A,B\subset[0,1]} \left| \int_{A\times B} f(x,y)\mathrm{d}x\mathrm{d}y - \int_{A\times B} g(x,y)\mathrm{d}x\mathrm{d}y \right|.$$

- Assumption (A1):

  There is a function $W : [0,\infty)^2 \mapsto [0,1]$, such that for every K fixed, the function W_{G_n} converges in strong cut distance to the function W on $[0,K]^2$.

- In some sense W captures the limit of the dense part of the graph.

First assumption (A1)

- Given two bounded measurable functions f, g from $[0,K]^2 \mapsto [0,1]$, define the strong cut distance between f and g by

$$\sup_{A,B\subset[0,1]} \left| \int_{A\times B} f(x,y)\mathrm{d}x\mathrm{d}y - \int_{A\times B} g(x,y)\mathrm{d}x\mathrm{d}y \right|.$$

- Assumption (A1):

  There is a function $W : [0,\infty)^2 \mapsto [0,1]$, such that for every K fixed, the function W_{G_n} converges in strong cut distance to the function W on $[0,K]^2$.

- In some sense W captures the limit of the dense part of the graph. Note that this assumption, along with Fatou's lemma automatically implies

$$\int_{[0,\infty)^2} |W(x,y)|\mathrm{d}x\mathrm{d}y < \infty.$$

Second assumption (A2)

- Assumption (A2):

  There is a function $d : [0,\infty) \mapsto [0,\infty)$ such that for every $M, K > 0$ we have

$$\min(d_{W_{G_n}}, M) \stackrel{L_1[0,K]}{\to} \min(d(.), M).$$

Second assumption (A2)

- Assumption (A2):

  There is a function $d : [0, \infty) \mapsto [0, \infty)$ such that for every $M, K > 0$ we have

  $$\min(d_{W_{G_n}}, M) \overset{L_1[0,K]}{\to} \min(d(.), M).$$

- Again Fatou's Lemma gives $\int_{[0,\infty)} |d(x)| \mathrm{d}x < \infty$.

Second assumption (A2)

- Assumption (A2):

  There is a function $d : [0, \infty) \mapsto [0, \infty)$ such that for every $M, K > 0$ we have

$$\min(d_{W_{G_n}}, M) \stackrel{L_1[0,K]}{\rightarrow} \min(d(.), M).$$

- Again Fatou's Lemma gives $\int_{[0,\infty)} |d(x)| \mathrm{d}x < \infty$.

- Also, one has $\int_0^\infty W(x, y) \mathrm{d}y \leq d(x)$.

Second assumption (A2)

- Assumption (A2):

  There is a function $d : [0, \infty) \mapsto [0, \infty)$ such that for every $M, K > 0$ we have

$$\min(d_{W_{G_n}}, M) \overset{L_1[0,K]}{\to} \min(d(.), M).$$

- Again Fatou's Lemma gives $\int_{[0,\infty)} |d(x)| \mathrm{d}x < \infty$.

- Also, one has $\int_0^\infty W(x, y) \mathrm{d}y \leq d(x)$. Let

$$\Delta(x) := d(x) - \int_0^\infty W(x, y) \mathrm{d}y.$$

Second assumption (A2)

- Assumption (A2):

  There is a function $d : [0, \infty) \mapsto [0, \infty)$ such that for every $M, K > 0$ we have

$$\min(d_{W_{G_n}}, M) \stackrel{L_1[0,K]}{\to} \min(d(.), M).$$

- Again Fatou's Lemma gives $\int_{[0,\infty)} |d(x)| \mathrm{d}x < \infty$.

- Also, one has $\int_0^\infty W(x,y)\mathrm{d}y \leq d(x)$. Let

$$\Delta(x) := d(x) - \int_0^\infty W(x,y)\mathrm{d}y.$$

- In some sense $\Delta(x)$ counts the edges between the high and low degree vertices.

Third assumption (A3)+Theorem I

- Assumption: (A3)

$$\lim_{K\to\infty}\lim_{n\to\infty}\int_{[K,\infty)^2} W_{G_n}(x,y)\mathrm{d}x\mathrm{d}y = 2\lambda_0.$$

Third assumption (A3)+Theorem I

- Assumption: (A3)

$$\lim_{K\to\infty}\lim_{n\to\infty}\int_{[K,\infty)^2} W_{G_n}(x,y)\mathrm{d}x\mathrm{d}y = 2\lambda_0.$$

- This assumption estimates the edges between the low degree vertices.

Third assumption (A3)+Theorem I

- Assumption: (A3)

$$\lim_{K\to\infty}\lim_{n\to\infty}\int_{[K,\infty)^2} W_{G_n}(x,y)\mathrm{d}x\mathrm{d}y = 2\lambda_0.$$

- This assumption estimates the edges between the low degree vertices.

Theorem (Bhattacharya-Mukherjee-M., AAP 20)

If (A1), (A2), (A3) hold, then

$$T_n \xrightarrow{D} Q_1 + Q_2 + Q_3.$$

Third assumption (A3)+Theorem I

- Assumption: (A3)

$$\lim_{K\to\infty}\lim_{n\to\infty}\int_{[K,\infty)^2} W_{G_n}(x,y)\mathrm{d}x\mathrm{d}y = 2\lambda_0.$$

- This assumption estimates the edges between the low degree vertices.

Theorem (Bhattacharya-Mukherjee-M., AAP 20)

If (A1), (A2), (A3) hold, then

$$T_n \xrightarrow{D} Q_1 + Q_2 + Q_3.$$

- *Here* $Q_3 \sim Pois(\lambda_0)$.

Third assumption (A3)+Theorem I

- Assumption: (A3)

$$\lim_{K\to\infty}\lim_{n\to\infty}\int_{[K,\infty)^2} W_{G_n}(x,y)\mathrm{d}x\mathrm{d}y = 2\lambda_0.$$

- This assumption estimates the edges between the low degree vertices.

Theorem (Bhattacharya-Mukherjee-M., AAP 20)

If (A1), (A2), (A3) hold, then

$$T_n \xrightarrow{D} Q_1 + Q_2 + Q_3.$$

- *Here* $Q_3 \sim Pois(\lambda_0)$.
- (Q_1, Q_2) *is jointly independent of* Q_3.

Third assumption (A3)+Theorem I

- Assumption: (A3)

$$\lim_{K\to\infty}\lim_{n\to\infty}\int_{[K,\infty)^2} W_{G_n}(x,y)\mathrm{d}x\mathrm{d}y = 2\lambda_0.$$

- This assumption estimates the edges between the low degree vertices.

Theorem (Bhattacharya-Mukherjee-M., AAP 20)

If (A1), (A2), (A3) hold, then

$$T_n \xrightarrow{D} Q_1 + Q_2 + Q_3.$$

- *Here $Q_3 \sim Pois(\lambda_0)$.*
- *(Q_1, Q_2) is jointly independent of Q_3.*
- *The joint Mgf of (Q_1, Q_2) appears on the next slide.*

Mgf of (Q_1, Q_2)

- For $t_1, t_2 > 0$, $\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2}$ equals

$$\mathbb{E}\exp\Big\{\frac{1}{2}\int_{[0,\infty)^2}\phi_{W,t_1}(x,y)dN(x)dN(y)-(1-e^{-t_2})\int_{[0,\infty)}\Delta(x)dN(x)\Big\}.$$

Mgf of (Q_1, Q_2)

- For $t_1, t_2 > 0$, $\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2}$ equals

$$\mathbb{E}\exp\left\{\frac{1}{2}\int_{[0,\infty)^2}\phi_{W,t_1}(x,y)dN(x)dN(y) - (1-e^{-t_2})\int_{[0,\infty)}\Delta(x)dN(x)\right\}.$$

- Here
  - $$\phi_{W,t_1}(x,y) = \log\left(1 - W(x,y) + W(x,y)e^{-t_1}\right)$$

    is integrable,

Mgf of (Q_1, Q_2)

- For $t_1, t_2 > 0$, $\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2}$ equals

$$\mathbb{E}\exp\Big\{\frac{1}{2}\int_{[0,\infty)^2}\phi_{W,t_1}(x,y)dN(x)dN(y) - (1-e^{-t_2})\int_{[0,\infty)}\Delta(x)dN(x)\Big\}.$$

- Here
  - $$\phi_{W,t_1}(x,y) = \log\Big(1 - W(x,y) + W(x,y)e^{-t_1}\Big)$$
    is integrable,
  - $\{N(t)\}_{t\geq 1}$ is a homogenous Poisson process of rate 1.

Mgf of (Q_1, Q_2)

- For $t_1, t_2 > 0$, $\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2}$ equals

$$\mathbb{E}\exp\left\{\frac{1}{2}\int_{[0,\infty)^2}\phi_{W,t_1}(x,y)dN(x)dN(y) - (1-e^{-t_2})\int_{[0,\infty)}\Delta(x)dN(x)\right\}.$$

- Here
  - $$\phi_{W,t_1}(x,y) = \log\left(1 - W(x,y) + W(x,y)e^{-t_1}\right)$$

    is integrable,
  - $\{N(t)\}_{t\geq 1}$ is a homogenous Poisson process of rate 1.

- $Q_2 \stackrel{D}{=} Pois(\int_{[0,\infty)}\Delta(x)dN(x))$.

Mgf of (Q_1, Q_2)

- For $t_1, t_2 > 0$, $\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2}$ equals

$$\mathbb{E}\exp\left\{\frac{1}{2}\int_{[0,\infty)^2}\phi_{W,t_1}(x,y)dN(x)dN(y)-(1-e^{-t_2})\int_{[0,\infty)}\Delta(x)dN(x)\right\}.$$

- Here
  - $$\phi_{W,t_1}(x,y) = \log\Big(1 - W(x,y) + W(x,y)e^{-t_1}\Big)$$

    is integrable,
  - $\{N(t)\}_{t\geq 1}$ is a homogenous Poisson process of rate 1.
- $Q_2 \overset{D}{=} Pois(\int_{[0,\infty)}\Delta(x)dN(x))$. This arises from the edges between the high and low degree.

Mgf of (Q_1, Q_2)

- For $t_1, t_2 > 0$, $\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2}$ equals

$$\mathbb{E}\exp\left\{\frac{1}{2}\int_{[0,\infty)^2}\phi_{W,t_1}(x,y)dN(x)dN(y) - (1-e^{-t_2})\int_{[0,\infty)}\Delta(x)dN(x)\right\}.$$

- Here
  - $$\phi_{W,t_1}(x,y) = \log\left(1 - W(x,y) + W(x,y)e^{-t_1}\right)$$
    is integrable,
  - $\{N(t)\}_{t\geq 1}$ is a homogenous Poisson process of rate 1.

- $Q_2 \stackrel{D}{=} Pois(\int_{[0,\infty)}\Delta(x)dN(x))$. This arises from the edges between the high and low degree.

- Q_1 arises from the edges between the high degree vertices, i.e. the dense part of the graph.

Stochastic Integral (Itô)

- Suppose $f : [0,\infty)^2 \mapsto \mathbb{R}$ is such that

$$\int_{[0,\infty)^2} |f(x,y)| \mathrm{d}x \mathrm{d}y < \infty.$$

Stochastic Integral (Itô)

- Suppose $f : [0,\infty)^2 \mapsto \mathbb{R}$ is such that
$$\int_{[0,\infty)^2} |f(x,y)| \mathrm{d}x \mathrm{d}y < \infty.$$
- Let $\{N(t)\}_{t \geq 1}$ is a homogenous Poisson process of rate 1.

Stochastic Integral (Itô)

- Suppose $f : [0, \infty)^2 \mapsto \mathbb{R}$ is such that

$$\int_{[0,\infty)^2} |f(x, y)| \mathrm{d}x\mathrm{d}y < \infty.$$

- Let $\{N(t)\}_{t \geq 1}$ is a homogenous Poisson process of rate 1.

- We want to define the bivariate stochastic integral

$$I_2(f) := \int_{[0,\infty)^2} f(x, y) dN(x)N(y).$$

Stochastic Integral (Itô)

- Suppose $f : [0,\infty)^2 \mapsto \mathbb{R}$ is such that
$$\int_{[0,\infty)^2} |f(x,y)| \mathrm{d}x \mathrm{d}y < \infty.$$
- Let $\{N(t)\}_{t \geq 1}$ is a homogenous Poisson process of rate 1.
- We want to define the bivariate stochastic integral
$$I_2(f) := \int_{[0,\infty)^2} f(x,y) dN(x) N(y).$$
- First assume
$$f = \sum_{i,j=1}^{k} c_{ij} 1_{A_i \times A_j},$$
where $\{A_1, \ldots, A_k\}$ is a measurable partition of $[0,1]$, and $c_{ii} = 0$ for $1 \leq i \leq k$.

Stochastic Integral (Itô)

- Suppose $f : [0, \infty)^2 \mapsto \mathbb{R}$ is such that

$$\int_{[0,\infty)^2} |f(x,y)| \mathrm{d}x \mathrm{d}y < \infty.$$

- Let $\{N(t)\}_{t \geq 1}$ is a homogenous Poisson process of rate 1.
- We want to define the bivariate stochastic integral

$$I_2(f) := \int_{[0,\infty)^2} f(x,y) dN(x) N(y).$$

- First assume

$$f = \sum_{i,j=1}^{k} c_{ij} 1_{A_i \times A_j},$$

where $\{A_1, \ldots, A_k\}$ is a measurable partition of $[0,1]$, and $c_{ii} = 0$ for $1 \leq i \leq k$.

- Define

$$I_2(f) := \sum_{i,j=1}^{k} c_{ij} N(A_i) N(A_j).$$

Stochastic Integral (Itô)

- For a general $f \in L^1[0,\infty)^2$, approximate f by a sequence $\{f_k\}$ of simple functions from the last slide in $L^1[0,\infty)^2$.

Stochastic Integral (Itô)

- For a general $f \in L^1[0,\infty)^2$, approximate f by a sequence $\{f_k\}$ of simple functions from the last slide in $L^1[0,\infty)^2$.
- Then $I_2(f_k)$ is Cauchy in $L^1(\Omega)$, where Ω contains the Poisson process.

Stochastic Integral (Itô)

- For a general $f \in L^1[0,\infty)^2$, approximate f by a sequence $\{f_k\}$ of simple functions from the last slide in $L^1[0,\infty)^2$.

- Then $I_2(f_k)$ is Cauchy in $L^1(\Omega)$, where Ω contains the Poisson process.

- Define $I_2(f) = \lim_{k\to\infty} I_2(f_k)$, where the limit is in $L_1(\Omega)$.

Stochastic Integral (Itô)

- For a general $f \in L^1[0,\infty)^2$, approximate f by a sequence $\{f_k\}$ of simple functions from the last slide in $L^1[0,\infty)^2$.

- Then $I_2(f_k)$ is Cauchy in $L^1(\Omega)$, where Ω contains the Poisson process.

- Define $I_2(f) = \lim_{k\to\infty} I_2(f_k)$, where the limit is in $L_1(\Omega)$.

- Then $I_2(f)$ is well defined, and

$$\mathbb{E}|I_2(f)| \le \|f\|_{L_1[0,\infty)^2}.$$

Example I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter q, and $p_n = \frac{1}{n}$.

Example I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter q, and $p_n = \frac{1}{n}$.
- Then the conditions of our theorem hold with

$$W(x,y) = q1\{[x,y] \in [0,1]^2\}, \quad \Delta(x) = 0, \quad \lambda_0 = 0.$$

Example I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter q, and $p_n = \frac{1}{n}$.
- Then the conditions of our theorem hold with

$$W(x,y) = q1\{[x,y] \in [0,1]^2\}, \quad \Delta(x) = 0, \quad \lambda_0 = 0.$$

- Thus our theorem gives $T_n \xrightarrow{D} Q_1$, where

$$\mathbb{E}e^{-t_1 Q_1} = \mathbb{E}\exp\left\{\frac{1}{2}\int_{[0,\infty)^2} \phi_{q,t_1}(x,y)dN(x)dN(y)\right\}.$$

Example I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter q, and $p_n = \frac{1}{n}$.

- Then the conditions of our theorem hold with

$$W(x,y) = q1\{[x,y] \in [0,1]^2\}, \quad \Delta(x) = 0, \quad \lambda_0 = 0.$$

- Thus our theorem gives $T_n \xrightarrow{D} Q_1$, where

$$\mathbb{E}e^{-t_1 Q_1} = \mathbb{E}\exp\left\{\frac{1}{2}\int_{[0,\infty)^2} \phi_{q,t_1}(x,y)dN(x)dN(y)\right\}.$$

- Here

$$\phi_{q,t_1}(x,y) = \log(1 - q + qe^{-t_1}) \text{ if } (x,y) \in [0,1]^2.$$

Example I (Erdős-Rényi)

- Thus we need to compute

$$\frac{1}{2}\log(1-q+qe^{-t_1})\int_{[0,1]^2} dN(x)dN(y).$$

Example I (Erdős-Rényi)

- Thus we need to compute

$$\frac{1}{2}\log(1-q+qe^{-t_1})\int_{[0,1]^2} dN(x)dN(y).$$

- The inside integral equals $N(1)^2 - N(1)$.

Example I (Erdős-Rényi)

- Thus we need to compute

$$\frac{1}{2}\log(1-q+qe^{-t_1})\int_{[0,1]^2} dN(x)dN(y).$$

- The inside integral equals $N(1)^2 - N(1)$.

- Combining we have

$$\mathbb{E}e^{-tQ_1} = \mathbb{E}\exp\left\{\binom{N(1)}{2}\log(1-q+qe^{-t_1})\right\}.$$

Example I (Erdős-Rényi)

- Thus we need to compute

$$\frac{1}{2}\log(1-q+qe^{-t_1})\int_{[0,1]^2} dN(x)dN(y).$$

- The inside integral equals $N(1)^2 - N(1)$.
- Combining we have

$$\mathbb{E}e^{-tQ_1} = \mathbb{E}\exp\left\{\binom{N(1)}{2}\log(1-q+qe^{-t_1})\right\}.$$

- This is the Mgf of a $Bin\left(\binom{N(1)}{2}, q\right)$.

Example I (Erdős-Rényi)

- Thus we need to compute

$$\frac{1}{2}\log(1-q+qe^{-t_1})\int_{[0,1]^2} dN(x)dN(y).$$

- The inside integral equals $N(1)^2 - N(1)$.

- Combining we have

$$\mathbb{E}e^{-tQ_1} = \mathbb{E}\exp\left\{\binom{N(1)}{2}\log(1-q+qe^{-t_1})\right\}.$$

- This is the Mgf of a $Bin\left(\binom{N(1)}{2}, q\right)$.

- Since $N(1) \sim Pois(1)$, we have the same limit distribution as before.

Example I′: Stochastic Block Model

- Suppose $p_n = \frac{1}{n}$, and G_n is a sequence of random graphs such that

$$\begin{aligned}\mathbb{P}(G_n(i,j)=1) =& a_{11} \text{ if } i,j < \frac{n}{2}\\ =& a_{12} \text{ if } i < \frac{n}{2}, j \geq \frac{n}{2} \text{ or } i \geq \frac{n}{2}, j < \frac{n}{2},\\ =& a_{22} \text{ if } i,j \geq \frac{n}{2}.\end{aligned}$$

Example I′: Stochastic Block Model

- Suppose $p_n = \frac{1}{n}$, and G_n is a sequence of random graphs such that

$$\begin{aligned}\mathbb{P}(G_n(i,j)=1) =& a_{11} \text{ if } i,j < \frac{n}{2}\\ =& a_{12} \text{ if } i < \frac{n}{2}, j \geq \frac{n}{2} \text{ or } i \geq \frac{n}{2}, j < \frac{n}{2},\\ =& a_{22} \text{ if } i,j \geq \frac{n}{2}.\end{aligned}$$

- In this case the conditions of our theorem hold with $\Delta \equiv 0, \lambda_0 = 0$, and W on $[0,1]^2$ given by

$$\begin{aligned}W(x,y) =& a_{11} \text{ if } x,y < \frac{1}{2},\\ =& a_{12} \text{ if } x < \frac{1}{2}, y \geq \frac{1}{2} \text{ or } x \geq \frac{1}{2}, y < \frac{1}{2},\\ =& a_{22} \text{ if } x,y \geq \frac{1}{2}.\end{aligned}$$

Example I′: Stochastic Block Model

- Our theorem gives $T_n \xrightarrow{D} Q_1$, where $\mathbb{E}e^{-t_1 Q_1}$ is in terms of a stochastic integral with respect to a block function.

Example I′: Stochastic Block Model

- Our theorem gives $T_n \stackrel{D}{\to} Q_1$, where $\mathbb{E}e^{-t_1 Q_1}$ is in terms of a stochastic integral with respect to a block function.

- Analyzing this stochastic integral gives

$$Q_1 \stackrel{D}{=} \frac{1}{2} Bin\left(\binom{S_1}{2}, a_{11}\right) + \frac{1}{2} Bin\left(\binom{S_2}{2}, a_{22}\right) + Bin\left(S_1 S_2, a_{12}\right).$$

Example I′: Stochastic Block Model

- Our theorem gives $T_n \stackrel{D}{\to} Q_1$, where $\mathbb{E}e^{-t_1 Q_1}$ is in terms of a stochastic integral with respect to a block function.

- Analyzing this stochastic integral gives

$$Q_1 \stackrel{D}{=} \frac{1}{2} Bin\left(\binom{S_1}{2}, a_{11}\right) + \frac{1}{2} Bin\left(\binom{S_2}{2}, a_{22}\right) + Bin\left(S_1 S_2, a_{12}\right).$$

- Here $S_1 = N(1/2)$ and $S_2 = N(1) - N(1/2)$ are iid Poissons with mean $1/2$.

Example I′: Stochastic Block Model

- Our theorem gives $T_n \stackrel{D}{\to} Q_1$, where $\mathbb{E}e^{-t_1 Q_1}$ is in terms of a stochastic integral with respect to a block function.

- Analyzing this stochastic integral gives

$$Q_1 \stackrel{D}{=} \frac{1}{2}Bin\left(\binom{S_1}{2}, a_{11}\right) + \frac{1}{2}Bin\left(\binom{S_2}{2}, a_{22}\right) + Bin\left(S_1 S_2, a_{12}\right).$$

- Here $S_1 = N(1/2)$ and $S_2 = N(1) - N(1/2)$ are iid Poissons with mean 1/2.

- Similar results apply to unequal blocks, or more than 2 blocks.

Example-II (Disjoint union of stars)

- Suppose G_n is a disjoint union of $\sqrt{n}$ many $K_{1,\sqrt{n}}$, and $p_n = \frac{1}{\sqrt{n}}$.

Example-II (Disjoint union of stars)

- Suppose G_n is a disjoint union of $\sqrt{n}$ many $K_{1,\sqrt{n}}$, and $p_n = \frac{1}{\sqrt{n}}$.
- This satisfies the conditions of our theorem with

$$W \equiv 0, \quad \Delta(x) = 1\{x \in [0,1]\}, \quad \lambda_0 = 0.$$

Example-II (Disjoint union of stars)

- Suppose G_n is a disjoint union of $\sqrt{n}$ many $K_{1,\sqrt{n}}$, and $p_n = \frac{1}{\sqrt{n}}$.

- This satisfies the conditions of our theorem with

$$W \equiv 0, \quad \Delta(x) = 1\{x \in [0,1]\}, \quad \lambda_0 = 0.$$

- Using our theorem, it follows that

$$T_n \xrightarrow{D} Pois\Big(\int_{[0,\infty)} \Delta(x) dN(x) \Big)$$

Example-II (Disjoint union of stars)

- Suppose G_n is a disjoint union of $\sqrt{n}$ many $K_{1,\sqrt{n}}$, and $p_n = \frac{1}{\sqrt{n}}$.
- This satisfies the conditions of our theorem with

$$W \equiv 0, \quad \Delta(x) = 1\{x \in [0,1]\}, \quad \lambda_0 = 0.$$

- Using our theorem, it follows that

$$T_n \stackrel{D}{\to} Pois\Big(\int_{[0,\infty)} \Delta(x) dN(x)\Big) = Pois(N(1)).$$

Example-II (Disjoint union of stars)

- Suppose G_n is a disjoint union of $\sqrt{n}$ many $K_{1,\sqrt{n}}$, and $p_n = \frac{1}{\sqrt{n}}$.
- This satisfies the conditions of our theorem with

$$W \equiv 0, \quad \Delta(x) = 1\{x \in [0,1]\}, \quad \lambda_0 = 0.$$

- Using our theorem, it follows that

$$T_n \stackrel{D}{\to} Pois\Big(\int_{[0,\infty)} \Delta(x) dN(x)\Big) = Pois(N(1)).$$

- Since $N(1) \sim Pois(1)$, we have the same limit distribution as before.

Example-III (Co-Existence)

- In all examples so far, either W or Δ is 0, and so either Q_1 or Q_2 is 0.

Example-III (Co-Existence)

- In all examples so far, either W or Δ is 0, and so either Q_1 or Q_2 is 0.
- We now construct an example where the terms (Q_1, Q_2) both survive in the limit.

Example-III (Co-Existence)

- In all examples so far, either W or Δ is 0, and so either Q_1 or Q_2 is 0.
- We now construct an example where the terms (Q_1, Q_2) both survive in the limit.
- Let K_n be the complete graph on n vertices.

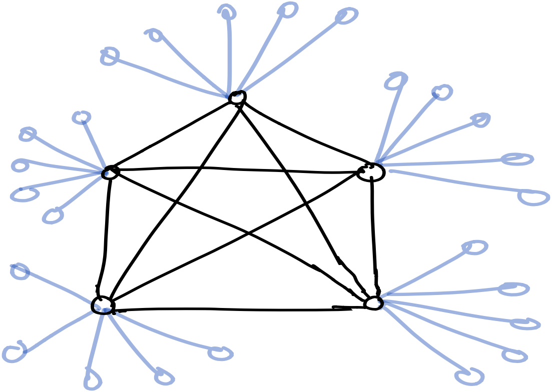
Example-III (Co-Existence)

- In all examples so far, either W or Δ is 0, and so either Q_1 or Q_2 is 0.
- We now construct an example where the terms (Q_1, Q_2) both survive in the limit.
- Let K_n be the complete graph on n vertices.
- Put a star graph $K_{1,n}$ centered on every vertex of K_n.

Example-III (Co-Existence)

- In all examples so far, either W or Δ is 0, and so either Q_1 or Q_2 is 0.
- We now construct an example where the terms (Q_1, Q_2) both survive in the limit.
- Let K_n be the complete graph on n vertices.
- Put a star graph $K_{1,n}$ centered on every vertex of K_n.
- Then the entire graph G_n has $n + n^2 \sim n^2$ vertices, and $\binom{n}{2} + n^2 \sim \frac{3n^2}{2}$ edges.

Co-existence example



EG: $n = 5$

TOTAL VERTICES $= 5 + 5 \times 5 = 30$

TOTAL EDGES $= \binom{5}{2} + 5 \times 5 = 35$

Example-III (Co-Existence)

- In this case, with $p_n = \frac{1}{n}$ the conditions of our theorem hold with

$$W(x,y) = 1\{(x,y) \in [0,1]^2\}, \quad \Delta(x) = 1\{x \in [0,1]\}, \quad \lambda_0 = 0.$$

Example-III (Co-Existence)

- In this case, with $p_n = \frac{1}{n}$ the conditions of our theorem hold with

$$W(x,y) = 1\{(x,y) \in [0,1]^2\}, \quad \Delta(x) = 1\{x \in [0,1]\}, \quad \lambda_0 = 0.$$

- Applying our theorem gives $T_n \xrightarrow{D} Q_1 + Q_2$, where

$$\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2} = \mathbb{E}\exp\Big\{-t_1\binom{S}{2} - (1-e^{-t_2})S\Big\}.$$

Example-III (Co-Existence)

- In this case, with $p_n = \frac{1}{n}$ the conditions of our theorem hold with

$$W(x,y) = 1\{(x,y) \in [0,1]^2\}, \quad \Delta(x) = 1\{x \in [0,1]\}, \quad \lambda_0 = 0.$$

- Applying our theorem gives $T_n \stackrel{D}{\to} Q_1 + Q_2$, where

$$\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2} = \mathbb{E}\exp\Big\{ - t_1 \binom{S}{2} - (1 - e^{-t_2})S \Big\}.$$

- Here $S = N(1) \sim Pois(1)$.

Example-III (Co-Existence)

- In this case, with $p_n = \frac{1}{n}$ the conditions of our theorem hold with

$$W(x,y) = 1\{(x,y) \in [0,1]^2\}, \quad \Delta(x) = 1\{x \in [0,1]\}, \quad \lambda_0 = 0.$$

- Applying our theorem gives $T_n \stackrel{D}{\to} Q_1 + Q_2$, where

$$\mathbb{E}e^{-t_1 Q_1 - t_2 Q_2} = \mathbb{E}\exp\Big\{ - t_1 \binom{S}{2} - (1 - e^{-t_2})S \Big\}.$$

- Here $S = N(1) \sim Pois(1)$.

- This gives

$$Q_1 + Q_2 \stackrel{D}{\to} \binom{S}{2} + Pois(S), \text{ where } S \sim Pois(1).$$

Example-IV (Non-compact support)

- In all examples so far, the functions W and Δ have compact support.

Example-IV (Non-compact support)

- In all examples so far, the functions W and Δ have compact support.
- We will now construct an example where support of both W and Δ are non-compact.

Example-IV (Non-compact support)

- In all examples so far, the functions W and Δ have compact support.
- We will now construct an example where support of both W and Δ are non-compact.
- Set $a_k = \frac{1}{16^k}, b_k = 4^k$, and $c_k = \frac{1}{32^k}$ for $k \geq 1$.

Example-IV (Non-compact support)

- In all examples so far, the functions W and Δ have compact support.
- We will now construct an example where support of both W and Δ are non-compact.
- Set $a_k = \frac{1}{16^k}, b_k = 4^k$, and $c_k = \frac{1}{32^k}$ for $k \geq 1$.
- Fix $k \in \{1, 2, \ldots, \log_4 n\}$, and draw nb_k many na_k stars (i.e K_{1,na_k}).

Example-IV (Non-compact support)

- In all examples so far, the functions W and Δ have compact support.

- We will now construct an example where support of both W and Δ are non-compact.

- Set $a_k = \frac{1}{16^k}, b_k = 4^k$, and $c_k = \frac{1}{32^k}$ for $k \geq 1$.

- Fix $k \in \{1, 2, \ldots, \log_4 n\}$, and draw nb_k many na_k stars (i.e K_{1,na_k}).

- These nb_k many stars have nb_k many central vertices.

Example-IV (Non-compact support)

- Connect the edges between these nb_k many central vertices independently, with probability c_k.

Example-IV (Non-compact support)

- Connect the edges between these nb_k many central vertices independently, with probability c_k.
- Do this for every $k \in \{1, 2, \dots \log_4 n\}$.

Example-IV (Non-compact support)

- Connect the edges between these nb_k many central vertices independently, with probability c_k.
- Do this for every $k \in \{1, 2, \ldots \log_4 n\}$.
- This ensures that G_n has $\Theta(n^2)$ vertices, and $\Theta_P(n^2)$ many edges.

Example-IV (Non-compact support)

- Connect the edges between these nb_k many central vertices independently, with probability c_k.
- Do this for every $k \in \{1, 2, \dots \log_4 n\}$.
- This ensures that G_n has $\Theta(n^2)$ vertices, and $\Theta_P(n^2)$ many edges.
- For the choice $p_n = \frac{1}{n}$, our theorem applies with

$$\begin{aligned} W(x,y) =& c_k \text{ if } x, y \in (r_{k-1}, r_k] \text{ for some } k \geq 1, \\ =& 0 \text{ otherwise }. \end{aligned}$$

Example-IV (Non-compact support)

- Also,

$$\Delta(x) = a_k \text{ if } x \in (r_{k-1}, r_k] \text{ for some } k \geq 1.$$

Example-IV (Non-compact support)

- Also,

$$\Delta(x) = a_k \text{ if } x \in (r_{k-1}, r_k] \text{ for some } k \geq 1.$$

- Here $r_k := \sum_{i=0}^k b_i \to \infty$, and so both W, Δ have non-compact support.

Example-IV (Non-compact support)

- Also,

$$\Delta(x) = a_k \text{ if } x \in (r_{k-1}, r_k] \text{ for some } k \geq 1.$$

- Here $r_k := \sum_{i=0}^{k} b_i \to \infty$, and so both W, Δ have non-compact support.

- Using our theorem gives

$$T_n \xrightarrow{D} \sum_{k=1}^{\infty} Bin\left(\binom{S_k}{2}, c_k\right) + Pois\left(\sum_{k=1}^{\infty} a_k S_k\right),$$

where $S_k \sim Pois(b_k)$ are mutually independent.

Outline

Theorem II

- Our first result shows that under (A1), (A2), (A3), the limit of T_n can be expressed as $\psi(W, d, \lambda)$ for suitable W, d, λ.

Theorem II

- Our first result shows that under (A1), (A2), (A3), the limit of T_n can be expressed as $\psi(W, d, \lambda)$ for suitable W, d, λ.

Theorem (Bhattacharya-Mukherjee-M., AAP-2020)

Suppose $\mathbb{E}T_n = O(1)$. Then (under no other assumption) if T_n converges in distribution to a limit, then the limit must be in the closure of $\psi(W, d, \lambda)$.

Theorem II

- Our first result shows that under (A1), (A2), (A3), the limit of T_n can be expressed as $\psi(W, d, \lambda)$ for suitable W, d, λ.

Theorem (Bhattacharya-Mukherjee-M., AAP-2020)

Suppose $\mathbb{E}T_n = O(1)$. Then (under no other assumption) if T_n converges in distribution to a limit, then the limit must be in the closure of $\psi(W, d, \lambda)$.

- Essentially, after permuting the remaining vertices, the assumptions of the theorem does (approximately) hold along a subsequence.

Theorem II

- Our first result shows that under (A1), (A2), (A3), the limit of T_n can be expressed as $\psi(W, d, \lambda)$ for suitable W, d, λ.

Theorem (Bhattacharya-Mukherjee-M., AAP-2020)

Suppose $\mathbb{E}T_n = O(1)$. Then (under no other assumption) if T_n converges in distribution to a limit, then the limit must be in the closure of $\psi(W, d, \lambda)$.

- Essentially, after permuting the remaining vertices, the assumptions of the theorem does (approximately) hold along a subsequence.

- We claim that the class $\psi(W, d, \lambda)$ is closed under weak topology.

Theorem III

- So far we have not shown any examples where the limit is just a Poisson.

Theorem III

- So far we have not shown any examples where the limit is just a Poisson.

- To address this, let's first state a characterization result for Poisson convergence.

Theorem III

- So far we have not shown any examples where the limit is just a Poisson.
- To address this, let's first state a characterization result for Poisson convergence.

Theorem (Bhattacharya-Mukherjee-M., AAP-2020)

If $\mathbb{E}T_n = O(1)$, the following are equivalent:

(i) $T_n \xrightarrow{D} Pois(\lambda)$.

(ii)
$$\lim_{M\to\infty}\lim_{n\to\infty}\mathbb{E}(T_{n,M}) = \lambda, \quad \lim_{M\to\infty}\lim_{n\to\infty} Var(T_{n,M}) = \lambda.$$

Theorem III

- Here

$$T_{n,M} := \sum_{i<j} G_n(i,j) X_i X_j 1\{d_i \leq M c_n, d_j \leq M c_n\}$$

  is a truncated version of T_n.

Theorem III

- Here

$$T_{n,M} := \sum_{i<j} G_n(i,j) X_i X_j 1\{d_i \leq Mc_n, d_j \leq Mc_n\}$$

  is a truncated version of T_n.

- The last theorem can be viewed as a truncated second moment phenomenon.

Theorem III

- Here

$$T_{n,M} := \sum_{i<j} G_n(i,j) X_i X_j 1\{d_i \le Mc_n, d_j \le Mc_n\}$$

is a truncated version of T_n.

- The last theorem can be viewed as a truncated second moment phenomenon.

- This immediately implies the following simpler second moment phenomenon:

Corollary (Bhattacharya-Mukherjee-M., AAP-2020)

If $\mathbb{E}T_n \to \lambda$ *and* $Var(T_n) \to \lambda$, *then* $T_n \xrightarrow{D} Pois(\lambda)$.

Theorem III

- Here
$$T_{n,M} := \sum_{i<j} G_n(i,j) X_i X_j 1\{d_i \le Mc_n, d_j \le Mc_n\}$$
is a truncated version of T_n.

- The last theorem can be viewed as a truncated second moment phenomenon.

- This immediately implies the following simpler second moment phenomenon:

Corollary (Bhattacharya-Mukherjee-M., AAP-2020)

If $\mathbb{E}T_n \to \lambda$ *and* $Var(T_n) \to \lambda$, *then* $T_n \xrightarrow{D} Pois(\lambda)$.

- Compare this with the more well studied fourth moment phenomenon for the Gaussian distribution (see Ivan Nourdin's webpage for a list of papers on this topic).

Example-I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter $q_n \to 0$ such that $q_n \gg \frac{1}{n^2}$

Example-I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter $q_n \to 0$ such that $q_n \gg \frac{1}{n^2}$ (needed to ensure $|E(G_n)| \stackrel{P}{\to} \infty$).

Example-I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter $q_n \to 0$ such that $q_n \gg \frac{1}{n^2}$ (needed to ensure $|E(G_n)| \xrightarrow{P} \infty$).

- Then it has approximately $\binom{n}{2} q_n \sim \frac{n^2 q_n}{2}$ edges with high probability.

Example-I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter $q_n \to 0$ such that $q_n \gg \frac{1}{n^2}$ (needed to ensure $|E(G_n)| \stackrel{P}{\to} \infty$).

- Then it has approximately $\binom{n}{2} q_n \sim \frac{n^2 q_n}{2}$ edges with high probability.

- Let p_n be such that $p_n^2 \frac{n^2 q_n}{2} \to \lambda$.

Example-I (Erdős-Rényi)

- Suppose G_n is an Erdős-Rényi graph with parameter $q_n \to 0$ such that $q_n \gg \frac{1}{n^2}$ (needed to ensure $|E(G_n)| \stackrel{P}{\to} \infty$).
- Then it has approximately $\binom{n}{2} q_n \sim \frac{n^2 q_n}{2}$ edges with high probability.
- Let p_n be such that $p_n^2 \frac{n^2 q_n}{2} \to \lambda$.
- Then $T_n \stackrel{d}{\to} Pois(\lambda)$.

Example-I (Erdős-Rényi)

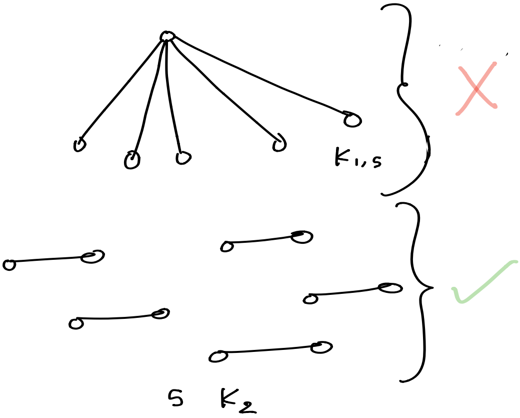
- Suppose G_n is an Erdős-Rényi graph with parameter $q_n \to 0$ such that $q_n \gg \frac{1}{n^2}$ (needed to ensure $|E(G_n)| \stackrel{P}{\to} \infty$).
- Then it has approximately $\binom{n}{2} q_n \sim \frac{n^2 q_n}{2}$ edges with high probability.
- Let p_n be such that $p_n^2 \frac{n^2 q_n}{2} \to \lambda$.
- Then $T_n \stackrel{d}{\to} Pois(\lambda)$.
- Similar results hold for sparse block models, and random regular graphs.

Example-II (Why Truncate?)

- Suppose G_n is a disjoint union of $K_{1,n}$, and n disjoint edges.

Example-II (Why Truncate?)

- Suppose G_n is a disjoint union of $K_{1,n}$, and n disjoint edges.



Example-II (Why Truncate?)

- Then G_n has $n + 1 + 2n \sim 3n$ vertices, and $n + n = 2n$ edges.

Example-II (Why Truncate?)

- Then G_n has $n + 1 + 2n \sim 3n$ vertices, and $n + n = 2n$ edges.
- Then with $p_n = \frac{1}{n}$, we have $\mathbb{E}T_n = p_n|E(G_n)| \to 2$.

Example-II (Why Truncate?)

- Then G_n has $n + 1 + 2n \sim 3n$ vertices, and $n + n = 2n$ edges.
- Then with $p_n = \frac{1}{n}$, we have $\mathbb{E}T_n = p_n|E(G_n)| \to 2$.
- However, it is not hard to check that $T_n \stackrel{D}{\to} Pois(1)$.

Example-II (Why Truncate?)

- Then G_n has $n + 1 + 2n \sim 3n$ vertices, and $n + n = 2n$ edges.
- Then with $p_n = \frac{1}{n}$, we have $\mathbb{E}T_n = p_n|E(G_n)| \to 2$.
- However, it is not hard to check that $T_n \xrightarrow{D} Pois(1)$.
- This is because with high probability, there are no monochromatic edges of color 1 in the star graph.

Example-II (Why Truncate?)

- Then G_n has $n+1+2n \sim 3n$ vertices, and $n+n=2n$ edges.
- Then with $p_n = \frac{1}{n}$, we have $\mathbb{E}T_n = p_n|E(G_n)| \to 2$.
- However, it is not hard to check that $T_n \stackrel{D}{\to} Pois(1)$.
- This is because with high probability, there are no monochromatic edges of color 1 in the star graph.
- The truncated second moment result captures this behavior automatically.

Outline

Proof idea of Theorem I

- With c_n the number of colors, we split the vertices into three groups,
  (i) degree less that εc_n (low degree vertices);

Proof idea of Theorem I

- With c_n the number of colors, we split the vertices into three groups,
  (i) degree less that εc_n (low degree vertices);

  (ii) degree between εc_n and Mc_n (high degree vertices);

Proof idea of Theorem I

- With c_n the number of colors, we split the vertices into three groups,
  (i) degree less that εc_n (low degree vertices);
  (ii) degree between εc_n and Mc_n (high degree vertices);
  (iii) degree greater than Mc_n (super high degree vertices).

Proof idea of Theorem I

- With c_n the number of colors, we split the vertices into three groups,
  (i) degree less that εc_n (low degree vertices);
  (ii) degree between εc_n and Mc_n (high degree vertices);
  (iii) degree greater than Mc_n (super high degree vertices).
- By using a first moment computation using Markov's inequality, we show that the super high degree vertices do not contribute for M large.

Proof idea of Theorem-I

- Thus there are three terms which contribute to the limit:

Proof idea of Theorem-I

- Thus there are three terms which contribute to the limit:

  (i) $T_{3,n}$, which counts monochromatic edges between low degree and low degree vertices.

Proof idea of Theorem-I

- Thus there are three terms which contribute to the limit:

  (i) $T_{3,n}$, which counts monochromatic edges between low degree and low degree vertices. This gives $Q_3 \sim Poisson$.

Proof idea of Theorem-I

- Thus there are three terms which contribute to the limit:

  (i) $T_{3,n}$, which counts monochromatic edges between low degree and low degree vertices. This gives $Q_3 \sim Poisson$.

  (ii) $T_{2,n}$, which counts monochromatic edges between low and high degree vertices.

Proof idea of Theorem-I

- Thus there are three terms which contribute to the limit:

  (i) $T_{3,n}$, which counts monochromatic edges between low degree and low degree vertices. This gives $Q_3 \sim Poisson$.

  (ii) $T_{2,n}$, which counts monochromatic edges between low and high degree vertices. This gives Q_2, the Poisson mixture.

Proof idea of Theorem-I

- Thus there are three terms which contribute to the limit:

  (i) $T_{3,n}$, which counts monochromatic edges between low degree and low degree vertices. This gives $Q_3 \sim Poisson$.

  (ii) $T_{2,n}$, which counts monochromatic edges between low and high degree vertices. This gives Q_2, the Poisson mixture.

  (iii) $T_{1,n}$, which counts monochromatic edges between high and high degree vertices.

Proof idea of Theorem-I

- Thus there are three terms which contribute to the limit:

  (i) $T_{3,n}$, which counts monochromatic edges between low degree and low degree vertices. This gives $Q_3 \sim Poisson$.

  (ii) $T_{2,n}$, which counts monochromatic edges between low and high degree vertices. This gives Q_2, the Poisson mixture.

  (iii) $T_{1,n}$, which counts monochromatic edges between high and high degree vertices. This gives Q_1, which in many examples is the sum of Binomials with quadratic Poisson parameters.

Proof idea of Theorem-I

- Thus there are three terms which contribute to the limit:

  (i) $T_{3,n}$, which counts monochromatic edges between low degree and low degree vertices. This gives $Q_3 \sim Poisson$.

  (ii) $T_{2,n}$, which counts monochromatic edges between low and high degree vertices. This gives Q_2, the Poisson mixture.

  (iii) $T_{1,n}$, which counts monochromatic edges between high and high degree vertices. This gives Q_1, which in many examples is the sum of Binomials with quadratic Poisson parameters.

- We argue using method of moments that $T_{3,n}$ is asymptotically independent from $T_{1,n}$ and $T_{2,n}$.

Proof idea of Theorem-I

- Using

  (i) (A1): strong cut metric convergence of W_{G_n} on $[0, K]^2$

  $+$

  (ii) (A2): the convergence of the degree function d_{G_n} in $L^1[0, K]$,

  we argue that the joint moments of $T_{1,n}$ and $T_{2,n}$ converge.

Proof idea of Theorem-I

- Using

  (i) (A1): strong cut metric convergence of W_{G_n} on $[0, K]^2$

  $+$

  (ii) (A2): the convergence of the degree function d_{G_n} in $L^1[0, K]$,

  we argue that the joint moments of $T_{1,n}$ and $T_{2,n}$ converge.

- We also show that the limiting moments determine their joint distribution.

Proof idea of Theorem-I

- Using

  (i) (A1): strong cut metric convergence of W_{G_n} on $[0, K]^2$

  $+$

  (ii) (A2): the convergence of the degree function d_{G_n} in $L^1[0, K]$,

  we argue that the joint moments of $T_{1,n}$ and $T_{2,n}$ converge.

- We also show that the limiting moments determine their joint distribution.

- To identify the distribution of (Q_1, Q_2), we compute the Mgf along a well chosen sequence of inhomogeneous random graphs.

Proof idea of Theorem-I

- Using

  (i) (A1): strong cut metric convergence of W_{G_n} on $[0, K]^2$

  $+$

  (ii) (A2): the convergence of the degree function d_{G_n} in $L^1[0, K]$,

  we argue that the joint moments of $T_{1,n}$ and $T_{2,n}$ converge.

- We also show that the limiting moments determine their joint distribution.

- To identify the distribution of (Q_1, Q_2), we compute the Mgf along a well chosen sequence of inhomogeneous random graphs.

- This gives the joint Mgf of (Q_1, Q_2), thereby proving Theorem I.

Outline

Summary of our results

- Motivated by graph coloring problems, we study asymptotic distribution of quadratic forms of Bernoulli random variables.

Summary of our results

- Motivated by graph coloring problems, we study asymptotic distribution of quadratic forms of Bernoulli random variables.
- We characterize the class of all possible limits of Bernoulli quadratic forms.

Summary of our results

- Motivated by graph coloring problems, we study asymptotic distribution of quadratic forms of Bernoulli random variables.
- We characterize the class of all possible limits of Bernoulli quadratic forms.
- As an application, we characterize exactly when is the limit a Poisson random variable.

Summary of our results

- Motivated by graph coloring problems, we study asymptotic distribution of quadratic forms of Bernoulli random variables.
- We characterize the class of all possible limits of Bernoulli quadratic forms.
- As an application, we characterize exactly when is the limit a Poisson random variable.
- We apply our theorem to several examples, which includes both deterministic and random graphs.

Future Scope

- Can one characterize the class of all possible limit distributions for the number of monochromatic triangles of color 1?

Future Scope

- Can one characterize the class of all possible limit distributions for the number of monochromatic triangles of color 1? What about other subgraphs?

Future Scope

- Can one characterize the class of all possible limit distributions for the number of monochromatic triangles of color 1? What about other subgraphs?

- Here we consider the case when the number of colors $c = c_n \to \infty$.

Future Scope

- Can one characterize the class of all possible limit distributions for the number of monochromatic triangles of color 1? What about other subgraphs?

- Here we consider the case when the number of colors $c = c_n \to \infty$. What can be said if c is fixed, and does not grow with n?

Future Scope

- Can one characterize the class of all possible limit distributions for the number of monochromatic triangles of color 1? What about other subgraphs?

- Here we consider the case when the number of colors $c = c_n \to \infty$. What can be said if c is fixed, and does not grow with n? Here we expect a similar result, but with Weiner process/Brownian motion replacing the Poisson process.

Future Scope

- Can one characterize the class of all possible limit distributions for the number of monochromatic triangles of color 1? What about other subgraphs?

- Here we consider the case when the number of colors $c = c_n \to \infty$. What can be said if c is fixed, and does not grow with n? Here we expect a similar result, but with Weiner process/Brownian motion replacing the Poisson process.

- Finally, our quadratic form is (in terms of) the adjacency matrix of a simple graph.

Future Scope

- Can one characterize the class of all possible limit distributions for the number of monochromatic triangles of color 1? What about other subgraphs?
- Here we consider the case when the number of colors $c = c_n \to \infty$. What can be said if c is fixed, and does not grow with n? Here we expect a similar result, but with Weiner process/Brownian motion replacing the Poisson process.
- Finally, our quadratic form is (in terms of) the adjacency matrix of a simple graph. Does a similar analysis apply for general quadratic forms?


The End