



UNIVERSIDAD MICHOACANA DE
SAN NICOLAS DE HIDALGO

*Facultad de Ciencias Físico-Matemáticas "Mat.
Luís Manuel Rivera Gutiérrez"*

El Método Probabilístico en Matemáticas Discretas

T E S I S

PARA OBTENER EL TITULO DE

LIC. EN CIENCIAS FÍSICO MATEMÁTICAS

PRESENTA

HÉCTOR MANUEL TÉLLEZ GÓMEZ

ASESOR

DRA. MARÍA LUISA PÉREZ SEGUÍ

MORELIA, MICH., SEPTIEMBRE 2011



El Método Probabilístico en Matemáticas Discretas

Héctor Manuel Téllez Gómez

5 de Septiembre de 2011

Agradecimientos

Antes que a alguien más, quiero agradecer a mi familia: a mis padres Manuel y Arcelia Angélica (a quienes rara vez me refiero con tales nombres) por todo su apoyo; a mis hermanos Luis Alberto y Gloria por siempre estar dispuestos a ayudarme cuando lo necesité. A mis tios paternos Francisco, Gilberto, Guadalupe, Gloria, Ignacio y Pilar por siempre animarme, y mis tios maternos Leonel y Caridad quienes siempre tuvieron confianza en mi. A mis dos abuelitas, Gloria y Jovita, quienes cuando gozaban de buena salud siempre desearon lo mejor para mi. También a todo el conjunto de primos, especialmente Berenice y Héctor (otro). A Sally, por acompañarme todo el tiempo.

A mi familia académica; a mi asesora y madre académica Malú, por su paciencia y enseñanzas no acotadas. A mis hermanos de generación Julián, Luis Antonio, Alfonso y Carlos Alberto. A mis hermanos mayores Atziri, David Becerril, Juan Cristobal, Daniel, Juan Pablo y Abdón. A mis tios académicos Fernando Hernández, David Meza y Edgar Chávez de quienes he aprendido muchas cosas y espero no dejar de hacerlo. A mis hermanos más pequeños Tero, Ana Cristina, Rob y Larissa. A mis amigos Soffer, Khépani, Manuel, Jonathan, Lucas, Saúl, Poke, Mikohol, Pelayo, José Luis, Piris. A mis viejos amigos con quienes no he perdido contacto: Enrique, Julio César, Emmanuel, Isaias.

A todos los compañeros que convivieron conmigo durante la carrera: Yesenia, Lalo, Bombón, Gabino, Nayeli, Vania, Jessica (ambas), Ulises, David (Griego), Gaby, Antonio Tavera, Tania Oyuki, Heather, Chonso, Alma, Zareth, Lua, Checo, y todos cuyos nombres ahora no recuerdo sin que esto signifique que sean menos importantes. Por hacer entretenido el tiempo que estuve en la facultad.

A todos los que me han ayudado con mis trámites de titulación ahora que mi tiempo libre es escaso: Carlos Alberto Mendoza, Alfonso Ortiz, Luis Antonio Ruiz, Antonio Montero, Khépani Raya, Luis Alberto Téllez, Arcelia Angélica Gómez, Héctor Téllez (mi padre), Adán Cortez, María Luisa Pérez, Sergio Albarrán, Salvador Lucas, Manuel Sedano y Jonathan Cancino. A todos ustedes, gracias por “titularme”.

A Liliana Medrano y Fabiola Bautista por ser especiales. A Liliana de nuevo por convencerme de estudiar esta carrera.

A Paul Erdős, Noga Alon y Joel H. Spencer por dejar a mi alcance un tema de tesis tan bonito.

A CONACyT por apoyar este trabajo, proyecto 61351, “Métodos asintóticos en propagación y difracción de ondas”.

Índice general

1. Preliminares	7
1.1. Probabilidad	7
1.2. Teoría de la complejidad computacional	12
1.3. Tasa de crecimiento de funciones	15
2. El Método Básico	21
2.1. El método probabilístico	21
2.2. Aplicaciones a Teoría de Gráficas	21
2.3. Aplicaciones a Combinatoria	30
2.4. Parejas ajenas	34
3. Aplicaciones De La Linealidad De La Esperanza	41
3.1. Caminos hamiltonianos en gráficas dirigidas	41
3.2. Subgráficas bipartitas	42
3.3. Más resultados sobre coloraciones	45
3.4. Geometría analítica	46
3.5. Alto cuello y alto número cromático	48
4. Alteraciones	53
4.1. Más sobre números de Ramsey	53
5. El método del segundo momento	57
5.1. Una versión débil del Teorema de Turán	58
5.2. Funciones umbral	60
Notación	63
Bibliografía	65

Introducción

El siguiente trabajo trata sobre el método probabilístico. El método probabilístico es una herramienta no constructiva que utiliza teoría de la probabilidad para demostrar la existencia de “objetos” matemáticos. Como no existe una forma estándar de aplicar estas demostraciones muchas personas prefieren llamarles métodos probabilísticos (en plural).

Estos métodos se han usado durante varias décadas, aunque, a las primeras demostraciones que se hicieron utilizándolos no se les prestó mucha atención. No fue sino hasta que Paul Erdős empezó a hacer demostraciones con estas técnicas que se les dio importancia. Gracias a estas técnicas por fin se pudo dar solución a problemas que por mucho tiempo habían sido clasificados como meras conjeturas.

Estas técnicas son principalmente aplicadas en problemas de matemáticas discretas. Además, éstas ofrecen bastante comodidad al tratar con teoría de gráficas. Es por esto que la mayoría de los resultados que se muestran en el trabajo son sobre teoría de gráficas y matemáticas discretas en general. En todos los capítulos se muestra al menos un resultado de teoría de gráficas.

De manera breve, se puede decir que estos métodos consisten en lo siguiente: Definir un espacio muestral apropiado, medir la probabilidad de que cierto objeto existe y, si ésta es positiva, entonces el objeto existe.

Cuando el espacio muestral es finito, el problema de calcular una probabilidad se vuelve en un problema de conteo. Muchas veces los problemas de conteo se vuelven difíciles de manejar, en vez de resolverlos completamente se utilizan cotas que ofrecen los teoremas de teoría de la probabilidad. Esto es muy útil cuando se trata con problemas de gráficas donde usualmente las hipótesis son escasas y las posibilidades para las gráficas son muchas y difíciles de manejar.

En el trabajo se discuten y analizan diferentes maneras en que se pueden utilizar métodos probabilísticos y qué resultados de teoría de la probabilidad son útiles. En otras palabras, se discute sobre varios caminos para explotar una idea.

En el primer capítulo se escribió lo básico sobre teoría de la probabilidad. Muchos resultados que se muestran en esa sección se utilizan a lo largo de todo el trabajo. También en el primer capítulo se discute sobre teoría de la complejidad computacional. Esto es porque muchos de los resultados que se muestran en el resto del trabajo son cotas que facilitan el cómputo de ciertos problemas. Por último, en el primer capítulo también se habla sobre tasas de crecimiento de funciones. Esto es porque este tipo de análisis es útil cuando se quiere acotar funciones de probabilidad.

En el segundo capítulo se tratan problemas básicos resueltos con métodos probabilísticos y, usando únicamente mediciones de probabilidad por medio de conteo, cotas que son resultados de cálculo y cotas que son resultados de teoría de la probabilidad.

En el tercer capítulo se hacen demostraciones utilizando la esperanza y resultados sobre ella. Se puede pensar en la esperanza como un promedio. Una de las ideas centrales en este capítulo es una generalización de que si el promedio de un conjunto de números es r entonces alguno de los números tiene valor mayor o igual a r o alguno tiene valor menor o igual a r . La generalización de esto en términos de esperanza es que dada una variable aleatoria X con esperanza e , debe existir un elemento s en el espacio muestral tal que $X(s) \geq e$ ó $X(s) \leq e$. Con un poco de cuidado se puede notar que esto es otra vez la idea base de que un objeto existe si la probabilidad de que exista es positiva. En este capítulo se demuestra un teorema muy importante en teoría de gráficas. El teorema dice que existen gráficas con cuello y número cromático arbitrariamente grandes. Este resultado fue primordial para que se tomaran en cuenta los métodos probabilísticos como una herramienta poderosa para hacer demostraciones.

En el cuarto capítulo, titulado Alteraciones, se utilizan varias técnicas ya antes usadas. En este capítulo todos los resultados son sobre teorías de gráficas. La idea principal consiste en contar o acotar los casos que no nos interesan, modificar la gráfica para que estos casos no puedan ocurrir. Generalmente esto se utiliza para obtener cotas.

En el quinto capítulo se utiliza en particular una desigualdad que involucra a la varianza de una variable aleatoria. Aquí los resultados no son contundentes, sin embargo, se demuestra que ciertos resultados se pueden obtener con probabilidad tan cercana a 1 como uno quiera. También aquí se define lo que es una función umbral, nombradas así porque son el parteaguas entre que algo ocurra “casi siempre” o “casi nunca”.

Capítulo 1

Preliminares

1.1. Probabilidad

Definimos la probabilidad de que un determinado suceso ocurra, como el cociente del número de casos favorables entre el número total de casos. Pondremos aquí estos conceptos con una mayor formalidad para poder extender la teoría. Daremos sólo algunas definiciones básicas para poder emplear el *método probabilístico*, el cual consistirá en usar resultados de probabilidad para resolver problemas combinatorios. Sólo nos interesa el caso finito; trabajaremos dentro de un conjunto (finito) Ω que llamaremos *espacio muestral* y cuyos elementos se llaman *resultados* que consideraremos igualmente probables. A cada subconjunto A de Ω se le llama *suceso* y definimos la *probabilidad de que A ocurra* como

$$P(A) = \frac{|A|}{|\Omega|}.$$

Observemos que P es una función del conjunto de sucesos de Ω en el conjunto de racionales entre 0 y 1, además tenemos $P(\emptyset) = 0$, $P(\Omega) = 1$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$, y si $B = \Omega \setminus A$ (o sea, si B es el complemento de A) entonces $P(B) = 1 - P(A)$.

De aquí en adelante Ω representará a un espacio muestral.

Ejemplo. ¿Cuál es la probabilidad de que al escoger dos subconjuntos de 4 elementos dentro de un conjunto de 10 elementos, los subconjuntos tengan al menos un elemento en común?

Solución. El espacio muestral Ω consta de las parejas de subconjuntos de 4 elementos de $\{1, 2, 3, 4, 5, 6, 7, 8, 9, 10\}$. Queremos calcular $P(A)$ donde A es el subconjunto de elementos (X, Y) de Ω tales que $X \cap Y \neq \emptyset$. Es más fácil contar la probabilidad del complemento B de A . Tenemos que $|B| = \binom{10}{4} \binom{6}{4}$, así que

$$P(A) = 1 - P(B) = 1 - \frac{\binom{10}{4} \binom{6}{4}}{\binom{10}{4}^2} = 1 - \frac{6 \cdot 5 \cdot 4 \cdot 3}{10 \cdot 9 \cdot 8 \cdot 7} = 1 - \frac{1}{14} = \frac{13}{14}.$$

1.1.1 Definición. Dados dos sucesos A y B dentro de un espacio muestral Ω , decimos que A y B son **independientes** si $P(A \cap B) = P(A)P(B)$. La idea es que A y B son independientes si la proporción de elementos de B dentro de A es la misma que la de elementos de B dentro de todo X , es decir,

$$\frac{|B \cap A|}{|A|} = \frac{|B|}{|X|}.$$

En caso contrario decimos que los sucesos son **dependientes**.

Por ejemplo, si Ω es el conjunto de posibles resultados al escoger un número del 1 al 60, A es el suceso de escoger un número múltiplo de 4, B es el suceso de escoger un múltiplo de 5 y C es el suceso de escoger un múltiplo de 6 (podemos pensar $\Omega = [60]$, $A = \{4, 8, \dots, 60\}$, $B = \{5, 10, \dots, 60\}$ y $C = \{6, 12, \dots, 60\}$) entonces A y B son independientes pero A y C son dependientes ($P(A) = \frac{1}{4}$, $P(B) = \frac{1}{5}$, $P(C) = \frac{1}{6}$, $P(A \cap B) = \frac{1}{20}$ y $P(A \cap C) = \frac{1}{12}$).

1.1.2 Definición. Una **variable aleatoria** es una función $f : \Omega \rightarrow \mathbb{R}$. Donde Ω es un espacio muestral.

Una variable aleatoria es una manera de asignar un peso a los elementos del espacio muestral. Haciendo esto se puede destacar qué elementos satisfacen mejor cierta propiedad.

1.1.3 Definición. Sean X y Y variables aleatorias en Ω . Decimos que X y Y son **independientes** si para cada $r, s \in \mathbb{R}$, se tiene que los eventos A, B definidos por $A = \{\omega \in \Omega : X(\omega) = r\}$ y $B = \{\omega \in \Omega : Y(\omega) = s\}$ son independientes.

Decimos que variables aleatorias X_1, X_2, \dots, X_n en un conjunto muestral Ω son independientes si para cualquier colección de reales r_1, r_2, \dots, r_n se tiene que los eventos A_1, A_2, \dots, A_n definidos por $A_i = \{\omega \in \Omega : X_i(\omega) = r_i\}$ son independientes.

1.1.4 Definición. Dada una variable aleatoria $X : \Omega \rightarrow \mathbb{R}$ (donde Ω es un espacio muestral), se define la **esperanza** de X como

$$E[X] = \sum_{\omega \in \Omega} P(\{\omega\})X(\omega).$$

Nótese que si $\omega, \omega' \in \Omega$ son distintos, entonces $P(\{\omega\}) + P(\{\omega'\}) = P(\{\omega, \omega'\})$. Pues $\{\omega\}$ y $\{\omega'\}$ son ajenos. Entonces, si en la definición de esperanza, en vez de sumar sobre los resultados, agrupamos los resultados con imágenes iguales tenemos que

$$E[X] = \sum_{r \in \mathbb{R}} rP(X^{-1}(r)).$$

La esperanza de una variable aleatoria se puede pensar como el promedio de los valores que toma (esto es más fácil de observar en la definición que en la equivalencia que acabamos de dar). Esto es útil para hacer afirmaciones sobre existencia

sin necesidad de una construcción. Por ejemplo, si obtenemos que la esperanza de una variable es menor que cierto $r \in \mathbb{R}$ se puede asegurar que existe algún resultado $\omega \in \Omega$ tal que $X(\omega) \leq r$.

La esperanza tiene la propiedad de ser lineal sobre las variables aleatorias. Esto la vuelve mucho más fácil de calcular.

1.1.5 Teorema. *La esperanza es lineal. Es decir, dadas $X, Y : \Omega \rightarrow \mathbb{R}$ variables aleatorias en Ω y $r \in \mathbb{R}$ se tiene que $E[X + rY] = E[X] + rE[Y]$.*

Demostración.

$$\begin{aligned} E[X + rY] &= \sum_{\omega \in \Omega} (X + rY)(\omega)P(\{\omega\}) = \sum_{\omega \in \Omega} (X(\omega) + rY(\omega))P(\{\omega\}) = \\ &= \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) + r \sum_{\omega \in \Omega} Y(\omega)P(\{\omega\}) = \sum_{\omega \in \Omega} X(\omega)P(\{\omega\}) + r \sum_{\omega \in \Omega} Y(\omega)P(\{\omega\}) = \\ &= E[X] + rE[Y]. \end{aligned}$$

□

A continuación presentamos un ejemplo de la fortaleza del resultado anterior. Contar la cantidad de estructuras cíclicas de las permutaciones de n elementos es un problema que no ha sido resuelto, pero una sencilla aplicación del resultado anterior nos puede decir cuál es la cantidad de ciclos que se espera de una permutación tomada al azar.

1.1.6 Teorema. *La esperanza de la cantidad de ciclos de una permutación de n elementos es $\sum_{i=1}^n \frac{1}{i}$.*

Demostración. Dado un ciclo $\sigma \in S_n$ de longitud i , la probabilidad de que σ se extienda por una permutación $\gamma \in S_n$ es $\frac{(n-i)!}{n!}$, pues basta dejar fijo el ciclo σ y ver las maneras posibles de revolver el resto de los $n - i$ elementos. Entonces la variable aleatoria $X_\sigma : S_n \rightarrow \mathbb{R}$ definida como $X_\sigma(\gamma) = 1$ si γ extiende a σ y $X_\sigma(\gamma) = 0$ en otro caso, tiene por esperanza precisamente $E[X_\sigma] = \frac{(n-i)!}{n!}$. Sea $A_i \subset S_n$ el conjunto de ciclos de longitud i . Definamos la variable aleatoria $X_i = \sum_{\sigma \in A_i} X_\sigma$. Definida así, X_i cuenta cuantos ciclos de longitud i tiene un γ dado. Por linealidad de la esperanza tenemos que

$$\begin{aligned} E[X_i] &= E\left[\sum_{\sigma \in A_i} X_\sigma\right] = \sum_{\sigma \in A_i} E[X_\sigma] = \sum_{\sigma \in A_i} \frac{(n-i)!}{n!} = \\ &= \binom{n}{i} (i-1)! \frac{(n-i)!}{n!} = \frac{n!}{i!(n-i)!} (i-1)! \frac{(n-i)!}{n!} = \frac{1}{i}. \end{aligned}$$

Por último definamos la variable $X = \sum_{i=1}^n X_i$. Definida de esta manera X cuenta la cantidad de ciclos en γ . Usando nuevamente linealidad de la esperanza

tenemos que

$$E[X] = E\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n E[X_i] = \sum_{i=1}^n \frac{1}{i}.$$

□

Después de observar las propiedades de la esperanza sobre la suma de variables aleatorias, es natural preguntarse si la esperanza también tiene propiedades parecidas con el producto. No se puede afirmar mucho sobre la esperanza de producto de variables aleatorias en términos generales, pero si las variables aleatorias son independientes entonces sí podemos hacer afirmaciones parecidas a las que hicimos con la suma.

1.1.7 Teorema. Sean $X, Y : \Omega \rightarrow \mathbb{R}$ variables aleatorias. Si X y Y son independientes, entonces $E[XY] = E[X]E[Y]$.

Demostración. Notemos que podemos escribir

$$E(XY) = \sum_{r \in \mathbb{R}} rP((XY)^{-1}(r)) = \sum_{r \in \mathbb{R}} \sum_{r' \in \mathbb{R}} rr'P(X^{-1}(r) \cap Y^{-1}(r')).$$

Usando que X y Y son independientes, tenemos

$$\begin{aligned} E(XY) &= \sum_{r \in \mathbb{R}} \sum_{r' \in \mathbb{R}} rr'P(X^{-1}(r))P(Y^{-1}(r')) = \\ &= \sum_{r \in \mathbb{R}} rP(X^{-1}(r)) \sum_{r' \in \mathbb{R}} r'P(Y^{-1}(r')) = E(X)E(Y). \end{aligned}$$

□

1.1.8 Lema (Desigualdad de Markov). Sean $X : \Omega \rightarrow \mathbb{R}$ una variable aleatoria y $a \in \mathbb{R}$ tal que $a > 0$. Entonces

$$P(|X| \geq a) \leq \frac{E(|X|)}{a}.$$

Demostración. Sea $\omega \in \Omega$ y sea $I : \Omega \rightarrow \mathbb{R}$ la variable aleatoria que toma el valor 1 si $|X(\omega)| \geq a$ y el valor 0 si no es así. Entonces tenemos que $aI(\omega) \leq |X(\omega)|$, pues si ocurre que $a \leq |X(\omega)|$ tenemos $a \cdot I(\omega) = a \cdot 1 \leq |X|$. Si ocurre que $|X(\omega)| < a$ tenemos $a \cdot I(\omega) = a \cdot 0 \leq |X|$. Entonces tenemos que $E(aI) \leq E(|X|)$. Por linealidad de la esperanza obtenemos $aE(I) \leq E(|X|)$. Pero $E(I) = P(|X| \geq a)$, por lo tanto $aP(|X| \geq a) \leq E(|X|)$. □

1.1.9 Definición. La **varianza** de una variable aleatoria X es:

$$\text{Var}[X] = E[(X - E[X])^2] = E[X^2] - (E[X])^2.$$

La varianza, de alguna manera, describe qué tanto se aleja una variable aleatoria de su esperanza.

1.1.10 Lema (Desigualdad de Chebyshev). *Sea X una variable aleatoria y $a \in \mathbb{R}$ tal que $a > 0$. Entonces*

$$P(|X - E[X]| \geq a) \leq \frac{\text{Var}[X]}{a^2}.$$

Demostración. Primero notemos que

$$P(|X - E[X]| \geq a) = P((X - E[X])^2 \geq a^2).$$

La desigualdad de Markov nos dice que

$$P((X - E[X])^2 \geq a^2) \leq \frac{E[(X - E[X])^2]}{a^2}.$$

Por definición de varianza tenemos

$$\text{Var}[X] = E[(X - E[X])^2],$$

por lo tanto

$$P(|X - E[X]| \geq a) = P((X - E[X])^2 \geq a^2) \leq \frac{E[(X - E[X])^2]}{a^2} = \frac{\text{Var}[X]}{a^2}.$$

□

1.1.11 Definición. Dadas dos variables aleatorias $X, Y : \Omega \rightarrow \mathbb{R}$ su **covarianza** se define como

$$\text{Cov}[X, Y] = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y].$$

La covarianza puede pensarse como una manera de cuantificar qué tan relacionadas están dos variables aleatorias. Observemos que si $X = Y$ la covarianza no es más que la varianza de X . También nótese que si X y Y son variables independientes, entonces su covarianza es cero, pues $\text{Cov}[X, Y] = E[XY] - E[X]E[Y] = E[X]E[Y] - E[X]E[Y] = 0$, gracias a 1.1.7

Esta última observación es importante. Nos permitirá calcular de manera más fácil la varianza de sumas de variables aleatorias independientes al combinarla con el siguiente resultado.

La varianza no cuenta con las cualidades de linealidad de la esperanza. Sin embargo, para variables aleatorias independientes, la varianza abre sumas.

1.1.12 Lema. *Sean $X_1, X_2, \dots, X_n : \Omega \rightarrow \mathbb{R}$ variables aleatorias. Entonces*

$$\text{Var} \left[\sum_{i=1}^n X_i \right] = \sum_{i=1}^n \text{Var}[X_i] + \sum_{i \neq j} \text{Cov}[X_i, X_j].$$

Demostración.

$$\begin{aligned}
\text{Var} \left[\sum_{i=1}^n X_i \right] &= E \left[\sum_{i=1}^n X_i \sum_{j=1}^n X_j \right] - E \left[\sum_{i=1}^n X_i \right] E \left[\sum_{j=1}^n X_j \right] = \\
&= \sum_{i=1}^n E[X_i^2] + \sum_{i \neq j} E[X_i X_j] - \sum_{i=1}^n (E[X_i])^2 - \sum_{i \neq j} E[X_i] E[X_j] = \\
&= \left(\sum_{i=1}^n E[X_i^2] - \sum_{i=1}^n (E[X_i])^2 \right) + \left(\sum_{i \neq j} E[X_i X_j] - \sum_{i \neq j} E[X_i] E[X_j] \right) = \\
&= \sum_{i=1}^n \text{Var}[X_i] + \sum_{i \neq j} \text{Cov}[X_i, X_j].
\end{aligned}$$

□

1.1.13 Corolario. Sean X, Y variables aleatorias independientes. Entonces $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y]$.

Demostración. El lema anterior nos dice que $\text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] + \text{Cov}[X, Y] + \text{Cov}[Y, X]$. Pero los últimos dos términos son cero por ser X, Y independientes. □

1.2. Teoría de la complejidad computacional

Para poder justificar mejor las cotas que se obtienen como resultado en esta tesis, será necesario hablar un poco sobre teoría de la complejidad computacional, sin embargo sólo se darán las nociones necesarias para familiarizarse sin formalizar.

Pensemos en que una **Máquina de Turing (determinista)** es una cabeza que recuerda una cantidad finita de estados (entre los cuales se encuentran los que llamamos estados finales y un estado inicial) y que se recorre a lo largo de una cinta infinita. La cinta bien puede ser infinita en ambas direcciones o sólo en una; para evitar complicaciones pensemos que siempre son infinitas en ambas direcciones. La cinta infinita está segmentada y en cada segmento puede guardar un símbolo o carácter; además existe un único símbolo conocido como espacio en blanco y es el único que se puede repetir infinitamente.

El conjunto de símbolos es finito y se le conoce como alfabeto. Cada vez que la cabeza se encuentra en un segmento, lo primero que hace es leer el carácter que se encuentra en dicho segmento; después de hacerlo tiene cuatro posibles acciones: reemplazar el carácter que encontró por otro, moverse al segmento de la derecha, moverse al segmento de la izquierda, o detenerse. Estas acciones son determinadas por el estado en que se encuentra y el carácter que encuentra. Al finalizar una acción la cabeza cambia de estado. La cabeza se detiene si llega a un estado final. Con

sólo esto una máquina de Turing sería capaz de hacer cualquier cosa que haga una computadora moderna.

Formalmente, una máquina de Turing puede definirse como una quinteta $(Q, \Sigma, s, F, \delta)$, donde Q es un conjunto finito de estados, Σ es un conjunto finito de símbolos al que pertenece el símbolo que representa el espacio en blanco, $s \in Q$ es el estado inicial, $F \subset Q$ es el conjunto de estados finales y $\delta : Q \times \Sigma \rightarrow Q \times (\Sigma \cup L, R)$ es una función, donde $L, R \notin \Sigma$. L significa moverse a la izquierda y R a la derecha (esta definición no es la que se da regularmente, pero para nuestros fines es suficiente).

Para asentar ideas veamos un ejemplo de una máquina de Turing que suma. Recibirá dos naturales en “unario” (es decir el 1 se escribe 1, el 2 se escribe 11, el 3 se escribe 111, etc.) separados por un # (así denotaremos el espacio en blanco). Se supone que antes del primer número la cinta está llena de puros # y después del segundo número también. Es decir, la cinta se ve de la forma $\cdots \#\#\#11 \cdots 1\#11 \cdots 1\#\#\# \cdots$. La posición inicial de la cabeza será el # que se encuentra a la izquierda del primer número. Los estados y acciones determinadas por éstos son los siguientes:

(en donde escribimos - por cualquier símbolo del alfabeto)

Estado actual	Símbolo en la cinta	Acción sobre la cinta	Estado al que pasa
Q_1	-	Moverse a la derecha	Q_2
Q_2	1	Moverse a la derecha	Q_2
Q_2	#	Moverse a la derecha	Q_3
Q_3	1	Escribir #	Q_4
Q_3	#		Q_f
Q_4	-	Moverse a la izquierda	Q_5
Q_5	-	Escribir 1	Q_2
Q_f	-	Detenerse	-

La idea aquí es ir recorriendo hacia la derecha el # que separa a los números para al final, tener un solo número que resulta de sumar los números dados. Veamos el ejemplo de cómo resolvería la suma $2 + 3$. El carácter subrayado indica la posición de la cabeza. La cinta al inicio se ve de la forma $\cdots \# \# \# \underline{11} \# 111 \# \# \cdots$. En la tabla siguiente se modela el comportamiento de la máquina de turing definida:

Cinta	Estado actual	Carácter leído	Acción sobre la cinta	Estado al que se pasa
$\dots \# \# 11 \# 111 \# \# \dots$	Q_1	$\#$	Moverse a la derecha	Q_2
$\dots \# \# \underline{1} 1 \# 111 \# \# \dots$	Q_2	1	Moverse a la derecha	Q_2
$\dots \# \# 1 \underline{1} \# 111 \# \# \dots$	Q_2	1	Moverse a la derecha	Q_2
$\dots \# \# 11 \# \underline{1} 11 \# \# \dots$	Q_2	$\#$	Moverse a la derecha	Q_3
$\dots \# \# 11 \# \underline{1} 1 \# \# \dots$	Q_3	1	Escribir un $\#$	Q_4
$\dots \# \# 11 \# \# \underline{1} 1 \# \# \dots$	Q_4	$\#$	Moverse a la izquierda	Q_5
$\dots \# \# 11 \# \# \underline{1} 1 \# \# \dots$	Q_5	$\#$	Escribir un 1	Q_2
$\dots \# \# 1 \underline{1} 1 \# 11 \# \# \dots$	Q_2	1	Moverse a la derecha	Q_2
$\dots \# \# 11 \underline{1} \# 11 \# \# \dots$	Q_2	$\#$	Moverse a la derecha	Q_3
$\dots \# \# 111 \# \underline{1} 1 \# \# \dots$	Q_3	1	Escribir un un $\#$ y	Q_4
$\dots \# \# 111 \# \# \underline{1} \# \# \dots$	Q_4	$\#$	Moverse a la izquierda	Q_5
$\dots \# \# 111 \# \# \underline{1} \# \# \dots$	Q_5	$\#$	Escribir un 1	Q_2
$\dots \# \# 111 \underline{1} \# 1 \# \# \dots$	Q_2	1	Moverse a la derecha	Q_2
$\dots \# \# 1111 \# \underline{1} \# \# \dots$	Q_2	$\#$	Moverse a la derecha	Q_3
$\dots \# \# 1111 \# \# \underline{1} \# \# \dots$	Q_3	1	Escribir un un $\#$	Q_4
$\dots \# \# 1111 \# \# \# \underline{1} \# \# \dots$	Q_4	$\#$	Moverse a la izquierda	Q_5
$\dots \# \# 1111 \# \# \# \underline{1} \# \# \dots$	Q_5	$\#$	Escribir un 1	Q_2
$\dots \# \# 1111 \underline{1} \# \# \# \dots$	Q_2	1	Moverse a la derecha	Q_2
$\dots \# \# 11111 \# \# \# \dots$	Q_2	$\#$	Moverse a la derecha	Q_3
$\dots \# \# 11111 \# \# \# \dots$	Q_3	$\#$		Q_f

En teoría de la complejidad los problemas se clasifican de acuerdo a la “cantidad” de pasos que necesita una máquina de Turing para resolverlos. Para ser más precisos, se clasifican de acuerdo a qué tanto crece la cantidad de pasos necesarios para resolverlo con respecto al tamaño de la entrada. Por ejemplo, si pidiéramos a una máquina de Turing encontrar todos los subconjuntos de un conjunto con n elementos, ésta al menos tardaría 2^n pasos para poder escribirlos (tardaría más, pero sólo nos interesa el radio de crecimiento); la complejidad de este problema es exponencial con respecto al tamaño de la entrada. Si pidiéramos a una máquina de Turing que tan sólo llegara hasta el final de una cadena de longitud n que le damos como entrada, ésta tardará solamente $n + 1$ pasos: n para recorrerla y uno para entrar al estado final. La complejidad de este problema es de orden lineal.

Existe otro tipo de máquinas de Turing a las que se les llama **no deterministas**. Para no entrar en detalles, sólo digamos que estas máquinas de Turing pueden crear copias de ellas mismas. En éstas se considera un solo paso a todos los pasos simultáneos que realizan las copias y la acción de copiarse no se considera un paso. Para una máquina de Turing no determinista el problema de verificar si es posible obtener un número entero k como la suma de elementos de un subconjunto de un conjunto A de n números enteros (no necesariamente distintos) es de orden lineal. Basta con que la máquina genere todos los subconjuntos de A y verifique si la suma de los elementos de cada uno es k . Para generar los subconjuntos lo único que tendría que hacer es enumerar a los elementos y luego hacer dos copias de sí misma: una que escribe al primer elemento y otra que no; para cada una de estas hacer dos copias de sí misma, una que escribe al segundo elemento y otra que no; y así sucesivamente

hasta tomar en cuenta al n -ésimo elemento. Generar a los subconjuntos de esta manera apenas utiliza un tiempo lineal, sumar los elementos de un conjunto también es realizable en tiempo lineal, y verificar si el resultado de una suma es igual a k también es lineal. Por lo tanto el tiempo necesario para resolver este problema es de orden lineal.

Decimos que un problema pertenece a la **clase de complejidad P** si una máquina de Turing determinista puede resolverlo en una cantidad polinomial de pasos. Decimos que un problema pertenece a la clase de **complejidad NP** si una máquina de Turing no determinista puede resolverlo en una cantidad polinomial de pasos. En general se tiene $P \subset NP$ pues una máquina de Turing no determinista puede resolver lo mismo que una determinista sin necesidad de crear sus “copias”. Hasta la fecha no se ha podido demostrar si P es subconjunto propio de NP o no, es decir, no se ha podido demostrar si es cierto o falso que $P = NP$ e incluso existe una grata recompensa para quien lo consiga.

En ocasiones es posible traducir un problema en otro, es decir tomar la entrada de una máquina de Turing que resuelve un problema y convertirla en una entrada para una máquina de Turing que resuelve otro problema de tal manera que resolver el problema en la segunda implica haberlo resuelto en la primera. Para fines de estudiar la complejidad sólo se permiten “traducciones” que se puedan realizar en un tiempo polinomial por una máquina de Turing determinista. El hecho de poder traducir un problema en otro significa que el segundo es más complejo que el primero y entonces decimos que el segundo reduce al primero.

Existe otra clase de complejidad que será muy de nuestro interés; se le llama “NP-completo”. Se define como sigue:

Un problema es **NP-completo** si es **NP** y reduce a cualquier otro problema **NP**. Es decir, la clase NP-completo es la clase de los problemas NP más difíciles. Visto así, para demostrar que $P = NP$ bastaría encontrar un problema P que pueda reducir a un problema NP-completo.

1.3. Tasa de crecimiento de funciones

Introducimos aquí notación utilizada en análisis de algoritmos para el estudio de crecimiento de funciones. El **costo** de un algoritmo es el número de operaciones necesarias para resolver un problema dado.

Para tener un estudio detallado hay que tomar en cuenta el lenguaje de programación en que se escribe y qué máquina se encargará de ejecutarlo. Esto se vuelve muy poco práctico si queremos hablar en términos generales sobre los costos.

Nos gustaría poder decir algo acerca del costo de un algoritmo de una manera más simple, sin involucrarnos en los lenguajes de programación o el tipo de computadora que se encargará de ejecutar los algoritmos. De esta manera consideramos suficiente conocer el mínimo número de operaciones necesarias para que un algoritmo dado pueda resolver un problema.

Supongamos que dado un conjunto A de n números enteros queremos conocer la suma de todos sus elementos. Diseñar un algoritmo para resolver este problema es simple, no hay más que sumar todos los elementos. Este algoritmo, sin importar la máquina que lo ejecute o el lenguaje en que fue escrito, al menos necesita realizar $n - 1$ operaciones (las $n - 1$ sumas). De esta manera, sin saber hasta los últimos detalles del algoritmo, podemos hablar sobre su costo.

La función de costo de un algoritmo es una función f que asigna a cada natural que representa el tamaño de la entrada, el número de operaciones necesarias para llegar a un resultado. Por ejemplo, para el algoritmo anterior su función de costo es $f(n) = n - 1$.

Supongamos que las funciones de costo de dos algoritmos son $f(n) = n^2$ y $g(n) = 3n$. Pongamos los costos para los primeros valores de n :

n	1	2	3	4	5	6	7	8	9	10	11	12
$f(n) = n^2$	1	4	9	16	25	36	49	64	81	100	121	144
$g(n) = 3n$	3	6	9	12	15	18	21	24	27	30	33	36

Observemos que, para $n \leq 3$, $g(n) \geq f(n)$, pero para $n > 3$ tenemos que $g(n) < f(n)$. Si queremos hablar de cuál de los costos es menor, tendríamos que comparar los costos para cada n . Sin embargo, en este ejemplo se puede notar fácilmente que f crece mucho más rápido que g .

Para hacer este tipo de análisis introducimos la siguiente notación.

1.3.1 Definición. Dadas dos funciones $f, g : \mathbb{N} \rightarrow \mathbb{N}$ decimos que $g \in O(f)$ si existen $r \in \mathbb{R}^+$ y $n_0 \in \mathbb{N}$ tales que para toda $n > n_0$, $g(n) \leq rf(n)$.

Observemos que cuando $g \in O(f)$ tenemos que “ g no crece más rápido que f ”.

1.3.2 Teorema. Dadas dos funciones $f, g : \mathbb{N} \rightarrow \mathbb{N}$, si $\lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = k$, con $k \in \mathbb{R}$. Entonces $g \in O(f)$.

Demostración. $\lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = k$ quiere decir que para todo $\epsilon > 0$ existe n_0 tal que si $n > n_0$ entonces $|\frac{g(n)}{f(n)} - k| < \epsilon$. Ya que esto es cierto para todo ϵ , en particular lo es

para $\epsilon = 1$ y tenemos

$$\forall n > n_0 \quad k - 1 < \frac{g(n)}{f(n)} < k + 1,$$

lo cual implica que $g(n) < (k + 1)f(n)$, $\forall n > n_0$. \square

A pesar de que parecería intuitivo que las condiciones del teorema anterior son necesarias y suficientes para que dos funciones f, g cumplieran que $g \in O(f)$, no es así. Un ejemplo claro donde la necesidad falla son las funciones $f, g : \mathbb{N} \rightarrow \mathbb{N}$, $f(n) = 1$, $g(n) = \frac{2}{2^{(-1)^n}}$ pues tenemos que

$$\forall n \in \mathbb{N} \quad g(n) \leq 4 = 4f(n),$$

así que $g \in O(f)$, pero $\lim_{n \rightarrow \infty} \frac{g(n)}{f(n)}$ no existe.

1.3.3 Teorema. Si $g(n)$ es un polinomio de grado m , entonces $g \in O(n^m)$.

Demostración. Escribamos a g de la siguiente manera $g(n) = a_m n^m + a_{m-1} n^{m-1} + \dots + a_0$. Entonces

$$\frac{g(n)}{n^m} = \frac{a_m n^m + a_{m-1} n^{m-1} + \dots + a_0}{n^m} = a_m + \frac{a_{m-1}}{n} + \dots + \frac{a_0}{n^m}.$$

Haciendo el límite tenemos

$$\lim_{n \rightarrow \infty} a_m + \frac{a_{m-1}}{n} + \dots + \frac{a_0}{n^m} = a_m.$$

Usando el teorema anterior, tenemos que $g(n) \in O(n^m)$ \square

Esto último nos dice que todos los polinomios del mismo grado crecen igual de rápido.

Si tenemos dos funciones f, g tales que $f \in O(g)$ y $g \in O(f)$, podemos decir que f y g esencialmente crecen igual de rápido. Cuando esto ocurre lo denotamos como $f \in \Theta(g)$.

Otro teorema fácil de demostrar pero importante para simplificar cálculos es el siguiente.

1.3.4 Teorema. Sean $f \in O(r)$, $g \in O(s)$ y $c \in \mathbb{R}^+$. Entonces

(i) $fg \in O(rs)$.

(ii) $f \in O(cf)$. En particular si $c_1, c_2 \in \mathbb{R}^+$ entonces $c_1 f \in \Theta(c_2 f)$

(iii) $f + g \in O(r + s)$,

(iv) Si $s \in O(r)$ entonces $f + g \in O(r)$.

Demostración. (i) Por definición existen naturales N_1, N_2 y reales positivos k_1, k_2 tales que

$$\begin{aligned} \forall n > N_1 & \quad f(n) \leq k_1 r(n), \\ \forall n > N_2 & \quad g(n) \leq k_2 s(n). \end{aligned}$$

Por lo tanto

$$f(n)g(n) \leq k_1 k_2 r(n)s(n) \quad \forall n > \max(N_1, N_2),$$

así que basta elegir a $k = k_1 k_2$ y a $N = \max(N_1, N_2)$ para encontrar el real y el natural que cumplen con la definición de que $fg \in O(rs)$.

(ii) Tenemos que $f \in O(f)$, pues

$$\forall n \in \mathbb{N} \quad f(n) \leq f(n).$$

Entonces basta elegir $k = \frac{1}{c}$ pues

$$\forall n \in \mathbb{N} \quad f(n) = \frac{1}{c} c f(n).$$

(iii) Por definición existen naturales N_1, N_2 y reales positivos k_1, k_2 tales que

$$\begin{aligned} \forall n > N_1 & \quad f(n) \leq k_1 r(n), \\ \forall n > N_2 & \quad g(n) \leq k_2 s(n). \end{aligned}$$

Por lo tanto

$$\forall n > \max(N_1, N_2) \quad f(n) + g(n) \leq k_1 r(n) + k_2 s(n).$$

Entonces basta elegir $k = \max(k_1, k_2)$ y $N = \max(N_1, N_2)$ para que

$$\forall n > N \quad f(n) + g(n) \leq k_1 r(n) + k_2 s(n) \leq k(r(n) + s(n)).$$

(iv) Otra vez, por definición existen naturales N_1, N_2, N_3 y reales positivos k_1, k_2, k_3 tales que

$$\begin{aligned} \forall n > N_1 & \quad f(n) \leq k_1 r(n), \\ \forall n > N_2 & \quad g(n) \leq k_2 s(n), \\ \forall n > N_3 & \quad s(n) \leq k_3 r(n). \end{aligned}$$

Por lo tanto

$$\forall n > \max\{N_1, N_2, N_3\} \quad f(n) + g(n) \leq k_1 r(n) + k_2 s(n) \leq k_1 r(n) + k_2(k_3 r(n)).$$

Entonces basta elegir $k = k_1 + k_2 k_3$ y $N = \max(N_1, N_2, N_3)$ para que

$$\forall n > N \quad f(n) + g(n) \leq k(r(n)).$$

□

1.3.5 Teorema. Sean $f, g, g', h : \mathbb{N} \rightarrow \mathbb{R}^+$ funciones tales que $\frac{f}{g} \in O(h)$, $g \in O(g')$. Entonces $\frac{f}{g'} \in O(h)$.

Demostración. Sean $N \in \mathbb{N}$ y $k \in \mathbb{R}^+$ tales que $g(n) \leq kg'(n)$ para toda $n > N$. Multiplicando por f en ambos lados, tenemos que

$$\begin{aligned} \forall n > N \quad f(n)g(n) &\leq kf(n)g'(n), \text{ por lo tanto} \\ \forall n > N \quad \frac{f(n)}{g'(n)} &\leq k \frac{f(n)}{g(n)}. \end{aligned}$$

Ahora, sean $N' \in \mathbb{N}$ y $k' \in \mathbb{R}^+$ tales que $\frac{f(n)}{g(n)} \leq k'h(n)$ para toda $n > N'$. Multiplicando por k en ambos lados tenemos que

$$\forall n > N' \quad \frac{f(n)}{g'(n)} \leq k \frac{f(n)}{g(n)} \leq kk'h(n),$$

y con esto termina la demostración. \square

1.3.6 Definición. Dadas dos funciones $f, g : \mathbb{N} \rightarrow \mathbb{N}$, decimos que $\mathbf{g} \in \mathbf{o}(f)$ si $\lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = 0$.

Observemos que $\mathbf{g} \in \mathbf{o}(f)$ nos dice que “ f crece más rápido que g ”.

En vista de 1.3.2 tenemos que $g \in o(f) \Rightarrow g \in O(f)$. El recíproco es claramente falso, pues $f \notin o(f)$.

Nótese que si $g \in o(f)$ y $h \in o(f)$, entonces $g + h \in o(f)$ y también que si $g \in O(f)$ y $h \in O(f)$ entonces $g + h \in O(f)$. De esta manera el uso de esta notación nos permite considerar únicamente el término que crezca más rápido de una función e ignorar al resto.

1.3.7 Teorema. Si $g(n)$ es un polinomio de grado m y $m < l$, entonces $g(n) \in o(n^l)$.

Demostración. Escribamos $g(n) = a_m n^m + a_{m-1} n^{m-1} + \dots + a_0$.

Entonces

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{g(n)}{n^l} &= \lim_{n \rightarrow \infty} \frac{a_m n^m + a_{m-1} n^{m-1} + \dots + a_0}{n^l} = \\ &= \lim_{n \rightarrow \infty} \frac{a_m}{n^{l-m}} + \frac{a_{m-1}}{n^{l-m+1}} + \dots + \frac{a_0}{n^l} = 0. \end{aligned}$$

\square

En particular si g y f son polinomios de grados m y l , respectivamente, y $m < l$, Entonces $g \in o(f)$.

Las definiciones y resultados anteriores se extienden fácilmente a funciones con dominio y codominio los reales positivos. Esto hace mucho más sencillo el estudio

de las tasas de crecimiento puesto que ahora podemos hacer uso de las herramientas que el cálculo ofrece. Después de todo, si $g \in O(f)$ o $g \in o(f)$ en las versiones continuas de f y g , también ocurrirá en sus versiones discretas.

Recordemos que $\log_a(b) = \frac{\ln(b)}{\ln(a)}$ (esto debido a que $a^{\log_a(b)} = b$ y, por lo tanto, $\log_a(b)\ln(a) = \ln(a^{\log_a(b)}) = \ln(b)$). De aquí se deduce inmediatamente que $\log_a(n) \in O(\log_b(n))$ y que $\log_a(n) \notin o(\log_b(n))$ para cualesquiera $a, b > 0$, $a \neq 1 \neq b$. En otras palabras, todos los logaritmos crecen igual de rápido.

1.3.8 Lema. $\log_a(n) \in o(n^r)$ para toda $r \in \mathbb{R}^+$.

Demostración. Trabajaremos con las versiones continuas de $\log_a(x)$ y x^r . Para empezar, $\log_a(x) = \frac{\ln(x)}{\ln(a)}$. Ahora, tanto $\ln(x)$ como x^r tienden a infinito conforme x crece y, por lo tanto, aplicamos L'Hôpital en el siguiente límite:

$$\lim_{x \rightarrow \infty} \frac{\ln(x)}{\ln(a)x^r} = \frac{1}{\ln(a)r} \lim_{x \rightarrow \infty} \frac{1}{x^r} = 0.$$

□

De la misma manera que antes se puede demostrar que para $q, r \in \mathbb{R}^+$, si $q < r$ entonces $x^q \in o(x^r)$. Esto, junto con el resultado anterior, nos dice que los logaritmos crecen muy lento.

Capítulo 2

El Método Básico

2.1. El método probabilístico

El método probabilístico es una herramienta poderosa para atacar problemas en matemáticas discretas particularmente, aunque, se ha utilizado en distintas áreas. El método se explica mejor a través de ejemplos, pero a grandes rasgos funciona de la siguiente manera: Para intentar probar la existencia de una estructura con propiedades específicas, se define un espacio muestral apropiado y luego se muestra que dicha estructura yace en el espacio con probabilidad positiva.

2.2. Aplicaciones a Teoría de Gráficas

Muchos de los teoremas que se demuestran usando el método probabilístico pertenecen a la teoría de las gráficas. Siempre que se mencione gráfica nos estaremos refiriendo a una gráfica simple.

2.2.1 Definición. Una **gráfica simple** G es un par ordenado (V, E) donde V es un conjunto finito a cuyos elementos se les llama *vértices*, y E es una familia de subconjuntos de V con exactamente dos elementos a los que se les llama *aristas*. Decimos que $x, y \in V$ son **vecinos** o **adyacentes** si $\{x, y\} \in E$. Esto también suele denotarse simplemente como $xy \in E$. Una **subgráfica** G' de G es un par ordenado (V', E') tal que $V' \subset V$, $E' \subset E$ y $e \subset V'$ para toda $e \in E'$. En ocasiones, si G' es subgráfica de G , lo denotamos por $G' \subset G$.

2.2.2 Definición. Una **gráfica completa de n vértices** es una gráfica $G = (V, E)$ donde $|V| = n$ y $E = \{\{x, y\} \mid x \neq y \in V\}$. También se denota como \mathbf{K}_n .

Una coloración no es más que una manera de clasificar los objetos de un conjunto. De manera formal lo definimos como sigue:

2.2.3 Definición. Una **coloración** sobre un conjunto A es una función $f : A \rightarrow C$, donde C es un conjunto a cuyos elementos se les llama colores. Si C es un conjunto finito con n elementos decimos que es una **n -coloración**. En ocasiones usaremos los nombres de colores del español para denotar elementos de C .

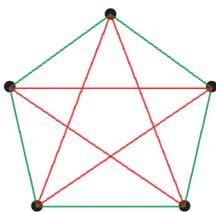
2.2.4 Definición. El número de Ramsey $\mathbf{R}(\mathbf{k}, \mathbf{l})$ es el mínimo número n con que se puede asegurar que dada una 2-coloración (rojo y verde) en las aristas de K_n , se puede encontrar siempre un K_k de aristas rojas o un K_l de aristas verdes como subgráfica.

Como ejemplo, veamos que $R(3, 3) = 6$. Para demostrarlo necesitamos que en K_6 siempre podamos encontrar un triángulo rojo o un triángulo verde y que hay coloraciones de K_5 donde no hay triángulos monocromáticos.

Ejemplo. $R(3, 3) = 6$.

Demostración. Sean $G = (V, E) = K_6$, $v \in V$ y asignemos una 2-coloración a G . En vista de que v pertenece a 5 aristas, al menos 3 de ellas deben ser del mismo color, sin pérdida de generalidad digamos que son rojas. Consideremos 3 de los vecinos de v con los que forma aristas rojas. Si entre cualesquiera dos de ellos existe una arista roja ya formarán un triángulo completamente rojo. Si no es así, las 3 aristas entre ellos tienen que ser verdes y entonces ellos forman un triángulo verde.

Para ver que $R(3, 3) > 5$ basta observar que en la siguiente coloración no hay triángulos monocromáticos.



□

La primera vez que Frank P. Ramsey mencionó los números que ahora llevan su nombre, no fueron tomados muy en cuenta. Fue hasta 1935 (5 años después de la muerte de Ramsey) que Erdős los utilizó junto con G. Szekeres para darle solución a un problema de geometría y combinatoria (conocido como el problema del final feliz). Después de esto, los matemáticos empezaron a dar importancia a la teoría de Ramsey. Hoy en día, las aplicaciones de la teoría de Ramsey, se extiende a, la teoría de números, geometría computacional, análisis armónico, espacios métricos, teoría ergódica, teoría de la información, etc. También usando el teorema de Ramsey finito, el cual dice que existe el número de Ramsey, se dio la primera afirmación indecidible en teoría de conjuntos finitos.

Calcular los números de Ramsey es una tarea particularmente difícil. Se conocen con precisión únicamente 9 de ellos y el más grande es $R(3, 9) = 36$. Se tienen cotas

relativamente pequeñas para cerca de otros 110 donde $n \leq 10$ y $m \leq 26$.

Se sabe que para $k \geq 2$ se tiene que $R(k, k) \leq 4^{k-1}$. A continuación demostraremos un teorema que ofrece una cota inferior para $R(k, k)$. Este fue el primer teorema que Erdős demostró usando el método probabilístico.

2.2.5 Teorema. Si $\binom{n}{k} 2^{1-\binom{k}{2}} < 1$, entonces $R(k, k) > n$. En particular, $R(k, k) > \lfloor 2^{\frac{k}{2}} \rfloor$ para todo $k \geq 3$.

Demostración. Consideremos una 2-coloración aleatoria obtenida coloreando cada arista de manera independiente, ya sea de rojo o verde, y cada color es igualmente probable. Sea S cualquier conjunto de k vértices y sea A_S el suceso en que la subgráfica completa K_k inducida por S es monocromática (es decir totalmente roja o totalmente verde). Entonces tenemos que $P(A_S) = \frac{2}{2^{\binom{k}{2}}} = 2^{1-\binom{k}{2}}$ pues sólo tenemos 2 casos favorables, que todas las aristas sean rojas, o que todas las aristas sean verdes, y los casos posibles son $2^{\binom{k}{2}}$ ya que cada una de las $\binom{k}{2}$ aristas puede ser roja o verde, y tenemos $\binom{n}{k}$ elecciones posibles para S y, por lo tanto, $P(\bigcup_{S \subset V} A_S) \leq \sum_{S \subset V} P(A_S) = \binom{n}{k} 2^{1-\binom{k}{2}} < 1$ (la desigualdad es por hipótesis); de donde existe al menos un caso en el que ningún S induce una subgráfica K_k monocromática, y así $R(k, k) > n$. Ahora bien, si $k \geq 3$ y tomamos $n = \lfloor 2^{\frac{k}{2}} \rfloor$ entonces

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

y, entonces, $R(k, k) > \lfloor 2^{\frac{k}{2}} \rfloor$. □

Este resultado puede parecer muy burdo, pero su radio de crecimiento es bastante alto (exponencial), es decir, puede no ser una muy buena cota inferior, pero al menos si intentáramos diseñar un algoritmo para calcular $R(k, k)$, este resultado nos ahorraría bastantes cálculos innecesarios. Es decir, para calcular $R(10, 10)$ podemos evitar revisar las gráficas con menos de $\lfloor 2^{10/2} \rfloor = 32$ vértices.

2.2.6 Definición. Una **gráfica dirigida** sin aristas simétricas o digráfica sin aristas simétricas es un par ordenado $G = (V, E)$ donde V es el conjunto finito de vértices y E es un conjunto de parejas ordenadas sobre V tal que si $(x, y) \in E$ entonces $(y, x) \notin E$.

2.2.7 Definición. Un **torneo de n jugadores** T es una gráfica completa K_n dirigida y sin aristas simétricas. A cada vértice se le llama jugador y se puede pensar que $(x, y) \in E$ significa "x derrota a y".

2.2.8 Definición. Decimos que un torneo T tiene la propiedad S_k si para cualquier conjunto de k jugadores existe un jugador que los derrota a todos ellos.

Por ejemplo, el torneo $T = (V, E)$, donde $V = \{1, 2, 3\}$ y $E = \{(1, 2), (2, 3), (3, 1)\}$, tiene la propiedad S_1 .

2.2.9 Teorema. Si $\binom{n}{k}(1 - 2^{-k})^{n-k} < 1$ entonces existe un torneo de n vértices que tiene la propiedad S_k .

Demostración. Considérese un torneo de n elementos $T = (V, E)$ arbitrario, y para cada $K \subset V$ tal que $|K| = k$, sea A_K el suceso donde ningún vértice derrota a todos los miembros de K . Tenemos que la probabilidad de que un elemento fuera de K derrote a todos los miembros de K es $\frac{1}{2^k}$ y, por lo tanto, la probabilidad de que no los derrote a todos es $1 - \frac{1}{2^k}$; además el que un elemento dado no derrote a todo K es independiente de que algún otro no lo derrote, de donde $P(A_K) = (1 - \frac{1}{2^k})^{n-k}$. Ahora bien, a K lo pudimos haber escogido de $\binom{n}{k}$ maneras; entonces, la probabilidad de que todo subconjunto de k elementos no sea derrotado por un elemento es

$$P\left(\bigcup_{\substack{K \subset V \\ |K|=k}} A_K\right) \leq \sum_{\substack{K \subset V \\ |K|=k}} P(A_K) = \binom{n}{k}(1 - 2^{-k})^{n-k} < 1.$$

Así, existe al menos un caso en el que todo $K \subset V$ tal que $|K| = k$ es derrotado por algún elemento. \square

Después de diseñar un programa para calcular los valores de $\binom{n}{k}(1 - 2^{-k})^{n-k}$ para gráficas de hasta 5000 vértices, este último teorema lo único que pudo asegurar es que existe un torneo de 5000 vértices con la propiedad S_6 .

2.2.10 Definición. En una gráfica G , se dice que una m -ada de vértices distintos (x_1, \dots, x_m) es un **ciclo** si para $1 \leq i < n$ se tiene que $\{x_i, x_{i+1}\} \in E$ y $\{x_n, x_1\} \in E$. A m se le llama **longitud** del ciclo.

2.2.11 Definición. En una gráfica $G = (V, E)$, si $v \in V$, el **grado** de v es la cardinalidad del conjunto de sus vecinos.

2.2.12 Definición. El **grado mínimo** de una gráfica es el mínimo de los grados de sus vértices.

2.2.13 Definición. Un **conjunto dominante** de una gráfica $G = (V, E)$ es un subconjunto $U \subseteq V$ tal que cada vértice $v \in V \setminus U$ tiene al menos un vecino en U .

Nótese que la importancia de un conjunto dominante radica en que para llegar a cualquier vértice de G partiendo de un conjunto dominante a lo más se necesita usar una arista. En teoría de juegos de estados finitos, este tipo de conjuntos sería muy útil pues, conociendo un conjunto dominante, a lo más se necesitaría un paso para conocer cualquier otro estado del juego. También nótese que entre más grande es un subconjunto de V es más fácil que sea dominante, y menos útil. El siguiente resultado nos garantiza la existencia de conjuntos dominantes “chicos” en caso de que el grado mínimo de la gráfica sea mayor que uno. Antes necesitaremos un pequeño resultado.

2.2.14 Lema. $1 - p \leq e^{-p} \quad \forall p \geq 0.$

Demostración. Veamos que la función $f : [0, 1] \rightarrow \mathbb{R}$ definida por $f(p) = e^{-p} - (1 - p)$ es creciente, lo cual, en vista de que $f(0) = 0$, terminaría la demostración. Para ver que es creciente veamos que su derivada es siempre no negativa

$$\frac{df(p)}{dp} = -e^{-p} + 1 = 1 - \frac{1}{e^p} \geq 0,$$

pues $e^p \geq 1$ para toda $p \geq 0$. □

2.2.15 Teorema. *Sea $G = (V, E)$ una gráfica de n vértices con grado mínimo $\delta > 1$. Entonces G tiene un conjunto dominante de a lo más $n \frac{1 + \ln(\delta + 1)}{\delta + 1}$ vértices.*

Demostración. Sea $p \in [0, 1]$ arbitrario por el momento. Lo que haremos será construir un conjunto dominante de G de manera arbitraria según p , estimar cardinalidad esperada de dicho conjunto y optimizar para p . Primero creamos un subconjunto $X \subseteq V$ seleccionando cada vértice de V con probabilidad p (es decir, la probabilidad de que un vértice dado haya sido seleccionado para estar en X es p); X elegido de esta manera no es necesariamente dominante, así definimos Y_X como los vértices en $V \setminus X$ que no tienen vecinos en X ; ahora sí $U = X \cup Y_X$ claramente es dominante. El valor esperado de $|X|$ es np , y la probabilidad para cada $v \in V$ de haber caído en Y_X es

$$P(v \in Y_X) = P(v \text{ y sus vecinos no cayeron en } X) \leq (1 - p)^{\delta + 1};$$

por lo tanto, el valor esperado para $|Y|$ es a lo más $n(1 - p)^{\delta + 1}$ y el de $|X| + |Y_X| = |U|$ es a lo más $np + n(1 - p)^{\delta + 1}$. Ahora, en vista de 2.2.14, tenemos

$$|U| \leq np + ne^{-p(\delta + 1)}.$$

Derivamos el lado derecho respecto a p e igualamos a cero para minimizar

$$\begin{aligned}
n - ne^{-p(\delta+1)}(\delta+1) &= 0 \Rightarrow \\
1 - e^{-p(\delta+1)}(\delta+1) &= 0 \Rightarrow \\
1 &= e^{-p(\delta+1)}(\delta+1) \Rightarrow \\
\frac{1}{(\delta+1)} &= e^{-p(\delta+1)} \Rightarrow \\
\ln\left(\frac{1}{(\delta+1)}\right) &= \ln(e^{-p(\delta+1)}) \Rightarrow \\
-\ln(\delta+1) &= -p(\delta+1) \Rightarrow \\
p &= \frac{\ln(\delta+1)}{\delta+1}
\end{aligned}$$

y, por lo tanto, con $p = \frac{\ln(\delta+1)}{\delta+1}$ tenemos

$$|U| \leq np + ne^{-p(\delta+1)} = n\left(\frac{\ln(\delta+1)}{\delta+1}\right) + ne^{-\left(\frac{\ln(\delta+1)}{\delta+1}\right)(\delta+1)} = n\frac{1 + \ln(\delta+1)}{\delta+1}.$$

Entonces existe un conjunto dominante con a lo más $n\frac{1+\ln(\delta+1)}{\delta+1}$ vértices. \square

Daremos una segunda prueba para este resultado sin usar métodos probabilísticos,

Segunda demostración de 2.2.15. Ésta será una prueba algorítmica y consistirá de en cada paso ir tomando el vértice que tenga más vecinos. Sea $v \in V$. Denotemos por $C(v)$ al conjunto que consiste de v y todos sus vecinos. Llamemos r al número de vértices u que durante el proceso no han sido cubiertos, es decir, que no han sido elegidos y no son vecinos de los ya elegidos. Ya que el grado mínimo es δ , la suma de las cardinalidades de los conjuntos $C(u)$ es al menos $r(\delta+1)$; promediando, esto nos dice que existe un vértice v que pertenece a al menos $\frac{r(\delta+1)}{n}$ de los conjuntos $C(u)$. Notemos que este vértice no pertenece a los ya elegidos pues, de hecho, ninguno de los u o sus vecinos puede pertenecer ya que los u no han sido cubiertos. Agreguemos este vértice v a los elegidos, y ahora la cantidad de vértices no cubiertos es a lo más $r(1 - \frac{\delta+1}{n})$. Procediendo de esta manera, cada paso por lo menos reduciremos la cantidad de vértices no cubiertos en un factor $(1 - \frac{\delta+1}{n})$. Entonces, antes de empezar a elegir vértices teníamos n vértices sin cubrir, después de s pasos nos quedarán a lo más $n(1 - \frac{\delta+1}{n})^s$ vértices sin cubrir, si elegimos $s = \frac{n}{\delta+1}\ln(\delta+1)$ entonces tendremos a lo más $n(1 - \frac{\delta+1}{n})^{\frac{n}{\delta+1}\ln(\delta+1)} = n\left((1 - \frac{\delta+1}{n})^n\right)^{\frac{\ln(\delta+1)}{\delta+1}}$, pero el lema 2.2.14 anterior nos dice que $(1 - \frac{\delta+1}{n})^n < e^{-(\delta+1)}$, por lo tanto $n\left((1 - \frac{\delta+1}{n})^n\right)^{\frac{\ln(\delta+1)}{\delta+1}} < ne^{-(\delta+1)\frac{\ln(\delta+1)}{\delta+1}} = \frac{n}{\delta+1}$; es decir, que a lo más tenemos $\frac{n}{\delta+1}$ vértices sin cubrir. Si añadimos éstos a los ya elegidos obtendremos un conjunto dominante de a lo más $\frac{n}{\delta+1}\ln(\delta+1) + \frac{n}{\delta+1} = n\frac{1+\ln(\delta+1)}{\delta+1}$. \square

Observando el segundo factor, $\frac{1+\ln(\delta+1)}{\delta+1}$, vemos que el denominador es de orden lineal sobre δ mientras que el numerador es logarítmico y, como muestra 1.3.8, el cociente tiende a cero conforme δ crece y el resultado nos garantiza conjuntos dominantes pequeños. Esto era de esperarse puesto que entre más alto grado tengan los vértices, menos serán necesarios para construir un conjunto dominante. Sin embargo el resultado hace notar un comportamiento “global”, sin importar el valor de n siempre existen conjuntos dominantes cuyo tamaño está acotado proporcionalmente según δ .

A pesar de la hipótesis del teorema, se pueden encontrar conjuntos dominantes con $n\frac{1+\ln(\delta+1)}{\delta+1}$ aún si $\delta = 0$ o $\delta = 1$. Si $\delta = 0$ tenemos $n\frac{1+\ln(\delta+1)}{\delta+1} = n\frac{1+\ln(0+1)}{0+1} = n$. Obviamente si tomamos todos los vértices de una gráfica obtendremos un conjunto dominante y para el caso de una gráfica sin aristas es el conjunto dominante más pequeño que podemos encontrar. Si $\delta > 0$ siempre puede encontrarse un conjunto dominante con $\frac{n}{2}$ vértices. La demostración es realmente sencilla, pero hay que hacer unas cuantas definiciones. Sea $G = (V, E)$ y sean $x, y \in V$ distintos. Una **xy -trayectoria** es un conjunto de vértices $\{x, x_1, x_2, \dots, x_t, y\}$ con $x \neq x_i \neq y$ y $x_i \neq x_j$ si $i \neq j$, tales que $xx_1, x_1x_2, \dots, x_ix_{i+1}, \dots, x_t y \in E$. En otras palabras podemos decir que si existe una xy -trayectoria, entonces x y y están conectados por medio de aristas con otros vértices. Decimos que G es **conexa** si para cualesquiera dos vértices distintos x, y existe una xy -trayectoria. Una componente conexa de G es una subgráfica conexa que no está contenida propiamente en otra subgráfica conexa. Un **árbol** es una gráfica conexa sin ciclos. Toda gráfica conexa contiene un árbol que contiene a todos los vértices de la gráfica. Además, los árboles son gráficas bipartitas, es decir, sus vértices se pueden dividir en dos conjuntos donde dentro de cada conjunto no existan vértices adyacentes (véase 3.2.1). Ahora tenemos todos los elementos necesarios para demostrar que toda gráfica con grado mínimo $\delta > 0$ tiene un conjunto dominante con a lo más $\frac{n}{2}$ vértices. Sea G una gráfica con grado mínimo $\delta > 0$, y sea $C = \{H_i = (V_i, E_i) | H_i \text{ es componente conexa de } G\}$ el conjunto de sus componentes conexas. Como H_i es conexa contiene un árbol T_i que contiene a todos sus vértices. Este árbol es bipartito. Sean A_i, B_i conjuntos de vértices que forman la partición de T_i . Cualquiera de A_i o B_i es un conjunto dominante. Veamos que A_i es dominante. Sea $x \in B_i$, como $\delta > 0$ y T_i es una gráfica conexa x debe tener al menos un vecino. Por definición de bipartición, ninguno de los vecinos de x está en B_i , por lo tanto x tiene un vecino en A_i y esto demuestra que A_i es dominante. La demostración de que B_i es dominante es análoga. Ahora, alguno de A_i y B_i tiene a lo más la mitad de los vértices de H_i . Entonces con a lo más $\frac{|V_i|}{2}$ vértices podemos dar un conjunto dominante de H_i . Uniendo los conjuntos dominantes de cada H_i obtendremos un conjunto dominante de G . Entonces podemos dar un conjunto dominante para G con a lo más $\sum_{H_i \in C} \frac{|V_i|}{2} = \frac{1}{2} \sum_{H_i \in C} |V_i| = \frac{|V|}{2}$ vértices.

Esto último vuelve inútil al teorema siempre que $\frac{1+\ln(\delta+1)}{\delta+1} > \frac{1}{2}$. El primer valor de δ para el que este número es menor que $\frac{1}{2}$ es $\delta = 5$.

Como ejemplo, si $\delta = 5$ entonces $n \frac{1+\ln(1+5)}{1+5} \approx n(0.47)$. Para $n = 1000$, el teorema dice que podemos encontrar un conjunto dominante con a lo más 470 vértices, es decir 30 vértices menos que lo que logramos con el razonamiento anterior.

2.2.16 Definición. Una **hipergráfica** es un par ordenado $H = (V, E)$, donde V es un conjunto finito a cuyos elementos se les llama *vértices* y E es una familia de subconjuntos de V llamados *aristas*.

2.2.17 Definición. Se dice que una hipergráfica es **k -uniforme** si cada una de sus aristas tiene exactamente k vértices.

2.2.18 Definición. Decimos que una hipergráfica H es **2-coloreable** si hay una 2-coloración de V tal que ninguna arista es monocromática.

Por ejemplo si H es una gráfica 2-uniforme (una gráfica usual), entonces H es 2-coloreable si y sólo si es bipartita, es decir si el conjunto de vértices se puede partir en dos conjuntos tales que no haya aristas dentro de ellos (véase 3.2.1), pues si la gráfica es bipartita, basta colorear cada conjunto en que se parte con un color distinto y así no habrá aristas monocromáticas. Por otro lado si la gráfica es 2-coloreable, basta con proponer la partición en donde cada conjunto de la partición tiene todos los vértices de un mismo color.

2.2.19 Definición. Denotamos por $\mathbf{m}(k)$ al menor número posible de aristas de una hipergráfica k -uniforme que no es 2-coloreable.

Por ejemplo $m(2) = 3$ pues, por un lado, un triángulo siempre tiene 2 vértices del mismo color, y, por otro lado, las gráficas de dos aristas sí son 2-coloreables. Para ver esto notemos que podemos clasificar a las gráficas de dos aristas en dos tipos: cuando las aristas no comparten vértices y cuando las aristas comparten un vértice. En las que no comparten vértices basta con colorear de color distinto los vértices de cada arista, y en las que comparten un vértice basta con colorear de un color al vértice que comparten y con el otro color a los otros vértices.

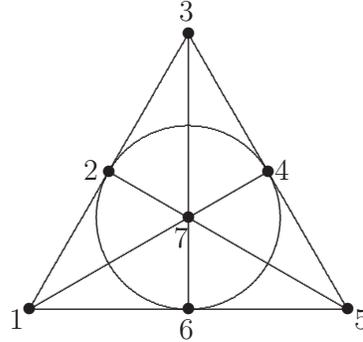
Para que sea más claro veamos el siguiente ejemplo.

2.2.20 Teorema. $m(3) = 7$.

Demostración. Primero veamos que $m(3) \leq 7$. Para esto basta proponer una gráfica 3-uniforme con 7 aristas que no sea 2-coloreable.

Proponemos como gráfica el **plano de fano**, en el que $V = \{1, 2, 3, 4, 5, 6, 7\}$ y

$$E = \{\{1, 2, 3\}, \{1, 4, 7\}, \{1, 5, 6\}, \{3, 4, 5\}, \{3, 6, 7\}, \{2, 5, 7\}, \{2, 4, 6\}\}$$



Supongamos que sí es 2-coloreable y tomemos una 2-coloración que no deje aristas monocromáticas. Llamemos rojo y verde a los colores. Primero hay que notar que cualesquiera dos vértices se encuentran en exactamente una arista. Ahora, en cualquier coloración que se dé, existirá un color que tenga más vértices, sin pérdida de generalidad digamos que es el rojo. Si hay 4 vértices rojos, como supusimos que no hay aristas monocromáticas, cualquier par de vértices rojos está acompañado de uno verde en la arista que determinan, es decir que cada par de vértices rojos determina una arista distinta, por lo cual el conjunto de vértices rojos determina $\binom{4}{2} = 6$ aristas. Usando el mismo argumento con los 3 vértices verdes que quedan, como no forman una arista, entonces cada par de vértices verdes determina una arista distinta, es decir $\binom{3}{2} = 3$ aristas. Nótese que las aristas generadas por el conjunto de vértices rojos son distintas de las generadas por el conjunto de vértices verdes, pues unas tienen dos vértices rojos y uno verde, y las otras dos vértices verdes y uno rojo, por lo tanto en total tendríamos 9 aristas distintas. Pero esto es una contradicción pues sólo existen 7 aristas. Si hay al menos 5 vértices rojos, usando el mismo argumento que antes, en vista de que supusimos que no hay aristas monocromáticas, tenemos que cada par de vértices rojos determina una arista distinta. Entonces tendríamos al menos $\binom{5}{2} = 10$ aristas pero sólo existen 7.

Ahora veamos que $m(3) \geq 7$. Para esto basta demostrar que cualquier hipergráfica 3-uniforme con 6 aristas es 2-coloreable.

Sea $H = (V, E)$ cualquier hipergráfica 3-uniforme con 6 aristas. Partiremos la demostración en 2 casos: cuando $|V| \leq 6$ y cuando $|V| > 6$. Si $|V| \leq 6$ podemos suponer que $|V| = 6$ pues si no es así podemos agregar vértices de tal manera que $|V| = 6$ y esto no afecta las condiciones de coloración. Para cualquier conjunto de 3 vértices definimos la coloración que pinta de rojo a estos 3 vértices y de verde al resto; de esta manera tenemos $\binom{6}{3} = 20$ coloraciones. Para una arista e dada tenemos que la probabilidad de que sea monocromática bajo una de estas coloraciones es $P(e \text{ es monocromática}) = \frac{2}{20} = \frac{1}{10}$ pues puede ser totalmente roja o totalmente verde. Dado que tenemos 6 aristas, la probabilidad de que alguna sea monocromática bajo una de estas coloraciones es $P(\bigcup_{e \in E} (e \text{ es monocromática})) \leq \sum_{e \in E} P(e \text{ es monocromática}) = \frac{6}{10} < 1$ y, por lo tanto, existe una coloración en

la que ninguna arista resulta ser monocromática. Usaremos esta última parte como base de inducción para la demostración del otro caso. Nuestra hipótesis de inducción es que para cierta $n \geq 6$, cualquier hipergráfica $H = (V, E)$ con $|V| = n \geq 6$ es 2-coloreable. Sea $H = (V, E)$ con $|V| = n + 1$. Notemos que siempre existen dos vértices que no están en una misma arista, pues al menos tenemos 7 vértices, y por lo tanto tenemos al menos $\binom{7}{2} = 21$ parejas de vértices, pero sólo 18 parejas se encuentran en una misma arista ya que cada arista solo tiene $\binom{3}{2} = 3$ parejas de vértices. Tomemos entonces $x, y \in V$ que no se encuentren ambos en una misma arista y construyamos la gráfica que resulta de “pegar” estos vértices en un vértice auxiliar z , es decir $V' = V \setminus \{x, y\} \cup \{z\}$ y $E' = \{e \in E : e \cap \{x, y\} = \emptyset\} \cup \{e \setminus \{x, y\} \cup \{z\} : e \in E, e \cap \{x, y\} \neq \emptyset\}$. Notemos que $H' = (V', E')$ es tal que $|V'| = n$ y $|E'| = |E| = 6$ y ésta, por hipótesis de inducción, es 2-coloreable. Dada una 2-coloración que no deje aristas monocromáticas podemos extenderla a H poniendo a x y a y del color que le correspondía a z , y ésta será una 2-coloración que no deja aristas monocromáticas porque x y y no estaban en una misma arista. □

Nótese que entre menos aristas tenga una hipergráfica, es más fácil que ésta sea 2-coloreable, sin embargo, es difícil garantizar cuántas aristas son necesarias para que una hipergráfica no sea 2-coloreable. El siguiente resultado nos da una cota inferior para $m(n)$.

2.2.21 Teorema (Erdős). *Toda hipergráfica n -uniforme con menos de 2^{n-1} aristas es 2-coloreable. En particular $m(n) \geq 2^{n-1}$.*

Demostración. Sea $e \in E$ y definamos el suceso A_e en el que e es monocromática, obviamente coloreando los vértices aleatoriamente con probabilidad uniforme $P(A_e) = \frac{2}{2^n} = 2^{1-n}$. Entonces la probabilidad de que alguna arista sea monocromática cumple que:

$$P\left(\bigcup_{e \in E} A_e\right) \leq \sum_{e \in E} P(A_e) < 2^{n-1} \cdot 2^{1-n} = 1,$$

por lo tanto existen coloraciones en las que no hay aristas monocromáticas. □

La cota inferior que nos ofrece el resultado anterior tiene una tasa de crecimiento exponencial; esto nos dice que $m(n)$ crece muy rápido cuando en un principio ni siquiera era muy claro en que momento se podía garantizar que una hipergráfica no era 2-coloreable.

2.3. Aplicaciones a Combinatoria

2.3.1 Definición. Sea $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ un conjunto de parejas ordenadas de subconjuntos de un conjunto arbitrario. Decimos que \mathcal{F} es un **sistema-(\mathbf{k}, \mathbf{l})** si $|A_i| = k$ y $|B_i| = l$ para todo $1 \leq i \leq h$, y $A_i \cap B_i = \emptyset$, y $A_i \cap B_j \neq \emptyset$ para todos los

i, j con $1 \leq i < j \leq h$.

Por ejemplo

$$\mathcal{F} = \{(\{1, 2\}, \{3, 4\}), (\{1, 3\}, \{2, 4\}), (\{2, 3\}, \{1, 4\})\}$$

es un sistema-(2, 2).

El siguiente teorema nos proporciona una cota superior para la cantidad de elementos que puede llegar a tener un sistema-(k, l).

2.3.2 Teorema. Si $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ es un sistema-(k, l) entonces $h \leq \binom{k+l}{k}$.

Demostración. Sea $X = \bigcup_{i=1}^h (A_i \cup B_i)$ y sea π un orden lineal arbitrario para X . Para cada $1 \leq i \leq h$ sea E_i el suceso en el que todos los elementos de A_i preceden a los de B_i según π . Tenemos que $P(E_i) = \frac{1}{\binom{k+l}{k}}$, pues dentro de los $k+l$ espacios que ocupan los elementos de $A_i \cup B_i$ el único acomodo en donde todos los elementos de A_i preceden a los de B_i , es cuando los elementos de A_i son los primeros k , y los elementos de A_i pudieron acomodarse de $\binom{k+l}{k}$ formas dentro de los $k+l$ espacios. Veamos que los sucesos E_i son ajenos dos a dos. Suponiendo que no lo son, sean i, j distintos tales que todos los elementos A_i preceden a los de B_i , y todos los de A_j preceden a los de B_j . Sin pérdida de generalidad digamos que el elemento más grande de A_j es mayor o igual a todos los de A_i , pero entonces todos los elementos de B_j son más grandes que todos los elementos de A_i , contradiciendo el que $A_i \cap B_j \neq \emptyset$. Ahora, la probabilidad de que todos los elementos de algún A_i precedan a los de B_i cumple que

$$1 \geq P\left(\bigcup_{i=1}^h E_i\right) = \sum_{i=1}^h P(E_i) = h \cdot \frac{1}{\binom{k+l}{k}} \Rightarrow$$

$$\binom{k+l}{k} \geq h,$$

con lo que termina la demostración. □

Si ahora debilitamos las hipótesis considerando $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ un conjunto de parejas ordenadas de subconjuntos de un conjunto arbitrario tales que $A_i \cap B_i = \emptyset$, y $A_i \cap B_j \neq \emptyset$ para todos los i, j con $1 \leq i < j \leq h$ (esta vez no condicionamos el tamaño de los conjuntos), la misma demostración que usamos en el teorema anterior nos sirve para asegurar el siguiente resultado.

2.3.3 Lema. Si $\mathcal{F} = \{(A_i, B_i)\}_{i=1}^h$ es un conjunto de parejas ordenadas de subconjuntos de un conjunto arbitrario tales que $A_i \cap B_i = \emptyset$ y $A_i \cap B_j \neq \emptyset$ para todos los i, j con $1 \leq i < j \leq h$, entonces $\sum_{i=1}^h \left(\frac{|A_i|+|B_i|}{|A_i|}\right)^{-1} \leq 1$.

Demostración. Otra vez consideremos $X = \bigcup_{i=1}^h (A_i \cup B_i)$ y π un orden lineal arbitrario para X , y sea E_i el suceso en el que todos los elementos de A_i preceden a los de B_i según π . Una vez más tenemos que los sucesos E_i son ajenos, pues si E_i y E_j con $i \neq j$ ocurrieran simultáneamente, sin pérdida de generalidad, digamos que el elemento más grande de A_j es mayor o igual a todos los elementos de A_i , pero entonces todos los elementos de B_j son más grandes que todos los elementos de A_i , contradiciendo el que $A_i \cap B_j \neq \emptyset$. Tenemos que $P(E_i) = \frac{1}{\binom{|A_i|+|B_i|}{|A_i|}}$ pues dentro de los $|A_i| + |B_i|$ espacios que ocupan los elementos de $A_i \cup B_i$ el único acomodo en donde todos los elementos de A_i preceden a los de B_i es cuando los elementos de A_i son los primeros $|A_i|$, y los elementos de B_i pudieron acomodarse de $\binom{|A_i|+|B_i|}{|A_i|}$ formas dentro de los $|A_i| + |B_i|$ espacios. Entonces, por ser los E_i sucesos ajenos, tenemos que

$$1 \geq P\left(\bigcup_{i=1}^h E_i\right) = \sum_{i=1}^h P(E_i) = \sum_{i=1}^h \frac{1}{\binom{|A_i|+|B_i|}{|A_i|}}.$$

□

Este lema demuestra fácilmente el teorema de Sperner.

2.3.4 Definición. Dados A y B conjuntos, se dice que A y B son **incomparables** si $A \not\subset B$ y $B \not\subset A$.

2.3.5 Teorema (Sperner). Sea \mathcal{S} una familia de subconjuntos incomparables por parejas de un conjunto X de n elementos. Entonces $|\mathcal{S}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$.

Demostración. Sea $\mathcal{F} = \{(A, X \setminus A) : A \in \mathcal{S}\}$. Tenemos que $|\mathcal{F}| = |\mathcal{S}|$ y que \mathcal{F} cumple las hipótesis del lema anterior puesto que $A \cap (X \setminus A) = \emptyset$ para todo $A \in \mathcal{S}$ y si $A, B \in \mathcal{S}$ son distintos entonces $A \cap (X \setminus B) \neq \emptyset$ dado que A y B son incomparables. Por lo tanto, el lema anterior nos dice que $\sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \leq 1$. Por otro lado tenemos que $\binom{n}{k} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$ para todo k tal que $0 \leq k \leq n$. Combinando esto tenemos que $\sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{|A|}} \leq \sum_{A \in \mathcal{F}} \frac{1}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} = \frac{|\mathcal{F}|}{\binom{n}{\lfloor \frac{n}{2} \rfloor}} \leq 1$ y entonces $|\mathcal{S}| = |\mathcal{F}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. □

Parte de la importancia del método probabilístico radica en la sencillez de las demostraciones que se pueden dar. Para hacer notar esto haremos una segunda demostración del teorema anterior que no usa el método probabilístico.

Segunda demostración de 2.3.5. Sea \mathcal{S} una familia de subconjuntos incomparables por parejas de un conjunto X de n elementos. Llamemos cadena a cualquier familia \mathcal{C} de subconjuntos de X en la que cualesquiera dos elementos sean comparables, es decir, si $A, B \in \mathcal{C}$ entonces $A \subset B$ ó $B \subset A$. Digamos que una cadena \mathcal{C} es máxima si no existe otra cadena que la contenga propiamente, es decir, si $A \subset X$ y $A \notin \mathcal{C}$ entonces $\mathcal{C} \cup \{A\}$ no es cadena. Tenemos que el número de cadenas máximas es $n!$

pues una cadena máxima se obtiene de ir agregando uno a uno los elementos de X , además, si $A \subset X$, $|A| = r$ el número c_r de cadenas máximas que tienen a A es $c_r = r!(n-r)!$ pues una cadena máxima que tenga a A se obtiene de ir agregando a A uno a uno elementos para obtener a los elementos de la cadena que lo contendrán y de ir quitando uno a uno los elementos de A para obtener los elementos que estarán contenidos en A . Por otro lado tenemos que $\binom{n}{r} \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$ para todo r tal que $0 \leq r \leq n$, de donde $\frac{n!}{r!(n-r)!} \leq \frac{n!}{\lfloor \frac{n}{2} \rfloor!(n-\lfloor \frac{n}{2} \rfloor)!}$ y entonces $\lfloor \frac{n}{2} \rfloor!(n-\lfloor \frac{n}{2} \rfloor)! \leq r!(n-r)! = c_r$. Notemos que cualquier cadena máxima contiene a lo más un elemento de \mathcal{S} pues los elementos de \mathcal{S} son incomparables por parejas. Ya que hay $n!$ cadenas máximas en total y cada elemento de \mathcal{S} está en al menos $\lfloor \frac{n}{2} \rfloor!(n-\lfloor \frac{n}{2} \rfloor)!$ cadenas máximas tenemos que $|\mathcal{S}| \lfloor \frac{n}{2} \rfloor!(n-\lfloor \frac{n}{2} \rfloor)! \leq n!$, de donde $|\mathcal{S}| \leq \binom{n}{\lfloor \frac{n}{2} \rfloor}$. \square

2.3.6 Definición. Se dice que un subconjunto A de un grupo abeliano G es **libre de suma** si $(A+A) \cap A = \emptyset$ (donde $+$ representa a la operación del grupo), es decir, si no existen $a_1, a_2, a_3 \in A$ tales que $a_1 + a_2 = a_3$.

Darse a la tarea de encontrar un subconjunto libre de suma dentro de un grupo abeliano parece no ser una tarea fácil si lo que esperamos es obtener un subconjunto con un “buen tamaño”.

A continuación se demuestra un teorema que garantiza que de cualquier conjunto de enteros finito (que no contenga al cero) podemos extraer un subconjunto libre de suma (visto como subconjunto de los enteros como grupo) de cardinalidad mayor a un tercio del primero. Necesitamos un lema.

2.3.7 Lema. *Existe una infinidad de primos de la forma $3k + 2$.*

Demostración. Supongamos lo contrario. Sea $D = \{p_1, \dots, p_n\}$ el conjunto de todos los primos de la forma $3k + 2$; sin pérdida de generalidad digamos que $p_1 = 2$ y sea $m = 3(\prod_{i=2}^n p_i) + 2$. Tenemos que $m \equiv 2 \pmod{3}$. Como el producto de dos números de la forma $3k + 1$ es de esa misma manera y en D se encuentran todos los primos congruentes con 2 módulo 3, alguno de ellos tendría que dividir a m . Primero veamos que 2 no puede ser pues $2 \nmid 3(\prod_{i=2}^n p_i)$; ahora, para p_i con $1 < i \leq n$, tenemos que $p_i \mid 3(\prod_{i=2}^n p_i)$ pero $p_i \nmid 2$ y por lo tanto $p_i \nmid m$. Esto es una contradicción. \square

2.3.8 Teorema (Erdős). *Todo conjunto $B = \{b_1, \dots, b_n\}$ de n enteros distintos de cero contiene un subconjunto libre de suma A tal que $|A| > \frac{n}{3}$.*

Demostración. Sea $p = 3k + 2$ un primo tal que $p > 2 \cdot \max_{1 \leq i \leq n} |b_i|$ y definamos $C = \{k+1, k+2, \dots, 2k+1\} \subset \mathbb{Z}_p$. Veamos que C es un subconjunto libre de suma en \mathbb{Z}_p . En caso de que no lo fuera existirían $a, b, c \in C$ tales que $a + b \equiv c \pmod{p}$ o, lo que es lo mismo, $a + b - c \equiv 0 \pmod{p}$. Escribamos $a = k+i, b = k+j$, y $c = k+l$. Entonces $a + b - c = (k+i) + (k+j) - (k+l) = k+i+j-l$. Pero $1 \leq i, j, k \leq k+1$,

así que $a+b-c \leq k+(k+1)+(k+1)-1 = 3k+1$ y $a+b-c \geq k+(1)+(1)-(k+1) = 1$, de donde $a+b-c$ no puede ser $0 \pmod{p}$. Tomemos un entero aleatorio x , $1 \leq x < p$, y definamos d_1, d_2, \dots, d_n como sigue: d_i es el residuo de xb_i bajo la división entre p . Haber tomado a p mayor que $2 \cdot \max_{1 \leq i \leq n} |b_i|$ nos garantiza que $b_i \not\equiv b_j \pmod{p}$ si $i \neq j$ y, por lo tanto, $d_i \not\equiv d_j \pmod{p}$. Al variar los residuos x sobre $\mathbb{Z}_p \setminus \{0\}$, d_i también varía sobre los residuos módulo p . Entonces $P(d_i \in C) = \frac{|C|}{p-1} = \frac{k}{3k+1} > \frac{1}{3}$. Así el número esperado de elementos b_i tales que $d_i \in C$ es mayor a $\frac{n}{3}$ y por lo tanto existe $x \in \mathbb{Z}_p \setminus \{0\}$ y $A \subset B$ con $|A| > \frac{n}{3}$ tal que el residuo de la división entre p de xa es un elemento de C para todo $a \in A$. Claramente este A es libre de suma, de lo contrario existirían $a_1, a_2, a_3 \in A$ tales que $a_1 + a_2 = a_3$ pero esto implicaría $xa_1 + xa_2 \equiv xa_3 \pmod{p}$ lo cual contradice que C es libre de suma en \mathbb{Z}_p . \square

2.4. Parejas ajenas

2.4.1 Definición. Sea \mathcal{F} una familia de m subconjuntos de $X = \{1, 2, \dots, n\}$. Definimos $d(\mathcal{F})$ como el número de parejas ordenadas ajenas en \mathcal{F} , es decir

$$d(\mathcal{F}) = |\{(F, F') : F, F' \in \mathcal{F}, F \cap F' = \emptyset\}|.$$

Por ejemplo para $\mathcal{F} = \mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$ tenemos $d(\mathcal{F}) = 9$.

El siguiente teorema garantiza que la cantidad de subconjuntos ajenos de un conjunto finito crece más lento que la cantidad de parejas posibles. Necesitamos antes un lema de cálculo.

2.4.2 Definición. Decimos que una función $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ es **convexa** si para cualesquiera $x_1, x_2 \in D$ y $t \in [0, 1]$ se tiene que $f(t(x_1) + (1-t)(x_2)) \leq t(f(x_1) + (1-t)f(x_2))$. En otras palabras, el conjunto de puntos del plano que quedan por encima de la gráfica de la función es un conjunto convexo. Nótese que esta definición implica que D es convexo.

2.4.3 Lema (Desigualdad de Jensen). Si $\varphi : D \subset \mathbb{R} \rightarrow \mathbb{R}$ es una función convexa, $x_1, x_2, \dots, x_n \in D$ y a_1, a_2, \dots, a_n son reales positivos tales que $\sum_{i=1}^n a_i = 1$, entonces

$$\varphi \left(\sum_{i=1}^n a_i x_i \right) \leq \sum_{i=1}^n a_i \varphi(x_i),$$

Demostración. Procedamos por inducción. Para $n = 1$ es obvio y, para $n = 2$ para el el resultado coincide con la definición de función convexa. Supongamos que el resultado es cierto para algún entero $n \geq 2$. Sean $x_1, x_2, \dots, x_{n+1} \in D$ y $a_1, a_2, \dots, a_{n+1} \in \mathbb{R}^+$ tales que $\sum_{i=1}^{n+1} a_i = 1$.

$$\varphi \left(\sum_{i=1}^{n+1} a_i x_i \right) = \varphi \left(\sum_{i=1}^n a_i x_i + a_{n+1} x_{n+1} \right) = \varphi \left((1 - a_{n+1}) \sum_{i=1}^n \frac{a_i x_i}{1 - a_{n+1}} + a_{n+1} x_{n+1} \right).$$

Tenemos que tanto $\sum_{i=1}^n \frac{a_i x_i}{1-a_{n+1}}$ como x_{n+1} pertenecen a D y, además $(1-a_{n+1}) + a_{n+1} = 1$. Así que, por convexidad de φ tenemos:

$$\varphi \left((1-a_{n+1}) \sum_{i=1}^n \frac{a_i x_i}{1-a_{n+1}} + a_{n+1} x_{n+1} \right) \leq (1-a_{n+1}) \varphi \left(\sum_{i=1}^n \frac{a_i x_i}{1-a_{n+1}} \right) + a_{n+1} \varphi(x_{n+1})$$

Ahora, como $\sum_{i=1}^{n+1} a_i = 1$, entonces $\sum_{i=1}^n a_i = 1 - a_{n+1}$, de donde $\sum_{i=1}^n \frac{a_i}{1-a_{n+1}} = 1$ y, por lo tanto, podemos aplicar la hipótesis de inducción a $\varphi \left(\sum_{i=1}^n \frac{a_i x_i}{1-a_{n+1}} \right)$, de donde

$$\begin{aligned} (1-a_{n+1}) \varphi \left(\sum_{i=1}^n \frac{a_i x_i}{1-a_{n+1}} \right) + a_{n+1} \varphi(x_{n+1}) &\leq \\ (1-a_{n+1}) \sum_{i=1}^n \frac{a_i}{1-a_{n+1}} \varphi(x_i) + a_{n+1} \varphi(x_{n+1}) &= \\ \sum_{i=1}^n a_i \varphi(x_i) + a_{n+1} \varphi(x_{n+1}) &= \sum_{i=1}^{n+1} a_i \varphi(x_i). \end{aligned}$$

□

Nótese que en particular si $a_1, a_2, \dots, a_n \in \mathbb{R}^+$ y $x_1, x_2, \dots, x_n \in D$, entonces $\varphi \left(\frac{\sum_{i=1}^n a_i x_i}{\sum_{i=1}^n a_i} \right) \leq \frac{\sum_{i=1}^n a_i \varphi(x_i)}{\sum_{i=1}^n a_i}$.

2.4.4 Teorema. Sea \mathcal{F} una familia de $m = 2^{\frac{1}{2}+\delta n}$ subconjuntos de $X = \{1, 2, \dots, n\}$, donde $0 < \delta \leq \frac{1}{2}$. Entonces

$$d(\mathcal{F}) < m^{2-\frac{\delta^2}{2}}. \quad (1)$$

Demostración. Supongamos falso (1) y tomemos aleatoriamente $t = \lceil 1+1/\delta \rceil$ miembros A_1, A_2, \dots, A_t de \mathcal{F} con posible repetición,

Mostraremos que, con probabilidad positiva, $|A_1 \cup A_2 \cup \dots \cup A_t| > \frac{n}{2}$ y que esta unión es ajena a más de $2^{\frac{n}{2}}$ subconjuntos de X , lo cual es una contradicción, pues $|X - (A_1 \cup A_2 \cup \dots \cup A_t)| < \frac{n}{2}$ y, por lo tanto, en este subconjunto de X hay menos de $2^{\frac{n}{2}}$ subconjuntos ajenos a $A_1 \cup A_2 \cup \dots \cup A_t$. Tenemos

$$\begin{aligned} P(|A_1 \cup A_2 \cup \dots \cup A_t|) &\leq \sum_{\substack{S \subset X \\ |S| \leq \frac{n}{2}}} P(A_i \subset S, i = 1, 2, \dots, t) \leq 2^n \left(\frac{2^{\frac{n}{2}}}{m} \right)^t = \\ 2^n \left(\frac{2^{\frac{n}{2}}}{2^{\frac{1}{2}+\delta n}} \right)^t &= 2^n \left(2^{-n\delta} \right)^t = 2^n \left(2^{-n\delta t} \right) = 2^{n-n\delta t} = 2^{n(1-\delta t)}. \end{aligned} \quad (2)$$

Ahora definamos

$$v(B) = |\{A \in \mathcal{F} : A \cap B = \emptyset\}|.$$

Es fácil notar que $\sum_{B \in \mathcal{F}} v(B) = 2 \cdot d(\mathcal{F}) \geq 2m^{2-\frac{\delta^2}{2}}$. La última desigualdad se cumple debido a que supusimos falso (1). Sea Y una variable aleatoria cuyo valor es el número de elementos $B \in \mathcal{F}$ que son ajenos a todos los A_i ($1 \leq i \leq t$). Damos una cota mínima para $E(Y)$

$$E(Y) = \sum_{B \in \mathcal{F}} \left(\frac{v(B)}{m} \right)^t = \frac{1}{m^t} \cdot m \left(\frac{\sum_{B \in \mathcal{F}} v(B)^t}{m} \right).$$

Por la convexidad de la función x^t , y la desigualdad de Jensen (2.4.3) tenemos que

$$\begin{aligned} \frac{1}{m^t} \cdot m \left(\frac{\sum_{B \in \mathcal{F}} v(B)^t}{m} \right) &\geq \frac{1}{m^t} \cdot m \left(\frac{\sum_{B \in \mathcal{F}} v(B)}{m} \right)^t = \frac{1}{m^t} \cdot m \left(\frac{2d(\mathcal{F})}{m} \right)^t \geq \\ &\frac{1}{m^t} \cdot m \left(\frac{2m^{2-\frac{\delta^2}{2}}}{m} \right)^t = \frac{1}{m^t} \cdot m \cdot \frac{2^t m^{2t-\frac{t\delta^2}{2}}}{m^t} = 2^t m^{1-\frac{t\delta^2}{2}} \geq 2m^{1-\frac{t\delta^2}{2}}. \end{aligned} \quad (3)$$

Con esto y con el hecho de que $Y \leq m$ obtenemos que

$$P\left(Y \geq m^{1-\frac{t\delta^2}{2}}\right) \geq m^{-\frac{t\delta^2}{2}}, \quad (4)$$

pues de lo contrario tendríamos

$$\begin{aligned} E(Y) &= E\left(Y < m^{1-\frac{t\delta^2}{2}}\right) + E\left(Y \geq m^{1-\frac{t\delta^2}{2}}\right) < \\ &m^{1-\frac{t\delta^2}{2}} + m \cdot m^{-\frac{t\delta^2}{2}} = m^{1-\frac{t\delta^2}{2}} + m^{1-\frac{t\delta^2}{2}} = 2m^{1-\frac{t\delta^2}{2}} \end{aligned}$$

lo cual contradice (3).

Ahora queremos mostrar que la probabilidad de que $|A_1 \cup A_2 \cup \dots \cup A_t| \leq \frac{n}{2}$ es mayor que la probabilidad de que $Y > 2^{\frac{n}{2}}$. Notemos que $0 < \delta \leq 1/2$ pues $m = 2^{(\frac{1}{2}+\delta)n}$ a lo más puede ser 2^n , y también notemos que $x \leq \lceil x \rceil < x + 1$. Entonces tenemos que $m^{(1-\frac{\delta^2}{2})} > 2^{\frac{n}{2}}$ pues obtenemos las siguientes igualdades y

desigualdades, cada una a partir de la anterior.

$$\begin{aligned}
1 - \left(\delta + \delta^2 + \frac{1}{4} \right) &\geq 1 - \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{4} \right) = 0, \\
\delta - \left(\delta^2 + \delta^3 + \frac{\delta}{4} \right) &\geq 0, \\
\frac{1}{2} + \delta - \left(\delta^2 + \delta^3 + \frac{\delta}{4} \right) &\geq \frac{1}{2}, \\
\frac{1}{2} + \delta - \left(\frac{\delta^2}{2} + \delta^3 + \frac{\delta}{4} + \frac{\delta^2}{2} \right) &\geq \frac{1}{2}, \\
\left(\frac{1}{2} + \delta \right) \left(1 - \frac{(2 + \frac{1}{\delta}) \delta^2}{2} \right) &\geq \frac{1}{2}, \\
\left(\frac{1}{2} + \delta \right) \left(1 - \frac{[1 + \frac{1}{\delta}] \delta^2}{2} \right) &> \left(\frac{1}{2} + \delta \right) \left(1 - \frac{(2 + \frac{1}{\delta}) \delta^2}{2} \right) \geq \frac{1}{2}, \\
1 - \left(\frac{\delta}{4} + \frac{\delta^2}{2} \right) &> 1 - \left(\frac{1}{8} + \frac{1}{8} \right) = 3/4 > 0, \\
\delta - \left(\frac{\delta^2}{4} + \frac{\delta^3}{2} \right) &> 0, \\
\frac{1}{2} + \delta - \left(\frac{\delta^2}{4} + \frac{\delta^3}{2} \right) &> \frac{1}{2}, \\
\left(\frac{1}{2} + \delta \right) \left(1 - \frac{\delta^2}{2} \right) &> \frac{1}{2}, \\
n \left(\frac{1}{2} + \delta \right) \left(1 - \frac{\delta^2}{2} \right) &> \frac{n}{2}, \\
\left(2^{n(\frac{1}{2} + \delta)} \right)^{\left(1 - \frac{\delta^2}{2} \right)} &> 2^{\frac{n}{2}}, \\
m^{\left(1 - \frac{\delta^2}{2} \right)} &> 2^{\frac{n}{2}}.
\end{aligned}$$

En vista de (4) esto demuestra que $Y > 2^{n/2}$ con probabilidad positiva, es decir la cantidad de subconjuntos ajenos a todos los A_i es mayor que $2^{n/2}$. Ahora veamos que esta probabilidad es más grande que la de que $|A_1 \cup A_2 \cup \dots \cup A_t| \leq \frac{n}{2}$. Para esto veamos que el lado derecho de (4) es más grande que el lado derecho de (2).

Usando que $\delta \leq \frac{1}{2}$:

$$\begin{aligned}
-\frac{\delta^2}{2} - \frac{3\delta}{4} + \frac{3}{4} &\geq -\frac{1}{8} - \frac{3}{8} + \frac{3}{4} = \frac{1}{4} > 0, \\
-\frac{\delta^3}{2} - \frac{3\delta^2}{4} + \frac{3\delta}{4} &> 0, \\
-\frac{\delta^3}{2} - \frac{3\delta^2}{4} + \frac{3\delta}{4} + 1 &> 1, \\
(\delta + 1) \left(-\frac{\delta^2}{2} - \frac{\delta}{4} + 1 \right) &> 1, \\
\left(1 + \frac{1}{\delta} \right) \left(-\frac{\delta^3}{2} - \frac{\delta^2}{4} + \delta \right) &> 1, \\
t \left(-\frac{\delta^3}{2} - \frac{\delta^2}{4} + \delta \right) = \left[1 + \frac{1}{\delta} \right] \left(-\frac{\delta^3}{2} - \frac{\delta^2}{4} + \delta \right) &\geq \left(1 + \frac{1}{\delta} \right) \left(-\frac{\delta^3}{2} - \frac{\delta^2}{4} + \delta \right), \\
-\frac{\delta^3 t}{2} - \frac{\delta^2 t}{4} + \delta t &> 1, \\
-\frac{\delta^3 t}{2} - \frac{\delta^2 t}{4} &> 1 - \delta t, \\
-\frac{\delta^3 t}{2} - \frac{\delta^2 t}{4} &> 1 - \delta t, \\
\left(\frac{1}{2} + \delta \right) \left(-t \frac{\delta^2}{2} \right) &> 1 - \delta t, \\
\left(\frac{1}{2} + \delta \right) n \left(-t \frac{\delta^2}{2} \right) &> n(1 - \delta t), \\
2^{\left(\frac{1}{2} + \delta\right)n \left(-t \frac{\delta^2}{2}\right)} &> 2^{n(1 - \delta t)}, \\
m^{-t \frac{\delta^2}{2}} &> 2^{n(1 - \delta t)} > P \left(|A_1 \cup A_2 \cup \dots \cup A_t| \leq \frac{n}{2} \right).
\end{aligned}$$

Ya que la probabilidad de que $Y > 2^{\frac{n}{2}}$ es más grande que la de que $|A_1 \cup A_2 \cup \dots \cup A_t| \leq \frac{n}{2}$, tenemos que existe un caso en donde $|A_1 \cup A_2 \cup \dots \cup A_t| > \frac{n}{2}$ y además el número de subconjuntos ajenos a $A_1 \cup A_2 \cup \dots \cup A_t$ es mayor que $2^{\frac{n}{2}}$, lo cual es una contradicción, y con esto termina la demostración. \square

Sabemos que $\binom{m}{2}$ es un polinomio cuadrático pues $\binom{m}{2} = \frac{m(m-1)}{2}$, y el último teorema garantiza que $d(\mathcal{F}) < m^{2 - \frac{\delta^2}{2}}$, cuyo lado derecho tiene una tasa de crecimiento menor a la de un polinomio cuadrático. Esto quiere decir que si tomamos cualquier δ fijo, el número de subconjuntos puede ir creciendo exponencialmente (m es exponencial con respecto a n), y sin embargo el número de subconjuntos ajenos ni siquiera crecerá cuadráticamente.

Por ejemplo, si $n = 4$ y $\delta = 1/2$, tenemos $m = 2^4 = 16$ y $d(\mathcal{F}) < 16^{2 - \frac{(1/2)^2}{2}} = 16^{\frac{15}{8}} \approx 181.0193$. Contemos la cantidad de conjuntos ajenos en el conjunto potencia de $\{0, 1, 2, 3\}$. Para esto, coloreamos los números del uno al cuatro con tres colores,

digamos que los verdes pertenecen a A , los verdes a B y los rojos no se usan. Obviamente $A \cap B = \emptyset$ bajo cualquiera de estas coloraciones, y tenemos $3^4 = 81$ coloraciones.

Esto parecería no presentar mucha ganancia en vista de que la diferencia entre el resultado que da el teorema y la cantidad de conjuntos ajenos que encontramos es muy grande. La importancia de este teorema radica en su historia. Erdős conjeturó que este número d pertenecía al conjunto de funciones $o(m^2)$. La veracidad de esta conjetura es un caso particular de este teorema, el cual demostró Noga Alon de la misma manera en que se muestra aquí.

2.4.5 Definición. Decimos que una familia \mathcal{F} de conjuntos es **intersectante** si para cualesquiera $A, B \in \mathcal{F}$ se tiene que $A \cap B \neq \emptyset$.

Si queremos una familia intersectante \mathcal{F} de subconjuntos con k elementos de un conjunto con $n \geq 2k$ elementos y queremos que tenga la mayor cardinalidad posible, lo primero que se ocurre es tomar un punto fijo que esté en todos los elementos de \mathcal{F} y elegir sin cuidado los $k - 1$ restantes para cada uno de ellos. Eso nos deja con una familia de $\binom{n-1}{k-1}$. La pregunta natural es si este número es lo más grande que podemos construir. El teorema de Erdős-Ko-Rado dice que, efectivamente, este número es lo más grande que se puede obtener. Nótese que el caso $n < 2k$ no es interesante pues, por principio de casillas, cualquier par de k -conjuntos (conjunto con k elementos) tendría intersección no vacía y entonces el número máximo alcanzable es $\binom{n}{k}$. Antes de poder probar el teorema necesitamos el siguiente lema.

2.4.6 Lema. Sean n, k naturales tales que $n \geq 2k$, $0 \leq s \leq n - 1$ y $A_s = \{s, s + 1, \dots, s + k - 1\}$, donde la suma es módulo n . Si \mathcal{F} es una familia de k -conjuntos intersectante, entonces \mathcal{F} a lo más contiene k de los conjuntos A_s .

Demostración. La demostración es sencilla; supongamos que existe al menos un $A_{s_0} \in \mathcal{F}$. El resto de los A_s que intersectan a A_{s_0} los podemos acomodar en pares de la forma $\{A_{s_0-i}, A_{s_0+k-i}\}$ con $1 \leq i \leq k-1$. Por ser $n \geq 2k$, los elementos de cada uno de estos pares son ajenos, por lo tanto \mathcal{F} a lo más puede tener un elemento de cada par y con esto termina la demostración. \square

2.4.7 Teorema. Sean n, k naturales tales que $n \geq 2k$. Si \mathcal{F} es una familia intersectante de k -subconjuntos de $\{1, 2, \dots, n\}$, entonces $|\mathcal{F}| \leq \binom{n-1}{k-1}$.

Demostración. Sea $\sigma \in S_n$ una permutación escogida al azar y de manera uniforme y sea $A = \{\sigma(i), \sigma(i+1), \dots, \sigma(i+k-1)\}$, donde la suma es módulo n . A es esencialmente uno de los conjuntos A_s del lema anterior y, por lo tanto, $P(A \in \mathcal{F}) \leq \frac{k}{n}$. Por otro lado, σ fue escogido uniformemente sobre las permutaciones y entonces A fue escogido uniformemente sobre los conjuntos con k elementos. De aquí $P(A \in \mathcal{F}) = \frac{|\mathcal{F}|}{\binom{n}{k}}$, así que

$$\frac{|\mathcal{F}|}{\binom{n}{k}} \leq \frac{k}{n} \quad \text{y} \quad |\mathcal{F}| \leq \frac{k}{n} \binom{n}{k} = \binom{n-1}{k-1}.$$

\square

Capítulo 3

Aplicaciones De La Linealidad De La Esperanza

En este capítulo se utiliza la esperanza y algunos de los resultados sobre ella para cambiar un poco la idea básica que usábamos para asegurar la existencia de objetos matemáticos. En el capítulo anterior, la idea central era que si la probabilidad de que un objeto exista es positiva entonces el objeto existe. Ahora, la idea central es que si una propiedad se verifica en “promedio” m veces, entonces existe un objeto en el que al menos (o a lo más) se verifica m veces. Esta idea facilita mucho el manejo de ciertas propiedades sobre gráficas sobre el que no se tiene mucho control con gráficas concretas. Es por eso que todos los resultados en este capítulo excepto dos son sobre teoría de gráficas.

3.1. Caminos hamiltonianos en gráficas dirigidas

3.1.1 Definición. Un camino hamiltoniano en una gráfica dirigida $G = (V, E)$ de n vértices (véase 2.2.6) es una permutación $\sigma : \{1, 2, \dots, n\} \rightarrow V$ de los vértices tal que para toda $1 \leq i < n$ se tiene que $(\sigma(i), \sigma(i+1)) \in E$.

El problema de encontrar un camino hamiltoniano en una gráfica dirigida dada es un problema NP-completo.

El siguiente teorema nos garantiza que al menos la cantidad de caminos hamiltonianos posibles en un torneo crece considerablemente con respecto al número de vértices.

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

Demostración. Sea $\sigma \in S_n$ una permutación. Decimos que σ define un camino hamiltoniano si para toda $1 \leq i < n$ se tiene que $(\sigma(i), \sigma(i+1)) \in E$. Sea X_σ la variable aleatoria que toma el valor 1 si σ define un camino hamiltoniano en T y toma el valor 0 de no ser así. Entonces tenemos que la variable aleatoria X definida

por $X = \sum_{\sigma \in S_n} X_\sigma$ nos indica precisamente cuantos caminos hamiltonianos fueron definidos por las permutaciones. Calculemos la esperanza de X .

$$E(X) = E\left(\sum_{\sigma \in S_n} X_\sigma\right) = \sum_{\sigma \in S_n} E(X_\sigma) = \sum_{\sigma \in S_n} \frac{1}{2^{n-1}} = \frac{n!}{2^{n-1}}.$$

Esto nos dice que existe una manera de dirigir a T de tal manera que existen al menos $\frac{n!}{2^{n-1}}$ caminos hamiltonianos. □

Intentar construir torneos con muchos caminos hamiltonianos es particularmente difícil. Por ejemplo, sea $G = (V, E)$ una gráfica completa con 10 vértices y sean $x_1, x_2, x_3 \in V$. Si (x_1, x_3) y (x_2, x_3) son aristas de nuestro torneo, ninguna de las ternas (x_1, x_3, x_2) , (x_2, x_3, x_1) podrá formar parte de un camino hamiltoniano; eso nos deja sin $2(7!)8 = 80640$ posibles caminos. Cada vez que pase algo así aumentará este número de caminos no realizables y aumentará muy rápido. Sin embargo, el teorema nos dice que existe un torneo con al menos $\frac{10!}{2^9} > 7087$ caminos hamiltonianos.

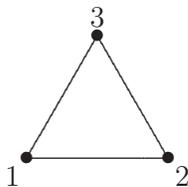
Tenemos que $n!$ crece mucho más rápido que 2^{n-1} ; esto, en vista del resultado anterior, nos dice que la cantidad de caminos hamiltonianos posibles en un torneo de n vértices aumenta considerablemente rápido según el tamaño del torneo.

3.2. Subgráficas bipartitas

3.2.1 Definición. Decimos que una gráfica $G = (V, E)$ es **bipartita** si es posible partir a V en dos subconjuntos de tal manera que ninguno de ellos contenga aristas, es decir, si existen $A, B \in V$ tales que $A \cup B = V$, $A \cap B = \emptyset$, y dada $\{x, y\} \in E$ arista, se tiene que $x \in A$ y $y \in B$, o que $y \in A$ y $x \in B$. Llamamos **partición** a $\{A, B\}$.

Ejemplos:

- La gráfica $G = (\{1, 2\}, \{\{1, 2\}\})$ es bipartita; basta definir $A = \{1\}$ y $B = \{2\}$, y así $\{1, 2\} \notin A$ y $\{1, 2\} \notin B$.
- La gráfica $G = (\{1, 2, 3\}, \{\{1, 2\}, \{2, 3\}, \{1, 3\}\})$ no es bipartita, pues en cualquier manera de partir a V en dos subconjuntos tenemos que alguno de ellos tiene al menos dos vértices, y en esta gráfica cualesquiera dos vértices forman una arista.



- Con el mismo argumento que en el ejemplo anterior podemos decir que si $n \geq 3$, entonces K_n no es bipartita.

Las gráficas bipartitas son importantes en matemáticas. Se aplican en problemas de apareamiento, e incluso en decodificación. En una gráfica bipartita $G = (V, E)$ donde $A, B \subset V$ forman una partición, cualquier subconjunto de A o B un conjunto independiente. Además si no hay vértices aislados (que no formen parte de una arista) cualquiera de A o B forman un conjunto dominante. Este último hecho se usó en el capítulo anterior para demostrar que en una gráfica sin vértices aislados existen conjuntos dominantes de tamaño al menos $\frac{|V|}{2}$.

Se tiene como resultado que en una gráfica cualquiera se puede encontrar una subgráfica bipartita de buen tamaño.

3.2.2 Teorema. *Sea $G = (V, E)$ una gráfica con a aristas. Entonces existe una subgráfica bipartita de G con al menos $\frac{a}{2}$ aristas.*

Demostración. Sea $A \subset V$ un subconjunto de vértices arbitrario dado por $P(x \in A) = \frac{1}{2}$, y sea $B = V \setminus A$. Digamos que una arista $\{x, y\} \in E$ es dividida por A y B si se tiene que $x \in A$ y $y \in B$ o que $y \in A$ y $x \in B$. Notemos que la probabilidad de que una arista dada $\{x, y\} \in E$ sea dividida por A y B es $\frac{1}{2}$ pues

$$\begin{aligned} P((x \in A \text{ y } y \in B) \vee (y \in A \text{ y } x \in B)) &= \\ P(x \in A \text{ y } y \in B) + P(y \in A \text{ y } x \in B) &= \\ P(x \in A) \cdot P(y \in B) + P(y \in A) \cdot P(x \in B) &= \\ P(x \in A) \cdot P(y \notin A) + P(y \in A) \cdot P(x \notin A) &= \\ \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}. \end{aligned}$$

Para una arista $\{x, y\} \in E$ definamos la variable aleatoria $X_{\{x,y\}}$ que toma el valor 1 si $\{x, y\}$ es dividida por A y B , y, que toma el valor 0 si no es el caso. Entonces la variable aleatoria X definida como $X = \sum_{\{x,y\} \in E} X_{\{x,y\}}$ nos indica precisamente cuantas aristas fueron divididas por A y B . Calculemos la esperanza de X .

$$\begin{aligned} E(X) = E\left(\sum_{\{x,y\} \in E} X_{\{x,y\}}\right) &= \sum_{\{x,y\} \in E} E(X_{\{x,y\}}) = \\ \sum_{\{x,y\} \in E} \frac{1}{2} &= \frac{a}{2}. \end{aligned}$$

Esto nos indica que existe una selección de A y B en la que al menos $\frac{a}{2}$ quedan divididas: basta tomar la subgráfica con todos los vértices de G pero sólo las $\frac{a}{2}$ aristas (al menos) que este resultado nos garantiza. □

El resultado anterior se puede mejorar un poco si procedemos de una manera distinta para elegir la partición de V y tomamos en cuenta la paridad de $|V|$. Además de poder garantizar un poco más de aristas, podemos garantizar que en la subgráfica bipartita los subconjuntos de vértices que forman la partición son casi del mismo tamaño.

3.2.3 Teorema. Sea $G = (V, E)$ una gráfica,

- (i) Si G tiene $2n$ vértices y a aristas, entonces contiene una subgráfica bipartita con al menos $\frac{an}{2n-1}$ aristas.
- (ii) Si G tiene $2n+1$ vértices y a aristas, entonces contiene una subgráfica bipartita con al menos $\frac{a(n+1)}{2n+1}$ aristas.

Demostración. (i) Sea $A \subset V$ un subconjunto de vértices con n elementos elegido aleatoriamente y sea $B = V \setminus A$. Nótese que, a diferencia de la prueba del teorema anterior, esta vez estamos eligiendo aleatoriamente sobre subconjuntos de V y no sobre sus elementos. Ahora, dada una arista $\{x, y\} \in E$ la probabilidad de que sea dividida por A y B es la probabilidad de que A tenga a x pero no a y o que A tenga a y pero no a x , para calcular esto procedamos de la manera siguiente: Queremos encontrar las maneras posibles de elegir a A de tal manera que tenga a x pero no a y (o al contrario). Entonces apartemos a x y y de V ; de los vértices restantes escojamos $n-1$ para nuestro subconjunto A y para completarlo le adherimos x o y ; de esta manera conseguimos a todos los subconjuntos A con n elementos que tienen a x o y pero no a ambos. Las maneras de hacer esto son $2 \cdot \binom{2n-2}{n-1}$. Entonces la probabilidad de que A tenga a x pero no a y y viceversa es

$$2 \cdot \frac{\binom{2n-2}{n-1}}{\binom{2n}{n}} = 2 \cdot \frac{\frac{(2n-2)!}{(n-1)!(n-1)!}}{\frac{(2n)!}{(n)!(n)!}} = 2 \cdot \frac{n \cdot n}{(2n)(2n-1)} = \frac{n}{2n-1}.$$

Ahora, para una arista dada $\{x, y\} \in E$ definimos la variable aleatoria $X_{\{x,y\}}$ que toma el valor 1 si $\{x, y\}$ es dividida por A y B y el valor 0 si no es así. Entonces la variable aleatoria X definida como $X = \sum_{\{x,y\} \in E} X_{\{x,y\}}$ nos indica cuantas aristas fueron separadas por A y B . Veamos cuál es la esperanza de la cantidad de aristas divididas por A y B :

$$E(X) = E\left(\sum_{\{x,y\} \in E} X_{\{x,y\}}\right) = \sum_{\{x,y\} \in E} E(X_{\{x,y\}}) = \sum_{\{x,y\} \in E} \frac{n}{2n-1} = \frac{an}{2n-1}.$$

Esto nos dice que existe una manera de escoger a A y B de tal manera que al menos queden $\frac{an}{2n-1}$ aristas divididas por A y B .

(ii) Otra vez sea $A \subset V$ un subconjunto de vértices con n elementos elegido aleatoriamente y sea $B = V \setminus A$. Para calcular la probabilidad de que una arista $\{x, y\} \in E$ sea dividida por A y B procedemos de manera análoga a la demostración del inciso anterior: Quitamos a x y a y y de los vértices restantes escogemos $n-1$ para A y para completarlo le adherimos x o y . Esto lo podemos hacer de $2 \cdot \binom{2n-1}{n-1}$; de esta manera conseguimos a todos los subconjuntos A de n elementos que tienen a x o y pero no a ambos. Entonces la probabilidad de que $\{x, y\}$ sea separada por A y B es

$$2 \cdot \frac{\binom{2n-1}{n-1}}{\binom{2n+1}{n}} = 2 \cdot \frac{\frac{(2n-1)!}{(n-1)!(n)!}}{\frac{(2n+1)!}{(n)!(n+1)!}} = 2 \cdot \frac{(n+1)(n)}{(2n+1)(2n)} = \frac{(n+1)}{(2n+1)}.$$

Ahora para cada arista $\{x, y\} \in E$ definimos la variable aleatoria $X_{\{x,y\}}$ que toma el valor 1 si $\{x, y\}$ es dividida por A y B y el valor 0 si no es así, y definimos la variable aleatoria $X = \sum_{\{x,y\} \in E} X_{\{x,y\}}$. Definida de esta manera X nos dice precisamente cuantas aristas fueron divididas por A y B . Calculemos la esperanza de la cantidad de aristas divididas por A y B .

$$E(X) = E\left(\sum_{\{x,y\} \in E} X_{\{x,y\}}\right) = \sum_{\{x,y\} \in E} E(X_{\{x,y\}}) = \sum_{\{x,y\} \in E} \frac{(n+1)}{(2n+1)} = \frac{a(n+1)}{(2n+1)}.$$

Entonces es posible encontrar una elección de A y B que deje al menos $\frac{a(n+1)}{(2n+1)}$ aristas divididas.

□

3.3. Más resultados sobre coloraciones

3.3.1 Teorema. *Existe una 2-coloración de K_n con a lo más $\binom{n}{a} \cdot \frac{2}{2^{\binom{a}{2}}}$ subgráficas completas con a vértices monocromáticas.*

Demostración. Demos a K_n una 2-coloración arbitraria. Para cada subgráfica $K_a \subset K_n$ completa con a vértices definamos la variable aleatoria X_{K_a} que toma el valor 1 si dicha K_a es monocromática y el valor 0 si no es así. La probabilidad de que una K_a sea monocromática es $\frac{2}{2^{\binom{a}{2}}}$. Ahora definamos la variable aleatoria $X = \sum_{K_a \subset K_n} X_{K_a}$. Esta nos dice cuantas subgráficas completas con a vértices resultaron ser monocromáticas. Calculemos la esperanza de X :

$$E(X) = E\left(\sum_{K_a \subset K_n} X_{K_a}\right) = \sum_{K_a \subset K_n} E(X_{K_a}) = \sum_{K_a \subset K_n} \frac{2}{2^{\binom{a}{2}}} = \binom{n}{a} \frac{2}{2^{\binom{a}{2}}}.$$

Esto nos asegura que podemos dar una coloración que deje a lo más $\binom{n}{a} \frac{2}{2^{\binom{a}{2}}}$ subgráficas completas con a vértices monocromáticas.

□

Por ejemplo, sean $n = 10$ y $a = 5$. El teorema dice que podemos dar coloraciones donde a lo más existan $\binom{10}{5} \cdot \frac{2}{2^{\binom{5}{2}}} = 252 \frac{2}{1024} = \frac{252}{512} = 0.4921875$ subgráficas completas en 5 vértices monocromáticas. En particular, esto dice que $R(5, 5) > 10$.

3.3.2 Definición. Una **gráfica bipartita completa** es una gráfica bipartita con partición $\{A, B\}$ tal que para todo $a \in A$ y todo $b \in B$, $\{a, b\}$ es una arista. Si $|A| = m$ y $|B| = n$ denotamos a esta gráfica por $\mathbf{K}_{m,n}$.

Procediendo de la misma manera que en la demostración del teorema anterior, obtenemos el siguiente.

3.3.3 Teorema. *Existe una 2-coloración de $K_{m,n}$ con a lo más $\binom{m}{a} \binom{n}{b} \frac{2}{2^{ab}}$ $K_{a,b}$ monocromáticas.*

Demostración. Demos a $K_{m,n}$ una 2-coloración arbitraria. Para cada $K_{a,b} \subset K_{m,n}$ definamos la variable aleatoria $X_{K_{a,b}}$ que toma el valor 1 si dicha $K_{a,b}$ es monocromática y el valor 0 si no es así. La probabilidad de que una $K_{a,b}$ sea monocromática es $\frac{2}{2^{ab}}$. Ahora definimos la variable aleatoria $X = \sum_{K_{a,b} \subset K_{m,n}} X_{K_{a,b}}$, la cual nos dice cuantas $K_{a,b}$ resultaron ser monocromáticas. Calculemos la esperanza de X .

$$E(X) = E\left(\sum_{K_{a,b} \subset K_{m,n}} X_{K_{a,b}}\right) = \sum_{K_{a,b} \subset K_{m,n}} E(X_{a,b}) = \sum_{K_{a,b} \subset K_{m,n}} \frac{2}{2^{ab}} = \binom{m}{a} \binom{n}{b} \frac{2}{2^{ab}},$$

con lo que aseguramos que podemos encontrar una coloración que deja a lo más $\binom{m}{a} \binom{n}{b} \frac{2}{2^{ab}}$ $K_{a,b}$ monocromáticas. □

El siguiente resultado es interesante porque, a pesar de que no parece tener relación alguna con probabilidad, su demostración probabilística es sencilla. Aquí $\|\cdot\|$ denota la norma euclidiana.

3.4. Geometría analítica

3.4.1 Teorema. *Sean $v_1, \dots, v_n \in \mathbb{R}^n$, tales que $\|v_i\| = 1$ para todo $1 \leq i \leq n$. Entonces existen $\lambda_1, \dots, \lambda_n = \pm 1$ tales que $\|\lambda_1 v_1 + \dots + \lambda_n v_n\| \leq \sqrt{n}$, y también existen $\lambda_1, \dots, \lambda_n = \pm 1$ tales que $\|\lambda_1 v_1 + \dots + \lambda_n v_n\| \geq \sqrt{n}$.*

Demostración. Sean $\lambda_1, \dots, \lambda_n$ seleccionadas aleatoriamente en $\{-1, 1\}$. Definimos la variable aleatoria $X = \|\lambda_1 v_1 + \dots + \lambda_n v_n\|^2$. Tenemos que

$$\begin{aligned} \|\lambda_1 v_1 + \dots + \lambda_n v_n\|^2 &= (\lambda_1 v_1 + \dots + \lambda_n v_n) \cdot (\lambda_1 v_1 + \dots + \lambda_n v_n) = \\ &= \sum_{i=1}^n \sum_{j=1}^n (\lambda_i v_i) \cdot (\lambda_j v_j) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (v_i \cdot v_j). \end{aligned}$$

Entonces la esperanza de X es

$$\begin{aligned} E(X) &= E\left(\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j (v_i \cdot v_j)\right) = \sum_{i=1}^n \sum_{j=1}^n E(\lambda_i \lambda_j (v_i \cdot v_j)) = \\ &= \sum_{i=1}^n \sum_{j=1}^n E(\lambda_i \lambda_j) (v_i \cdot v_j). \end{aligned}$$

Pero tenemos que si $i \neq j$ entonces $E(\lambda_i \lambda_j) = E(\lambda_i)E(\lambda_j) = (\frac{1}{2} - \frac{1}{2})(\frac{1}{2} - \frac{1}{2}) = 0$ y también tenemos que si $i = j$ entonces $\lambda_i \lambda_j = \lambda_i^2 = 1$ y, por lo tanto, $E(\lambda_i^2) = 1$. Entonces la esperanza de X se reduce a

$$E(X) = \sum_{i=1}^n v_i \cdot v_i = n.$$

Ahora que tenemos que la esperanza del cuadrado de la norma es n , tenemos que la esperanza de la norma es \sqrt{n} y, por lo tanto, existen $\lambda_1, \dots, \lambda_n = \pm 1$ tales que $\|\lambda_1 v_1 + \dots + \lambda_n v_n\| \leq \sqrt{n}$, y también existen $\lambda_1, \dots, \lambda_n = \pm 1$ tales que $\|\lambda_1 v_1 + \dots + \lambda_n v_n\| \geq \sqrt{n}$. \square

Si $v_1, \dots, v_n \in \mathbb{R}^n$, nótese que $v = b_1 v_1 + \dots + b_n v_n$, con $b_1, \dots, b_n \in \{0, 1\}$, representa una esquina del paralelepípedo de dimensión n que resulta de tomar v_1, \dots, v_n como algunos de sus lados, y el resto obtenerlos por traslaciones de éstos. También nótese que $w = p_1 v_1 + \dots + p_n v_n$, con $p_1, \dots, p_n \in [0, 1]$, es un punto interior de dicho paralelepípedo. El siguiente resultado se puede interpretar como sigue: En un paralelepípedo de dimensión n con lados de longitud menor o igual a 1 cualquier punto dista a lo más $\frac{\sqrt{n}}{2}$ de una esquina. El caso $p_1 = p_2 = \dots = p_n = \frac{1}{2}$ y $\|v_1\| = \|v_2\| = \dots = \|v_n\| = 1$ no es más que el resultado anterior pero trasladando el origen a $\frac{1}{2}v_1 + \frac{1}{2}v_2 + \dots + \frac{1}{2}v_n$ y con las normas iguales a $\frac{1}{2}$.

3.4.2 Teorema. Sean $v_1, \dots, v_n \in \mathbb{R}^n$ tales que $\|v_i\| \leq 1$, $p_1, \dots, p_n \in [0, 1]$, y $w = p_1 v_1 + \dots + p_n v_n$. Entonces existen $b_1, \dots, b_n \in \{0, 1\}$ tales que $\|w - v\| \leq \frac{\sqrt{n}}{2}$, donde $v = b_1 v_1 + \dots + b_n v_n$.

Demostración. Escojamos los b_i con las siguientes probabilidades: $P(b_i = 1) = p_i$. Definimos la variable aleatoria $X(v)$ como $X(v) = \|w - v\|^2$. Expandiendo obtenemos:

$$\|w - v\|^2 = \left\| \sum_{i=1}^n (p_i - b_i) v_i \right\|^2 = \sum_{i=1}^n \sum_{j=1}^n v_i \cdot v_j (p_i - b_i)(p_j - b_j).$$

Entonces, la esperanza de X es:

$$E(X) = \sum_{i=1}^n \sum_{j=1}^n v_i \cdot v_j E[(p_i - b_i)(p_j - b_j)].$$

Ahora, para $i \neq j$, por ser p_i y p_j independientes, tenemos que $E[(p_i - b_i)(p_j - b_j)] = E(p_i - b_i)E(p_j - b_j)$. Pero $E(p_i - b_i) = p_i(p_i - 1) + (1 - p_i)p_i = 0$, de donde $E[(p_i - b_i)(p_j - b_j)] = 0$. Si $i = j$ entonces $E[(p_i - b_i)(p_j - b_j)] = E[(p_i - b_i)^2] = p_i(p_i - 1)^2 + (1 - p_i)p_i^2 = p_i((p_i - 1)^2 + (1 - p_i)p_i) = p_i(p_i^2 - 2p_i + 1 + p_i - p_i^2) = p_i(1 - p_i)$. Veamos que $p_i(1 - p_i)$ está acotado por $\frac{1}{4}$. Derivando $x(1 - x) = x - x^2$ queda $1 - 2x$ el cual es igual a cero cuando $x = 1/2$. La segunda derivada de $x(1 - x)$ es -1 ,

por lo tanto $x(1-x)$ alcanza su máximo en $x = \frac{1}{2}$ y es $\frac{1}{2}(1 - \frac{1}{2}) = \frac{1}{4}$. Entonces $E[(p_i - b_i)^2] = p_i(1 - p_i) \leq \frac{1}{4}$, de donde

$$E(X) = \sum_{i=1}^n v_i \cdot v_i E[(p_i - b_i)^2] = \sum_{i=1}^n \|v_i\|^2 E[(p_i - b_i)^2] \leq \frac{1}{4} \sum_{i=1}^n \|v_i\|^2 \leq \frac{n}{4}.$$

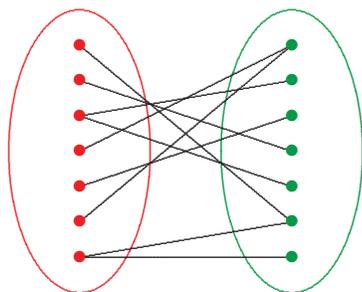
Por lo tanto existen $b_1, \dots, b_n \in \{0, 1\}$ tales que $\|w - v\|^2 \leq \frac{n}{4}$, lo cual implica que $\|w - v\| \leq \frac{\sqrt{n}}{2}$. \square

3.5. Alto cuello y alto número cromático

Una **coloración propia** (en vértices) es una asignación de colores a los vértices de una gráfica de manera que vértices adyacentes siempre tengan color distinto. Claramente, si G es una gráfica con n vértices, al colorear con n colores (un color distinto para cada vértice) obtendremos una coloración propia. Entonces hablar de muchos colores no tiene sentido, siempre con n o más colores se puede garantizar trivialmente una coloración propia. El problema interesante es cuál es la cantidad mínima de colores necesaria para garantizar la existencia de una coloración propia.

Pensemos en un ejemplo. Cualquier gráfica con al menos una arista necesitará al menos dos colores para tener una coloración propia, pues, al menos los dos vértices unidos por la arista necesitan color distinto uno del otro.

En una gráfica bipartita no trivial dos colores siempre son suficientes para dar una coloración propia. Para ver esto tomemos una partición de los vértices en dos conjuntos tales que en cada conjunto no hay vértices adyacentes. A los vértices de un conjunto los pintamos de rojo y a los del otro de azul. Como dentro de cada conjunto no hay adyacencias, no hay vértices adyacentes con el mismo color.



Al mínimo número de colores necesario para dar una coloración propia a una gráfica G , se le llama **número cromático** de G y se denota por $\chi(\mathbf{G})$.

Nótese que en una coloración propia para una gráfica G , el conjunto de vértices determinado por cada color es un conjunto tal que ningún par de vértices sobre ellos

es adyacente. Este tipo de conjuntos de vértices adyacentes es conocido como **conjunto independiente de vértices** de G . Trivialmente tenemos que un solo vértice conforma un conjunto independiente de vértices puesto que no existe otro vértice en el conjunto con el que pudiera ser adyacente. En este caso lo que es interesante es el tamaño del conjunto de vértices independiente más grande que podamos encontrar en G . A éste se le conoce como **número de independencia** de G y se le denota por $\alpha(G)$.

El número cromático y el número de independencia se encuentran relacionados de una manera bastante sencilla.

3.5.1 Lema. *Sea $G = (V, E)$ una gráfica con n vértices. Entonces $\chi(G)\alpha(G) \geq n$.*

Demostración. Tomemos una coloración propia $f : V \rightarrow [\chi(G)]$, es decir, si $x, y \in V$ son tales que $f(x) = f(y)$, entonces $\{x, y\} \notin E$. Así $f^{-1}(i)$, con $i \in [\chi(G)]$, es un subconjunto independiente de vértices. Por lo tanto, $\forall i \in [\chi(G)]$, $\alpha(G) \geq |f^{-1}(i)|$, de donde

$$\chi(G)\alpha(G) = \sum_{i \in [\chi(G)]} \alpha(G) \geq \sum_{i \in [\chi(G)]} |f^{-1}(i)| = n.$$

□

Intuitivamente uno esperaría que el número cromático de una gráfica estuviera relacionado con qué tan enredada se encuentra. Después de todo, gráficas muy conectadas deben tener número cromático grande, como es el caso de K_n , donde forzosamente se necesitan n colores puesto que cualesquiera dos vértices se encuentran conectados.

Una manera de interpretar qué tan enredada se encuentra una gráfica es fijarse en la longitud de su ciclo más pequeño. Entre más pequeño sea éste, más enredada estará la gráfica. A la longitud del ciclo más pequeño de una gráfica G se le conoce como **cuello** de G .

Si tomamos un ciclo de longitud mínima, cada uno de los vértices sólo tendrá dos vecinos entre los vértices del ciclo. Si tuviera alguno más, entonces tendríamos un ciclo más pequeño. Esto nos da una cota para el número de independencia de una gráfica.

3.5.2 Lema. *Sea $G = (V, E)$ una gráfica. Si el cuello c de G es finito, entonces $\alpha(G) \geq \lfloor \frac{c}{2} \rfloor$.*

Demostración. Para demostrar el lema basta encontrar un subconjunto independiente de vértices con $\lfloor \frac{c}{2} \rfloor$ elementos. Sea $A = (x_1, x_2, \dots, x_c)$ un ciclo de longitud mínima en G . Veamos que el subconjunto $B = \{x_1, x_3, \dots, x_{2m-1}, \dots, x_{2\lfloor \frac{c}{2} \rfloor - 1}\} \subset V$ con $1 \leq m \leq \lfloor \frac{c}{2} \rfloor$ es independiente. Supongamos lo contrario y, sin pérdida de generalidad, digamos que x_1 es vecino de x_{2s-1} para algún $1 < s \leq \lfloor \frac{c}{2} \rfloor$. Entonces tenemos

que $(x_1, x_{2s-1}, x_{2s+1}, \dots, x_c)$ es un ciclo de longitud $c - 2s < c$, lo cual contradice la minimalidad de la longitud de A . Por lo tanto B es un conjunto independiente con $\lfloor \frac{c}{2} \rfloor$ vértices. \square

Si el cuello de una gráfica $G = (V, E)$ es ∞ , entonces no tiene ciclos, por lo tanto es bipartita y tiene un conjunto independiente con al menos $\frac{|V|}{2}$ vértices. Además, como habíamos dicho antes, si G no es trivial, su número cromático es 2.

Ahora encadenamos desigualdades para poder asegurar algo sobre el número cromático. Tenemos que para una gráfica G con n vértices y con cuello finito c , $\alpha(G) \geq \lfloor \frac{c}{2} \rfloor$, que es lo mismo que $1 \geq \frac{\lfloor \frac{c}{2} \rfloor}{\alpha(G)}$. Por otro lado teníamos que $\chi(G)\alpha(G) \geq n$, y, por lo tanto, $\chi(G) \geq \frac{n}{\alpha(G)}$. Relacionando esto tenemos que $\chi(G) \geq \frac{n \lfloor \frac{c}{2} \rfloor}{(\alpha(G))^2}$.

Esto ya no va tan de acuerdo a nuestra intuición. Dice que si dejamos fijo el número de independencia, entonces el número cromático crece según el cuello crece. Pero, después de todo, el número de independencia sí está muy condicionado al cuello.

Por mucho tiempo se conjeturó que cuello grande implicaba número cromático pequeño. Sin embargo, la conjetura era falsa. Con suficientes vértices se pueden tener gráficas con número cromático tan grandes como se quiera y cuellos tan grandes como se quiera. El siguiente teorema fue demostrado inicialmente por Erdős usando métodos probabilísticos y fue esencial para que los matemáticos se dieran cuenta de la utilidad de estos.

3.5.3 Teorema. *Dados $s, t \in \mathbb{N}$ existe una gráfica G con cuello c , tal que $\chi(G) > s$ y $c > t$.*

Demostración. Sean $s, t, n \in \mathbb{N}$, $\varepsilon = \frac{1}{2t}$, $p = n^{\varepsilon-1}$, y sea $G = (n, p)$ una gráfica aleatoria. Sea X la variable aleatoria que cuenta los ciclos de G con longitud menor que t . Primero acotamos el número de ciclos con longitud a lo más t :

$$\frac{1}{2} \binom{n}{k} (k-1)! = \frac{1}{2} \frac{n!}{k!(n-k)!} (k-1)! = \frac{1}{2} \frac{n(n-1) \cdots (n-k+1)}{k} \leq n^k.$$

Ahora acotamos la esperanza de X :

$$E[X] = \sum_{k=3}^t \frac{1}{2} \binom{n}{k} (k-1)! p^k \leq \sum_{k=3}^t n^k p^k = \sum_{k=3}^t n^{\varepsilon k} = \sum_{k=3}^t n^{\frac{k}{2t}}.$$

Esto último es sublineal pues el valor más alto que alcanza $\frac{k}{2t}$ es $\frac{1}{2}$. Entonces tenemos que $E[X] \in o(n)$. Por lo tanto existe un n_0 tal que si $n > n_0$ entonces $E[X] < \frac{n}{4}$. Por desigualdad de Markov, y para $n > n_0$ (1.1.8), tenemos

$$P(X \geq \frac{n}{2}) \leq \frac{E[X]}{\frac{n}{2}} \leq \frac{\frac{n}{4}}{\frac{n}{2}} = \frac{1}{2}.$$

Como trabajar sobre el número cromático es muy complicado, trabajaremos sobre el número de independencia, después de todo, el número cromático y el número de independencia están relacionados por 3.5.1.

Sea $a \in \mathbb{N}$, acotemos la probabilidad de que una gráfica aleatoria $G = (n, p)$ tenga un número de independencia mayor o igual que a . Si una gráfica tiene número de independencia mayor o igual a a , en particular debe tener un subconjunto independiente de vértices de tamaño a . Por lo tanto,

$$P(\alpha(G) \geq a) \leq \binom{n}{a} (1-p)^{\binom{a}{2}} \leq n^a e^{-p \binom{a}{2}} = e^{(\ln(n) - p(a-1)/2)a}.$$

La desigualdad es cierta en vista de 2.2.14. Si elegimos $a = \lceil \frac{3}{p} \ln(n) \rceil$ tenemos que

$$\begin{aligned} \ln(n) - p(a-1) &= \ln(n) - p \left(\left\lceil \frac{3}{p} \ln(n) \right\rceil - 1 \right) \leq \\ \ln(n) - p \left(\frac{3}{p} \ln(n) - 1 \right) &= \ln(n) - 3 \ln(n) - p \rightarrow -\infty, \end{aligned}$$

y, por lo tanto, $P(\alpha(G) \geq a) \rightarrow 0$.

De aquí tenemos que existe n'_0 tal que, para $n > n'_0$, $P(\alpha(G) \geq a) < \frac{1}{2}$. Por lo tanto, para $n > \max(n_0, n'_0)$, tenemos:

$$P(\alpha(G) \geq a) + P(X \geq \frac{n}{2}) < 1.$$

Por lo tanto existe una gráfica G con n vértices con número de independencia menor que a y con menos de $n/2$ ciclos con longitud a lo más t . Si tomamos una de estas gráficas y le quitamos un vértice de cada ciclo de longitud a lo más t , obtendremos una gráfica sin ciclos de longitud menor o igual a t y no vacía, pues a lo más quitamos $n/2$ vértices. A esta nueva gráfica le llamaremos G^* .

El lema anterior nos asegura que el número de independencia no crece al borrar vértices. Ahora sólo falta medir el número cromático. Como G^* tiene número de independencia menor que a , su número de independencia es a lo más $a-1$. Recordando 3.5.1 y haciendo las sustituciones necesarias llegamos a que

$$\chi(G^*) \geq \frac{n/2}{a-1} \geq \frac{pn}{6 \ln(n)} = \frac{n^\epsilon}{6 \ln(n)} \rightarrow \infty.$$

Por lo tanto basta tomar n suficientemente grande para garantizar que G^* tiene número cromático más grande que s y cuello más grande que t . \square

Capítulo 4

Alteraciones

La idea básica en este capítulo consiste en demostrar probabilísticamente la existencia de un objeto que no cumpla con las hipótesis y luego alterarlo obligándolo a cumplirlas. Regularmente este método se usa para demostrar desigualdades aunque también puede usarse para demostrar la existencia de objetos.

4.1. Más sobre números de Ramsey

4.1.1 Teorema. *Para cualesquiera $n, k \in \mathbb{N}$ se tiene que:*

$$R(k, k) > n - \binom{n}{k} \frac{2}{2^{\binom{k}{2}}}.$$

Demostración. Tomemos la gráfica completa de n vértices K_n y una coloración arbitraria sobre ella eligiendo el color de cada vértice con probabilidad $\frac{1}{2}$. Para cualquier subconjunto R de k vértices definamos la variable aleatoria $X_R : 2^{|V(K_n)|} \rightarrow \{0, 1\}$, que toma el valor 1 si el K_k generado por R es monocromático y el valor 0 si no es así (Nótese que $2^{|V(K_n)|}$ denota al conjunto de coloraciones de los vértices de K_n). Definimos $X = \sum_R X_R$. Definida de esta manera X nos indica cuantas subgráficas completas de k vértices resultaron ser monocromáticas. La esperanza de X es:

$$E[X] = \sum_R E[X_R] = \sum_R \frac{2}{2^{\binom{k}{2}}} = \binom{n}{k} \frac{2}{2^{\binom{k}{2}}}.$$

Esto quiere decir que existe una coloración de K_n que deja a lo más $\binom{n}{k} \frac{2}{2^{\binom{k}{2}}}$ subgráficas completas de k vértices monocromáticas. Usando esta coloración, si quitamos un vértice de cada K_k que quedó monocromática, habremos construido una gráfica con al menos $n - \binom{n}{k} \frac{2}{2^{\binom{k}{2}}}$ vértices que no contiene alguna K_k monocromática y, por lo tanto, $R(k, k) > n - \binom{n}{k} \frac{2}{2^{\binom{k}{2}}}$. \square

Este teorema nos da una segunda cota para $R(k, k)$ (la primera fue el resultado de 2.2.5).

El problema de encontrar una subgráfica completa de k vértices con todas sus aristas de un mismo color en una gráfica completa con n vértices cuyas aristas están coloreadas con dos colores es un problema NP-completo. A pesar de esto, el teorema anterior y 2.2.5 nos dan cotas fáciles de calcular que nos evitarían hacer bastantes intentos inútiles. También el problema de determinar si la gráfica tiene una subgráfica completa de k vértices con todas sus aristas del primer color o una subgráfica completa de l vértices del segundo color es un problema NP-completo y los siguientes dos teoremas nos ayudan a no hacer intentos vanos para calcular $R(k, l)$.

4.1.2 Teorema. *Dado $n \in \mathbb{N}$, si existe $p \in [0, 1]$ tal que*

$$\binom{n}{k} p^{\binom{k}{2}} + \binom{n}{l} (1-p)^{\binom{l}{2}} < 1,$$

entonces $R(k, l) > n$.

Demostración. Supongamos que existe un p como dice la hipótesis. Queremos ver que existe una coloración para K_n con la cual no hay ningún K_k rojo ni ningún K_l azul. Tomemos la gráfica completa de n vértices K_n y, coloreemos de rojo a cada arista con probabilidad p y, si la arista no resultó ser roja, entonces la coloreamos de azul. Dado un subconjunto A de k vértices llamemos S_A al suceso en que éste genera una subgráfica K_k con todas sus aristas rojas; tenemos que $P(S_A) = p^{\binom{k}{2}}$. Dado un subconjunto B de l vértices llamemos S_B al evento en que éste genera una subgráfica K_l con todas sus aristas azules. Tenemos que $P(S_B) = (1-p)^{\binom{l}{2}}$. Entonces la probabilidad de que K_n contenga algún K_k con todas sus aristas rojas es la probabilidad de $\bigvee S_A$, donde $A \subset V$ es tal que $|A| = k$, y ésta cumple que

$$P\left(\bigvee_{\substack{A \subset V \\ |A|=k}} S_A\right) \leq \sum_{\substack{A \subset V \\ |A|=k}} P(S_A) = \binom{n}{k} p^{\binom{k}{2}}.$$

También tenemos que la probabilidad de que K_n contenga a algún K_l con todas sus aristas azules es la probabilidad de $\bigvee S_B$, donde $B \subset V$ es tal que $|B| = l$, y esta cumple que

$$P\left(\bigvee_{\substack{B \subset V \\ |B|=l}} S_B\right) \leq \sum_{\substack{B \subset V \\ |B|=l}} P(S_B) = \binom{n}{l} (1-p)^{\binom{l}{2}}.$$

Por lo tanto la probabilidad de que K_n contenga a algún K_k con todas sus aristas rojas o algún K_l con todas sus aristas azules cumple que

$$P\left(\bigvee_{\substack{A \subset V \\ |A|=k}} S_A \vee \bigvee_{\substack{B \subset V \\ |B|=l}} S_B\right) \leq P\left(\bigvee_{\substack{A \subset V \\ |A|=k}} S_A\right) + P\left(\bigvee_{\substack{B \subset V \\ |B|=l}} S_B\right) \leq \binom{n}{k} p^{\binom{k}{2}} + \binom{n}{l} (1-p)^{\binom{l}{2}} < 1,$$

lo cual, significa que existe una 2-coloración que no deja algún K_k con todas sus aristas rojas o algún K_l con todas sus aristas azules como subgráfica y, por lo tanto, $R(k, l) > n$. \square

Con un razonamiento análogo obtenemos el siguiente resultado.

4.1.3 Teorema. *Para todo $n \in \mathbb{N}$ y todo $p \in [0, 1]$, se tiene que:*

$$R(k, l) > n - \binom{n}{k} p^{\binom{k}{2}} - \binom{n}{l} (1-p)^{\binom{l}{2}}$$

Demostración. Tomamos la gráfica completa de n vértices K_n y coloreamos sus aristas de rojo y azul, siendo cada una roja con probabilidad p . Sea $X_A : \{\text{Coloraciones de } K_n\} \rightarrow \mathbb{Z}$ la variable aleatoria que cuenta el número de subgráficas K_k con todas sus aristas rojas y sea $X_B : \{\text{Coloraciones de } K_n\} \rightarrow \mathbb{Z}$ la variable aleatoria que cuenta el número de subgráficas K_l con todas sus aristas azules. Definamos $X = X_A + X_B$. Midiendo la esperanza de X tenemos que

$$E[X] = E[X_A] + E[X_B] = \binom{n}{k} p^{\binom{k}{2}} + \binom{n}{l} (1-p)^{\binom{l}{2}}.$$

Llamemos s a este número; entonces existe una 2-coloración que deja a lo más s subgráficas “malas”. Si quitamos un vértice de cada una de ellas y las aristas que están unidas a éstos, habremos construido una gráfica con al menos $n - s$ vértices que no contiene algún K_k con todas sus aristas rojas o algún K_l con todas sus aristas azules como subgráfica y, por lo tanto, $R(k, l) > n - s$. \square

Notemos que encontrar un conjunto independiente es análogo a encontrar una subgráfica completa de aristas “faltantes”; dicho problema, como ya habíamos mencionado antes, es un problema que pertenece a la clase NP-completo. Sin embargo el teorema siguiente nos dice que podemos encontrar conjuntos independientes relativamente grandes sin necesidad de examinar muchos detalles de la gráfica.

4.1.4 Teorema. *Sea $G = (V, E)$ una gráfica con n vértices y $\frac{nd}{2}$ aristas, con $d \geq 1$. Entonces $\alpha(G) \geq \frac{n}{2d}$.*

Demostración. Sea $U \subset V$ elegido aleatoriamente con probabilidad p , es decir $P(v \in U) = p$ (p será determinado más tarde). Definamos $X : \{U : U \subset V\} \rightarrow \mathbb{Z}$ como $X(U) = |U|$. Para cada $e \in E$, sea $Y_e : \{U : U \subset V\} \rightarrow \{0, 1\}$ definida como $Y_e(U) = 1$ si $e \subset U$ y $Y_e(U) = 0$ si no es así. Por último sea $Y = \sum_{e \in E} Y_e(U)$.

Dado un $e = \{x, y\} \in E$ tenemos que $E(Y_e) = P(x, y \in U) = p^2$ y, por linealidad de la esperanza,

$$E(Y) = \sum_{e \in E} E(Y_e) = \frac{nd}{2} p^2.$$

Por otra parte tenemos que $E(X) = np$ y, otra vez por linealidad de la esperanza, tenemos que

$$E(X - Y) = np - \frac{nd}{2} p^2.$$

Derivando esto último respecto a p tenemos $n - ndp$, el cual se anula para $p = \frac{1}{d}$. La segunda derivada es $-nd < 0$ por lo tanto es un máximo. Elijiendo $p = \frac{1}{d}$ tenemos

que $E(X) = \frac{n}{2d}$. Por lo tanto existe algún U para el cual su número de vértices menos el número de aristas contenidas en U es al menos $\frac{n}{2d}$. Para cada arista de U quitamos uno de los vértices que la componen; de esta manera obtenemos un conjunto U^* independiente con al menos $\frac{n}{2d}$ vértices y con esto termina la demostración. \square

Capítulo 5

El método del segundo momento

En este capítulo se harán demostraciones utilizando implicaciones del lema de Chebyshev (1.1.10), el cual enuncia un resultado de la varianza. La varianza es el segundo momento de una variable aleatoria (el primero es la esperanza). Es por eso que a estas demostraciones se les llama el método del segundo momento.

Un caso particular del teorema de Turán nos dice que si n es el número de vértices de una gráfica G , entonces es suficiente que el número de aristas de G sea $\frac{n^2}{4}$ para que G tenga un triángulo.

Una buena pregunta es cuál es la probabilidad de que una gráfica tenga un triángulo. Es natural pensar que en una gráfica donde existan muchas aristas sea muy probable encontrar un triángulo. De igual manera se podría pensar que pocas aristas existentes nos producen gráficas libres de triángulos.

El siguiente teorema va en contra de esta intuición; esencialmente dice que basta que nuestra proporción de aristas existentes entre aristas posibles decrezca menos rápido que $\frac{1}{n}$ para que exista un triángulo con probabilidad tan cercana a 1 como queramos (también se suele decir que existe un triángulo **casi siempre**). Lo interesante aquí es que nuestra proporción también puede hacerse muy pequeña, siempre y cuando no decrezca tan rápido como $\frac{1}{n}$. Esto es sorprendente, la contundencia del teorema de Turán exige $\frac{n^2}{4}$ aristas para asegurar un triángulo, ahora podemos apostar por un triángulo con tanta seguridad como queramos con muchas menos aristas. Antes de demostrar el teorema, necesitaremos el siguiente lema.

5.0.5 Lema. *Sea X_1, X_2, X_3, \dots una sucesión de variables aleatorias no negativas tales que*

$$\lim_{n \rightarrow \infty} \frac{\text{Var}[X_n]}{(E[X_n])^2} = 0.$$

Entonces

$$\lim_{n \rightarrow \infty} P(X_n > 0) = 1.$$

Demostración. Elegimos $a = E[X_n]$ en la desigualdad de Chebyshev (1.1.10). Entonces

$$P(|X_n - E[X_n]| \geq E[X_n]) \leq \frac{\text{Var}[X_n]}{(E[X_n])^2}.$$

Notemos que el suceso $|X_n - E[X_n]| \geq E[X_n]$ es la unión de los sucesos ajenos $X_n - E[X_n] \geq E[X_n]$ y $E[X_n] - X_n \geq E[X_n]$. Por lo tanto, la probabilidad del primero es mayor o igual que la probabilidad de cualquiera de los otros dos. Es decir

$$\begin{aligned} \lim_{n \rightarrow \infty} P(X_n \leq 0) &= \lim_{n \rightarrow \infty} P(X_n - E[X] \leq -E[X]) = \lim_{n \rightarrow \infty} P(E[X] - X_n \geq E[X]) \leq \\ &\lim_{n \rightarrow \infty} P(|X_n - E[X_n]| \geq E[X_n]) \leq \lim_{n \rightarrow \infty} \frac{\text{Var}[X_n]}{(E[X_n])^2} = 0. \end{aligned}$$

Por lo tanto

$$\lim_{n \rightarrow \infty} P(X_n > 0) = \lim_{n \rightarrow \infty} 1 - P(X_n \leq 0) = 1 - \lim_{n \rightarrow \infty} P(X_n \leq 0) = 1.$$

□

5.1. Una versión débil del Teorema de Turán

En lo siguiente, $G(n, p)$ denota una gráfica aleatoria en n vértices en la que cada una de sus aristas existe con probabilidad p .

5.1.1 Teorema. Sean $p : \mathbb{N} \rightarrow [0, 1]$ tal que $\frac{1}{n} \in o(p)$ y $\mathcal{G} = \{G_n = G(n, p(n)) : n \in \mathbb{N}\}$ un conjunto de gráficas aleatorias. Entonces

$$\lim_{n \rightarrow \infty} P(G_n \text{ tiene un triángulo}) = 1.$$

Demostración. Para cada $n \in \mathbb{N}$ sea X_n la variable aleatoria que cuenta los triángulos en G_n . Midamos la esperanza de cada X_n .

$$E[X_n] = \sum_{i=1}^{\binom{n}{3}} p(n)^3 = \binom{n}{3} p(n)^3.$$

Ahora acotemos la varianza de cada X_n . Para cada terna de vértices en G_n , definamos T_i como la variable aleatoria que toma el valor 1 si la i -ésima terna forma un triángulo y el valor 0 si no. Notemos que $X_n = \sum_{i=1}^{\binom{n}{3}} T_i$. Recordando 1.1.12 tenemos

$$\text{Var}[X_n] = \sum_{i=1}^{\binom{n}{3}} \binom{n}{3} \text{Var}[T_i] + \sum_{1 \leq i, j \leq \binom{n}{3}, i \neq j} \text{Cov}[T_i, T_j].$$

Acotemos sumando a sumando. Para el primero tenemos

$$\text{Var}[T_i] = E[(T_i)^2] - (E[T_i])^2 = p(n)^3 - p(n)^6 \leq p(n)^3.$$

Entonces

$$\sum_{i=1}^{\binom{n}{3}} \text{Var}[T_i] \leq \binom{n}{3} p(n)^3.$$

Para acotar el siguiente sumando recordemos que la covarianza de variables aleatorias independientes es 0. En este caso, T_i es independiente de T_j cuando los triángulos determinados por las ternas i -ésima y j -ésima no comparten aristas, después de todo, la existencia de una arista no depende de la existencia de otra. Lo mismo ocurre con triángulos ajenos en aristas. Entonces sólo nos interesan los casos donde los triángulos comparten una arista (no pueden compartir más sin ser iguales). Para esto necesitamos 4 vértices y escoger una de las 6 aristas posibles como la arista que compartirán los triángulos, esto determina al par de triángulos posibles. Tenemos

$$\sum_{1 \leq i, j \leq \binom{n}{3}, i \neq j} \text{Cov}[T_i, T_j] = \binom{n}{4} \binom{4}{2} \text{Cov}[T_i, T_j],$$

donde las ternas i, j definen triángulos que sólo comparten una arista. Por otro lado

$$\text{Cov}[T_i, T_j] = E(T_i \cdot T_j) - E(T_i)E(T_j) = p(n)^5 - p(n)^6 \leq p(n)^5,$$

de donde

$$\sum_{1 \leq i, j \leq \binom{n}{3}, i \neq j} \text{Cov}[T_i, T_j] \leq 2 \binom{n}{4} \binom{4}{2} p(n)^5 = 12 \binom{n}{4} p(n)^5.$$

Del último término notemos que $\binom{n}{4} \in O(n^4)$, por lo tanto existen $N_1 \in \mathbb{N}$ y $c_1 \in \mathbb{R}^+$ tales que para $n > N_1$, $\binom{n}{4} \leq c_1 n^4$. De igual manera, $\binom{n}{3} \in O(n^3)$ y entonces existen $N_2 \in \mathbb{N}$ y $c_2 \in \mathbb{R}^+$ tales que para $n \geq N_2$ se tiene que $\binom{n}{3} \leq c_2 n^3$. Tomando $N = \max\{N_1, N_2\}$ y $c = \max\{c_1, c_2\}$ llegamos a que para cada $n > N$

$$\text{Var}[X_n] \leq \binom{n}{3} p(n)^3 + 2 \binom{n}{4} \binom{4}{2} p(n)^5 \leq c(n^3 p(n)^3 + n^4 p(n)^5)$$

Ahora, para aplicar el lema anterior, consideremos el siguiente cociente para $n > N$

$$\frac{\text{Var}[X_n]}{(E[X_n])^2} \leq \frac{\binom{n}{3} p(n)^3 + 2 \binom{n}{4} \binom{4}{2} p(n)^5}{\left(\binom{n}{3} p(n)^3\right)^2} \leq c \frac{n^3 p(n)^3 + n^4 p(n)^5}{\left(\binom{n}{3} p(n)^3\right)^2}.$$

Notemos que si $p(n)$ es una función constante distinta de cero, el término de la derecha tiende a cero conforme n crece puesto que el denominador es un polinomio de grado 6 y el numerador es de grado 4 en n . Hasta aquí, con $p(n)$ constante, se satisfacen ya las condiciones del lema anterior, por lo tanto, $\lim_{n \rightarrow \infty} P(G_n \text{ tiene un triángulo}) =$

1. Para $p(n)$ no constante, basta notar que $\binom{n}{3} \in O(n^3)$ y, por lo tanto, $\binom{n}{3}^2 p(n)^6 \in O(n^6 p(n)^6)$. Recordando 1.3.5, vemos que

$$c \frac{n^3 p(n)^3 + n^4 p(n)^5}{\left(\binom{n}{3} p(n)^3\right)^2} \in O\left(\frac{n^3 p(n)^3 + n^4 p(n)^5}{n^6 p(n)^6}\right) = O\left(\frac{1}{n^3 p(n)^3} + \frac{1}{n^2 p(n)}\right).$$

Como $\frac{1}{n} \in o(p(n))$, $\frac{1}{np(n)} \rightarrow 0$ y entonces $np(n) \rightarrow \infty$. Entonces $n^3 p(n)^3$ y $n^2 p(n)$ tienden ambos a infinito conforme n crece, y, por lo tanto, $\frac{\text{Var}[X_n]}{(E[X_n])^2} \rightarrow 0$. Otra vez, se satisfacen las condiciones del lema. Entonces $\lim_{n \rightarrow \infty} P(G_n \text{ tiene un triángulo}) = 1$ como queríamos demostrar. \square

Por ejemplo, si $p(n) = \frac{1}{\ln(n)}$, podemos apostar por triángulos casi siempre. La esperanza de aristas para una gráfica aleatoria donde sus aristas existen con probabilidad $p(n)$ es de orden $O\left(\frac{n^2}{\ln(n)}\right)$. El teorema de Turán exige $\frac{n^2}{4}$ aristas para garantizar un triángulo. Lo interesante aquí es que $\frac{n^2}{\ln(n)} \in o\left(\frac{n^2}{4}\right)$, en otras palabras, con muchas menos aristas podemos apostar por un triángulo con tanta seguridad como queramos.

5.2. Funciones umbral

En el teorema anterior usamos la condición $\frac{1}{n} \in o(p(n))$ para decir que se podían tener triángulos “casi siempre”. Una pregunta natural es ¿qué condiciones se necesitan sobre $p(n)$ para poder “casi nunca” garantizar triángulos? Sorprendentemente es suficiente que $p(n) \in o\left(\frac{1}{n}\right)$ para poder “casi nunca” garantizar triángulos. Este es un hecho fácil de verificar, pues, $E[X_n] = \binom{n}{3} p(n)^3 \in O(n^3 p(n)^3)$ y, si $p(n) \in o\left(\frac{1}{n}\right)$ entonces $np(n) \rightarrow 0$. Por lo tanto $n^3 p(n)^3 \rightarrow 0$.

Este fenómeno de que sea muy aguda la transición entre que una gráfica tenga un triángulo “casi siempre” o “casi nunca” puede generalizarse de la siguiente manera. Una propiedad de una gráfica se dice **monótona** si dada una gráfica G que la satisface, cualquier gráfica que la contenga también la satisface. Por ejemplo, tener un triángulo es una propiedad monótona y ser bipartita no lo es. Una **función umbral** para una propiedad monótona A es una función $f : \mathbb{N} \rightarrow [0, 1]$ tal que si $f \in o(p)$ entonces $\lim_{n \rightarrow \infty} P(A \text{ se satisface en } G(n, p(n))) = 1$ (casi siempre) y si $p(n) \in o(f(n))$ entonces $\lim_{n \rightarrow \infty} P(A \text{ se satisface en } G(n, p(n))) = 0$ (casi nunca). Como ejemplo, la propiedad de tener un triángulo tiene a $f(n) = \frac{1}{n}$ como función umbral según acabamos de demostrar.

5.2.1 Definición. Dada una gráfica $G = (V, E)$ se define su **densidad** como $\rho(G) = \frac{|E|}{|V|}$.

5.2.2 Definición. Decimos que una gráfica G es **balanceada** si para cada subgráfica H (no necesariamente inducida) se tiene que $\rho(H) \leq \rho(G)$.

El siguiente teorema es una generalización del anterior.

5.2.3 Teorema. *Sea H una gráfica balanceada con v vértices y e aristas. Entonces, la propiedad monótona “ $G_n = G(n, p(n))$ contiene a H como subgráfica” tiene como función umbral a $n^{-1/\rho(H)} = \frac{1}{n^{v/e}}$.*

Demostración. Sea $\{a_1, a_2, \dots, a_v\}$ el conjunto de vértices de H . Para cada v -tupla $\beta = (b_1, b_2, \dots, b_v)$ donde b_1, b_2, \dots, b_v son vértices distintos de G_n , sea A_β el suceso en el que β induce un homomorfismo de gráficas de H a G_n , es decir, A_β ocurre si $a_i \rightarrow b_i$ es un homomorfismo de gráficas. Llamemos X_β a la variable aleatoria que toma el valor 1 si A_β ocurre y el valor 0 si no. En este caso $X_n = \sum X_\beta$ no cuenta necesariamente el número de copias de H que se encuentran en G_n debido a las posibles simetrías de H . Sin embargo $X_n = 0$ nos indica la ausencia de copias de H , y $X_n > 0$ nos indica la presencia de copias de H y esto nos es suficiente.

El que ocurra A_β sólo depende de la existencia de las e aristas apropiadas, es decir, $P(A_\beta) = p(n)^e$. Entonces, por linealidad de la esperanza, tenemos que $E[X_n] = \sum E[X_\beta] = \binom{n}{v} p(n)^e$. Como $\binom{n}{v}$ es un polinomio de grado v , tenemos que $\binom{n}{v} \in \Theta(n^v)$ y, por lo tanto, $E[X_n] \in \Theta(n^v p(n)^e)$. En caso de que $p(n) \in o(\frac{1}{n^{v/e}})$, por definición, $p(n)n^{v/e} \rightarrow 0$ y, entonces, $(p(n)n^{v/e})^e = p(n)^e n^v \rightarrow 0$. Es decir que encontraremos una copia de H casi nunca, con lo que terminamos la primera parte de la demostración.

Ahora supongamos que $\frac{1}{n^{v/e}} \in o(p(n))$. Calculemos la varianza de X . Recordando 1.1.12 tenemos

$$Var[X_n] = \sum_{\beta, \gamma} Cov[X_\beta, X_\gamma].$$

La existencia de una arista es independiente de la existencia de otra; entonces la covarianza de X_β y X_γ es cero si las copias de H que β y γ inducen no comparten aristas. En especial, si β y γ comparten a lo más un vértice, su covarianza será cero. Como no tenemos información sobre cómo están distribuidas las aristas en H , sería difícil trabajar para copias con un número de aristas compartidas. En vez de eso trabajaremos con cotas para el número de aristas que podrían compartir las copias de H que inducen β y γ en función de cuántos vértices comparten.

Sea $2 \leq t \leq v$ el número de vértices que comparten β y γ . Sea e' el número de aristas que comparten. Ya que H es balanceada y que la intersección de las copias de H que inducen β y γ es una subgráfica de H , tenemos que $\frac{e'}{t} \leq \frac{e}{v}$, de donde e' es a lo más $\frac{te}{v}$. Ahora, la unión de las copias a lo más tiene $2e$ aristas (cada una tiene exactamente e), y sabemos que a lo más comparten $\frac{te}{v}$ aristas. Entonces, la unión al menos debe tener $2e - \frac{te}{v}$ aristas. Llamemos e'' al número de aristas en la unión. De esto último tenemos

$$Cov[X_\beta, X_\gamma] = E[X_\beta, X_\gamma] - E[X_\beta]E[X_\gamma] \leq E[X_\beta, X_\gamma] = p^{e''} \leq p^{2e - \frac{te}{v}}.$$

Ahora contemos cuántos pares β, γ comparten exactamente t vértices. Primero, como están compartiendo t vértices, en total los vértices que usan son $2v - t$. Las maneras de escoger esta cantidad de vértices son $\binom{n}{2v-t}$. De estos vértices escogidos hay que elegir los t que compartirán, esto se puede hacer de $\binom{2v-t}{t}$ maneras. Ahora necesitamos escoger los $v - t$ que faltan para β , esto se puede hacer de $\binom{2v-2t}{v-t}$ maneras, y los que faltan para γ ya quedan determinados. Luego, como el orden importa, falta revolver todas las entradas de β y γ , y eso se puede hacer de $v!v!$ maneras. Por último hay que dividir entre 2 porque no importa cuál de β y γ fue elegido primero. En total la cantidad de pares es $\binom{n}{2v-t} \binom{2v-t}{t} \binom{2v-2t}{v-t} v!v!/2$. De todo esto sólo nos interesa el primer factor, puesto que el resto es constante para t fija. Es decir, el número de pares β, γ que comparten exactamente t vértices pertenece a $O\left(\binom{n}{2v-t}\right) = O(n^{2v-t})$. Por 1.3.4(i), tenemos

$$\sum_{|\beta \cap \gamma|=t} Cov[X_\beta, X_\gamma] \in O(n^{2v-t} p(n)^{2e-\frac{te}{v}}) = O((n^v p(n)^e)^{2-\frac{t}{v}}).$$

Sumando sobre t tenemos

$$Var[X_n] \in O\left(\sum_{t=2}^v (n^v p(n)^e)^{2-\frac{t}{v}}\right).$$

Ahora, recordando 1.3.5, tenemos

$$\frac{Var[X_n]}{(E[X_n])^2} \in O\left(\frac{\sum_{t=2}^v (n^v p(n)^e)^{2-\frac{t}{v}}}{(n^v p(n)^e)^2}\right) = O\left(\sum_{t=2}^v (n^v p(n)^e)^{-\frac{t}{v}}\right).$$

Para poder aplicar el lema 5.0.5 faltaría ver que el término de la derecha tiende a cero conforme n crece. Puesto que el número de sumandos es constante, basta ver que cada sumando tiende a cero conforme n crece. Entonces tomemos una t fija. El que $\frac{1}{n^{v/e}} \in o(p(n))$ significa que $\frac{1}{n^{v/e} p(n)} \rightarrow 0$, es decir que $n^{v/e} p(n) \rightarrow \infty$ y, por lo tanto, $(n^{v/e} p(n))^e = n^v p(n)^e \rightarrow \infty$. Como v es constante y estamos trabajando para una t fija, tenemos

$$(n^v p(n)^e)^{-\frac{t}{v}} = \frac{1}{(n^v p(n)^e)^{\frac{t}{v}}} \rightarrow 0.$$

Aplicando el lema 5.0.5, vemos que $\lim_{n \rightarrow \infty} P(X_n > 0) = 1$, es decir que encontraremos al menos una copia de H casi siempre, como queríamos demostrar. \square

Notación

$\lfloor x \rfloor$	Función parte entera o función piso, mayor entero menor o igual a x .
$\lceil x \rceil$	Función techo, menor entero mayor o igual a x .
$\binom{n}{m}$	Cantidad de subconjuntos de m elementos contenidos en un conjunto de n elementos. Se sabe que este número está dado por $\frac{n!}{m!(n-m)!}$.
$n!$	Factorial de n , donde $n \in \mathbb{N} \cup \{0\}$. La función factorial se define recursivamente por $0! = 1$, $n! = n \cdot (n-1)!$.
$ A $	Cardinalidad del conjunto A .
$P(A)$	Probabilidad del suceso A .
$E[X]$	Esperanza de la variable aleatoria X . Véase 1.1.4.
$Var[X]$	Varianza de la variable aleatoria X . Véase 1.1.9.
$Cov[X, Y]$	Covarianza de las variables aleatorias X, Y . Véase 1.1.11.
$O(f)$	Conjunto de funciones $g \in \mathbb{N}^{\mathbb{R}^+}$ (ó $\mathbb{R}^{+\mathbb{R}^+}$) tales que existen $k > 0$, $n_0 > 0$ con $g(n) \leq f(n) \cdot k$ para toda $n > n_0$. Véase 1.3.1.
$\Theta(f)$	Conjunto de funciones $g \in \mathbb{N}^{\mathbb{R}^+}$ (ó $\mathbb{R}^{+\mathbb{R}^+}$) tales que $g \in O(f)$ y $f \in O(g)$. Véase página 17.
$o(f)$	Conjunto de funciones $g \in \mathbb{N}^{\mathbb{R}^+}$ (ó $\mathbb{R}^{+\mathbb{R}^+}$) tales que $\lim_{n \rightarrow \infty} \frac{g(n)}{f(n)} = 0$. Véase 1.3.6.

- $G = (V, E)$ En este trabajo, la letra G siempre se refiere a una gráfica simple y, las letras V, E representan a sus conjuntos de vértices y aristas respectivamente. Véase 2.2.1.
- K_n Gráfica completa de n vértices. Véase 2.2.2.
- $R(k, l)$ Número de Ramsey para k y l . Véase 2.2.4.
- $K_{m,n}$ Gráfica bipartita completa. Véase 3.3.2.
- $\chi(G)$ Número cromático de la gráfica G . Véase página 48.
- $\alpha(G)$ Número de independencia de la gráfica G . Véase página 49.

Bibliografía

- [1] María Luisa Pérez Seguí. *Combinatoria Avanzada*. Instituto de Matemáticas, Universidad Nacional Autónoma de México, 2010.
- [2] Gonzalo Navarro. *Fundamentos de la Ciencia de la Computación (Lenguajes Formales, Computabilidad y Complejidad)*. Universidad de Chile, 2008.
- [3] Harry R. Lewis; Christos H. Papadimitriou. *Elements of Theory of Computation*. PRENTICE-HALL, 1998.
- [4] Noga Alon; Joel H. Spencer. *The Probabilistic Method*. Wiley-Interscience, 2000.
- [5] J. A. Bondy; U. S. R. Murty. *Graph Theory With Application*. NORTH-HOLLAN, 1976.
- [6] Jiří Matoušek; Jan Vondrák. *The Probabilistic Method*. Department of Applied Mathematics, Charles University, Czech Republic, 2008
- [7] Gregory J. E. Rawlins. *Compared to what? an introduction to the analysis of algorithms*. W. H. Freeman & Co. 1991.
- [8] Sheldon M. Ross; *Introduction to Probability Models*. University of California, 1997.
- [9] Charles M. Grinstead; J. Laurie Snell. *Introduction to Probability*. American Mathematical Society, 1997.
- [10] Robert M. Gray. *Probability, Random Processes and Ergodic Propierties*. Stanford University, 2001.