Quantitative and Local Central Limit Theorems

Abstract

We prove a quantitative local limit theorem for the number of descents in a random permutation. We conjecture that there is no local limit theorem for the number of length 3 arithmetic progressions in a random subset of $\mathbb{Z}/n\mathbb{Z}$, but a proof of this remains elusive. A promising avenue of proof is to condition on the size of the subset and show that the resulting distributions are too far apart for different sizes. This has proven difficult because the distances between these conditioned distributions are extremely close to their standard deviations.

1 Introduction

Let $\{X_n\}$ be a sequence of discrete random variables taking integer values with mean μ_n and standard deviation σ_n. For example, X_n could be the number of heads that occur in n coinflips. The central question is of convergence of the X_n. In the example of coinflips, the X_n themselves do not converge in any meaningful way, since μ_n is increasing in n, so we consider convergence of $Y_n = \frac{X_n - \mu_n}{\sigma_n}$ to the standard normal distribution Z instead.

One notion of convergence of a sequence of random variables is convergence in distribution [3]. This concerns pointwise convergence of cumulative distribution functions. We write that $\{Y_n\}$ converges in distribution to Z, or $Y_n \xrightarrow{d} Z$, if

$$|\mathbb{P}(Y_n \leq t) - \mathbb{P}(Z \leq t)| \to 0$$

for each $t \in \mathbb{R}$. If this condition holds, the sequence $\{X_n\}$ satisfies a **central limit theorem**. The example of coinflips satisfies a central limit theorem. But what if we wanted to know, for instance, the probability of getting exactly half heads, that is, $\mathbb{P}(X_n = n/2)$ (for even n)? The central limit theorem only tells us the probability of having at most half heads, that is, $\mathbb{P}(X_n < \frac{n}{2}) = \mathbb{P}(Y_n < 0) \to \mathbb{P}(Z < 0) = \frac{1}{2}$. We could try to extract a pointwise probability from the central limit theorem by observing

$$\mathbb{P}(X_n = k) = \mathbb{P}(k - \frac{1}{2} < X_n < k + \frac{1}{2})$$
$$= \mathbb{P}((k - \frac{1}{2} - \mu_n)/\sigma_n < Y_n < (k + \frac{1}{2} - \mu_n)/\sigma_n)$$
$$\to \mathbb{P}((k - \frac{1}{2} - \mu_n)/\sigma_n < Z < (k + \frac{1}{2} - \mu_n)/\sigma_n).$$

But this is not very helpful, since each quantity simply goes to 0 as n gets large and there is no bound on the rate of convergence. Another way to phrase it is that

$$\left| \mathbb{P}(X_n = k) - \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\left(\frac{k-\mu_n}{\sigma_n}\right)^2} \right| \to 0,$$

but this is trivial since both terms go to 0. To find an error bound that makes the convergence meaningful, observe that a normal distribution with standard deviation σ has height $\Theta(\frac{1}{\sigma})$. Therefore, we want

$$\left| \mathbb{P}(X_n = k) - \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\left(\frac{k-\mu_n}{\sigma_n}\right)^2} \right| = o\left(\frac{1}{\sigma_n}\right)$$

uniformly in k. If this condition holds, the sequence $\{X_n\}$ satisfies a **local limit theorem**. The existence of local limit theorems is our main area of interest.

1.1   Descents

One variable of interest to us is the number of descents in a random permutation. A permutation π has a descent at index i if $\pi(i) > \pi(i+1)$. The number of descents in π, denoted $D(\pi)$, is the number of such indices i. We define the indicator random variable for the ith descent

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

so the number of descents can be written

$$D_n = \sum_{i=1}^{n-1} X_i.$$

To sample random permutations, uniformly and independently pick random integers $1 \le a_i \le n-i+1$ for each $1 \le i \le n$. The a_i correspond to permutations like so: let $S = \{1, \ldots, n\}$, and for each i, in order from 1 to n, define $\pi(i)$ to be the ith remaining element of S, and remove $\pi(i)$ from S. Under this correspondence, $\pi(i) > \pi(i+1)$ if and only if $a_i > a_{i+1}$. Therefore, X_i can be equivalently defined by

$$X_i = \begin{cases} 1 & a_i > a_{i+1} \\ 0 & \text{else} \end{cases}.$$

With this method of sampling, we can compute some basic facts about D. At each index i, there is either a descent or an ascent ($\pi(i) < \pi(i+1)$) and these both occur with equal probability, so $\mathbb{E}[X_i] = \frac{1}{2}$. Therefore, $\mathbb{E}[D] = \frac{n-1}{2}$. With similar reasoning as in computing the expectation of X_i, we get $\mathbb{E}[X_i X_{i+1}] = \frac{1}{3!} = \frac{1}{6}$. Also, it is important to note that X_i is

2

independent of all other X_j except for X_{i-1} and X_{i+1}. Now we compute

$$
\begin{aligned}
\operatorname{Var}(D) &= \mathbb{E}[D^2] - \mathbb{E}[D]^2 \\
&= \mathbb{E}\left[\sum_{i=1}^{n-1}\sum_{j=1}^{n-1} X_i X_j\right] - \left(\frac{n-1}{2}\right)^2 \\
&= \sum_{i=1}^{n-1}\mathbb{E}[X_i^2] + \sum_{|i-j|=1}\mathbb{E}[X_i X_j] + \sum_{|i-j|>1}\mathbb{E}[X_i X_j] - \frac{(n-1)^2}{4} \\
&= \frac{n-1}{2} + \frac{2(n-2)}{6} + \frac{(n-1)^2 - (n-1) - 2(n-2)}{4} - \frac{(n-1)^2}{4} \\
&= \frac{n+1}{12}.
\end{aligned}
$$

1.2 Arithmetic Progressions

The other random variable of interest is the number of length 3 arithmetic progressions (APs) in a random subset of $\mathbb{Z}/n\mathbb{Z}$. We require n odd and $n > 3$ to avoid issues of divisibility. For $1 \le i \le n$, let x_i be the indicator random variable for $i \in S \subseteq \mathbb{Z}/n\mathbb{Z}$. The x_i are independent and each has expectation $\frac{1}{2}$. We define the number of arithmetic progressions in S as

$$
A_n = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n-1} x_i x_{i+j} x_{i+2j}.
$$

This definition does not consider triples that contain the same element twice to be an arithmetic progression, and the factor of $\frac{1}{2}$ causes multiple triples with the same elements to be considered the same progression. The mean of A_n is $\mathbb{E}[A_n] = \frac{1}{8}\binom{n}{2}$. The variance of A_n is computed in section 4.

2 Background

2.1 Number of Triangles in a Random Graph

Previous papers have studied the distribution of the number of triangles T_n in a random graph $G_{n,p}$, which is the undirected graph on n vertices where each of the $\binom{n}{2}$ edges have probability p of appearing in the graph. Let $R_n = \frac{T_n - \mu}{\sigma}$ be the normalized T_n. The standard normal probability density function $\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$. Gilmer and Kopparty (ADD CITATION) establish the existence of an LLT for T_n. In particular, they show that the distribution of T_n is pointwise approximated by a discrete Gaussian distribution.

Theorem 2.1 (Gilmer, Kopparty, 2014). *If $L = \{\frac{k-\mu_n}{\sigma_n} : k \in \mathbb{Z}\}$, as $n \to \infty$,*

$$
\sup_{x \in L} |\sigma_n P(R_n = x) - \mathcal{N}(x)| \to 0.
$$

We note that this is a qualitative result, where the error $= o(n^{-2}) = o(\frac{1}{\sigma})$. Berkowitz (ADD CITATION) expanded on their work and established a quantitative bound on the distance between the distribution of T_n and the Gaussian distribution. In particular, he shows the following result:

Theorem 2.2 (Berkowitz, 2016). *For all $k \in \mathbb{Z}$,*

$$P(T_n = k) = \frac{1}{\sqrt{2\pi}\sigma_n}\exp\left(-((k-\mu_n)/\sigma_n)^2/2\right) + o(n^{-2.5+\epsilon}).$$

2.1.1 Methods used to establish an LLT for T_n

Here we summarize the methods used by Gilmer and Kopparty (ADD CITATION) in proving that an LLT exists for T_n. With some calculations, we have the mean $\mathbb{E}[T_n] = \mu_n = p^3\binom{n}{3}$ and the variance $\sigma^2 = \Theta(n^4)$. A crucial formula for the proof is the following:

Proposition 2.3 (Fourier Inversion Formula). *If Y is a random variable with support in the discrete lattice $L = \frac{1}{b}(\mathbb{Z}-a)$ for $a, b \in \mathbb{R}$, and $\phi(t) = \mathbb{E}[e^{itY}]$ is the characteristic function of Y, then $\forall\ y \in L$,*

$$P(Y = y) = \frac{1}{2\pi b}\int_{-\pi b}^{\pi b} e^{-ity}\phi(t)\ dt.$$

If we let $\phi_n(t) = \mathbb{E}[e^{itR_n}]$, then $\sigma_n P(R_n = x) = \frac{1}{2\pi}\int_{-\pi\sigma}^{\pi\sigma} e^{-itx}\phi(t)\ dt$. By the standard Fourier inversion formula, $\mathcal{N}(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}e^{x^2/2}\ dt$. Thus,

$$|\sigma_n P(R_n = x)| \leq \int_{-\pi\sigma_n}^{\pi\sigma_n}\left|\phi_n(t) - e^{-t^2/2}\right| dt + 2\int_{\pi\sigma_n}^{\infty} e^{-t^2/2}\ dt$$

Thus, since the second term goes to 0 as n goes to infinity, we just have to show

$$\int_{-\pi\sigma_n}^{\pi\sigma_n}\left|\phi_n(t) - e^{-t^2/2}\right| dt = o(1)$$

.

For any constant A, we can write

$$\int_{-\pi\sigma}^{\pi\sigma}\left|\phi_n(t) - e^{-t^2/2}\right| dt \leq \int_{-A}^{A}\left|\phi_n(t) - e^{-t^2/2}\right| dt + \int_{A \leq |t| \leq \pi\sigma}\left(|\phi_n(t)| + |e^{-t^2/2}|\right)$$

Hence, we have reduced the problem of proving an LLT to one of sufficiently bounding the characteristic function $|\phi_n(t)|$. The overall strategy used to bound the $|\phi_n(t)|$ involves finding an event that occurs with high probability and allows R_n to be written as the sum of n i.i.d. random variables X_i after conditioning on the event. Let F be all the edges except for those in a perfect matching M, and let X_F be the indicator vector for edges in F that appear in G. For any edge $e = \{u, v\}$ in M, let C_e be the number of paths from u to v of length 2 that appear in G. Now C_e and C, the number of triangles with edges only in F, are both constants conditioned on X_F. Then using a few lemmas (see G-K for details), we can conclude

$$\left|\mathbb{E}[e^{itR_n}|X_F]\right| = \left|\mathbb{E}[e^{it(\sum_{e \in M} C_e X_e)/\sigma_n}]\right| \approx e^{-\Theta(t^2/n)}.$$

The remaining elements and details of the proof are omitted here.

Using the p-biased Fourier basis, Berkowitz shows $\int_{-\pi\sigma_n}^{\pi\sigma_n}\left|\phi_n(t) - e^{-t^2/2}\right| dt = O(n^{-1/2+\epsilon})$. In section 4 below, we set up and apply this tool for 3-term arithmetic progressions.

3 Descents in a Permutation

Let $\pi \in S_n$ be a permutation. Define $D(\pi) = |\{1 \leq i < n \mid \pi(i+1) < \pi(i)\}|$ to be the number of descents in π. Viewing D_n as a random variable on uniformly distributed permutations, a central limit theorem is known and we establish a local limit theorem.

As another way of viewing a permutation π, consider a sequence $a_1, \ldots, a_n$ with $1 \leq a_i \leq n - i + 1$. To get a permutation π from such a sequence, start with $S = \{1, \ldots, n\}$. For $i = 1, \ldots, n$, let b_i be the a_ith remaining element of S, let $\pi(i) = b_i$, and remove b_i from S. Then $\pi(i + i) < \pi(i)$ if and only if $a_{i+1} < a_i$, so the problem of counting descents in π is equivalent to the problem of counting descents in $\{a_i\}$. This is nice since the a_i are distributed independently.

Now, define the random variable X_i for $1 \leq i < n$ as the indicator for the ith potential descent, that is, 1 if $a_{i+1} < a_i$ and 0 otherwise. We can write $D_n = X_1 + \ldots + X_n$. Since X_i depends only on a_i and a_{i+1}, X_i is independent of all other X_j except for X_{i-1} and X_{i+1}. We calculate

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

To bound the characteristic function of D_n, we first fix the values of X_i for odd i. For simplicity, assume n is even. (When n is odd, one must also observe that $\mathbb{P}(X_i = 1 \mid X_{i-1} = 1) = 1/3$ and $\mathbb{P}(X_i = 1 \mid X_{i-1} = 0) = 2/3$ are constant.) The even X_i are independent of each other, so conditioned on the odd X_i, D_n is a sum of independent random variables. The values of the odd X_i can be viewed as a binary string, and since the distribution of the even X_i depend only on the adjacent odd X_{i-1} and X_{i+1}, the distribution of the sum of the even X_i depends entirely on the number of occurrences of each length 2 substring in the binary string.

Write N_{11} for the number of substrings 11, N_{10} for the number of substrings 10, and the same for N_{01} and N_{00}. Let $B_j^{(p)}$ denote an independent Bernoulli random variable with expectation p. We compute

$$|\phi_n(t)| = \left| \mathbb{E}\left[e^{itD_n/\sigma} \right] \right| = \left| \mathbb{E}\left[\exp\left(it \left(\sum_{j=1}^{N_{11}} B_j^{(\frac{1}{6})} + \sum_{j=1}^{N_{10}} B_j^{(\frac{1}{2})} + \sum_{j=1}^{N_{01}} B_j^{(\frac{1}{2})} + \sum_{j=1}^{N_{00}} B_j^{(\frac{5}{6})} \right) / \sigma \right) \right] \right|$$

$$= \left| \mathbb{E}\left[e^{it \sum_{j=1}^{N_{11}} B_j^{(\frac{1}{6})}/\sigma} \right] \mathbb{E}\left[e^{it \sum_{j=1}^{N_{10}} B_j^{(\frac{1}{2})}/\sigma} \right] \mathbb{E}\left[e^{it \sum_{j=1}^{N_{01}} B_j^{(\frac{1}{2})}/\sigma} \right] \mathbb{E}\left[e^{it \sum_{j=1}^{N_{00}} B_j^{(\frac{5}{6})}/\sigma} \right] \right|$$

Using an argument similar to that of Gilmer-Kopparty [1], we bound each expectation sep-

arately.

$$\left| \mathbb{E}\left[e^{it\sum_{j=1}^{N} B_j^{(p)}/\sigma}\right] \right| = \left| \prod_{j=1}^{N} \mathbb{E}\left[e^{itB_j^{(p)}/\sigma}\right] \right|$$

$$\leq \left(1 - 8p(1-p)\left\|\frac{t}{2\pi\sigma_n}\right\|^2\right)^N$$

$$= \left(1 - 8p(1-p)\left(\frac{t}{2\pi\sigma_n}\right)^2\right)^N \qquad \text{(when } t < \pi\sigma_n)$$

$$\leq \left(1 - C\left(\frac{t}{2\pi\sigma_n}\right)^2\right)^N \qquad \text{(taking } C = \min\{8p(1-p)\} = 80/9)$$

Since $N_{11} + N_{10} + N_{01} + N_{00} = \frac{n}{2} - 1$, we get

$$|\phi_n(t)| \leq \left(1 - C\left(\frac{t}{2\pi\sigma_n}\right)^2\right)^{(\frac{n}{2}-1)}$$

$$= \left(1 - \Theta(t^2/n)\right)^{(\frac{n}{2}-1)}$$

$$\leq e^{-\Theta(t^2)}.$$

Now we obtain the final bound for the local limit theorem. As shown in [2], we have a central limit theorem

$$\sup_{-\infty < x < \infty}\left|\mathbb{P}(\frac{D_n - \mu}{\sigma} < t) - \mathbb{P}(Z < t)\right| \leq \frac{12}{\sqrt{n}}$$

.

For $y < 0$, $P(D_n - \mu \leq y\sigma) = \frac{1}{2}P(|D_n - \mu| \geq |y|\sigma) \leq \frac{1}{2y^2}$ by Chebyshev's Inequality. Similarly, $P(D_n - \mu \leq y\sigma) \leq \frac{1}{2y^2}$.

Thus,

$$\left|\mathbb{P}(\frac{D_n - \mu}{\sigma} < t) - \mathbb{P}(Z < t)\right| = \begin{cases} \frac{12}{\sqrt{n}} \\ \mathbb{P}(\frac{D_n - \mu}{\sigma} \leq y) + \mathbb{P}(Z \leq y) \leq \frac{1}{2y^2} + e^{-\Theta(y^2)} & \text{for } y < 0 \\ \mathbb{P}(\frac{D_n - \mu}{\sigma} \geq y) + \mathbb{P}(Z \geq y) \leq \frac{1}{2y^2} + e^{-\Theta(y^2)} & \text{for } y > 0 \end{cases}$$

.

Hence, we have

$$\left|\phi_n(t) - e^{-t^2/2}\right| \leq |t|\int_R \left|\mathbb{P}(\frac{D_n - \mu}{\sigma} < y) - \mathbb{P}(Z < y)\right| dy \tag{1}$$

$$\leq |t|\int_{|y|>k\sigma}\left|\mathbb{P}(\frac{D_n - \mu}{\sigma} < y) - \mathbb{P}(Z < y)\right| dy + |t|\int_{|y|\leq k\sigma}\left|\mathbb{P}(\frac{D_n - \mu}{\sigma} < y) - \mathbb{P}(Z < y)\right| dy \tag{2}$$

$$\leq |t|\int_{|y|>k\sigma}\frac{1}{2y^2} + e^{-\Theta(y^2)}dy + |t|2k\sigma\sqrt{\frac{12}{n}}. \tag{3}$$

6

If we take $k = O(n^{-\frac{1}{2}+\epsilon})$. Since $\sigma = O(n)$, we have $\left|\phi_n(t) - e^{-t^2/2}\right| \leq |t| O(n^{-\frac{1}{2}+\epsilon})$. Thus, for any $\varepsilon > 0$, we compute

$$
\int_{-\pi\sigma}^{\pi\sigma} \left|\phi_n(t) - e^{-t^2/2}\right| \mathrm{d}t \leq \int_{-n^\varepsilon}^{n^\varepsilon} \left|\phi_n(t) - e^{-t^2/2}\right| \mathrm{d}t + \int_{n^\varepsilon < |t| < \pi\sigma} (|\phi_n(t)| + |e^{-t^2/2}|)\mathrm{d}t
$$
$$
\leq \int_{-n^\varepsilon}^{n^\varepsilon} \left|\phi_n(t) - e^{-t^2/2}\right| \mathrm{d}t + \int_{n^\varepsilon < |t| < \pi\sigma} e^{-\Theta(t^2)}\mathrm{d}t
$$
$$
= O(n^{-\frac{1}{2}+\varepsilon}).
$$

4   3-Term Arithmetic Progressions

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

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

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

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

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

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

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

$$\sigma^2 = \sum_{\substack{S \subseteq \mathbb{F}_n \\ S \neq \varnothing}} \hat{\mathcal{A}}(S)^2$$

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

$$= \frac{9}{4}n(n-1)^2 p^5 (1-p) + 9\binom{n}{2}p^4(1-p)^2 + \binom{n}{2}p^3(1-p)^3$$

$$= \Theta(n^3)$$

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* $|t| \ll \sqrt{n}$,

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

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 $t \leq \frac{1}{4L_n}$, we have

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

9

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

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

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 $\hat{\mathcal{A}}$,

$$n\sigma\hat{Z}^2(k) - \sigma^2 = n\hat{\mathcal{A}}^2(k) - \sigma^2 = O(n^2)$$

$$n\hat{Z}^2(k) - 1 = O\left(\frac{1}{n}\right)$$

$$\hat{Z}^2(k) - \frac{1}{n} = O\left(\frac{1}{n^2}\right).$$

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 $t < \frac{1}{4L_n} = O(\sqrt{n})$,

$$\left|\phi_Z(t) - e^{-t^2/2}\right| = \left|\mathbb{E}[e^{itZ} - e^{-t^2/2}]\right| = \left|\mathbb{E}[e^{it(X+Y)} - e^{-t^2/2}]\right|$$

$$\leq \left|\mathbb{E}[e^{it(X+Y)} - e^{itX}]\right| + \left|\mathbb{E}[e^{itX} - e^{-t^2/2}]\right|$$

$$\leq \mathbb{E}[tY] + 16L_n t^3 e^{-t^2/3} \qquad \text{(using the Mean Value Theorem on } e^{itX})$$

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

$\square$

4.0.1   Comments on small t bound

4.1   Experimental Results and Conjectures

As the bound achieved in the previous section was not particularly promising, we decided to simulate the number of arithmetic progressions in a random subset of $\mathbb{Z}/n\mathbb{Z}$ in order to verify whether or not an LLT actually held.
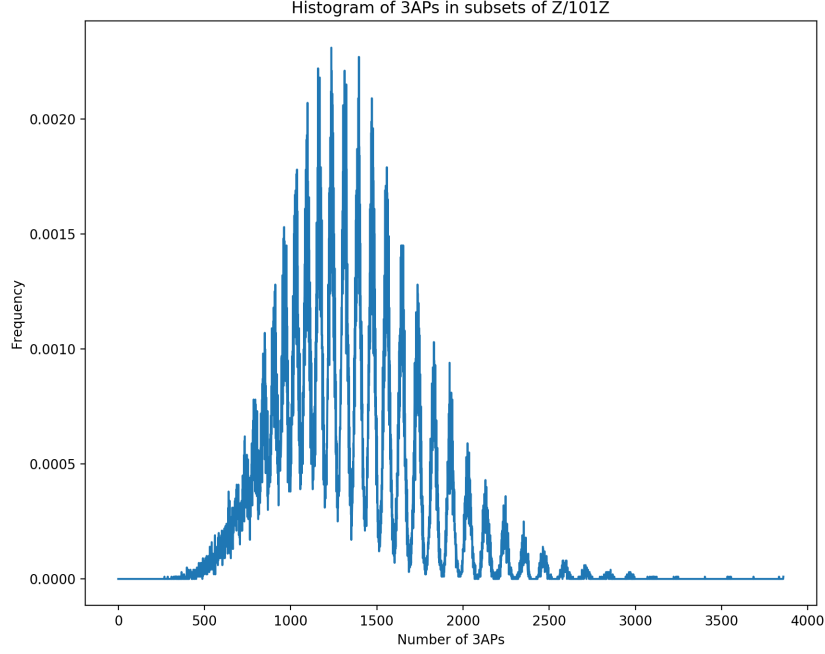
Figure 1: Histogram of the number of 3-term arithmetic progressions in 100,000 random subsets of $\mathbb{Z}/101\mathbb{Z}$.

As figure 1 demonstrates, the number of arithmetic progressions in a random subset of $\mathbb{Z}/n\mathbb{Z}$ does not follow the Gaussian distribution pointwise. However, after some inspection, one might notice that the oscillations in the histogram are spaced out almost evenly. Furthermore, the oscillations themselves appear fairly Gaussian, almost as if the entire distribution is the sum of several spaced out Gaussians. Given that the spacings appear to be $O(n)$ and the entire distribution ranges from 0 to $\binom{n}{2} = O(n^2)$, we would expect there to be $O(n)$ of these smaller Gaussians. It is thus reasonable to conclude that the number of 3-term arithmetic progressions in a random subset of $\mathbb{Z}/n\mathbb{Z}$ depends on some other random variable which can take $O(n)$ values.

A good guess for this other random variable is the size of a random set of $\mathbb{Z}/n\mathbb{Z}$. From figure 2 we can see that the number of 3-term arithmetic progressions in a random subset of $\mathbb{Z}/n\mathbb{Z}$ with fixed size does, in fact, follow the Gaussian distribution pointwise. From now on, we call this random variable $A_{n,k}$, where k is the fixed size of the random subset of $\mathbb{Z}/n\mathbb{Z}$.

Therefore, we conjecture that the distance between these smaller distributions, $\mathbb{E}[A_{n,k+1}] - \mathbb{E}[A_{n,k}]$, is sufficiently smaller than the standard deviation of the distributions themselves, $\sqrt{Var(A_{n,k})}$. In other words, there is a non-trivial interval between $\mathbb{E}[A_{n,k}]$ and $\mathbb{E}[A_{n,k+1}]$ where for all x in the interval, $P(A_{n,k} = x)$ and $P(A_{n,k+1} = x)$ are both very small. In particular, we suspect that those two probabilities are significantly smaller than the Gaussian approximation of $P(A_n = x)$ would suggest.

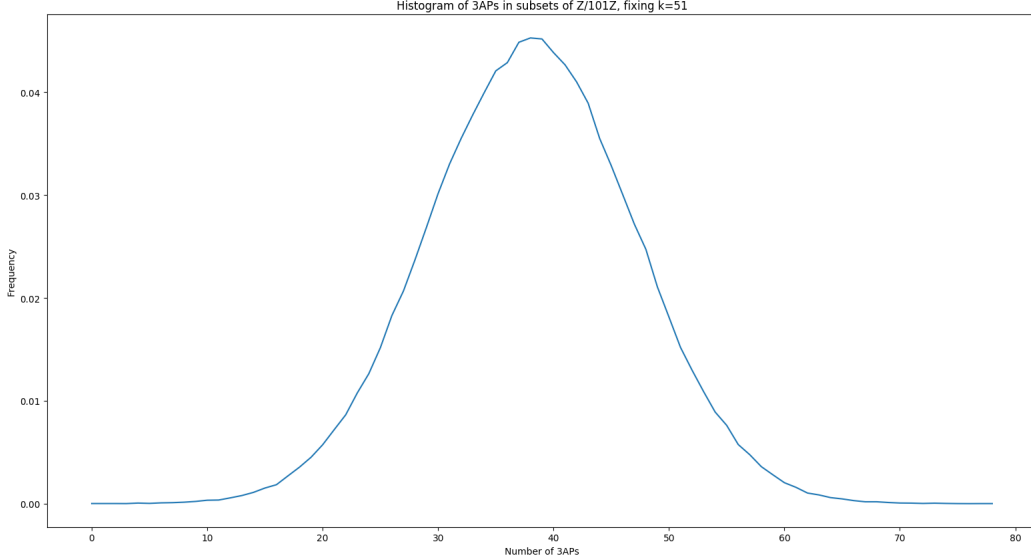This train of thought will be made more precise in sections 6 and 7.

11

Figure 2: Histogram of the number of 3-term arithmetic progressions in 200,000 random subsets of size 51 of $\mathbb{Z}/101\mathbb{Z}$

4.1.1 Continuous sets

The first thing that should come to one's mind after observing that A_n does not have an LLT due to the dependence on the size of the random set, is whether A_n would have an LLT if the size of the set was made continuous. That is to say, the variable x_i which denotes whether an element i of $\mathbb{Z}/n\mathbb{Z}$ is in our random set S, is now a uniform random variable on $[0, 1]$ instead of a Bernoulli random variable on $\{0, 1\}$ with $p = \frac{1}{2}$. The number of arithmetic progressions $\bar{A}_n$ is still defined in a similar manner,

$$\bar{A}_n = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n-1} x_i x_{i+j} x_{i+2j}$$

We simulated this to see if it possibly followed the Gaussian distribution pointwise. From figure 3, we conclude that it is very likely that there is an LLT for $\bar{A}_n$. However, we have not pursued a proof of this.

5 Attempts to get CLT or LLT for $A_{n,k}$

Attempt to get CLT for $A_{n,k}$ using Stein's Method

In order to prove a CLT for $A_{n,k}$, we use the method of exchangeable pairs. Let $A_{n,k}$ be the number of 3-term arithmetic progressions given the subset size $= k$, and W be $A_{n,k}$ standardized. We want to show that W converges to a normal distribution. We construct W' as follows. We have a random subset $S \subseteq \mathbb{F}_n$ with size k. Randomly choose two elements $I, J \in \mathbb{F}_n$, swap their status regarding their inclusion in S, and call the new subset S' (i.e. if $I \in S$ and $J \notin S$, then $I \notin S'$ and $J \in S'$). Now define A to be the number of arithmetic
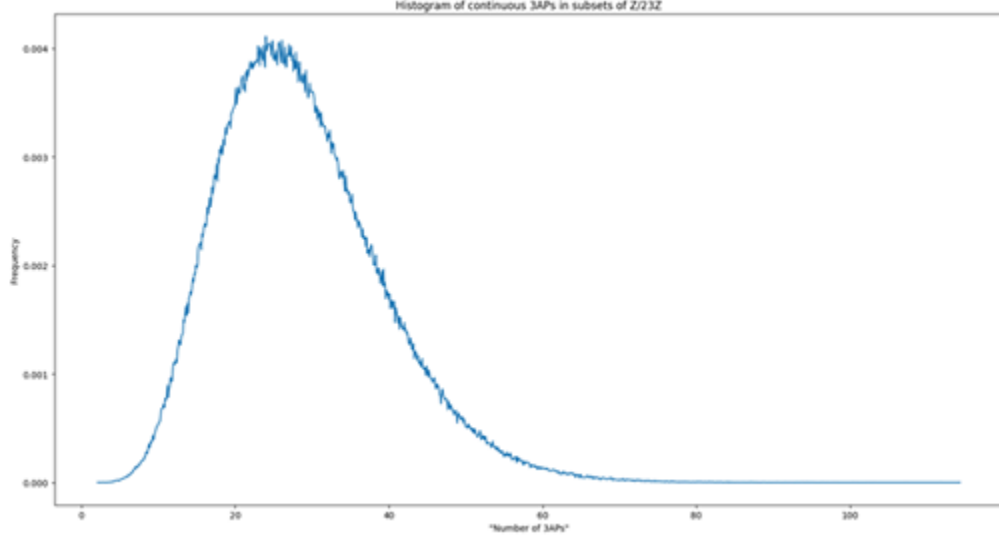
Figure 3: Number of 3-term "arithmetic progressions" in 1,000,000 random "continuous sets" of $\mathbb{Z}/23\mathbb{Z}$

progressions in S and define W to be standardized A, namely $\frac{A-\mu_{n,k}}{\sigma_{n,k}}$. Similarly, define A' to be the number of arithmetic progressions in S' and define W' to be $\frac{A'-\mu_{n,k}}{\sigma_{n,k}}$.

Remark 5.1. W and W' are an exchangeable pair, each with mean 0 and variance 1.

Let Λ be a 3-term arithmtic progression in $\mathbb{F}_n$. Define $1_{\Lambda \subseteq S} = \begin{cases} 1 & \Lambda \subseteq S \\ 0 & \text{else} \end{cases}$

Then $A = \sum_{\Lambda \subseteq \mathbb{F}_n} 1_{\Lambda \subseteq S}$.

Also note that $\mu_{n,k} = \mathbb{E}[\sum_{\Lambda \subseteq \mathbb{F}_n} 1_{\Lambda \subseteq S}] = \sum_{\Lambda \subseteq \mathbb{F}_n} P(\Lambda \in S) = \binom{n}{2}\frac{\binom{k}{3}}{\binom{n}{3}}$.

We first must show that $\mathbb{E}[W' - W \mid W] = -\lambda W$ for some $\lambda \in (0, 1)$.

Lemma 5.2. $\lambda = \frac{3(n-k)}{\binom{n}{2}}$

Proof. We have

$$\mathbb{E}[A' \mid A] = \sum_{\Lambda \subseteq \mathbb{F}_n} \mathbb{E}[1_{\Lambda \subseteq S'} \mid A] = \sum_{\Lambda \subseteq \mathbb{F}_n} P(\Lambda \subseteq S' \mid A).$$

Now

$$P(\Lambda \subseteq S' \mid A) = P(\Lambda \subseteq S' \mid \Lambda \subseteq S, A)P(\Lambda \subseteq S \mid A) + P(\Lambda \subseteq S' \mid \Lambda \not\subseteq S, A)P(\Lambda \not\subseteq S \mid A).$$

Note that

$$P(\Lambda \subseteq S \mid A) = \frac{A}{\binom{n}{2}}$$

$$P(\Lambda \not\subseteq S \mid A) = 1 - \frac{A}{\binom{n}{2}}$$

13

as there are A arithmetic progressions in S and $\binom{n}{2}$ arithmetic progressions total in $\mathbb{F}_n$.

We have

$$P(\Lambda \subseteq S' \mid \Lambda \subseteq S, A) = 1 - \frac{3(n-k)}{\binom{n}{2}}$$

since the probability of $\Lambda \not\subseteq S'$ is the probability of one of its 3 elements being selected to be swapped along with one of the elements outside of S.

Taking this into account in our previous expression, we have

$$P(\Lambda \subseteq S' \mid A) = \left(1 - \frac{3(n-k)}{\binom{n}{2}}\right)\left(\frac{A}{\binom{n}{2}}\right) + P(\Lambda \subseteq S' \mid \Lambda \not\subseteq S, A)\left(1 - \frac{A}{\binom{n}{2}}\right)$$

Finally, we examine

$$P(\Lambda \subseteq S' \mid \Lambda \not\subseteq S, A) = \sum_{i=0}^{3} P(\Lambda \subseteq S' \text{ and } |\Lambda \cap S| = i \mid \Lambda \not\subseteq S \text{ and } A)$$

$$= \sum_{i=0}^{2} P(\Lambda \subseteq S' \mid \Lambda \not\subseteq S \text{ and } |\Lambda \cap S| = i \text{ and } A)P(|\Lambda \cap S| = i \mid \Lambda \not\subseteq S \text{ and } A)$$

$$= P(\Lambda \subseteq S' \mid |\Lambda \cap S| = 2, \text{ and } A)P(|\Lambda \cap S| = 2 \mid \Lambda \not\subseteq S \text{ and } A).$$

We have $P(\Lambda \subseteq S' \mid |\Lambda \cap S| = 2, \text{ and } A) = \frac{k-2}{\binom{n}{2}}$ since the only way for Λ to be contained in S' if $|\Lambda \cap S| = 2$ is if the one element of Λ that is not in S is chosen, along with one other element of $S \backslash \Lambda$.

Substituting this in, we get

$$P(\Lambda \subseteq S' \mid \Lambda \not\subseteq S, A) = \left(\frac{k-2}{\binom{n}{2}}\right) P(|\Lambda \cap S| = 2 \mid \Lambda \not\subseteq S \text{ and } A)$$

$$P(|\Lambda \cap S| = 2 \mid \Lambda \not\subseteq S \text{ and } A) = 3P(\Lambda_1 \in S, \Lambda_2 \in S, \Lambda_3 \notin S \mid \Lambda \not\subseteq S \text{ and } A)$$

$$= 3P(\Lambda_3 \notin S \mid \Lambda_1 \in S, \Lambda_2 \in S, \Lambda \not\subseteq S, A)$$

$$\cdot P(\Lambda_2 \in S \mid \Lambda \not\subseteq S, \Lambda_1 \in S, A)P(\Lambda_1 \in S \mid \Lambda \not\subseteq S, A)$$

$$= 3P(\Lambda_2 \in S \mid \Lambda \not\subseteq S, \Lambda_1 \in S, A)P(\Lambda_1 \in S \mid \Lambda \not\subseteq S, A)$$

Here we have two non-trivial quantities to examine. First we see by Bayes's rule that

$$P(\Lambda_1 \in S \mid \Lambda \not\subseteq S, A) = \frac{P(\Lambda \not\subseteq S \mid \Lambda_1 \notin S, A)P(\Lambda_1 \not\subseteq S, A)}{P(\Lambda \not\subseteq S, A)}$$

$$= \frac{P(\Lambda \not\subseteq S \mid \Lambda_1 \notin S, A)P(\Lambda_1 \notin S \mid A)}{P(\Lambda \not\subseteq S \mid A)}$$

$$= \frac{P(\Lambda \not\subseteq S \mid \Lambda_1 \notin S, A)\left(\frac{k}{n}\right)}{1 - \frac{A}{\binom{n}{2}}}$$

14

Further, we also have by Bayes's rule

$$
\begin{aligned}
P(\Lambda_2 \in S \mid \Lambda \not\subseteq S, \Lambda_1 \in S, A) &= \frac{P(\Lambda \not\subseteq S \mid \Lambda_2 \in S, \Lambda_1 \in S, A) P(\Lambda_2 \in S, \Lambda_1 \in S, A)}{P(\Lambda \not\subseteq S, \Lambda_1 \in S, A)} \\
&= \frac{P(\Lambda \not\subseteq S \mid \Lambda_2 \in S, \Lambda_1 \in S, A) P(\Lambda_2 \in S \mid \Lambda_1 \in S, A)}{P(\Lambda \not\subseteq S \mid \Lambda_1 \in S, A)} \\
&= \frac{P(\Lambda_3 \notin S \mid \Lambda_2 \in S, \Lambda_1 \in S, A) P(\Lambda_2 \in S \mid \Lambda_1 \in S, A)}{P(\Lambda \not\subseteq S \mid \Lambda_1 \in S, A)} \\
&= \frac{\left(1 - \frac{k-2}{n-2}\right)\left(\frac{k-1}{n-1}\right)}{P(\Lambda \not\subseteq S \mid \Lambda_1 \in S, A)}
\end{aligned}
$$

So,

$$
\begin{aligned}
P(|\Lambda \cap S| = 2 \mid \Lambda \not\subseteq S \text{ and } A) &= 3\left(\frac{P(\Lambda \not\subseteq S \mid \Lambda_1 \notin S, A)\left(\frac{k}{n}\right)}{1 - \frac{A}{\binom{n}{2}}}\right)\left(\frac{\left(1 - \frac{k-2}{n-2}\right)\left(\frac{k-1}{n-1}\right)}{P(\Lambda \not\subseteq S \mid \Lambda_1 \in S, A)}\right) \\
&= 3\,\frac{\left(1 - \frac{k-2}{n-2}\right)\left(\frac{k-1}{n-1}\right)\left(\frac{k}{n}\right)}{1 - \frac{A}{\binom{n}{2}}}
\end{aligned}
$$

Thus, we have

$$
P(\Lambda \subseteq S' \mid \Lambda \not\subseteq S, A) = 3\left(\frac{\left(1 - \frac{k-2}{n-2}\right)\left(\frac{k-1}{n-1}\right)\left(\frac{k}{n}\right)}{\left(1 - \frac{A}{\binom{n}{2}}\right)}\right)\left(\frac{k-2}{\binom{n}{2}}\right) = 3\left(\frac{M}{\binom{n}{2} - A}\right)
$$

, where $M = (n-k)\frac{\binom{k}{3}}{\binom{n}{3}}$. Hence,

$$
P(\Lambda \subseteq S' \mid A) = \left(1 - \frac{3(n-k)}{\binom{n}{2}}\right)\left(\frac{A}{\binom{n}{2}}\right) + 3\left(\frac{M}{\binom{n}{2}}\right)
$$

, and so

$$
\begin{aligned}
\mathbb{E}[A' - A \mid A] &= -A + \sum_{\Lambda \subseteq \mathbb{F}_n} P(\Lambda \subseteq S' \mid A) = -A + \left(1 - \frac{3(n-k)}{\binom{n}{2}}\right)(A) + 3\,(M) \\
&= -\left(\frac{3(n-k)}{\binom{n}{2}}\right)(A) + 3(n-k)\frac{\binom{k}{3}}{\binom{n}{3}} \\
&= -\frac{3(n-k)}{\binom{n}{2}}\left(A - \binom{n}{2}\frac{\binom{k}{3}}{\binom{n}{3}}\right) \\
&= -\frac{3(n-k)}{\binom{n}{2}}(A - \mu).
\end{aligned}
$$

Thus, $\mathbb{E}[A' - A|A] = -\lambda(A - \mu)$, where $\lambda = \frac{3(n-k)}{\binom{n}{2}}$. Hence, $\mathbb{E}[W' - W|W] = \mathbb{E}[\frac{A'-\mu}{\sigma} - \frac{A-\mu}{\sigma}] = \frac{1}{\sigma}\mathbb{E}[A' - A] = -\lambda\left(\frac{A-\mu}{\sigma}\right) = -\lambda W$.

$\square$

Thus, using the method of exchangeable pairs, we can bound the Wasserstein distance $Wass(W, Z)$, where Z is a standard Gaussian random variable:

$$Wass(W, Z) \leq \sqrt{\left(\frac{2}{\pi}\right)\text{Var}(\mathbb{E}[\frac{1}{2\lambda}(W' - W)^2|W])} + \frac{1}{3\lambda}\mathbb{E}[|W' - W|^2]$$

. Now $\mathbb{E}[\frac{1}{2\lambda}(W' - W)^2|W] = \mathbb{E}[\frac{\binom{n}{2}}{6(n-k)}(W' - W)^2|W] = \frac{\binom{n}{2}}{6(n-k)}(\mathbb{E}[(W')^2|W] - 2\mathbb{E}[WW'|W] + W^2)$.

5.1 Attempt to get LLT for $A_{n,k}$

The main issue that we run into when trying to apply the same technique used to show LLTs for T_n and D_n to $A_{n,k}$ is that we cannot find an event to condition on such that $A_{n,k}$ will depend

6 General Framework for no LLT for other R.V.'s

We want to develop an overarching framework for what properties a random variable X_n must have so that it follows a CLT but not an LLT. Let X_n be conditioned on some event with values denoted by k.

Let gap $g = (\mu_{k+1} - \mu_k) \sim f(n, k) = \Omega(\sigma_{n,k}\sigma_n^\epsilon)$ for some $\epsilon > 0$. We proceed to show the following short proposition, which establishes a condition for which X_n does not satisfy a LLT:

Proposition 6.1. *If we pick x such that $|x - \mu_n| = \Omega(f(n, k)) = \omega(\sigma_{n,k}) \, \forall k$, then $|P(X_n = x) - \mathcal{N}_n(x)| = \Omega(\frac{1}{\sigma_n})$.*

Proof. We have

$$P(X_n = x) = \sum_{k=1}^{j} P(Y_k)P(X_n = x|Y)k) \leq \sum_{k=1}^{j} P(|X_n - \mu_k| \geq |x - \mu_k|)P(Y_k) \leq \sum_{k=1}^{j} \left(\frac{\sigma_{n,k}}{|x - \mu_k|}\right)^2 P(Y_k)$$

by Chebyshev's Inequality. Now $\left(\frac{\sigma_{n,k}}{|x-\mu_k|}\right) = o(\frac{1}{\sigma_n^\epsilon})$, so $\left(\frac{\sigma_{n,k}}{|x-\mu_k|}\right)^2 = o(\frac{1}{\sigma_n})$. Now $\mathcal{N}_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n}e^{-\left(\frac{x-\mu_n}{\sigma_n}\right)^2/2} = O(\frac{1}{\sigma_n})$. Hence, $|P(X_n = x) - \mathcal{N}_n(x)| = \Omega(\frac{1}{\sigma_n})$ and thus X_n does not follow a LLT.

$\square$

7 Attempt to prove no LLT for A_n

To show that there is no LLT for A_n, our strategy is to choose a point x in the middle between two $A_{n,k}$ distributions and showing that $|P(A_n = x) - \mathcal{N}_n(x)| = \Omega(\frac{1}{\sigma_n})$, where $\mathcal{N}_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{((x-\mu_n)/\sigma_n)^2/2}$.

Set $p = \frac{1}{2}$, and let $x \geq \mu_{n,k} \ \forall k \leq j$ and $x \leq \mu_{n,k} \ \forall k > j$. Then we have

$$
\begin{aligned}
P(A_n = x) &= \sum_{k=0}^{n} P(A_n = x|\ |S| = k) P(|S| = k) \\
&= \sum_{k=0}^{n} P(A_{n,k} = x) \cdot \frac{\binom{n}{k}}{2^n} \\
&\leq \sum_{k=0}^{j} P(A_{n,k} \geq x) \cdot \frac{\binom{n}{k}}{2^n} + \sum_{k=j+1}^{n} P(A_{n,k} \leq x) \cdot \frac{\binom{n}{k}}{2^n}
\end{aligned}
$$

Now we must sufficiently bound the tails of $A_{n,k}$. We know that $\mu_{A_n} = \frac{1}{8}\binom{n}{2} \sigma_{A_n}^2 = \Theta(n^3), \mu_{n,k} = \frac{3}{n-2}\binom{k}{3}$, and $\sigma_{A_{n,k}}^2 = \frac{k^3(n-k)^3}{2n^4}$, but we do not know anything about the shape of the distribution of $A_{n,k}$. Thus, we first try using Chebyshev's Inequality...

References

[1] Justin Gilmer and Swastik Kopparty. A local central limit theorem for the number of triangles in a random graph. *ArXiv e-prints*, November 2014.

[2] Jason Fulman. Stein's Method and Non-Reversible Markov Chains. *ArXiv e-prints*, December 1997.

[3] Sourav Chatterjee. Lecture Notes in Stein's method and applications. August 2007. Retrieved August 14, 2018 from https://statweb.stanford.edu/~souravc/Lecture1.pdf.