

Quantitative and Local Central Limit Theorems

Bryce Cai, Annie Chen, Ben Heller, Eyob Tsegaye

August 29, 2018

Abstract

We prove a quantitative local limit theorem for the number of descents in a random permutation. Our proof uses a conditioning argument and is based on bounding the characteristic function $\phi(t)$ of the number of descents.

We also establish a central limit theorem for the number of 3-term arithmetic progressions (3-APs) in a random subset of $\mathbb{Z}/n\mathbb{Z}$. We conjecture that there is no local limit theorem for 3-APs, 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 the same order as their standard deviations such that the constant multiple between them is not very large.

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. Our goal is to understand to what degree such a sequence converges to a normal distribution. 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 normalize by considering 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. 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

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

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, $\mathbb{P}(X_n < \frac{n}{2}) = \mathbb{P}(Y_n < 0) \rightarrow \mathbb{P}(Z < 0) = \frac{1}{2}$. To satisfy this, we would need

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

but this is not meaningful since both terms go to 0 trivially. 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

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

uniformly in k . If this condition holds, the sequence $\{X_n\}$ satisfies a **local limit theorem**. Equivalently, we require

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

Geometrically, this is equivalent since when the Gaussian is normalized, its height is scaled by σ_n . If there is a bound on this last rate of convergence, the local limit theorem is said to be *quantitative*. Equivalently, there is a quantitative bound if the rate of convergence of the first difference is better than $o(1/\sigma_n)$.

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 j if $\pi(j) > \pi(j+1)$. The number of descents in π , denoted $D(\pi)$, is the number of such indices j . We define the indicator random variable for the j th descent

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

so the number of descents can be written

$$D_n = \sum_{j=1}^{n-1} X_j.$$

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

$$X_j = \begin{cases} 1 & a_j > a_{j+1} \\ 0 & \text{else} \end{cases}.$$

With this method of sampling, we can compute some basic facts about D . At each index j , there is either a descent or an ascent ($\pi(j) < \pi(j+1)$) and these both occur with equal probability, so $\mathbb{E}[X_j] = \frac{1}{2}$. Therefore, $\mathbb{E}[D] = \frac{n-1}{2}$. With similar reasoning as in computing

the expectation of X_j , we get $\mathbb{E}[X_j X_{j+1}] = \frac{1}{3!} = \frac{1}{6}$. Also, it is important to note that X_j is independent of all other X_k except for X_{j-1} and X_{j+1} . Now we compute

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

In Section 3, we establish a quantitative local limit theorem for D_n :

Theorem 1.1.

$$\left| P(D_n = x) - \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\left(\frac{x-\mu_n}{\sigma_n}\right)^2/2} \right| = O(n^{-1+\varepsilon}) = O\left(\frac{1}{\sigma_n} n^{-\frac{1}{2}+\varepsilon}\right).$$

1.2 3-term arithmetic progressions

The other random variable of interest is the number of 3-term arithmetic progressions (3-APs) in a random subset of $\mathbb{Z}/n\mathbb{Z}$, with the presence of each element determined by a coinflip. We require n odd and not a multiple of 3 to avoid issues of divisibility. For $1 \leq i \leq n$, 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 3-term 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 expectation of A_n is $\mathbb{E}[A_n] = \frac{1}{8} \binom{n}{2}$. The variance of A_n is computed in section 4.

In section 4, we prove a central limit theorem for A_n , stated as follows:

Theorem 1.2.

$$\left| P\left(\frac{A_n - \mu_n}{\sigma_n} \leq x\right) - \mathbb{P}(Z \leq x) \right| = O(n^{-1/4}).$$

Based on experimental results and heuristics in section 4, we conjecture that there is no local limit theorem for A_n .

Conjecture 1.3.

$$\left| P(A_n = x) - \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\left(\frac{x-\mu_n}{\sigma_n}\right)^2/2} \right| \neq o\left(\frac{1}{\sigma_n}\right).$$

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 and $\mathcal{N}(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$ be the standard normal probability density function. Gilmer and Kopparty [1] establish the existence of an LLT for T_n , so T_n is pointwise approximated by a discrete Gaussian distribution.

Theorem 2.1. (*Gilmer, Kopparty; 2014*) *Uniformly for all $k \in \mathbb{Z}$,*

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

We note that this is a qualitative result, where the error is $o(n^{-2}) = o(\frac{1}{\sigma})$. Berkowitz [2] 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; 2017*) *Uniformly for all $k \in \mathbb{Z}$,*

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

2.1.1 Methods used to establish an LLT for T_n

Here we summarize the methods used by Gilmer and Kopparty [1] 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 for all $y \in L$,*

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

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

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

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

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

For any constant A , we can write

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

Since we have a CLT, the first integral on the right goes to 0. Hence, we have reduced the problem of proving an LLT to one of sufficiently bounding the characteristic function $|\phi_n(t)|$. To get a quantitative LLT, we must also bound $\left| \phi_n(t) - e^{-t^2/2} \right|$ for small values of 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. The full proof can be found in [1].

Using the p -biased Fourier basis, Berkowitz [2] 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 in order to bound $\left| \phi_n(t) - e^{-t^2/2} \right|$ for small values of t , although we conjecture that there is no bound for larger t .

3 Descents in a permutation

Let $\pi \in S_n$ be a permutation. Define $D(\pi) = |\{1 \leq j < n \mid \pi(j+1) < \pi(j)\}|$ 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 [3] and in this section, we prove Theorem 1.1, a local limit theorem. In the proof, we will apply the following bound.

Lemma 3.1. (*Gilmer, Kopparty; 2014*) *Let B be a Bernoulli random variable that is 1 with probability p . Then*

$$|e^{i\theta B}| \leq 1 - 8p(1-p) \cdot \left\| \frac{\theta}{2\pi} \right\|^2$$

where $\|x\|$ is the closest integer to x . As a result, for $|\theta| < \pi$,

$$|e^{i\theta B}| \leq 1 - 8p(1-p) \cdot \left(\frac{\theta}{2\pi} \right)^2.$$

As covered in 1.1, we write $D_n = \sum_{j=1}^{n-1} X_j$, where X_j is the indicator random variable for $a_j > a_{j+1}$. Each X_j depends only on a_j and a_{j+1} , so X_j is independent of all other X_k except for X_{j-1} and X_{j+1} . This lemma computes the dependence:

Lemma 3.2. *Given the values of X_{j-1} and X_{j+1} , the distribution of X_j does not depend on j or n .*

Proof. The descents X_{j-1} to X_{j+1} depend only on a_{j-1} to a_{j+2} . Every combination of values of a_{j-1} to a_{j+2} is equally likely, so it suffices to consider the case $n = 4$, $j = 2$.

Although it is not necessary for the local limit theorem, the exact distribution of X_j can be computed. In the $n = 4$ case, each pair of values of X_1 and X_3 appears in exactly 6

permutations. The following probabilities are established by looking at all 6 cases for each pair.

$$\begin{aligned}\mathbb{P}(X_i = 1 \mid X_{j-1} = 1, X_{j+1} = 1) &= 1/6, \\ \mathbb{P}(X_i = 1 \mid X_{j-1} = 1, X_{j+1} = 0) &= 1/2, \\ \mathbb{P}(X_i = 1 \mid X_{j-1} = 0, X_{j+1} = 1) &= 1/2, \\ \mathbb{P}(X_i = 1 \mid X_{j-1} = 0, X_{j+1} = 0) &= 5/6.\end{aligned}$$

□

Theorem 3.3. *The sequence of random variables D_n satisfies a local limit theorem. Quantitatively,*

$$\mathbb{P}(D_n = x) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-((x-\mu_n)/\sigma_n)^2/2} + O(n^{-1+\varepsilon}).$$

Proof. To bound the characteristic function of D_n , we condition on the values of X_j for odd j . For simplicity, assume n is even. (When n is odd, one must also observe that $\mathbb{P}(X_j = 1 \mid X_{j-1} = 1) = 1/3$ and $\mathbb{P}(X_j = 1 \mid X_{j-1} = 0) = 2/3$ are constant.) The even X_j are independent of each other, so conditioned on the odd X_j , D_n is a sum of independent random variables. After conditioning, we compute

$$\begin{aligned}|\mathbb{E}[e^{itD_n/\sigma}]| &= \left| \mathbb{E}[e^{it(C + \sum_{\text{odd } j} X_j)/\sigma}] \right| \\ &= \prod_{\text{odd } j} |\mathbb{E}[e^{itX_j/\sigma}]| \\ &\leq \prod_{\text{odd } j} \left(1 - 8p_j(1-p_j) \left(\frac{t}{2\pi\sigma} \right)^2 \right)\end{aligned}$$

by Lemma 3.1, where p_j is the probability associated to X_j by the given values of the odd X_j . But there are only finitely many possible values of p_j (4 of them), so there is some value p that maximizes the quantity above, and we can bound

$$\begin{aligned}|\mathbb{E}[e^{itD_n/\sigma}]| &\leq \prod_{\text{odd } j} \left(1 - 8p(1-p) \left(\frac{t}{2\pi\sigma} \right)^2 \right) \\ &\leq \prod_{\text{odd } j} (1 - \Theta(t^2/n)) \\ &= (1 - \Theta(t^2/n))^{n/2} \\ &\leq e^{-\Theta(t^2)/2} \\ &= e^{-\Theta(t^2)}.\end{aligned}$$

Now we obtain the final bound for the local limit theorem. As shown in [3], we have a central limit theorem

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

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

Thus,

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

For a characteristic function $\phi(t)$ for a variance X , using integration by parts we can write

$$\phi(t) = \int_{\mathbb{R}} e^{itx} \mathbb{P}(X = x) dx = -it \int_{\mathbb{R}} e^{itx} \mathbb{P}(X \leq x) dx.$$

Hence, we have the following (where $|\dots|$ is a repetition of the same integrand):

$$\begin{aligned} \left| \phi_n(t) - e^{-t^2/2} \right| &\leq |t| \int_{\mathbb{R}} \left| \mathbb{P}\left(\frac{D_n - \mu}{\sigma} < y\right) - \mathbb{P}(Z < y) \right| dy \\ &\leq |t| \int_{|y| > k\sigma} \left| \mathbb{P}\left(\frac{D_n - \mu}{\sigma} < y\right) - \mathbb{P}(Z < y) \right| dy + |t| \int_{|y| \leq k\sigma} |\dots| dy \\ &\leq |t| \int_{|y| > k\sigma} \left(\frac{1}{2y^2} + \frac{e^{-\Theta(y^2)}}{y} \right) dy + |t| 2k\sigma \sqrt{\frac{12}{n}}. \end{aligned}$$

Take $k = O(n^{-\frac{1}{2}+\epsilon})$. Then since $\sigma = \Theta(n)$, we have $\left| \phi_n(t) - e^{-t^2/2} \right| \leq |t| O(n^{-\frac{1}{2}+\epsilon})$.

Thus, for any $\epsilon > 0$, we compute

$$\begin{aligned} \int_{-\pi\sigma}^{\pi\sigma} \left| \phi_n(t) - e^{-t^2/2} \right| dt &\leq \int_{-n^\epsilon}^{n^\epsilon} \left| \phi_n(t) - e^{-t^2/2} \right| dt + \int_{n^\epsilon < |t| < \pi\sigma} (|\phi_n(t)| + |e^{-t^2/2}|) dt \\ &\leq \int_{-n^\epsilon}^{n^\epsilon} \left| \phi_n(t) - e^{-t^2/2} \right| dt + \int_{n^\epsilon < |t| < \pi\sigma} e^{-\Theta(t^2)} dt \\ &= O(n^{-\frac{1}{2}+\epsilon}). \end{aligned}$$

□

4 3-term arithmetic progressions

Throughout this section, we define a random variable A_n to be the number of 3-term arithmetic progressions in a randomly chosen subset $S \subset \mathbb{Z}/n\mathbb{Z}$, where each element in $\mathbb{Z}/n\mathbb{Z}$ has probability $\frac{1}{2}$ (also denoted 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}_i is 1 with probability p and 0 with probability $1 - p$. Further, the random variable A_n can be thought of as a function $A_n : \{0, 1\}^n \rightarrow \mathbb{N}$.

4.1 Variance of A_n

In this section, we describe a p -biased Fourier basis for functions on the probability space, exactly analogous to the method of Berkowitz [2], which we use to find the variance of A_n and will use in the next section to try to prove an LLT for A_n . In order to do this, we first define $\chi_i : \{0, 1\}^n \rightarrow \mathbb{R}$ by

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

so that χ_i is a normalized version of \mathbf{x}_i . Further, we can extend this to define, for an arbitrary set $S \subseteq \mathbb{Z}/n\mathbb{Z}$,

$$\chi_S := \prod_{i \in S} \chi_i.$$

Note that if we take the inner product of two functions $f, g : \{0, 1\}^n \rightarrow \mathbb{R}$ to be $\mathbb{E}[fg]$, then $\{\chi_S \mid S \subseteq \mathbb{Z}/n\mathbb{Z}\}$ forms an orthonormal basis for functions on our probability space. Then if we define the Fourier transform $\hat{f} : \mathbb{Z}/n\mathbb{Z} \rightarrow \mathbb{R}$ of an arbitrary function $f : \{0, 1\}^n \rightarrow \mathbb{R}$ by

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

from the orthonormality of our basis we get

$$f(\mathbf{x}) = \sum_{S \subseteq \mathbb{Z}/n\mathbb{Z}} \hat{f}(S) \chi_S(\mathbf{x}).$$

We will use this expansion to calculate the variance of A_n and bound the pointwise distance of the characteristic function from that of the discrete Gaussian for small t .

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

$$Z = Z_n := \frac{A_n - \mu_n}{\sigma_n}$$

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 $\hat{A}_n(S)$, we will first note that there are $\binom{n}{2}$ possible (non-trivial) 3-term arithmetic progressions in $\mathbb{Z}/n\mathbb{Z}$. 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 3-term 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

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

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

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

$$\sigma^2 = \sum_{S \neq \emptyset} \hat{A}_n(S)^2.$$

We now work with the 3-term arithmetic progression indicator functions 1_Λ a bit more. We use $i \in \Lambda$ to denote that i is a term in the 3-term arithmetic progression Λ . Therefore, the indicator function can be expressed as

$$\begin{aligned} 1_\Lambda(\mathbf{x}) &= \prod_{i \in \Lambda} \mathbf{x}_i = \prod_{i \in \Lambda} \left(\sqrt{p(1-p)} \chi_i + p \right) \\ &= p^3 + p^2 \sqrt{p(1-p)} \sum_{i \in \Lambda} \chi_i + p^2(1-p) \sum_{i_1 \neq i_2 \in \Lambda} \chi_{\{i_1, i_2\}} + p^{3/2}(1-p)^{3/2}. \end{aligned}$$

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

$$\begin{aligned} A_n &= p^3 \binom{n}{2} + \frac{3}{2}(n-1)p^2 \sqrt{p(1-p)} \sum_{i \in \mathbb{Z}/n\mathbb{Z}} \chi_i \\ &\quad + 3p^2(1-p) \sum_{\{i_1, i_2 | i_1 \neq i_2\} \subset \mathbb{Z}/n\mathbb{Z}} \chi_{\{i_1, i_2\}} + \sum_{\Lambda} p^{3/2}(1-p)^{3/2} \chi_\Lambda. \end{aligned}$$

Thus, we have the Fourier Transform of A_n :

$$\hat{A}_n(S) = \begin{cases} p^3 \binom{n}{2} & \text{if } S = \emptyset \\ \frac{3}{2}(n-1)p^2 \sqrt{p(1-p)} & \text{if } |S| = 1 \\ 3p^2(1-p) & \text{if } S = \{i_1, i_2\} \\ p^{3/2}(1-p)^{3/2} & \text{if } S = \Lambda \\ 0 & \text{else} \end{cases}$$

and we can use Parseval's Theorem to give us the variance of A_n :

$$\begin{aligned} \sigma_n^2 &= \sum_{\substack{S \subset \mathbb{Z}/n\mathbb{Z} \\ S \neq \emptyset}} \hat{A}_n(S)^2 \\ &= \sum_{i \in \mathbb{Z}/n\mathbb{Z}} \left(\frac{3}{2}(n-1)p^2 \sqrt{p(1-p)} \right)^2 + \sum_{\{i_1, i_2\} \in \mathbb{Z}/n\mathbb{Z}} (3p^2(1-p))^2 + \sum_{\Lambda} (p^{3/2}(1-p)^{3/2})^2 \\ &= \frac{9}{4}n(n-1)^2 p^5(1-p) + 9 \binom{n}{2} p^4(1-p)^2 + \binom{n}{2} p^3(1-p)^3 \\ &= \Theta(n^3). \end{aligned}$$

This means that $\sigma_n = \Theta(n^{3/2})$.

4.2 A CLT for 3-term APs

In this section, we establish a central limit theorem for the number of 3-term arithmetic progressions in $\mathbb{Z}/n\mathbb{Z}$, A_n .

Theorem 4.1.

$$\left| P\left(\frac{A_n - \mu_n}{\sigma_n} \leq x\right) - \Phi(x) \right| = O(n^{-1/4}).$$

We prove this theorem using the method of dependency graphs detailed in Chatterjee's notes [4] (Lecture 6). This method bounds the Wasserstein distance of a distribution from the normal based on the dependency graph. We define the Wasserstein distance as such, where Ω is the probability space:

$$\text{Wass}(\mu, \nu) = \sup \left\{ \left| \int f d\mu - \int f d\nu \right| \mid f : \Omega \rightarrow \mathbb{R} \text{ is 1-Lipschitz} \right\}.$$

The Wasserstein distance is a metric that is stronger than the L_1 distance of the cumulative distribution functions of two distributions [4]. In other words, a vanishing bound on the Wasserstein distance between a distribution and the standard normal would immediately establish a central limit theorem for that distribution.

More formally, we define the Kolmogorov distance as:

$$\text{Kolm}(\mu, \nu) = \sup_{x \in \mathbb{R}} |\mu((-\infty, x]) - \nu((-\infty, x])|.$$

So the Kolmogorov distance is the largest difference in the cumulative distribution functions. We write the Wasserstein distance between two random variables $X \sim \mu$ and $Y \sim \nu$ as $\text{Wass}(X, Y) = \text{Wass}(\mu, \nu)$, and similarly, $\text{Kolm}(X, Y) = \text{Kolm}(\mu, \nu)$. A useful bound for the Kolmogorov distance between a random variable and the standard normal is stated and proved in [4]:

Lemma 4.2 (Chatterjee, 2007). *For a pair of r.v.'s W, Z where $Z \sim N(0, 1)$,*

$$\text{Kolm}(W, Z) \leq \sqrt{\frac{2}{\pi}} \text{Wass}(W, Z).$$

Thus, the distance between the cdf of a random variable W and the standard normal, Z is bounded by the square root of the Wasserstein distance between W and Z (with a constant factor). So the problem of deriving a central limit theorem for W is reduced to bounding $\text{Wass}(W, Z)$.

A formal statement of the method of dependency graphs is as follows. Suppose there are a collection of random variables, $\{X_i \mid i \in V\}$, indexed by the vertices of a graph $G = (V, E)$ such that $(i, j) \in E$ iff X_i, X_j are dependent. Such a graph is referred to as a dependency graph. Let $D = 1 + \Delta(G)$, where $\Delta(G)$ is the maximum degree of G . We have the following:

Lemma 4.3 (Chatterjee, 2007). Suppose that $\mathbb{E}[X_i] = 0$, $\sigma^2 = \text{Var}(\sum X_i)$, $W = \frac{\sum X_i}{\sigma}$ and $Z \sim N(0, 1)$. Then,

$$\text{Wass}(W, Z) \leq \frac{4}{\sqrt{\pi}\sigma^2} \sqrt{D^3 \sum \mathbb{E}|X_i|^4} + \frac{D^2}{\sigma^3} \sum \mathbb{E}|X_i|^3.$$

For the problem of 3-term arithmetic progressions, we will write Λ to denote a particular 3-term AP. We define $V = \{\Lambda \subseteq \mathbb{Z}/n\mathbb{Z} \mid \Lambda \text{ is a 3-AP}\}$ and $X_\Lambda = 1_{\Lambda \in S} - \mathbb{E}[1_{\Lambda \in S}] = 1_{\Lambda \in S} - p^3$. Note that there are $\binom{n}{2}$ 3-term APs in $\mathbb{Z}/n\mathbb{Z}$, not counting degenerate 3-term arithmetic progressions, so $|V| = \binom{n}{2}$. Further, note that X_{Λ_1} and X_{Λ_2} are independent iff $\Lambda_1 \cap \Lambda_2 = \emptyset$. We have the following properties:

Lemma 4.4. The degree of the dependency graph is $D = O(n)$.

Proof.

$$\begin{aligned} \deg(\Lambda_1) &= |\{\Lambda_2 \mid \Lambda_1 \cap \Lambda_2 \neq \emptyset\}| \\ &= |\{\Lambda_2 \mid (\Lambda_2)_1 \in \Lambda_1 \vee (\Lambda_2)_2 \in \Lambda_1 \vee (\Lambda_2)_3 \in \Lambda_1\}| \\ &\leq 3 |\{\Lambda_2 \mid (\Lambda_2)_1 \in \Lambda_1\}| \\ &= \frac{9}{2}(n-1). \end{aligned}$$

So $\Delta(G) \leq \frac{9}{2}(n-1)$ and $D \leq \frac{9}{2}(n-1) + 1$. Thus, $D = O(n)$. \square

Proposition 4.5. The m th moment of X_Λ is bounded by $\mathbb{E}|X_\Lambda|^m \leq 1$ for all $m \geq 1$.

Proof. X_Λ is a random variable which takes values at $1 - p^3$ and p^3 , so its absolute value is bounded by 1. Thus, its m th absolute moment is also bounded by 1. \square

Furthermore, we note that $\text{Var}(\sum X_i) = \text{Var}(\sum 1_\Lambda - p^3) = \sigma_n^2$ and $W = \frac{\sum X_i}{\sigma} = \frac{A_n - \mu_n}{\sigma_n}$. We now proceed with the proof of theorem 4.1.

Proof. (Theorem 4.1)

Applying lemma 4.3 to $W = \frac{A_n - \mu_n}{\sigma_n}$ and using the bound on the Kolmogorov distance from 4.2 we get:

$$\begin{aligned} \left| P\left(\frac{A_n - \mu_n}{\sigma_n} \leq x\right) - \Phi(x) \right| &\leq \sqrt{\frac{2}{\pi} \text{Wass}(W, Z)} \\ &\leq \sqrt{\frac{2}{\pi}} \sqrt{\frac{4}{\sqrt{\pi}\sigma^2} \sqrt{D^3 \sum \mathbb{E}|X_\Lambda|^4} + \frac{D^2}{\sigma^3} \sum \mathbb{E}|X_\Lambda|^3} \end{aligned}$$

Now, using the established properties that $|V| = \binom{n}{2} \leq n^2$, $D = O(n)$, $\mathbb{E}|X_\Lambda|^m \leq 1$ for all $m \geq 1$, and $\sigma^2 = \Theta(n^3)$, we get

$$\begin{aligned} \left| P\left(\frac{A_n - \mu_n}{\sigma_n} \leq x\right) - \Phi(x) \right| &= O\left(\sqrt{\frac{1}{n^3} \sqrt{n^3 n^2} + \frac{n^2}{n^{9/2}} n^2}\right) \\ &= O(n^{-1/4}). \end{aligned}$$

\square

4.3 Attempt to show LLT for A_n

We recall from section 2.1.1 that since we have a CLT, we can reduce the problem of proving an LLT to one of bounding $\left| \phi_Z(t) - e^{-t^2/2} \right|$. Using the p -biased Fourier basis developed above, we proceed to show the following bound for small values of t using the same method as Berkowitz [2]:

Proposition 4.6. *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

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

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

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

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

$$\mathbb{E}[|Y|] \leq \sqrt{\mathbb{E}[|Y|^2]} = \sqrt{\text{Var}(Y)} = \sqrt{\sum_{k \in \mathbb{Z}/n\mathbb{Z}} (\hat{Z}(k) - Q)^2 + \sum_{|S| \geq 2} \hat{Z}^2(S)}.$$

Now using our above calculation of the variance σ_n^2 of A_n ,

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

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

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

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

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

□

Remark 4.7. The bound we find for the absolute distance between the characteristic function of Z and that of the standard normal is a decent bound for small t if we want to show an LLT for A_n . However, we are unable to prove any meaningful bound for larger t , and below we explore the reasons for why this is impossible.

4.4 Experimental results and conjectures

As we could not decrease the bound achieved in the previous section and prove a bound for larger values of t , we decided to simulate the number of 3-term arithmetic progressions in a random subset of $\mathbb{Z}/n\mathbb{Z}$ in order to verify whether or not an LLT actually holds and found that one does not.

As figure 1 demonstrates, the number of 3-term 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 oscillations in the histogram that are close together are spaced out almost evenly. Furthermore, the oscillations themselves appear fairly Gaussian, 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.

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 size of the subset of $\mathbb{Z}/n\mathbb{Z}$.

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, $\sqrt{\text{Var}(A_{n,k})}$. 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

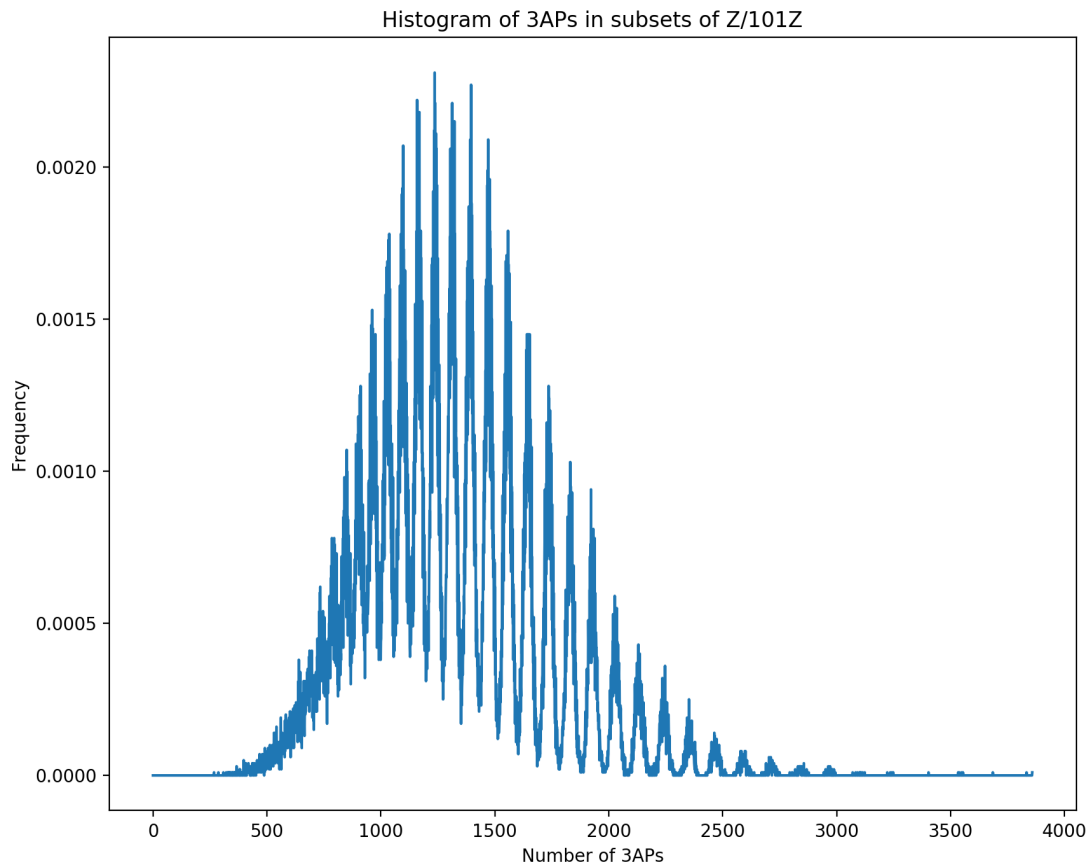


Figure 1: Histogram of the number of 3-term arithmetic progressions in 100000 random subsets of $\mathbb{Z}/101\mathbb{Z}$.

in the interval, $\mathbb{P}(A_{n,k} = x)$ and $\mathbb{P}(A_{n,k+1} = x)$ are both very small. In particular, we suspect that those two probabilities are significantly smaller than the Gaussian approximation of $\mathbb{P}(A_n = x)$ would suggest.

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

4.4.1 Continuous sets

One thing that comes to 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 existence of each element 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 3-term arithmetic progressions \bar{A}_n is still defined in a similar manner,

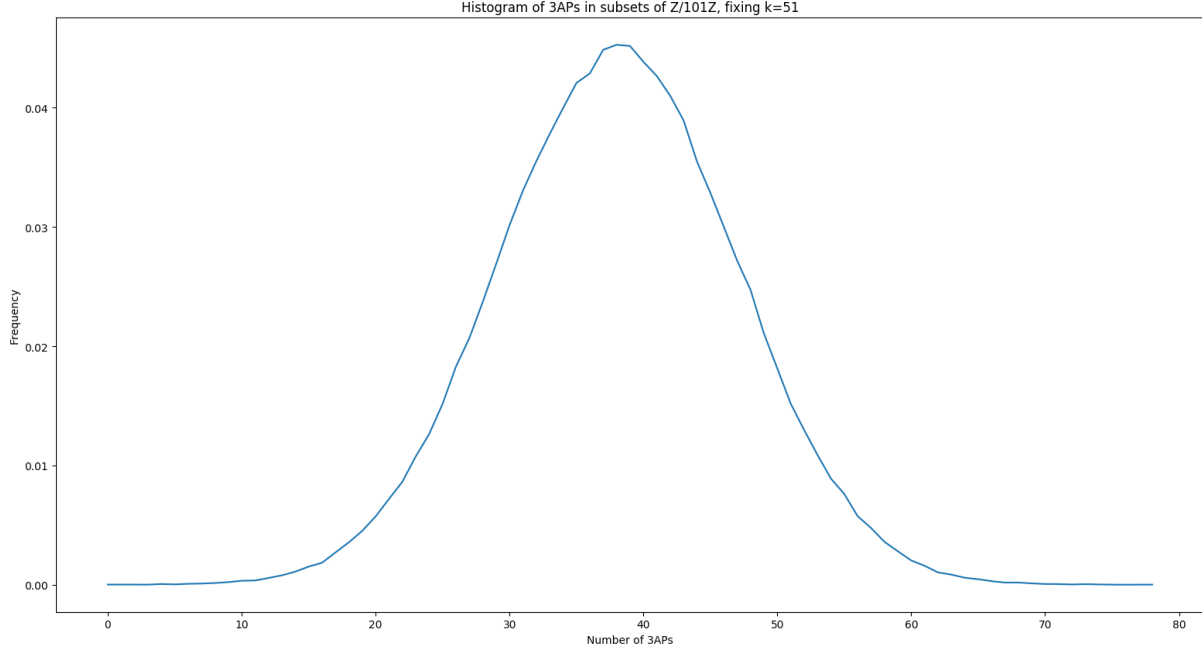


Figure 2: Histogram of the number of 3-term arithmetic progressions in random subsets of $\mathbb{Z}/53\mathbb{Z}$ of size k for $k = 0$ to 53, separated by color and . Each k used 10000 samples. (The graph should extend to 1378 on the x axis, but there is very little content in the large k that are very far out so we zoomed in on the bulk of the distribution.)

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

We simulated this to see if it possibly followed the Gaussian distribution pointwise. From figure 3, we conjecture that it is very likely that there is an LLT for \bar{A}_n . However, we successfully can show that there is a CLT for \bar{A}_n . Namely, define $\bar{\mu}_n = \mathbb{E}[\bar{A}_n]$ and $\bar{\sigma}_n^2 = \text{Var}(\bar{A}_n)$. Then the following theorem holds:

Theorem 4.8.

$$\left| P \left(\frac{\bar{A}_n - \bar{\mu}_n}{\bar{\sigma}_n} \leq x \right) - \Phi(x) \right| = O(n^{-1/4}).$$

Notably, this is exactly the same bound achieved in Theorem 4.1. This CLT is a consequence of the fact that the dependency graph of the indicator functions 1_Λ does not change if we make the underlying x_i uniform on $[0, 1]$ instead of Bernoulli on $\{0, 1\}$. However, the following lemma is what gives us the exact same bound as for the standard case, A_n :

Lemma 4.9. $\bar{\sigma}_n^2 = \Theta(n^3)$.

Proof. First, note that

$$\bar{\mu}_n = \mathbb{E} \left[\sum_{\Lambda} 1_{\Lambda} \right] = \sum_{\Lambda} \mathbb{E}[1_{\Lambda}] = \frac{1}{8} \binom{n}{2}.$$

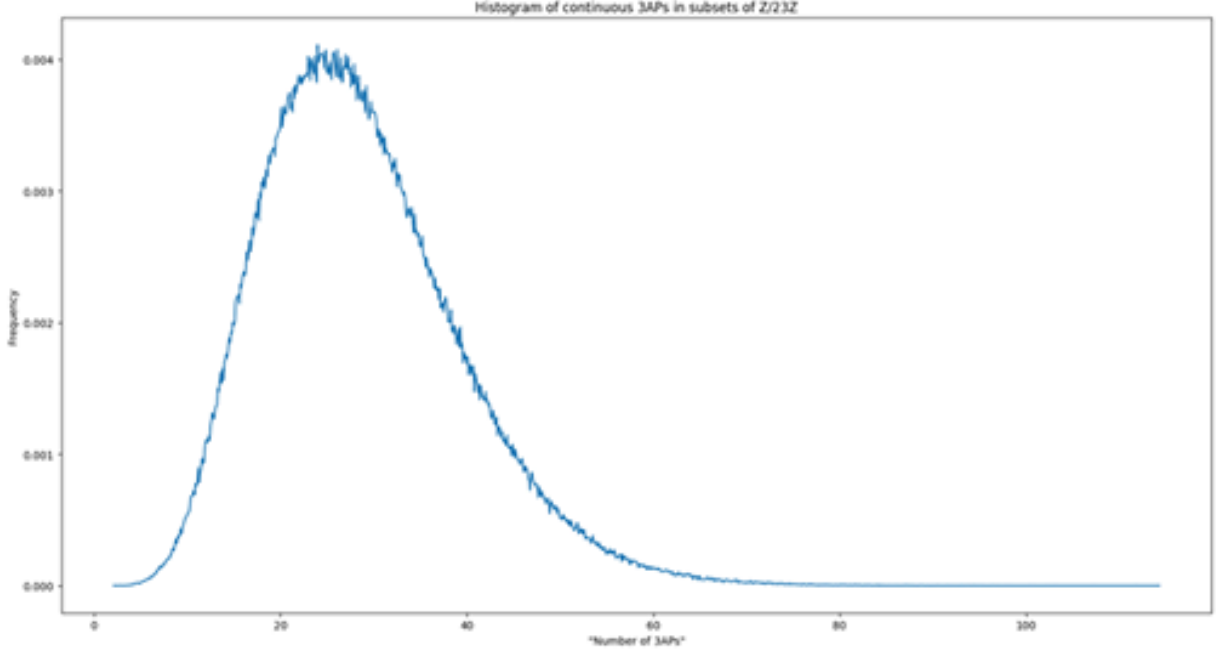


Figure 3: Number of 3-term "arithmetic progressions" in 1000000 random "continuous sets" of $\mathbb{Z}/23\mathbb{Z}$

Now, we can also calculate

$$\begin{aligned}
\mathbb{E} [\bar{A}_n^2] &= \mathbb{E} \left[\left(\sum_{\Lambda} 1_{\Lambda} \right)^2 \right] = \mathbb{E} \left[\sum_{\Lambda_1} \sum_{\Lambda_2} 1_{\Lambda_1} 1_{\Lambda_2} \right] \\
&= \mathbb{E} \left[\sum_{\Lambda_1} \sum_{\Lambda_2} \prod_{j \in \Lambda_1 \cup \Lambda_2} x_j \right] \\
&= \mathbb{E} \left[\sum_{\Lambda_1} \sum_{i=0}^3 \sum_{\substack{\Lambda_2 \text{ s.t.} \\ |\Lambda_1 \cap \Lambda_2| = i}} \prod_{j \in \Lambda_1 \cup \Lambda_2} x_j \right] \\
&= \sum_{\Lambda_1} \sum_{i=0}^3 \sum_{\substack{\Lambda_2 \text{ s.t.} \\ |\Lambda_1 \cap \Lambda_2| = i}} \left(\frac{1}{2} \right)^{6-i} \\
&= \sum_{\Lambda_1} \sum_{i=0}^3 \left(\frac{1}{2} \right)^{6-i} (\#\{\Lambda_2 \mid |\Lambda_1 \cap \Lambda_2| = i\}).
\end{aligned}$$

The particular quantity $\#\{\Lambda_2 \mid |\Lambda_1 \cap \Lambda_2| = i\}$ is examined further and calculated in section 6.2. For now we take the particular values we need as given and note that the quantity does not depend on our choice of Λ_1 .

$$\mathbb{E}[\bar{A}_n^2] = \binom{n}{2} \left(\frac{1}{8} + \frac{1}{16}(6) + \frac{1}{32} \left(\frac{9}{2}n - \frac{39}{2} \right) + \frac{1}{64} \left(\frac{1}{2}n^2 - 5n + \frac{25}{2} \right) \right).$$

Now, given that $\bar{\sigma}_n^2 = \mathbb{E}[\bar{A}_n^2] - \bar{\mu}_n^2$, we have

$$\bar{\sigma}_n^2 = \frac{1}{64} \binom{n}{2} \left(n - \frac{25}{54} \right) = \Theta(n^3).$$

□

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

In order to prove a CLT for $A_{n,k}$, we use the method of exchangeable pairs detailed in Chatterjee's notes [4] (Lecture 7). Like the method of dependency graphs used in section 4, this method bounds the Wasserstein distance of a distribution from the normal distribution.

Definition 5.1. (W, W') is an exchangeable pair of random variables if $(W, W') \stackrel{d}{=} (W', W)$.

The following is a formal statement of the method of exchangeable pairs.

Lemma 5.2. (Chatterjee, 2007) If (W, W') is an exchangeable pair, $\mathbb{E}[W^2] = 1$ and there exists $\lambda \in (0, 1)$ such that $\mathbb{E}[W - W'|W] = -\lambda W$, then

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

where $Z \sim N(0, 1)$.

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 in distribution to a normal distribution. We construct W' as follows. We have a random subset $S \subseteq \mathbb{Z}/n\mathbb{Z}$ with size k . Randomly choose two elements $I, J \in \mathbb{Z}/n\mathbb{Z}$, swap their status regarding their inclusion in S , and call the new subset S' (i.e. if $I \in S$ iff $J \in S'$ and $J \in S$ iff $I \in S'$). Now define A to be the number of 3-term arithmetic 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 3-term arithmetic progressions in S' and define W' to be $\frac{A' - \mu_{n,k}}{\sigma_{n,k}}$.

Remark 5.3. Let W and W' be an exchangeable pair, each with mean 0 and variance 1, and let Λ be a 3-term arithmetic progression in $\mathbb{Z}/n\mathbb{Z}$. Define $1_{\Lambda \subseteq S} = \begin{cases} 1 & \Lambda \subseteq S \\ 0 & \text{else} \end{cases}$.

Then $A = \sum_{\Lambda \subseteq \mathbb{Z}/n\mathbb{Z}} 1_{\Lambda \subseteq S}$. Also note that

$$\mu_{n,k} = \mathbb{E} \left[\sum_{\Lambda \subseteq \mathbb{Z}/n\mathbb{Z}} 1_{\Lambda \subseteq S} \right] = \sum_{\Lambda \subseteq \mathbb{Z}/n\mathbb{Z}} \mathbb{P}(\Lambda \in S) = \binom{n}{2} \frac{\binom{k}{3}}{\binom{n}{3}}.$$

In order to apply the method, we first must show that $\mathbb{E}[W' - W \mid W] = -\lambda W$ for some $\lambda \in (0, 1)$.

Lemma 5.4. *Let $\lambda = \frac{3(n-k)}{\binom{n}{2}}$. Then $\mathbb{E}[W' - W \mid W] = -\lambda W$.*

Proof. We have

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

Now

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

Note that

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

as there are A 3-term arithmetic progressions in S and $\binom{n}{2}$ 3-term arithmetic progressions total in $\mathbb{Z}/n\mathbb{Z}$.

We have

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

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

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

Finally, we examine

$$\begin{aligned} \mathbb{P}(\Lambda \subseteq S' \mid \Lambda \not\subseteq S, A) &= \sum_{i=0}^3 \mathbb{P}(\Lambda \subseteq S' \text{ and } |\Lambda \cap S| = i \mid \Lambda \not\subseteq S \text{ and } A) \\ &= \sum_{i=0}^2 \mathbb{P}(\Lambda \subseteq S' \mid \Lambda \not\subseteq S \text{ and } |\Lambda \cap S| = i \text{ and } A) \mathbb{P}(|\Lambda \cap S| = i \mid \Lambda \not\subseteq S \text{ and } A) \\ &= \mathbb{P}(\Lambda \subseteq S' \mid |\Lambda \cap S| = 2, \text{ and } A) \mathbb{P}(|\Lambda \cap S| = 2 \mid \Lambda \not\subseteq S \text{ and } A). \end{aligned}$$

We have $\mathbb{P}(\Lambda \subseteq S' \mid |\Lambda \cap S| = 2, \text{ and } A) = \frac{k-2}{\binom{n}{2}}$ 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 $S \setminus \Lambda$.

Substituting this in, we get

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

$$\begin{aligned} \mathbb{P}(|\Lambda \cap S| = 2 \mid \Lambda \not\subseteq S \text{ and } A) &= 3\mathbb{P}(\Lambda_1 \in S, \Lambda_2 \in S, \Lambda_3 \notin S \mid \Lambda \not\subseteq S \text{ and } A) \\ &= 3\mathbb{P}(\Lambda_3 \notin S \mid \Lambda_1 \in S, \Lambda_2 \in S, \Lambda \not\subseteq S, A) \\ &\quad \cdot \mathbb{P}(\Lambda_2 \in S \mid \Lambda \not\subseteq S, \Lambda_1 \in S, A) \mathbb{P}(\Lambda_1 \in S \mid \Lambda \not\subseteq S, A) \\ &= 3\mathbb{P}(\Lambda_2 \in S \mid \Lambda \not\subseteq S, \Lambda_1 \in S, A) \mathbb{P}(\Lambda_1 \in S \mid \Lambda \not\subseteq S, A). \end{aligned}$$

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

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

Further, we also have by Bayes's rule,

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

So,

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

Thus, we have

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

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

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

and so

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

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

Thus, we can bound the Wasserstein distance $\text{Wass}(W, Z)$, where Z is a standard Gaussian random variable using the method of exchangeable pairs. We have

$$\begin{aligned} \mathbb{E}\left[\frac{1}{2\lambda}(W' - W)^2 \mid W\right] &= \mathbb{E}\left[\frac{\binom{n}{2}}{6(n - k)}(W' - W)^2 \mid W\right] \\ &= \frac{\binom{n}{2}}{6(n - k)} (\mathbb{E}[(W')^2 \mid W] - 2\mathbb{E}[WW' \mid W] + W^2). \end{aligned}$$

6 Attempt to prove nonexistence of LLT for A_n

6.1 General framework for nonexistence of an LLT for other R.V.'s

We develop more general theorems for what properties a random variable X_n can have to ensure it does not follow an LLT. Let X_n (with mean μ_n and standard deviation σ_n) be conditioned on some event with value k , denoted by $X_{n,k}$. Let $\mu_{n,k} = \mathbb{E}[X_{n,k}]$ and $\sigma_{n,k}$ be the standard deviation of $X_{n,k}$.

The following theorem says that X_n does not satisfy a LLT when the distances between the conditioned distributions $X_{n,k}$ are of a sufficiently larger order than their standard deviations $\sigma_{n,k}$. We show that at a point, the distribution of X_n does not converge to the normal distribution.

Theorem 6.1. Suppose $|x - \mu_{n,k}| = \omega(\sigma_{n,k}\sqrt{\sigma_n})$. Then

$$|\mathbb{P}(X_n = x) - \mathcal{N}_n(x)| = \Omega\left(\frac{1}{\sigma_n}\right)$$

and there is no local limit theorem for $\{X_n\}$.

Proof. Since we make no assumptions about the concentration of $X_{n,k}$, the best we can do is use Chebyshev's inequality.

$$\begin{aligned} \mathbb{P}(X_n = x) &= \sum_{k=1}^j \mathbb{P}(X_n = x \mid Y_k) \mathbb{P}(Y_k) \\ &\leq \sum_{k=1}^j \mathbb{P}(|X_{n,k} - \mu_{n,k}| \geq |x - \mu_{n,k}|) \mathbb{P}(Y_k) \\ &\leq \sum_{k=1}^j \left(\frac{\sigma_{n,k}}{|x - \mu_{n,k}|} \right)^2 \mathbb{P}(Y_k) \\ &= \sum_{k=1}^j o\left(\frac{1}{\sigma_n}\right) \mathbb{P}(Y_k) \\ &= o\left(\frac{1}{\sigma_n}\right). \end{aligned}$$

□

In the previous theorem, the required distance between consecutive $X_{n,k}$ is rather large, and this is likely due to the inefficiency of using Chebyshev's inequality. If we assume a CLT on the $X_{n,k}$, we can use a better concentration inequality and require less of a distance between the $X_{n,k}$.

Theorem 6.2. Suppose $X_{n,k}$ follows a CLT and $|x - \mu_{n,k}| = \omega(\sigma_{n,k}\sigma_n^\epsilon)$. Then

$$|\mathbb{P}(X_n = x) - \mathcal{N}_n(x)| = \Omega\left(\frac{1}{\sigma_n}\right)$$

and there is no local limit theorem for $\{X_n\}$.

Proof. We have

$$\begin{aligned}
\mathbb{P}(X_n = x) &= \sum_{k=1}^j \mathbb{P}(Y_k) \mathbb{P}(X_n = x \mid Y_k) \\
&\leq \sum_{k=1}^j \mathbb{P}(|X_{n,k} - \mu_{n,k}| \geq |x - \mu_{n,k}|) \mathbb{P}(Y_k) \\
&\leq \sum_{k=1}^j \left(\frac{\sigma_{n,k}}{|x - \mu_{n,k}| \sqrt{2\pi}} e^{-\left(\frac{x - \mu_{n,k}}{\sigma_{n,k}}\right)^2 / 2} \right) \mathbb{P}(Y_k) \\
&= \sum_{k=1}^j o(\sigma_n^{-\epsilon}) e^{-o(\sigma_n^{2\epsilon})} \mathbb{P}(Y_k) \\
&= o(\sigma_n^{-\epsilon}) o(\sigma_n^{-1+\epsilon}) = o(\sigma_n^{-1}).
\end{aligned}$$

Thus, $\mathbb{P}(X_n = x) = o\left(\frac{1}{\sigma_n}\right)$. Now $\mathcal{N}_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\left(\frac{x - \mu_n}{\sigma_n}\right)^2 / 2} = \Theta\left(\frac{1}{\sigma_n}\right)$. Hence,

$$|\mathbb{P}(X_n = x) - \mathcal{N}_n(x)| = \Omega\left(\frac{1}{\sigma_n}\right)$$

and thus X_n does not follow a LLT. □

6.2 Variance of $A_{n,k}$

Lemma 6.3. *The variance of $A_{n,k}$ is $\sigma_{n,k}^2 = \frac{k^3(n-k)^3}{2n^4} + O(n)$.*

Proof. We have that $\sigma_{n,k}^2 = \mathbb{E}[A_{n,k}^2] - (\mathbb{E}[A_{n,k}])^2 = \mathbb{E}[A_{n,k}^2] - \left(\frac{n}{2}\right)\left(\frac{k}{n}\right) = \mathbb{E}[A_{n,k}^2] - \frac{3}{n-2}\binom{k}{3}$. We proceed to calculate $\mathbb{E}[A_{n,k}^2]$. Now $A_{n,k}^2 = (\sum_{\Lambda} 1_{\Lambda \subset S})^2 = \sum_{\Lambda_1} \sum_{\Lambda_2} 1_{\Lambda_1 \subset S} 1_{\Lambda_2 \subset S}$. So

$$\begin{aligned}
\mathbb{E}[A_{n,k}^2] &= \sum_{\Lambda_1} \sum_{\Lambda_2} \mathbb{E}[1_{\Lambda_1 \subset S} 1_{\Lambda_2 \subset S}] \\
&= \sum_{i=0}^3 \frac{\binom{k}{6-i}}{\binom{n}{6-i}} \sum_{\Lambda_1} \sum_{\substack{\Lambda_2 \text{ s.t.} \\ |\Lambda_1 \cap \Lambda_2| = i}} 1.
\end{aligned}$$

We now consider the quantity $\#\{\Lambda_2 \mid |\Lambda_1 \cap \Lambda_2| = i\}$. Since we know there are i elements pre-determined from Λ_1 , we take each subset of i elements from Λ_1 and perform the following procedure. Check how many arithmetic progressions contain those many numbers (which themselves come from an arithmetic progression), given the function $f : \mathbb{N} \times \{0, 1, 2, 3\} \rightarrow \mathbb{N}$ defined by $f(0) = \binom{n}{2}$, $f(1) = \frac{3}{2}(n-1)$, $f(2) = 3$, and $f(3) = 1$. In particular, we check $f(i)$. However, this overcounts the quantity we desire because it also counts arithmetic progressions whose intersection with Λ_1 is greater than 1. Thus, we check all the ways of adding another element from Λ_1 to our current subset of i elements and subtract out $f(i+1)$

arithmetic progressions. We apply continue applying inclusion-exclusion until we arrive at the quantity we desire. Thus, we have

$$\#\{\Lambda_2 \mid |\Lambda_1 \cap \Lambda_2| = i\} = \binom{3}{i} \sum_{j=0}^{3-i} (-1)^j \binom{i}{j} f(j-i).$$

Now, counting the number of pairs of 3-term arithmetic progressions $\{\Lambda_1, \Lambda_2\}$ that intersect 0, 1, 2, and 3 times, we have:

$$\begin{aligned} \#\{\{\Lambda_1, \Lambda_2\} \mid |\Lambda_1 \cap \Lambda_2| = 3\} &= \binom{n}{2} \\ \#\{\{\Lambda_1, \Lambda_2\} \mid |\Lambda_1 \cap \Lambda_2| = 2\} &= 6 \binom{n}{2} \\ \#\{\{\Lambda_1, \Lambda_2\} \mid |\Lambda_1 \cap \Lambda_2| = 1\} &= \frac{1}{2} \binom{n}{2} (9n - 39) \\ \#\{\{\Lambda_1, \Lambda_2\} \mid |\Lambda_1 \cap \Lambda_2| = 0\} &= \frac{1}{2} \binom{n}{2} (n^2 - 10n + 25). \end{aligned}$$

After using these to calculate the second moment of $A_{n,k}$ by plugging the resulting expression into Mathematica, we find

$$\sigma_{n,k}^2 = \frac{(k-2)(k-1)k(k^3 - 3k^2(n-1) - n(n^2 - 3n + 2) + k(3n^2 - 6n + 2))}{2(n-4)(n-3)(n-2)^2}.$$

Simplifying, we have $\sigma_{n,k}^2 = \frac{k^3(n-k)^3}{2n^4} + O(n)$. □

6.3 Nonexistence of LLT for A_n

The reason that we cannot use the general theorem in section 6 to prove that no LLT exists for A_n is that the distances between the conditioned distributions $A_{n,k}$ are the same order as their standard deviations $\sigma_{n,k}$, since $\mu_{n,k+1} - \mu_{n,k} = \Theta(\sigma_{n,k})$ for reasonably likely values of k (when $k = 1$, this breaks down but values of k this extreme are very unlikely).

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 $|\mathbb{P}(A_n = x) - \mathcal{N}_n(x)| = \Omega(\frac{1}{\sigma_n})$, 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 $x \geq \mu_{n,k}$ for all $k \leq j$ and $x \leq \mu_{n,k}$ for all $k > j$. Then we have

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

Now we must sufficiently bound the tails of $A_{n,k}$. We know that $\mu_n = \frac{1}{8}\binom{n}{2}$, $\sigma_n^2 = \Theta(n^3)$, $\mu_{n,k} = \frac{3}{n-2}\binom{k}{3}$, and $\sigma_{n,k}^2 = \frac{k^3(n-k)^3}{2n^4}$, but we do not know anything about the shape of the distribution of $A_{n,k}$.

Remark 6.4. We conjecture that $A_{n,k}$ follows a local limit theorem. 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 be the sum of independent random variables, since the constraint on size k makes the presence of each element of $\mathbb{Z}/n\mathbb{Z}$ dependent on the presence of every other element. If we could prove an LLT for $A_{n,k}$, that would allow us to conclude that no LLT exists for A_n , as we show below. It might be possible to prove that there is no LLT for A_n without proving the LLT for $A_{n,k}$ by a specialized argument, but using the following proposition would be the most satisfying way since an LLT for $A_{n,k}$ would give the most complete picture of the histogram of A_n .

Proposition 6.5. *Suppose $A_{n,k}$ follows an LLT for each k . Then there exists no LLT for A_n .*

Proof. We show that the height at the center of an $A_{n,k}$ distribution is greater than the Gaussian distribution at that point as n gets large. Let $k = \frac{n+1}{2}$ and $x = \mu_{n,k}$. Then since $A_{n,k}$ follows an LLT,

$$\begin{aligned} \mathbb{P}(A_n = x) &= \mathbb{P}(|S| = k) \cdot \mathbb{P}(A_{n,k} = x) = \frac{\binom{n}{k}}{2^n} \cdot \frac{1}{\sqrt{2\pi}\sigma_{n,k}} e^{-\left(\frac{x-\mu_{n,k}}{\sigma_{n,k}}\right)^2/2} \\ &= \left(\frac{1}{\sqrt{2\pi n}} e^{-\frac{(k-\frac{n}{2})^2}{2n}} \right) \cdot \left(\frac{1}{\sqrt{2\pi\sigma_{n,k}^2}} \right) \\ &\approx \frac{1}{2\pi} \sqrt{\frac{2n^4}{nk^3(n-k)^3}} \\ &= \frac{\sqrt{2}}{\pi} \left(\frac{n}{k(n-k)} \right)^{3/2} \\ &\approx \frac{8\sqrt{2}}{\pi} n^{-3/2}. \end{aligned}$$

Now we turn our attention to the height of the Gaussian distribution at $x = \mu_{n,k}$. Recall that $\sigma_n^2 \approx \frac{9}{4}n^3$ and let \mathcal{N}_n be the probability density function for the normal distribution with mean μ_n and variance σ_n^2 . Then we have

$$\mathcal{N}_n(x) = \frac{1}{\sqrt{2\pi}\sigma_n} e^{-\left(\frac{x-\mu_n}{\sigma_n}\right)^2/2} \leq \frac{1}{\sqrt{2\pi}\sigma_n} \approx \left(\frac{1}{3} \sqrt{\frac{2}{\pi}} \right) n^{-3/2}.$$

Therefore,

$$|\mathbb{P}(X_n = x) - \mathcal{N}_n(x)| \approx \left(\frac{8\sqrt{2}}{\pi} - \frac{1}{3} \sqrt{\frac{2}{\pi}} \right) n^{-3/2} = \Omega\left(\frac{1}{\sigma_n}\right),$$

so there exists no LLT for A_n . □

7 Acknowledgements

We would like to thank Felipe Hernandez for his close mentorship and for providing the topic of this project. We would also like to thank George Schaeffer for his helpful guidance and input on this project and for organizing SURIM, without which this project would not have been possible.

8 Appendix

8.1 3-APs in a subset of $\mathbb{Z}/n\mathbb{Z}$ and its complement

In this section, we prove a fun fact regarding 3-term APs. Let $n > 3$ be odd. Consider a subset $S \subseteq \mathbb{Z}/n\mathbb{Z}$. Define $A(S)$ to be the number of 3-term arithmetic progressions in S . That is,

$$A(S) = \frac{1}{2} |\{a, b \in \mathbb{Z}/n\mathbb{Z} \mid b \neq 0; a, a+b, a+2b \in S\}|.$$

Let $k = |S|$, and denote the complement of S in $\mathbb{Z}/n\mathbb{Z}$ as S^c , so $|S^c| = n - k$. We prove the following proposition.

Proposition 8.1. *The sum of the number of 3-APs in S and S^c is constant conditioning on n and k . That is, $A(S) + A(S^c)$ depends only on n and k .*

Note: this method of counting considers (x, y, z) and (z, y, x) to be the same arithmetic progression (due to the factor of $1/2$), and does not consider (x, x, x) to be an arithmetic progression. However, the result still holds under every permutation of these settings.

Proof. The total number of 3-term arithmetic progressions $A(\mathbb{Z}/n\mathbb{Z})$ is $n(n-1)$. These arithmetic progressions can be split into four groups based on the locations of their elements: all in S , all in S^c , exactly one in S , and exactly one in S^c .

There are $A(S)$ progressions all in S and $A(S^c)$ all in S^c . The number of progressions with exactly one element in S is $\frac{3}{2}k(k-1) - \frac{3}{2}A(S)$ since each pair of elements of S is contained in 3 total progressions, but we need to subtract the contribution of the progressions entirely contained in S , each of which contains 3 pairs. Similarly, the number of progressions with exactly one element in S^c is $\frac{3}{2}(n-k)(n-k-1) - \frac{3}{2}A(S^c)$. Adding everything together, we get

$$n(n-1) = A(S) + \left(\frac{3}{2}k(k-1) - \frac{3}{2}A(S) \right) + \left(\frac{3}{2}(n-k)(n-k-1) - \frac{3}{2}A(S^c) \right) + A(S^c).$$

Simplifying,

$$A(S) + A(S^c) = \frac{1}{4} (3k(k-1) + 3(n-k)(n-k-1) - n(n-1)).$$

Surprisingly, this sum depends only on n and k .

□

Intuition tells us that $A(S)$ is a measure of structure in S and the amount of structure in S is the same as the amount of structure in S^c . But this intuition must be wrong since this calculation implies that either the amount of structure in S^c is the inverse of the amount of structure in S , or $A(S)$ does not really measure structure.

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] Ross Berkowitz. A Quantitative Local Limit Theorem for Triangles in Random Graphs. *ArXiv e-prints*, March 2017.
- [3] Jason Fulman. Stein’s Method and Non-Reversible Markov Chains. *ArXiv e-prints*, December 1997.
- [4] Sourav Chatterjee. Lecture Notes in Stein’s method and applications. August 2007. Retrieved August 14, 2018 from <https://statweb.stanford.edu/~souravc/Lecture7.pdf>.