SURIM 2018 - Quantitative and Local Central Limit Theorems

1 General Background

1.1 Probability Theory

CLT:

$$|P(X_n < t) - P(Z < t)| \xrightarrow{n\to\infty} 0$$

LLT:

$$P(X_n = k) = \Phi(k) + o(1) \text{ where } \Phi(k) = \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-((k-\mu_n)/\sigma_n)^2/2\right)$$

Berry-Esseen Theorem: If the third moment is finite, given i.i.d. random variables, gives a quantitative bound for CLT

Levy's Continuity Theorem: If a sequence of characteristic functions converge pointwise to $\phi(t)$, X_n converges to random variable X, whose characteristic function is $\phi(t)$.

Stein's theorem (1986): If finite third and fourth moments, the dependency graph has finite degree $d < \infty$, we have a general CLT

1.2 Fourier Analysis

Fourier transform: physically, a complex valued function of frequency, magnitude represents amount of frequency present in original function

$$\hat{f}(\xi) := \int_{-\infty}^{+\infty} f(x)e^{-2\pi i x\xi}dx$$

Inverse transform:

$$f(x) := \int_{-\infty}^{+\infty} \hat{f}(\xi)e^{-2\pi i x\xi}d\xi$$

Dirac delta function: loosely, $\delta(x) = \begin{cases} +\infty & x = 0 \\ 0 & x \neq 0 \end{cases}$

Parseval's Theorem: Fourier transform is unitary (loosely, integral of square of function equals the integral of square of its transform)

1.3 Measure Theory

Dominated Convergence Theorem (used in Gilmer Kopparty small t): almost everywhere convergence of a sequence of functions implies convergence in the L^1 norm.

2 Papers

Gilmer and Kopparty (2014)

- Establishes that a local limit theorem for the number of triangles in a random graph exists
- More specifically, shows $\int_{-\pi\sigma_n}^{\pi\sigma_n} \left|\phi_n(t) - e^{-t^2/2}\right| dt = \text{o}(1)$ (tends to 0).

- Proof splits cases into small, intermediate, and big t. Small comes relatively easily from CLT but authors provide no bound on rate of convergence; intermediate: splits the $\binom{n}{2}$ edges up into those that are contained in a perfect matching M and those that are not; big t has a good t^{-50} bound.

Berkowitz (2017)

- Provides a quantitative bound on how far the distribution of the number of triangles in a random graph can vary from a discrete normal distribution
- Shows $\int_{-\pi\sigma_n}^{\pi\sigma_n} \left|\phi_n(t) - e^{-t^2/2}\right| dt = \mathrm{O}(n^{-1/2+\epsilon})$

Possibly related papers

2.0.1 Chatterjee and Diaconis (2017) A Central Limit Theorem for a New Statistic on Permutations

- Establishes a CLT for the random variable $T(\pi) := D(\pi) + D(\pi^{-1})$
- Gives 6 approaches for establishing CLT for descents $D(\pi)$.

2.0.2 Chatterjee (2014) A Short Survey of Stein's Method

- More ideas for possible random variables for which CLT's haven't been shown: Euclidean salesman problem (choose set of n random points from a unit square, let P be the shortest path that visits all points and ends with the same point, P should satisfy a CLT), minimum matching problem (suppose n is even, pair the points into $n/2$ pairs such that the sum S of the pairwise distances is minimized, S should satisfy a CLT), analogous problems with weights on edges

3 New Random Variables

3.1 Arithmetic Progressions

Let A_n be the number of arithmetic progressions in a random subset of F_n.

$A_n = \frac{1}{2}\sum_{k\in\mathbb{F}_n}\sum_{j\neq 0\in\mathbb{F}_n} X_k X_{k+j} X_{k+2j}$ for n prime and not 2 or 3.

Given a fixed $k \in \mathbb{F}_n$, there are $\frac{3}{2}(n-1)$ arithmetic progressions that contain k.

LLT Bound:

For small t: for $|t| << \sqrt{n}$,

$$\left|\phi_Z(t) - e^{-t^2/2}\right| \leq O\left(\frac{t^3 e^{-t^2/3}}{\sqrt{n}} + \frac{t}{n^{5/4}}\right)$$

.

3.2 Number of Descents in a Permutation

Let $D(\pi) = \#\{i : 1 \leq i \leq n-1, \pi(i) > \pi(i+1)\}$.

$$\mathbb{E}[D(\pi)] = \frac{n-1}{2}; Var(D(\pi)) = \frac{n+1}{12}$$

Since dependency graph for $D(\pi)$ has degree $2 < \infty$, by Stein's method, we have a generic CLT for $D(\pi)$.

Let $X_i = \begin{cases} 1 & \pi(i) > \pi(i+1) \\ 0 & \text{else} \end{cases}$

Probability

$$P(X_i X_{i+1} \ldots X_{i+k}) = \frac{1}{(k+2)!}$$

Eulerian numbers: $A(n,m)$ is the number of permutations of the numbers from 1 to n that have exactly m descents (ascents).
Possible strategy: Consider the following decomposition: $D(\pi) = Odd(\pi) + Even(\pi)$ where

$$Odd(\pi) = \#\{i : i \text{ odd}, 1 \le i \le n-1, \pi(i) > \pi(i+1)\}$$
$$Even(\pi) = \#\{i : i \text{ even}, 1 \le i \le n-1, \pi(i) > \pi(i+1)\}$$

Then we essentially have a random variable conditioned on a binary string of the values of the odd X_i, with each string equally likely since $\mathbb{E}(X_i) = 1/2$.

$$P(X_{i+1} = 1 \mid X_i = 1, X_{i+2} = 1) = 1/6$$
$$P(X_{i+1} = 1 \mid X_i = 1, X_{i+2} = 0) = 1/2$$
$$P(X_{i+1} = 1 \mid X_i = 0, X_{i+2} = 1) = 1/2$$
$$P(X_{i+1} = 1 \mid X_i = 0, X_{i+2} = 0) = 5/6$$

If we condition on the values of X_i for odd i, the only remaining variables are the X_i for even i, which are all independent. Furthermore, the distribution of an even X_i is determined entirely by the odd X_{i-1} and X_{i+1} surrounding it, so the only important data from the odd X_i, aside from the total number that are 1, is the number N_{11} of subsequences 11 of the binary string, the numbers N_{10} of subsequences 10 of the binary string, etc. Once the odd X_i are determined and these are fixed, D is the sum of four binomial distributions (plus a constant).

To establish a local limit theorem, we want to prove a quantitative bound for $\int_{-\pi\sigma_n}^{\pi\sigma_n} \left|\phi_n(t) - e^{-t^2/2}\right| dt$.

$$\left|\mathbb{E}e^{itD/\sigma}\right| = \left|\mathbb{E} \exp\left(it\left(\sum_{j=1}^{N_{11}} B_j + \dots\right)/\sigma\right)\right|$$

4 AP Stuff small t or whatever

We define a random variable $\mathcal{A}$ to be the number of 3-term arithmetic progressions in a randomly chosen subset of $\mathbb{F}_n$, where each element in $\mathbb{F}_n$ has probability p of appearing in the subset. The elements of the probability space can thus be described as $\mathbf{x} \in \{0,1\}^n$ where $\mathbf{x}_k$ is 1 with probability p and 0 with probability $1-p$. Further, the random variable $\mathcal{A}$ can be thought of as the function $\mathcal{A} : \{0,1\}^n \to \mathbb{N}$. Although $\mathcal{A}$ depends on both n and p, we take p to be fixed and constrain our analysis to n.

We define a p-biased Fourier basis for functions on the probability space. In order to do this we first define $\chi_k : \{0,1\}^n \to \mathbb{R}$ by

$$\chi_k := \chi_k(\mathbf{x}) := \frac{\mathbf{x}_k - p}{\sqrt{p(1-p)}} = \begin{cases} -\sqrt{\frac{p}{1-p}} & \text{if } \mathbf{x}_p = 0 \\ \sqrt{\frac{1-p}{p}} & \text{if } \mathbf{x}_p = 1 \end{cases}$$

so that χ_k is a normalized version of x_k. Further, we can extend this to define, for an arbitrary set $S \subset \mathbb{F}_n$,

$$\chi_S := \prod_{k \in S} \chi_k$$

Note that if we take the inner product of two functions $f, g : \{0,1\}^n \to \mathbb{R}$ to be $\mathbb{E}[fg]$, then $\{\chi_S : S \subset \mathbb{F}_n\}$ forms an orthonormal basis for functions on our probability space.

So if we define the Fourier transform $\hat{f} : \{0,1\}^n \to \mathbb{R}$ of an arbitrary function $f : \{0,1\}^n \to \mathbb{R}$ by

$$\hat{f}(S) := \mathbb{E}[f(\mathbf{x})\chi_S(\mathbf{x})$$

then from the orthonormality of our basis we get

$$f(\mathbf{x}) = \sum_{S \subset \mathbb{F}_n} \hat{f}(S)\chi_S(\mathbf{x})$$

We will use this expansion to calculate the variance of $\mathcal{A}$ and bound the pointwise distance of the characteristic function from that of the discrete Gaussian.

It will now be useful to normalize the random variable $\mathcal{A}$. We take the mean of $\mathcal{A}$ to be $\mu := \mathbb{E}[\mathcal{A}]$. We write the variance of $\mathcal{A}$ as $\sigma^2 := \mathbb{E}[\mathcal{A}^2] - \mathbb{E}[\mathcal{A}]^2$. So we define $Z : \{0,1\}^n \to \mathbb{R}$ by

$$Z := \frac{\mathcal{A} - \mu}{\sigma}$$

and we will often refer to the characteristic function of Z defined by $\phi_Z(t) := \mathbb{E}[e^{itZ}]$.

Before moving on to calculate the Fourier coefficients $\hat{\mathcal{A}}(S)$, we will first note that there are $\binom{n}{2}$ possible (non-trivial) arithmetic progressions in $\mathbb{F}_n$. There are first n choices for the start of the arithmetic progression, then $n-1$ choices for a non-trivial separation distance d, and finally both d and $-d$ will have counted the same arithmetic progression from different starting points, so we divide by 2 to yield $\frac{n(n-1)}{2} = \binom{n}{2}$. Additionally, each 3-term arithmetic progression occurs with probability p^3 (each of the three terms in the progression occur independently with probability p). This allows us to calculate

$$\mu = \mathbb{E}\left[\sum_{\Lambda} 1_\Lambda\right] = \sum_{\Lambda} \mathbb{E}[1_\Lambda] = \sum_{\Lambda} p^3 = p^3\binom{n}{2}$$

where 1_Λ is the indicator function for a particular 3-term arithmetic progression Λ in $\mathbb{F}_n$.

Furthermore, the fact that our basis for functions on the probability space is orthonormal allows us to calculate variance according to this formula, commonly known as Parseval's Theorem.

$$\sigma^2 = \sum_{S \neq \varnothing} \hat{\mathcal{A}}(S)^2$$

We now work with the arithmetic progression indicator functions 1_Λ a bit more. We use $k \in \Lambda$ to denote that k is a term in the arithmetic progression Λ. Therefore, the indicator function can be expressed as

$$\begin{aligned} 1_\Lambda(\mathbf{x}) &= \prod_{k \in \Lambda} \mathbf{x}_k = \prod_{k \in \Lambda} \left(\sqrt{p(1-p)}\chi_k + p\right) \\ &= p^3 + p^2\sqrt{p(1-p)}\sum_{k \in \Lambda} \chi_k + p^2(1-p) \sum_{k_1 \neq k_2 \in \Lambda} \chi_{\{k_1,k_2\}} + p^{3/2}(1-p)^{3/2} \end{aligned}$$

Note that any two elements of $\mathbb{F}_n$ appear in exactly 3 arithmetic progressions and any one element appears in exactly $\frac{3}{2}(n-1)$ arithmetic progressions. Hence, by summing over all arithmetic progressions, we have

$$\mathcal{A} = p^3\binom{n}{2} + \frac{3}{2}(n-1)p^2\sqrt{p(1-p)}\sum_{k \in \mathbb{F}_n} \chi_k + 3p^2(1-p) \sum_{\{k_1,k_2\} \in \mathbb{F}_n} \chi_{\{k_1,k_2\}} + \sum_{\Lambda} p^{3/2}(1-p)^{3/2}\chi_\Lambda$$

Thus, we have the Fourier Transform of $\mathcal{A}$:

$$\hat{\mathcal{A}}(S) = \begin{cases} p^3\binom{n}{2} & \text{if } S = \varnothing \\ \frac{3}{2}(n-1)p^2\sqrt{p(1-p)} & \text{if } |S| = 1 \\ 3p^2(1-p) & \text{if } S = \{k_1, k_2\} \\ p^{3/2}(1-p)^{3/2} & \text{if } S = \Lambda \\ 0 & \text{else} \end{cases}$$

and we can use Parseval's Theorem to give us the variance of $\mathcal{A}$ from this

$$\begin{aligned}
\sigma^2 &= \sum_{\substack{S\subset\mathbb{F}_n\\ S\neq\varnothing}} \hat{\mathcal{A}}(S)^2 \\
&= \sum_{k\in\mathbb{F}_n}\left(\frac{3}{2}(n-1)p^2\sqrt{p(1-p)}\right)^2 + \sum_{\{k_1,k_2\}\in\mathbb{F}_n}\left(3p^2(1-p)\right)^2 + \sum_{\Lambda}\left(p^{3/2}(1-p)^{3/2}\right)^2 \\
&= \frac{9}{4}n(n-1)^2p^5(1-p) + 9\binom{n}{2}p^4(1-p)^2 + \binom{n}{2}p^3(1-p)^3 \\
&= \Theta(n^3)
\end{aligned}$$

We also derive the fact that $\sigma = \Theta(n^{3/2})$ from this calculation. We proceed to show the following bound for small values of t:

Proposition 4.1. *For* $|t| \ll \sqrt{n}$,

$$\left|\phi_Z(t) - e^{-t^2/2}\right| \leq O\left(\frac{t^3e^{-t^2/3}}{\sqrt{n}} + \frac{t}{\sqrt{n}}\right).$$

Proof. To begin, we decompose $Z = X + Y$, where

$$Q = \frac{1}{\sqrt{n}}, X = \sum_{k\in\mathbb{F}_n} Q\chi_k, Y = \sum_{k\in\mathbb{F}_n}(\hat{Z}(k) - Q)\chi_k + \sum_{|S|\geq 2}\hat{Z}(k)\chi_k.$$

We first bound the distance from the characteristic function of X to the normal distribution: Since we normalized X using Q, X is a random variable with mean 0 and variance 1. Since

$$L_n := n\mathbb{E}[|Q\chi_k|^3] = O\left(\frac{1}{\sqrt{n}}\right) < \infty,$$

then by Berry-Esseen, if $t \leq \frac{1}{4L_n}$, then

$$|\mathbb{E}[e^{itX}] - e^{-t^2/2}| \leq 16L_nt^3e^{-t^2/3}.$$

Next we consider Y. By Cauchy-Schwarz, orthogonality of our basis, and Parseval's Theorem, we have

$$\mathbb{E}|Y| \leq \sqrt{\mathbb{E}|Y|^2} = \sqrt{var(Y)} = \sqrt{\sum_{k\in\mathbb{F}_n}(\hat{Z}(k) - Q) + \sum_{|S|\geq 2}\hat{Z}(S).}$$

Now using our above calculation of the variance σ^2 of $\mathcal{A}$,

$$\sum_{|S|\geq 2}\hat{Z}(S) = \frac{9\binom{n}{2}p^4(1-p)^2 + \binom{n}{2}p^3(1-p)^3}{\sigma^2} = \Theta\left(\frac{1}{n}\right).$$

In addition, using our calculation of $\hat{\mathcal{A}}$,

$$\begin{aligned}
n\sigma\hat{Z}^2(k) - \sigma^2 = n\hat{\mathcal{A}}^2(k) - \sigma^2 &= O(n^2) \\
n\hat{Z}^2(k) - 1 &= O\left(\frac{1}{n}\right) \\
\hat{Z}^2(k) - \frac{1}{n} &= O\left(\frac{1}{n^2}\right).
\end{aligned}$$

Since $\hat{Z}(k) + Q = O\left(\frac{1}{\sqrt{n}}\right)$ and $Q^2 = \frac{1}{n}$,

$$\left|\hat{Z}(k) - Q\right| \leq \left|\frac{\hat{Z}(k) - \frac{1}{n}}{\hat{Z}(k) + Q}\right| = O\left(\frac{1}{n^{3/2}}\right).$$

Hence, we have $\mathbb{E}[Y] = O\left(\frac{1}{\sqrt{n}}\right)$. Thus, we conclude that if $t < \frac{1}{4L_n} = O(\sqrt{n})$,

$$\begin{aligned}
\left|\phi_Z(t) - e^{-t^2/2}\right| &= \left|\mathbb{E}[e^{itZ} - e^{-t^2/2}]\right| = \left|\mathbb{E}[e^{it(X+Y)} - e^{-t^2/2}]\right| \\
&\leq \left|\mathbb{E}[e^{it(X+Y)} - e^{itX}]\right| + \left|\mathbb{E}[e^{itX} - e^{-t^2/2}]\right| \\
&\leq \mathbb{E}[tY] + 16L_n t^3 e^{-t^2/3} \qquad \text{(using the Mean Value Theorem on } e^{itX}\text{)} \\
&= O\left(\frac{t^3 e^{-t^2/3}}{\sqrt{n}} + \frac{t}{\sqrt{n}}\right).
\end{aligned}$$

□