Asymptotics of a model problem from sieve theory

Zarathustra Brady

Sieve theory

- Take an interval A of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.

Sieve theory

- Take an interval A of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.
- Remove the multiples of some collection of primes $\mathcal{P}$ from this interval. Call the set that remains $\mathcal{S}(A, \mathcal{P})$.

Sieve theory

- Take an interval A of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.
- Remove the multiples of some collection of primes $\mathcal{P}$ from this interval. Call the set that remains $\mathcal{S}(A, \mathcal{P})$.
- For instance, if $\mathcal{P} = \{2, 3\}$, then $\mathcal{S}([5, 9], \{2, 3\}) = \{5, 7\}$.

Sieve theory

- Take an interval A of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.
- Remove the multiples of some collection of primes $\mathcal{P}$ from this interval. Call the set that remains $\mathcal{S}(A, \mathcal{P})$.
- For instance, if $\mathcal{P} = \{2, 3\}$, then $\mathcal{S}([5, 9], \{2, 3\}) = \{5, 7\}$.
- The big question:

  What can we say about $|\mathcal{S}(A, \mathcal{P})|$?

Sieve theory

- Take an interval A of consecutive whole numbers, such as $[5, 9] = \{5, 6, 7, 8, 9\}$.
- Remove the multiples of some collection of primes $\mathcal{P}$ from this interval. Call the set that remains $\mathcal{S}(A, \mathcal{P})$.
- For instance, if $\mathcal{P} = \{2, 3\}$, then $\mathcal{S}([5, 9], \{2, 3\}) = \{5, 7\}$.
- The big question:

$$\text{What can we say about } |\mathcal{S}(A, \mathcal{P})|?$$

- Pretend that we know $\mathcal{P}$, and that we know the length of A, but we don't know the endpoints of A.

Probabilistic version

- Suppose we pick a uniformly random number n from the interval A.

Probabilistic version

- Suppose we pick a uniformly random number n from the interval A.
- Although we don't know exactly what A is, we do know that

$$\frac{1}{2} - \frac{1}{|A|} \leq \mathbb{P}[2 \text{ divides } n] \leq \frac{1}{2} + \frac{1}{|A|}.$$

Probabilistic version

- Suppose we pick a uniformly random number n from the interval A.
- Although we don't know exactly what A is, we do know that

$$\frac{1}{2} - \frac{1}{|A|} \leq \mathbb{P}[2 \text{ divides } n] \leq \frac{1}{2} + \frac{1}{|A|}.$$

- We also know that

$$\frac{1}{3} - \frac{1}{|A|} \leq \mathbb{P}[3 \text{ divides } n] \leq \frac{1}{3} + \frac{1}{|A|},$$
$$\frac{1}{6} - \frac{1}{|A|} \leq \mathbb{P}[6 \text{ divides } n] \leq \frac{1}{6} + \frac{1}{|A|}.$$

Probabilistic version

- Suppose we pick a uniformly random number n from the interval A.
- Although we don't know exactly what A is, we do know that

$$\frac{1}{2} - \frac{1}{|A|} \leq \mathbb{P}[2 \text{ divides } n] \leq \frac{1}{2} + \frac{1}{|A|}.$$

- We also know that

$$\frac{1}{3} - \frac{1}{|A|} \leq \mathbb{P}[3 \text{ divides } n] \leq \frac{1}{3} + \frac{1}{|A|},$$
$$\frac{1}{6} - \frac{1}{|A|} \leq \mathbb{P}[6 \text{ divides } n] \leq \frac{1}{6} + \frac{1}{|A|}.$$

- So we can say that

$$\mathbb{P}[n \in \mathcal{S}(A, \{2,3\})] \geq 1 - \left(\frac{1}{2} + \frac{1}{|A|}\right) - \left(\frac{1}{3} + \frac{1}{|A|}\right) + \left(\frac{1}{6} - \frac{1}{|A|}\right).$$

The naïve approaches don't work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] \overset{?}{\approx} \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

The naïve approaches don't work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] \stackrel{?}{\approx} \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

- This guess is completely wrong!

The naïve approaches don't work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] \overset{?}{\approx} \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

- This guess is completely wrong!
- Take $A = [1, N]$ and take $\mathcal{P}_{\sqrt{N}}$ to be the set of primes below $\sqrt{N}$.

The naïve approaches don't work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] \overset{?}{\approx} \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

- This guess is completely wrong!
- Take $A = [1, N]$ and take $\mathcal{P}_{\sqrt{N}}$ to be the set of primes below $\sqrt{N}$.
- The guess above predicts that

$$\mathbb{P}[n \in \mathcal{S}([1, N], \mathcal{P}_{\sqrt{N}})] \overset{?}{\approx} \prod_{p < \sqrt{N}} \left(1 - \frac{1}{p}\right) \approx \frac{e^{-\gamma}}{\log(\sqrt{N})}.$$

The naïve approaches don't work

- If we ignore the $1/|A|$ error terms, we can use P.I.E. to predict

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] \overset{?}{\approx} \prod_{p \in \mathcal{P}} \left(1 - \frac{1}{p}\right).$$

- This guess is completely wrong!
- Take $A = [1, N]$ and take $\mathcal{P}_{\sqrt{N}}$ to be the set of primes below $\sqrt{N}$.
- The guess above predicts that

$$\mathbb{P}[n \in \mathcal{S}([1, N], \mathcal{P}_{\sqrt{N}})] \overset{?}{\approx} \prod_{p < \sqrt{N}} \left(1 - \frac{1}{p}\right) \approx \frac{e^{-\gamma}}{\log(\sqrt{N})}.$$

- But the true value is

$$\mathbb{P}[n \in \mathcal{S}([1, N], \mathcal{P}_{\sqrt{N}})] \approx \frac{1}{\log(N)}.$$

The naïve approaches don't work

- So we can't ignore the error terms.

The naïve approaches don't work

- So we can't ignore the error terms.
- Let's be really conservative this time, and try the union bound:

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] \geq 1 - \sum_{p \in \mathcal{P}} \left(\frac{1}{p} + \frac{1}{|A|} \right).$$

The naïve approaches don't work

- So we can't ignore the error terms.
- Let's be really conservative this time, and try the union bound:

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] \geq 1 - \sum_{p \in \mathcal{P}} \left(\frac{1}{p} + \frac{1}{|A|} \right).$$

- Now the error terms are under control, and at first this seems to be working well...

The naïve approaches don't work

- So we can't ignore the error terms.
- Let's be really conservative this time, and try the union bound:

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] \geq 1 - \sum_{p \in \mathcal{P}} \left(\frac{1}{p} + \frac{1}{|A|} \right).$$

- Now the error terms are under control, and at first this seems to be working well...
- The problem is that

$$\sum_{p \leq N} \frac{1}{p} \approx \log(\log(N))$$

diverges. This kills most simple variants of the above idea.

Bucketing approach

- Since $\sum_p \frac{1}{p}$ diverges, a good strategy is to put primes in *buckets*:
$$\mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \cdots \sqcup \mathcal{P}_k.$$

Bucketing approach

- Since $\sum_p \frac{1}{p}$ diverges, a good strategy is to put primes in *buckets*:
$$\mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \cdots \sqcup \mathcal{P}_k.$$
- We choose our buckets so that each sum
$$\sum_{p \in \mathcal{P}_i} \frac{1}{p}$$
is of size $\asymp 1$.

Bucketing approach

- Since $\sum_p \frac{1}{p}$ diverges, a good strategy is to put primes in *buckets*:
$$\mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \cdots \sqcup \mathcal{P}_k.$$
- We choose our buckets so that each sum
$$\sum_{p \in \mathcal{P}_i} \frac{1}{p}$$
is of size $\asymp 1$.
- This corresponds to taking buckets of the form
$$\mathcal{P}_i = \mathcal{P} \cap [|A|^{1/s}, |A|^{1/t}].$$

Bucketing approach

- Since $\sum_p \frac{1}{p}$ diverges, a good strategy is to put primes in *buckets*:

$$\mathcal{P} = \mathcal{P}_1 \sqcup \mathcal{P}_2 \sqcup \cdots \sqcup \mathcal{P}_k.$$

- We choose our buckets so that each sum

$$\sum_{p \in \mathcal{P}_i} \frac{1}{p}$$

is of size $\asymp 1$.

- This corresponds to taking buckets of the form

$$\mathcal{P}_i = \mathcal{P} \cap [|A|^{1/s}, |A|^{1/t}].$$

- Buckets corresponding to smaller primes $\to$ smaller error terms $\to$ naïve P.I.E. guess is a better approximation.

The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.

The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.
- The model problem asks: what if that was the only bucket?

The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.

- The model problem asks: what if that was the only bucket?
- Suppose we have

$$p \in \mathcal{P} \implies p \in [|A|^{1/(k+1)}, |A|^{1/k}].$$

The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.
- The model problem asks: what if that was the only bucket?
- Suppose we have

$$p \in \mathcal{P} \implies p \in [|A|^{1/(k+1)}, |A|^{1/k}].$$

- Then for any $p_1, ..., p_k \in \mathcal{P}$, we know that

$$\mathbb{P}[p_1 \cdots p_k \text{ divides } n] = \frac{1}{p_1 \cdots p_k} + O\Big(\frac{1}{|A|}\Big).$$

The model problem

- Most of the asymptotic error comes from the bucket containing the largest primes.
- The model problem asks: what if that was the only bucket?
- Suppose we have

$$p \in \mathcal{P} \implies p \in [|A|^{1/(k+1)}, |A|^{1/k}].$$

- Then for any $p_1, ..., p_k \in \mathcal{P}$, we know that

$$\mathbb{P}[p_1 \cdots p_k \text{ divides } n] = \frac{1}{p_1 \cdots p_k} + O\Big(\frac{1}{|A|}\Big).$$

- So the primes in $\mathcal{P}$ are uncorrelated when considered at most k at a time.

Simplification of model problem

- Since the primes all have roughly the same size, we treat them as interchangeable.

Simplification of model problem

- Since the primes all have roughly the same size, we treat them as interchangeable.
- Define a random variable X by

$$X = \#\{p \in \mathcal{P} \text{ such that } p \text{ divides } n\}.$$

Simplification of model problem

- Since the primes all have roughly the same size, we treat them as interchangeable.
- Define a random variable X by

$$X = \#\{p \in \mathcal{P} \text{ such that } p \text{ divides } n\}.$$

- The expected size of X is

$$\mathbb{E}[X] \approx \sum_{p \in \mathcal{P}} \frac{1}{p}.$$

Simplification of model problem

- Since the primes all have roughly the same size, we treat them as interchangeable.
- Define a random variable X by

$$X = \#\{p \in \mathcal{P} \text{ such that } p \text{ divides } n\}.$$

- The expected size of X is

$$\mathbb{E}[X] \approx \sum_{p \in \mathcal{P}} \frac{1}{p}.$$

- The second moment of X is given by

$$\mathbb{E}\left[\binom{X}{2}\right] \approx \sum_{p<q \in \mathcal{P}} \frac{1}{pq} \approx \frac{1}{2}\Big(\sum_{p \in \mathcal{P}} \frac{1}{p}\Big)^2.$$

A Poisson imitator appears

- For each $i \leq k$, we see that

$$\mathbb{E}\left[\binom{X}{i}\right] \approx \frac{\mathbb{E}[X]^i}{i!}.$$

A Poisson imitator appears

- For each $i \leq k$, we see that
$$\mathbb{E}\left[\binom{X}{i}\right] \approx \frac{\mathbb{E}[X]^i}{i!}.$$
- These are exactly the first k moments of a Poisson distribution!

A Poisson imitator appears

- For each $i \leq k$, we see that

$$\mathbb{E}\left[\binom{X}{i}\right] \approx \frac{\mathbb{E}[X]^i}{i!}.$$

- These are exactly the first k moments of a Poisson distribution!
- (We have no idea about the higher moments of X.)

A Poisson imitator appears

- For each $i \leq k$, we see that

$$\mathbb{E}\left[\binom{X}{i}\right] \approx \frac{\mathbb{E}[X]^i}{i!}.$$

- These are exactly the first k moments of a Poisson distribution!
- (We have no idea about the higher moments of X.)

- We want to estimate

$$\mathbb{P}[n \in \mathcal{S}(A, \mathcal{P})] = \mathbb{P}[X = 0].$$

Our problem

- Forget all the previous stuff.

Our problem

- Forget all the previous stuff.
- We have a random variable $X \in \mathbb{N}$, a Poisson parameter $\nu \in \mathbb{R}^+$, and $k \in \mathbb{N}$, s.t.

$$i \leq k \implies \mathbb{E}\left[\binom{X}{i}\right] = \frac{\nu^i}{i!}.$$

Our problem

- Forget all the previous stuff.
- We have a random variable $X \in \mathbb{N}$, a Poisson parameter $\nu \in \mathbb{R}^+$, and $k \in \mathbb{N}$, s.t.

$$i \leq k \implies \mathbb{E}\left[\binom{X}{i}\right] = \frac{\nu^i}{i!}.$$

- What are the best bounds we can put on

$$\mathbb{P}[X = 0]?$$

Our problem

- Forget all the previous stuff.
- We have a random variable $X \in \mathbb{N}$, a Poisson parameter $\nu \in \mathbb{R}^+$, and $k \in \mathbb{N}$, s.t.

$$i \leq k \implies \mathbb{E}\left[\binom{X}{i}\right] = \frac{\nu^i}{i!}.$$

- What are the best bounds we can put on

$$\mathbb{P}[X = 0]?$$

- For which ν, k can we prove that

$$\mathbb{P}[X = 0] > 0?$$

Markov's inequality

- How do we use the moment information?

Markov's inequality

- How do we use the moment information?
- Consider a polynomial $\theta(x)$ of degree k:

$$\theta(x) = \lambda_0 + \lambda_1 x + \lambda_2 \binom{x}{2} + \cdots + \lambda_k \binom{x}{k}.$$

Markov's inequality

- How do we use the moment information?
- Consider a polynomial $\theta(x)$ of degree k:

$$\theta(x) = \lambda_0 + \lambda_1 x + \lambda_2 \binom{x}{2} + \cdots + \lambda_k \binom{x}{k}.$$

- Our moment information tells us that

$$\mathbb{E}[\theta(X)] = \lambda_0 + \lambda_1 \nu + \lambda_2 \frac{\nu^2}{2} + \cdots + \lambda_k \frac{\nu^k}{k!}.$$

Markov's inequality

- How do we use the moment information?
- Consider a polynomial $\theta(x)$ of degree k:

$$\theta(x) = \lambda_0 + \lambda_1 x + \lambda_2 \binom{x}{2} + \cdots + \lambda_k \binom{x}{k}.$$

- Our moment information tells us that

$$\mathbb{E}[\theta(X)] = \lambda_0 + \lambda_1 \nu + \lambda_2 \frac{\nu^2}{2} + \cdots + \lambda_k \frac{\nu^k}{k!}.$$

- If $\theta(x) \leq 0$ for $x \in \{1, 2, ...\}$, we get

$$\mathbb{E}[\theta(X)] \leq \mathbb{P}[X = 0]\theta(0).$$

Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:

Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - θ has degree at most k,

Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - θ has degree at most k,
  - $\theta(0) = 1$,

Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - θ has degree at most k,
  - $\theta(0) = 1$,
  - for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,

Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - θ has degree at most k,
  - $\theta(0) = 1$,
  - for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,
- and to conclude that

$$\mathbb{P}[X = 0] \geq \mathbb{E}[\theta(X)] = e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!}.$$

Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - θ has degree at most k,
  - $\theta(0) = 1$,
  - for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,
- and to conclude that

$$\mathbb{P}[X = 0] \geq \mathbb{E}[\theta(X)] = e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!}.$$

- Are there any better ways to prove a lower bound on $\mathbb{P}[X = 0]$?

Convex optimization

- Our proof method is to write down a polynomial $\theta(x)$ such that:
  - θ has degree at most k,
  - $\theta(0) = 1$,
  - for all $x \in \mathbb{N}^+$, $\theta(x) \leq 0$,
- and to conclude that

$$\mathbb{P}[X = 0] \geq \mathbb{E}[\theta(X)] = e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!}.$$

- Are there any better ways to prove a lower bound on $\mathbb{P}[X = 0]$?
- A general duality result in convex optimization says that the **best lower bound** using this strategy is equal to the **least possible value** of $\mathbb{P}[X = 0]$.

Optimizing our choice of θ

- Selberg was able to compute the optimal choices of θ by hand for single digit values of the degree k.

Optimizing our choice of θ

- Selberg was able to compute the optimal choices of θ by hand for single digit values of the degree k.
- How?

Optimizing our choice of θ

- Selberg was able to compute the optimal choices of θ by hand for single digit values of the degree k.
- How?
- To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write θ in terms of its *roots*:

$$\theta(x) = \left(1 - \frac{x}{r_1}\right)\left(1 - \frac{x}{r_2}\right) \cdots \left(1 - \frac{x}{r_k}\right).$$

Optimizing our choice of θ

- Selberg was able to compute the optimal choices of θ by hand for single digit values of the degree k.
- How?
- To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write θ in terms of its *roots*:

$$\theta(x) = \left(1 - \frac{x}{r_1}\right)\left(1 - \frac{x}{r_2}\right)\cdots\left(1 - \frac{x}{r_k}\right).$$

- If there are any complex roots, replacing them with their real parts strictly improves our objective function.

Optimizing our choice of θ

- Selberg was able to compute the optimal choices of θ by hand for single digit values of the degree k.
- How?
- To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write θ in terms of its *roots*:

$$\theta(x) = \left(1 - \frac{x}{r_1}\right)\left(1 - \frac{x}{r_2}\right)\cdots\left(1 - \frac{x}{r_k}\right).$$

- If there are any complex roots, replacing them with their real parts strictly improves our objective function.
- Removing negative roots also strictly improves our objective function.

Optimizing our choice of θ

- Selberg was able to compute the optimal choices of θ by hand for single digit values of the degree k.
- How?
- To ensure that $\theta(x) \leq 0$ for $x \in \mathbb{N}^+$, we write θ in terms of its *roots*:

$$\theta(x) = \left(1 - \frac{x}{r_1}\right)\left(1 - \frac{x}{r_2}\right)\cdots\left(1 - \frac{x}{r_k}\right).$$

- If there are any complex roots, replacing them with their real parts strictly improves our objective function.
- Removing negative roots also strictly improves our objective function.
- Since coefficients of θ are linear in $1/r_i$, each r_i may be taken to be a whole number.

A simplex algorithm you can run by hand

- Our function θ can now be completely described by listing out its (integer) roots.

A simplex algorithm you can run by hand

- Our function θ can now be completely described by listing out its (integer) roots.
- Such a θ satisfies our requirements if:

A simplex algorithm you can run by hand

- Our function θ can now be completely described by listing out its (integer) roots.
- Such a θ satisfies our requirements if:
  - 1 is the least root of θ, and

A simplex algorithm you can run by hand

- Our function θ can now be completely described by listing out its (integer) roots.
- Such a θ satisfies our requirements if:
  - 1 is the least root of θ, and
  - the remaining roots of θ can be paired up so that each pair of roots are at most 1 apart.

A simplex algorithm you can run by hand

- Our function θ can now be completely described by listing out its (integer) roots.
- Such a θ satisfies our requirements if:
  - 1 is the least root of θ, and
  - the remaining roots of θ can be paired up so that each pair of roots are at most 1 apart.
- Our objective function is $e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!} = \sum_i \lambda_i \frac{\nu^i}{i!}$.

A simplex algorithm you can run by hand

- Our function θ can now be completely described by listing out its (integer) roots.
- Such a θ satisfies our requirements if:
  - 1 is the least root of θ, and
  - the remaining roots of θ can be paired up so that each pair of roots are at most 1 apart.
- Our objective function is $e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!} = \sum_i \lambda_i \frac{\nu^i}{i!}$.
- We can "pivot" our choice of θ by moving one of its roots, while keeping the other roots fixed.

A simplex algorithm you can run by hand

- Our function θ can now be completely described by listing out its (integer) roots.
- Such a θ satisfies our requirements if:
  - 1 is the least root of θ, and
  - the remaining roots of θ can be paired up so that each pair of roots are at most 1 apart.
- Our objective function is $e^{-\nu} \sum_n \theta(n) \frac{\nu^n}{n!} = \sum_i \lambda_i \frac{\nu^i}{i!}$.
- We can "pivot" our choice of θ by moving one of its roots, while keeping the other roots fixed.

- Proposition

  If no pivot increases the objective value, then θ is (globally) optimal.

...or by computer

k	critical ν_k	roots of the optimal θ
1	1	1
3	2	$1, \{3,4\}$ *or* $1, \{4,5\}$
5	3.11714	$1, \{3,4\}, \{7,8\}$
7	4.14377	$1, \{3,4\}, \{6,7\}, \{11,12\}$
9	5.23808	$1, \{3,4\}, \{6,7\}, \{10,11\}, \{14,15\}$
1001	≈ 503.37	$1, \{3,4\}, \{5,6\}, \{7,8\}, \ldots$
2001	≈ 1004	$1, \{3,4\}, \{5,6\}, \{7,8\}, \ldots$

...or by computer

k	critical ν_k	roots of the optimal θ
1	1	1
3	2	$1, \{3,4\}$ *or* $1, \{4,5\}$
5	3.11714	$1, \{3,4\}, \{7,8\}$
7	4.14377	$1, \{3,4\}, \{6,7\}, \{11,12\}$
9	5.23808	$1, \{3,4\}, \{6,7\}, \{10,11\}, \{14,15\}$
1001	≈ 503.37	$1, \{3,4\}, \{5,6\}, \{7,8\}, \ldots$
2001	≈ 1004	$1, \{3,4\}, \{5,6\}, \{7,8\}, \ldots$

- Selberg conjectured that $\nu_k \asymp \frac{k}{2}$ based on hand calculations.

...or by computer

k	critical ν_k	roots of the optimal θ
1	1	1
3	2	$1, \{3,4\}$ *or* $1, \{4,5\}$
5	3.11714	$1, \{3,4\}, \{7,8\}$
7	4.14377	$1, \{3,4\}, \{6,7\}, \{11,12\}$
9	5.23808	$1, \{3,4\}, \{6,7\}, \{10,11\}, \{14,15\}$
1001	≈ 503.37	$1, \{3,4\}, \{5,6\}, \{7,8\}, \ldots$
2001	≈ 1004	$1, \{3,4\}, \{5,6\}, \{7,8\}, \ldots$

- Selberg conjectured that $\nu_k \asymp \frac{k}{2}$ based on hand calculations.
- Selberg was able to prove that

$$\left\lfloor \frac{k+1}{2} \right\rfloor \leq \nu_k \leq k$$

for all k.

Selberg's lower bound

- Selberg has a famous construction of a "good enough" sieve which is easy to work with.

Selberg's lower bound

- Selberg has a famous construction of a "good enough" sieve which is easy to work with.
- In this context, we try θ of the form

$$\theta(x) = (1 - x) f(x)^2,$$

for an arbitrary polynomial $f(x)$ of degree $d = \frac{k-1}{2}$.

Selberg's lower bound

- Selberg has a famous construction of a "good enough" sieve which is easy to work with.
- In this context, we try θ of the form

$$\theta(x) = (1-x)f(x)^2,$$

  for an arbitrary polynomial $f(x)$ of degree $d = \frac{k-1}{2}$.
- The objective becomes a quadratic function of the coefficients of $f(x)$.

Selberg's lower bound

- Selberg has a famous construction of a "good enough" sieve which is easy to work with.
- In this context, we try θ of the form

$$\theta(x) = (1-x)f(x)^2,$$

  for an arbitrary polynomial $f(x)$ of degree $d = \frac{k-1}{2}$.
- The objective becomes a quadratic function of the coefficients of $f(x)$.
- By a miracle, we can optimize this quadratic form by hand!

Selberg's lower bound: the quadratic form

- Write out f in the binomial basis as

$$f(n) = \sum_{r \leq d} \ell_r \binom{n}{r}.$$

Selberg's lower bound: the quadratic form

- Write out f in the binomial basis as

$$f(n) = \sum_{r \le d} \ell_r \binom{n}{r}.$$

- We change coordinates to y_r given by

$$y_r = (-1)^r \sum_{i \ge 0} \ell_{r+i} \frac{\nu^i}{i!}.$$

Selberg's lower bound: the quadratic form

- Write out f in the binomial basis as

$$f(n) = \sum_{r \le d} \ell_r \binom{n}{r}.$$

- We change coordinates to y_r given by

$$y_r = (-1)^r \sum_{i \ge 0} \ell_{r+i} \frac{\nu^i}{i!}.$$

- Our objective function is

$$e^{-v} \sum_{n \ge 0} (1-n) f(n)^2 \frac{\nu^n}{n!} = \sum_r \frac{\nu^r}{r!} y_r^2 - \sum_r \frac{\nu^{r+1}}{r!} (y_r - y_{r+1})^2.$$

Selberg's lower bound: the quadratic form

- Write out f in the binomial basis as

$$f(n) = \sum_{r \le d} \ell_r \binom{n}{r}.$$

- We change coordinates to y_r given by

$$y_r = (-1)^r \sum_{i \ge 0} \ell_{r+i} \frac{\nu^i}{i!}.$$

- Our objective function is

$$e^{-v} \sum_{n \ge 0} (1-n) f(n)^2 \frac{\nu^n}{n!} = \sum_r \frac{\nu^r}{r!} y_r^2 - \sum_r \frac{\nu^{r+1}}{r!} (y_r - y_{r+1})^2.$$

- This becomes negative semidefinite when $\nu = d + 1 = \frac{k+1}{2}$.

Can we make a better sieve?

- I want to know how much we can improve Selberg's construction.

Can we make a better sieve?

- I want to know how much we can improve Selberg's construction.
- **Idea:** We know the optimal θ has the form

$$\theta(x) = (1 - x)f(x)f(x + 1)$$

for some polynomial f with integer roots.

Can we make a better sieve?

- I want to know how much we can improve Selberg's construction.
- **Idea:** We know the optimal θ has the form

$$\theta(x) = (1 - x)f(x)f(x + 1)$$

  for some polynomial f with integer roots.
- What if we drop the condition that f has integer roots?

Can we make a better sieve?

- I want to know how much we can improve Selberg's construction.
- **Idea:** We know the optimal θ has the form

$$\theta(x) = (1-x)f(x)f(x+1)$$

  for some polynomial f with integer roots.
- What if we drop the condition that f has integer roots?
- This will **over-estimate** the best possible lower bound on $\mathbb{P}[X=0]$.

A more difficult quadratic form

- We use the same change of variables y_r as in Selberg's construction.

A more difficult quadratic form

- We use the same change of variables y_r as in Selberg's construction.

- Our objective function is

$$e^{-v}\sum_{n\geq 0}(1-n)f(n)f(n+1)\frac{\nu^n}{n!} =$$

$$\sum_r \frac{\nu^r}{r!}y_r(y_r - y_{r+1}) - \sum_r \frac{\nu^{r+1}}{r!}(y_r - y_{r+1})(y_r - 2y_{r+1} + y_{r+2}).$$

A more difficult quadratic form

- We use the same change of variables y_r as in Selberg's construction.
- Our objective function is

$$e^{-v}\sum_{n\geq 0}(1-n)f(n)f(n+1)\frac{\nu^n}{n!} =$$

$$\sum_r \frac{\nu^r}{r!}y_r(y_r-y_{r+1})-\sum_r \frac{\nu^{r+1}}{r!}(y_r-y_{r+1})(y_r-2y_{r+1}+y_{r+2}).$$

- Selberg had to deal with a tridiagonal matrix, I have to deal with a pentadiagonal matrix!

Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for ν large.

Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for ν large.
- I computed the Cholesky decomposition for numerical examples to get a hint.

Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for ν large.
- I computed the Cholesky decomposition for numerical examples to get a hint.
- Eventually I found a (somewhat) clean proof that that it is negative semidefinite for $\nu \geq (\sqrt{d} + 1)^2$.

Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for ν large.
- I computed the Cholesky decomposition for numerical examples to get a hint.
- Eventually I found a (somewhat) clean proof that that it is negative semidefinite for $\nu \geq (\sqrt{d} + 1)^2$.

- Theorem
  For $k = 2d + 1$, we have $\nu_k \leq d + 2\sqrt{d} + 1$.

Everything somehow works out

- I want to prove that this horrible pentadiagonal symmetric matrix is negative semidefinite for ν large.
- I computed the Cholesky decomposition for numerical examples to get a hint.
- Eventually I found a (somewhat) clean proof that that it is negative semidefinite for $\nu \geq (\sqrt{d} + 1)^2$.

- Theorem

  For $k = 2d + 1$, we have $\nu_k \leq d + 2\sqrt{d} + 1$.

  - This result is not best-possible: numerical calculations indicate it can be improved to $\nu_k \leq d + \frac{\sqrt{d}}{2} + O(1)$.

Can we really get a square-root improvement?

- In our relaxed setting, it is possible to construct a polynomial $f(x)$ of degree d such that

$$\sum_{n\geq 0}(1-n)f(n)f(n+1)\frac{\nu^n}{n!} > 0$$

with $\nu \geq d + \Omega(\sqrt{d})$.

Can we really get a square-root improvement?

- In our relaxed setting, it is possible to construct a polynomial $f(x)$ of degree d such that

$$\sum_{n\geq 0}(1-n)f(n)f(n+1)\frac{\nu^n}{n!} > 0$$

with $\nu \geq d + \Omega(\sqrt{d})$.

- Does this mean that $\nu_{2d+1} \geq d + \Omega(\sqrt{d})$?

Can we really get a square-root improvement?

- In our relaxed setting, it is possible to construct a polynomial $f(x)$ of degree d such that

$$\sum_{n \geq 0} (1-n) f(n) f(n+1) \frac{\nu^n}{n!} > 0$$

  with $\nu \geq d + \Omega(\sqrt{d})$.

- Does this mean that $\nu_{2d+1} \geq d + \Omega(\sqrt{d})$?

- The first few roots of such an f (for $d \sim 500$) are

$$1, \{2.53, 3.53\}, \{5.19, 6.19\}, \{7.43, 8.43\}, \ldots$$

Can we really get a square-root improvement?

- In our relaxed setting, it is possible to construct a polynomial $f(x)$ of degree d such that

$$\sum_{n\geq 0}(1-n)f(n)f(n+1)\frac{\nu^n}{n!} > 0$$

with $\nu \geq d + \Omega(\sqrt{d})$.

- Does this mean that $\nu_{2d+1} \geq d + \Omega(\sqrt{d})$?

- The first few roots of such an f (for $d \sim 500$) are

$$1, \{2.53, 3.53\}, \{5.19, 6.19\}, \{7.43, 8.43\}, \ldots$$

- Most of the improvement can be traced back to allowing the second and third roots to be at 2.5 and 3.5.

How much of an improvement can we really get?

- I don't believe in a square-root improvement, but I want to show there is a real, definite improvement we can make.

How much of an improvement can we really get?

- I don't believe in a square-root improvement, but I want to show there is a real, definite improvement we can make.
- **Idea:** Take the roots from Selberg's construction, and round each multiplicity-two root up and down.

How much of an improvement can we really get?

- I don't believe in a square-root improvement, but I want to show there is a real, definite improvement we can make.
- **Idea:** Take the roots from Selberg's construction, and round each multiplicity-two root up and down.
- Numerically, this seems to give us a (small) improvement.

How much of an improvement can we really get?

- I don't believe in a square-root improvement, but I want to show there is a real, definite improvement we can make.
- **Idea:** Take the roots from Selberg's construction, and round each multiplicity-two root up and down.
- Numerically, this seems to give us a (small) improvement.
- **Problem:** we can't guarantee that doing this rounding won't make things *worse*.

How to make an improvement safely

- Recall our objective function (up to scale):

$$\sum_n \theta(n) \frac{\nu^n}{n!}.$$

How to make an improvement safely

- Recall our objective function (up to scale):

$$\sum_n \theta(n) \frac{\nu^n}{n!}.$$

- Every single summand, other than $\theta(0)$, is negative (or 0).

How to make an improvement safely

- Recall our objective function (up to scale):

$$\sum_n \theta(n) \frac{\nu^n}{n!}.$$

- Every single summand, other than $\theta(0)$, is negative (or 0).
- **Idea:** To guarantee that the objective increases, we try to *decrease* the absolute value $|\theta(n)|$ for all $n \in \mathbb{N}^+$.

Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1-x)\left(1-\frac{x}{r_1}\right)^2 \cdots \left(1-\frac{x}{r_d}\right)^2.$$

Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1-x)\left(1-\frac{x}{r_1}\right)^2 \cdots \left(1-\frac{x}{r_d}\right)^2.$$

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:

Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1-x)\left(1-\frac{x}{r_1}\right)^2 \cdots \left(1-\frac{x}{r_d}\right)^2.$$

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,

Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1-x)\left(1-\frac{x}{r_1}\right)^2 \cdots \left(1-\frac{x}{r_d}\right)^2.$$

- Replace each factor $(1 - x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,

Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1-x)\left(1-\frac{x}{r_1}\right)^2 \cdots \left(1-\frac{x}{r_d}\right)^2.$$

- Replace each factor $(1-x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq (1-x/r_i)^2$ for $x \in \mathbb{N}^+$, and

Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1-x)\left(1-\frac{x}{r_1}\right)^2 \cdots \left(1-\frac{x}{r_d}\right)^2.$$

- Replace each factor $(1-x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq (1-x/r_i)^2$ for $x \in \mathbb{N}^+$, and
  - at least one of $\lfloor r_i \rfloor$, $\lceil r_i \rceil$ is a root of $q_i(x)$.

Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1-x)\left(1-\frac{x}{r_1}\right)^2 \cdots \left(1-\frac{x}{r_d}\right)^2.$$

- Replace each factor $(1-x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq (1-x/r_i)^2$ for $x \in \mathbb{N}^+$, and
  - at least one of $\lfloor r_i \rfloor$, $\lceil r_i \rceil$ is a root of $q_i(x)$.

- This definitely doesn't hurt us. Does it help?

Safer rounding

- Write Selberg's $\theta(x)$ as a product:

$$\theta(x) = (1-x)\left(1-\frac{x}{r_1}\right)^2 \cdots \left(1-\frac{x}{r_d}\right)^2.$$

- Replace each factor $(1-x/r_i)^2$ by a quadratic $q_i(x)$ such that:
  - $q_i(0) = 1$,
  - $q_i(x) \geq 0$ for $x \in \mathbb{N}^+$,
  - $q_i(x) \leq (1-x/r_i)^2$ for $x \in \mathbb{N}^+$, and
  - at least one of $\lfloor r_i \rfloor$, $\lceil r_i \rceil$ is a root of $q_i(x)$.

- This definitely doesn't hurt us. Does it help?
- We can now guarantee that at least one of $\theta(\lfloor r_i \rfloor), \theta(\lceil r_i \rceil)$ has been replaced with 0!

An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

$$\sum_{r_i} \min\left(\left|\theta(\lfloor r_i \rfloor)\right| \frac{\nu^{\lfloor r_i \rfloor}}{\lfloor r_i \rfloor !},\ \left|\theta(\lceil r_i \rceil)\right| \frac{\nu^{\lceil r_i \rceil}}{\lceil r_i \rceil !} \right).$$

An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

$$\sum_{r_i} \min\left(\left|\theta(\lfloor r_i \rfloor)\right| \frac{\nu^{\lfloor r_i \rfloor}}{\lfloor r_i \rfloor!},\ \left|\theta(\lceil r_i \rceil)\right| \frac{\nu^{\lceil r_i \rceil}}{\lceil r_i \rceil!}\right).$$

- So now we need to understand two things:

An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

$$\sum_{r_i} \min\left(\left|\theta(\lfloor r_i \rfloor)\right| \frac{\nu^{\lfloor r_i \rfloor}}{\lfloor r_i \rfloor !},\ \left|\theta(\lceil r_i \rceil)\right| \frac{\nu^{\lceil r_i \rceil}}{\lceil r_i \rceil !}\right).$$

- So now we need to understand two things:
  - Where are the roots of Selberg's function θ?

An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

$$\sum_{r_i} \min\left(\left|\theta(\lfloor r_i \rfloor)\right| \frac{\nu^{\lfloor r_i \rfloor}}{\lfloor r_i \rfloor !},\ \left|\theta(\lceil r_i \rceil)\right| \frac{\nu^{\lceil r_i \rceil}}{\lceil r_i \rceil !} \right).$$

- So now we need to understand two things:
  - Where are the roots of Selberg's function θ?
  - How big is θ at the nearby integers?

An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

$$\sum_{r_i} \min\left(\left|\theta(\lfloor r_i \rfloor)\right| \frac{\nu^{\lfloor r_i \rfloor}}{\lfloor r_i \rfloor!},\ \left|\theta(\lceil r_i \rceil)\right| \frac{\nu^{\lceil r_i \rceil}}{\lceil r_i \rceil!} \right).$$

- So now we need to understand two things:
  - Where are the roots of Selberg's function θ?
  - How big is θ at the nearby integers?

- We have exact, combinatorial formulas for the coefficients of Selberg's function.

An understandable improvement

- If we perform the safer rounding, we guarantee improving our (rescaled) objective function by at least

$$\sum_{r_i} \min\left(\left|\theta(\lfloor r_i \rfloor)\right| \frac{\nu^{\lfloor r_i \rfloor}}{\lfloor r_i \rfloor!},\ \left|\theta(\lceil r_i \rceil)\right| \frac{\nu^{\lceil r_i \rceil}}{\lceil r_i \rceil!}\right).$$

- So now we need to understand two things:
  - Where are the roots of Selberg's function θ?
  - How big is θ at the nearby integers?

- We have exact, combinatorial formulas for the coefficients of Selberg's function.

- Slight wrinkle: Selberg's function is optimized for $\nu = d + 1$. So we modify it for larger ν, before rounding.

Explicit formula for Selberg's function

- Selberg's function is $\theta(x) = (1-x)f(x)^2$, where f is given by

$$f(n+2) = \frac{1}{(d+1)^{n+1}} \sum_i (-1)^i a(n,i) d^i .$$

Explicit formula for Selberg's function

- Selberg's function is $\theta(x) = (1-x)f(x)^2$, where f is given by

$$f(n+2) = \frac{1}{(d+1)^{n+1}} \sum_i (-1)^i a(n,i) d^i.$$

- Here $a(n,i)$ is the number of permutations of an n-set having exactly i cycles of size greater than 1.

Explicit formula for Selberg's function

- Selberg's function is $\theta(x) = (1-x)f(x)^2$, where f is given by

$$f(n+2) = \frac{1}{(d+1)^{n+1}} \sum_i (-1)^i a(n,i) d^i.$$

- Here $a(n,i)$ is the number of permutations of an n-set having exactly i cycles of size greater than 1.

- For $\nu > d+1$, we use the function f_ν given by

$$f_\nu(n+2) = \frac{1}{\nu^{n+1}} \sum_i (-1)^i a_q(n,i) d^i,$$

where $q = \nu - d$ and

$$a_q(n,i) = \sum_{\sigma \in S_n, i \text{ nontrivial cycles}} q^{\# \operatorname{Fix}(\sigma)}.$$

Let's at least understand $f(3)$ and $f(4)$

- To understand the contribution from rounding at the smallest root, we compute $f_\nu(3)$ and $f_\nu(4)$.

Let's at least understand $f(3)$ and $f(4)$

- To understand the contribution from rounding at the smallest root, we compute $f_\nu(3)$ and $f_\nu(4)$.
- We have

$$f_\nu(1+2) = \frac{1}{\nu^{1+1}}(a_q(1,0)d^0) = \frac{q}{\nu^2},$$

and

$$f_\nu(2+2) = \frac{1}{\nu^{2+1}}(a_q(2,0)d^0 - a_q(2,1)d^1) = -\frac{d-q^2}{\nu^3}.$$

Let's at least understand $f(3)$ and $f(4)$

- To understand the contribution from rounding at the smallest root, we compute $f_\nu(3)$ and $f_\nu(4)$.
- We have

$$f_\nu(1+2) = \frac{1}{\nu^{1+1}}(a_q(1,0)d^0) = \frac{q}{\nu^2},$$

and

$$f_\nu(2+2) = \frac{1}{\nu^{2+1}}(a_q(2,0)d^0 - a_q(2,1)d^1) = -\frac{d-q^2}{\nu^3}.$$

- These have opposite sign, so f_ν has a root between 3 and 4, and both $|f_\nu(3)|, |f_\nu(4)|$ are $\gg \frac{1}{d^2}$.

Let's at least understand $f(3)$ and $f(4)$

- To understand the contribution from rounding at the smallest root, we compute $f_\nu(3)$ and $f_\nu(4)$.
- We have

$$f_\nu(1+2) = \frac{1}{\nu^{1+1}}(a_q(1,0)d^0) = \frac{q}{\nu^2},$$

and

$$f_\nu(2+2) = \frac{1}{\nu^{2+1}}(a_q(2,0)d^0 - a_q(2,1)d^1) = -\frac{d-q^2}{\nu^3}.$$

- These have opposite sign, so f_ν has a root between 3 and 4, and both $|f_\nu(3)|, |f_\nu(4)|$ are $\gg \frac{1}{d^2}$.
- Most of the contribution to $f_\nu(n)$ comes from permutations which are almost entirely 2-cycles, so the result depends heavily on whether n is even or odd.

Saddle point method

- I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in i in order to get strong enough approximations...

Saddle point method

- I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in i in order to get strong enough approximations...
- My advisor (Sound) suggested a different approach.

Saddle point method

- I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in i in order to get strong enough approximations...

- My advisor (Sound) suggested a different approach.
- We can compute f_ν via a contour integral:

$$f_\nu(n+2) = \frac{n!}{2\pi i} \int_C e^{\nu z}(1-z)^d \frac{dz}{z^{n+1}}.$$

Saddle point method

- I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in i in order to get strong enough approximations...

- My advisor (Sound) suggested a different approach.
- We can compute f_ν via a contour integral:

$$f_\nu(n+2) = \frac{n!}{2\pi i} \int_C e^{\nu z}(1-z)^d \frac{dz}{z^{n+1}}.$$

- The integrand has saddle points at $z_0, \bar{z}_0$ solving the quadratic

$$\nu z_0^2 - (n+q)z_0 + n = 0.$$

Saddle point method

- I continued with the combinatorial analysis, eventually proving that $a_q(n, i)$ is log-concave in i in order to get strong enough approximations...

- My advisor (Sound) suggested a different approach.
- We can compute f_ν via a contour integral:

$$f_\nu(n+2) = \frac{n!}{2\pi i} \int_C e^{\nu z}(1-z)^d \frac{dz}{z^{n+1}}.$$

- The integrand has saddle points at $z_0, \bar{z}_0$ solving the quadratic

$$\nu z_0^2 - (n+q)z_0 + n = 0.$$

- Either way, we get a somewhat complicated sinusoidal expression for f_ν.

The dust settles

Theorem

If $k = 2d + 1$ *then*

$$\nu_k - d \geq (c + o(1))\sqrt[3]{d},$$

where $c \approx \frac{1}{12.14}$ *is the greatest positive solution of the inequality*

$$\int_0^\infty \frac{1}{x^{3/2}} \min\left(\sin^2\left((\tfrac{x}{3} + c)\sqrt{x}\right), \cos^2\left((\tfrac{x}{3} + c)\sqrt{x}\right)\right) dx \geq 2\pi c.$$

Thank you for your attention.