Universidad Nacional Autónoma de México

Facultad de Ciencias

El método probabilístico y el teorema de Hales-Jewett.

TESIS

que para obtener el grado académico de:

Actuario

presenta:

Fernando Campos García

Director de tesis: Dra. María Emilia Caballero Acosta

México, D.F. 25 de mayo del 2012.

1.-Datos del alumno:
Apellido Paterno: Campos
Apellido Materno: García
Nombre: Fernando
Tel: 56130252
Universidad Nacional Autónoma de México, Facultad de Ciencias
Carrera: Actuaría
de cuenta: 305123235
2.-Datos del tutor:
Grado: Dra.
Nombre: María Emilia
Apellido paterno: Caballero
Apellido materno: Acosta

3.- Datos del propietario 1:
Grado: Dr.
Nombre: Gerónimo Francisco
Apellido paterno: Uribe
Apellido materno: Bravo

4.- Datos del propietario 2:
Grado: M. en C.
Nombre: Guillermo
Apellido paterno: Garro
Apellido materno: Gómez

5.- Datos del suplente 1:
Grado: Dr.
Nombre: Ricardo
Apellido paterno: Gómez
Apellido materno: Aiza

6.- Datos del suplente 2:
Grado: Mat.
Nombre: Luis Alberto
Apellido paterno: Briseño
Apellido materno: Aguirre

7.-Datos de trabajo escrito:
Título: El método probabilístico y el teorema de Hales-Jewett.
de páginas : 52
Año: 2012

Agradecimientos

Primero que nada quisiera agradecer a mi familia; a mi padre, Angel, Marco y en especial a mi madre, sin ellos esta tesis asi como muchas otras cosas en mi vida no hubieran sido posibles.

En esta vida hay muchas cosas que agradecer y aunque hay muchas personas a las que debo decirle gracias, aquí pondré las más relevantes para mis estudios en los últimos 5 años. Pido una enorme disculpa a todos aquellos que no aparecen explícitamante esperando lleguen a identificarse con el ”gracias a todos”.

Dra. Maria Emilia Caballero, le agradezco por todo el apoyo a lo largo de los últimos años y por haber sido mi tutora, es sin duda la persona que más ha influido en mi formación universitaria, gracias por todo.

A mis sinodales. Gerónimo, gracias por tu enorme aportación a esta tesis en todo momento, sin duda alguna fuiste clave en su elaboración. Ricardo, un profesor que siempre me agradó, gracias por las correcciones. Memo, sinodal y amigo, gracias por todo. Luis, niñote, ocho años de conocernos, gracias por todo lo que me has enseñado en este tiempo. De igual manera quisiera agradecer a Toño y a Natalia por haber aceptado ser propuestos como sinodales.

En esta ruta de aprendizaje, quisiera agradecer a mis profesores, en especial a Carmen Arrillaga, Alfredo Cobián, Alfredo Amor, Ana Guzmán, Ricardo Gómez, Nelson Muriel, Mónica Clapp, Erick Mier, Carmen Hernández, Verónica de Jesús, Ruth Fuentes, María Emilia Caballero, Carlos Contreras y Mogens Bladt.

Gracias a mis ayudantes. Asímismo quisiera agradecer a Alfredo, Juan José y Lalo por permitirme ser su ayudante.

Gracias a:

Mis amigos olímplicos, en especial a Leo, Marco, Valente y Gogo, por todo lo que estudiamos, trabajamos, aprendimos y nos divertimos, vaya, por todo lo vivido. Mis amigos de toda la carrera, Lucy, Dení y Jorge, gracias por su apoyo y gracias por aguantarme, con ustedes fue muy amena la carrera. Mis compañeros y amigos, en especial a Pau y Pau, Jonathan, Oscar, Ale, Héctor, María, Diego, Lupita, Vivis, Mariana. Lucía, por ser fuerte de inspiración y un apoyo constante durante la escritura de este trabajo.

Finalmente quisiera dar un agradecimiento a las siguientes instituciones. Facultad de Ciencias, Instituto de Matemáticas, SEP, Fundación Telmex, Fundación Andrade y Fundación Hernández-Salicrup.

En resumen:

Gracias a todos.

Prefacio

Este trabajo que realizo para obtener mi título, ha sido todo un capítulo en el libro de mi vida. A continuación les presento un resumen de dicho capítulo.

Decidí que quería hacer mi tesis con la Dra. María Emilia Caballero, había sido mi profesora año y medio y me había gustado mucho trabajar con ella, se lo solité y María Emilia amablemente me aceptó como su tesista. La primer parte fue buscar un tema para la tesis, eso nos llevó un poco de tiempo, yo quería algo con probabilidad y teoría de números, en respuesta a esta solicitud mi asesora me entregó el artículo [5], yo me quedé muy sorprendido por que la doctora asumió que yo podría leer el artículo en fránces, después descubrí que leer matemáticas en fránces (y en general en cualquier idioma) es más sencillo de lo que imaginaba. Este artículo fue la motivación para abordar el método probabilístico en general y no sólo alguna aplicación a teoría de números.

El siguiente paso fue buscar más información acerca del tema, gran sorpresa, casi no había información, sin embargo encontré la biblia de este tema, [1], la cual me maravilló a tal grado que decidí dedicar gran parte de la tesis a exhibir estas joyas de las matemáticas abordadas con el método probabilístico.

Después aparece en la historia el Dr. Gerónimo Uribe. Él nos sugirió el artículo [6] pues ahí se aplicaba el método probabilístico y además se hacía de una manera novedosa, el artículo era vanguardista, de hecho en ese momento ni siquiera había sido aceptado para su publicación. Es hasta finales de abril del 2012 que dicho trabajo es publicado.

Como resultado se obtuvo una tesis que en su primer capítulo exhibe la belleza y poder del método y en el segundo aborda [6] tanto como pudimos. Esperamos que los capítulos sirvan como una introducción al método y al artículo y que al leer este trabajo disfruten tanto como yo disfruté al escribirlo.

Fernando Campos García.

Índice general

Capítulo 1

El Método Probabilístico

1.1. Descripción e historia

En inglés; *Probabilistic Method*, en francés; *Méthode Probabiliste*, desafortunadamente en español no he encontrado el método abordado en la tesis, sin embargo, al parecer clara la traducción, me referiré a él bajo el nombre de *Método Probabilístico.*

El método probabilístico es una herramienta para resolver problemas de matemáticas discretas, la primer demostración usando este método se remonta a 1943 y fue hecha por Szele, sin embargo es Paul Erdös el que lo populariza y realmente comprende el potencial del argumento a lo largo de la segunda mitad del siglo pasado, tal fue su aportación que algunos autores nombran al método como "Método de Erdös".

Antes de describir el método, les comentaré brevemente sobre este gran ícono de las matemáticas; Paul Erdös. Nacido en Hungría en 1913 y con un talento para las matemáticas demostrado desde edad precoz, Paul Erdös es argumentablemente el mejor matemático del siglo XX, con alrededor de 1500 publicaciones y 500 colaboradores (números que siguen creciendo apesar de su fallecimiento) sólo superado por Euler. Fue una persona que dedicó su vida a las matemáticas, no se casó, no tuvo hijos y casi no tenía posesiones materiales, el creía que las matemáticas era una actividad social y paso mucho tiempo de su vida alojado con sus colaboradores.

La biografía de Erdös, [2], es digna de ser un libreto de una película al estilo Hollywood por sus impresionantes situaciones que pueden llevar a la risa e incluso al llanto, sin embargo, sería una película con un derroche de conocimiento inmenso pues el *tío Paul*, como lo conocían sus allegados, publicó muchos artículos y de mucha calidad. Es tanto el legado de Erdös; sus nomenclaturas, sus frases, sus historias y por supuesto, su trabajo que hoy en día los matemáticos usan el número de Erdös para expresar su cercanía con él. Si publicaste con Erdös eres Erdös 1, si publicaste con alguien que publicó con Erdös, eres Erdös 2 y así sucesivamente. El tío Paul falleció al pie del cañón (la película debía tener un final a la altura), en un congreso en Varsovia en el año de 1996.

Regresando a la explicación. El método esencialmente se basa en dos observaciones: si hay un número finito de elementos en un conjunto y ninguno cumple la propiedad P, entonces la probabilidad de que un elemento elegido al azar del conjunto cumpla la propiedad P, es cero. Una definición habitual de la primer parte del método es la siguiente: para probar la existencia de una estructura combinatoria con cierta propiedad, se construye un espacio de probabilidad apropiado y se demuestra que un elemento escogido aleatoriamente en este espacio tiene probabilidad positiva de satisfacer la propiedad.

La segunda observación es que en un conjunto de números, hay números que son mayores o iguales al promedio y números que son menores o iguales al promedio. Formalmente podemos enunciar estas observaciones así:

- Sea A un conjunto finito, $a \in A$ y P una propiedad. Si $\mathbb{P}(a$ cumpla $P) > 0$ entonces existe $a^* \in A$ tal que cumple la propiedad P.
- Si X es una variable aleatoria no constante entonces $\mathbb{P}(X > \mathbb{E}(X)) > 0$ y $\mathbb{P}(X < \mathbb{E}(X)) > 0$.

Hay que notar es que el método probabilístico es determinista, esto en el sentido de que su conclusión lo es, es decir, la conclusión es "*existe* a^*" no "*con probabilidad cercana a* 1 (o 1) *existe* a^*" como pasa en los métodos probabilísticos de cálculo (e.g. Monte Carlo) los cuales son los que en español se mencionan habitualmente como métodos probabilísticos.

En un principio el método tiene su mayor aparición en gráficas, ahí, como en el resto de sus primeras aplicaciones, se puede ver que el argumento usado se puede traducir a un argumento de conteo (algunas veces de manera más sencilla que otras), sin embargo la ventaja de usar probabilidad es simplificar los argumentos, por ejemplo, cuando uno resuelve un problema de geometría, todo se puede traducir a los axiomas, sin embargo preferimos usar la teoría ya desarrollada para explicar una solución, algo similar pasa al usar el método.

En años más recientes, los alcances del método han crecido en varios sentidos, los temas abordados y las herramientas empleadas. Hoy en día hay aplicaciones del método cuya traducción a argumentos más simples es cada vez más difícil, esto debido a que se han incorporado cada vez más elementos de probabilidad que van desde esperanza, varianza, desigualdad de Chebyshev hasta cadenas de Markov, martingalas, leyes cero-uno, etc. De las aplicaciones del método en diversas áreas hablaremos más adelante.

1.2. División por tipo de problema

Las aplicaciones del método pueden ser divididas en tres grandes grupos. El primero se enfoca en el estudio de objetos aleatorios como las gráficas aleatorias o las matrices aleatorias, los resultados en estás áreas son escencialmente resultados en probabilidad pero motivados por problemas en combinatoria. El segundo consiste de aplicaciones para probar la existencia de estructuras combinatorias que satisfacen ciertas propiedades, de este es del más ejemplos se darán a lo largo de la tesis, en algunas ocasiones al probar la existencia de estas estructuras se pueden obtener construcciones de configuraciones extremas que dan solución a otras preguntas relacionadas a los problemas originales. El tercer grupo es quizá donde el método es más sorprendente ya que se compone de los problemas con enunciados deterministas cuya formulación no da el menor indicio de que la aleatorización pueda ayudar en su estudio.

1.2.1. Estructuras Aleatorias

Una gráfica aleatoria $G(n, p)$ es una gráfica sobre un conjunto fijo de n vértices etiquetados donde entre cada par de vértices hay una arista con probabilidad p y el evento de que hay una arista es independiente para cada par de vértices de la gráfica. Cada propiedad A es un evento en este espacio de probabilidad y es posible estudiar $\mathbb{P}(A)$, es decir, la probabilidad de que una gráfica aleatoria $G(n, p)$ satisfaga la propiedad A. Hay propiedades que cuando n es grande, la probabilidad de que $G(n, p)$ la satisfaga cambia rápidamente de casi 0 a casi 1 cuando p incrementa. Un ejemplo es el de la componente gigante: si $p < \frac{c}{n}$ y $c < 1$ entonces un elemento típico del espacio de probabilidad tiene todas sus componentes de tamaño a lo más logarítmico en n, sin embargo si $c > 1$, $G(n, p)$ es más propensa a tener una componente de tamaño lineal en n, conocida como componente gigante. Otros ejemplos son el *clique number* y el número cromático.

1.2.2. Construcciones probabilísticas

La primera demostración usando el método fue la del siguiente problema que plantearemos en una versión más coloquial. Se tienen n personas, en un torneo de todos contra todos cada pareja se enfrenta exactamente una vez y en cada juego hay un perdedor y un ganador, no hay empates. El teorema afirma que existe un torneo para el cual hay al menos $\frac{n!}{2^{n-1}}$ formas de acomodar a las n personas en una línea de tal manera que cada una le ganó al de su derecha. El resultado de la prueba no es constructivista, se prueba que dicho torneo existe pero no se da una construcción explícita de él, este fenónomeno se observa frecuentemente en las demostraciones hechas con este método.

Otro ejemplo sería el del torneo absurdo, ahora supongamos que tenemos un torneo como el descrito anteriormente y que al final del torneo se desea premiar a k personas (k fijo), imaginemos lo siguiente: que no importa a que k personas se premie, siempre existe una persona que les ganó a esas k, eso sería un absurdo puesto que uno premia a los *mejores* y si hay alguien que les gana a todos ellos, entonces *debería* ser premiado, resulta que si dejamos k fijo y hacemos n crecer, ¡siempre será posible encontrar un torneo absurdo! Esto es realmente sorprendente pues la estructura que tiene está gráfica es realmente compleja y parece poco plausible un torneo como este, sin embargo, el método probabilístico ayuda a probar su existencia. Problemas de este estilo y relacionados con la teoría de Ramsey fueron los que impulsaron el desarrollo del método probablístico. En la siguiente sección daremos la solución de este problema y algunos otros que hacen construcciones probabilísticas.

1.2.3. Enunciados deterministas

Consideremos el siguiente problema: para todo natural k existe un natural n tal que para todo conjunto A de n elementos hay una coloración de todos los enteros, cada uno de ellos pintado de alguno de los k colores, tal que cualquier traslación del conjunto A tiene enteros de todos los colores. El enunciado de entrada no da la impresión de tener que ver con probabilidad sin embargo su solución usa el lema local de Lovasz que afirma que para eventos çasi independientes"de probabilidad baja, hay probabilidad positiva de que ninguno ocurra.

Otro ejemplo es el teorema de Sperner del cual daremos una demostración en la siguiente sección, el teorema dice que si $\boldsymbol{F}$ es un conjunto de subconjuntos de $N = \{1, 2, \ldots, n\}$ tal que no existen $A, B \in \boldsymbol{F}$ con $A \subset B$, entonces $|\boldsymbol{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$. Este teorema es importante para esta tesis ya que representa el caso $k = 2$ del teorema Hales-Jewett el cual es de gran relevancia en este trabajo.

1.3. Aplicaciones del método en diversas áreas.

Como mencionamos anteriormente, el método ha aparecido cada vez en más áreas de las matemáticas, a continuación presentaremos algunos ejemplos que son representativos o por lo menos estéticos y hemos decidido presentarlos en este trabajo.

1.3.1. Gráficas

Realmente fue difícil la elección de los problemas para esta sección, como hemos comentado, es en teoría de gráficas donde ocurre el primer desarrollo del método y es grande el número de problemas que se han resuelto usándolo, sin embargo hemos decidido que este trabajo no podría estar completo sin el ejemplo por excelencia, una solución que apareció en un artículo de Erdös titulado *Some remarks on the theory of graphs* en el año de 1947 y da una cota inferior para los números diagonales de Ramsey, antes de enunciar el teorema definiremos estos números.

Definición 1.1. *Una gráfica completa K_n es un conjunto de n vértices con todas las aristan entre ellos.*

Definición 1.2. *El número de Ramsey $R(r,s)$ es el menor natural n tal que para cualquier coloración de las aristas de una gráfica completa K_n, cada una de ellas de rojo o azul, hay una gráfica completa K_r con todas sus aristas rojas o hay una gráfica completa K_s con todas sus aristas azules.*

Teorema 1.1. *Si $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ entonces $R(k,k) > n$, en particular $R(k,k) > \lfloor 2^{k/2} \rfloor$ para todo $k > 2$.*

Demostración. Primero demostraremos un lema que será útil en otras demostraciones.

Lema 1.1. ***Desigualdad de Boole.*** *Sean* $\{A_1, A_2, \ldots\}$ *una sucesión de eventos, entonces se cumple que:*

$$\mathbb{P}\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Demostración. Sea $\{B_1, B_2, \ldots\}$ una secuencia definida como

$$B_i = A_i - \bigcup_{k=1}^{i-1} A_k$$

entonces claramente los B_i son disjuntos y satisfacen que

$$\bigcup_{n=1}^{i} A_n = \bigcup_{n=1}^{i} B_n,\ \forall i \in \mathbb{N}$$

$$\mathbb{P}\left(\bigcup_{n=1}^{i} B_n\right) = \sum_{n=1}^{i} \mathbb{P}\left(B_n\right),\ \forall i \in \mathbb{N}.$$

Claramente $B_j \subseteq A_j$ entonces por monotonía se tiene que $P(B_j) \leq P(A_j)$ lo que implica que

$$\mathbb{P}\left(\bigcup_{n=1}^{i} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{i} B_n\right) \leq \sum_{n=1}^{i} \mathbb{P}\left(A_n\right),\ \forall i \in \mathbb{N}$$

haciendo $n \to \infty$ tenemos que por continuidad de la probabilidad

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) \leq \sum_{n=1}^{\infty} \mathbb{P}\left(A_n\right)$$

$\square$

Ahora continuamos con la demostración.

Consideremos una coloración aleatoria y equiprobable de K_n, con esto queremos decir que cada una de las aristas se va a pintar de azul o de rojo, que cada arista se pinta de manera independiente y que la probabilidad de que sea azul es la misma que sea roja, es decir, que ambas son iguales a $\frac{1}{2}$.

Sea B un subconjunto de k vertices de V, donde V es el conjunto de vértices de K_n, denotamos como A_B el evento de que la gráfica que consiste de los vértices en B sea monocromática, es decir, que todos sus vértices sean del mismo color. Notemos que $\mathbb{P}[A_B] = 2^{1-\binom{k}{2}}$ ya que después de pintar la primer arista, para que todas sean del mismo color deben ser iguales a la primera y eso pasa, para cada arista de las $\binom{k}{2} - 1$ restantes con probabilidad $\frac{1}{2}$, puesto que los eventos son independientes entonces la probabilidad de que esta gráfica inducida sea monocromática es el producto de las probabilidades que es efectivamente $2^{1-\binom{k}{2}}$. Ahora bien, la probabilidad de que alguno de estos eventos se cumpla es

$$\mathbb{P}\left(\bigcup_{B\subset V|B|=k} A_B\right) \leq \sum_{B\subset V|B|=k} \mathbb{P}(A_B) = \binom{n}{k} 2^{1-\binom{k}{2}} < 1,$$

donde hemos aplicado la desigualdad de Boole y la hipótesis del problema, por lo tanto la probabilidad de que al menos una de las subgráficas completas de tamaño k sea monocromática es menor a 1, por lo tanto existe una coloración para la cual no existe ninguna gráfica completa de k vértices que sea monocromática.

Para la segunda parte del resultado observemos que

$$\begin{aligned}\binom{n}{k}2^{1-\binom{k}{2}} &= \frac{n(n-1)\cdots(n-k+1)}{k!}2^{1+\frac{k}{2}-\frac{k^2}{2}}\\ &< \frac{n^k}{k!}2^{1+\frac{k}{2}-\frac{k^2}{2}}\\ &= \frac{2^{1+\frac{k}{2}}}{k!}\frac{n^k}{2^{\frac{k^2}{2}}},\end{aligned}$$

si $k \geq 3$ tenemos que $\frac{2^{1+\frac{k}{2}}}{k!} < 1$ y si además $n = \lfloor 2^{k/2} \rfloor$ entonces $n^k 2^{\frac{-k^2}{2}} \leq 1$ de donde $R(k,k) > \lfloor 2^{k/2} \rfloor$ para todo $k > 2$.

$\square$

Si el lector está interesado en más aplicaciones del método probabilístico a los números de Ramsey, les resultará muy interesanta la siguiente referencia: [7].

Retomemos el ejemplo del torneo absurdo comentado en 2,2 y ahora lo enunciaremos formalmente.

Definición 1.3. *Un torneo aleatorio es una gráfica en la que para todo par* $\{(i,j) : 1 \leq i < j \leq n\}$*, con probabilidad* $1/2$ *aparece la arista* (i,j) *y con probabilidad* $1/2$ *aparece la* (j,i) *de manera independiente.*

Nótese que hay $2^{\binom{n}{2}}$ torneos pues cada arista tiene dos opciones y hay $\binom{n}{2}$ aristas, asímismo los torneos son equiprobables.

Teorema 1.2. *Sea* k *un natural fijo entonces existe* n *un número natural para el cual hay un torneo de* n *personas en las que para cualquier subconjunto de* k *personas hay una que le ganó a todas ellas.*

Demostración. Si i y j son vértices de la gráfica decimos que i le gana a j si existe la arista (i,j) en otro caso j le gana a i, recordemos que no hay empates, primero demostraremos que si $\binom{n}{k}(1-2^{-k})^{n-k} < 1$ entonces hay un torneo con n vértices que cumple la condición. Sea V el conjunto de vértices y consideremos $\binom{n}{k}$ eventos, uno por cada subconjunto I de k vértices a los cuales denotamos A_I, dicho evento es que, para ese subconjunto no hay ningún otro jugador que les haya ganado a todos ellos.

Tenemos que $\mathbb{P}[A_I] = \mathbb{P}[$ Para todo vértice que no está en I, no le gana a los vértices en $I] = \prod_{v \notin I} \mathbb{P}[v$ no le gane a todos los vértices de $I] = \prod_{v \notin I}(1 - \mathbb{P}[v$ le gane a todos los vértices de $I]) = (1-2^{-k})^{n-k}$, ahora bien, la probabilidad de que ningún evento A_I suceda es lo mismo que para todo subconjunto de k elementos hay algún otro vértice que los vence a todos es decir

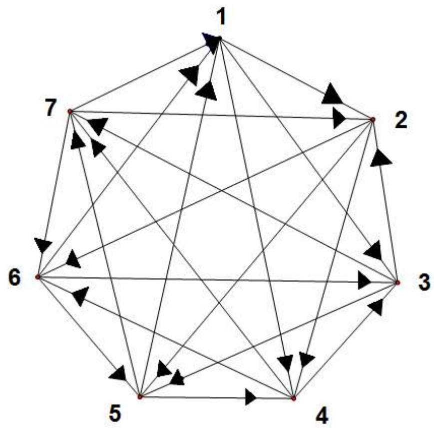
$$
\begin{aligned}
1 - \mathbb{P}[\bigcup_{I \subset V, |I|=k} A_I] &> 1 - \sum_{I \subset V, |I|=k} \mathbb{P}[A_I] \\
&= 1 - \binom{n}{k}(1-2^{-k})^{n-k} > 0,
\end{aligned}
$$

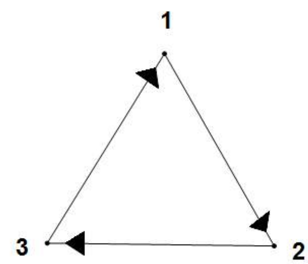
que es lo primero que queríamos demostrar.

Para finalizar la prueba falta ver que dichas n existen para lo cual basta notar que $\lim_{n\to\infty} \binom{n}{k}(1-2^{-k})^{n-k} = 0$.

$\square$

No es difícil probar que si $f(k)$ es el mínimo valor posible de n para el cual existe un torneo tal que cualquier subconjunto de k jugadores es vencido por alguno de los $n-k$ restantes entonces $f(1) = 3$ y $f(2) = 7$ como lo muestran los siguientes diagramas y el siguiente teorema.





Teorema 1.3. *Para todo k número natural se cumple que* $f(k) \geq 2^{k+1} - 1$.

Demostración. La demostración la haremos por inducción, para $k = 1$ es claro que 2 no es suficiente y ya dimos un acomodo con $f(k) = 3 \geq 2^{1+1}-1$ por lo tanto la base de inducción está hecha, ahora supongamos que $f(k) \geq 2^{k+1}-1$ para algún k y veamos que $f(k+1) \geq 2^{(k+1)+1} - 1$. Consideremos $f(k+1)$ personas numeradas a partir del 1 y digamos sin pérdida de generalidad que el jugador 1 es el que ganó más juegos. Para todo i definimos g_i y p_i como el número de juegos ganados y perdidos por el jugador i, respectivamente, entonces $g_1 - p_1 \geq 0$ pues de lo contrario al sumar sobre todos los i se obtendría que el número de juegos ganados es menor que el número de juegos perdidos lo cual es una contradicción ya que deben ser iguales, por lo tanto el que más gana, gana al menos tantos partidos como los que pierde. Nos fijamos en los que le ganan al jugador 1 y denotamos a ese conjunto por G y los que pierden contra él y lo denotamos por P, veamos que al menos $|G| > k$, de lo contrario al elegirlo a 1 junto con los elementos de G (y quizá algunos otros) se tendrían $k+1$ personas tales que nadie les gana a todos, pues ya nadie le gana al 1. Ahora tomamos cualquier subconjunto de k elementos de G y lo unimos con el 1 entonces tenemos $k+1$ jugadores y entonces alguien le gana a todos, pero ese alguien no puede estar en P pues nadie de ellos le gana a 1 entonces debe estar en G, por lo que tenemos que para todo conjunto de k jugadores de G hay un jugador en G que les gana a todos lo que implica que $|G| \geq f(k) \geq 2^{k+1} - 1$

por lo que $f(k+1) = 1 + |G| + |P| \geq 1 + 2^{k+1} - 1 + 2^{k+1} - 1 = 2^{(k+1)+1} - 1$ lo que concluye la inducción y la prueba del teorema.

$\square$

Este resultado me pareció muy bonito y no tan complicado técnicamente que decidí proponerlo para la Olimpiada Mexicana de Matemáticas como problema, sin embargo, contrario a lo que pensaba, el problema resultó ser conocido exactamente como se enuncia arriba y el autor del artículo donde aparece es nada más ni nada menos que Erdös. Esta es una anécdota más en mi vida acerca de la vida de Erdös.

Finalizaremos esta sección de gráficas con uno de los ejemplos más claros del poder del método, en 2,2 dimos la versión coloquial y aquí lo enunciaremos formalmemente, pero antes un par de definiciones.

Definición 1.4. *Un torneo es una gráfica completa donde cada arista se orienta en uno de los dos sentidos posibles.*

Definición 1.5. *Un camino hamiltoniano en una gráfica es aquél que pasa por todos los vértices de la gráfica exactamente una vez, empezando en un vértice y moviéndose por las aristas (respetando la orientación en caso de que la haya).*

Teorema 1.4. *Existe un torneo con n jugadores con al menos $\frac{n!}{2^{n-1}}$ caminos hamiltonianos.*

Demostración. Tomemos un torneo aleatorio T y una permutación al azar σ de los números del 1 al n (hay $n!$ permutaciones) entonces definimos H como el número de caminos hamiltonianos en T así como H_σ la variable aleatoria de que sigma genere un camino hamiltoniano, es decir, que $\forall i$, $1 \leq i < n$, $(\sigma(i), \sigma(i+1)) \in T$ de esta forma $H = \sum_\sigma H_\sigma$ por lo que

$$\mathbb{E}[H] = \mathbb{E}[\sum_\sigma H_\sigma] = \sum_\sigma \mathbb{E}[H_\sigma]$$

$$= \sum_{\sigma} \mathbb{P}[\sigma \text{ genera un camino hamiltoniano}] = \sum_{\sigma} (\frac{1}{2})^{n-1} = n!2^{-(n-1)},$$

donde hemos usado la definición de H, la linealidad de la esperanza, el hecho de que la esperanza de una indicadora es la probabilidad del evento que indica, que cada arista está con probabilidad $\frac{1}{2}$ y que hay $n!$ permutaciones.

Con esto concluimos la sección y que existe T con más de $n!2^{-(n-1)}$ caminos hamiltonianos. □

1.3.2. Teoría de números

En gran medida esta tesis fue inspirada por la teoría de los números, en la búsqueda de un tema que combinara probabilidad y teoría de números nos encontramos con aplicaciones del método probabilístico a esta rama en el artículo [5], en particular una argumentación heurística de porque el teorema de Green-Tao (2004) es cierto. El teorema escencialmente dice que para todo $k \in \mathbb{N}$ existen k primos en progresión aritmética, esto constituye en cierto sentido una generalización al teorema de Szemerédi el cual abordaremos en el próximo capítulo ya que el teorema de Hales-Jewett también es una generalización de esta joya de las matemáticas.

Empezaremos con un teorema de Erdös de 1965.

Definición 1.6. *Un subconjunto S de un grupo abeliano G es llamado libre de suma si $(S + S) \cap S = \emptyset$, es decir, no existen $s_1, s_2, s_3 \in S$ tales que $s_1 + s_2 = s_3$.*

Teorema 1.5. *Cualquier conjunto $A = \{a_1, \ldots, a_n\}$ de n enteros diferentes de cero contiene un subconjunto B libre de suma con $|B| > \frac{n}{3}$. Aquí G es los números enteros.*

Demostración. Por el teorema de Dirichlet que afirma que si a y b son enteros con $mcd(a, b) = 1$ entonces la sucesión $c_n = a + bn$ contiene una infinidad de primos, podemos asegurar que hay una infinidad de primos de la forma $3k+2$, sin embargo por completez daremos una demostración de este último hecho, similar a la de Euclides para demostrar que hay una infinidad de primos.

Lema 1.2. *Hay una infinidad de primos de la forma* $3k + 2$.

Demostración. Veamos que el conjunto de los primos de la forma $3k + 2$, P, es no vacío pues $2 \in P$. Supongamos que P es finito entonces sea $\pi = \prod_{p \in P} p$ y consideremos $\pi^2 + 1 \equiv 2 \pmod 3$ pues $3 \notin P$ y $1^2 \equiv 1$, $2^2 \equiv 1 \pmod 3$ por lo tanto $\pi^2 + 1$ debe tener algún factor primo de la forma $3k + 2$, de lo contrario, si todos sus factores primos fueran de la forma $3k + 1$ su producto sería de esta misma forma. Sea p^* un factor de la forma $3k + 2$ entonces si $p^* \in P$ como $p^* \left| \pi \right.$ se tiene que $p^* \left| \pi^2 \right.$ pero $p^* \left| \pi^2 + 1 \right.$ por lo que $p^* \left| 1 \right.$ que es una contradicción, por lo tanto $p^* \notin P$, esto contradice el supuesto de que en P estaban todos los primos de la forma $3k + 2$, por lo tanto este conjunto es infinito. $\square$

Por el lema podemos considerar un primo p de la forma $3k + 2$ primo relativo con todos los a_i, pues los a_i son finitos y los primos de la forma $3k+2$ son infinitos. Sea $C = \{k + 1, k + 2, \ldots 2k + 1\}$. Notemos que C es libre de suma en $\mathbb{Z}_p$ ya que $k+i+k+j \equiv 2k+i+j$ pero $2k+1 < 2k+i+j < 4k+3$ ya que $1 \leq i, j \leq k + 1$, además

$$\frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}.$$

Elijamos un entero x, $1 \leq x < p$ con distribución uniforme sobre las $p - 1$ posibilidades y definimos x_i como xa_i (mod$p$), $0 \leq x_i < p$. Veamos que cuando x varía sobre las $p-1$ posibilidades, x_i varía sobre el mismo conjunto. Supongamos que no, que existen $p > r > s > 0$ tales que $ra_i \equiv sa_i$ (mod

p) entonces como $mcd(a_i, p) = 1$ se tiene que $r \equiv s \pmod p$ lo cual es una contradicción, por lo tanto $P[x_i \in C] = |C|/(p-1) > 1/3$.
Sea X la variable aleatoria del número de subíndices i para los cuales $xa_i \in C$ y definimos para toda i la variable X_i como la indicadora de que $xa_i \in C$, es decir, $X_i = 1$ si $xa_i \in C$, en otro caso $X_i = 0$ entonces se tiene que

$$X = \sum_i X_i \Rightarrow \mathbb{E}[X] = \mathbb{E}[\sum_i X_i] = \sum_i \mathbb{E}[X_i] = \sum_i \mathbb{P}[X_i = 1] > \frac{n}{3}$$

por lo tanto existe una x y un subconjunto B de cardinalidad mayor a $\frac{n}{3}$ tal que $xb \pmod p \in C \ \forall b \in B$, este conjunto funciona pues si $b_1 + b_2 = b_3$ entonces $xb_1 + xb_2 = xb_3$ en particular $xb_1 + xb_2 \equiv xb_3 \pmod p$ lo cual es una contradicción con que C es libre de suma por lo tanto no existen dichos b_1, b_2 y b_3, lo que concluye la prueba. □

El ejercicio 3 del capítulo 2 de [1] plantea la siguiente generalización:

Teorema 1.6. *Cualquier conjunto $A = \{a_1, \ldots, a_n\}$ de n números reales diferentes de cero contiene un subconjunto B libre de suma con $|B| \geq \frac{n}{3}$.*

Demostración. Primero veamos que se puede suponer que los números en $A = \{a_1, a_2, \ldots a_n\}$ son todos irracionales. Consideremos $Q_i = \mathbb{Q}/a_i$ para $i = 1, 2, \ldots, n$ es decir Q_i es el conjunto de los números q_i tales que $q_i a_i$ es racional, tomemos $q \in \mathbb{R}$ tal que no pertenezca a la unión de los Q_i, esto es posible pues la unión de los Q_i es numerable ya que $a_i \neq 0$, entonces el conjunto $Aq = \{a_1 q, a_2 q, \ldots a_n q\}$ contiene únicamente números irracionales y si Bq es un conjunto libre de suma para Aq entonces al dividir entre q los elementos de Bq se obtiene un conjunto libre de suma para A, por lo tanto, sin pérdida de generalidad en A sólo hay números irracionales. Ahora probaremos un par de lemas:

Lema 1.3. *Teorema de Kronecker: Si v es irracional, el conjunto de puntos $\{nv\}$ con n en los naturales, es denso en el intervalo $(0, 1)$, donde $\{x\}$ denota la parte fraccionaria de x.*

Demostración. Supongamos que existen don puntos $P_1 < P_2$ en el intervalo $(0,1)$ que no contienen otro punto de la forma $\{nv\}$ y que están a distancia $\epsilon > 0$, como la sucesión $\{nv\}$ es infinita, existen $n_1 < n_2$ tales que $|\{n_1v\} - \{n_2v\}| < \epsilon/2$. Si $N = n_2 - n_1$ entonces la sucesión $n_1 + kNv$ con k en los naturales va dando *saltos* de a lo más $\epsilon/2$ en la misma dirección, por lo tanto en algún momento la sucesión caerá en el intervalo (P_1, P_2) lo cual es una contradicción, por lo tanto $\{nv\}$ es denso. $\square$

Definición 1.7. *Decimos que un conjunto de puntos P_n es uniformemente distribuido, si para todo intervalo I contenido en el $(0,1)$ se tiene que*

$$\lim_{n\to\infty} \frac{n_I}{n} = |I|\,,$$

donde n_I es la cantidad de puntos P_j con $1 \leq j \leq n$ que pertenecen al intervalo I.

Lema 1.4. *Teorema de distribución uniforme: Si v es irracional, el conjunto de puntos $\{nv\}$ con n en los naturales, es uniformemente distribuido.*

Demostración. Sea $\epsilon > 0$, por el lema 1 podemos elegir j tal que $0 < \{jv\} = \delta < \epsilon$. Sea $K = [1/\delta]$. Si $0 \leq h \leq K$, el intervalo I_h es el conjunto de puntos x tales que $\{hjv\} < x \leq \{(h+1)jv\}$. El intervalo I_K se extenderá más allá del 1 y consideraremos la representación circular (si se pasa del 1 continua desde el 0). Sea $f_h(n)$ el número de valores r con $r = 1, 2, \ldots n$ tales que $\{rv\} \in I_h$. Notemos que $\{tv\} \in I_0 \Leftrightarrow \{tv + hjv\} \in I_h$, entonces si $n > hj$ se tiene que

$$f_h(n) - f_h(hj) = f_0(n - hj).$$

Sin embargo, es claro que $f_h(hj) \leq hj$ y $f_0(n-hj)+hj \geq f_0(n)$, entonces

$$f_0(n) - hj \leq f_h(n) \leq f_0(n) + hj \Longrightarrow \lim_{n\to\infty} \frac{f_h(n)}{f_0(n)} = 1 \;\; (0 \leq h \leq K).$$

Notemos que

$$\sum_{h=0}^{K-1} f_h(n) \leq n \leq \sum_{h=0}^{K} f_h(n)$$

y del límite calculado deducimos que

$$\frac{1}{K+1} \leq \varliminf_{n\to\infty} \frac{f_0(n)}{n} \leq \varlimsup_{n\to\infty} \frac{f_0(n)}{n} \leq \frac{1}{K}.$$

Si I es el intervalo (a,b) con $b-a \geq \epsilon$ entonces existen enteros u y k tales que

$$0 \leq \{ujv\} \leq a \leq \{(u+1)jv\} \leq \{(u+k)jv\} \leq b \leq \{(u+k+1)jv\}$$

lo que implica que

$$\sum_{h=u+1}^{u+k-1} f_h(n) \leq n_I \leq \sum_{h=u}^{u+k} f_h(n).$$

Usando que $\lim_{n\to\infty} \frac{f_h(n)}{f_0(n)} = 1$, tenemos que

$$k-1 \leq \varliminf_{n\to\infty} \frac{n_I}{f_0(n)} \leq \varlimsup_{n\to\infty} \frac{n_I}{f_0(n)} \leq k+1$$

esto combinado con la cadena de desigualdades de arriba nos da

$$\frac{k-1}{K+1} \leq \varliminf_{n\to\infty} \frac{n_I}{n} \leq \varlimsup_{n\to\infty} \frac{n_I}{n} \leq \frac{k+1}{K}.$$

Pero

$$K\delta \leq 1 \leq (K+1)\delta, (k-1)\delta < I < (k+1)\delta.$$

Entonces

$$\frac{I-2\delta}{1+\delta} \leq \varliminf_{n\to\infty} \frac{n_I}{n} \leq \varlimsup_{n\to\infty} \frac{n_I}{n} \leq \frac{I+2\delta}{1-\delta}.$$

Como nosotros escogemos ϵ (y por lo tanto δ) tan pequeño como queramos entonces haciendo tender δ a 0 tenemos que $\lim_{n\to\infty} \frac{n_I}{n} = |I|$ como queríamos. $\square$

Para concluir la demostración del problema tomemos nuestro conjunto A de n números irracionales y nos fijamos en el intervalo $(1/3, 2/3)$, ahora consideremos la sucesión de conjuntos A, $2A\dots$, y sea $g(r)$ la cantidad de números x en rA tales que $1/3 < \{rx\} < 2/3$, por el lema que probamos, existe R tal que $g(R) \geq n/3$ pues si esta cantidad fuera menor siempre que $n/3$, siempre habría un número cuya parte fraccionaria de sus múltiplos aparece en menos de la tercera parte en el intervalo $(1/3, 2/3)$ lo cual no es posible. Entonces al considerar los números cuya parte fraccionaria aparece en el intervalo $(1/3, 2/3)$ en RA tenemos el subconjunto B buscado. $\square$

1.3.3. Geometría

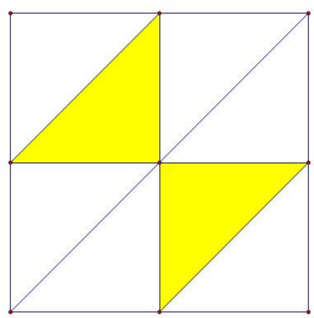
La relación de la probabilidad y la geometría es muy antigua, en un curso de probabilidad básico se estudia la probabilidad geométrica y se abordan problemas bastante asombrosos, nuestro primer ejemplo es un problema que se aborda cuando se aprende probabilidad geométrica y del cual presentaremos una generalización que, aunque no usa el método probabilístico, nos ejemplifica como una idea ingeniosa puede facilitarnos la demostración de un problema, cosa que hace el método probabilístico. El segundo ejemplo es una aplicación del método al problema de determinar el máximo de triángulos en un conjunto de n puntos en posición general tales que no contienen a ningún otro punto en su interior.

Problema 1.1. *Calcular la probabilidad de que al elegir tres puntos al azar sobre la circunferencia unitaria, el triángulo determinado tenga al origen en su interior.*

Demostración. Recordemos que un triángulo contiene a su circuncentro si y sólo si es acutángulo y que un triángulo sea acutángulo es equivalente a que los tres arcos determinados por los vértices sean menores que una semicircunferencia. Dada la simetría, se puede asumir sin pérdida de generalidad que un

punto es fijo, cortemos la circunferencia por dicho punto, entonces nuestro problema se traduce a encontrar la probabilidad de que al elejir aleatoriamente y con probabilidad uniforme dos puntos sobre un segmento dado, los tres segmentos así formados tengan longitud menor a la mitad del segmento original.
También sin pérdida de generalidad podemos suponer que el segmento es el $[0,1]$, Identifiquemos las coordenadas de los otros dos puntos como un punto aleatorio y con probabilidad uniforme (x,y) dentro del cuadrado $[0,1]\times[0,1]$. El área total es 1 y la región favorable es

$$\left\{(x,y)\,\middle|\,0<x<\frac{1}{2},\ \frac{1}{2}<y<x+\frac{1}{2}\right\}\cup\left\{(x,y)\,\middle|\,\frac{1}{2}<x<1,\ x-\frac{1}{2}<y<\frac{1}{2}\right\}$$



que tiene como se puede ver en la figura, área $1/4$, de donde la probabilidad buscada es $1/4$. □

Ahora supongamos que queremos resolver el problema para cuando se eligen 4 puntos al azar, o mejor cuando se eligen n puntos, entonces por un argumento similar al caso 3 podemos ver que es lo mismo que dividir el segmento unitario con $n-1$ puntos en n segmentos de tal manera que cada una de las longitudes sea menor a $1/2$ y entonces el problema apartir de aquí ya es técnico, *conocemos* la región de integración y conocemos la función de densidad conjunta por lo que en teoría podemos calcular la probabilidad, sin embargo, hay que ser cuidadosos con los límites de integración y como se van resolviendo esas integrales si se quiere llegar al resultado correcto. A continuación presentaremos una bella solución a este problema que evita hacer cálculos de ese estilo.

Problema 1.2. *Calcular la probabilidad de que al elegir* n *puntos* $P_1, P_2, \ldots, P_n$ *al azar sobre la circunferencia unitaria, el casco convexo tenga al origen en su interior.*

Demostración. Elijamos n puntos de manera aleatoria de la siguiente forma. Escojamos n líneas que pasen por el centro de manera uniforme y conideremos los extremos que denotamos por $Q_1, Q_{n+1},\ Q_2, Q_{n+2},\ \ldots\ Q_n, Q_{2n}$, después escogemos P_i aleatoriamente y equiprobablemente de entre Q_i y Q_{n+i}. La probabilidad de que el origen no este dentro del casco convexo de los puntos P_i dados los puntos distintos Q_j (esto pasa con probabilidad 1) es precisamente $x/2^n$ donde x es el número de subconjuntos de los puntos Q_j contenidos en un semiplano determinado por una línea que pasa por el origen y que no pasa por ningún punto de los Q_j. Veamos que $x = 2n$, esto sucede ya que si volvemos a numerar los puntos Q_j en el sentido de las manecillas del reloj tenemos que $Q_{n+i} = -Q_i$ entonces los subconjuntos contenidos en tales semiplanos son precisamente los de la forma $\{Q_r, Q_{r+1}, \ldots, Q_{r+n-1}\}$ donde los subíndices están pensados módulo $2n$, de donde la probabilidad buscada es $1 - 2n/2^n$. □

Ahora abordemos otro interesante problema. Dado un conjunto X de n puntos en el plano, no tres de ellos en la misma recta, sea $f(X)$ el número de triángulos determinados por esos puntos tales que en su interior no contienen puntos de X. Se define $f(n)$ como el mínimo valor de $f(X)$ cuando X varía sobre todas las posiciones posibles. En 1988, Katchalski y Meir probaron que

$$\binom{n-1}{2} \leq f(n) < 200n^2,$$

ahora presentaremos un teorema que da una mejor cota superior.

Teorema 1.7. *Sea* $I_i = \{(x, y) : x = i, y \in [0, 1]\}$ *para* $i = 1, 2, \ldots, n$. *Para cada* i *se elige un punto* P_i *en el intervalo* I_i *de manera independiente a los demás intervalos y con distribución uniforme. Sea* X *el conjunto de los* n *puntos entonces el número esperado de triángulos sin puntos de* X *en su interior es a lo más* $2n^2 + O(nlogn)$.

Demostración. Primero daremos una cota para la probabilidad de que el triángulo determinado por los puntos P_i, P_{i+a}, P_{i+k} para una i fija y con $k = a + b \geq 3$ (no nos interesa el caso $a = b = 1$ pues los triángulos $P_iP_{i+1}P_{i+2}$ son vacíos). Sea m la distancia entre P_{i+a} y $\overline{P_iP_{i+k}} \cap I_{i+a}$. Como cada uno de los puntos P_j para $i < j < i + k$ son escogidos aleatoriamente con una distribución uniforme en I_j se tiene que

$$\begin{aligned}
&\mathbb{P}(\triangle P_iP_{i+1}P_{i+2} \text{ es vacío} \,|m) \\
&= \prod_{j=i+1}^{i+a-1} \mathbb{P}(P_j \notin \triangle P_iP_{i+a}P_{i+k} \,|m) \prod_{j=i+a+1}^{i+k-1} \mathbb{P}(P_j \notin \triangle P_iP_{i+a}P_{i+k} \,|m) \\
&= \left(1 - \frac{m}{a}\right)\left(1 - 2\frac{m}{a}\right) \cdots \left(1 - (a-1)\frac{m}{a}\right)\left(1 - (b-1)\frac{m}{b}\right)\left(1 - \frac{m}{b}\right) \\
&\leq \exp\left(-\frac{m}{a} - 2\frac{m}{a} - \cdots - (a-1)\frac{m}{a} - (b-1)\frac{m}{b} - \cdots - \frac{m}{b}\right) \\
&= \exp\left(-\frac{a(a-1)}{2}\frac{m}{a} - \frac{b(b-1)}{2}\frac{m}{b}\right) = \exp\left(-(k-2)\frac{m}{2}\right),
\end{aligned}$$

donde se ha usado el teorema de Thales, la desigualdad $1 - \frac{1}{x} \leq e^{-\frac{1}{x}}$ y suma de Gauss.

Ahora calculemos la probabilidad no condicional, para esto veamos que la probabilidad de que la distancia entre P_{i+a} y $\overline{P_iP_{i+k}} \cap I_{i+a}$ sea a lo más d es menor o igual que $2d$ (puede estar de cualquier lado) de donde

$$\mathbb{P}(\triangle P_iP_{i+1}P_{i+2} \text{ es vacío}) \leq 2\int_0^\infty \mathbb{P}(\triangle P_iP_{i+1}P_{i+2} \text{ es vacío} \,|m) f(m)dm$$

$$\leq 2\int_0^\infty \exp\left(-(k-2)\frac{m}{2}\right)dm = \frac{4}{k-2}.$$

Sea X la variable aleatoria que cuenta el número de triángulos vacios entonces X es suma de $\binom{n}{3}$ indicadoras, utilizando la linealidad de la esperanza, el hecho de que los triángulos $P_iP_{i+1}P_{i+2}$ son vacíos con $i = 1, 2, \ldots, n-2$ y la probabilidad antes calculada se tiene que

$$\begin{aligned}
\mathbb{E}[X] &\leq n-2+\sum_{i=1}^{n-3}\sum_{k=3}^{n-i}\sum_{a=1}^{k-1}\frac{4}{k-2}\\
&= n-2+\sum_{i=1}^{n-3}\sum_{k=3}^{n-i}(k-1)\frac{4}{k-2}\\
&= n-2+\sum_{k=3}^{n-1}\sum_{n=1}^{n-k}(k-1)\frac{4}{k-2}\\
&= n-2+\sum_{k=3}^{n-1}(n-k)(k-1)\frac{4}{k-2}\\
&= n-2+\sum_{k=3}^{n-1}(n-k)\frac{4}{k-2}+4\sum_{k=3}^{n-1}(n-k)\\
&= n-2+\sum_{k=3}^{n-1}(n-k)\frac{4}{k-2}+4\sum_{k=3}^{n-1}n-4\sum_{k=3}^{n-1}k\\
&= n-2+\sum_{k=3}^{n-1}(n-k)\frac{4}{k-2}+4n^2-12n-4\left(\frac{(n-1)(n-2)}{2}-3\right)\\
&\leq 4n\sum_{k=3}^{n-1}\frac{1}{k-2}+2n^2-5n+6\\
&\leq 4n\ln n+2n^2 = 2n^2+O(n\log n).
\end{aligned}$$

$\square$

1.3.4. Computación

La probabilidad desempeña un papel cada vez más importante en la implementación de algoritmos, en algunos casos se cuentan con demostraciones de la existencia de una estructura con cierta propiedad (demostraciones con el método probabilístico) sin embargo no hay formas prácticas de construir de construir dicha estructura entonces si es *fácil* verificar si se cumple la propiedad o no resulta en ocasiones más conveniente elegir un espacio de probabilidad en donde al escoger un elemento al azar de él, la probabilidad de que cumpla dicha propiedad sea grande de está manera en vez de hacer un complejo algoritmo cnstructivista tendremos que hacer un algoritmo de revisión y usarlo unas cuantas veces. En otros casos, los algoritmos probabilísticos que provee una prueba de este estilo pueden ser *desaleatorizados* para crear un algoritmo determinista que en *poco tiempo* construya la estructura deseada.

El problema que presentamos en esta sección es referente a la teoría de códigos, y es conocido como la desigualdad de Kraft ($L = 2$), por ser fácilmente generalizable daremos la prueba del caso general.

Teorema 1.8. *Consideremos un abecedario con L letras y sea F una colección finita de palabras (concatenación de letras del abecedario) de longitud finita tal que ninguna palabra es prefijo de alguna otra en F. Sea N_i el número de palabras de longitud i en F, entonces*

$$\sum_i \frac{N_i}{L^i} \leq 1$$

Demostración. Supongamos que $\sum_i \frac{N_i}{L^i} > 1$, elijamos una palabra W de M letras donde M es el máximo de letras de una palabra en F y calculemos la probabilidad de que W empiece con una palabra de F. Definimos $A_i =$

$\{la\ palabra\ W\ tiene\ \ prefijo\ en\ F\ de\ i\ letras\}$ entonces la probabilidad de que W empiece con alguna de las palabras de F es igual a

$$\mathbf{P}\left(\bigcup_i A_i\right) = \sum_i \mathbf{P}(A_i)$$

esto debido a que los eventos A_i son disjuntos ya que si W empieza con una palabra de r letras y con una de s con $r < s$ entonces hay en F una palabra que tiene como prefijo una palabra que también está en F, lo cual es imposible. Ahora bien $\mathbf{P}(A_i) = \frac{N_i}{L^i} \Rightarrow \sum_i \mathbf{P}(A_i) > 1$ que es imposible ya que es una probabilidad y las probabilidades son menores que 1, por lo tanto nuestro supuesto es falso y se debe tener que $\sum_i \frac{N_i}{L^i} \leq 1$ como queríamos. $\square$

1.3.5. Olimpiada

Un trabajo tan importante para mí no podría estar completo sin reflejar una parte tan importante de mi vida. Las competencias de matemáticas se han realizado desde hace muchos años, por ejemplo desde 1984 se realiza en Hungría la *competencia Eötvös* para estudiantes que aún no hayan entrado a la universidad, sin embargo no es hasta 1959 que aparece la *Olimpiada Internacional de Matemáticas*, sin duda alguna el concurso más importante de matemáticas en el mundo por el que han pasado grandes matemáticos como Terence Tao, Timothy Gowers, Grigory Margulis y Po-Shen Loh, todos ellos relevantes en esta tesis. Los primeros dos son pilares del polymath base del capítulo 2, el tercero fue el primero en construir las *expander graphs* y el último la motivación de esta sección.

Po-Shen Loh participó por Estados Unidos en la IMO de 1999, durante su Ph.D. en Princeton se familiarizó con el método probabilístico el cual enseña a los seleccionados de ese país para la IMO desde 2008, año en el que se reincorporó como entrenador del equipo estadounidense. Él ha observado que hay varios problemas de olímpiada que es posible resolver usando este método, sobre todo explotando la linealidad de la esperanza. Para mayor información de problemas tipo olimpiada que pueden ser abordados con esta herramienta se recomienda al lector consultar [8].

Los dos ejemplos que presentaremos en esta sección son problemas que aparecieron en listas cortas de la IMO, la lista corta es una lista reducida de los problemas que proponen todos los países para conformar el examen de la IMO y consiste de entre 25 y 30 problemas divididos por áreas (álgebra, combinatoria, geometría y teoría de números), de esta lista, los líderes de cada país deciden cuáles serán los 6 problemas que conformarán el examen de ese año.

Problema 1.3. *Lista corta de la IMO de 1999. Sea A un conjunto de N residuos (mod N^2). Probar que existe un conjunto B de N residuos (mod N^2) tal que el conjunto $A+B=\{a+b\,|a\in A, b\in B\}$ contiene al menos la mitad de todos los residuos (mod N^2).*

Demostración. Hagamos N elecciones aleatorias con probabilidad uniforme sobre los N^2 residuos y tomemos ese conjunto como B, dado que las elecciones fueron independientes puede que en B haya menos de N elementos pero no importa, si dicho B cumple, al agregarle elementos para que tenga cardinalidad N seguirá cumpliendo. Sea S el número de valores que se pueden expresar como $a+b$ con $a\in A, b\in B$. Notemos que como A tiene N elementos, entonces para cada residuo i hay exactamente N residuos b tales que $i\in A+b$ entonces

$$\mathbb{P}[i\in A+B]=1-\mathbb{P}[i\notin A+B]=1-\left(\frac{N^2-N}{N^2}\right)^N$$

y por linealidad de la esperanza, usando que X no es más que la suma de N^2 indicadoras (las indicadoras de que $i\in A+B$, $i=1,2,\ldots N^2-1$), se tiene que

$$\mathbb{E}[X] = N^2\left(1-\left(\frac{N^2-N}{N^2}\right)^N\right) \geq N^2/2 \Leftrightarrow 1-\left(\frac{N^2-N}{N^2}\right)^N \geq 1/2$$

$$\Leftrightarrow \left(\frac{1}{2}\right)^{1/N} \geq 1-\frac{1}{N}$$

que es cierta ya que $e > 2$ y $e^{-1/N} \geq 1 - \frac{1}{N}$ pues $e^{-1} = \text{lím}_{N\to\infty}\left(1-\frac{1}{N}\right)^N$ y $\left(1-\frac{1}{x}\right)^x$ es una función creciente para $x \geq 1$. $\square$

Problema 1.4. *Lista corta de la IMO de 2006. Sea S un conjunto finito de puntos en el plano tales que no hay tres de ellos colineales. Para cada polígono convexo P cuyos vértices están en S, sea $a(P)$ el número de vértices de P y sea $b(P)$ el número de puntos de S que están afuera de P. Probar que*

$$\sum_P x^{a(P)}(1-x)^{b(P)} = 1$$

donde la suma es tomada sobre todos los polígonos convexos con vértices en S.
Nota: *Un segmento, un punto y el conjunto vacío son considerados polígonos convexos de 2, 1 y 0 vértices, respectivamente.*

Demostración. Consideremos primero el caso en que $0 \leq x \leq 1$ y pintemos aleatoriamente con probabilidad x los puntos de rojo, por lo tanto con probabilidad $1-x$ no están pintados, digamos que son negros si no están pintados. Para cada polígono convexo P sea I_P el evento de que todos los puntos en el perímetro de P son rojos y todos los puntos afuera de P son negros. Estos eventos son ajenos pues si P y P^* son dos polígonos para los cuales sucede I_P e I_{P^*} entonces eso significa que fuera de P no hay puntos rojos de donde $P \supseteq P^*$, también se tiene que fuera de P^* no hay puntos rojos de donde $P^* \supseteq P$, que implica que el casco convexo de P y de P^* coinciden pero al no haber tres puntos colineales en S lo anterior implica que $P = P^*$, esto nos dice que el lado izquierdo de la igualdad corresponde a la probabilidad de que algún I_P suceda que es igual a la esperanza del número de polígonos P que cumplen que sus vértices son rojos y fuera de P los vértices son negros, pero sabemos que esto pasa exactamente una vez, cuando P es la envolvente convexa de los puntos rojos, por lo tanto, dicha variable aleatoria resulta ser la constante 1 que tiene esperanza 1 y de ahí el resultado en este caso.

Para finalizar notemos que el polinomio $\sum_P x^{a(P)} (1-x)^{b(P)} - 1$ es de grado finito y se anula en una infinidad de valores (el intervalo $(0,1)$) por lo tanto debe ser el polinomio constante 0 lo que nos dice que $\sum_P x^{a(P)} (1-x)^{b(P)} = 1$ es cierto para todo real x. $\square$

1.3.6. Conjuntos

Hay en la literatura muchos problemas de conjuntos que pueden ser abordados usando el método probabilístico, Bollobás y Lovász tienen teoremas *parecidos*, está la desigualdad de Tuza y un poco más famoso, el teorema de Erdös-Ko-Rado, todos ellos perfectos ejemplos de la aplicación y belleza del método, sin embargo, a manera de introducción al siguiente capítulo, finalizaremos este con el lema de Sperner, el cual representa como ya hemos mencionado el caso trivial del teorema Hales-Jewett pero no por esto es un problema fácil, obtendremos la mayor información posible sobre este teorema, cuando se alcanza la igualdad y un corolario.

Teorema 1.9. *Sea $\boldsymbol{F}$ un conjunto de subconjuntos de $N = \{1, 2, \ldots, n\}$ tal que no existen $A, B \in \boldsymbol{F}$ tales que $A \subset B$, entonces $|\boldsymbol{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.*

Demostración. Elijamos un subconjunto aleatorio de N de la siguiente forma: primero elegimos una permutación σ de N y después elegimos un entero $m \in N \cup \{0\}$, ambas elecciones independientes con probabilidad uniforme sobre las posibilidades y construimos el conjunto aleatorio $A_{\sigma,m} = \{\sigma(1), \ldots, \sigma(m)\}$, ahora calculamos $\mathbb{P}(A_{\sigma,m} \in \boldsymbol{F})$. Para cada elección de σ hay a lo más una elección de m para la cual el conjunto aleatorio generado pertenece a $\boldsymbol{F}$, ya que si hubiera dos valores i, j con $i < j$ para los cuales $A_{\sigma,i}, A_{\sigma,i} \in \boldsymbol{F}$ entonces $A_{\sigma,j} \supset A_{\sigma,i}$ lo cual sería una contradicción, por lo tanto, como para cada σ hay a lo más una elección de m, $\mathbb{P}(A_{\sigma,m} \in \boldsymbol{F}) \leq \frac{1}{n+1}$.

Ahora calculemos la probabilidad de elegir un subconjunto en particular A^* con k elementos, ésta probabilidad es $\frac{1}{(n+1)\binom{n}{k}}$ ya que $\mathbb{P}(m = k) = \frac{1}{n+1}$ y $\mathbb{P}(A = A^*|m = k) = \frac{1}{\binom{n}{k}}$ ya que entre los conjuntos de k elementos la probabilidad es uniforme.

Tenemos que

$$\sum_{B \in \boldsymbol{F}} \frac{1}{(n+1)\binom{n}{|B|}} = \sum_{B \in \boldsymbol{F}} \mathbb{P}(A = B) = \mathbb{P}(A \in \boldsymbol{F}) \leq \frac{1}{n+1},$$

multiplicando por $n+1$ hemos demostrado la desigualdad de Lubell-Yamamoto-Meshalkin (LYM) la cual dice que, bajo las mismas hipótesis del teorema, se cumple que $\displaystyle\sum_{B \in \boldsymbol{F}} \frac{1}{\binom{n}{|B|}} \leq 1$.
Usando LYM y el hecho de que $\binom{n}{k}$ es máximo cuando $k = \lfloor n/2 \rfloor$ tenemos que

$$\frac{|\boldsymbol{F}|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = \sum_{B \in \boldsymbol{F}} \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq \sum_{B \in \boldsymbol{F}} \frac{1}{\binom{n}{|B|}} \leq 1,$$

de donde $\boldsymbol{F} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. $\square$

Una pregunta natural es: ¿para cuáles conjuntos se alcanza la cota? LYM responde en parte esta pregunta, si $1 = \frac{|\boldsymbol{F}|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq \sum_{B \in \boldsymbol{F}} \frac{1}{\binom{n}{|B|}} \leq 1$, entonces $\binom{n}{|B|} = \binom{n}{\lfloor \frac{n}{2} \rfloor}$ $\forall B \in \boldsymbol{F}$, de donde $|B| = \lfloor \frac{n}{2} \rfloor$ o $|B| = \lceil \frac{n}{2} \rceil$.

Corolario 1.1. *Sea n par y sea $\boldsymbol{F}$ un conjunto de subconjuntos de $N = \{1, 2, \ldots, n\}$ tal que no existen $A, B \in \boldsymbol{F}$ tales que $A \subset B$ y además $|\boldsymbol{F}| = \binom{n}{\lfloor \frac{n}{2} \rfloor}$, entonces $\boldsymbol{F} = N_{\frac{n}{2}}$, donde N_k denota el conjunto de los subconjuntos de N de exactamente k elementos.*

Demostración. Si $n = 2s$ entonces hay exactamente $\binom{2s}{s}$ subonjuntos de N de cardinalidad $\lfloor \frac{n}{2} \rfloor = s$ o cardinalidad $\lceil \frac{n}{2} \rceil = s$ entonces el único conjunto posible de $\binom{n}{\lfloor \frac{n}{2} \rfloor} = \binom{2s}{s}$ subconjuntos de N es aquél que contiene todos los subconjuntos de s elementos de N, es decir N_s, el cual claramente cumple que no existen $A, B \in N_k$ tales que $A \subset B$. □

En el caso $n = 2s + 1$, los elementos de $\boldsymbol{F}$ son conjuntos de cardinalidad $\lfloor \frac{n}{2} \rfloor = s$ o de cardinalidad $\lceil \frac{n}{2} \rceil = s + 1$, entonces en este caso lo único que nos dice LYM es que los conjuntos $\boldsymbol{F}$ que alcanzan la cota sólo tienen como elementos conjuntos de s o $s + 1$ elementos de N, es decir, un subconjunto de cardinalidad $\binom{2s+1}{s+1}$ del conjunto $N_s \cup N_{s+1}$, un total de $\binom{2d}{d}$ posibilidades, donde $\binom{2s+1}{s+1} = d$, sin embargo la demostración presentada, a diferencia de otras, permite probar de manera no tan complicada el siguiente resultado:

Corolario 1.2. *Sea n impar y sea $\boldsymbol{F}$ un conjunto de subconjuntos de $N = \{1, 2, \ldots, n\}$ tal que no existen $A, B \in \boldsymbol{F}$ tales que $A \subset B$ y además $|\boldsymbol{F}| = \binom{n}{\lfloor \frac{n}{2} \rfloor}$, entonces $\boldsymbol{F} = N_{\frac{n-1}{2}}$ o $\boldsymbol{F} = N_{\frac{n+1}{2}}$.*

Demostración. Antes de dar la demostración del corolario 2, daremos una definición y probaremos un lema.

Definición 1.8. *Sean A y B dos subconjuntos diferentes de s elementos de $N = \{1, 2, \ldots, s, s+1, \ldots, 2s+1\}$, decimos que A y B son vecinos si difieren en exactamente un elemento, es decir, existen $a, b \in N$ tales que $A \cup \{b\} - \{a\} = B$.*

Lema 1.5. *Si $X \notin \boldsymbol{F}$, $|X| = s$ y X y Y son vecinos, entonces $Y \notin \boldsymbol{F}$.*

Demostración. Sea $n = 2s + 1$. Regresando a la demostración del Teorema de Sperner, notemos que si en LYM se da la igualdad entonces dividiendo entre $n + 1$ en $(^1)$ tenemos que $\mathbb{P}(A \in \boldsymbol{F}) = \frac{1}{n+1}$ pero entonces eso significa que para cada σ existe **exactamente** un m tal que $A_{\sigma,m} \in \boldsymbol{F}$ y ese m, por lo dicho anteriormente, debe ser s o $s + 1$ entonces si hubiera un conjunto E de $s + 1$ elementos tales que $X \supset E$ y $X, E \notin \boldsymbol{F}$ al tomar una permutación σ^* tal que $\sigma^*(s) = X$ y $\sigma^*(s + 1) = E$ tendríamos una permutación tal que para ningún valor de m, $A_{\sigma^*,m} \in \boldsymbol{F}$ por lo que $\mathbb{P}(A \in \boldsymbol{F}) < \frac{1}{n+1}$, lo cual sería una contradicción, por lo tanto como $X \notin \boldsymbol{F}$ entonces $X \cup \{c\} \in \boldsymbol{F}$ $\forall c \in N - X$. Sea $X \cup \{x\} - \{y\} = Y$ entonces $X \cup \{x\} \in \boldsymbol{F}$ (nótese que $x \in N - X$, de lo contrario Y no tendría s elementos o X sería Y), ahora nos fijamos en $X \cup \{x\}$ que tiene $s+1$ elementos y $Y = X \cup \{x\} - \{y\}$ que tiene s elementos, ambos subconjuntos no pueden pertencer a $\boldsymbol{F}$ pues $X \cup \{x\} \supset Y$, por lo tanto $Y \notin \boldsymbol{F}$ que es lo que queríamos demostrar. □

Continuamos con la demostración del corolario. Supongamos que existe $D \notin \boldsymbol{F}$ tal que $|D| = s$, consideremos cualquier otro conjunto G de cardinalidad s. Supongamos $|D \cap G| = r$ y sean $D - G = \{d_1, d_2, \ldots, d_{s-r}\}$, $G - D = \{g_1, g_2, \ldots, g_{s-r}\}$, $D_i = (D_{i-1} \cup \{g_i\}) - \{d_i\}$, donde definimos $D_0 = D$. Notemos que D_i y D_{i+1} son *vecinos* para $i = 0, 1, \ldots, s - r - 1$ y que $D - s - r - 1 = G$ entonces aplicando el lema a las parejas D_0 y D_1, D_1 y $D_2 \ldots D_{s-r-1}$ y D_{s-r} se obtiene que $D_i \notin \boldsymbol{F}$ para $i = 0, 1, \ldots, s - r$, en particular $G \notin \boldsymbol{F}$, por lo tanto ningún conjunto de cardinalidad s pertenece a $\boldsymbol{F}$ y entonces como $|\boldsymbol{F}| = \binom{2s+1}{s+1}$ y sólo quedan los $\binom{2s+1}{s+1}$ conjuntos de cardinalidad $s + 1$, entonces $\boldsymbol{F} = N_{\frac{n+1}{2}}$. Si nuestro supuesto de existencia no es cierto, eso significa que todo conjunto de cardinalidad s pertenece a $\boldsymbol{F}$ y en ese caso tenemos que $\boldsymbol{F} = N_{\frac{n-1}{2}}$. □

Claramente $N_{\frac{n+1}{2}}$ y $N_{\frac{n-1}{2}}$ cumplen las hipótesis del teorema pues tienen cardinalidad $\binom{2s+1}{s+1}$ y dentro de ellos no hay un conjunto que contenga a otro diferente pues, al ser conjuntos de la misma cardinalidad, una contención implicaría la igualdad.

Entonces una vez probados estos corolarios podemos enunciar el teorema de Sperner de la siguiente manera:

Teorema 1.10. *Sea* $\boldsymbol{F}$ *un conjunto de subconjuntos de* $N = \{1, 2, \ldots, n\}$ *tal que no existen* $A, B \in \boldsymbol{F}$ *tales que* $A \subset B$*, entonces* $|\boldsymbol{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$ *y la igualdad se alcanza si y sólo si* $\boldsymbol{F} = N_{\lceil \frac{n-1}{2} \rceil}$ *o* $\boldsymbol{F} = N_{\lfloor \frac{n+1}{2} \rfloor}$.

Antes de enunciar nuestro siguiente corolario recordemos la notación asintótica usual.

Para dos funciones f y g decimos que $f = O(g)$ si $f \leq cg$ para valores suficientemente grandes de las variables en ambas funciones donde c es una constante positiva. Si el límite de la razón $\frac{f}{g}$ tiende a cero cuando las variables de las funciones tienden a infinito, decimos que $f = o(g)$, finalmente si f/g tiende a uno cuando las variables de las funciones tienden a infinito, es decir $f = (1 + o(1))g$ decimos que $f \sim g$.

Corolario 1.3. *Lema de Littlewood-Offord. Sean* $x_1, x_2, \ldots, x_n$ *números reales diferentes de cero. Supongamos que* $c_1, c_2, \ldots, c_n$ *son variables aleatorias independientes cada una de las cuales es* ± 1 *con la misma probabilidad. Probar que* $\mathbb{P}\{c_1x_1 + c_2x_2 + \ldots + c_nx_n = 0\} \leq O\left(\frac{1}{\sqrt{n}}\right)$.

Demostración. Sin pérdida de generalidad podemos asumir que $x_i > 0$ pues $\mathbb{P}(c_ix_i > 0) = \mathbb{P}(c_ix_i < 0) = \frac{1}{2}$. A cada conjunto de valores de los c_i le asociamos el subconjunto de $N = \{1, 2, \ldots, n\}$ que corresponde a cuáles índices i cumplen que $c_i = 1$ y en $\boldsymbol{F}$ ponemos a todos los subconjuntos asociados que cumplen que $\sum c_i x_1 = 0$ entonces $\boldsymbol{F}$ es una anticadena pues si hubiera dos conjuntos diferentes tales que $A \subset B$ entonces la suma asociada al conjunto B es estrictamente mayor a la suma asociada al conjunto A pues los x_i son positivos. Por el lema de Sperner $|\boldsymbol{F}| \leq \binom{n}{\lfloor n/2 \rfloor}$ y entonces la probabilidad buscada es a lo más $\binom{n}{\lfloor n/2 \rfloor}/2^n$ que usando la fórmula de Stirling es aproximadamente $\sqrt{2/n\pi}$ que es de orden $\frac{1}{\sqrt{n}}$ como queríamos demostrar. $\square$

Es importante notar cómo un argumento probabilistico fue útil para probar algo determinista y cómo ese lema determinista fue útil para contestar una pregunta planteada en términos probabilísticos, esa es la idea que se quiere plasmar con este capítulo, que si a lo largo de nuestra aprendizaje de la probabilidad hemos usado herramientas de otras áreas, también es posible usar la probabilidad como herramientas en muchas otras ramas de las matemáticas.

Capítulo 2

El teorema de Hales-Jewett

2.1. Nueva forma de hacer matemáticas; proyecto polymath.

El teorema que abordaremos en este capítulo usa el método probabilístico de una manera más sofisticada, dicho teorema fue tratado en el primer *proyecto polymath*. Antes de entrar en el tema discutiremos un poco acerca de qué es un proyecto polymath y acerca de la historia del paper en el cual se basa este capítulo de la tesis.

Polymath es una palabra que proviene del inglés y significa: *una persona que sabe bastante acerca de muchos diferentes temas, erudito*. Sin embargo, podemos ver que la palabra también puede ser vista como la unión de dos palabras; *poly* y *math* que nos dan un significado de *muchas matemáticas* el cual degenera naturalmente en *muchos matemáticos* que es precisamente la idea central de lo que consiste un proyecto polymath y la razón por la que usaré la palabra polymath como préstamo lingüistico en vez de usar una traducción literal de ella.

La idea de los proyectos polymath fue pensada inicialmente por el ganador de la medalla Fields en 1998, el matemático inglés Timothy Gowers quién en enero de 2009 publica en su blog un post llamado *¿Es posible la colaboración matemática masiva?*, en este post se discute la idea de una nueva posibilidad de hacer matemáticas, grupalmente, pero no se refiere a un grupo de 3 o 4 colaboradores, el piensa en un proyecto donde primero que nada, el problema no sea claramente dividible en varios subproblemas y que contenga un importante número de personas trabajando en él.

Gowers afirma que no sabe si esto tendrá ventajas pero da razones para creer que existen, entre ellas menciona que por pura probabilidad, algunas veces se necesita de un poco de suerte para tener la idea adecuada y entre más personas hayan pensando en el problema es más probable que se le ocurra a alguien, pero no sólo eso, sino que el conocimiento que tiene un grupo de personas es casi siempre mayor que el de cualquiera de sus integrantes. Otra de las ventajas es que en algunas ocasiones, los matemáticos pasan mucho tiempo tratando de demostrar algo que conjeturaron y puede que sea falso o de hecho que ya este probado, el conocimiento de este hecho puede reducir considerablemente el tiempo en la resolución del problema.

Esta idea fue rápidamente difundida y al día siguiente, otro matemático prominente, Terence Tao, ganador de la medalla Fields en 2006, decide unirse a la discusión aportando su experiencia, sus ideas y su trabajo convirtiéndose así en uno de los representates más importantes de los proyectos polymath.

A noviembre del 2011 se han realizado 6 proyectos polymath y 3 mini-proyectos polymath:

- Polymath1 Nueva prueba y cotas para el teorema de Hales Jewett. Inició el 1 de febrero de 2009 y ya hay un artículo con los resultados.
- Polymath2 Espacios de Banach definibles. Inició el 17 de febrero del 2009.
- Polymath3 La conjetura del polinomio de Hirsh. Propuesto el 17 de julio del 2009 pero empezó el 30 de septiembre del 2010.
- Polymath4 Una manera determinista de encontrar primos. Propuesto el 27 de julio del 2009 y empezó el 9 de agosto del mismo año. Los resultados de la investigación están en proceso de publicación.
- Polymath5 El problema de discrepancia de Erdös. Empezó el 19 de enero del 2010.
- Polymath6 Mejora en las cotas del teorema de Roth. Propuesto el 2 de febrero del 2011.

Los mini-proyectos polymath han sido encabezados por Terence Tao y se han realizado en 2009, 2010 y 2011. La duración de ellos ha sido corta, de dos días, cinco horas y un poco más de una hora (hasta la primer solución completa), respectivamente y se han basado en problemas de la Olimpiada Internacional de Matemáticas, IMO, del respectivo año. El objetivo de este tipo de proyectos es experimentar con ciertos aspectos del trabajo colectivo. El del último año tuvo como objetivo analizar la colaboración masiva en un problema geométrico.

- Mini-polymath1 Problema 6 de la IMO 2009.
- Mini-polymath1 Problema 5 de la IMO 2010.
- Mini-polymath1 Problema 2 de la IMO 2011.

Sin duda alguna estos proyectos revolucionarán la forma de hacer matemáticas y quizá otros tipos de investigaciones científicas, en un mundo globalizado y con los avances tecnólogicos, este tipo de investigaciones tendrán un gran auge en un futuro no muy lejano.

Ahora nos centraremos en el proyecto relevante para la tesis, el proyecto polymath1. Este proyecto empieza el 1 de febrero del 2009, ese mismo día, Gowers escribe en su blog sobre los objetivos del proyecto, afirma que la intención no es encontrar una solución combinatoria para el teorema de Hales-Jewett en el caso $k = 3$, que a él le encantaría que sucediera eso pero que el objetivo del experimento es más modesto: obtener un avance en la solución o dar argumentos convicentes para concluir que dicho avance no es posible. Sin embargo, seis semanas después, el 6 de marzo, Tim anuncia que piensa que el problema ya está resuelto y que además hay indicios para creer que no sólo el caso 3 se puede resolver, cree que es posible dar una solución en general logrando así la prueba más elemental del teorema de Szemerédi. Finalmente en octubre del 2009 terminan la primera versión del artículo en el que exponen la solución completa al problema.

2.2. Introducción a *una nueva prueba del teorema DHJ.*

En esta parte abordaremos el artículo titulado: *A new proof of the density Hales-Jewett theorem* (arXiv:0910.3926v2), daremos un panorama general del artículo y demostraremos detalladamente algunos de los teoremas que ahí aparecen, esperando que esto sirva como una buena introducción al artículo así como guía si se desea hacer una lectura completa del mismo.

El teorema principal del artículo fue demostrado por primera vez por

Furstenberg y Katznelson en 1991 y generaliza al teorema de Hales-Jewett en la misma forma en la que el teorema de Szemerédi generaliza al teorema de van der Waerden. Enunciaremos estos dos últimos teoremas.

Teorema 2.1. *Teorema de van der Waerden. Para todo par de naturales k y r, existe N tal que en cualquier r-coloración de $[N]$ existe una progresión aritmética monocromática de longitud k.*

En el enunciado del teorema se usa $[N]$ para denotar el conjunto $\{1, 2, \ldots, n\}$, una r-coloración de un conjunto X es una función $f : X \to [r]$, un subconjunto Y de X es monocromático si $f(y)$ es el mismo valor para todo $y \in Y$.

Al menor valor de N que satisface la condición del teorema se le conoce como número de van der Waerden y se le denota $W(k, r)$, al igual que en casi todos los teoremas estilo Ramsey, se conocen muy pocos valores no triviales, así, a pesar de que este teorema fue probado hace casi 75 años sólo se conocen 6 de los números de van der Waerden: $W(2, 3) = 9$, $W(2, 4) = 35$, $W(2, 5) = 178$, $W(2, 6) = 1132$, $W(3, 3) = 27$ y $W(4, 3) = 76$.

Teorema 2.2. *Teorema de Szemerédi. Para toda pareja (k, δ) con k entero positivo y $\delta > 0$, existe N tal que cualquier subconjunto A de $[N]$ de cardinalidad al menos $N\delta$ contiene una progresión aritmética de longitud k.*

Conjeturado en 1936 por Erdös y Turán, Szemerédi prueba en 1975 la versión densidad del teorema de van der Waerden Se le llama así pues si consideramos $\frac{|A|}{N}$ como la densidad de A dentro de $[N]$, la condición del teorema es que A tiene densidad al menos δ y teniendo en mente el teorema de van der Waerden, la conclusión es que un color que se usa *frecuentemente* contiene una progresión aritmética.

Proposición 2.1. *Szemerédi $\Rightarrow$ van der Waerden.*

Demostración. Sean k y r números naturales, hagamos $\delta = 1/r$ entonces por Szemerédi existe N_0 tal que cualquier subconjunto de $[N_0]$ de densidad al menos $1/r$ contiene una progresión aritmética entonces podemos concluir que hacer $N = N_0$ en el teorema de van der Waerden funciona, pues claramente existe un color tal que la cardinalidad de los números de ese color es al menos $1/r$ y en ese subconjunto hay una progresión aritmética de longitud k por Szemerédi, que de hecho resulta ser monocromática. □

Antes de enunciar el teorema de Hales-Jewett es necesario hacer una definición.

Definición 2.1. *Una línea combinatoria es un subconjunto de* $[k]^n$ *con* k *elementos llamados puntos y puede ser pensado apartir de un elemento del conjunto* $([k] \cup \{*\})^n$ *con al menos una coordenada que tome el valor comodín* $*$*. Para obtener los* k *puntos de la línea combinatoria, se hace variar* j *desde* 1 *hasta* k *y se sustituyen todas las coordenas con comodín por el valor* j*. Por ejemplo, para* $k = 3$ *y* $n = 8$ *podemos considerar el punto* $(*, 3, *, 2, 2, *, 1, 2) \in ([k] \cup \{*\})^n$ *que genera la línea*

$$\{(\boldsymbol{1}, 3, \boldsymbol{1}, 2, 2, \boldsymbol{1}, 1, 2), (\boldsymbol{2}, 3, \boldsymbol{2}, 2, 2, \boldsymbol{2}, 1, 2), (\boldsymbol{3}, 3, \boldsymbol{3}, 2, 2, \boldsymbol{3}, 1, 2)\}$$

que es un subconjunto de k *elementos* $[k]^n$*.*

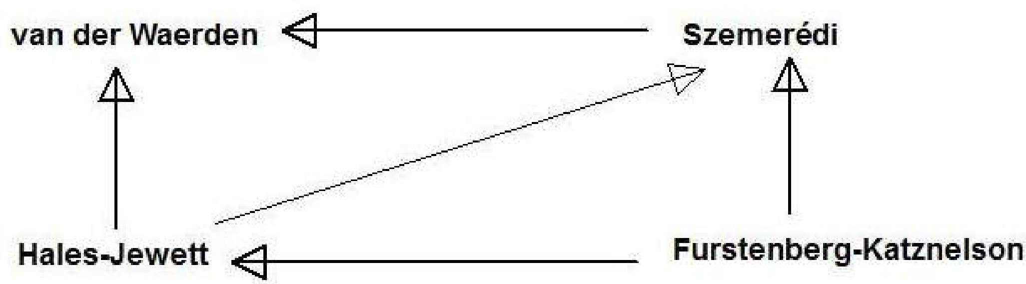
Es claro que si uno permite líneas degeneradas en las cuales no haya coordenadas con comodín, entonces habría una correspondencia uno a uno entre líneas combinatorias en $[k]^n$ y puntos en $[k+1]^n$, pero como se exije que al menos una coordenada tome el valor comodín, ya no hay biyección y sólo se trata de una inyección.

Teorema 2.3. *Teorema de Hales-Jewett. Para todo par de enteros positivos* k *y* r*, existe un entero positivo* $HJ(k, r)$ *tal que para cualquier* $n \geq HJ(k, r)$ *y cualquier* r*-coloración del conjunto* $[k]^n$ *existe una línea combinatoria monocromática.*

Así como en el teorema de van der Waerden, podemos considerar la versión densidad, donde la densidad de $A \subseteq [k]^n$ es $|A| / k^n$. Este teorema también es conocido como teorema de Furstenberg-Katznelson.

Teorema 2.4. *Versión densidad del teorema de Hales-Jewett. Para todo entero positivo k y todo $\delta > 0$ existe un entero positivo $DHJ(k, \delta)$ tal que si $n \geq DHJ(k, \delta)$ y A es cualquier subconjunto de $[k]^n$ de densidad al menos δ, entonces A contiene una línea combinatoria.*

De manera análoga a la prueba de que Szemerédi implica van der Waerden, se ve que la versión densidad de Hales-Jewett implica Hales-Jewett. Sin embargo esa no es la única relación que guardan entre sí estos cuatro teoremas, también es cierto que Hales-Jewett implica van der Waerden de la misma forma en la que la versión densidad de Hales-Jewett implica la versión densidad de van der Waerden, es decir, el teorema de Szemerédi. Además, aunque de este hecho no daremos prueba en este trabajo, es posible demostrar que Hales-Jewett implica el teorema de Szemerédi, lo cual completa el siguiente esquema jerárquico:



Las implicaciones horizontales ya fueron discutidas, ahora daremos la prueba de una de las implicaciones verticales, la otra vertical es análoga.

Proposición 2.2. *Hales-Jewett $\Rightarrow$ van der Waerden.*

Demostración. Para demostrar esto (y también para demostrar la otra implicación) la idea básica será asociar a una línea combinatoria una progresión aritmética. Para esto pensemos a $[m]$ como $\{0, 1, 2, \ldots, m-1\}$ en lugar de $\{1, 2, 3, \ldots, m\}$ e identifiquemos a los números en $[N]$ con su representación en base k en $[k]^n$, entonces es claro que una línea combinatoria en $[k]^n$ corresponde a una progresión aritmética de longitud k en $[N]$ pues si S es el conjunto de coordenadas con comodín en la línea entonces la diferencia común de la progresión sería $\sum_{i\in S} k^{n-i}$.
Sean k y r números naturales, por Hales-Jewett existe un entero positivo $HJ(k, r)$ tal que para cualquier $n \geq HJ(k, r)$ y cualquier r-coloración del conjunto $[k]^n$ existe una línea combinatoria monocromática. Entonces en el teorema de van der Waerden proponemos $N = k^{HJ(k,r)}$, por Hales-Jewett sabemos que en $[k]^{HJ(k,r)}$ hay una línea combinatoria y por la observación de arriba entonces en $[N]$ hay una progresión aritmética monocromática. $\square$

Una de las razones más importantes para encontrar una nueva prueba para la versión densidad de Hales-Jewett es que implica directamente el teorema de Szemerédi lo cual *parece que siempre ilumina*, afirman, además esta nueva forma de probar Szemerédi es quizá la forma más sencilla de demostrarlo y realmente es una prueba diferente pues en ella se ve gran similitud en probar DHJ_3, osea cuando $k = 3$, y la prueba en general, cosa que no se observa en ninguna de las pruebas antes conocidas.

Por último en esta subsección, enunciaremos la versión multidimensional de la versión densidad del teorema Hales-Jewett a la cual nos referiremos como $MDHJ$ dejando DHJ para referirnos a la versión densidad del teorema Hales-Jewett y HJ para la primer versión.

De manera natural uno puede plantearse por la existencia de espacios de dimensión mayor en conjuntos densos de $[k]^n$. Sabiendo que existen líneas, ¿por qué no planos o subespacios d-dimensionales? Furstenberg y Katznelson notaron que esto es consecuencia directa de DHJ, daremos la demostración de este hecho, pero antes definiremos los espacios que estamos buscando.

Definición 2.2. *Un subespacio combinatorio d-dimensional en* $[k]^n$ *es la generalización de una línea combinatoria, ahora hay d comodines en lugar de solo uno. Hagamos una partición de* $[n]$ *en* $k+d$ *conjuntos* $X_1, X_2, \ldots, X_k, C_1, C_2 \ldots, C_d$ *tal que* $C_1, C_2 \ldots, C_d$ *son no vacíos y entonces el subespacio consistiría de todos los puntos* x *tales que* $x_i = j$ *si* $i \in X_j$ *y* x *es constante en cada conjunto* W_r *y toma todos los posibles valores desde* 1 *hasta* k *generando así los* k^d *puntos del subespacio, por ejemplo, si* $k = 3$*,* $d = 2$*,* $n = 7$ *y tenemos* $X_1 = \{3, 4\}$*,* $X_2 = \emptyset$*,* $X_3 = \{1, 7\}$*,* $C_1 = \{2, 6\}$ *y* $C_2 = \{5\}$ *entonces nuestro subespacio* 2*-dimensional estaría conformado por los puntos:* $(3, \boldsymbol{1}, 1, 1, \boldsymbol{1}, \boldsymbol{1}, 3)$, $(3, \boldsymbol{1}, 1, 1, \boldsymbol{2}, \boldsymbol{1}, 3)$, $(3, \boldsymbol{1}, 1, 1, \boldsymbol{3}, \boldsymbol{1}, 3)$, $(3, \boldsymbol{2}, 1, 1, \boldsymbol{1}, \boldsymbol{2}, 3)$, $(3, \boldsymbol{2}, 1, 1, \boldsymbol{2}, \boldsymbol{2}, 3)$, $(3, \boldsymbol{2}, 1, 1, \boldsymbol{3}, \boldsymbol{2}, 3)$, $(3, \boldsymbol{3}, 1, 1, \boldsymbol{1}, \boldsymbol{3}, 3)$, $(3, \boldsymbol{3}, 1, 1, \boldsymbol{2}, \boldsymbol{3}, 3)$, $(3, \boldsymbol{3}, 1, 1, \boldsymbol{3}, \boldsymbol{3}, 3)$.

Notemos que hay una inyección del conjunto de los subespacios combinatorios de $[k]^n$ en $[k+d]^n$ pues basta tomar los d comodines como los valores $k+1.$ $k+2, \ldots, k+d$, también notemos que si se permite que los conjuntos C_i sean vacíos, entonces la inyección se convierte en una biyección.

Teorema 2.5. *MDHJ* (k, d, δ) *Para toda* $\delta > 0$ *y cada par de enteros positivos* k *y* d*, existe un entero* $MDHJ(k, d, \delta)$ *tal que para todo* $n \geq MDHJ(k, d, \delta)$ *y cualquier subconjunto* $A \subseteq [k]^n$ *de densidad al menos* δ*, resulta que* A *contiene un subespacio combinatorio* $d-$*dimensional de* $[k]^n$*.*

De manera similar a DHJ, denotaremos $MDHJ_k$ cuando la k esté fija.

Proposición 2.3. *Para toda* k*,* $DHJ_k \Rightarrow MDHJ_k$*.*

Demostración. Probaremos el resultado por inducción en d. Para $d = 1$ se tiene que $MDKJ_k$ y DHJ_k son de hecho los mismos teoremas entonces ya tenemos la base de inducción. Ahora supongamos que sabemos que $MDHJ_k$ es cierto para dimensión $d-1$ y sea $A \subseteq [k]^n$ de densidad al menos δ. Sea $m = MDHJ(k, d-1, \delta/2)$ el cual por hipótesis de inducción sabemos que existe. Escribamos un punto $z \in [k]^n$ como (x, y) donde $x \in [k]^m$ y $y \in [k]^{n-m}$.

Decimos que un punto $y \in [k]^{n-m}$ es *bueno* si $A_y = \{x \in [k]^m : (x, y) \in A\}$ tiene densidad al menos $\delta/2$ en $[k]^m$. Sea $G \subseteq [k]^{n-m}$ el conjunto de todos los puntos buenos. Entonces la densidad de G en $[k]^{n-m}$ es al menos $\delta/2$ o de lo contrario se tiene que $|A| = |A \cap \{y \in G\}| + |A \cap \{y \notin G\}| < (\delta/2) * k^{n-m} * 1 * k^n + (1-\delta/2) * k^{n-m} * (\delta/2) * k^n = (\delta/2)(2-\delta/2)k^n < k^n\delta$, lo cual es una contradicción pues la densidad de A en $[k]^n$ es al menos δ.

Por hipótesis de inducción, para cada y bueno, el conjunto A_y contiene un subespacio combinatorio $(d-1)-$dimensional en $[k]^m$. Por la inyección antes mencionada, sabemos que a lo más hay $M = (k+d-1)^n$ subespacios combinatorios $(d-1)-$dimensionales en $[k]^m$, por lo tanto debe existir un subespacio $\sigma \subseteq [k]^m$ tal que el conjunto
$G_\sigma = \left\{y \in [k]^{n-m} : (x, y) \in A \ \forall x \in \sigma\right\}$ tiene densidad al menos $(\delta/2)/M$ en $[k]^{n-m}$. Si tomamos $n \geq m + DHJ(k, \delta/2M) := MDHJ_k$, podemos concluir por DHJ_k que G_σ contiene una línea combinatoria, λ. Entonces $\sigma \times \lambda$ es el subespacio combinatorio $d-$dimensional contenido en A que buscamos, lo que concluye la prueba. $\square$

2.3. Teorema de Sperner y su versión multidimensional.

El caso $k = 2$ de DHJ es equivalente a lo siguiente: para cada $\delta > 0$ existe n tal que si $\mathbf{F}$ es una colección de al menos $\delta 2^n$ subconjuntos de $[n]$ entonces existen dos conjuntos distintos A y B contenidos en $\mathbf{F}$ tales que $A \subset B$. La equivalencia se ve si uno observa que un subconjunto C de $[n]$ se puede pensar como un punto en $\{0,1\}^n$ donde en la $j-$ésima coordenada hay un 1 si y sólo si $j \in C$, de esta forma vemos que el par $\{A, B\}$ con $A \subset B$ corresponde a una línea combinatoria en $\{0,1\}^n$.

En el capítulo anterior probamos que si $|\mathbf{F}| \geq \binom{n}{\lfloor n/2 \rfloor}$ entonces habrá dos conjuntos diferentes tales que uno de ellos contiene al otro y puesto que $\lim_{n\to\infty} \binom{n}{\lfloor n/2 \rfloor}/2^n = 0$, $(\binom{n}{\lfloor n/2 \rfloor}/2^n$ es de orden $1/\sqrt{n})$, entonces se tiene demostrado DHJ para el caso $k = 2$.

Ahora enunciaremos la versión multidimensional del teorema de Sperner la cual fue demostrada por Gunderson, Rödl y Sidorenko en 1999. La prueba que se usó como base es la de [6], la que presentamos es una versión más detallada y con unas pequeñas correcciones a la original.

Teorema 2.6. *Sperner multidimensional. Sea* $\boldsymbol{F}$ *una colección de conjuntos de* $[n]$ *que no contiene un subespacio combinatorio* $d-$*dimensional. Entonces la densidad de* $\boldsymbol{F}$ *es a lo más* $(25/n)^{1/2^d}$.

Demostración. En este caso será más útil caracterizar a los subespacios combinatorios $d-$dimensionales de la siguiente forma; tomemos una colección de conjuntos disjuntos no vacíos $C_0, C_1, \ldots, C_d$ tales que $C_0 \cup \bigcup_{i\in E} C_i \in \mathbf{F}$ para todo $E \subseteq \{1, 2, \ldots, d\}$, ($C_0$ son las coordenadas fijas y C_j son las coordenadas con el $j-$ésimo comodín).

Sea δ la densidad de $\mathbf{F}$, es decir, $\mathbf{F}$ tiene cardinalidad $\delta 2^n$, veamos que podemos suponer que $\delta < 1/2$, pues en caso contrario tomamos un subconjunto de $\mathbf{F}$ que sí lo cumpla, dicho subconjunto seguirá sin tener subespacios combinatorios $d-$dimensionales. La idea de la demostración es probar que si $n \geq 25/\delta^{2^d}$ entonces podremos construir los conjuntos C_j mencionados anteriormente y entonces el subconjunto de **F** tendría el subespacio y por lo tanto **F** tendría un subespacio combinatorio $d-$dimensional, lo cual por hipótesis es falso, entonces esto nos llevará a concluir que la densidad de **F** es a lo más $(25/n)^{1/2^d}$.

Para $i = 1, 2, \ldots, d-1$ definimos $n_i = \lfloor n/4^{d-i} \rfloor$ ($\lfloor x \rfloor$ denota la parte entera de x) y definimos $n_d = n - (n_1 + n_2 + \cdots + n_{d-1})$, notemos que $n_d \geq n - (n/4^{d-1} + n/4^{d-2} + \cdots + n/4) \geq n - n\sum_{i=1}^{\infty} 1/4^i = 2n/3$. Además notemos que la sucesión de los n_i es no decreciente.

Ahora hagamos una partición de $[n]$ en conjuntos $J_1 \cup J_2 \cup \cdots \cup J_{d-1} \cup E$ con $|J_i| = n_i$ y entonces $|E| = n_d$. Consideremos la siguiente forma de elegir un subconjunto aleatorio A de $[n]$. Primero escojemos una permutación aleatoria y con probabilidad uniforme π de $[n]$. Luego elegimos un entero s de manera aleatoria con distribución binomial de parámetros n_1 y $1/2$, es decir $\mathbb{P}(s = k) = \binom{n_1}{k}(\frac{1}{2})^{n_1}$. Después, elegimos B como un subconjunto aleatorio y con probabilidad uniforme del conjunto $V = \{\pi(n_1+1), \ldots, \pi(n)\}$. Finalmente, hacemos $A = \{\pi(1), \ldots, \pi(s)\} \cup B = A_{\pi,s} \cup B$. Lo que resulta es que elegir al conjunto A de esta forma es lo mismo que elegir el conjunto A aleatoriamente y con probabilidad uniforme de entre los subconjuntos de $[n]$ lo cual probaremos a continuación. Notemos que en la forma en que elegimos al conjunto A hay involucradas tres medidas de probabilidad; la de π, la de s y la de B a las que denotaremos $\mathbb{P}_\pi$, $\mathbb{P}_s$ y $\mathbb{P}_B$, respectivamente.

Para probar que A tiene distribución uniforme probaremos una equivalencia, probaremos que $\mathbb{P}_T(a \in A) = 1/2\ \forall a \in [n]$ donde la medida de probabilidad $\mathbb{P}_T$ es la medida producto de las tres medidas de probabilidad mencionadas.

$$
\begin{aligned}
\mathbb{P}_T(a \in A) &= \mathbb{P}_s(\mathbb{P}_B(\mathbb{P}_\pi(a \in A))) \\
&= \mathbb{P}_s(\mathbb{P}_B(\mathbb{P}_\pi(a \in A_{\pi,s} \cup B))) \\
&= \mathbb{P}_s(\mathbb{P}_B(\mathbb{P}_\pi(a \in A_{\pi,s}) + \mathbb{P}_\pi(a \in B))) \\
&= \mathbb{P}_s(\mathbb{P}_B(\mathbb{P}_\pi(a \in A_{\pi,s} \mid a \in V^c)\mathbb{P}_\pi(a \in V^c) \\
&\qquad + \mathbb{P}_\pi(a \in B \mid a \in V)\mathbb{P}_\pi(a \in V))) \\
&= \mathbb{P}_s(\mathbb{P}_B(\mathbb{P}_\pi(a \in A_{\pi,s} \mid a \in V^c)(n_1/n) \\
&\qquad + \mathbb{P}_\pi(a \in B \mid a \in V)\,[(n-n_1)/n])) \\
&= (n_1/n)\mathbb{P}_B(\mathbb{P}_\pi(\mathbb{P}_s(a \in A_{\pi,s} \mid a \in V^c))) \\
&\qquad + [(n-n_1)/n]\,\mathbb{P}_s(\mathbb{P}_\pi(\mathbb{P}_B(a \in B \mid a \in V))) \\
&= (n_1/n)\mathbb{P}_B(\mathbb{P}_\pi(\mathbb{P}_s(a \in A_{\pi,s} \mid a \in V^c))) + [(n-n_1)/n]\,(1/2) \\
&= (n_1/n)\mathbb{P}_B(\mathbb{P}_\pi(\sum_{k=0}^{n_1}\mathbb{P}_s(a \in A_{\pi,s} \mid a \in V^c, s=k)\mathbb{P}_s(s=k)) + \frac{n-n_1}{2n} \\
&= (n_1/n)\mathbb{P}_B(\sum_{k=0}^{n_1}\mathbb{P}_\pi(\mathbb{P}_s(a \in A_{\pi,s} \mid a \in V^c, s=k)\frac{1}{2^{n_1}}\binom{n_1}{k}) + \frac{n-n_1}{2n} \\
&= (n_1/n)\frac{1}{2^{n_1}}\mathbb{P}_B(\sum_{k=0}^{n_1}\mathbb{P}_s(\mathbb{P}_\pi(a \in A_{\pi,s} \mid a \in V^c, s=k)\binom{n_1}{k}) + \frac{n-n_1}{2n} \\
&= (n_1/2^{n_1}n)\mathbb{P}_B(\sum_{k=0}^{n_1}\mathbb{P}_s(\frac{k}{n_1}\binom{n_1}{k})) + \frac{n-n_1}{2n} \\
&= (1/n)\frac{1}{2^{n_1}}\sum_{k=0}^{n_1}k\binom{n_1}{k} + \frac{n-n_1}{2n} \\
&= (1/n)\frac{1}{2^{n_1}}\left[n_1 2^{n_1-1}\right] + \frac{n-n_1}{2n} \\
&= \frac{1}{2}.
\end{aligned}
$$

Hemos usado la definición de P_T, definición de A, que $A_{\pi,s}$ y B son ajenos, que la medida π es uniforme y Fubini, además usamos entre otras cosas que s es binomial, Fubini y que π es uniforme, además en la penúltima igualdad hemos usado una idéntidad conocida que puede ser demostrada contando de dos formas distintas el número de comités con un presidente elegidos de un grupo de n_1 personas.

Continuando con la demostración, definimos $X_{\pi,s}$ como el conjunto de todos los $B \subseteq V$ tales que $A = A_{\pi,s} \cup B \in \mathbf{F}$. Ahora calculemos $\mathbb{E}_T\left[|X_{\pi,s}|\right]$, es decir, la cardinalidad promedio de $X_{\pi,s}$.

$$\begin{aligned}
\mathbb{E}_T\left[|X_{\pi,s}|\right] &= \mathbb{E}_T\left[|\{B : A_{\pi,s} \cup B \in \mathbf{F}\}|\right] \\
&= \sum_{D \subset V^c} \mathbb{E}_T\left[|\{B : A_{\pi,s} \cup B \in \mathbf{F}\} \mid A_{\pi,s} = D|\right] \mathbb{P}_T\left(A_{\pi,s} = D\right) \\
&= \sum_{D \subset V^c} \sum_{B \subset V} \mathbb{E}_T\left[\mathbb{I}_{A_{\pi,s} \cup B \in \mathbf{F}} \mid A_{\pi,s} = D\right] \mathbb{P}_T\left(A_{\pi,s} = D\right) \\
&= \sum_{D \subset V^c} \sum_{B \subset V} \mathbb{P}_T\left(A_{\pi,s} \cup B \in \mathbf{F} \mid A_{\pi,s} = D\right) \mathbb{P}_T\left(A_{\pi,s} = D\right) \\
&= \sum_{B \subset V} \sum_{D \subset V^c} \mathbb{P}_T\left(A_{\pi,s} \cup B \in \mathbf{F} \mid A_{\pi,s} = D\right) \mathbb{P}_T\left(A_{\pi,s} = D\right) \\
&= \sum_{B \subset V} \mathbb{P}_T\left(A_{\pi,s} \cup B \in \mathbf{F}\right) = 2^{n-n_1}\delta.
\end{aligned}$$

Si $\mu_V(X_{\pi,s}) = |X_{\pi,s}|/2^{|V|}$ representa la densidad del conjunto $X_{\pi,s}$ dentro del conjunto potencia de V, entonces tenemos que $\mathbb{E}_T\left[\mu_V\left(X_{\pi,s}\right)\right] = \mathbb{E}_T\left[|X_{\pi,s}|/2^{|V|}\right] = \frac{1}{2^{|V|}}\mathbb{E}_T\left[|X_{\pi,s}|\right] = \frac{1}{2^{n-n_1}}2^{n-n_1}\delta = \delta$.

Lema 2.1. *Sea X un conjunto finito y sea X_γ un conjunto aleatorio de X donde γ es una variable aleatoria definida en el espacio de probabilidad Γ, es decir, γ determina el conjunto X_γ. Supongamos que $\mathbb{E}_\gamma\left[\mu\left(X_\gamma\right)\right] = \delta$. Sean γ y γ_1 dos variables aleatorias independientes idénticamente distribuidas. Entonces $\mathbb{E}_{\gamma,\gamma_1}\left[\mu\left(X_\gamma \cap X_{\gamma_1}\right)\right] \geq \delta^2$.*

Demostración. Sea ξ_γ la función característica de X_γ, entonces

$$\begin{aligned}
\delta^2 &= \mathbb{E}_\gamma \left[\mu\left(X_\gamma\right)\right]^2 \ \ (Hip.)\\
&= \mathbb{E}_\gamma \left[\mathbb{E}_x \left[\xi_\gamma\left(x\right)\right]\right]^2\\
&= \mathbb{E}_x \left[\mathbb{E}_\gamma \left[\xi_\gamma\left(x\right)\right]\right]^2 \ \ (Fubini)\\
&\leq \mathbb{E}_x \left[1^2\right] \mathbb{E}_x \left[\mathbb{E}_\gamma \left[\xi_\gamma\left(x\right)\right]^2\right] \ \ (Cauchy-Schwarz)\\
&= \mathbb{E}_x \left[\mathbb{E}_\gamma \left[\xi_\gamma\left(x\right)\right] \mathbb{E}_{\gamma_1} \left[\xi_{\gamma_1}\left(x\right)\right]\right]\\
&= \mathbb{E}_x \left[\mathbb{E}_{\gamma,\gamma_1} \left[\xi_\gamma\left(x\right) \xi_{\gamma_1}\left(x\right)\right]\right] \ \ (Independencia)\\
&= \mathbb{E}_x \left[\mathbb{E}_{\gamma,\gamma_1} \left[\xi_{\gamma\cap\gamma_1}\left(x\right)\right]\right]\\
&= \mathbb{E}_{\gamma,\gamma_1} \left[\mathbb{E}_x \left[\xi_{\gamma\cap\gamma_1}\left(x\right)\right]\right] \ \ (Fubini)\\
&= \mathbb{E}_{\gamma,\gamma_1} \left[\mu\left(X_\gamma \cap X_{\gamma_1}\right)\right] \blacksquare
\end{aligned}$$

$\square$

Con este lema podemos concluir que si primero escogemos π aleatoriamente y luego escogemos s y t independientemente y con distribución binomial entonces la desnsidad promedio de $X_{\pi,s} \cap X_{\pi,t}$ es al menos δ^2.
Queremos que s y t sean distintos. La probabilidad de que $s = t$ es

$$\sum_{i=0}^{n_1} \mathbb{P}\left[s = t = i\right] = \sum_{i=0}^{n_1} \mathbb{P}\left[s = i\right] \mathbb{P}\left[t = i\right] = \frac{1}{2^{2n_1}} \sum_{i=0}^{n_1} \binom{n_1}{i}^2 = 2^{-2n_1} \binom{2n_1}{n_1}$$

donde de nuevo hemos usado una idéntidad combinatoria conocida que corresponde a contar caminos que sólo van a la derecha o hacia arriba en una cuadricula de $n_1 \times n_1$ del vétice inferior izquierdo al vértice superior derecho. Aquí en el artículo cometen un error pues afirman que dicha probabilidad es $2^{-n_1}\binom{2n_1}{n_1}$, sin embargo, el error debe ser de escritura ya que afirman que dicha cantidad es menor que $1/\sqrt{n_1}$ lo cual no es cierto pero si se considera la cantidad correcta que es $2^{-2n_1}\binom{2n_1}{n_1}$, sí se cumple que es menor que $1/\sqrt{n_1}$ sin embargo aquí hay otro error pues el artículo afirma que $1/\sqrt{n_1} \leq 2^{d-1}n^{-1/2}$ lo que implicaría que $n \leq 4^{d-1}n_1 = 4^{d-1}\left\lfloor n/4^{d-1}\right\rfloor$ lo cual únicamente es cierto cuando n es múltiplo de 4^{d-1} (si $n = 4^{d-1}K + r$ con $0 < r < 4^{d-1}$ se tiene que $n = 4^{d-1}K + r > 4^{d-1}K = 4^{d-1}n_1$) por lo tanto aquí hay un error que debe ser corregido. La primer opción que pensamos fue cambiar la función piso por la función techo en la definición de los n_i, eso solucionaría esta desigualdad, sin embargo, con ese cambio ya no se cumpliría que $|E| = n_d \geq 2n/3$ y esto lo necesitaremos más adelante. Lo que se nos ocurrió es no poner el término $1/\sqrt{n_1}$ entre los dos valores e intentar probar directamente que $2^{-2n_1}\binom{2n_1}{n_1} \leq 2^{d-1}n^{-1/2}$, esta desigualdad resultó ser cierta, la prueba usa las siguientes cotas para $n!$;

$$\sqrt{2\pi}\ n^{n+1/2}e^{-n} \leq n! \leq en^{n+1/2}e^{-n}$$

de esta manera tenemos que

$$2^{-2n_1}\binom{2n_1}{n_1} = 2^{-2K}\binom{2K}{K} \leq 2^{-2K}e(2K)^{2K+1/2}e^{-2K}/2\pi K^{2K+1}e^{-2K}$$

$$= e/\pi\sqrt{2K} \leq 1/\sqrt{2K} \leq 1/\sqrt{K+1} \leq 2^{d-1}/\sqrt{4^{d-1}K+r} = 2^{d-1}/\sqrt{n}.$$

Estas observaciones se las envié a Terence Tao, en la versión publicada, ya no aparecen los errores.

Regresando a la demostración, notemos que así como se probó la desigualdad con n_1 puede hacerse con cualquier n_i, es decir probar que $2^{-2n_i}\binom{2n_i}{n_i} \leq 2^{d-i}/\sqrt{n}$. Entonces utilizando el lema y la desigualdad anterior se tiene que

$$\begin{aligned}
\delta^2 &= \mathbb{E}\left[\mu\left(X_{\pi,s} \cap X_{\pi,t}\right)\right] \\
&= \mathbb{E}\left[\mu\left(X_{\pi,s} \cap X_{\pi,t}\right) \mid s = t\right]\mathbb{P}\left(s = t\right) + \mathbb{E}\left[\mu\left(X_{\pi,s} \cap X_{\pi,t}\right) \mid s \neq t\right]\mathbb{P}\left(s \neq t\right) \\
&\leq \mathbb{E}\left[\mu\left(X_{\pi,s} \cap X_{\pi,t}\right) \mid s \neq t\right](1) + (1)\,\mathbb{P}\left(s = t\right).
\end{aligned}$$

de donde

$$\mathbb{E}\left[\mu\left(X_{\pi,s} \cap X_{\pi,t}\right) \mid s \neq t\right] \geq \delta^2 - 2^{d-1}n^{-1/2}.$$

Ahora aplicamos el método probabilístico para concluir que podemos elegir $s < t$ tales que $\mu\left(X_{\pi,s} \cap X_{\pi,t}\right) \geq \delta^2 - 2^{d-1}n^{-1/2}$ y denotamos $A_0^{(1)}$ y $A_1^{(1)}$ a $A_{\pi,s}$

y $A_{\pi,t}$, respectivamente. Notemos que $A_0^{(1)} \subset A_1^{(1)}$ que ambos son disjuntos de V y que $A_0^{(1)} \cup B$, $A_1^{(1)} \cup B \in \mathbf{F}$ $\forall B \in X_{\pi,s} \cap X_{\pi,t}$.

Ahora repitamos el argumento con $n - n_1$ en lugar de n, n_2 en lugar de n_1 y $F_1 = X_{\pi,s} \cap X_{\pi,t}$ en lugar de $\mathbf{F}$. Eso nos da conjuntos $A_0^{(2)}$ y $A_1^{(2)}$ y un conjunto F_2 de subconjuntos de $\{\pi(n_2+1), \ldots, \pi(n)\}$ tales que $A_0^{(2)} \subset A_1^{(2)}$, ambos son disjuntos de $\{\pi(n_2+1), \ldots, \pi(n)\}$ y tales que $A_0^{(2)} \cup B$, $A_1^{(2)} \cup B \in F_1$ $\forall B \in F_2$, además la densidad de F_2 en $\{\pi(n_2+1), \ldots, \pi(n)\}$ es al menos $\left(\delta^2 - 2^{d-1}n^{-1/2}\right)^2 - 2^{d-2}n^{-1/2} \geq \delta^4 - 2^d n^{-1/2}\delta^2 - 2^{d-2}n^{-1/2} \geq \delta^4 - 2^{d-1}n^{-1/2}$, donde en la última desigualdad usamos que $\delta < 1/2$.

Al continuar con este proceso tendremos que F_r tiene densidad al menos $\delta^{2^r} - 2^{d-r+1}n^{-1/2}$ y en el siguiente paso obtendremos que F_{r+1} tiene densidad al menos

$$\left(\delta^{2^r} - 2^{d-r+1}n^{-1/2}\right)^2 - 2^{d-r-1}n^{-1/2} \geq \delta^{2^{r+1}} - 2^{d-r+2}n^{-1/2}\delta^{2^r} - 2^{d-r-1}n^{-1/2}$$

$$\geq \delta^{2^{r+1}} - 2^{d-r}n^{-1/2}.$$

Si repetimos esto hasta que $r = d-2$ y resulta que se cumple que

$$\delta^{2^{d-1}} - 4n^{-1/2} \geq \frac{1}{2\sqrt{2n/3}} \geq \binom{n_d}{\lfloor \frac{n_d}{2} \rfloor}$$

entonces por el teorema de Sperner, $F_{d-1} = E$ contiene dos conjuntos $A_0^{(d)}$ y $A_1^{(d)}$ con $A_0^{(d)} \subset A_1^{(d)}$. Esto nos da el subespacio combinatorio $d-$dimensional buscado, el cual consiste de los 2^d conjuntos de la forma $A_{\epsilon_1}^{(1)} \cup \cdots \cup A_{\epsilon_d}^{(d)}$ donde cada ϵ_i es 1 o 0, aquí hay otro error en el artículo pues en su notación intercambian subíndices con superíndices de tal forma que el conjunto que ellos afirman que funciona ni siquiera está bien definido También podemos definir $C_0 = \bigcup_{i=1}^{d} A_0^{(d)}$ y $C_i = A_1^{(i)} - A_0^{(i)}$ para $i = 1, 2, \ldots, d$ y utilizar la interpretación dada al principio de la demostración de lo que es un subespacio combinatorio $d-$dimensional. La existencia de este conjunto contradice nuestra hipótesis sobre $\mathbf{F}$ por lo tanto se debe cumplir que $\delta^{2^{d-1}} - 4n^{-1/2} < 1/2\sqrt{2n/3}$ lo cual sucede si $\delta^{2^d} < (4n^{-1/2} + 1/2\sqrt{2n/3})^2 < 25/n$ lo que implica que $n \geq 25/\delta^{2^d}$, de donde $\delta \leq (25/n)^{1/2^d}$ como queríamos demostrar. $\square$

Bibliografía

[1] N. Alon y J. Spencer. The probabilistic method. Wiley-Interscience; 3a edición, 2008.

[2] P. Hoffman. The Man Who Loved Only Numbers. Hyperion; 1a edición, 1999.

[3] S. Ross. A First Course in Probability. Prentice Hall; 8a edición, 2009.

[4] G. Hardy y E. Wright. An Introduction to the Theory of Numbers. Oxford University Press; 6a edición, 2008.

[5] M. Heyvaert y F.T. Bruss. La méthode probabiliste. Gazette des mathématiciens; SMF 124, 2010, abril del 2010.

[6] D. H. J. Polymath. A new proof of the density Hales.Jewett theorem. Annals of mathematics; Volumen 175, abril del 2012.

[7] I. Matosevic. The Probabilistic Method and the Bounds of Classical Ramsey Numbers. http://www.eecg.toronto.edu/ imatos/ramsey.pdf, 2005.

[8] P. Loh. The probabilistic method. http://www.math.cmu.edu/ploh/docs/math/mop2010/prob-comb-soln.pdf, junio del 2010.