



Universidad Nacional Mayor de San Marcos

Universidad del Perú. Decana de América

Facultad de Ciencias Matemáticas

Escuela Profesional de Matemática

**Desvío moderado de conteo de triángulos en los
modelos de grafos aleatorios uniforme y binomial**

TESIS

Para optar el Título Profesional de Licenciado en Matemática

AUTOR

José Diego ALVARADO MORALES

ASESOR

Dra. Yolanda Silvia SANTIAGO AYALA

Lima, Perú

2023



Reconocimiento - No Comercial - Compartir Igual - Sin restricciones adicionales

<https://creativecommons.org/licenses/by-nc-sa/4.0/>

Usted puede distribuir, remezclar, retocar, y crear a partir del documento original de modo no comercial, siempre y cuando se dé crédito al autor del documento y se licencien las nuevas creaciones bajo las mismas condiciones. No se permite aplicar términos legales o medidas tecnológicas que restrinjan legalmente a otros a hacer cualquier cosa que permita esta licencia.

Referencia bibliográfica

Alvarado, J. (2023). *Desvío moderado de conteo de triángulos en los modelos de grafos aleatorios uniforme y binomial*. [Tesis de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Matemática]. Repositorio institucional Cybertesis UNMSM.

Metadatos complementarios

Datos de autor	
Nombres y apellidos	José Diego Alvarado Morales,
Tipo de documento de identidad	DNI
Número de documento de identidad	46847040
URL de ORCID	https://orcid.org/0000-0002-8570-2487
Datos de asesor	
Nombres y apellidos	Yolanda Silvia Santiago Ayala
Tipo de documento de identidad	DNI
Número de documento de identidad	06445705
URL de ORCID	https://orcid.org/0000-0003-2516-0871
Datos del jurado	
Presidente del jurado	
Nombres y apellidos	Martha Hilda Timoteo Sánchez
Tipo de documento	DNI
Número de documento de identidad	25702279
Miembro del jurado 1	
Nombres y apellidos	José Raúl Luyo Sanchez
Tipo de documento	DNI
Número de documento de identidad	09394743
Datos de investigación	
Línea de investigación	A.3.1.1 Ecuaciones Diferenciales (Ordinarias, Parciales) y Análisis Funcional A.3.1.4 Matemática Aplicada (Biomatemática, Economía, Optimización y otros)

Grupo de investigación	Grupo de Ecuaciones Diferenciales, Análisis y Aplicaciones - GEDAAP
Agencia de financiamiento	Sin financiamiento
Ubicación geográfica de la investigación	País: Perú Departamento: Lima Provincia: Lima Distrito: Cercado de Lima Latitud: -12.063897 Altitud: -77.064461
Año o rango de años en que se realizó la investigación	2021-2023
URL de disciplinas OCDE	MATEMATICA PURA https://purl.org/pe-repo/ocde/ford#1.01.00



UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS

(Universidad del Perú, DECANA DE AMÉRICA)

FACULTAD DE CIENCIAS MATEMÁTICAS

Ciudad Universitaria - Av. Venezuela S/N cuadra 34

Teléfono: 619-7000, Anexo 1610

Correo Postal: 05-0021, E-mail: eapmat@unmsm.edu.pe

Lima - Perú

Escuela Profesional de Matemática

***ACTA DE SUSTENTACIÓN DE TESIS PARA OPTAR EL TÍTULO PROFESIONAL DE
LICENCIADO EN MATEMÁTICA
(Modalidad Virtual)***

*En la UNMSM - Ciudad Universitaria - Facultad de Ciencias Matemáticas, siendo las 20:00 horas del jueves 14 de diciembre del 2023, se reunieron los docentes designados como Miembros del Jurado Evaluador de Tesis: Dra. Martha Hilda Timoteo Sánchez (PRESIDENTE), Dr. José Raúl Luyo Sanchez (MIEMBRO) y la Dra. Yolanda Silvia Santiago Ayala (MIEMBRO ASESOR), para la sustentación de la tesis titulada: “**DESVÍO MODERADO DE CONTEO DE TRIÁNGULOS EN LOS MODELOS DE GRAFOS ALEATORIOS UNIFORME Y BINOMIAL**”, presentado por el señor Bachiller **José Diego Alvarado Morales**, para optar el Título Profesional de Licenciado en Matemática.*

Luego de la exposición de la tesis, el Presidente invitó al expositor a dar respuesta a las preguntas formuladas.

Realizada la evaluación correspondiente por los miembros del Jurado Evaluador, el expositor mereció la aprobación de sobresaliente con mención con un calificativo promedio de 19(diecinove).

*A continuación, el Presidente del Jurado, Dra. Martha Hilda Timoteo Sánchez, manifestó que el señor Bachiller **José Diego Alvarado Morales**, en vista de haber aprobado la sustentación de su tesis, será propuesto para que se le otorgue el Título Profesional de Licenciado en Matemática.*

Siendo las 9:04 pm horas se levantó la sesión firmando para constancia la presente Acta en tres (2) copias originales.

Dra. Martha Hilda Timoteo Sánchez
PRESIDENTE

Dr. José Raúl Luyo Sánchez
MIEMBRO

Dra. Yolanda Silvia Santiago Ayala
MIEMBRO ASESOR



CERTIFICADO DE SIMILITUD

Yo Dra. Yolanda Silvia Santiago Ayala en mi condición de asesor acreditado con la Resolución Decanal N°001356-2023-D-FCM/UNMSM de la tesis/monografía/informe de investigación/trabajo académico, cuyo título es **“DESVÍO MODERADO DE CONTEO DE TRIÁNGULOS EN LOS MODELOS DE GRAFOS ALEATORIOS UNIFORME Y BINOMIAL”**, presentado por el bachiller José Diego Alvarado Morales, para optar el título Profesional de Licenciado en Matemática CERTIFICO que se ha cumplido con lo establecido en la Directiva de Originalidad y de Similitud de Trabajos Académicos, de Investigación y Producción Intelectual. Según la revisión, análisis y evaluación mediante el software de similitud textual, el documento evaluado cuenta con el porcentaje de 2% de similitud, nivel **PERMITIDO** para continuar con los trámites correspondientes y para su **publicación en el repositorio institucional.**

Se emite el presente certificado en cumplimiento de lo establecido en las normas vigentes, como uno de los requisitos para la obtención del grado/ título/ especialidad correspondiente.

Firma del Asesor

DNI: 06445705

Nombres y apellidos del asesor:

YOLANDA SILVIA SANTIAGO AYALA



*A mis padres Víctor y Rosa,
a mi familia,
y a Dios.*

AGRADECIMIENTOS

Primeramente quisiera agradecer a Dios y a mi familia. Principalmente a mis padres que gracias a su apoyo fue posible que yo consiguiese tener y culminar mis estudios universitarios. Sin olvidar a mi querida hermana y a mi sobrina.

A mi asesora, la Dr. Yolanda Silvia Santiago Ayala, que desde estudiante siempre supo despertar en mi (y con seguridad en todos sus alumnos) ese interés por la matemática, principalmente en su forma analítica más pura. Es una alegría para mi poder contar con su asesoría y consejos tanto académicos como profesionales.

Al departamento de matemática de la UNMSM, tanto profesores, amigos y administrativos, pues gracias a ellos adquirí los primeros conocimientos que me permitieron avanzar como matemático y como persona. También agradezco al Profesor Simon Griffiths y al Dr. Leonardo G. de Oliveira por el intercambio de ideas y los trabajos en conjunto realizados.

Finalmente, me gustaría agradecer a mi pareja Fabiola Cordero, quien me a brindado su compañía incondicional y fue un gran soporte durante los años difíciles de pandemia que nos tocó vivir a todos.

RESUMEN

DESVÍO MODERADO DE CONTEO DE TRIÁNGULOS EN LOS MODELOS DE GRAFOS ALEATORIOS UNIFORME Y BINOMIAL

JOSÉ DIEGO ALVARADO MORALES

Asesora : Dra. Yolanda Silvia Santiago Ayala.

Título Obtenido : Licenciado en Matemática.

En este trabajo discutimos el problema de determinar la probabilidad de desvíos moderados respecto al número de triángulos en dos modelos de grafos aleatorios, a saber, el modelo uniforme y el binomial.

Principalmente nos enfocaremos en el trabajo de Goldschmidt, Griffiths y Scott (2020), donde se explora la conexión entre la Teoría de Martingalas y este tipo de problemas. En dicho trabajo se resuelve satisfactoriamente este problema en el ambiente denso.

Finalmente presento un esquema de como aplicar esta técnica, junto con nuevas ideas, para resolver el problema en el ambiente disperso, siguiendo [1].

Palabras Claves: Desvío moderado, grafos aleatorios, conteo de subgrafos, teoría de martingalas.

ABSTRACT

MODERATE DEVIATION FOR TRIANGLES COUNT IN UNIFORM AND BINOMIAL RANDOM GRAPH MODELS

JOSÉ DIEGO ALVARADO MORALES

Advisor : Dra. Yolanda Silvia Santiago Ayala.

Obtained Title : Licentiate in Mathematics.

In this work, we discuss the problem of determining the probability of moderate deviations concerning the number of triangles in two models of random graphs, namely the uniform model and the binomial model.

Our main focus is on the work of Goldschmidt, Griffiths, and Scott (2020), where the connection between Martingale Theory and this type of problem is explored. In that work, this problem is successfully solved in the dense setting.

Finally, I present an outline of how to apply this technique, along with new ideas, to solve the problem in the sparse setting, following [1].

Keywords: Moderate deviation, random graphs, subgraph counts, martingale theory.

Índice general

Resumen	III
Abstract	IV
Glosario de Símbolos Matemáticos	VIII
1. Introducción	1
1.1. Historia	1
2. Preliminares	8
2.1. Notación Asintótica	8
2.2. Teoría de Grafos	9
2.3. Probabilidades	13
2.4. Combinatoria Probabilística	24
2.4.1. Modelos Probabilísticos Combinatorios	24
2.4.2. Aplicación de Martingalas en $\mathbf{G}(n, m)$	27
2.4.3. Análisis	28
3. Sobre los resultados de Goldschmidt, Griffiths y Scott	31
3.1. Panorama General de Goldschmidt, Griffiths y Scott	31
3.2. Desvíos Moderados en $\mathbf{G}(n, m)$	32
3.2.1. Representación Martingala de $D_{\wedge}(\mathbf{G}(n, m))$ y $D_{\Delta}(\mathbf{G}(n, m))$	33
3.2.2. Sobre Grados y Cogrados en $\mathbf{G}(n, m)$	39
3.2.3. Probabilidad del Desvío de Conteo de Cerezas y Triángulos en $\mathbf{G}(n, m)$	51
4. Nuevos resultados	58
4.1. Resultados	58
4.2. Nuevas Ideas	63
4.3. Ingredientes Principales	71
4.3.1. Sobre Grados en $\mathbf{G}(n, m)$: Modelo Disperso	72
4.3.2. Sobre Cogrados en $\mathbf{G}(n, m)$: Modelo Disperso	76

4.4.	Desvío positivo de triángulos en $\mathbf{G}(n, m)$: límite superior	84
4.4.1.	Controlando $D'_{\Delta}(G_m)$	85
4.4.2.	Controlando $N^*_{\Delta}(G_m)$	90
5.	Conclusiones	99
	Bibliografía	103

Índice de figuras

2.1. Esta figura ilustra un grafo completo con 10 vértices (los vértices se encuentran etiquetados).	10
2.2. La figura ilustra un <i>biclique</i> con partes A y B de tamaño 8 y 5, respectivamente. Por comodidad, hemos denotado por a_i los elementos de A y por b_j los de B	10
2.3. Un matching (izquierda) y una estrella (derecha) formadas por tres aristas cada una.	11
2.4. Diseño de un endomorfismo que ‘sumerge’ un triángulo en un grafo <i>diamante</i>	11
2.5. La figura muestra un grafo con grado máximo 5, donde el único vértice con dicho grado es colorido de rojo.	13
2.6. La figura ilustra una parte de una realización de un proceso aleatorio (indexado por tiempo discreto).	15
2.7. En este arreglo podemos ver una experiencia de lanzar 10 monedas independientes (cara= 1, sello= 0).	17
2.8. La imagen ilustra la distribución de una variable aleatoria. Aquí podemos contemplar el comportamiento de la variable, en particular, sus desvíos.	18
2.9. La imagen muestra una realización de un subconjunto aleatorio uniforme colorido en azul.	25

Glosario de Símbolos Matemáticos

$(x)_k$	$:= x(x-1)\dots(x-k+1)$.
$\sigma(X)$	sigma álgebra de la variable aleatoria (función medible) X .
$\binom{U}{2}$	conjunto de pares $\{u, w\}$ donde $u, w \in U$ y $u \neq w$.
n	variable entera positiva que tiende a $+\infty$.
$\lfloor \cdot \rfloor$	función piso que asigna a cada real su entero inferior más próximo (se acostumbra, como haremos aquí, que la función piso sea tácita).
$[r]$	$:= \{1, 2, \dots, r\}$.
$[r, R]$	$:= \{r, r+1, \dots, R\}$.
N	$:= \binom{n}{2}$.
$\mathbb{Z}_{\geq 0}$	conjunto de los números enteros no-negativos.
$ S $	cardinalidad del conjunto (finito) S .
$\mathbf{G}(n, m), G_m$	grafo aleatorio uniforme con n vértices y m aristas.
$\mathbf{G}(n, p)$	grafo aleatorio p -binomial con n vértices.
$N_H(G)$	número de copias inmersas del grafo H en G .
$D_H(G_m)$	$:= N_H(G_m) - \mathbb{E}[N_H(G_m)]$.
$A_H(G_{i-1} \cup e_i)$	número de copias inmersas de H que aparecen al ser escogida la arista e_i .
$X_H(G_{i-1} \cup e_i)$	$:= A_H(G_{i-1} \cup e_i) - \mathbb{E}[A_H(G_{i-1} \cup e_i) G_{i-1}]$.
t	$:= \frac{m}{N}$.
ℓ	$:= \ln(1/t)$.
ℓ_b	$:= \ln(eb/tn)$.
$d_u(G)$	grado del vértice u en el grafo aleatorio G_m .
$D_u(G)$	$:= d_u(G_m) - \mathbb{E}[d_u(G_m)]$
$d_{uv}(G)$	cogrado del par de vértices $\{u, v\}$ en el grafo aleatorio G_m .
$D_{uv}(G)$	$:= d_{uv}(G_m) - \mathbb{E}[d_{uv}(G_m)]$.

Capítulo 1

Introducción

En una era marcada por la interconexión y sistemas complejos, la teoría de grafos aleatorios, en particular el estudio subestructuras (*clusters*), emerge como una herramienta útil con aplicaciones de amplio alcance (vea por ejemplo [2]).

Dentro del entramado de las redes sociales, los grafos aleatorios son modelos que ofrecen información sobre la formación de grupos cohesivos (y la dinámica de la propagación de la información, ver [40]). Al adentrarnos en la neurociencia, estos modelos matemáticos nos ayudan a entender las conexiones neuronales, proporcionando una aproximación de la compleja interacción de las neuronas en el cerebro.

De hecho, esta influencia es ubicua y se extiende a nuestro mundo digital, siendo una herramienta de gigantes tecnológicos como Facebook y Google. Los algoritmos que respaldan sugerencias de amigos, conexiones y recomendaciones de contenido están modelados por grafos aleatorios (*large networks*), adaptando las experiencias de usuario a preferencias individuales (vea [30]).

Así, desde la modelización de enfermedades hasta la neurociencia y los algoritmos que dan forma a nuestras experiencias en línea, la exploración de grafos aleatorios se muestra como un tópico importante, y es lo que pasaremos a explorar en este trabajo (por lo menos una de sus tantas facetas).

1.1. Historia

La historia de los grafos aleatorios comienza en 1959-1960 con los trabajos iniciales y fundamentales de los matemáticos Húngaros Erdős y Rényi [16, 17]. Ellos introducen el modelo aleatorio uniforme, denotado por $G(n, m)$, con distribución uniforme respecto a todos los grafos con n vértices (etiquetados) y m aristas. Este grafo aleatorio también puede ser generado por un proceso estocástico, conocido actualmente como el proceso Erdős-Rényi, que describimos a seguir:

- comenzamos con el grafo vacío con n vértices, denotado por G_0 ,
- en el paso i -ésimo del proceso, adicionamos una arista escogida uniformemente de las $\binom{n}{2} - i + 1$ sobrantes. Denotemos por \mathbf{G}_i el grafo aleatorio obtenido en este paso.

De este modo, el grafo aleatorio \mathbf{G}_m tiene la misma distribución $\mathbf{G}(n, m)$. Paralelamente (e independientemente) a los trabajos de Erdős y Rényi, el matemático Gilbert introdujo el modelo aleatorio binomial [20], denotado por $\mathbf{G}(n, p)$, generado al conectar con probabilidad $p \in [0, 1]$ cada par de vértices (distintos) de forma independiente. Aunque, en naturaleza, ambos modelos aleatorios son diferentes, estos se encuentran fuertemente relacionados, por lo que no es sorprendente que estudiando uno de estos modelos, podamos transferir ese conocimiento al otro modelo.

En este trabajo, cuando hablemos de *desvíos*, nos referiremos a los eventos en que una variable (cantidad) aleatoria se encuentra por encima (o por debajo) de su esperanza (valor esperado). Así, nuestro interés es estimar, lo mejor posible, la probabilidad de ciertos desvíos. El asunto de desvíos en la Teoría de Probabilidades es un foco muy importante en esta área de las matemáticas, y abarcan los ramos conocidos como Teorema Central del Límite hasta la Teoría de los Grandes Desvíos. El primero estudia los desvíos en torno del desvío padrón, mientras que el segundo estudia los desvíos en torno a su esperanza. Los primeros resultados importantes sobre desvíos fueron mostrados para combinaciones lineales de variables aleatorias independientes. Naturalmente, el siguiente paso sería estudiar sistemas con cierta dependencia. Un ejemplo clásico e importante para nosotros es la cuestión sobre el número de triángulos (e otros subgrafos fijos) en un grafo aleatorio binomial $\mathbf{G}(n, p)$. Para el lector interesado en los primeros trabajos de esta área, le referimos a [38, 39].

Denotemos por $N_\Delta(G)$ (respectivamente, $N_H(G)$) el número ‘inmerso’ de triángulos (respectivamente, copias ‘inmersas’ de H) en el grafo G . El foco central de estudio al que nos referimos, es el de determinar el comportamiento asintótico de la “función de tasa”

$$r(\delta, p, n) := -\log \mathbb{P} \left(N_\Delta(\mathbf{G}(n, p)) > (1 + \delta) p^3(n)_3 \right).$$

Note que $\mathbb{E} [N_\Delta(\mathbf{G}(n, p))] = p^3(n)_3$. Los primeros resultados en esta dirección se encuentran en los artículos de Vu [39], Kim y Vu [28], y Janson, Oleszkiewicz y Ruciński [26] en los inicios del 2000. Ellos mostraron que, para $\delta > 0$ fijo,

$$c(\delta)p^2n^2 \leq r(\delta, p, n) \leq C(\delta)p^2n^2 \log(1/p). \quad (1.1)$$

Sin embargo, quedaba responder cual era el orden de magnitud “correcto” de la función $r(\delta, p, n)$.

Si imaginamos que la causa del desvío $\delta p^3(n)_3$ en el número de triángulos se debe “esencialmente” a la presencia de una estructura de tipo ‘clique’¹ con $s = \Theta(\delta^{1/3}pn)$ vértices

¹El término ‘clique’ es usado para referirnos a un subgrafo completo dentro de un grafo mayor.

(fenómeno conocido como “fenómeno de localización”), es plausible esperar que el orden correcto de $r(\delta, p, n)$ sea dado por el limite superior en la desigualdad (1.1), pues la probabilidad de “plantar” una tal estructura es $p^{\binom{s}{2}} = \exp(-\Theta_\delta(p^2 n^2 \log(1/p)))$.

Esta conjetura fue resuelta posteriormente en dos trabajos publicados independiente por Chatterjee [7] y por DeMarco y Kahn [11]. Específicamente, ellos mostraron que, para $\delta > 0$ fijo,

$$c(\delta)p^2 n^2 \log(1/p) \leq r(\delta, p, n) \leq C(\delta)p^2 n^2 \log(1/p).$$

Esto último resolvía el problema de determinar el orden de magnitud de la función $r(\delta, p, n)$ (pensando en δ como una cantidad fija).

La siguiente cuestión natural sobre la función $r(\delta, p, n)$ sería determinar su orden asintótico (cuando $n \rightarrow \infty$). En esta línea de investigación, los probabilistas Chatterjee y Varadhan [9] mostraron que este problema para el caso denso, es decir con $p \in (0, 1)$ fijo, puede ser reducido a un problema variacional para unos objetos analíticos llamados graphons². Posteriormente Chatterjee y Dembo [8] extendieron este resultado para el caso disperso, es decir cuando $p = o_n(1)$, con la condición que p no sea demasiado “disperso”, específicamente $p \geq n^{-1/42} \log n$. Así, restaba resolver estos problemas variacionales. Fueron los matemáticos Lubetzky y Zhao [32] los que consiguieron resolver este problema, obteniendo que

$$r(\delta, p, n) = (1 + o_n(1)) \min \left\{ \frac{\delta^{2/3}}{2}, \frac{\delta}{3} \right\} p^2 n^2 \log(1/p) \quad (1.2)$$

cuando $n^{-1/42} \log n \leq p = o_n(1)$ (con $\delta > 0$ fijo).

En trabajos más recientes diferentes autores, entre ellos Eldan [15] y Cook y Dembo [10] consiguieron extender estos resultados para valores de p del orden $\omega(n^{-1/2}(\log n)^2)$.

Para valores más dispersos de p no se tenía una prueba de la validez de (1.2). Sorprendentemente, usando métodos diferentes³ Harel, Mousset y Samotij [22] consiguieron extender estos resultados para valores “muy dispersos” de p , respondiendo esencialmente por completo el problema de la “cola superior” (grandes desvíos) en el conteo de triángulos para el modelo binomial $\mathbf{G}(n, p)$. En particular, consiguieron mostrar la validez de la fórmula (1.2) para $n^{-1/2} \ll p = o_n(1)$, más aún, descubrieron el comportamiento de $r(\delta, p, n)$ para los valores más dispersos posibles de p (esto es, para $n^{-1} \ll p \ll n^{-1/2}$) además de estudiar el problema más complejo cuando $p = cn^{-1/2}$ (para $c \in \mathbb{R}$ positivo) donde muestran que los desvíos son causados por ciertas estructuras “híbridas”.

²La terminología “graphons” viene de la combinación de las palabras en inglés “graph” y “functions”. Los graphons pueden ser pensados como objetos límites de los grafos, así como los números reales (en el intervalo $[0, 1]$) son objetos límites de los racionales (en este intervalo).

³Siendo más precisos, usaron los métodos conocidos como el “método de los momentos” y el “método de la entropía”.

Lo anterior expuesto trató sobre grandes desvíos (positivos). Un otro problema de sumo interés tanto en Probabilidades como en Combinatoria Probabilística es el del “Teorema Central del Límite”⁴. Esta sub-área de las probabilidades intenta entender el comportamiento en distribución (asintótica) de una secuencia de cantidades aleatorias (debidamente re-escalada). Así, dada una secuencia aleatoria real $(\mathbf{X}_n)_n$, la primera pregunta a estudiar es si la variable aleatoria

$$\bar{\mathbf{X}}_n := \frac{\mathbf{X}_n - \mathbb{E}[\mathbf{X}_n]}{\sigma(\mathbf{X}_n)}$$

converge en distribución (cuando n tiende a $+\infty$) a $\mathcal{N}(0, 1)$ ⁵. Una vez respondida afirmativamente esta pregunta (si es el caso) el siguiente problema natural es estudiar la tasa de convergencia, esto es, entender el comportamiento asintótico de la secuencia:

$$r_n := \sup_{x \in \mathbb{R}} |\mathbb{P}(\bar{\mathbf{X}}_n \leq x) - \mathbb{P}(\mathbf{Z} \leq x)|$$

donde \mathbf{Z} es una variable aleatoria real con distribución $\mathcal{N}(0, 1)$.

Actualmente disponemos de una amplia literatura sobre el Teorema Central del Límite relacionada al problema de estudio de este trabajo. Por citar algunas referencias, en [38] Ruciński demostró que la variable aleatoria que cuenta el número de triángulos⁶ en $\mathbf{G}(n, p)$, llamémosla provisoriamente $\mathbf{X}_\Delta(n, p)$, satisface el “teorema central del límite” (como explicado en el párrafo anterior) siempre que $p = p_n$ satisface las condiciones $p^3 n, (1 - p)n^2 \rightarrow +\infty$ (cuando $n \rightarrow +\infty$).

Como indicado, la siguiente cuestión a investigar es el comportamiento (asintótico) de la tasa de convergencia r_n . En la literatura existen varios trabajos que fueron mejorando estas cotas, aquí nos limitamos a listar algunos de ellos [3, 29, 36, 37]. Más aún, Janson [25] (trabajando sobre los primeros resultados de Janson [23] y Janson y Nowicki [24]) consiguió probar un resultado más general que el Teorema Central del Límite, conocido como “*Functional Central Limit Theorem*” (o como el “Principio Invariante de Donker”), respecto al conteo de triángulos (y otros subgrafos) en el grafo aleatorio $\mathbf{G}(n, m)$.

Si bien lo anteriormente expuesto está relacionado con el contenido de este trabajo, nuestro foco de estudio son los desvíos moderados, esto es, aquellos que varían “entre” el desvío estándar y la esperanza. Esta es una parte de la Combinatoria Probabilística menos explorada.

Uno de los primeros resultados relevantes es el de Doering y Eichelsbacher [13] donde

⁴Abreviado por sus siglas en inglés ‘CLT’.

⁵Es usual usar la notación $\mathcal{N}(0, 1)$ para referirnos a la distribución normal (o Gaussiana) centralizada en 0 y con desvío estándar 1.

⁶De hecho, en su trabajo Ruciński estudia otros subgrafos además de triángulos.

fue probado el siguiente resultado asintótico:

$$r(\delta_n, p, n) \sim \frac{\delta_n^2 p n^2}{36(1-p)}$$

siempre que $p^{-1/2} n^{-1} \ll \delta_n \ll p^7$ (vea también [14]). Note que esta última condición solo aplica cuando $p \gg n^{-2/15}$, dejando un gran espacio por cubrir tanto para p como para δ_n (en el caso disperso). Por otro lado, para el modelo denso (esto es, cuando $p \in (0, 1)$ está fijo) un resultado sorprendente debido a Féray, Méliot y Nikeghbali [18] nos proporciona no sólo un valor asintótico de $r(\delta, p, n)$, sino que nos da el valor asintótico de la probabilidad de ciertos desvíos moderados. A saber, en [18] los autores muestran que:

$$\mathbb{P} \left(\mathbf{X}_\Delta(n, p) > (1 + \delta) p^3 \binom{n}{3} \right) \sim \sqrt{\frac{9(1-p)}{\pi p}} \exp \left(-\frac{\delta^2 p n^2}{36(1-p)} + \frac{(7-8p)\delta^3 p n^2}{324(1-p)^2} - \log(n\delta) \right)$$

siempre que $n^{-1} \ll \delta_n \ll n^{-1/2}$.

Volvamos por un momento al modelo aleatorio uniforme. En [21], los autores enfocaron su investigación en el grafo aleatorio $\mathbf{G}(n, m)$ visto como el grafo generado por el proceso Erdős-Rényi en la posición m -ésima, como explicado al inicio de esta introducción. Denotemos por $D_\Delta(G_m)$ la variable aleatoria dada al “centralizar” $N_\Delta(G_m)$ (donde $G_m \sim \mathbf{G}(n, m)$), esto es

$$D_\Delta(G_m) := N_\Delta(G_m) - L_\Delta(m)$$

con $L_\Delta(m) := \mathbb{E}_{G_m \sim \mathbf{G}(n, m)}[N_\Delta(G_m)]$. La idea clave en [21] es exhibir una representación martingala de $D_\Delta(G_m)$ para todo valor de $m \in [N]$, de modo que los incrementos de esta representación tengan cierto “contenido combinatorio” de modo que podamos manipularlos con suceso mediante técnicas combinatorio-probabilísticas.

Para el lector familiarizado con la Teoría de Martingalas, es sabido que dado una variable aleatoria centralizada \mathbf{Y} (esto es, $\mathbb{E}[\mathbf{Y}] = 0$) y una filtración $(\mathcal{F}_i)_{i=0}^m$ de σ -álgebras donde $\mathcal{F}_0 = \{\emptyset, \Omega\}$ y $\mathcal{F}_m = \sigma(\mathbf{Y})$, uno puede escribir \mathbf{Y} como una representación martingala del siguiente modo

$$\mathbf{Y} = \sum_{i=1}^m (\mathbb{E}[\mathbf{Y} | \mathcal{F}_i] - \mathbb{E}[\mathbf{Y} | \mathcal{F}_{i-1}]),$$

este tipo de descomposición también es llamado *Descomposición de Doob*, en honor al probabilista Joseph L. Doob quien desarrolló esta teoría⁷. Ya que el proceso de Erdős-Rényi

⁷El resultado es más general y, a grosso modo, permite descomponer un proceso estocástico adaptado (L^1 -limitado) en una única suma (c.c) de un proceso martingala y una proceso predecible (integrable).

induce naturalmente una filtración (\mathcal{F}_i) ⁸, el lector observador puede percibir que bastaría aplicar la descomposición anterior con $D_\Delta(G_m)$ en el lugar de \mathbf{Y} . Si bien esto es plausible, los incrementos producidos por esta descomposición resultan algo difíciles de manipular pues dependen de promediar muchos eventos a futuro. De forma ingeniosa los autores en [21] exhiben una representación martingala más fácil de manipular (con significado combinatorio, y con una prueba meramente inductiva):⁹,

$$D_\Delta(G_m) = \sum_{i=1}^m \left\{ \frac{(N-m)_2(m-i)}{(N-i)_3} X_\wedge(G_{i-1} \cup e_i) + \frac{(N-m)_3}{(N-i)_3} X_\Delta(G_{i-1} \cup e_i) \right\}. \quad (1.3)$$

En esta expresión, el “incremento” $X_\Delta(G_{i-1} \cup e_i)$ es dado por la expresión $A_\Delta(G_{i-1} \cup e_i) - \mathbb{E}[A_\Delta(G_{i-1} \cup e_i) | G_{i-1}]$, donde la variable (aleatoria) $A_\Delta(G_{i-1} \cup e_i)$ cuenta el número de triángulos (inmersos) creados al adicionar la arista ‘seleccionada’ e_i al grafo G_{i-1} (en el proceso Erdős-Rényi) y la esperanza condicionada se toma respecto de la arista ‘variable’ e_i (similarmente para $X_\wedge(G_{i-1} \cup e_i)$ substituyendo “ Δ ” por “ \wedge ”).

Con la ecuación (1.3) en nuestras manos y con la ayuda de la “desigualdad de Freedman”, es plausible esperar que seamos capaces de estimar el comportamiento (asintótico) de

$$\log \mathbb{P}(D_\Delta(\mathbf{G}(n, m)) > \alpha) \quad (1.4)$$

para $n^{3/2} \ll \alpha \ll n^2$ (en el modelo denso).¹⁰

Para finalizar esta introducción, nos gustaría resaltar algunos aspectos de nuestra propia investigación. Esta comenzó motivado por el deseo de mejorar/extender los resultados de Goldschmidt, Griffiths y Scott en el mundo disperso, debido a que sus resultados eran óptimos sólo en el mundo (relativamente) denso.

La estrategia central que ellos siguieron es mantenida, pero importando otras ideas para lidiar con las “obstrucciones” que aparecen ya al inicio de su argumento (en el mundo disperso). Recordemos que una de las ideas centrales es “pensar” en el modelo binomial $\mathbf{G}(n, p)$ como un promedio ponderado de los modelos uniformes $\mathbf{G}(n, m)$. Así, entender un evento (en nuestro caso “desvío”) en el modelo binomial se “reduce” a entender este mismo evento en los modelos uniformes. Ahora bien, para estudiar desvíos en el modelo uniforme $\mathbf{G}(n, m)$, Goldschmidt, Griffiths y Scott usaron la teoría de Martingalas bajo la luz del proceso Erdős-Rényi que induce de forma natural una filtración (lo que es requerido en esta teoría). En este punto del argumento, la herramienta principal es la Desigualdad de Freedman (o “tipo” Freedman) que, esencialmente, nos permite estimar la

⁸Esta filtración puede verse como los eventos al revelar las aristas seleccionadas hasta la posición i -ésima.

⁹En [21] los autores no solo consiguen una tal representación martingala para el conteo de triángulos, sino que consiguen tal representación para el conteo de cualquier grafo fijo H .

¹⁰De hecho, la técnica nos permite estimar (1.4) para valores de α entorno a n^3 (para el caso denso) pero con menor precisión, dando su valor salvo constantes (en lugar de valores asintóticos exactos). Esto lo veremos con más detalles en el Capítulo 4, donde además se extiende el estudio al caso disperso.

probabilidad de desvío (de un martingala) en términos de ciertas cantidades estadísticas como la varianza (del proceso), la magnitud máxima de los incrementos, que denotamos por R , además del propio desvío considerado.

Pues bien, en el caso disperso que nos interesa mejorar, el valor de R puede ser demasiado grande, lo que es un obstáculo cuando queremos aplicar la desigualdad de Freedman a nuestro desvío “esperado”¹¹. Intuitivamente lo que sucede es que si bien los incrementos pueden ser muy grandes estos eventos son poco probables. Así esperamos que, en el caso que ocurran, estos no tengan una contribución en el desvío mayor que la explicada por el resto del proceso. Desde un punto de vista técnico, una de las maneras de formalizar esto es a través del método de truncamiento, o mejor conocido para este tipo de problemas como *Divide and Conquer* (en inglés)¹². Al truncar los incrementos en la magnitud “correcta” obtenemos un nuevo proceso (super-martingala) en el que la estrategia anterior descrita (usando la desigualdad de Freedman) se aplica satisfactoriamente. Resta estimar el resto del proceso, esto es, la contribución de todos los incrementos de magnitud “grande”. Afortunadamente esta “grandeza” es tal que, en este punto, nos enfrentamos a grandes desvíos, en lugar de desvíos moderados, lo que nos permite evocar argumentos de esta área de la Combinatoria Probabilística. Esta parte será descrita sucintamente a lo largo de este trabajo por ser muy técnica, mas a groso modo, como estamos esencialmente contando triángulos “atípicos”¹³ lo que se hará será clasificarlos de acuerdo a cuán atípicos son, lo que nos permitirá controlar cada contribución por separado.

Todo el trabajo descrito anteriormente nos permitirá entender por completo (salvo factores constantes en el desvío) el orden de magnitud del desvío del conteo de triángulos en $\mathbf{G}(n, m)$ para $t = m/N = \Omega(n^{-1/2}(\log n)^{1/2})$ a lo largo del desvío estándar hasta el máximo desvío posible (esto es, $t^{3/2}n^3$), lo cual es sorprendente considerando que inicialmente nuestro objetivo era controlar el desvío hasta la orden del valor esperado (esto es, t^3n^3). Como consecuencia, cuando aplicamos estos resultados al modelo $\mathbf{G}(n, p)$ (con $p \gg n^{-1/2}(\log n)^{1/2}$) conseguimos extender el resultado de Doering y Eichelsbacher [13, 14] para toda la región “no localizada” dando un resultado exacto (asintótico) para la función de tasa. Mientras que en la región “localizada” conseguimos cubrir todo el desvío del número de triángulos en $\mathbf{G}(n, p)$, con el costo de un factor constante en la función de tasa. Recalamos que, si bien este trabajo usa el modelo uniforme $\mathbf{G}(n, m)$ como “puente” para llegar al modelo binomial $\mathbf{G}(n, p)$, ambos modelos de grafos aleatorios son de importancia en Combinatoria Probabilística, así este trabajo puede ser visto como un estudio de ambos modelos, $\mathbf{G}(n, m)$ y $\mathbf{G}(n, p)$.

¹¹Estos desvíos ‘esperados’ tienen su justificación heurística en el fenómeno conocido como “localización”.

¹²Esta idea que incorporamos a la estrategia de Goldschmidt, Griffiths y Scott es una de las más importantes y servirá como eje de nuestro análisis.

¹³‘Atípico’ quiere decir para nosotros que su existencia es poco probable. En este caso, un triángulo es “atípico” por causa de una de sus partes (vértice y/o arista) tener un alto desvío asociado (grado o cogrado).

Capítulo 2

Preliminares

En este capítulo proporcionaremos al lector las definiciones y resultados esenciales que serán usados en el resto de este trabajo. Para una mejor presentación, dividiremos este capítulo en secciones que agrupan distintas áreas de la matemática.

Sin embargo, como el lector observará, en este trabajo estas diversas áreas se “entrelazan”, lo que se ve reflejado en cierta interdependencia entre las secciones.

2.1. Notación Asintótica

Un grupo de notaciones ubicuas en este trabajo será la notación asintótica, por eso consideramos adecuado ya introducirla al inicio.

Existen diversos símbolos y variantes que se usan en este lenguaje, sin embargo la idea detrás es esencialmente la misma, a saber, reducir/simplificar la carga de parámetros que aparecen al escribir desigualdades donde nos permitimos colocar cantidades muy grandes o muy pequeñas.

Notación Big O

Sea $f(n)$ y $g(n)$ dos secuencias reales (positivas) que dependen de $n \in \mathbb{N}$.

Escribimos $f(n) = O(g(n))$ cuando existe una constante $C > 0$ (suficientemente grande) tal que

$$f(n) \leq C g(n).$$

Una notación alternativa será escribir $f(n) \lesssim g(n)$. Similarmente, si existe una constante

$c > 0$ tal que

$$f(n) \geq c g(n).$$

Alternativamente escribimos $f(n) = \Omega(g(n))$. En algunas oportunidades también escribiremos $f(n) = \Theta(g(n))$, o simplemente $f(n) \asymp g(n)$, para referirnos a ambas desigualdades simultáneas, esto es, $cg(n) \leq f(n) \leq Cg(n)$ para ciertas constantes $c, C > 0$.

El lector prevenido verá que a lo largo de los capítulos escribimos enunciados de la forma “si $f(n) = \Omega(g(n))$, entonces cierta propiedad dependiente de f y g ocurre”. En estas expresiones abusamos ligeramente la notación para decir que “si $f(n) \geq Cg(n)$ para $C > 0$ suficientemente grande, entonces cierta propiedad (que depende de f y g) es válida”. Lo mismo para $O(\cdot)$.

También usaremos notación “en el límite” (cuando n tiende a ∞). Escribiremos que $f(n) = \omega(g(n))$ si

$$\lim_{n \rightarrow \infty} f(n)/g(n) = \infty.$$

Una notación alternativa será escribir $f(n) \gg g(n)$. Complementariamente escribiremos $f(n) = o(g(n))$ si $g(n) \gg f(n)$. Alternativamente escribimos $f(n) \ll g(n)$.

Como advertimos al inicio, no toda esta notación es estándar (depende que área de la matemática que la utiliza) y más aún, existen generalizaciones y/o variantes. Una de estas es la ‘notación asintótica probabilística’, donde las desigualdades tienen cierta información probabilística adicional. Esto lo veremos mejor en la sección de Probabilidades (de este capítulo).

2.2. Teoría de Grafos

Cuando nos referimos a un grafo G , nos estamos refiriendo a un objeto combinatorio que se define como un par ordenado $(V(G), E(G))$ donde $V(G)$ es un conjunto (finito) y $E(G)$ es un subconjunto de $\binom{V(G)}{2}$. Fijado un grafo G , los elementos de $V(G)$ son llamados vértices y los elementos de $E(G)$ son llamados de aristas. Por definición, las aristas “ e ” son conjuntos de tamaño 2, digamos $e = \{u, w\}$, así que convendremos en llamar a estos vértices como *endpoints* de la arista e . Adicionalmente, para simplificar la escrita, cuando nos referimos a aristas escribiremos uw en lugar de $\{u, w\}$.

Existen algunos ejemplos importantes de grafos que pasaremos a describir pues además de ilustrativos, estos serán mencionados a lo largo de este trabajo. Un k -clique (o grafo completo) es un grafo, que denotamos por K_k , donde el conjunto de vértices tiene cardinalidad k y todo par (distinto) de vértices es una arista, esto es, $E(K_k) = \binom{V(K_k)}{2}$.

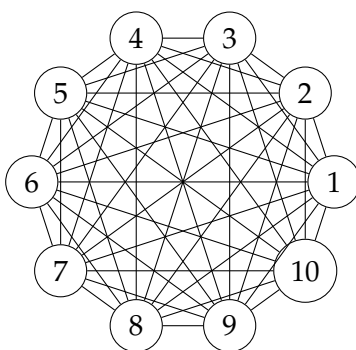


Figura 2.1: Esta figura ilustra un grafo completo con 10 vértices (los vértices se encuentran etiquetados).

Otro ejemplo importante, que define toda una clase de grafos, son los grafos bipartidos. Un grafo bipartido G con partes A y B (disjuntos) es un grafo cuyo conjunto de vértices es $A \cup B$ y cuyas aristas $e \in E(G)$ están restringidas a tener un endpoint en A y el otro en B . Similarmente, diremos que un grafo bipartido es completo (o *biclique*, en inglés) si tiene todas las aristas posibles preservando la propiedad de bipartición, esto es, un grafo bipartido con partes A y B es completo si el conjunto de aristas coincide con $\{uw : u \in A \text{ y } w \in B\}$.

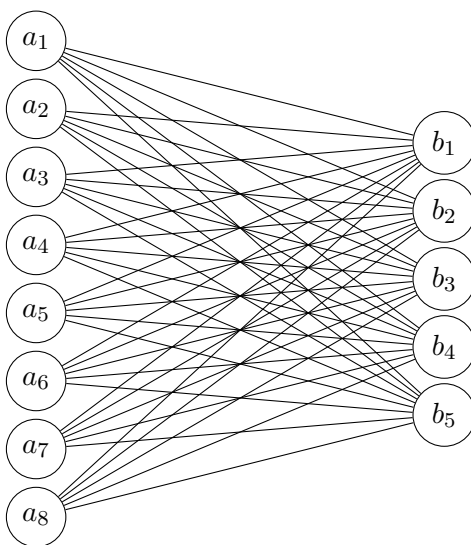


Figura 2.2: La figura ilustra un *biclique* con partes A y B de tamaño 8 y 5, respectivamente. Por comodidad, hemos denotado por a_i los elementos de A y por b_j los de B .

Otro ejemplo importante son los emparejamientos (*matchings*). Como el nombre sugiere un matching G es un grafo donde el conjunto de aristas $E(G)$ la propiedad de no intersección, esto es, para cada par distinto $e, f \in E(G)$ se cumple que $e \cap f = \emptyset$. Como último ejemplo, tenemos las estrellas. Una estrella de tamaño n , denotado por S_n es un biclique donde una de las partes es unitaria (singletón). De antemano, también hacemos notar al lector que por simplicidad y estética, usaremos los símbolos \triangle (triángulo) y \wedge (cereza) para referirnos al grafo completo K_3 y a la estrella S_2 , respectivamente.



Figura 2.3: Un matching (izquierda) y una estrella (derecha) formadas por tres aristas cada una.

Definidos los objetos (grafos) de los que disertaremos, pasaremos a definir ciertas relaciones que conciernen este estudio. Un morfismo en nuestra ‘categoría’ de grafos, es un mapa que va de un grafo H en un grafo G tal que preserva conectividad (estructura). En otras palabras, un homomorfismo $\psi : H \rightarrow G$ es un mapa de $V(H)$ en $V(G)$ tal que

$$\psi(u)\psi(w) \in E(G)$$

siempre que $uw \in E(H)$. Si el homomorfismo ψ es adicionalmente inyector (como mapa), diremos que se trata de un endomorfismo. En este trabajo, optaremos por seguir la notación de [21], así denotaremos por $N_H(G)$ al número de endomorfismos de H en G ¹.

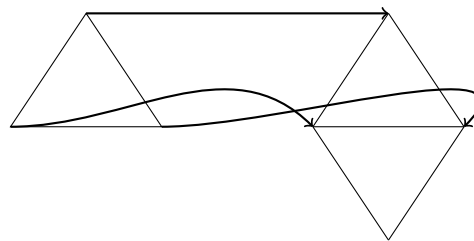


Figura 2.4: Diseño de un endomorfismo que ‘sumerge’ un triángulo en un grafo *diamante*.

¹El lector debe notar la sutil diferencia entre la imagen de un endomorfismo (que es un grafo) con el endomorfismo en sí (que es un mapa).

Dado dos grafos G' e G , escribiremos $G' \subseteq G$ si y sólo si

$$V(G') \subseteq V(G) \text{ y } E(G') \subseteq E(G),$$

o en otras palabras, cuando el mapa inclusión es un endomorfismo. En tal caso, diremos que G' es un subgrafo de G . Note que la relación “ \subseteq ” entre grafos es una relación de orden parcial (reflexiva). Con esto en mente, podemos hablar de objetos maximales. Por ejemplo, dado un grafo G y un subconjunto de vértices A , denotamos por $G[A]$ el grafo maximal $H = (A, E(H))$ con respecto a la propiedad $H \subset G$. Una forma alternativa de decir esto es que $G[A]$ es el grafo con conjunto de vértices A tal que $ab \in \binom{A}{2}$ es una arista si y solo si $ab \in E(G)$. El grafo $G[A]$ es llamado el subgrafo (de G) inducido por A . En el mundo bipartido tenemos el análogo de $G[A]$. Sean A y B dos subconjuntos disjuntos de vértices (de G), denotamos por $G[A, B]$ el grafo bipartido con partes A y B que contiene todas las aristas posibles de G con endpoints en A y B (ambos).

Para referirnos a la clase (o familia) de grafos con n vértices usaremos la notación \mathcal{G}_n . Si además nos restringimos a los grafos con m aristas, entonces usaremos la notación $\mathcal{G}_{n,m}$.

Ahora definiremos las propiedades intrínsecas de un grafo. Sea pues G un grafo. Dado un vértice v de G definimos la *vecindad de v* (en G) como el subconjunto de vértices

$$N_G(v) := \{w \in V(G) : vw \in E(G)\}.$$

La cardinalidad de $N_u(G)$, denotado por $d_u(G)$, es llamado el *grado de u* (en G). De forma similar, dado un par distinto de vértices uv , definimos la *vecindad común de uv* como siendo

$$N_{uv}(G) := \{w \in V(G) : vw \in E(G) \text{ y } uw \in E(G)\}.$$

La cardinalidad de $N_{uv}(G)$, denotado por $d_{uv}(G)$, es llamado el *cogrado de uv* (en G). Note que se cumple la relación $N_{uv}(G) = N_u(G) \cap N_v(G)$. La secuencia de grados

$$(d_v(G) : v \in V(G))$$

es una información importante del grafo, y dentro de ella tenemos tres parámetros fundamentales, a saber, el grado máximo, el grado mínimo y el grado medio. Aquí sólo necesitaremos introducir el *grado máximo* $\Delta(G)$, que como el nombre sugiere es dado por

$$\Delta(G) := \max_{v \in V(G)} d_v(G).$$

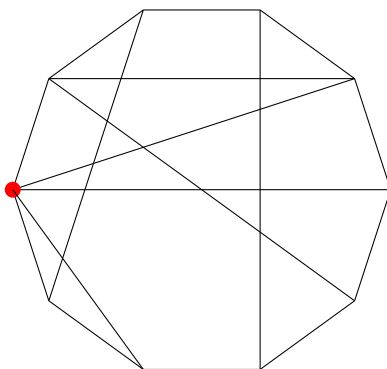


Figura 2.5: La figura muestra un grafo con grado máximo 5, donde el único vértice con dicho grado es colorido de rojo.

La teoría pura de grafos es un vasto campo con muchos resultados y áreas aguardando nuevos frutos. Nuestro objeto de estudio en este trabajo es un híbrido entre teoría de grafos y probabilidades, con un mayor foco en las probabilidades. Sin embargo, ciertos resultados de la teoría de grafos serán necesarios para conseguir nuestros resultados. Aquí presentaremos sólo estos hechos necesarios y dejamos para el lector curioso, revisar las referencias [4, 12] para profundar en grafos.

Lema 2.1 (Lema Star-Matching). *Para cualquier grafo $G \in \mathcal{G}_{n,m}$ se cumple que G contiene una estrella (star) o un emparejamiento (matching) de tamaño $\Omega(\sqrt{m})$.*

Note que en este lema, la información del número de vértices del grafo no es relevante.

2.3. Probabilidades

En esta sección introduciremos brevemente las nociones y resultados probabilísticos que usaremos de forma natural en este trabajo. Recalamos al lector que aquí estudiaremos los grafos “aleatorios” en lugar de sólo grafos. Así, la teoría se torna más rica en estructura y técnicas. También aclaramos que a partir de esta sección se requerirá un cierto conocimiento (básico) de análisis matemático, teoría de la medida (abstracta) y, posiblemente, algo de combinatoria.

Espacio Probabilístico

Así pues, diremos que un espacio medible $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio probabilístico si la medida \mathbb{P} satisface que $\mathbb{P}(\Omega) = 1$. Este espacio de probabilidades será usado para des-

cribir los “eventos” de nuestro interés y la probabilidad “ \mathbb{P} ” nos servirá para medirlos. De este modo, como es frecuente en probabilidades (a diferencia de la teoría de la medida), muchas veces este espacio será considerado tácito y lo suficientemente robusto para abarcar todas los objetos estocásticos que nos interesen (y sus relaciones). Este punto de vista es tratado en detalle en un curso “fundacional” de las probabilidades (vea por ejemplo [27]). Sin embargo, no debemos descartar que en algunos momentos un probabilista puede “cambiar” su punto de vista nuevamente y volver al espacio tácito para trabajar con él, por ejemplo como sucede en la teoría de los grandes desvíos (vea [18]) donde el Teorema de Radon-Nikodym es crucial.

Variable Aleatoria

La siguiente noción esencial es de variable aleatoria². Pues bien, una variable aleatoria \mathbf{X} no es otra cosa que una función medible definida en el espacio probabilístico $(\Omega, \mathcal{F}, \mathbb{P})$. Note que los valores de \mathbf{X} no están restringidos a ser números. Denotemos por \mathcal{X} el espacio (medible) de llegada de \mathbf{X} , esto es $\mathbf{X} \in \mathcal{X}$ (*casi ciertamente*). Es usual denominar a estas variables aleatorias por \mathcal{X} -rv, por ejemplo, si $\mathbf{X} \in \mathbb{R}$ podemos decir que \mathbf{X} es una \mathbb{R} -rv. También será permitido usar nominaciones en lugar de símbolos, así en el ejemplo anterior, podemos decir que \mathbf{X} es una variable aleatoria *real*. Otro ejemplo importante es cuando \mathcal{X} es $\mathcal{G}_{n,m}$ o \mathcal{G}_n . En estos casos, decimos que \mathbf{X} es un grafo aleatorio. Los ejemplos son innumerables, pero antes de detenernos daremos mas un ejemplo ilustrativo. Suponga que en lugar de escoger un grafo de forma aleatoria, escogemos las aristas de forma aleatoria (digamos una por una). De esta manera, obtenemos una secuencia de grafos aleatorios, esto es, $\mathbf{X} = (\mathbf{X}_{n,m})_{m=1}^{\binom{n}{2}}$. Aquí el espacio de llegada es un *arreglo* de grafos (con relación entre ellos), dicho arreglo puede ser pensado en $\mathcal{X} = \prod_{m=1}^{\binom{n}{2}} \mathcal{G}_{n,m}$. Tales ejemplos de variables aleatorias son de suma importancia y serán introducidos a continuación formalmente.

Un proceso aleatorio $\mathbf{M} = (\mathbf{M}_t)_{t \in T}$ indexado en T (que podemos pensarlo como el “tiempo”) es una secuencia de variables aleatorias \mathbf{M}_t (definidas en el mismo espacio probabilístico). El lector debe notar que no estamos imponiendo ninguna restricción para los “estados” \mathbf{M}_t , ni siquiera independencia³.

²Sinónimos para aleatoria son randómica, estocástica y probabilística. Aquí nos tomaremos la libertad de usar cualquier de estos sinónimos.

³Brevemente, desde el punto de vista de la teoría de la medida, dos variables aleatorias \mathbf{X} y \mathbf{Y} son independientes si la distribución (vea siguiente sección) de la variable (\mathbf{X}, \mathbf{Y}) es el producto de sus distribuciones marginales, esto es, $\mu_{(\mathbf{X}, \mathbf{Y})} = \mu_{\mathbf{X}} \otimes \mu_{\mathbf{Y}}$. Podemos extender esta definición a eventos via variables indicadoras.

Procesos Estocásticos : Martingalas

Una noción importante asociada a la de proceso aleatorio es la de filtración. Suponiendo que el conjunto T (*index set*) posee un orden total (como “el tiempo”), una filtración $(\mathcal{F}_t)_{t \in T}$ es una secuencia de sub- σ -álgebras de \mathcal{F} con la particularidad que $\mathcal{F}_s \subseteq \mathcal{F}_t$ siempre que $s \leq_T t$. Informalmente, podemos pensar en una filtración como la información aprendida en una experiencia (aleatoria) conforme el tiempo t avanza. Así pues, un proceso aleatorio $\mathbf{M} = (\mathbf{M}_t)_{t \in T}$ (con index set ordenado) posee una filtración, llamada *natural*, definida por $\mathcal{F}_t := \sigma(\mathbf{M}_s : s \leq t)$. Como el lector puede observar, esta definición de filtración natural de un proceso aleatorio, se corresponde con la idea de información aprendida hasta un tiempo t (si todo lo que deseamos saber es el resultado del proceso). En general, decimos que un proceso estocástico $(\mathbf{M}_t)_{t \in T}$ está *adaptado* a una filtración $(\mathcal{F}_t)_{t \in T}$ si \mathbf{M}_t es \mathcal{F}_t -medible para todo $t \in T$.

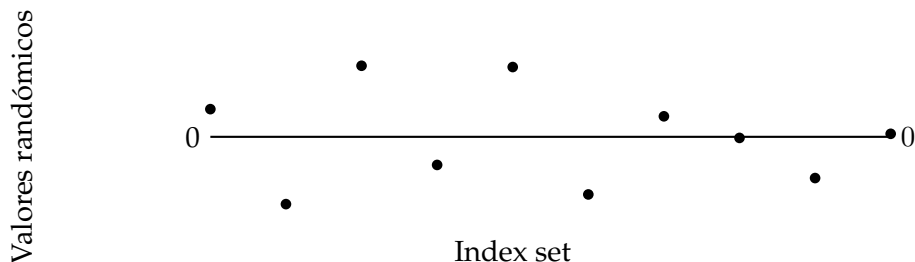


Figura 2.6: La figura ilustra una parte de una realización de un proceso aleatorio (indexado por tiempo discreto).

Uno de los procesos más estudiados y que será relevante a lo largo de este trabajo son las *martingalas* (o *martingales*). Intuitivamente, una martingala es una apuesta ‘justa’, esto significa que ‘en esperanza’ el jugador quedará con la misma cantidad de dinero al realizar la siguiente apuesta. Para formalizar esto, recurriremos al lenguaje introducido anteriormente. Un proceso aleatorio (discreto)⁴ $(\mathbf{M}_t)_{t \in \mathbb{Z}_{\geq 0}}$ adaptado a una filtración $(\mathcal{F}_t)_{t \in \mathbb{Z}_{\geq 0}}$ es una *martingala* se satisface que, para todo tiempo (discreto) $t \geq 0$,

$$\mathbb{E}[\mathbf{M}_{t+1} \mid \mathcal{F}_t] = \mathbf{M}_t.$$

Otra formulación equivalente, que usaremos en este trabajo, es mediante los ‘incrementos’. Denote por

$$\mathbf{X}_t := \mathbf{M}_t - \mathbf{M}_{t-1}.$$

⁴Imponemos la restricción de “discreto” en nuestra presentación sólo por simplicidad. De hecho, la definición se puede extender para tiempos continuos.

Estas variables aleatorias son llamadas *incrementos* por razones obvias. Note que la propiedad de martingala es equivalente a decir que

$$\mathbb{E}[\mathbf{X}_{t+1} | \mathcal{F}_t] = 0, \quad (2.1)$$

para todo tiempo discreto $t \geq 0$. De forma similar podemos definir las *supermartingalas* y *submartingalas* simplemente cambiando la igualdad “=” en (2.1) por las desigualdades “ \leq ” y “ ≥ 0 ”, respectivamente.

A lo largo de los capítulos centrales, veremos que no hacemos uso de la notación ‘por σ -álgebras’. Esto sucede por que preferimos adoptar la notación ‘por variable aleatoria’, debido a la naturaleza de nuestras álgebras. Específicamente, nuestras σ -álgebras $(\mathcal{F}_t)_{t \in \mathbb{Z}_{\geq 0}}$ serán generadas (naturalmente) por variables aleatorias \mathbf{Y}_t (con $t \in \mathbb{Z}_{\geq 0}$). Así, optaremos por escribir $\mathbb{E}[\mathbf{X}_{t+1} | \mathbf{Y}_1, \dots, \mathbf{Y}_t]$ en lugar de $\mathbb{E}[\mathbf{X}_{t+1} | \mathcal{F}_t]$.

Distribuciones : Binomial e Hipergeométrica.

Cuando nos referimos a la distribución (o *law*) de una variable aleatoria $\mathbf{X} \in \mathcal{X}$, nos estamos refiriendo a la medida natural en \mathcal{X} asociada a esta variable. En otras palabras, dada una variable aleatoria \mathbf{X} , podemos definir la medida (probabilística)

$$\mu_{\mathbf{X}}(A) := \mathbb{P}(\mathbf{X} \in A)$$

para cada conjunto mensurable A de \mathcal{X} . El lector familiarizado con Teoría Ergódica, ya habrá notado que $\mu_{\mathbf{X}}$ no es otra cosa que el *pushforward* de la medida \mathbb{P} mediante \mathbf{X} (denotado por $\mathbf{X}^*\mathbb{P}$). Otras notaciones estandarizadas de $\mu_{\mathbf{X}}$ también son $\mathbb{P}_{\mathbf{X}}$ y $\mathbb{P}_{x \sim \mathbf{X}}$. La relación de distribución sera denotada por \sim , esto es, al escribir $\mathbf{X} \sim \mu$ queremos decir que la distribución de \mathbf{X} es μ . No es difícil verificar que $\mu_{\mathbf{X}}$ es de hecho una medida probabilística en el espacio mensurable \mathcal{X} , así que lo dejaremos como ejercicio⁵. Existen muchas distribuciones importantes, las más básicas con valores reales, pero no restrictas a estos estados. La distribuciones de Dirac son las más simples pero no irrelevantes. estas son dadas de la siguiente manera. Sea x un elemento de \mathcal{X} , asociamos la medida probabilística $\delta_x(A) := 1_{x \in A}$.

A pesar del aspecto inocente de estas medidas, la combinación convexa de estas medidas produce la importante clase de medidas llamadas *empíricas*. En otras palabras, una

⁵Lo que no es fácil es la dirección opuesta, esto es, toda medida probabilística en un espacio ‘adecuado’ es la distribución de alguna variable aleatoria. Esto será admitido aquí como verdad, y puede ser estudiado al detalle en un curso de teoría de probabilidades.

medida (probabilística) en un espacio \mathcal{X} es llamada empírica si tiene la forma

$$\sum_x w_x \delta_x$$

donde x varia en un subconjunto finito de \mathcal{X} , y $w_x \geq 0$ son pesos con suma total cero. No es difícil convencerse que si el espacio \mathcal{X} es finito, entonces toda medida probabilística es una medida empírica. En general, si nuestro espacio es lo suficientemente rico, podemos aproximar (en cierto sentido) toda medida probabilística por medidas empíricas, pero nosotros no abordaremos estos tópicos (que se cubren en un curso de teoría ergódica). Pues bien, casos particulares de estas medidas o distribuciones son las de *Bernoulli* que modelan monedas (que pueden estar viciadas). Fijado un número $p \in [0, 1]$, la distribución $\text{ber}(p)$ es definida formalmente como la medida real $p\delta_1 + (1 - p)\delta_0$. La siguiente distribución de interés es conocida como la distribución *Binomial*. Intuitivamente, esta distribución binomial es la distribución asociada al número de sucesos de un experimento que se repite muchas veces de forma independiente.

0 0 1 1 0 1 1 1 1 1

Figura 2.7: En este arreglo podemos ver una experiencia de lanzar 10 monedas independientes (cara= 1, sello= 0).

Siendo más formales, la distribución binomial $\text{Bin}(n, p)$ con parámetros n y p es la distribución asociada a la variable aleatoria

$$\sum_{i=1}^n \xi_i$$

donde las variables aleatorias ξ_i son independientes e idénticamente distribuidas con $\xi_i \sim \text{ber}(p)$.

Algunas de las variables aleatorias que estudiaremos tendrán un comportamiento parecido a la distribución binomial, pero con una cierta carga de dependencia. A seguir describiremos este tipo de variables y enunciaremos sus respectivos resultados de concentración que serán utilizados más adelante. Eucarística-mente, una variable aleatoria Hipergeométrica es una variable aleatoria discreta que describe la probabilidad de k ‘sucesos’ en una selección de n individuos ‘sin reemplazo’ de una población total de N individuos donde exactamente K individuos son considerados como ‘sucesos’⁶.

⁶Note que si removemos la condición de ‘sin reemplazo’ de la definición, lo que obtenemos es una distribución binomial de parámetros n y $p = K/N$. Así, ambas distribuciones están extremadamente relacionadas como veremos más adelante, sin embargo no son exactamente iguales.

Así, una variable aleatoria discreta X tiene distribución Hypergeometric(N, K, n) si

$$\mathbb{P}(X = k) := \frac{\binom{K}{k} \times \binom{N-k}{n-k}}{\binom{N}{n}},$$

cuando $k \in [K]$.

Desigualdades de Concentración

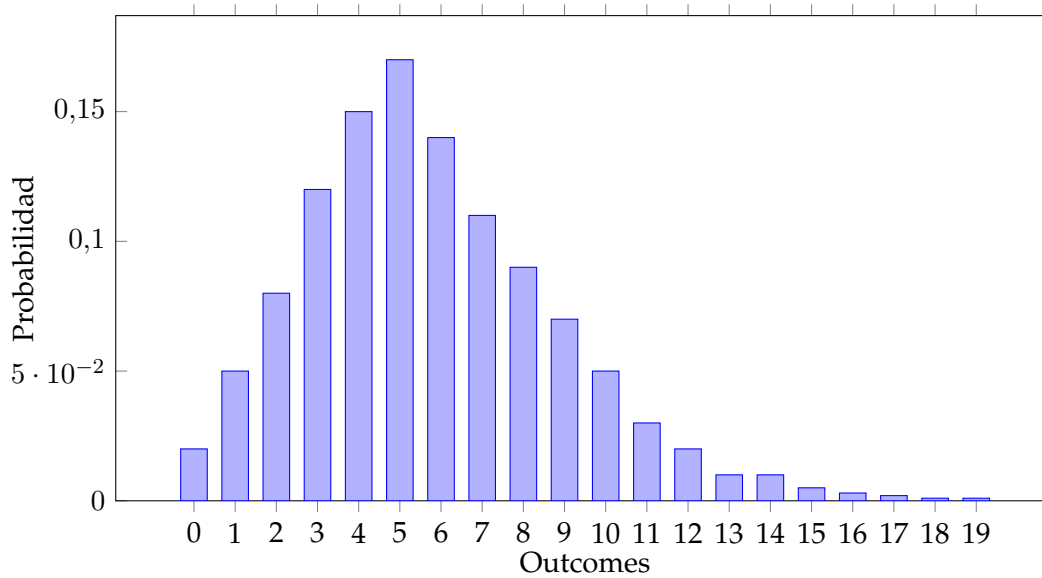


Figura 2.8: La imagen ilustra la distribución de una variable aleatoria. Aquí podemos contemplar el comportamiento de la variable, en particular, sus desvíos.

Un importante tópico en la teoría de probabilidades son las “desigualdades de concentración” (*Concentration Inequalities*). Estas desigualdades nos permiten estimar la probabilidad de que una variable aleatoria (numérica) se desvíe una cierta cantidad de su esperanza. Los desvíos pueden ser positivos o negativos, en cuyo caso nos referimos a estas estimativas como *upper tail* (cola superior) y *lower tail* (cola inferior). La utilidad de estas estimativas reside en que se hayan expresadas en función de valores estadísticos (relativamente) simples de computar : esperanza, varianza, supremos, y otros parámetros similares. Nosotros estamos interesados en estimativas exponenciales (o mejor dicho con decaimiento exponencial), pero antes mencionaremos la desigualdad más fundamental, esto es, la desigualdad de Markov (o primer momento). Sea pues $X \geq 0$ una variable

aleatoria con valores positivos. Entonces, para todo $a > 0$ se cumple que

$$\mathbb{P}(\mathbf{X} \geq a) \leq \frac{\mathbb{E}[\mathbf{X}]}{a}.$$

Note que si $a \gg \mathbb{E}[\mathbf{X}]$, entonces la probabilidad de desvío (positivo) converge para '0'. Note que esta probabilidad tiene decaimiento lineal, esto es, es limitada por la inversa de una función lineal. Sorprendentemente, aplicando este resultado de forma inteligente (más cierto trabajo analítico), muchas desigualdades con mejores decaimientos pueden ser deducidos, incluyendo los decaimientos exponenciales (pero de lejos todas).

A seguir, enunciaremos las desigualdades de concentración, conocidas como "tipo Chernoff", que obedecen las distribuciones binomial e hipergeométricas.

Teorema 2.1 (Desigualdades de Concentración para distribución Binomial e Hipergeométrica). *Sea \mathbf{X} una variable aleatoria con distribución binomial o hipergeométrica. Para todo real positivo a , se cumplen las siguientes desigualdades:*

- [Cola superior]

$$\mathbb{P}(\mathbf{X} \geq \mathbb{E}[\mathbf{X}] + a) \leq \exp\left(\frac{-a^2}{2\mathbb{E}[\mathbf{X}] + 2/3 a}\right).$$

- [Cola inferior]

$$\mathbb{P}(\mathbf{X} \leq \mathbb{E}[\mathbf{X}] - a) \leq \exp\left(\frac{-a^2}{2\mathbb{E}[\mathbf{X}]}\right).$$

Adicionalmente, tenemos la siguiente desigualdad para grandes desvíos (positivos): para cada $\theta \geq e$, se cumple que

$$\mathbb{P}(\mathbf{X} \geq \theta \mathbb{E}[\mathbf{X}]) \leq \exp(-\theta \mathbb{E}[\mathbf{X}] \ln(\theta/e)).$$

Desigualdades de Concentración para Martingalas

Como pudimos ver a groso modo en la introducción, un paso fundamental de nuestro enfoque es representar nuestra variable aleatoria como una martingala⁷. Esto se debe

⁷Por un resultado de Doob (dada una variable aleatoria y una filtración) podemos (bajo simples condiciones) representar esta variable como una martingala. Esto torna la teoría robusta. Sin embargo, no siempre es fácil entender estos incrementos. Así, descripciones concretas, como la que daremos aquí, suelen ser útiles.

a que el comportamiento de una martingala puede ser entendido conociendo sólo cierta información estadística de sus incrementos (específicamente, el supremo, esperanza y varianza condicionada). En particular, tenemos la siguiente desigualdad debida a Freedman [19] y que nos permite controlar los desvíos de la martingala eficientemente.

Teorema 2.2 (Desigualdad de Freedman). *Sea $(\mathbf{S}_m : m = 0, 1, \dots)$ una (super)martingala con incrementos $(\mathbf{X}_m : m = 1, 2, \dots)$ adaptado a la filtración $(\mathcal{F}_m : m = 0, 1, \dots)$. Asuma que los incrementos están limitados superiormente, esto es, para todo $m \geq 1$*

$$\mathbf{X}_m \leq R$$

donde R es algún número positivo. Denote por

$$V_m := \sum_{i=1}^m \mathbb{E} [|\mathbf{X}_i|^2 | \mathcal{F}_{i-1}].$$

Entonces, para cualquier $\alpha > 0$ y $\beta > 0$ se cumple la siguiente desigualdad,

$$\mathbb{P}(\exists m \geq 0 : \mathbf{S}_m - \mathbf{S}_0 > \alpha \text{ y } V_m \leq \beta) \leq \exp\left(\frac{-\alpha^2}{2(\beta + R\alpha/3)}\right). \quad (2.2)$$

Esta desigualdad suele ser muy útil en desvíos entorno al desvío padrón (moderados), que es estimado por el parámetro β (parámetro que se espera altamente ‘predecible’). Infelizmente, para desvíos muy grandes una aplicación directa de este teorema no nos proporciona, en muchas situaciones, la estimación ‘óptima’.

Sin embargo, como veremos más adelante, una estrategia sumamente útil es *divide-and-conquer*, una idea en que se descompone a la variable aleatoria objetivo en dos partes, una de tipo martingala (normal) y otra estructural. Así, en este argumento más sofisticado, podemos atacar la parte “normal” por el teorema de Freedman, mientras que la parte estructural es analizada por separado (con otros métodos).

De esta forma, a pesar de la no aplicabilidad directa de este teorema para grandes desvíos, aún resulta sumamente útil en nuestro análisis.

Acoplamiento : Dominación Estocástica

El tema de acoplamiento (y dominación estocástica) es un tópico amplio y de suma importancia en probabilidades. Infelizmente no podemos abarcar este tema a profundidad, pero si daremos lo básico para aplicar en este trabajo.

Algunos comentarios antes de comenzar las definiciones. La teoría de acoplamientos (*coupling*) suele ser desarrollada en el lenguaje de la teoría de la medida (un segundo curso). Sin embargo, por brevedad, aquí optaremos por darla en términos de variables aleatorias (un punto de vista más probabilístico). También optaremos por definir los conceptos en espacios generales (*posets*), con el fin de capturar procesos que aparecen naturalmente en este trabajo, como lo son los grafos aleatorios (en un curso inicial, se suele empezar por acoplar medidas en los reales).

Dada una variable aleatoria \mathbf{X} , decimos que otra variable $\hat{\mathbf{X}}$ es una copia de \mathbf{X} si $\hat{\mathbf{X}}$ tiene la misma distribución (o law) que \mathbf{X} . Repare que en esta definición, no forzamos que las variables estén definidas en el mismo espacio probabilístico. Un par de variables aleatorias $(\hat{\mathbf{X}}, \hat{\mathbf{Y}})$ se dicen acopladas (*coupled*) si están definidas en el mismo espacio probabilístico. Así, dado un par de variables aleatorias \mathbf{X} y \mathbf{Y} , un acoplamiento $(\hat{\mathbf{X}}, \hat{\mathbf{Y}})$ de \mathbf{X} y \mathbf{Y} es un acoplamiento tal que $\hat{\mathbf{X}}$ y $\hat{\mathbf{Y}}$ son copias de \mathbf{X} y \mathbf{Y} , respectivamente. Lo interesante de los acoplamientos es que al ser generados, estos pueden ser construidos manteniendo de modo que existan relaciones útiles entre el par de variables obtenidas. Un ejemplo de esto, y a la vez una clase sumamente importante de acoplamientos son los *monótonos*, en los cuales los acoplamientos son generados tal que una variable ‘domina’ a la otra (como elementos de un *poset*⁸). Formalmente, sea (\mathcal{X}, \leq) un poset con \mathbf{X} y \mathbf{Y} dos variables aleatorias en \mathcal{X} (no necesariamente acopladas). Decimos que un acoplamiento $(\hat{\mathbf{X}}, \hat{\mathbf{Y}})$ de \mathbf{X} y \mathbf{Y} es monótono si $\hat{\mathbf{X}} \geq \hat{\mathbf{Y}}$ (casi ciertamente). Note que, en tal caso, si $A \subseteq \mathcal{X}$ es creciente, entonces de la monotonía del acoplamiento deducimos la siguiente desigualdad útil,

$$\mathbb{P}(\hat{\mathbf{X}} \in A) \geq \mathbb{P}(\hat{\mathbf{Y}} \in A). \quad (2.3)$$

En general, cuando dos variables aleatorias \mathbf{X} y \mathbf{Y} (no necesariamente acopladas) satisfacen la desigualdad (2.3) decimos que \mathbf{X} *domina estocásticamente a* \mathbf{Y} , y lo denotamos por $\mathbf{X} \succeq \mathbf{Y}$. Así, hemos mostrado un hecho simple pero fundamental, a saber, si dos variables \mathbf{X} y \mathbf{Y} pueden ser monotonamente acopladas entonces $\mathbf{X} \succeq \mathbf{Y}$. En cierto espacio de estados \mathcal{X} (como los finitos), un resultado conocido como Strassen’s Theorem nos dice que el recíproco es cierto, a saber, si $\mathbf{X} \succeq \mathbf{Y}$ entonces \mathbf{X} y \mathbf{Y} pueden ser monotonamente acoplados.

A pesar de lo abstracto de la exposición, ejemplos naturales aparecen ya en este trabajo de forma temprana. Como el lector atento podrá haber notado, al producir los grafos uniformes usando el proceso de Erdős-Rényi, obtenemos un acoplamiento monótono en

⁸Cuando nos referimos a *poset* queremos decir un conjunto (parcialmente) ordenado.

la familia de los grafos ordenados por inclusión “ \subseteq ”, esto es,

$$\mathbf{G}(n, m_1) \subseteq \mathbf{G}(n, m_2)$$

siempre que $m_1 \leq m_2$. En particular, $\mathbf{G}(n, m_1) \preceq \mathbf{G}(n, m_2)$. Existen muchos más ejemplos cuando $\mathcal{X} = \mathbb{R}$ y las variables son Bernoulli o Binomiales, pero por simplicidad de la exposición, iremos directamente al resultado principal de este apartado. Para esto, imagine que usted da un paso al frente (en un camino recto) siempre que al lanzar una moneda resulta uno, caso contrario se queda quieto. En el primer escenario usted usa la misma moneda $\text{ber}(p)$ en cada lanzamiento, pero en el segundo escenario le cambian la moneda basados en los resultados anteriores, y lo único que usted sabe es que la probabilidad de la moneda dar uno (sabiendo el pasado) no supera a p . Si lanzamos las monedas el mismo número de veces para ambas caminatas, ¿quién tiene más chances de llegar ante a la meta?. Como existe una mayor tendencia en el primer escenario de dar un paso al frente, es razonable esperar que aquí haya más chances de llegar a la meta que en el segundo escenario. Es este tipo de resultados que el siguiente lema formaliza.

Lema 2.2 (Upper-Tail Criterium para suma de indicadores). *Sea $(\mathbf{X}_i)_{i=1}^m$ un proceso estocástico donde cada \mathbf{X}_i es una variable de Bernoulli (indicador) y sea $(\mathbf{Y}_i)_{i=1}^m$ una secuencia de variables indicadoras independientes.*

Si se cumple que $[\mathbf{X}_i \mid \mathbf{X}_1 \dots \mathbf{X}_{i-1}] \preceq \mathbf{Y}_i$, esto es,

$$\mathbb{P}(\mathbf{X}_i = 1 \mid \mathbf{X}_1 \dots \mathbf{X}_{i-1}) \leq \mathbb{P}(\mathbf{Y}_i = 1) = p_i$$

para cada $i \in [m]$, entonces $\mathbf{X} = \sum_{i=1}^m \mathbf{X}_i \preceq \mathbf{Y} = \sum_{i=1}^m \mathbf{Y}_i$.

En particular, si $p_i = p$ (esto es, p independiente de i), entonces $\mathbf{X} = \sum_{i=1}^m \mathbf{X}_i \preceq \text{Bin}(m, p)$.

Notación Asintótica Probabilística

Como ya vimos al inicio de este capítulo, en muchas ocasiones resulta conveniente introducir cierto lenguaje asintótico que favorece tanto la lectura de los enunciados como la eucarística a la hora de probarlos (más aún, la prueba puede verse simplificada). En este apartado definiremos algunas de ellas que vienen acompañadas con la noción de probabilidad.

Definición 2.1 (Eventos con alta probabilidad). Sea $E = E_n$ un evento (que depende de $n \in \mathbb{N}$). Decimos que E acontece con alta probabilidad (o asintótica casi-ciertamente) si

$$\lim_n \mathbb{P}(E_n) = 1.$$

En estos casos también se usa la abreviación whp (o aas) por sus siglas en inglés.

Como pudimos ver al inicio de este capítulo, la notación “big O” simplifica la escrita de constantes. Aquí introduciremos su análogo probabilístico.

Definición 2.2. Para cada número positivo b definimos una relación de orden entre variables reales “ \lesssim^{*b} ” del siguiente modo. Dos variables reales \mathbf{X} y \mathbf{Y} satisfacen que $\mathbf{X} \lesssim^{*b} \mathbf{Y}$ se existen constantes $C, K > 0$ tales que

$$\mathbb{P}(\mathbf{X} > C \mathbf{Y}) \leq \exp(-Kb).$$

Una propiedad de simple verificación pero de interés en esta relación es su transitividad “ $\mathbf{X} \lesssim^{*b} \mathbf{Y}$ y $\mathbf{Y} \lesssim^{*b} \mathbf{Z}$ implica $\mathbf{X} \lesssim^{*b} \mathbf{Z}$ ”. También señalamos que en la expresión “ $\mathbf{X} \lesssim^{*b} \mathbf{Y}$ ” no prohibimos que alguna de las variables sea determinista (ni las dos a la vez, aunque esto último carece de interés). De forma natural, definimos la relación “ \gtrsim^{*b} ” como la relación inversa de “ \lesssim^{*b} ”. Estos análogos probabilísticos de la notación asintótica se pueden seguir mimetizando para las relaciones de orden restantes, pero aquí no nos serán de utilidad, así que lo dejamos como un ejercicio lúdico para el lector.

Antes de finalizar esta sección, como ayuda al lector, listaremos algunas propiedades elementales referentes a la teoría de probabilidades. Así, cuando usadas, nos referiremos a ella por su nombre indexado o por su etiqueta.

Proposición 2.1 (Propiedades Básicas en Probabilidades). Sean E_1, \dots, E_k y F eventos, sean $\mathbf{X}_1, \dots, \mathbf{X}_k$ y \mathbf{Z} \mathbb{R} -variables aleatorias y sea \mathcal{F} una σ -subálgebra (de Ω). Se cumplen las siguientes propiedades:

- (Union Bound) $\mathbb{P}\left(\bigcup_{i=1}^k E_i\right) \leq \sum_{i=1}^k \mathbb{P}(E_i)$.
- (Probabilidad y Esperanza) $\mathbb{P}(F) = \mathbb{E}[\mathbf{1}_F]$.

- (Linealidad de la Esperanza) $\mathbb{E} \left[\sum_{i=1}^k \mathbf{X}_i \right] = \sum_{i=1}^k \mathbb{E} [\mathbf{X}_i]$.

- (Propiedades de la Varianza Condicionada)

$$\text{Var}[\mathbf{Z} \mid \mathcal{F}] := \mathbb{E} [(\mathbf{Z} - \mathbb{E}[\mathbf{Z} \mid \mathcal{F}])^2 \mid \mathcal{F}] = \mathbb{E} [\mathbf{Z}^2 \mid \mathcal{F}] - \mathbb{E} [\mathbf{Z} \mid \mathcal{F}]^2.$$

De la definición inferimos que $\text{Var}[\mathbf{Z} \mid \mathcal{F}] \geq 0$ y $\text{Var}[\mathbf{Z} \mid \mathcal{F}] \leq \mathbb{E} [\mathbf{Z}^2 \mid \mathcal{F}]$.

2.4. Combinatoria Probabilística

En esta sección describiremos los modelos probabilísticos que serán estudiados en este trabajo (y que de hecho ya venimos mencionando desde el título), junto con un resultado de concentración en particular que será de utilidad para uno de nuestros modelos introducidos.

2.4.1. Modelos Probabilísticos Combinatorios

Una vez introducida la noción de variable aleatoria (y distribución), ya podemos pensar en la colección de todas las variables aleatorias que toman valores en un espacio de estados fijo, digamos \mathcal{X} . De hecho, como el lector atento habrá notado, podemos pensar de forma ‘alternativa’ en el conjunto de todas las distribuciones (medidas de probabilidad) del espacio fijo \mathcal{X} (suponiendo ciertas condiciones en el espacio, como finitud). Cada una de estas distribuciones pueden ser vistas como un *modelo probabilístico en \mathcal{X}* . Por simplicidad y abusando un poco de la notación, también diremos que una variables aleatoria $\mathbf{X} \in \mathcal{X}$ es un *modelo probabilístico con distribución μ* , recordando que existen diversas formas de generar variables aleatorias con la misma distribución μ . Cuando el espacio \mathcal{X} es discreto⁹ nos referimos a sus modelos probabilísticos como “combinatorios”. Esto es sólo un convenio que adoptaremos aquí y no pretendemos ‘axiomatizar’ la noción de combinatoria de forma alguna.

Existen algunos ejemplos de modelos probabilísticos discretos, que daremos de forma no-rigurosa.

- Sea $\mathcal{X} = 2^S$, donde S es un conjunto finito. Para cada elemento de S asociamos una moneda ($\text{ber}(p)$) y coleccionamos los elementos de S para los cuales la experiencia resultó 1. Tal conjunto es un subconjunto aleatorio de S y lo denotaremos por \mathbf{S}_p . Claramente, $\mathbf{S}_p \in \mathcal{X}$ es un modelo probabilístico conocido como *p-binomial random subset of S*, referenciando el tipo de distribución que posee.

⁹Aquí, esto significará ‘finito’, aunque a veces también se suele entender como ‘enumerable’.

- Similarmente, sea $\mathcal{X} = \binom{S}{m}$ donde $m \leq |S|$ es un entero no-negativo. Ahora vamos a sortear uniformemente un elemento de \mathcal{X} . Tal modelo probabilístico es denotada por \mathbf{S}_m y es conocido como *m-uniform random subset of S*. De forma alternativa, podemos generar una copia de \mathbf{S}_m mediante el siguiente proceso. Escoja un elemento de S uniformemente para posteriormente repetir este proceso m veces sin reposición. Así, obtenemos una secuencia s_1, \dots, s_m de elementos retirados. El conjunto aleatorio $S_m = \{s_1, \dots, s_m\}$ formado por estos elementos tiene la misma distribución que \mathbf{S}_m . Es importante recordar que tanto el modelo binomial como el uniforme están relacionados estrechamente bajo la transformación $p = m/|S|$.
- Sea $\mathcal{X} = \mathbb{F}^d$, donde \mathbb{F} es un cuerpo finito. Pensando en \mathcal{X} como un \mathbb{F} -espacio vectorial, construimos la variable aleatoria $\mathbf{x} = (x_1, \dots, x_d)$ donde los x_i 's son los resultados de barajar uniformemente elementos de \mathbb{F} y de forma independiente (con reposición).
- Como ejemplo final, generaremos permutaciones aleatorias. Esta será construida nuevamente por un proceso aleatorio como el del segundo ítem. Sea \mathcal{X} el conjunto de mapas 1-a-1 de $[m]$ en sí mismo (automorfismos). A cada paso i escogemos un elemento uniformemente del conjunto restante (sin reposición). Al terminar este proceso, tenemos una asignación aleatoria $i \rightarrow \sigma(i)$ que es inyectora, y por tanto biyectora (como bien se sabe). Así mismo, este modelo probabilístico tiene distribución uniforme dentro de \mathcal{X} y es llamado de *uniform random permutation*.

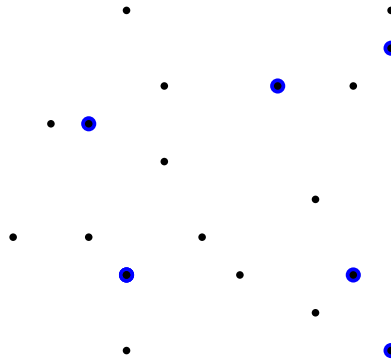


Figura 2.9: La imagen muestra una realización de un subconjunto aleatorio uniforme colorido en azul.

Ahora describiremos nuestros modelos aleatorios de grafos, junto con el proceso que los genera (acoplados), conocido como el proceso Erdős-Rényi. Estos también pudieron formar parte de los ejemplos, pero por su importancia en este trabajo, decidimos discutirlos aparte. Los dos modelos probabilísticos que estudiaremos tienen la peculiaridad que

los vértices están fijos (conjunto etiquetado) y son las aristas los objetos que serán escogidos de forma aleatoria. Denotemos por $\mathbf{G}(n, p)$ el grafo aleatorio donde (por simplicidad) $V(\mathbf{G}(n, p)) := [n]$ (determinístico) y $E(\mathbf{G}(n, p))$ es el subconjunto aleatorio p -binomial de $\binom{n}{2}$, esto es,

$$E(\mathbf{G}(n, p)) := \{ij : \xi_{ij} = 1\}$$

donde $(\xi_{ij} : ij \in \binom{n}{2})$ es una secuencia iid¹⁰ con distribución $\text{ber}(p)$. De forma semejante, denotaremos por $\mathbf{G}(n, m)$ el grafo aleatorio donde (por simplicidad) $V(\mathbf{G}(n, m)) = [n]$ y $E(\mathbf{G}(n, m))$ es el subconjunto aleatorio m -uniforme de $\binom{n}{2}$. Como vimos en los ejemplos de modelos (uniforme), este mismo modelo grafo aleatorio puede ser generado por un proceso estocástico de elecciones sin repetición, que por razones históricas lleva el nombre de proceso Erdős-Rényi: comenzamos con el conjunto vacío y a cada paso escogemos (uniformemente) una arista e_i de las restantes, así obtenemos una secuencia e_1, \dots, e_N (donde $N := \binom{n}{2}$). Así, el grafo aleatorio G_m formado por las m primeras aristas e_1, \dots, e_m tiene la misma distribución que $\mathbf{G}(n, m)$. Alternativamente, una permutación aleatoria σ sobre $[N]$ induce una copia de $\mathbf{G}(n, m)$ donde las aristas son $\{e_{\sigma(1)}, \dots, e_{\sigma(m)}\}$ donde $\{e_1, \dots, e_N\}$ era una ordenación (determinista) fija de las N -aristas totales (también denotado por \mathbf{G}_m , por tratarse de una construcción esencialmente equivalente)

Como el lector suspicaz percibirá, existen un gran espectro de modelos aleatorios para grafos. No es nuestra intención listar todos ellos pues aquí sólo nos referiremos a estos dos modelos. Lo que si haremos brevemente antes de pasar a la siguiente subsección será hablar brevemente de ellos. Una importante observación es notar que al condicionar el modelo binomial $\mathbf{G}(n, p)$ (con $p \in (0, 1)$) a un número fijo de aristas m , el grafo aleatorio generado es precisamente el uniforme $\mathbf{G}(n, m)$. En términos matemáticos, $[\mathbf{G}(n, p) \mid e(\mathbf{G}(n, p)) = m]$ es una copia de $\mathbf{G}(n, m)$. Si hacemos variar m en $\{0, 1, \dots, N\}$ podemos ver que los eventos " $e(\mathbf{G}(n, p)) = m$ " particionan el espacio de posibilidades. Así, por la ley total de probabilidades¹¹, tenemos que

$$\mathbb{P}(\mathbf{G}(n, p) \in P) = \sum_{m=0}^N \mathbb{P}(\mathbf{G}(n, m) \in P) \mathbb{P}(\text{Bin}(N, p) = m).$$

donde P es una clase de grafos (o 'propiedad' de grafos). Esta fórmula la podemos ver como una herramienta para 'transferir' resultados probabilísticos de un modelo al otro.

Específicamente, para obtener una estimativa en $\mathbf{G}(n, p)$ podemos ver esta como un 'promedio' (ponderado) de estimativas en el modelo uniforme. De hecho esta observación

¹⁰Escribiremos "iid" como abreviatura de "independiente e idénticamente distribuida".

¹¹Con esto, nos referimos a la siguiente propiedad fundamental de las probabilidades: si $\Omega = \bigcup_{i=1}^m E_i$ es una partición de Ω y F es un evento, entonces $\mathbb{P}(F) = \sum_{i=1}^m \mathbb{P}(F \mid E_i) \mathbb{P}(E_i)$.

la que nos permitirá deducir resultados sobre los desvíos en $\mathbf{G}(n, p)$ ¹².

Lo inverso también es válido, para obtener una estimativa en $\mathbf{G}(n, m)$ podemos estimar primero en $\mathbf{G}(n, p)$ y usar la desigualdad directa

$$\mathbb{P}(\mathbf{G}(n, m) \in P) \leq \frac{\mathbb{P}(\mathbf{G}(n, p) \in P)}{\mathbb{P}(\text{Bin}(N, p) = m)}$$

donde el valor de p se suele tomar entorno de m/N ¹³. Es por esta razón ambos modelos (con los parámetros correctos) suelen ser considerados ‘parecidos’ y se espera que satisfagan muchas propiedades probabilísticas mutuamente. Sin embargo, es bueno advertir que no es siempre este el caso. Basta notar que la variable aleatoria $e(\mathbf{G})$ puede tener distribución delta de Dirac para el modelo uniforme mientras que es binomial para el modelo binomial (valga la redundancia). Nosotros aprovecharemos este determinismo del modelo uniforme pues eventos que pueden verse afectados por el número de aristas se tornan relativamente fáciles de estudiar en el modelo uniforme, donde este fenómeno deja de acontecer.

Por último, introducimos la terminología “denso” y “disperso” para indicar que nuestros modelos aleatorios son escogidos con “densidad” constante $\Theta(1)$ o ínfima $o(1)$. Específicamente, para $\mathbf{G}(n, p)$ y $\mathbf{G}(n, m)$ decimos que son un modelo denso siempre que p y m/N son constantes, mientras que decimos que los modelos son dispersos siempre que p y m/N convergen a cero cuando n tiende a infinito (esto es, $p \ll 1$ y $m \ll N$). Esta noción de “dispersión” se puede extender a otros modelos combinatorios de forma natural, así que dejamos esto como ejercicio mental para el lector.

2.4.2. Aplicación de Martingalas en $\mathbf{G}(n, m)$

En la subsección sobre desigualdades de concentración vimos algunos resultados que nos permiten estimar la probabilidad de desvío de variables aleatorias $\mathbf{X} = f(\mathbf{X}_1, \dots, \mathbf{X}_m)$ donde $(\mathbf{X}_i : i \in [m])$ es una secuencia de variables independientes (o “casi” independientes) y f es la función “suma”. El matemático McDiarmid (vea [33]) contempló estos resultados como casos particulares de un teorema mayor que actualmente lleva su nombre. Brevemente, usando la teoría de martingalas él fue capaz de estimar la probabilidad de desvío de \mathbf{X} en términos de las tasas de cambio que produce cada coordenada x_i en $f(x_1, \dots, x_m)$. Cabe indicar que esto fue mostrado en el caso ‘independiente’. La pregunta que nos hacemos aquí es, ¿qué sucede en el caso ‘casi’ independiente? (esto es, cuando seleccionamos sin reposición). El siguiente resultado de nuestra autoría responde a esta

¹²El lector interesado puede dirigirse a la referencia [1].

¹³Esto se debe a que la ‘masa’ de la distribución binomial se concentra alrededor de su esperanza pN , por lo tanto es plausible escoger los parámetros que maximizan el denominador.

pregunta, y será utilizado en nuestro estudio del modelo uniforme $\mathbf{G}(n, m)$. Adicionalmente, remarcamos al lector que si bien este resultado no aparece en la literatura previa, si existen variantes que podemos clasificar como “Desigualdades Tipo McDiarmid”.

Proposición 2.2 (Desigualdad Tipo McDiarmid - versión uniforme). *Sea $f : \mathcal{S}_{N,m} \rightarrow \mathbb{R}$ y sea $\psi : \mathcal{S}_N \rightarrow \mathbb{R}^+$. Si f es una función ψ -Lipschitz, entonces para todo $a \geq 0$ se cumple que*

$$\mathbb{P}_{\mathcal{S}_{N,m}} (f(S_m) - \mathbb{E}_{\mathcal{S}_{N,m}}[f(S_m)] \geq a) \leq \exp\left(\frac{-a^2}{24t\|\psi\|_2^2 + 6a\|\psi\|_\infty}\right)$$

Más aún, el mismo limitante aplica para $\mathbb{P}_{\mathcal{S}_{N,m}} (f(S_m) - \mathbb{E}_{\mathcal{S}_{N,m}}[f(S_m)] < -a)$.

Este resultado es válido para cualquier modelo uniforme. Cuando especificamos muestra elección a aristas de grafo completo, obtenemos el siguiente resultado sobre $\mathbf{G}(n, m)$ que es lo que realmente nos interesa.

Corolario 2.1 (Desigualdad Tipo McDiarmid para $\mathbf{G}(n, m)$). *Sea $f : \mathcal{G}_{n,m} \rightarrow \mathbb{R}$ y sea $\psi : E(K_N) \rightarrow \mathbb{R}^+$. Si f es una función ψ -Lipschitz, entonces para todo $a \geq 0$ se cumple que*

$$\mathbb{P}_m (f(G_m) - \mathbb{E}_m (f(G_m)) \geq a) \leq \exp\left(\frac{-a^2}{24t\|\psi\|_2^2 + 6a\|\psi\|_\infty}\right)$$

Más aún, el mismo limitante aplica para $\mathbb{P}_m (f(G_m) - \mathbb{E}_m (f(G_m)) < -a)$.

Para finalizar esta subsección, nos gustaría indicar que la prueba de este corolario puede ser encontrada en [1]. De hecho, este corolario fue el primer resultado que obtuvimos del ‘tipo McDiarmid’ y sólo posteriormente nos reparamos de su extensión a cualquier modelo uniforme. Esta observación fue adherida a la versión final de nuestro trabajo pues la prueba sigue las mismas líneas.

2.4.3. Análisis

En este trabajo requeriremos ciertas técnicas y habilidades analíticas. Mucho de este conocimiento es estándar y es desarrollado en un curso usual de Análisis Matemático. Sin embargo existen principalmente dos puntos que no suelen ser tópicos regulares en primeros cursos de análisis. Uno de ellos es el llamado “argumento diádico” que veremos posteriormente. Lo que si es abordado en ciertos cursos de análisis son las estimaciones de series vía integrales. Un ejemplo bien conocido de esto es la aproximación de la serie

$\sum_{n=1}^{\infty} \frac{1}{n}$ mediante la integral $\int \frac{1}{t} dt$. Motivados por estas ideas, y debido a cierta necesidad de estimar series de exponenciales con muchas variables, desarrollamos el siguiente resultado que nos será de gran ayuda a la hora de estimar nuestras probabilidades.

Estimando Sumas con Integrales.

A seguir enunciamos el resultado analítico de este apartado.

Proposición 2.3 (Estimando Sumas con Rápido Decaimiento Exponencial). *Sea $d \in \mathbb{N}$, r un número real por lo menos \sqrt{d} y $\beta \geq 1$ un parámetro. Entonces se cumple la siguiente desigualdad,*

$$\sum_{\substack{x \in \mathbb{Z}^d: \\ \|x\|_2 \geq r}} \exp(-\beta \|x\|_2^2) \leq C^d \exp\left(\frac{-\beta(r - \sqrt{d})^2}{2}\right),$$

(donde C puede ser tomando como $\sqrt{8\pi}$, por ejemplo).

Observación 2.1. *En la práctica, la Proposición 2.3 es aplicada con r significativamente mayor que d (esto es, $r \geq K\sqrt{d}$ con $K > 0$ constante grande) de modo que, la estimativa superior tenga la forma $\exp(-\Omega(r^2))$.*

Este resultado es probado de forma general, para ciertas funciones monótonas con propiedades radiales. Los detalles pueden ser encontrados en [1].

El lector familiarizado con técnicas analíticas puede encontrar natural este resultado, pero a pesar de nuestras búsquedas, no encontramos una referencia directa para este resultado (ni su versión generalizada). Por esto decidimos colocar la prueba completa de esta desigualdad en uno de los apéndices de [1].

El Argumento Diádico.

En análisis, principalmente en análisis armónico, la técnica del argumento diádico es bien conocida. En este apartado explicaremos a grosso modo de que se trata, en lugar de colocar algún tipo de fórmula que la explique meramente.

Imagine que tiene una serie como $\sum_{x \in X} f(x)$, donde X es un conjunto finito. Suponga

que usted tiene la tarea de estimar esta serie, salvo constantes.

¿Cuál sería un método sistemático en que usted haría esto? El argumento diádico nos sugiere la siguiente estrategia. En lugar de considerar individualmente cada contribución, vamos a agrupar términos que tienen un cierto valor salvo un factor. Así, nuestra tarea se reduce a un problema de conteo, esto es, contar cuantos elementos tiene cada grupo. En lenguaje matemático, tenemos lo siguiente.

Sea

$$m := \text{mín } f \text{ y } M := \text{máx } f.$$

Por simplicidad consideraremos un factor de 2, así partimos nuestro intervalo $[m, M]$ en subintervalos de la forma

$$[2^i m, 2^{i+1} m)$$

donde i varía de 0 a $L := \log_2(M/m)$. Denotando por

$$X_i := f^{-1}[2^i m, 2^{i+1} m),$$

tenemos que

$$m \sum_{i=0}^L 2^i |X_i| \leq \sum_{x \in X} f(x) \leq 2m \sum_{i=0}^L 2^i |X_i|.$$

Así, hemos estimado nuestra serie $\sum_{x \in X} f(x)$ por la serie

$$\left(\sum_{i=0}^L 2^i |X_i| \right) \times m,$$

salvo un factor de 2, con la 'ventaja' que se ha reducido nuestro problema a un problema de conteo, esto es, entender los tamaños de los conjuntos X_i .

En muchos casos, esto suele ser suficiente para atingir nuestro objetivo, como veremos a lo largo de este trabajo.

Mas, como alerta, y ya que sucede en casi todas las aplicaciones de técnicas matemáticas, no siempre es posible aplicar directamente una estrategia para esta funcionar, mas sí, al aplicarla de una forma inteligente.

Capítulo 3

Sobre los resultados de Goldschmidt, Griffiths y Scott

En este capítulo “desmembraremos” el trabajo reciente de Goldschmidt, Griffiths y Scott [21] donde los autores estudian desvíos moderados del conteo de subgrafos en los modelos de grafos aleatorios uniforme y binomial. Es justo decir que en [21] también se da a conocer una conexión (no-determinista) estrecha entre el desvío del conteo de subgrafos y el desvío de dos grafos particulares, a saber, P_2 (cereza) y K_3 (triángulo). Sin embargo, en el presente trabajo “proyectaremos” el estudio realizado en [21] hacia los resultados que nos interesan, esto es, el estudio del desvío del conteo de triángulos (y cerezas) en los modelos aleatorios densos.¹

3.1. Panorama General de Goldschmidt, Griffiths y Scott

En Goldschmidt, Griffiths y Scott [21] los autores tienen como objetivo inicial estudiar los desvíos moderados del conteo de triángulos para el modelo aleatorio binomial $\mathbf{G}(n, p)$. Así, el lector curioso podría preguntarse porqué nos “desviamos” al modelo aleatorio uniforme $\mathbf{G}(n, m)$. La razón es que en el modelo binomial $\mathbf{G}(n, p)$ una de las causas naturales (quizá la más natural) de desvío en el conteo de triángulos (y otras estructuras) es un exceso en el número de aristas en el grafo aleatorio.

¹En [21] también se obtienen resultados para modelos “relativamente densos”, pero en el caso de los no-densos estos resultados no son óptimos.

Informalmente, cada arista forma parte (en promedio) de “ p^2n ” triángulos aproximadamente, luego si tenemos s aristas nuevas, esto nos aumenta (en promedio) sp^2n triángulos, por tanto debido a que la variable aleatoria que cuenta el número de aristas en $\mathbf{G}(n, p)$ es asintóticamente normal, si estamos dispuestos a pagar una probabilidad de $\exp(-b)$, conseguimos un exceso de aristas de orden $s = b^{1/2}p^{1/2}n$ y un desvío (positivo) de triángulos de orden $sp^2n = b^{1/2}p^{5/2}n^2$.

La observación crucial de los autores en [21] es que existen otras causas más sutiles que también causan desvíos en el conteo de triángulos. Luego, una forma fina de centrarnos en estas posibles causas, sin “interferencia” de la causa anterior (por exceso de aristas), es condicionar nuestro grafo aleatorio $\mathbf{G}(n, p)$ a tener un número fijo de aristas m . Así el grafo aleatorio resultante es el uniforme $\mathbf{G}(n, m)$, como esperado. Una vez entendido el caso uniforme, resta ponderar estos modelos para recuperar $\mathbf{G}(n, p)$. Esta tarea no es simple de ninguna forma, pero es más tecnicismo y la analizaremos al final del capítulo. A lo largo de este capítulo también veremos que, dentro del modelo $\mathbf{G}(n, m)$, otra posible causa de desvío en el conteo de triángulos es “el azar” (o “puro azar”). Pero aquí no acaba la historia, eventualmente será mostrado que existe otro grupo de causas que son conocidos como “fenómeno de localización”, mas esto sólo será visto en el Capítulo 4 donde explicaremos nuestras contribuciones.

Antes de pasar a los argumentos, listaremos las contribuciones más importantes en [21]:

- (1) Se proporciona una representación martingala del desvío de conteo de cualquier subgrafo fijo (en particular triángulos y cerezas) en el modelo uniforme $\mathbf{G}(n, m)$ (para cualquier $m \in \{0, 1, \dots, N\}$).
- (2) Se determina asintóticamente la función tasa asociada a desvíos moderados de conteo de subgrafos fijos en $\mathbf{G}(n, m)$, cuando m es (relativamente) denso y para ciertos desvíos moderados (todo la región en el caso denso).
- (3) Se deducen resultados sobre desvíos moderados de subgrafos (fijos) en el modelo $\mathbf{G}(n, p)$ que son más fuertes que enunciados anteriores a ese trabajo.

3.2. Desvíos Moderados en $\mathbf{G}(n, m)$

El objetivo de esta sección es mostrar los resultados centrales de desvíos moderados (de cerezas y triángulos) en el modelo (denso) uniforme $\mathbf{G}(n, m)$. Con este fin, dividiremos la sección en tres subsecciones. Las dos primeras son preparatorias y contienen los ingredientes principales sobre la representación Martingala y sobre los grados y cogrados. Para finalizar, en la última subsección encontraremos la prueba de los resultados deseados, donde se hará uso de los ingredientes obtenidos en las dos subsecciones anteriores.

3.2.1. Representación Martingala de $D_{\wedge}(\mathbf{G}(n, m))$ y $D_{\Delta}(\mathbf{G}(n, m))$

En esta subsección exhibiremos la representación Martingala de

$$D_{\wedge}(\mathbf{G}(n, m)) \text{ y } D_{\Delta}(\mathbf{G}(n, m))$$

respecto al Proceso Erdős-Rényi. El punto crucial es que esta representación será descrita de una forma natural y (relativamente) simple de modo que podamos manipularla usando herramientas de Combinatoria Probabilística.

Un ingrediente importante que usaremos en la deducción de nuestra representación martingala será computar el valor esperado del número de copias creadas de $H \in \{\Delta, \wedge\}$ al adicionar la arista m -ésima en el proceso aleatorio Erdős-Rényi.

Con esto en mente, sea

$$(\mathbf{G}_m : m = 0, \dots, N)$$

el proceso Erdős-Rényi (generado a partir de la permutación aleatoria uniforme σ). A partir de este proceso definimos las variables aleatorias²

$$A_H(\sigma_m) := N_H(\mathbf{G}_m) - N_H(\mathbf{G}_{m-1}) \quad (3.1)$$

$$X_H(\sigma_m) := A_H(\sigma_m) - \mathbb{E}[A_H(\sigma_m) | \sigma_{m-1}] \quad (3.2)$$

donde la variable aleatoria $A_H(\sigma_m)$ es precisamente la variable que cuenta el número de copias (inmersiones) de H en \mathbf{G}_m creadas al adicionar la arista e_m al grafo \mathbf{G}_{m-1} . Por otra parte, la variable aleatoria $X_H(\sigma_m)$ es la variable centralizada de $A_H(\sigma_m)$ una vez expuesta la secuencia σ_{m-1} . En particular,

$$\mathbb{E}[X_H(\sigma_m) | \sigma_{m-1}] = 0.$$

Ya estamos en condiciones de enunciar el teorema central de esta subsección. Como el enunciado es determinístico, será formulado para cada realización posible del proceso aleatorio.³

²Note que en la ecuación (3.2) la esperanza condicionada $\mathbb{E}[A_H(\sigma_m) | \sigma_{m-1}]$ es dada al revelar la información de σ_{m-1} , así dicha 'esperanza' es tomada respecto a la variable aleatoria e_m y prevenimos al lector de no 'evaluar' σ_m dentro de esta esperanza.

³En [21] los autores muestran una versión general de este teorema que aplica a cualquier subgrafo fijo H en lugar de sólo \wedge y Δ . Sin embargo el argumento es el mismo, salvo tecnicismos incorporados al trabajar con un grafo cualquiera.

Teorema 3.1 (Representación Martingala de los Desvíos del Conteo de Subgrafos). *Sea* $(G_m : m = 0, \dots, N)$ *una realización del proceso aleatorio Erdős-Rényi. Para cada* $m \in \{0, \dots, N\}$ *se cumple la siguiente igualdad,*

$$D_{\wedge}(G_m) = \sum_{i=1}^m \left\{ \frac{(N-m)_2}{(N-i)_2} X_{\wedge}(\sigma_i) \right\}, y \quad (3.3)$$

$$D_{\Delta}(G_m) = \sum_{i=1}^m \left\{ 3 \frac{(N-m)_2(m-i)}{(N-i)_3} X_{\wedge}(\sigma_i) + \frac{(N-m)_3}{(N-i)_3} X_{\Delta}(\sigma_i) \right\} \quad (3.4)$$

Como el lector podrá observar, dada la filtración (natural) G_0, \dots, G_m (o, equivalentemente, e_1, \dots, e_m) deducimos que la parte derecha de (3.3) y (3.4) son martingalas puesto que cada X 's es un incremento martingala (resp. de la filtración) y combinación lineal de incrementos martingala es un incremento martingala.

Antes de pasar a la prueba del Teorema 3.1 y para simplicidad del lector, vamos a enunciar un lema sobre la esperanza condicionada de $A_H(\sigma_m)$ (dado σ_{m-1}) como ya anticipamos al inicio de esta subsección, esto dejará la prueba del Teorema 3.1 más limpia.

Lema 3.1 (Esperanza del incremento de copias en una etapa). *Sea* $(G_m : m = 1, \dots, N)$ *una realización del proceso Erdős-Rényi. Tenemos que*

$$\begin{aligned} & L_H(m) - L_H(m-1) \\ &= \frac{1}{N-m+1} \sum_{f \in E(H)} (L_{H \setminus f}(m-1) - L_H(m-1)), \end{aligned} \quad (3.5)$$

y

$$\begin{aligned} & \mathbb{E}[A_H(\sigma_m) | \sigma_{m-1}] \\ &= \frac{1}{N-m+1} \sum_{f \in E(H)} (N_{H \setminus f}(G_{m-1}) - N_H(G_{m-1})) \end{aligned} \quad (3.6)$$

$$\begin{aligned} &= L_H(m) - L_H(m-1) \\ &+ \frac{1}{N-m+1} \sum_{f \in E(H)} (D_{H \setminus f}(m-1) - D_H(m-1)). \end{aligned} \quad (3.7)$$

Demostración del Lema 3.1:

Primeramente queremos hacer notar al lector que basta mostrar la ecuación (3.6) debido a la Ley de la esperanza condicionada. En efecto, para mostrar (3.5) observe que

$$\begin{aligned}
 L_H(m) - L_H(m-1) &= \mathbb{E}[A_H(\sigma_m)] \\
 &= \mathbb{E}[\mathbb{E}[A_H(\sigma_m) \mid \sigma_{m-1}]] \\
 &= \frac{1}{N-m+1} \sum_{f \in E(H)} (\mathbb{E}[N_{H \setminus f}(G_{m-1})] - \mathbb{E}[N_H(G_{m-1})]) \\
 &= \frac{1}{N-m+1} \sum_{f \in E(H)} (L_{H \setminus f}(m-1) - L_H(m-1)).
 \end{aligned}$$

Adicionalmente, de (3.5) y (3.6) podemos fácilmente deducir (3.7) por sustitución como sigue,

$$\begin{aligned}
 &\mathbb{E}[A_H(\sigma_m) \mid \sigma_{m-1}] \\
 &= \frac{1}{N-m+1} \sum_{f \in E(H)} (N_{H \setminus f}(G_{m-1}) - N_H(G_{m-1})) \\
 &= \frac{1}{N-m+1} \sum_{f \in E(H)} (L_{H \setminus f}(m-1) + D_{H \setminus f}(G_{m-1}) \\
 &\quad - (L_H(m-1) + D_H(G_{m-1}))) \\
 &= \frac{1}{N-m+1} \sum_{f \in E(H)} (L_{H \setminus f}(m-1) - L_H(m-1)) \\
 &\quad + \frac{1}{N-m+1} \sum_{f \in E(H)} (D_{H \setminus f}(G_{m-1}) - D_H(G_{m-1})) \\
 &= L_H(m) - L_H(m-1) + \frac{1}{N-m+1} \sum_{f \in E(H)} (D_{H \setminus f}(G_{m-1}) - D_H(G_{m-1})).
 \end{aligned}$$

Así, resta mostrar (3.6). Para esto introducimos la siguiente notación auxiliar. Sea $f \in E(H)$, denotemos por

$$A_{H,f}(\sigma_m) := |\{\phi \in \text{End}(H, G_m) : \phi[f] = e_m\}|.$$

Es inmediato notar que $A_{H,f}(\sigma_m) = |\{\psi \in \text{End}(H \setminus f, G_{m-1}) : \psi[f] = e_m\}|$, pues existe una correspondencia natural entre estos dos conjuntos. Por definición de $A_{H,f}(\sigma_m)$ y

$A_H(\sigma_m)$ tenemos además que

$$A_H(\sigma_m) = \sum_{f \in E(H)} A_{H,f}(\sigma_m)$$

Luego, habremos demostrado (3.6) se mostramos que, para cada $f \in E(H)$ se satisface la siguiente ecuación,

$$\mathbb{E}[A_{H,f}(\sigma_m) | \sigma_{m-1}] = \frac{1}{N - m + 1} (N_{H \setminus f}(G_{m-1}) - N_H(G_{m-1})) \quad (3.8)$$

Mas, tomando esperanza condicionada en $A_{H,f}(\sigma_m)$ tenemos que

$$\begin{aligned} & \mathbb{E}[A_{H,f}(\sigma_m) | \sigma_{m-1}] \\ &= \frac{1}{N - m + 1} \sum_{e \notin E(G_{m-1})} |\{\psi \in \text{End}(H \setminus f, G_{m-1}) : \psi[f] = e\}| \\ &= \frac{1}{N - m + 1} \left| \bigcup_{e \notin E(G_{m-1})} \{\psi \in \text{End}(H \setminus f, G_{m-1}) : \psi[f] = e\} \right| \\ &= \frac{1}{N - m + 1} |\{\psi \in \text{End}(H \setminus f, G_{m-1}) : \psi[f] \notin E(G_{m-1})\}| \\ &= \frac{1}{N - m + 1} (|\text{End}(H \setminus f, G_{m-1})| \\ &\quad - |\{\psi \in \text{End}(H \setminus f, G_{m-1}) : \psi[f] \in E(G_{m-1})\}|) \\ &= \frac{1}{N - m + 1} (|\text{End}(H \setminus f, G_{m-1})| - |\text{End}(H, G_{m-1})|) \\ &= \frac{1}{N - m + 1} (N_{H \setminus f}(G_{m-1}) - N_H(G_{m-1})), \end{aligned}$$

como deseábamos. ■

A continuación pasaremos a la prueba del Teorema 3.1. El argumento que usaremos está basado en una doble inducción. Así, primero será mostrado la ecuación (3.3) y luego usaremos este resultado para deducir (3.4).

Demostración del Teorema 3.1:

La prueba (en su versión general) pasa por una doble inducción respecto del número

de aristas e_H del subgrafo H a contar y de m . Así es conveniente hacer las siguientes aclaraciones:

(i) Si $e_H \in \{0, 1\}$ entonces

$$D_H(G_m) = 0 \text{ y } X_H(\sigma_m) = 0.$$

En efecto, esto se debe a que en el proceso estocástico considerado, el número de aristas es determinístico. En particular, también es determinístico el número de copias de H (cuando $e_H \in \{0, 1\}$), lo que resulta en fluctuación 0.

(ii) $D_H(G_0) = 0$.

En efecto, en nuestro proceso estocástico el grafo G_0 no presenta aristas y por tanto la variable $N_H(G_0)$ es determinístico, y como en el caso anterior no presenta fluctuaciones.

El ítem (i) y (ii) nos permite iniciar la inducción.

Por otra parte, el paso inductivo se soporta en la siguiente fórmula recursiva.

$$D_H(G_m) = D_H(G_{m-1}) + X_H(\sigma_m) + \frac{1}{N - m + 1} \sum_{f \in E(H)} (D_{H \setminus f}(G_{m-1}) - D_H(G_{m-1})) \quad (3.9)$$

Demostración de (3.9):

La prueba se sigue directamente de las definiciones de las variables aleatorias junto con el Lema 3.1, como se muestra a continuación.

$$\begin{aligned} D_H(G_m) &= N_H(G_m) - L_H(m) \\ &= N_H(G_{m-1}) + A_H(\sigma_m) - L_H(m) \\ &= D_H(G_{m-1}) + X_H(\sigma_m) + \mathbb{E}[A_H(\sigma_m) | \sigma_{m-1}] - (L_H(m) - L_H(m-1)) \\ &= D_H(G_{m-1}) + X_H(\sigma_m) + \frac{1}{N - m + 1} \sum_{f \in E(H)} (D_{H \setminus f}(G_{m-1}) - D_H(G_{m-1})). \end{aligned}$$

Ahora como ilustración (y por estar interesados sólo en estos dos casos), mostraremos como deducir la representación martingala para

$$H = \wedge \text{ y } H = \Delta.$$

Caso $H = \wedge$.

Procedemos por inducción en m . Por el item (ii) sabemos que para $m = 0$ el enunciado es trivial. Ahora mostraremos el paso inductivo.

Supongamos que nuestra representación se cumple para $m - 1$ (con $m \geq 1$).
 Substituyendo en la fórmula (3.9) y usando el item (i), tenemos que

$$\begin{aligned} D_{\wedge}(G_m) &= D_{\wedge}(G_{m-1}) + X_{\wedge}(\sigma_m) - \frac{2}{N - m + 1} D_{\wedge}(G_{m-1}) \\ &= \left(1 - \frac{2}{N - m + 1}\right) D_{\wedge}(G_{m-1}) + X_{\wedge}(\sigma_m) \\ &= \sum_{i=1}^{m-1} \left\{ \left(\frac{N - m - 1}{N - m + 1}\right) \times \frac{(N - m + 1)_2}{(N - i)_2} X_{\wedge}(\sigma_i) \right\} + X_{\wedge}(\sigma_m) \\ &= \sum_{i=1}^m \frac{(N - m)_2}{(N - i)_2} X_{\wedge}(\sigma_i) \end{aligned}$$

lo que demuestra la veracidad de nuestra representación para $H = \wedge$.

Caso $H = \Delta$.

Nuevamente procedemos por inducción, mas teniendo en cuenta el resultado para ⁴

$$H = \wedge.$$

Como antes, el caso $m = 0$ es trivial.

Supongamos válida la representación para $m - 1$. Así, juntando (3.9) y el caso anterior,

⁴En esta observación se encuentra escondida la doble inducción.

tenemos que $D_\Delta(G_m)$ puede ser escrito como

$$\begin{aligned}
& D_\Delta(G_{m-1}) + X_\Delta(\sigma_m) + \frac{1}{N-m+1} \sum_{f \in E(\Delta)} (D_{\Delta \setminus f}(G_{m-1}) - D_\Delta(G_{m-1})) \\
&= D_\Delta(G_{m-1}) + X_\Delta(\sigma_m) + \frac{3}{N-m+1} (D_\wedge(G_{m-1}) - D_\Delta(G_{m-1})) \\
&= \left(1 - \frac{3}{N-m+1}\right) D_\Delta(G_{m-1}) + \frac{3}{N-m+1} D_\wedge(G_{m-1}) + X_\Delta(\sigma_m) \\
&= \sum_{i=1}^{m-1} 3 \left(\frac{N-m-2}{N-m+1} \times \frac{(N-m+1)_2(m-1-i)}{(N-i)_3} \right. \\
&\quad \left. + \frac{1}{N-m+1} \times \frac{(N-m+1)_2}{(N-i)_2} \right) X_\wedge(\sigma_i) \\
&\quad + \sum_{i=1}^{m-1} \left(\frac{N-m-2}{N-m+1} \times \frac{(N-m+1)_3}{(N-i)_3} \right) X_\Delta(\sigma_i) + X_\Delta(\sigma_m) \\
&= \sum_{i=1}^{m-1} 3 \left(\frac{(N-m)_2(m-i)}{(N-i)_3} \right) X_\wedge(\sigma_i) + \sum_{i=1}^{m-1} \frac{(N-m)_3}{(N-i)_3} X_\Delta(\sigma_i) + X_\Delta(\sigma_m) \\
&= \sum_{i=1}^m \left\{ 3 \frac{(N-m)_2(m-i)}{(N-i)_3} X_\wedge(\sigma_i) + \frac{(N-m)_3}{(N-i)_3} X_\Delta(\sigma_i) \right\},
\end{aligned}$$

con lo que queda demostrado. ■

3.2.2. Sobre Grados y Cogrados en $G(n, m)$

El t3pico de grados y cogrados en grafos aleatorios ha sido ampliamente estudiado en la literatura, vea por ejemplo [5, 31, 34].

Sin embargo, tanto en este trabajo como en [21] estamos interesados en ciertos eventos particulares (relativos a grados/cogrados) pensados para ser aplicados en nuestro estudio de conteo de tri3ngulos (e cerezas). Por esta raz3n no llamaremos a las referencias anteriores, sino que nos abocaremos a la tarea de probar nuestros propios resultados, los cuales ser3n deducidos usando la versi3n de la Desigualdad de Concentraci3n de McDiarmid, enunciada en el Corolario 2.1.

Esta subsecci3n ser3 organizada en dos partes. En la primera parte estudiaremos los grados donde mostraremos dos resultados, uno referente al m3ximo desv3o de grados y la otra referente a la suma de cuadrados de los desv3os de grados. Posteriormente, en la

segunda parte, mostraremos los resultados análogos para cogrados. Tanto la subsección de grados como de cogrados serán revisadas en el Capítulo 4 pues ciertas mejoras en las estimativas serán requeridas, así los resultados mostrados en este capítulo sólo son óptimos para los modelos densos.

Recordamos al lector que en este trabajo, sólo estamos interesados en el orden de magnitud de las funciones involucradas, así no nos concentraremos en optimizar las constantes que irán apareciendo.

Sobre Grados en $\mathbf{G}(n, m)$.

Sea $\mathcal{G}_{n,m}$ el espacio de probabilidades de los grafos con n vértices y m aristas dotados de la probabilidad uniforme $\mathbb{P}_{n,m}$. En lo que se sigue sólo escribiremos \mathbb{P} en lugar de $\mathbb{P}_{n,m}$ para economizar notación, y en su lugar a los elementos de $\mathcal{G}_{n,m}$ los denotaremos por G_m . Para cada vértice $u \in [n]$, definimos la variable aleatoria *desvío de grado de u* (respecto a G_m) como siendo

$$\begin{aligned} D_u(G_m) &:= d_u(G_m) - \mathbb{E}[d_u(G_m)] \\ &= d_u(G_m) - 2 \frac{m}{n}. \end{aligned}$$

También definimos la variable aleatoria *máximo desvío de grado* (respecto a G_m) como siendo

$$D_{\max}(G_m) := \max_{u \in [n]} D_u(G_m).$$

El siguiente resultado, que es mostrado en [21], es una consecuencia inmediata del hecho que $d_u(G_m)$ tiene distribución Hypergeometric($N, n - 1, m$) junto con el resultado de concentración asociado a este tipo de distribución (Teorema 2.1).

Lema 3.2 (Máximo Desvío de Grado en $\mathbf{G}(n, m)$). *Sea $G_m \in \mathcal{G}_{n,m}$ con $t = m/N$ y $b = b_n \geq \log n$ una secuencia real. Entonces, se cumple la siguiente desigualdad no-determinista*

$$|D_{\max}(G_m)| \lesssim^{*b} b^{1/2} \sqrt{tn} + b$$

El siguiente lema nos permitirá controlar (superiormente) la varianza del proceso aleatorio asociado a los incrementos $X_\wedge(\sigma_i)$ en nuestras representaciones martingalas.

Lema 3.3 (Suma de Cuadrados de los Desvíos de Grados en $\mathbf{G}(n, m)$). Sea $G_m \in \mathcal{G}_{n,m}$ y $b = b_n \geq n$ una secuencia real. Entonces, se cumple que

$$\sum_{u \in [n]} D_u(G_m)^2 \lesssim^{*b} bn$$

Observación 3.1. En la demostración de este lema veremos como aplicar el argumento diádico (ver 2.4.3) para estimar ciertas variables aleatorias. Este tipo de argumentos son usuales en la teoría de grandes desvíos, sin embargo aquí veremos una cara distinta de esta técnica, pues los autores ([21]) presentan una variante “global” de este argumento, al considerar todas las contribuciones al mismo tiempo, en lugar de (como es usual), separar las contribuciones respecto a su orden y estimarlas independientemente.

Demostración del Lema 3.3:

En lugar de trabajar con la variable aleatoria original

$$\sum_{u \in [n]} D_u(G_m)^2$$

vamos a “linealizar” el problema substituyendo la variable cuadrática anterior por variables lineales de la forma

$$f_x(G_m) := \sum_{u \in [n]} x_u D_u(G_m)$$

donde $x = (x_u : u \in [n])$ y cada coordenada x_u toma valores de la forma $\pm 2^l$ con $l \in [\log n]$.⁵

Para relacionar estas variables aleatorias, nos apoyaremos en el argumento diádico. Sea G_m un grafo (realización) en $\mathcal{G}_{n,m}$. Definamos el siguiente vector “diádico” $x^* = (x_u^* : u \in [n])$ asociado a G_m ,

$$x_u^*(G_m) = \begin{cases} 0, & \text{si } |D_u(G_m)| \leq \sqrt{n}, \text{ y} \\ \text{sign}(D_u(G_m)) 2^j, & \text{si } 2^j \sqrt{n} \leq |D_u(G_m)| < 2^{j+1} \sqrt{n}. \end{cases} \quad (3.10)$$

⁵Note que esta sustitución particular está motivada por el argumento diádico (ver 2.4.3).

Esto es, $x_u^* = x_u^*(G_m)$ codifica el orden diádico (considerando signos) de $D_u(G_m)$ escalado por la varianza de grado (orden \sqrt{n}).⁶ Note que, de la definición dada, tenemos las siguientes desigualdades:

$$f_{x^*}(G_m) \geq \sum_{u \in [n]} (x_u^*(G_m))^2 \sqrt{n}, \text{ y} \quad (3.11)$$

$$\sum_{u \in [n]} D_u(G_m)^2 \leq 4 \sum_{u \in [n]} x_u^*(G_m)^2 n. \quad (3.12)$$

Ya estamos preparados para “linealizar” nuestra variable aleatoria original. Supongamos que se cumple el evento

$$\left\| \sum_{u \in [n]} D_u(G_m)^2 > Cbn \right\|$$

para alguna constante C .

Entonces, de (3.12) deducimos que (junto con (3.11))

$$\begin{aligned} f_x(G_m) &\geq \|x\|_2^2 \sqrt{n}, \\ \|x\|_2^2 &\geq C_1 b, \text{ y} \\ \|x\|_\infty &\leq \sqrt{n}, \end{aligned}$$

para algún vector diádico x .

Denotando por $F(x)$ el evento “ $f_x(G_m) \geq \|x\|_2^2 \sqrt{n}$ ”, hemos mostrado que

$$\begin{aligned} \mathbb{P} \left(\sum_{u \in [n]} D_u(G_m)^2 > Cbn \right) &\leq \mathbb{P} \left(\bigvee_{x: \|x\|_2 \geq C_2 \sqrt{b}} F(x) \right), \\ &\leq \sum_{x: \|x\|_2 \geq C_2 \sqrt{b}} \mathbb{P}(F(x)). \end{aligned} \quad (3.13)$$

La siguiente etapa de la prueba será estimar la probabilidad del evento $F(x)$, enunciada en la siguiente afirmación.

Afirmación. $\mathbb{P}(F(x)) \leq \exp(-K' \|x\|_2^2)$, para alguna constante $K' > 0$.

⁶Note que el vector x^* es un vector aleatorio, pues está definido en un espacio probabilístico, esto es $\mathcal{G}_{n,m}$.

Debido a la linealidad de la variable aleatoria f_x , disponemos de las técnicas apropiadas de concentración para estimar estos eventos.

Demostración de la Afirmación:

Deseamos demostrar que la variable aleatoria f_x está altamente concentrada en torno de 0 (esto es, su esperanza). Para ello usaremos el Corolario 2.1.⁷

Para esto, afirmamos que f_x es ψ -Lipschitz de la función

$$\psi(e) := \sum_{u \in e} x_u.$$

En efecto, sean $G, G' \in \mathcal{G}_{n,m}$ con aristas distinguidas e y e' respectivamente (esto es $e \in E(G) \setminus E(G')$ y $e' \in E(G') \setminus E(G)$).

Así, podemos escribir $f_x(G)$ y $f_x(G')$ de la siguiente manera:

$$\begin{aligned} f_x(G) &= \sum_{u \notin e \cup e'} x_u D_u(G) + \sum_{u \in e \cup e'} x_u D_u(G) \\ f_x(G') &= \sum_{u \notin e \cup e'} x_u D_u(G') + \sum_{u \in e \cup e'} x_u D_u(G') \end{aligned}$$

Restando las dos expresiones anteriores deducimos que

$$\begin{aligned} |f_x(G) - f_x(G')| &\leq \sum_{u \in e \cup e'} |x_u| |D_u(G) - D_u(G')|, \\ &= \sum_{u \in e \cup e'} |x_u| |d_u(G) - d_u(G')|, \\ &\leq \sum_{u \in e \cup e'} |x_u|, \\ &\leq \sum_{u \in e} |x_u| + \sum_{u \in e'} |x_u|. \end{aligned}$$

Así, acabamos de mostrar nuestra afirmación, es decir que f es ψ -Lipschitz.

⁷En [21] los autores usan una versión más débil de este Corolario, pero efectiva para el caso denso que ellos estudian.

Las siguientes estimativas también serán útiles al momento de aplicar el Corolario 2.1 para f_x :

$$\begin{aligned}
\|\psi\|_2^2 &= \sum_{e \in E(K_n)} \psi(e)^2 \\
&\leq 2 \sum_{e \in E(K_n)} \sum_{u \in e} x_u^2 \\
&= 2 \sum_{u \in [n]} \sum_{e \ni u} x_u^2 \\
&\leq 2n \sum_{u \in [n]} x_u^2 = 2n \|x\|_2^2, \text{ y} \\
\|\psi\|_\infty &= \max_{e \in E(K_n)} \sum_{u \in e} |x_u| \\
&\leq 2 \|x\|_\infty \\
&\leq 2\sqrt{n}.
\end{aligned}$$

Por tanto, del Corolario 2.1 y aplicando las estimativas obtenidas, concluimos que

$$\begin{aligned}
\mathbb{P}(F(x)) &= \mathbb{P}(f_x(G_m) > C \|x\|_2^2 \sqrt{n}) \\
&\leq \exp\left(\frac{-C' \|x\|_2^4 n}{\|\psi\|_2^2 + \|x\|_2^2 \sqrt{n} \|\psi\|_\infty}\right) \\
&\leq \exp\left(\frac{-C' \|x\|_2^4 n}{2n \|x\|_2^2 + 2 \|x\|_2^2 n}\right) \\
&\leq \exp(-K' \|x\|_2^2),
\end{aligned}$$

lo que muestra nuestra Afirmación inicial.

Para finalizar la prueba del Lema 3.3 basta mostrar que

$$\sum_{x: \|x\|_2 \geq C_2 \sqrt{b}} \exp(-K' \|x\|_2^2) \leq \exp(-K b) \tag{3.14}$$

para algún $K > 0$, pues combinando (3.13) con la Afirmación, obtendríamos la estimativa superior deseada.

La heurística detrás de esta desigualdad es que la suma de términos con decaimiento ex-

ponencial es “dominada” por el mayor de los términos. En [21] los autores exhiben un argumento combinatorio que justifica (3.14). En este trabajo la estimativa resulta inmediata en virtud de nuestro resultado analítico 2.3 que nos proporciona este tipo de estimativas de forma automática. Para ello, basta aplicar nuestra proposición 2.3 para los valores $d = n$, $r = \Omega(\sqrt{b})$, $\beta = \Omega(1)$ y verificar que r es más grande que \sqrt{d} . ■

Este argumento es elegante y automático. A partir de aquí, cuando necesitemos esta proposición sólo la referiremos, sin mostrar los detalles de la prueba, por ser todas parecidas con esta aplicación.

Sobre Cogrados en $\mathbf{G}(n, m)$.

Esta sección es parecida a la anterior, tanto en su estructura como en los enunciados (y pruebas). Así, nuestra revisión será un poco más sucinta.

Para cada par de vértices (diferentes) $u, w \in [n]$, definimos la variable aleatoria *desvío de cogrado de u* (respecto a G_m) como siendo

$$\begin{aligned} D_{uw}(G_m) &:= d_{uw}(G_m) - \mathbb{E}[d_{uw}(G_m)] \\ &= d_{uw}(G_m) - (n-2) \frac{\binom{m}{2}}{\binom{N}{2}}. \end{aligned}$$

También definimos la variable aleatoria *máximo desvío de cogrado* (respecto a G_m) como siendo

$$D'_{\max}(G_m) := \max_{u,w \in E(K_n)} D_{uw}(G_m).$$

El siguiente resultado, que es mostrado en [21], es una consecuencia inmediata del Teorema 2.1, similar al Lema 3.2 sobre grados.

Lema 3.4 (Máximo Desvío de Grado en $\mathbf{G}(n, m)$). *Sea $G_m \in \mathcal{G}_{n,m}$ con $t = m/N \in (0, 1)$ y $b = b_n \geq \log n$ una secuencia real. Entonces, se cumple la siguiente desigualdad no-determinista*

$$|D'_{\max}(G_m)| \lesssim^{*b} b^{1/2} \sqrt{n} + b$$

Observación 3.2. *En [21] es mostrado una estimativa más fuerte que en el Lema 3.4, mostrando*

que el limitante superior puede ser tomado de orden $b^{1/2}\sqrt{tn} + b$. Sin embargo, esta estimativa está lejos de ser la mejor posible, como veremos en el Capítulo 4. Además en el caso denso que estamos interesados en este momento, el factor constante \sqrt{t} no afecta el resultado.

Del mismo modo que el Lema 3.3, el siguiente lema nos permitirá controlar (superiormente) la varianza del proceso aleatorio asociado a los incrementos $X_{\Delta}(\sigma_i)$ en nuestra representación martingala.

Lema 3.5 (Suma de Cuadrados de los Desvíos de Cogrados en $\mathbf{G}(n, m)$). Sea $G_m \in \mathcal{G}_{n,m}$ y $b = b_n \geq n$ una secuencia real. Entonces, se cumple que

$$\sum_{uw \in E(K_n)} D_{uw}(G_m)^2 \lesssim^{*b} bn^2.$$

Demostración del Lema 3.5:

Una parte importante de esta prueba, es lidiar con la correlación que existe entre las variables $\{d_{uw}(\mathbf{G}_m) : uw \in E(K_n)\}$. En esta demostración particionaremos el grafo K_n (o sus aristas) en subestructuras donde la subcolección de variables aleatorias es *esencialmente independiente*. Esto nos permitirá aplicar nuestros resultados de concentración a cada grupo de variables aleatorias para finalmente sumar todas las contribuciones. A seguir implementaremos esta estrategia con ciertas modificaciones técnicas.

Sea $\{M_i : i \in I\}$ una partición en *matchings* de K_n como descrita en la sección 2.2. Denotemos por

$$q(G_m) := \sum_{uw \in E(K_n)} D_{uw}(G_m)^2$$

. Análogamente, para cada $i \in I$, denotemos por $q_i(G_m)$ la contribución de M_i a $q(G)$, esto es,

$$q_i(G_m) := \sum_{uw \in M_i} D_{uw}(G_m)^2.$$

Así, $q = \sum_i q_i$. Supongamos pues, que el evento " $q(G_m) > C bn^2$ " a ocurrido (para algún $C > 0$), entonces

$$q/|I| > C' bn$$

ya que $|I|$ es de orden n .

Luego por el principio del palomar tenemos garantizado que para algún $i \in I$,

$$q_i(G_m) > C' bn.$$

Hemos mostrado que, por Union Bound (ver Proposición 2.1),

$$\begin{aligned} \mathbb{P}(q(G_m) > Cbn^2) &\leq \mathbb{P}(\exists i \in I, q_i(G_m) > C' bn) \\ &\leq \sum_{i \in I} \mathbb{P}(q_i(G_m) > C' bn) \end{aligned} \quad (3.15)$$

Por tanto, resta mostrar que para cada $i \in I$ se tiene que

$$\mathbb{P}(q_i(G_m) > C' bn) \leq \exp(-K'b) \quad (3.16)$$

para alguna constante $K' > 0$. En efecto, (3.16) junto con (3.15) implican que

$$\begin{aligned} \mathbb{P}(q(G_m) > Cbn^2) &\leq \sum_{i \in I} \mathbb{P}(q_i(G_m) > C' bn) \\ &\leq |I| \exp(-K'b) \\ &\leq n \exp(-K'b) \\ &\leq \exp(-Kb) \end{aligned}$$

para cierto $K > 0$, como deseábamos.

Así, nos queda el trabajo de probar (3.16). Esta parte de la prueba es similar al argumento del Lema 3.3 pues ambos escenarios son parecidos. El lector atento podrá notar que la demostración a seguir le es familiar. Fijado un *matching* $M = M_i$, procederemos a “linealizar diádicamente” el problema. Para cada vector

$$x = (x_{uw} : uw \in M)$$

donde las coordenadas x_{uw} toman valores de la forma $\pm 2^l$ con $l \in [\log n]$, definimos la variable aleatoria f_x como siendo

$$f_x(G_m) := \sum_{uw \in M} x_{uw} D_{uw}(G_m).$$

También definamos el vector aleatorio diádico $x^* = (x_{uw}^* : uw \in M)$ (asociado a G_m) dado por

$$x_{uw}^*(G_m) := \begin{cases} 0, & \text{si } |D_{uw}(G_m)| \leq \sqrt{n}, \text{ y} \\ \text{sign}(D_{uw}(G_m)) 2^j, & \text{si } 2^j \sqrt{n} \leq |D_{uw}(G_m)| < 2^{j+1} \sqrt{n}. \end{cases} \quad (3.17)$$

De la definición dada, tenemos las desigualdades a seguir:

$$f_{x^*}(G_m) \geq \sum_{uw \in M} (x_{uw}^*(G_m))^2 \sqrt{n}, \text{ y} \quad (3.18)$$

$$\sum_{uw \in M} D_{uw}(G_m)^2 \leq 4 \sum_{uw \in M} x_{uw}^*(G_m)^2 n. \quad (3.19)$$

Ahora bien, supongamos que se cumple el evento " $\sum_{uw \in M} D_{uw}(G_m)^2 > C' bn$ ". Entonces, de (3.19) y (3.18) deducimos que se cumple el evento "existe un vector diádico x tal que

$$\begin{aligned} f_x(G_m) &\geq \|x\|_2^2 \sqrt{n}, \\ \|x\|_2^2 &\geq C_1 b, \text{ y} \\ \|x\|_\infty &\leq \sqrt{n}. \end{aligned}$$

Denotando por $F'(x)$ el evento

$$"f_x(G_m) \geq \|x\|_2^2 \sqrt{n}",$$

hemos mostrado que

$$\begin{aligned} \mathbb{P} \left(\sum_{uw \in M} D_{uw}(G_m)^2 > C' bn \right) &\leq \mathbb{P} \left(\bigvee_{x: \|x\|_2 \geq C_2 \sqrt{b}} F'(x) \right), \\ &\leq \sum_{x: \|x\|_2 \geq C_2 \sqrt{b}} \mathbb{P}(F'(x)). \end{aligned} \quad (3.20)$$

Ahora pasamos a la siguiente afirmación, en orden de estimar 3.20.

Afirmación. $\mathbb{P}(F'(x)) \leq \exp(-K' \|x\|_2^2)$, para alguna constante $K' > 0$.

Demostración de la Afirmación:

Queremos mostrar que la variable aleatoria f_x está altamente concentrada en torno de su esperanza (esto es, 0).

Con el objetivo de usar el Corolario 2.1 definimos la siguiente función

$$\psi' : E(K_n) \rightarrow \mathbb{R}^+$$

que asigna a cada $e \in E(K_n)$ la suma indexada por uw de los módulos de x_{uw} , donde uw varia en M y son incidentes a e (denotemos este conjunto por $M(e)$).

Afirmamos que f_x es ψ' -Lipschitz.

En efecto, sean $G, G' \in \mathcal{G}_{n,m}$ con aristas distinguidas e y e' respectivamente. Así, podemos escribir $f_x(G)$ y $f_x(G')$ de la siguiente manera:

$$\begin{aligned} f_x(G) &= \sum_{uw \in M \setminus (M(e) \cup M(e'))} x_{uw} D_{uw}(G) + \sum_{uw \in M(e) \cup M(e')} x_{uw} D_{uw}(G) \\ f_x(G') &= \sum_{uw \in M \setminus (M(e) \cup M(e'))} x_{uw} D_{uw}(G') + \sum_{uw \in M(e) \cup M(e')} x_{uw} D_{uw}(G') \end{aligned}$$

Restando las dos expresiones anteriores obtenemos que

$$\begin{aligned} |f_x(G) - f_x(G')| &\leq \sum_{uw \in M(e) \cup M(e')} |x_{uw}| |D_{uw}(G) - D_{uw}(G')|, \\ &= \sum_{uw \in M(e) \cup M(e')} |x_{uw}| |d_{uw}(G) - d_{uw}(G')|, \\ &\leq \sum_{uw \in M(e) \cup M(e')} |x_{uw}|, \\ &\leq \sum_{uw \in M(e)} |x_{uw}| + \sum_{uw \in M(e')} |x_{uw}| \\ &= \psi'(e) + \psi'(e'). \end{aligned}$$

Acabamos de mostrar nuestra afirmación.

Las siguientes estimativas también serán útiles al aplicar el Corolario 2.1 en f_x :

$$\begin{aligned}
\|\psi\|_2^2 &= \sum_{e \in E(K_n)} \psi(e)^2 \\
&\leq 2 \sum_{e \in E(K_n)} \sum_{uw \in M(e)} x_{uw}^2 \\
&= 2 \sum_{uw \in M} \sum_{e: M(e) \ni uw} x_{uw}^2 \\
&\leq 4n \sum_{uw \in M} x_{uw}^2 = 4n \|x\|_2^2, \text{ y} \\
\|\psi\|_\infty &= \max_{e \in E(K_n)} \sum_{uw \in M(e)} |x_{uw}| \\
&\leq 2 \|x\|_\infty \\
&\leq 2\sqrt{n}.
\end{aligned}$$

Del Corolario 2.1 y aplicando las estimativas anteriores, obtenemos que

$$\begin{aligned}
\mathbb{P}(F'(x)) &= \mathbb{P}(f_x(G_m) > C' \|x\|_2^2 \sqrt{n}) \\
&\leq \exp\left(\frac{-C'' \|x\|_2^4 n}{\|\psi\|_2^2 + \|x\|_2^2 \sqrt{n} \|\psi\|_\infty}\right) \\
&\leq \exp\left(\frac{-C'' \|x\|_2^4 n}{4n \|x\|_2^2 + 2 \|x\|_2^2 n}\right) \\
&\leq \exp(-K'' \|x\|_2^2),
\end{aligned}$$

lo que muestra nuestra Afirmación.

Para finalizar la prueba de (3.16) basta mostrar que

$$\sum_{x: \|x\|_2 \geq C_2 \sqrt{b}} \exp(-K'' \|x\|_2^2) \leq \exp(-K' b) \tag{3.21}$$

para algún $K' > 0$, pues al combinar (3.20) con la Afirmación, obtendríamos la estimativa deseada. Así como fue comentado en la subsección de grados, dejaremos esta verificación para el próximo capítulo, donde usaremos una herramienta analítica elegante.

3.2.3. Probabilidad del Desvío de Conteo de Cerezas y Triángulos en $\mathbf{G}(n, m)$

En esta subsección mostraremos uno de los resultados principales de [21], cuyo enunciado sigue a continuación.

Teorema 3.2 (Probabilidad del Desvío de Conteo de \wedge y \triangle en $\mathbf{G}(n, m)$: Modelo Denso).

Sea $m \in [N]$. Para todo $\alpha > 0$ (suficientemente grande) se cumple que

$$-\log(\mathbb{P}(D_{\wedge}(G_m) > \alpha \times n^{3/2})) \geq C_1 \alpha \times \min(\alpha, \sqrt{n}) \quad (3.22)$$

$$-\log(\mathbb{P}(D_{\triangle}(G_m) > \alpha \times n^{3/2})) \geq C_2 \alpha \times \min(\alpha, \sqrt{n}) \quad (3.23)$$

donde $C_1, C_2 > 0$ son constantes absolutas.

Antes de pasar a la prueba, a la que dedicamos esta subsección, nos gustaría hacer algunas observaciones y/o comentarios de este resultado.

Como el lector puede observar, en el enunciado del Teorema 3.2 no existe ninguna restricción sobre la densidad $t_n = m/N$ de nuestro modelo. Sin embargo, este teorema es sólo justo en el caso de modelos densos. Esto, en parte, puede ser contemplado en la re-escala de $n^{3/2}$ presente en el desvío. Este valor es tomado pues tiene el mismo orden del desvío estándar de nuestras variables aleatorias en el modelo denso, lo que no sucede en modelos dispersos, donde el orden del desvío estándar es $t_n^{3/2} n^{3/2}$ para conteo de triángulos (e $t_n n^{3/2}$ para conteo de cerezas).

Esta fue una de las motivaciones iniciales que nos llevaron al estudio del modelo disperso, que desarrollaremos en el Capítulo 4.

También queremos hacer notar al lector que la función tasa de orden de magnitud

$$\alpha \times \min(\alpha, \sqrt{n})$$

que aparece en ambas desigualdades del Teorema 3.2 resulta más familiar una vez escrita por partes. Pongamos como ejemplo el caso de la variable aleatoria $N_{\triangle}(\mathbf{G}(n, m))$.

Podemos reescribir la ecuación (3.23) de la siguiente forma:

$$\mathbb{P}\left(\frac{N_{\triangle}(\mathbf{G}(n, m)) - \mu_{n,m}}{\sigma_{n,m}} > \alpha\right) \leq \begin{cases} \exp(-C \alpha^2), & \text{si } \alpha \leq \sqrt{n}, \text{ y} \\ \exp(-C \alpha \sqrt{n}), & \text{si } \alpha \geq \sqrt{n}. \end{cases} \quad (3.24)$$

donde $\mu_{n,m}$ y $\sigma_{n,m}$ son el valor esperado y el desvío estándar de $N_{\Delta}(\mathbf{G}(n, m))$ respectivamente.

Note que en (3.24) inicialmente la función (con argumento α) es cuadrática (sub-Gaussiano) en el exponente para luego ralentizarse y tornarse lineal (grandes desvíos). Este tipo de comportamiento es usual para muchas variables aleatorias que pueden ser representadas como sumas de variables aleatorias con poca dependencia entre ellas y sirve como heurística al estudiar variables aleatorias en las que se ha mostrado un Teorema Central del Límite, sin embargo esta heurística debe ser tomada con cierto cuidado pues entre la transición de desvíos moderados a grandes desvíos ciertos fenómenos pueden formar parte del interludio, como veremos en el siguiente capítulo.

Para finalizar los comentarios y pasar a la prueba, consideramos importante hacer notar al lector que el Teorema 3.2 sólo busca entender el orden de la función tasa, sin embargo en [21] los autores también gastan esfuerzos en entender su valor asintótico, objetivo que consiguen en el modelo denso cuando $\alpha \ll \sqrt{n}$. Este es un punto interesante de investigación que captó también nuestra atención pero que no será considerado en el trabajo presente (para más detalle, vea [21]).

Otro estimulante dirección que podemos tomar es el estudio de desvíos negativos. Este problema también es abordado en [21] como el lector ya podría sospechar pues las técnicas de Martingalas se aplican también para estos problemas. Sin embargo, el resultado sólo es justo para el modelo denso y valores cercanos al desvío estándar, dejando el problema de desvíos negativos mayores aún por optimizar.

En esta dirección, recientemente en [35] se han hecho progresos para el conteo de triángulos, donde la teoría de matrices aleatorias y grandes desvíos ha mostrado su eficacia. Para más detalles y una conjetura en esta línea de investigación, recomendamos al lector el artículo anteriormente citado [35].

Demostración del Teorema 3.2

La prueba del Teorema 3.2 es en esencia (una vez obtenida la representación martingala Teorema 3.1) una aplicación de la Desigualdad de Freedman (Teorema 2.2). De esta forma, nuestra tarea se reduce a estimar tanto los incrementos como la varianza previsible asociada al proceso martingala.

Para beneficio del lector, organizaremos estas estimativas en un lema para finalmente aplicarlo en la ya mencionada desigualdad de Freedman.

Para ambas variables aleatorias $D_{\wedge}(\mathbf{G}(n, m))$ y $D_{\Delta}(\mathbf{G}(n, m))$ son usados los mismos argumentos, por tanto sólo nos enfocaremos en el conteo de triángulos.

En virtud de la representación martingala Teorema 3.1 (3.4), definimos las variables aleatorias,

$$\mathbb{X}_\Delta(i; m) := 3 \frac{(N-m)_2(m-i)}{(N-i)_3} X_\wedge(\sigma_i) + \frac{(N-m)_3}{(N-i)_3} X_\Delta(\sigma_i). \quad (3.25)$$

Como ya mencionamos en el mismo teorema, estas variables son incrementos respecto a la filtración natural dada por el proceso Erdős-Rényi. Así, podemos escribir

$$D_\Delta(G_m) = \sum_{i=1}^m \mathbb{X}_\Delta(i; m).$$

Luego, resta estimar el comportamiento (no-determinista) de los incrementos $\mathbb{X}_\Delta(i; m)$ y de su varianza previsible $V_m = \sum_{i=1}^m \mathbb{E} [\mathbb{X}_\Delta(i; m)^2 | G_{i-1}]$. Note que de (3.25) podemos observar que

$$\begin{aligned} |\mathbb{X}_\Delta(i; m)| &\leq 3 |X_\wedge(\sigma_i)| + |X_\Delta(\sigma_i)| \quad \text{y} \\ \mathbb{X}_\Delta(i; m)^2 &\leq 18 X_\wedge(\sigma_i)^2 + 2 X_\Delta(\sigma_i)^2, \end{aligned}$$

esto es inmediato pues $\frac{(N-m)_2(m-i)}{(N-i)_3}$ (resp. $\frac{(N-m)_3}{(N-i)_3}$) es limitada por 1, adicionalmente hemos usado que $(a+b)^2 \leq 2a^2 + 2b^2$. En particular, obtenemos que

$$\max_i |\mathbb{X}_\Delta(i; m)| = O(\max_i |X_\wedge(\sigma_i)| + \max_i |X_\Delta(\sigma_i)|) \quad (3.26)$$

$$\mathbb{E} [\mathbb{X}_\Delta(i; m)^2 | G_{i-1}] = O(\mathbb{E} [X_\wedge(\sigma_i)^2 | G_{i-1}] + \mathbb{E} [X_\Delta(\sigma_i)^2 | G_{i-1}]) \quad (3.27)$$

Ahora bien, por definición de las variables aleatorias $X_\wedge(\sigma_i)$ y $X_\Delta(\sigma_i)$ tenemos que

$$\begin{aligned} |X_\wedge(\sigma_i)| &\leq |A_\wedge(\sigma_i)| + |\mathbb{E} [A_\wedge(\sigma_i) | G_{i-1}]| \\ &\leq 2 \|A_\wedge(\sigma_i)\|_\infty = O(n) \end{aligned}$$

desde que $A_\wedge(\sigma_i)$ cuenta el número de cerezas (inmersas) creadas al adicionar la arista e_i . Similarmente obtenemos que $|X_\Delta(\sigma_i)| = O(n)$. Esto sugiere que podemos considerar $R = C_1 n$, donde $C_1 > 0$ es una constante suficientemente grande.

Para estimar probabilísticamente la varianza previsible V_m , necesitaremos el siguiente lema.

Lema 3.6 (Varianza Condicionada de $X_\wedge(\sigma_i)$ y $X_\Delta(\sigma_i)$: orden de magnitud). *Sea $i \in [N/2]$. Para todo $\lambda \geq 1$ se cumplen las siguientes desigualdades probabilistas,*

$$\mathbb{E} [X_\wedge(\sigma_i)^2 | G_{i-1}] \lesssim^{*\lambda\sqrt{n}} \max\{\lambda, \sqrt{n}\} \sqrt{n} \quad (3.28)$$

$$\mathbb{E} [X_\Delta(\sigma_i)^2 | G_{i-1}] \lesssim^{*\lambda\sqrt{n}} \max\{\lambda, \sqrt{n}\} \sqrt{n} \quad (3.29)$$

Observación 3.3. *Nuevamente, el uso de la función ‘máximo’ presente en el lema anterior es para compactar la información de modo a utilizarlo de una forma más limpia. Sin embargo, es esclarecedor hacer notar al lector que (3.28) es (esencialmente) equivalente a decir que:*

$$\mathbb{E} [X_\wedge(\sigma_i)^2 | G_{i-1}] \lesssim^{*\lambda\sqrt{n}} \lambda \sqrt{n},$$

para todo $\lambda \geq \sqrt{n}$. Lo que tiene un aspecto bastante más familiar con “grandes desvíos”, pero al requerir aplicar estas estimativas para valores de λ menores a \sqrt{n} , optamos por la notación usada en el Lema 3.6 (mismo motivo para la ecuación (3.29)).

Demostración del Lema 3.6:

Debido a la naturaleza combinatoria de las variables aleatorias $X_\wedge(\sigma_i)$ y $X_\Delta(\sigma_i)$ resulta plausible esperar una relación directa entre estas con las variables grado y cogrado. Así, nuestra primera afirmación exhibe esta relación.

Afirmación. Para todo $1 \leq i \leq N/2$, tenemos las siguientes igualdades deterministas

$$\mathbb{E} [X_\wedge(\sigma_i)^2 | G_{i-1}] = O \left(\frac{\sum_{u \in [n]} D_u (G_{i-1})^2}{n} \right) \quad (3.30)$$

$$\mathbb{E} [X_\Delta(\sigma_i)^2 | G_{i-1}] = O \left(\frac{\sum_{uw \in E(K_n)} D_{uw} (G_{i-1})^2}{n^2} \right) \quad (3.31)$$

Demostración de la Afirmación:

Ambas igualdades tienen esencialmente la misma prueba, por tanto sólo mostraremos el argumento de la primera.

Iniciamos recordando al lector la definición de $X_\wedge(\sigma_i)$, esto es,

$$X_\wedge(\sigma_i) = A_\wedge(\sigma_i) - \mathbb{E}[A_\wedge(\sigma_i) | G_{i-1}]$$

donde $A_\wedge(\sigma_i) = 2 \sum_{u \in e_i} d_u(G_{i-1})$ cuenta el número de copias (inmersas) de cerezas creadas al seleccionar la arista e_i . Así,

$$\begin{aligned} \mathbb{E}[X_\wedge(\sigma_i)^2 | G_{i-1}] &= \text{Var}[A_\wedge(\sigma_i) | G_{i-1}] \\ &= 2 \text{Var} \left[\sum_{u \in e_i} d_u(G_{i-1}) | G_{i-1} \right] \\ &= 2 \text{Var} \left[\sum_{u \in e_i} D_u(G_{i-1}) | G_{i-1} \right] \\ &\leq 2 \mathbb{E} \left[\left(\sum_{u \in e_i} D_u(G_{i-1}) \right)^2 | G_{i-1} \right] \\ &\leq 4 \mathbb{E} \left[\sum_{u \in e_i} D_u(G_{i-1})^2 | G_{i-1} \right] \\ &= 4 \mathbb{E} \left[\sum_{u \in [n]} D_u(G_{i-1})^2 \mathbf{1}_{\{u \in e_i\}} | G_{i-1} \right] \\ &= 4 \sum_{u \in [n]} D_u(G_{i-1})^2 \mathbb{P}(u \in e_i | G_{i-1}) \end{aligned}$$

donde hemos usado las propiedades listados en la Proposición 2.1.

Resumiendo, hemos llegado a la siguiente expresión cruda

$$\mathbb{E}[X_\wedge(\sigma_i)^2 | G_{i-1}] = O \left(\sum_{u \in [n]} D_u(G_{i-1})^2 \times \mathbb{P}(u \in e_i | G_{i-1}) \right).$$

Luego, basta verificar que $\mathbb{P}(u \in e_i | G_{i-1}) = O\left(\frac{1}{n}\right)$, para todo $u \in [n]$. A seguir mostramos las cuentas que culminan la prueba de la afirmación

$$\begin{aligned}\mathbb{P}(u \in e_i | G_{i-1}) &= \frac{n-1-d_u(G_{i-1})}{N-i+1} \\ &\leq \frac{n}{N/2} = O\left(\frac{1}{n}\right).\end{aligned}$$

Ahora veremos como al juntar la Afirmación anterior junto con los resultados de las secciones de grados y cogrados (Lemas 3.3 y 3.5) deducimos nuestro lema deseado.

Sea $b := \max\{\lambda, \sqrt{n}\} \times \sqrt{n}$. Note que

- (1) $b \geq n$, y
- (2) $b \geq \lambda \sqrt{n}$.

Debido a (1) podemos aplicar el Lema 3.3, con $i-1$ en lugar de m , de modo que se cumple la desigualdad (no-determinista)

$$\frac{\sum_{u \in [n]} D_u(G_{i-1})^2}{n} \lesssim^{*b} \max\{\lambda, \sqrt{n}\} \times \sqrt{n},$$

De las propiedades básicas del orden asintótico (probabilístico) " \lesssim^{*b} " y el item (2), inferimos que

$$\frac{\sum_{u \in [n]} D_u(G_{i-1})^2}{n} \lesssim^{*\lambda \sqrt{n}} \max\{\lambda, \sqrt{n}\} \times \sqrt{n}. \quad (3.32)$$

Para finalizar, basta juntar la Afirmación (3.30) y (3.32) para obtener la desigualdad (3.29), esto es,

$$\begin{aligned}\mathbb{E}[X_\wedge(\sigma_i)^2 | G_{i-1}] &= O\left(\frac{\sum_{u \in [n]} D_u(G_{i-1})^2}{n}\right) \\ &\lesssim^{*\lambda \sqrt{n}} \max\{\lambda, \sqrt{n}\} \times \sqrt{n}.\end{aligned}$$

Demostrado el Lema 3.6, podemos retomar la demostración del Teorema 3.2. Colectando

las estimativas obtenidas hasta el momento en términos de $\mathbb{X}_\Delta(i, m)$, obtenemos que:

$$\begin{aligned} \max_i \|\mathbb{X}_\Delta(i, m)\|_\infty &= O(\max_i \|X_\wedge(\sigma_i)\|_\infty + \max_i \|X_\Delta(\sigma_i)\|_\infty) \\ &= O(n), \text{ y} \end{aligned} \tag{3.33}$$

$$\begin{aligned} V_m &= \sum_{i=1}^m \mathbb{E} [\mathbb{X}_\Delta(i, m)^2 | G_{i-1}] \\ &= O\left(\sum_{i=1}^m \{\mathbb{E} [X_\wedge(\sigma_i)^2 | G_{i-1}] + \mathbb{E} [X_\Delta(\sigma_i)^2 | G_{i-1}]\}\right) \\ &\lesssim^* \lambda \sqrt{n} m \times \max\{\lambda, \sqrt{n}\} \sqrt{n} \\ &\leq \max\{\lambda, \sqrt{n}\} n^{5/2}. \end{aligned} \tag{3.34}$$

Por tanto, aplicando el Teorema 2.2 (Desigualdad de Freedman) con $\alpha = \lambda n^{3/2}$, $R = C_1 n$ y $\beta = C_2 \max\{\lambda, \sqrt{n}\} n^{5/2}$, obtenemos las siguientes desigualdades (como en la Desigualdad de Freedman, denote por S_i las sumas parciales del Martingal):

$$\begin{aligned} &\mathbb{P}\left(D_\Delta(G_m) > \lambda \times n^{3/2}\right) \\ &\leq \mathbb{P}\left(S_m - S_0 > \lambda \times n^{3/2} \text{ y } V_m \leq \beta\right) + \mathbb{P}(V_m > \beta) \\ &\leq \mathbb{P}\left(\exists i \geq 0 : S_i - S_0 > \lambda \times n^{3/2} \text{ y } V_i \leq \beta\right) + \exp(-\Omega(\lambda \sqrt{n})) \\ &\leq \exp\left(-\Omega\left(\frac{\lambda^2 n^3}{\max\{\lambda, \sqrt{n}\} n^{5/2} + n \times \lambda n^{3/2}}\right)\right) + \exp(-\Omega(\lambda \sqrt{n})) \\ &\leq \exp\left(-\Omega\left(\frac{\lambda^2 n^3}{\max\{\lambda, \sqrt{n}\} n^{5/2}}\right)\right) + \exp(-\Omega(\lambda \sqrt{n})) \\ &\leq \exp\left(-\Omega\left(\frac{\lambda^2 \sqrt{n}}{\max\{\lambda, \sqrt{n}\}}\right)\right) + \exp(-\Omega(\lambda \sqrt{n})) \\ &\leq \exp(-\Omega(\lambda \min\{\lambda, \sqrt{n}\})), \end{aligned}$$

donde hemos usado las estimativas dadas por (3.33) y (3.34).
Esto completa la demostración del teorema. ■

Capítulo 4

Nuevos resultados

En este capítulo extenderemos los resultados presentados en el Capítulo 3¹. Así como en [21] nuestro estudio es determinar la probabilidad del desvío de número de triángulos en $\mathbf{G}(n, m)$ (y $\mathbf{G}(n, p)$), pero con la diferencia que nosotros nos enfocaremos en el caso disperso ($n^{3/2}(\log n)^{1/2} \ll m \ll n^2$). Mas aún, para el modelo $\mathbf{G}(n, m)$, estudiamos (esencialmente) todos los desvíos posibles (moderados, grandes y más allá). Una de nuestras principales contribuciones en este estudio es mostrar que existen diferentes regiones (específicamente 4) donde el desvío es explicado por una de ellas : región “suerte” (región no-localizada), región “clique”, región “estrella” y región “centro” (regiones localizadas). Como veremos, para este estudio debemos eludir ciertas dificultades que no aparecen en el caso denso, lo que torna la prueba (rigurosa) muy técnica. Así, en el presente capítulo sólo mostraremos algunos resultados, asumiendo ciertos enunciados preparatorios, también nos tomaremos la libertad de no entrar en muchos detalles con el objetivo de pasar al lector la idea de la prueba de una forma más clara y menos agotadora.

4.1. Resultados

En esta sección enunciaremos y explicaremos los resultados de nuestro trabajo. Para ello necesitaremos introducir algunos objetos matemáticos. Por conveniencia, denotemos por $\ell := \log(1/t)$.

Definición 4.1 (Funciones y Regiones : N, S, H y C). *Sea $t \in (0, 1)$ y $b \geq 0$. Definimos las*

¹Estas nuevas ideas son consecuencia de una colaboración conjunta con S. Griffiths (PUC-RJ) y L. G. de Oliveira (PUC-RJ).

funciones $N(b, t)$, $S(b, t)$, $H(b, t)$ y $C(b, t)$ como siendo

$$\begin{aligned} N(b, t) &:= b^{1/2} t^{3/2} n^{3/2}, \\ S(b, t) &:= \frac{b^2}{\ell^2} t \mathbf{1}_{b \leq n\ell}, \\ H(b, t) &:= \frac{b t n}{\ell} \mathbf{1}_{b \geq n\ell}, \text{ y} \\ C(b, t) &:= \frac{b^{3/2}}{\ell^{3/2}}. \end{aligned}$$

Estas funciones serán llamadas de normal, estrella, centro (hub) y clique.

Denotaremos por $M(b, t)$ a la función máximo de estas cuatro, es decir

$$M(b, t) := \max\{N(b, t), S(b, t), H(b, t), C(b, t)\}.$$

Antes de continuar con las definiciones, nos gustaría dar una intuición de donde surgen estas funciones. En la literatura ya podemos observar que ciertos desvíos grandes pueden ser “explicados” por estructuras que aparecen dentro del grafo aleatorio, fenómeno conocido actualmente como *localización*. Por otra parte, entorno al desvío estándar, la causa es sólo debida al “azar” y al no tener una explicación estructural este tipo de fenómenos son conocidos como “no-localizados”. En nuestro estudio, que abarca desvíos moderados, grandes y más allá, parece plausible que aparezcan estos dos fenómenos, de hecho este es el caso, como pasamos a esbozar:

Pura Suerte.

Como ya indicamos, existe un TCL para conteo de triángulos. Esto sucede para desvíos de la ‘forma’ $b^{1/2} t^{3/2} n^{3/2}$ donde $b > 0$ está fijado (y es suficientemente grande). En este caso, el costo (probabilidad) por tal desvío es e^{-b} . La intuición por tras de la función

$$N(b, t) = b^{1/2} t^{3/2} n^{3/2}$$

es que este comportamiento se debería extender por más tiempo en el plano (b, t) . Debido a la representación martingala de $D_{\Delta}(G_m)$, de la cual disponemos, resulta plausible que un argumento en la línea de la desigualdad de Freedman nos permita mostrar esta

heurística por un tiempo significativo.²

Una Estrella.

Imagine que ‘plantamos’ una estrella grande, de grado $d < n$, en nuestro grafo aleatorio $\mathbf{G}(n, m)$. En promedio, esta estrella causaría un desvío de $\Theta(d^2 t)$ triángulos. Ya que el costo por plantar una tal estrella es de $\approx t^d = e^{-d\ell}$, el desvío que conseguiríamos ‘pagando’ e^{-b} es del orden

$$S(b, t) = d^2 t = \left(\frac{b}{\ell}\right)^2 t 1_{b \leq n\ell}.$$

Un Centro (Hub).

Un centro (hub) es una extensión natural del caso estrella cuando $d = n - 1$. Imagine que tenemos un grupo de k vértices con grado $n - 1$. En promedio, este *hub* causaría un desvío de orden ktn^2 . Ya que el costo por plantar un tal centro (hub) es de $\approx t^{kn} = e^{-kn\ell}$, el desvío que conseguiríamos ‘pagando’ e^{-b} es del orden

$$H(b, t) = \frac{bkn}{\ell} 1_{b \geq n\ell}.$$

Un Clique.

Otra forma eficiente de crear muchos triángulos es a través de cliques. Imagine que ‘plantamos’ una clique con k vértices en nuestro grafo aleatorio $\mathbf{G}(n, m)$. Esta clique crearía instantáneamente $\Theta(k^3)$ triángulos. Ya que el costo por plantar un tal clique es de $\approx t^{k^2} = e^{-k^2\ell}$, el desvío que conseguiríamos ‘pagando’ e^{-b} es del orden

$$C(b, t) = k^3 = \left(\frac{b}{\ell}\right)^{3/2}.$$

Ahora bien, si disponemos de un ‘saldo’ de e^{-b} para ‘pagar’ por un desvío, lo más natural sería ‘comprar’ el desvío más alto posible. De aquí la definición de M .

Para enunciar de una forma limpia nuestro resultado principal necesitamos de más una definición.

²Esto es compatible con el hecho que la varianza previsible de nuestro martingala es de orden $t^3 n^3$.

Definición 4.2 (Función Desvío). Sea $t \in (0, 1)$ y $b \geq 0$. Definimos la función $DEV_{\Delta}(b, t)$ como siendo

$$DEV_{\Delta}(b, t) := \min\{a \geq 0 : \mathbb{P}(D_{\Delta}(\mathbf{G}(n, m)) \geq a) \leq \exp(-b)\}.$$

De esta manera, nuestro resultado principal nos dice que $DEV_{\Delta}(b, t)$ tiene orden de magnitud $M(b, t)$, a través de un amplio margen de t y b .

Teorema 4.1. Para toda secuencia $t = \Omega(n^{-1/2}(\log n)^{1/2})$ y $3 \log n \leq b \leq tn^2\ell$, se cumple que

$$DEV_{\Delta}(b, t) = \Theta(M(b, t)).$$

Como una aplicación del Teorema 4.1, obtenemos estimativas para la probabilidad de desvíos de conteo de triángulos (y cerezas) en el modelo binomial $G(n, p)$. En la Introducción 1 dimos un panorama general de este problema para el modelo binomial. Para beneficio del lector, substraeremos los puntos cruciales para esta sección.

Definición 4.3 (Función Tasa para Conteo de Triángulos). Sean $p \in (0, 1)$ y $\delta \in (0, 1)$ dos números reales. Definimos la función tasa para conteo de triángulos $r = r(\delta, p, n)$ como siendo

$$r(\delta, p, n) := -\log \mathbb{P}(N_{\Delta}(\mathbf{G}(n, p)) > (1 + \delta)p^3(n)_3).$$

Análogamente podemos definir la función tasa para conteo de cerezas substituyendo la variable aleatoria $N_{\Delta}(\mathbf{G}(n, p))$ por $N_{\wedge}(\mathbf{G}(n, p))$.

Nosotros estamos interesados en estudiar el caso disperso $p = p_n = o_n(1)$ cuando $\delta = \delta_n$ varia en el intervalo $p^{-1/2}n^{-1} \ll \delta_n \leq 1$ (desde el desvío estándar hasta grandes desvíos).

Uno de los primeros resultados relevantes es el de Doering y Eichelsbacher [13] donde fue probado el siguiente resultado asintótico:

$$r(\delta_n, p, n) \sim \frac{\delta_n^2 p n^2}{36(1-p)}, \tag{4.1}$$

siempre que $p^{-1/2}n^{-1} \ll \delta_n \ll p^7$ (vea también [14]). Note que la expresión a la derecha en (4.1) cuando $p = o(1)$ es $(\frac{1}{36} + o(1))\delta_n^2 pn^2$. Nuestro primer resultado principal (para el modelo binomial) nos dice que (4.1) se satisface para toda la región que llamamos de “no-localizada” al no presentar una causa estructural que explique el desvío. Denotemos por

$$\begin{aligned}\delta_n^{(1)}(p) &= p^{-1/2}n^{-1}, \text{ y} \\ \delta_n^{(2)}(p) &= \min\{p^{3/4}n^{3/4}, n^{-1/3}(\log n)^{2/3} + p \log 1/p\}.\end{aligned}$$

Teorema 4.2 (Región no-localizada para Conteo de Triángulos en el Modelo Binomial Disperso). *Sea $n^{-1/2} \log n \ll p \ll 1$ y sea δ_n una secuencia de números reales satisfaciendo que*

$$\delta_n^{(1)}(p) \ll \delta_n \ll \delta_n^{(2)}(p).$$

Entonces, se cumple la siguiente igualdad asintótica

$$r(\delta_n, p, n) = \left(\frac{1}{36} + o(1)\right) \delta_n^2 pn^2. \quad (4.2)$$

Observación 4.1. *De hecho, nuestros métodos nos permiten estimar con mayor detalle el término erro $o(\delta_n^2 pn^2)$ que aparece en (4.2).*

Para la región localizado, esto es, donde la causa es explicada por la presencia de alguna estructura, fueron Lubetzky y Zhao [32] los que consiguieron un primer resultado de este problema, obteniendo que

$$r(\delta, p, n) = (1 + o_n(1)) \min\left\{\frac{\delta^{2/3}}{2}, \frac{\delta}{3}\right\} p^2 n^2 \log(1/p) \quad (4.3)$$

cuando $n^{-1/42} \log n \leq p = o_n(1)$ e $\delta > 0$ está fijo.³

³En trabajos más recientes extendió estos resultados para valores de p del orden $\omega(n^{-1/2}(\log n)^2)$.

En esta región, nuestras técnicas no implican una estimativa asintótica para la función tasa $r(\delta_n, p, n)$, pero si nos permiten determinar el orden de magnitud de dicha función en toda la región restante.

Teorema 4.3 (Región Localizada para Conteo de Triángulos en el Modelo Binomial Disperso). *Sea $n^{-1/2} \log n \ll p \ll 1$ y sea δ_n una secuencia de números reales satisfaciendo*

$$\delta_n^{(2)}(p) \ll \delta_n \leq 1.$$

Entonces, tenemos que

$$r(\delta_n, p, n) \asymp \min\{\delta_n^{2/3} p^2 n^2 \log n, \delta_n^{1/2} p n^{3/2} \log n + \delta_n p^2 n^2 \log(1/p)\} \quad (4.4)$$

La idea central para probar estos resultados es la misma que en [21], esto es, pensar en $\mathbf{G}(n, p)$ como un promedio ponderado de $\mathbf{G}(n, m)$, usar los resultados obtenidos para el modelo uniforme $\mathbf{G}(n, m)$, y finalmente estimar la probabilidad en el modelo original $\mathbf{G}(n, p)$ (vía este promedio). La diferencia crucial con nuestro trabajo, es que nosotros hemos mejorado y ampliado las estimativas de las probabilidades en el modelo uniforme $\mathbf{G}(n, m)$, lo que nos permite ir más allá en el modelo binomial $\mathbf{G}(n, p)$. Sin más, nosotros nos enfocaremos en el modelo uniforme, dejando los detalles de la transferencia al modelo binomial para el lector curioso, que puede hallar la demostración completa en [1].

4.2. Nuevas Ideas

Al final de la Introducción 1 explicamos en palabras la idea general de nuestros resultados para el modelo uniforme $\mathbf{G}(n, m)$. Antes de comenzar a enunciar y probar resultados, refrescaremos al lector los puntos cruciales de nuestra estrategia.

Dada la representación martingala de $D_\Delta(G_m)$, podemos intentar aplicar directamente la desigualdad de Freedman para estimar la probabilidad de desvío de la orden $M = M(b, t)$. Desafortunadamente, la cantidad “ RM ” que aparece en el exponente de la desigualdad puede ser demasiado grande, en el sentido que supera a la varianza previsible (del proceso). Esto es un obstáculo, pues el exponente resulta menor que lo esperado, esto es b . Para rodear este aparente inconveniente, usamos el método de truncamiento también conocido como “*Divide and Conquer*”, nombre que aparece en la literatura en [28]. La idea tras este método es separar la contribución de eventos inusuales (o raros) de los eventos

comunes. Así, la contribución de los eventos comunes son tratados con las técnicas de concentración tradicionales, en cuanto que la contribución de los eventos inusuales son tratados por separado, usualmente con técnicas relativas a grandes desvíos. En nuestro contexto, el truncamiento natural es el siguiente. Denotemos por

$$K_{\wedge} = K_{\wedge}(b, t) := \frac{M}{tb}$$

$$K_{\Delta} = K_{\Delta}(b, t) := \frac{M}{b}$$

Lo que deseamos ‘truncar’ son los incrementos $X_{\Delta}(\sigma_i)$ e de nuestra representación martingala de $D_{\Delta}(G_m)$ cuando estos sobrepasan el valor de K . Note que esto es el valor natural para definir K_{Δ} , pues deseamos que

$$\frac{M^2}{KM} \asymp b$$

en la desigualdad de Freedman (análogamente para K_{\wedge}). Así, definimos nuestros truncamientos

$$X'_{\wedge}(\sigma_i) := X_{\wedge}(\sigma_i) \mathbf{1}_{\{X_{\wedge}(\sigma_i) \leq K_{\wedge}\}} \text{ y}$$

$$X'_{\Delta}(\sigma_i) := X_{\Delta}(\sigma_i) \mathbf{1}_{\{X_{\Delta}(\sigma_i) \leq K_{\Delta}\}}.$$

También consideramos las variables aleatorias

$$Z'_{\wedge}(\sigma_i) := X_{\wedge}(\sigma_i) \mathbf{1}_{\{X_{\wedge}(\sigma_i) > K_{\wedge}\}} \text{ y}$$

$$Z'_{\Delta}(\sigma_i) := X_{\Delta}(\sigma_i) \mathbf{1}_{\{X_{\Delta}(\sigma_i) > K_{\Delta}\}}.$$

De este modo,

$$X_{\wedge}(\sigma_i) = X'_{\wedge}(\sigma_i) + Z_{\wedge}(\sigma_i) \text{ y}$$

$$X_{\Delta}(\sigma_i) = X'_{\Delta}(\sigma_i) + Z_{\Delta}(\sigma_i).$$

Así, podemos expresar $D_{\Delta}(G_m)$ como

$$D_{\Delta}(G_m) = D'_{\Delta}(G_m) + N_{\Delta}^*(G_m),$$

donde

$$D'_{\Delta}(G_m) := \sum_{i=1}^m \left\{ 3 \frac{(N-m)_2 (m-i)}{(N-i)_3} X'_{\wedge}(\sigma_i) + \frac{(N-m)_3}{(N-i)_3} X'_{\Delta}(\sigma_i) \right\}$$

y

$$N^*_{\Delta}(G_m) := 3 \sum_{i=1}^m \frac{(N-m)_2 (m-i)}{(N-i)_3} Z_{\wedge}(\sigma_i) + \sum_{i=1}^m \frac{(N-m)_3}{(N-i)_3} Z_{\Delta}(\sigma_i).$$

De esta forma, nuestra tarea se reduce a mostrar separadamente que

$$D'_{\Delta}(G_m) \lesssim^{*b} M \text{ y} \tag{4.5}$$

$$N^*_{\Delta}(G_m) \lesssim^{*b} M. \tag{4.6}$$

para $t \gg n^{-1/2} \log n$ y $3 \log n \leq b \leq tn^2 \ell$.

Para mostrar (4.5), note que $D'_{\Delta}(G_m)$ es representada por un supermartingala pues los incrementos sólo se han visto reducidos. Afortunadamente la desigualdad de Freedman se aplica también a supermartingalas. Luego, en esta etapa aplicamos directo nuestra desigualdad (previamente debemos estimar nuestra varianza previsible).

Ya para (4.6) es requerido un análisis más detallado del proceso Erdős-Rényi junto con algunas observaciones adicionales sobre el costo de las contribuciones. En la literatura previa, específicamente [6], el autor estudia grandes desvíos en el conteo de triángulos clasificando los triángulos de acuerdo con ciertas “anomalías” como pueden ser que uno de sus vértices tenga un grado muy grande o (caso contrario) que el cogrado sea grande. Esta forma de contar es inteligente, pues controlar los triángulos del primer tipo es relativamente fácil a través de los grados, ya para el segundo tipo un esfuerzo adicional es requerido debido a la correlación entre los cogrados, pero esto es resuelto al pasar a subestructuras donde los cogrados no presentan mucha correlación (y son lo suficientemente grandes para explicar el desvío).

Nosotros presentamos una idea diferente, pero no completamente desconectada de la anterior. En nuestro caso, vamos a clasificar pares de vértices (aristas y no-aristas) respecto a si su cogrado es explicado por el grado (de sus extremos) o (caso contrario) la grandeza del cogrado ‘realmente’ viene de él. Nuestros triángulos serán automáticamente clasificados por este criterio (si tiene una arista con gran desvío explicado o no el grado). Esta clasificación simplifica mucho el análisis de la variable aleatoria $N^*_{\Delta}(G_m)$, pues al igual que en la literatura previa podemos controlar los triángulos del primer tipo (con desvío de cogrado explicado por el grado) usando resultados de grados (que obtendremos en nuestra sección de grados a seguir), mientras que los de segundo podrían ser explicados

al pasar a subestructuras con cogrado casi independientes (sin embargo, un nuevo tipo de subestructura tiene que ser estudiada para obtener todos nuestros resultados deseados). Esto último no es del todo cierto! Como veremos, existen desvíos cuyo ‘costo’ se genera al seleccionar un gran número de pares de vértices durante el proceso Erdős-Rényi, y es esto lo que explica la probabilidad de e^{-b} . Afortunadamente, un argumento de acoplamiento (estocástico) nos permite comparar este proceso con una distribución binomial adecuada, lo que nos lleva a entender satisfactoriamente el comportamiento asintótico de esta contribución restante.

Antes de comenzar a enunciar los resultados preliminares (y juntar todo en la prueba final) vamos a detallar un poco más el argumento e introducir cierto lenguaje que nos será útil en la prueba de (4.6). Sea $i \in [m]$. Dado un par de vértices (distintos) f , decimos que f es i -bad (relativo a m) se

$$|D_f(G_i)| > Ct \sum_{u \in f} |D_u(G_i)|.$$

donde G_i es el i -ésimo grafo de una realización en el proceso Erdős-Rényi y C es una contante absoluta (grande). Por otro lado, se f no satisface la propiedad anterior diremos que f es i -good (relativo a m), i.e., $|D_f(G_i)| \leq Ct \sum_{u \in f} |D_u(G_i)|$. Note que las nociones de par *good* y *bad* se corresponden con nuestra intuición de ‘cogrado explicado por el grado’ y ‘cogrado que realmente viene de él’. Con estas nociones bien establecidas ya podemos descomponer el segundo sumando de la variable aleatoria $N_{\Delta}^*(G_m)$ de modo que una de las partes es controlada por los grados (grandes) mientras que la segunda parte será analizada por nuestros (nuevos) resultados de cogrados que presentaremos en breve (junto con argumentos de acoplamientos para estimar el proceso). Así, pasamos a definir las siguientes variables aleatorias.

$$N_{\text{good}}^*(m) := \sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{d_{e_i}(G_{i-1}) > K_{\Delta}/6\}} \mathbf{1}_{\{\exists j=i-1, \dots, m-1, e_i \text{ es } j\text{-good}\}} \quad (4.7)$$

$$N_{\text{bad}}^*(m) := \sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{d_{e_i}(G_{i-1}) > K_{\Delta}/6\}} \mathbf{1}_{\{\forall j=i-1, \dots, m-1, e_i \text{ es } j\text{-bad}\}} \quad (4.8)$$

Para ver la relación entre $N_{\Delta}^*(G_m)$, $N_{\text{good}}^*(m)$ y $N_{\text{bad}}^*(m)$, note que de la definición de $X_{\Delta}(\sigma_i)$, tenemos que

$$X_{\Delta}(\sigma_i) \leq 6 d_{e_i}(G_{i-1}).$$

Luego, si $X_{\Delta}(\sigma_i) > K_{\Delta}$, entonces $d_{e_i}(G_{i-1}) > K_{\Delta}/6$. De ahí que el segundo sumando de

$N_{\Delta}^*(G_m)$ es limitado (superiormente) por

$$\sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{d_{e_i}(G_{i-1}) > K_{\Delta}/6\}}$$

Por lo tanto, usando el hecho que $\mathbf{1} = \mathbf{1}_E + \mathbf{1}_{E^c}$, deducimos que el segundo sumando de $N_{\Delta}^*(G_m)$ es limitado por la suma de las variables aleatorias

$$N_{\text{good}}^*(m) + N_{\text{bad}}^*(m).$$

De la definición de “goodness”, es fácil mostrar que

$$N_{\text{good}}^*(m) = O\left(t \sum_v d_v(G_m)^2 \mathbf{1}_{\{d_v(G_m) = \Omega(K_{\wedge})\}}\right). \quad (4.9)$$

Efectivamente, de la definición de “good”, si e_i es j -good (para $i - 1 \leq j \leq m - 1$) y $d_{e_i}(G_{i-1}) > K_{\Delta}/6$, entonces (por monotonía)

$$\begin{aligned} d_{e_i}(G_{i-1}) &\leq d_{e_i}(G_j) \\ &\leq t^2 n + D_{e_i}(G_j) \\ &\leq t^2 n + C t \sum_{u \in e_i} (d_u(G_j) + t n) \\ &= O\left(t^2 n + t \sum_{u \in e_i} d_u(G_m)\right) \\ &= O\left(t \sum_{u \in e_i} d_u(G_m)\right) \end{aligned}$$

donde la última desigualdad es deducida de la anterior, pues esta implica que $\sum_{u \in e_i} d_u(G_m)$ es $\Omega(K_{\wedge})$ (y notando que $K_{\wedge} \gg t n$, cuando $t \ll 1$). Mas aún, si para algún $u \in e$ ($e \in E(G_m)$) tenemos que $d_u(G_m) \geq C K_{\wedge}$, entonces es fácil ver que

$$\sum_{u \in e} d_u(G_m) \leq 2 \sum_{u \in e} d_u(G_m) \mathbf{1}_{\{d_u(G_m) \geq C K_{\wedge}\}}.$$

Poniendo todo esto en una sola expresión, resulta que

$$\begin{aligned}
\sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{d_{e_i}(G_{i-1}) > K_{\Delta}/6\}} &= t O \left(\sum_{e \in E(G_m)} \sum_{u \in e} d_u(G_m) \mathbf{1}_{\{d_u(G_m) = \Omega(K_{\Delta})\}} \right) \\
&= t O \left(\sum_u \sum_{e \ni u} d_u(G_m) \mathbf{1}_{\{d_u(G_m) = \Omega(K_{\Delta})\}} \right) \\
&= t O \left(\sum_u d_u(G_m) \mathbf{1}_{\{d_u(G_m) = \Omega(K_{\Delta})\}} \sum_{e \ni u} 1 \right) \\
&= t O \left(\sum_v d_v(G_m)^2 \mathbf{1}_{\{d_v(G_m) = \Omega(K_{\Delta})\}} \right),
\end{aligned}$$

como deseábamos mostrar.

Debido a la expresión (4.9), para estimar $N_{\text{good}}^*(m)$ es suficiente entender el comportamiento de la variable aleatoria “ $\sum_{v: d_v(G_m) > Ctn} d_v(G_m)^2$ ”.

Afortunadamente disponemos de las herramientas para estudiar el comportamiento probabilístico de este tipo de variables aleatorias. Hecho todo esto, restaría estimar la variable $N_{\text{bad}}^*(m)$.

Una primera tentativa sería tentar controlar la variable aleatoria $N_{\text{bad}}^*(m)$ por la siguiente variable de conteo⁴:

$$\sum_{f \in E(K_n)} d_f(G_m) \mathbf{1}_{\{f \text{ is } m\text{-bad and } d_f(G_m) = \Omega(K_{\Delta})\}}$$

Usando nuestros resultados más detallados sobre cogrados (en G_m) junto con el argumento diádico (ver 2.4.3), podremos concluir que,

$$\begin{aligned}
\sum_{f \in E(K_n)} d_f(G_m) \mathbf{1}_{\{f \text{ is } m\text{-bad and } d_f(G_m) = \Omega(K_{\Delta})\}} &\lesssim^{*b} \frac{b^2}{\ell^2 K_{\Delta}} \\
&\lesssim \frac{b^3}{\ell^2 M}
\end{aligned}$$

⁴De hecho, esta variable cuenta triángulos con aristas ‘raras’ en G_m .

Como nuestro objetivo final es que $N_{\text{bad}}^*(m)$ sea acotado superiormente por $O(M)$, nos bastaría que

$$M \geq \frac{b^{3/2}}{\ell}.$$

Desafortunadamente esta desigualdad falla por un factor de $\ell^{1/2}$. Así, necesitamos un argumento menos crudo que capture este factor. Si analizamos el argumento anterior, podemos rescatar dos cosas:

- (i) Se rodamos el mismo argumento, para cogrados más grandes, las cuentas funcionan correctamente. Esencialmente, podemos mostrar que

$$\sum_{f \in E(K_n)} d_f(G_m) \mathbf{1}_{\{f \text{ es } m\text{-bad y } d_f(G_m) \geq L\}} \lesssim^{*b} M,$$

donde $L := \max\{\frac{b^2}{\ell^2 M}, K_\Delta\}$. Así, podemos controlar toda la contribución de $N_{\text{bad}}^*(m)$ que viene de los cogrados ‘gigantes’ (i.e., por lo menos L).

- (ii) De (i), deducimos que resta controlar la contribución de $N_{\text{bad}}^*(m)$ que viene de los cogrados muy grandes (i.e., por lo menos K_Δ) que no son gigantes, esto es,

$$\sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{K_\Delta/6 \lesssim d_{e_i}(G_{i-1}) \lesssim L \text{ y } e_i \text{ es } (i-1)\text{-bad}\}}. \quad (4.10)$$

Como ya vimos, en este punto de la prueba, tenemos que refinar el análisis. Para este fin, consideraremos el costo de seleccionar aristas (en el proceso Erdős-Rényi) de entre los pares de vértices que tienen cogrado muy grande. Como es usual, este argumento vendrá de la mano del argumento diádico, que reducirá nuestro problema de cogrados a contar aristas.

Así, por el argumento diádico, podemos estimar (4.10) por

$$K_\Delta \sum_{j=1}^J 2^j Z_j \quad (4.11)$$

donde $J := \log_2(L/K_\Delta)$,

$$Z_j = Z_{2^j K_\Delta} := \sum_{i=1}^m \mathbf{1}_{\{e_i \in H_j(i-1)\}}$$

y

$$\begin{aligned} H_j(i-1) &= H_{2^j K_\Delta} \\ &:= \{f \in E(K_n) \setminus E(G_{i-1}) : 2^j K_\Delta \leq d_f(G_{i-1}) \leq 2^{j+1} K_\Delta, f \text{ es } (i-1)\text{-bad}\}. \end{aligned}$$

Como veremos posteriormente, una de nuestras estimaciones cruciales será mostrar que

$$Z_j \lesssim^{*b} \frac{b}{\ell}. \quad (4.12)$$

Con esto podemos estimar (4.11) por

$$O\left(L \frac{b}{\ell}\right) = O(M)$$

Por tanto, tendremos completada nuestra prueba, una vez justifiquemos (4.12). Como el lector atento podrá haber notado, esta desigualdad (e la forma de la variable Z_j) se asemeja a la de la variable aleatoria binomial (cuando 'b' es un 'grande' desvío). Usaremos esta heurística para dar una intuición de esta desigualdad. Comenzaremos estimando $\mathbb{E} \left[\mathbf{1}_{\{e_i \in H_j(i-1)\}} \right] = \mathbb{P}(e_i \in H_j(i-1))$. Como el conjunto aleatorio $H_j(i-1)$ es \mathbf{G}_{i-1} -mensurable y e_i es escogido uniformemente de $E(K_n) \setminus E(G_{i-1})$, tenemos la siguiente relación determinista,

$$\begin{aligned} \mathbb{P}(e_i \in H_j(i-1) \mid G_{i-1}) &= \frac{|H_j(i-1)|}{N-i+1} \\ &\leq \frac{|H_j(i-1)|}{N-m+1} \\ &\lesssim \frac{|H_j(i-1)|}{n^2}, \end{aligned}$$

ya que $m \ll n^2$. Como ya observamos, en nuestra sección de cogrados será mostrado que, excepto con probabilidad $e^{-\Omega(b)}$, $|H_j(i-1)| \lesssim \frac{b^2}{2^{2j} K_\Delta^2 \ell^2} \leq \frac{b^2}{K_\Delta^2 \ell^2}$.

De aquí, resulta que promediando tenemos

$$\begin{aligned}\mathbb{P}(e_i \in H_j(i-1)) &\lesssim \frac{b^2}{n^2 K_\Delta^2 \ell^2} + e^{-\Omega(b)} \\ &\approx \frac{b^2}{n^2 K_\Delta^2 \ell^2}.\end{aligned}$$

Luego, por la linealidad de la esperanza (ver Proposición 2.1),

$$\mathbb{E}[Z_j] \lesssim m \frac{b^2}{n^2 K_\Delta^2 \ell^2} \lesssim t \frac{b^2}{K_\Delta^2 \ell^2}.$$

Más aún, note que $t \frac{b^2}{K_\Delta^2 \ell^2} \leq t \frac{b}{\ell} \ll \frac{b}{\ell}$. Por lo tanto, de la distribución binomial, deducimos que,

$$\begin{aligned}\mathbb{P}\left(Z_j > \frac{b}{\ell}\right) &\leq \exp\left(-\Omega\left(-b/\ell \ln\left(\frac{b/\ell}{tb/\ell}\right)\right)\right) \\ &= \exp(-\Omega(b)),\end{aligned}$$

como deseábamos. Para formalizar esta prueba precisaremos de dos técnicas: dominación estocástica e producto *dot* (de dos procesos estocásticos). A seguir, comenzaremos con la formalización de lo expuesto en esta sección.

4.3. Ingredientes Principales

En esta sección prepararemos las herramientas que serán utilizadas para la prueba del Teorema 4.1. Esta sección tiene el mismo espíritu que la subsección 3.2.2. Sin embargo, estas estimativas son débiles para nosotros⁵. Así, el primer trabajo a realizar es mejorar estos resultados al punto de ser óptimas (o casi). Con este fin, será requerido un análisis más delicado de concentración, además de implementar nuevos argumentos para estudiar grandes desvíos de grados y cogrados. Dicho esto, dividiremos esta sección en grados y cogrados, siguiendo el modelo de la subsección 3.2.2.

⁵Fueron estimativas suficientes en el caso denso estudiado en [21], pero se pierde mucha precisión en el caso disperso que nos interesa.

4.3.1. Sobre Grados en $\mathbf{G}(n, m)$: Modelo Disperso

Para nuestro estudio sobre grados en $\mathbf{G}(n, m)$, resulta conveniente tener cierta intuición. Ya que la variable aleatoria $d_u(\mathbf{G}(n, m))$ tiene distribución hipergeométrica, su función de distribución es bien conocida. Así, tenemos que,

$$\Delta(G_{n,m}) \lesssim^{*b} \frac{b}{\ell_b}$$

siempre que $b \geq 4tn$, $t \gg n^{-1} \ln(n)$ y donde hemos denotado por $\ell_b := \ln(b/etn)$. Note que este limitante superior es útil para $b = \Omega(n\ell)$, pues posterior a este valor se torna mayor que n .

Con esto en mente pasamos a una justificación heurística del valor de “ $\sum_u D_u(G_m)^2$ ”⁶. Dado que el desvío estándar de cada variable aleatoria D_u es aproximadamente \sqrt{tn} , solo por puro azar, se espera que $\sum_u D_u(G_m)^2$ sea de orden tn^2 . De hecho, el evento en que el que una cierta porción positiva de vértices u cumple que $D_u(G_m) = \Omega(\sqrt{tn})$ implica que $\sum_u D_u(G_m)^2 = \Omega(tn^2)$. Como se espera que el primer evento tenga probabilidad $e^{-\Omega(n)}$, esto sugiere que la estimativa de $D_u(G_m) = \Omega(\sqrt{tn})$ entorno de tn^2 sea válido hasta $b \leq n$. Así como el mero acaso puede generar desvíos, es bien conocido que otra causa principal de desvíos son los “fenómenos de localización”. Imaginemos que tenemos “plantamos” una estrella de con b/ℓ_b aristas (con $b = \Omega(tn)$). Este evento, que como vimos tiene probabilidad entorno de e^{-b} , contribuye con un desvío de $(b/\ell_b)^2$ para nuestra cantidad $\sum_u D_u(G_m)^2$. Finalmente, para valores de $b \geq n\ell$, comienzan a aparecer vértices de grado $n - 1$. De hecho, heurísticamente, el costo en la probabilidad por tener k vértices de grado $n - 1$ es aproximadamente $e^{-kn\ell}$. Así, con $k = b/(n\ell)$, podemos generar un desvío de bn/ℓ en $\sum_u D_u(G_m)^2$. Motivados por esta intuición, definimos la variable aleatoria $\kappa(b, t)$ que cuantificará el desvío de $\sum_u D_u(G_m)^2$ con un costo de e^{-b} . Sea pues

$$\kappa(b, t) = \begin{cases} tn^2 & \text{se } 1 \leq b < t^{1/2}n\ell, \\ (b/\ell)^2 & \text{se } t^{1/2}n\ell \leq b < n\ell, \\ bn/\ell & \text{se } n\ell \leq b \leq tn^2 \ell, \end{cases} \quad (4.13)$$

donde la división en intervalos se corresponde a los momentos donde un valor de κ supera al otro mientras $b = b_n$ recorre el intervalo $[1, tn^2\ell]$.

Similarmente podemos inferir heurísticamente el comportamiento de la variable aleatoria $\sum_u d_u(G_m) \mathbf{1}_{\{d_u(G_m) = \Omega(tn)\}}$. Esta cantidad será cuantificada por el parámetro κ^+ , que

⁶De hecho, necesitamos estimar esta suma para todo G_i , donde $i \in [m]$. Afortunadamente, el argumento es análogo y no amerita un esfuerzo significativo.

definimos directamente como

$$\kappa^+(b, t) = \begin{cases} (b/\ell_b)^2 & \text{se } 32tn \leq b < n\ell, \\ bn/\ell & \text{se } n\ell \leq b \leq tn^2 \ell. \end{cases} \quad (4.14)$$

Ahora pasamos a formalizar esta heurística en la siguiente proposición. Recordemos que denotamos por $(G_i)_{i=0}^N$ el proceso estocástico de Erdős-Rényi, y que consideramos la densidad $t = t_n$ como siendo $t = m/N$.

Proposición 4.1 (Suma del cuadrado de los (desvíos de) grados : versión dispersa). *Sea $t \geq 2n^{-1} \ln n$ y sea $b = b_n$ una secuencia real tal que $b \geq 32tn$. Entonces se cumplen las siguientes desigualdades probabilísticas:*

- (i) $\max_{i \in [m]} \sum_u D_u^2(G_i) \lesssim^{*b} \kappa(b, t)$, y
- (ii) $\sum_u d_u^2(G_m) \mathbf{1}_{\{d_u(G_m) \geq 32tn\}} \lesssim^{*b} \kappa^+(b, t)$.

Para mostrar esta proposición, necesitaremos de dos resultados útiles. El primero controla el tamaño de los conjuntos con vértices atípicos en los grafos uniformemente distribuidos (también colectamos el resultado sobre grado máximo). Sea $V_d(G_m) := \{u : d_u(G_m) \geq d\}$ el conjunto aleatorio de los vértices de G_m con grado por lo menos d .

Lema 4.1 (Tamaño de conjuntos de vértices atípicos). *Sea $t \geq 2n^{-1} \ln n$ y sea $b \geq 4tn$. Entonces para cada $d \geq 32tn$ se cumple que*

$$|V_d(G_m)| \lesssim^{*b} \frac{b}{d \ln(d/tn)}.$$

Más aún, $\Delta(G_m) \lesssim^{*b} b/\ell_b$.

Observación 4.2. *El lector podrá notar que ambas desigualdades probabilísticas son deducibles usando concentración de la distribución hipergeométrica (del tipo Chernoff) junto con Union Bound. Este tipo de argumentos será considerado estándar en lo que resta de este trabajo. Sólo para referencia colocaremos la expresión que estamos usando: Sea X una variable aleatoria del tipo*

binomial o hipergeométrica, entonces

$$\mathbb{P}(X > \lambda) \leq \exp(-\Omega(\lambda \ln(\lambda/\mathbb{E}[X]))) ,$$

donde $\lambda = \Omega(\mathbb{E}[X])$. Recalamos que la presencia del "ln" en la tasa de decaimiento es crucial en nuestras estimativas.

El segundo lema que usaremos controla la contribución de los D_u en $\sum_u D_u^2(G_m)$ cuando no ultrapasan el valor esperado, salvo factor constante.

Lema 4.2 (Suma del cuadrado de los (desvío de) grados no grandes : versión dispersa).

Sea $t \geq 2n^{-1} \ln n$ y sea $b \geq n$. Entonces se cumplen que:

$$\max_{i \in [m]} \sum_u D_u^2(G_i) \mathbf{1}_{\{d_u(G_m) \leq 32tn\}} \lesssim^{*b} btn.$$

Observación 4.3. La prueba del Lema 4.2 es parecida a la de su versión densa (Lema 3.3). Nuestra contribución en esta prueba fue una mejora de la Desigualdad Tipo McDiarmid (Enunciado 2.1), donde, a diferencia de la versión original dada en [21], conseguimos capturar la influencia de la densidad $t = m/N$ involucrada en la tasa de decaimiento (de la probabilidad de desvío). Adicionalmente, fue en este punto que conseguimos sistematizar una estimativa óptima para sumatorios del tipo $\sum_{\substack{x \in \mathbb{Z}^d \\ \|x\|_2 \geq r}} \exp(-\beta \|x\|_2^2)$ (vea enunciado 2.3). Previo a esto, en [21], la estimativa era extensa y particular a cada caso envolviendo métodos puramente combinatorios.

Ahora pasamos a esbozar la prueba de la Proposición 4.1.

Para comenzar, mostraremos la siguiente desigualdad determinista:

$$\sum_u D_u(G_i)^2 \leq \sum_u D_u(G_i)^2 \mathbf{1}_{\{d_u(G_i) < 32tn\}} + \sum_u d_u(G_m)^2 \mathbf{1}_{\{d_u(G_m) \geq 32tn\}}, \quad (4.15)$$

para todo $i \in [m]$. En efecto, dividiendo la contribución de los grados (desvíos) de acuerdo

al tamaño de los grados, basta mostrar que

$$\sum_u D_u(G_i)^2 \mathbf{1}_{\{d_u(G_i) \geq 32tn\}} \leq \sum_u d_u(G_m)^2 \mathbf{1}_{\{d_u(G_m) \geq 32tn\}}.$$

Pero esta desigualdad es una consecuencia directa del hecho que

$$D_u(G_i)^2 \mathbf{1}_{\{d_u(G_i) \geq 32tn\}} \leq d_u(G_m)^2 \mathbf{1}_{\{d_u(G_m) \geq 32tn\}}.$$

Para ver esto, suponga que $d_u \geq 32tn$. De ahí que $0 < D_u(G_i) < d_u(G_i) \leq d_u(G_m)$. Por tanto, $D_u(G_i)^2 \leq d_u(G_m)^2$ y $d_u(G_m) \geq 32tn$. En otras palabras (usando indicadores), hemos justificado la última desigualdad y en consecuencia la ecuación (4.15).

En particular, si asumimos que el segundo ítem de la Proposición 4.1 es válido, de la desigualdad (4.15) junto con el Lema 4.2, obtenemos que

$$\max_{i \in [m]} \sum_u D_u^2(G_i) \lesssim^{*b} (b tn + tn^2) + \kappa^+(b, t).$$

Note que el término “ tn^2 ” acrescentado al lado derecho, proviene de considerar valores de b inferiores a n . Dividiendo en intervalos, no es difícil ver que $(b tn + tn^2) + \kappa^+(b, t) \lesssim \kappa(b, t)$, siempre que $b \geq 32tn$. Con lo que habríamos mostrado la veracidad del primer ítem de la Proposición 4.1.

Así, restaría mostrar el segundo ítem. Para esto usaremos nuestro acostumbrado argumento diádico (ver 2.4.3) junto con el Lema 4.1.

$$\begin{aligned} \sum_u d_u(G_m)^2 \mathbf{1}_{\{d_u(G_m) \geq 32tn\}} &= \sum_{j=5}^J \sum_u d_u(G_m)^2 \mathbf{1}_{\{2^j tn \leq d_u(G_m) < 2^{j+1} tn\}} \\ &\lesssim \sum_{j=5}^J d_j^2 |V_{d_j}(G_m)| \\ &\lesssim^{*b} \sum_{j=5}^J d_j \frac{b}{\ln(d_j/tn)} \\ &= b tn \sum_{j=5}^J 2^j / j, \end{aligned}$$

donde $d_j := 2^j tn$ y $J = J(b) := \min\{\log(b/tn \ell_b), \log(1/t)\}$. Note que el valor de J es el

mayor valor j tal que $2^j tn \leq \min\{b/\ell_b, n\}$, donde este último término es el grado máximo de un grafo G_m (con probabilidad $1 - e^{-b}$), salvo factor de 2. Como la función $g(j) = 2^j/j$ ($j \geq 5$) es creciente con factor de crecimiento mayor estricto a 1, del análisis matemático, tenemos que la suma $\sum_{j=5}^J 2^j/j = O(2^J/J)$. Así, junto a la estimativa probabilística anterior, obtenemos que

$$\begin{aligned} \sum_u d_u(G_m)^2 \mathbf{1}_{\{d_u(G_m) \geq 32tn\}} &\lesssim^{*b} btn \times 2^J/J \\ &= (b/J) \times 2^J tn. \end{aligned}$$

Por tanto, si $32tn \leq b < n\ell$, J es un valor entorno a ℓ_b con $2^J tn = b/\ell_b$, luego el lado derecho de nuestra última desigualdad probabilística es de orden $(b/\ell_b)^2$. Similarmente, si $b \geq n\ell$, entonces J es un valor entorno de ℓ , luego el lado derecho de nuestra desigualdad es de orden bn/ℓ . Con lo que concluimos nuestro esbozo.

4.3.2. Sobre Cogrados en $G(n, m)$: Modelo Disperso

Esta subsección es una de las más importantes de nuestro trabajo. En particular, es una de las secciones que contiene más ideas adicionales respecto a [21]. Nuestra intención no es llenar al lector de demasiados detalles técnicos, pero si pasar las ideas y un esbozo digerido de los resultados. Dicho esto, comenzaremos la subsección.

Como indicado al inicio de esta sección, resulta conveniente separar la contribución de los pares de vértices que son “good” de los que son “bad”. Hecho esto, nos enfocaremos en los pares que son “bad” (pues los “good” se reducen al estudio de grados). Para el estudio de los pares “bad”, usaremos el argumento diádico. Para usar este argumento con eficacia, primeramente dividiremos los cogrados respecto a la magnitud de su desvío, esto es : desvío moderado y gran desvío. Para el caso de cogrados, el desvío moderado recorre el orden de magnitudes desde $t\sqrt{n}$ hasta t^2n , mientras que el gran desvío recorre el orden de t^2n hasta n (teniendo como modelo el grafo aleatorio G_m). Así, aplicaremos el argumento diádico separadamente de acuerdo a esta división. Con esto en mente, las siguientes definiciones son naturales.

1. (Régimen Moderado) Para cada $k \in K_1 := \{10, \dots, \log(t\sqrt{n}) + 10\}$, definimos

$$\begin{aligned} &F_k(G) \\ &:= \{f : D_f(G) \in [2^k t\sqrt{n}, 2^{k+1} t\sqrt{n}], \max\{D_u(G) : u \in f\} \leq 2^{k-5} \sqrt{n}\} \\ & \text{y} \\ &f_k(G) := |F_k(G)|. \end{aligned}$$

2. (Régimen Grande) Para cada $k \in K_2 := \{10, \dots, \log(1/t)\}$, definimos

$$H_k(G) := \{f : d_f(G) \in [2^k t^2 n, 2^{k+1} t^2 n] \text{ y } \max\{d_u(G) : u \in f\} \leq 2^{k-5} t n\} \text{ y}$$

$$h_k(G) := |H_k(G)|.$$

Con estas definiciones estamos listos para enunciar nuestra proposición central en esta subsección.

Proposición 4.2. Sea $t = \Omega(n^{-1/2}(\ln n)^{1/2})$.

- Si $b \geq n$, entonces se cumple que

$$\max\{f_k(G_i) : i \in [m], k \in K_1\} \lesssim^{*b} \left(\frac{b}{2^{2k}}\right)^2.$$

- Si $b = \Omega(\ln n)$, entonces se cumple que⁷

$$\max\{h_k(G_i) : i \in [m], k \in K_2\} \lesssim^{*b} \left(\frac{b}{(2^k t^2 n)k}\right)^2.$$

Como consecuencia de esta proposición y lo discutido anteriormente, obtenemos el siguiente corolario que nos permite “predecir” el comportamiento de la suma de los cuadrados de los desvíos de cogrados. Adicionalmente, incluimos una estimativa de la suma de (muy) grandes cogrados que nos será de utilidad eventualmente, como indicado en la discusión inicial de esta sección.

Corolario 4.1. Sea $t = \Omega(n^{-1/2}(\ln n)^{1/2})$. Si $b \geq n$, entonces se cumple que

$$\max \left\{ \sum_{f \in E(K_n)} D_f(G_i)^2 : i \in [m] \right\} \lesssim^{*b} \max\{bt^2 n^2, b^2\}. \quad (4.16)$$

⁷Note que es posible enunciar el resultado en términos del cogrado (digamos R) en lugar de enunciarlo en términos de k . Para resaltar esto en la desigualdad, escribimos el cogrado entre “paréntesis”. Una formulación similar también es posible para el ítem previo.

Más aún, para $b \geq 4tn$ tenemos que

$$\sum_{f \in E(K_n)} d_f(G_m) \mathbf{1}_{\{f \text{ is } m\text{-bad y } d_f(G_m) = \Omega(L)\}} \lesssim^{*b} M.$$

donde $L = \max\{K_\Delta, b^2/\ell^2 M\}$.

Demostración del Corolario 4.1:

Para comenzar, mostraremos como deducir (4.16). Note que basta mostrar dicha desigualdad probabilística para cada $i \in [m]$ debido a que $b \gg \ln n$. Así, nuestro objetivo es mostrar que, fijado $i \in [m]$,

$$\sum_{f \in E(K_n)} D_f(G_i)^2 \lesssim^{*b} \max\{bt^2n^2, b^2\}.$$

Note que podemos asumir que $|D_f(G_i)| = \Omega(t\sqrt{n})$, pues la contribución de los f 's que no satisfacen esta propiedad es $O(t^2n^3) = O(bt^2n)^2$. Ahora bien, de manera crucial, vamos a clasificar los pares $f \in E(K_n)$ en relación a ser i -bad o i -good. Para el caso de los pares f que son i -good, tenemos que

$$\begin{aligned} \sum_{f \text{ } i\text{-good}} D_f(G_i)^2 &\lesssim t^2 \sum_f \sum_{u \in f} D_u(G_i)^2 \\ &\leq t^2 n \sum_u D_u(G_i)^2 \end{aligned}$$

Luego, aplicando la Proposición 4.1 (i), inferimos que

$$\begin{aligned} \sum_{f \text{ } i\text{-good}} D_f(G_i)^2 &\lesssim^{*b} t^2 n \kappa(b, t) \\ &\leq bt^2 n^2 / \ell \\ &\leq bt^2 n^2, \end{aligned}$$

donde hemos usado que $\kappa(b, t) \leq bn/\ell$, para $b \geq n$. Con esto, hemos reducido el trabajo a estimar la contribución de los f 's i -bad. La estimativa de esta contribución no es inmediata. En particular, sub-clasificaremos los i -bad de acuerdo a la magnitud de su desvío. Por un lado analizaremos la contribución de los desvíos moderados, y por otro la contribución

de los grandes desvíos. Es decir, separaremos la contribución como sigue.

$$\begin{aligned} \sum_{f \text{ } i\text{-bad}} D_f(G_i)^2 &\leq \sum_f D_f(G_i)^2 \mathbf{1}_{\{2^{10}t\sqrt{n} \leq |D_f(G_i)| < 2^{11}t^2n \text{ y } f \text{ } i\text{-bad}\}} \\ &\quad + \sum_f d_f(G_i)^2 \mathbf{1}_{\{2^{10}t^2n \leq d_f(G_i) < n \text{ y } f \text{ } i\text{-bad}\}}. \end{aligned}$$

Aplicando el argumento diádico (ver 2.4.3) sobre los cogrados y usando la definición de i -bad, no es difícil ver que

$$\sum_{f \text{ } i\text{-bad}} D_f(G_i)^2 \lesssim \sum_{k \in K_1} 4^k t^2 n f_k(G_i) + \sum_{k \in K_2} 4^k t^4 n^2 h_k(G_i)$$

Para finalizar, basta mostrar que cada sumando de derecha está bien controlado (probabilísticamente). Por simplicidad de la exposición, sólo mostraremos los detalles del segundo término. Por tanto, aplicando la Proposición 4.2, obtenemos que

$$\begin{aligned} \sum_{k \in K_2} 4^k t^4 n^2 h_k(G_i) &\lesssim^{*b} \sum_{k \in K_2} 4^k t^4 n^2 \frac{b^2}{4^k t^4 n^2 k^2} \\ &= b^2 \sum_{k \in K_2} \frac{1}{k^2} \\ &\leq b^2 \sum_{k \geq 1} \frac{1}{k^2} = O(b^2), \end{aligned}$$

donde hemos usado que $\sum_{k \geq 1} \frac{1}{k^2} = O(1)$. Lo que concluye la prueba de la desigualdad probabilística (4.16). La prueba de la segunda desigualdad es aún más inmediata, así que no la detallaremos aquí. ■

Para finalizar esta subsección, daremos un boceto de la prueba de la Proposición 4.2. Como el argumento (que llamamos *star-matching*) usado para ambos parámetros (f_k y h_k) son similares (vea también [6]), nosotros nos enfocaremos en el menos técnico, es decir en h_k (que envuelve grandes desvíos de cogrados). Cabe aclarar, que el parámetro f_k necesita de un argumento adicional para valores de k “pequeños”. La razón de esto es técnica y es debido a que el argumento *star-matching* nos da una probabilidad que no es lo suficientemente fuerte para ganar al Union Bound (sobre todos los matchings de un tamaño fijo). Antes de pasar a la siguiente sección, también daremos una breve

descripción de la modificación realizada al argumento star-matching y como esto nos permite controlar f_k para valores de k “pequeños”.

Demstración de la Proposición 4.2 relativa a h_k .

Para comenzar, note que podemos asumir que $k \in K_2$ satisface que $k2^k > b/t^2n^2$. Caso contrario, el resultado se deduce trivialmente al ser $h_k(G_i) \leq n^2$.

Recalamos al lector que $h_k(G_i) = |H_k(G_i)|$ es una variable aleatoria que cuenta ciertos pares “atípicas” de cogrado entorno de $2^k t^2 n$ (donde el cogrado no es explicado por el grado) y $H_k(G_i)$ el el grafo (auxiliar) que codifica estos pares. Para analizar h_k vamos a aplicar un argumento diádico (o mejor dicho, una partición diádica) en el grado (máximo) de cada par. Con esta información extra de los grados, el análisis de h_k se vuelve viable. Con este fin, definimos los siguientes subgrafos (auxiliares) $H_{k,j}$:

$$\begin{aligned} &H_{k,0}(G) \\ &:= \{uw : d_{uw}(G) \in [2^k t^2 n, 2^{k+1} t^2 n], \max\{d_u(G_i), d_w(G_i)\} \leq 2^{k/2} tn\} \\ &\text{y} \\ &H_{k,j}(G) \\ &:= \{uw : d_{uw}(G) \in [2^k t^2 n, 2^{k+1} t^2 n], \max\{d_u(G_i), d_w(G_i)\} \in [2^j tn, 2^{j+1} tn]\}. \end{aligned}$$

donde $k/2 < j \leq \min\{k - 6, \log_2(1/t)\}$. Convenimos en denotar los índices por J , esto es, $J := \{0\} \cup \{j : k/2 < j \leq \min\{k - 6, \log_2(1/t)\}\}$. Así, definimos

$$h_{k,j}(G) := |H_{k,j}(G)|$$

para todo $j \in J$. Claramente tenemos que

$$h_k(G) = \sum_{j \in J} h_{k,j}(G)$$

Así, nuestro objetivo se ve reducido a limitar (estocásticamente) cada variable $h_{k,j}(G_i)$. Esto será realizado siguiendo una estrategia motivada por el argumento estándar “star-matching”. La filosofía tras este argumento es que si un grafo (auxiliar) es demasiado denso, este contiene una estructura (subgrafo) relativamente simple (e “casi-independiente” en cierto sentido) con muchas aristas. Por tanto, es plausible esperar que el costo de tener esta estructura es muy alto, tornando el evento altamente improbable. Para formalizar esta idea, enunciamos el siguiente lema que controla (probabilísticamente) el tamaño de

ciertas estructuras dentro del grafo $H_{k,j}(G_i)$.

Lema 4.3 (Lema del tipo *star-matching* para $H_{k,j}$). Dado $t \geq n^{-1/2} \ln^{1/2}(n)$ y $b = \Omega(\ln n)$.

Para cada $p \in (0, t)$ se cumple que con probabilidad por lo menos $1 - \exp(-b)$,

(i) $H_{k,0}(\mathbf{G}(n, p))$ no contiene una estrella de tamaño $O\left(\frac{b}{k 2^k t^2 n}\right)$

(ii) $H_{k,0}(\mathbf{G}(n, p))$ no contiene un matching (o emparejamiento) de tamaño $O\left(\frac{b}{k 2^k t^2 n}\right)$

(iii) $H_{k,j}(\mathbf{G}(n, p))$ no contiene una estrella de tamaño $O\left(\frac{b}{(k-j) 2^k t^2 n}\right)$

para todo $k \leq 2 \log_2(1/t)$ (suficientemente grande) y $j \in J \setminus \{0\}$.

Observación 4.4. Como el lector minucioso podrá haber notado el lema anterior está enunciado en términos del modelo $\mathbf{G}(n, p)$ en lugar de $\mathbf{G}(n, i)$ (con $i \leq m$). Esto es un convenio técnico, que nos permite simplificar la prueba del lema, aprovechándonos de la independencia del modelo binomial. Como corolario, obtendremos (esencialmente) el mismo resultado para el modelo uniforme $\mathbf{G}(n, i)$ usando la transferencia entre ambos modelos. Siendo más exactos, basta generar $\mathbf{G}(n, i)$ como $\mathbf{G}(n, i/N)$ condicionado al evento que $e(\mathbf{G}(n, i/N)) = i$.

Antes de pasar a la prueba del Lema 4.3, mostraremos como utilizar este resultado para controlar las variables $h_{k,j}$ y en consecuencia h_k .

Sea $j \in J$. Si $j = 0$, entonces por el Lema 4.3 (itens (i) y (ii)), sabemos que, excepto con probabilidad menor a e^{-b} , $H_{k,0}(\mathbf{G}(n, p))$ no contiene ni una estrella ni un matching de tamaño $O\left(\frac{b}{k 2^k t^2 n}\right)$. Así, dentro de este evento suponga que tenemos que $h_{k,0}(\mathbf{G}(n, p)) = \Omega\left(\left(\frac{b}{k 2^k t^2 n}\right)^2\right)$. Esto significa que el grafo (auxiliar) $H_{k,0}(\mathbf{G}(n, p))$ tiene $\Omega\left(\left(\frac{b}{k 2^k t^2 n}\right)^2\right)$ aristas. Por el Lema 2.1 inferimos que $H_{k,0}(\mathbf{G}(n, p))$ contiene o una estrella o un matching de tamaño $\Omega\left(\frac{b}{k 2^k t^2 n}\right)$, lo que contradice nuestro evento inicial. Por tanto,

$$h_{k,0}(\mathbf{G}(n, p)) \lesssim^{*b} \left(\frac{b}{k 2^k t^2 n}\right)^2.$$

Ahora bien, si $j \in J \setminus \{0\}$, entonces controlaremos el tamaño de $H_{k,j}(\mathbf{G}(n, p))$ contando aristas. Para esto, usaremos el Lema 4.3 (item (iii)) junto con un resultado de la sección de

grados (4.1). En efecto, es fácil ver que por definición

$$h_{k,j}(\mathbf{G}(n,p)) = O(|V_{2^j t n}(\mathbf{G}(n,p))| \times \Delta(H_{k,j}(\mathbf{G}(n,p)))).$$

Luego, si $b \geq 4tn$ ⁸, entonces por el Lema 4.1 tenemos que

$$|V_{2^j t n}(\mathbf{G}(n,p))| \lesssim^{*b} \frac{b}{j 2^j t n},$$

y el Lema 4.3 (item (iii)) nos dice que

$$\Delta(H_{k,j}(\mathbf{G}(n,p))) \lesssim^{*b} \frac{b}{(k-j) 2^k t^2 n}.$$

Juntando estos dos resultados, y por Union Bound (ver Proposición 2.1), deducimos que

$$h_{k,j}(\mathbf{G}(n,p)) \lesssim^{*b} \frac{b^2}{j(k-j) 2^{k+j} t^3 n^2}.$$

Nuevamente por Union Bound y sumando todas las contribuciones de $h_{k,j}$, obtenemos que

$$h_k(\mathbf{G}(n,p)) = \sum_{j \in J} h_{k,j}(\mathbf{G}(n,p)) \lesssim^{*b} \left(\frac{b}{k 2^k t^2 n} \right)^2 + \sum_{j \in J \setminus \{0\}} \frac{b^2}{j(k-j) 2^{k+j} t^3 n^2}$$

Para finalizar, basta mostrar que el segundo término de la suma es limitado asintóticamente por $\left(\frac{b}{k 2^k t^2 n} \right)^2$. En efecto, por análisis matemático⁹, podemos ver que

$$\sum_{j \in J \setminus \{0\}} \frac{1}{j(k-j) 2^j} = O\left(\frac{1}{k^2 2^{k/2}} \right)$$

⁸En el caso que $b < 4tn$ la estimativa es trivial, pues $H_{k,j}$ es el grafo vacío con probabilidad $1 - e^{-\Omega(b)}$, debido a la ausencia de vértices con alto grado.

⁹Note que si los términos de una serie decaen exponencialmente, entonces el primer término domina la serie (salvo constante).

Sustituyendo esta estimativa en el segundo término deseado, tenemos que

$$\begin{aligned} \sum_{j \in J \setminus \{0\}} \frac{b^2}{j(k-j)2^{k+j}t^3n^2} &\lesssim \frac{b^2}{k^2 2^{3k/2}t^3n^2} \\ &\leq \left(\frac{b}{k 2^{k/2}t^2n} \right)^2 \times (2^k t^2)^{1/2} \\ &\leq \left(\frac{b}{k 2^{k/2}t^2n} \right)^2, \end{aligned}$$

donde hemos usado que $2^k t^2 n < n$. Usando nuevamente Union Bound sobre k , obtenemos las estimativas deseadas para h_k cuando k varía en K_2 .

Lo que resta por mostrar en esta subsección es el Lema 4.3. Siguiendo la filosofía de este capítulo, sólo daremos un boceto de uno de los ítems (en este caso el ítem (i)) de este lema, con el único objetivo de dar una idea global de las pruebas. Supongamos pues que existe una estrella en $H_{k,0}(\mathbf{G}(n,p))$ de tamaño

$$s = \Omega \left(\frac{b}{k 2^{k/2} t^2 n} \right).$$

Ya que el número de estrellas S de tamaño s dentro de K_n es limitado por n^{s+1} , usando Union Bound, tenemos que

$$\begin{aligned} &\mathbb{P}(H_{k,0}(\mathbf{G}(n,p)) \text{ contiene una estrella de tamaño } s) \\ &\leq n^{s+1} \max_S \{\mathbb{P}(S \subseteq H_{k,0}(\mathbf{G}(n,p)))\} \end{aligned}$$

donde S varía entre las estrellas en K_n de tamaño s . Así, hemos reducido el problema a mostrar que, fijado una estrella S (en K_n) de tamaño s , se tiene que

$$\mathbb{P}(S \subseteq H_{k,0}(\mathbf{G}(n,p))) \leq e^{-\Omega(b)} \tag{4.17}$$

Sea pues S una tal estrella con centro en u y $W = \{w_1, \dots, w_s\}$ su conjunto de hojas. La estrategia que seguiremos será “revelar” la vecindad de u dentro de $\mathbf{G}(n,p)$. De este modo habremos reducido el problema de estudiar desvíos de cogrados a desvío de aristas entre dos conjuntos grandes, como veremos a seguir. Observe que el evento “ $S \subseteq H_{k,0}(\mathbf{G}(n,p))$ ” implica que el número total de aristas entre el conjunto aleatorio

$N_u(\mathbf{G}(n, p))$ y el conjunto W es $\Omega(2^k t^2 ns)$. Así, por la ley total de las probabilidades (o simplemente por independencia del modelo), resulta que

$$\begin{aligned}
& \mathbb{P}(S \subseteq H_{k,0}(\mathbf{G}(n, p))) \\
&= \sum_{\Gamma_u} \mathbb{P}(N_u(\mathbf{G}(n, p)) = \Gamma_u) \mathbb{P}(S \subseteq H_{k,0}(\mathbf{G}(n, p)) \mid N_u(\mathbf{G}(n, p)) = \Gamma_u) \\
&\leq \sum_{\Gamma_u} \mathbb{P}(N_u(\mathbf{G}(n, p)) = \Gamma_u) \mathbb{P}\left(e_{\mathbf{G}(n,p)}(W, \Gamma_u) \gtrsim 2^k t^2 ns \mid N_u(\mathbf{G}(n, p)) = \Gamma_u\right) \\
&\leq \max_{\Gamma_u} \mathbb{P}\left(e_p(W, \Gamma_u) \gtrsim 2^k t^2 ns\right)
\end{aligned}$$

donde Γ_u varia entre las posibles vecindades de u en K_n de tamaño menor o igual a $2^{k/2}tn$ (por definición de $H_{k,0}(\mathbf{G}(n, p))$). También hemos usado la notación $e_p(W, U)$ para denotar el número de aristas entre U y W escogidas aleatoriamente e independientemente con probabilidad p . Note además que los eventos " $N_u(\mathbf{G}(n, p)) = \Gamma_u$ " y " $e_{\mathbf{G}(n,p)}(W, \Gamma_u) = \Omega(2^k t^2 ns)$ " son independientes (en $\mathbf{G}(n, p)$), luego la última de las tres ecuaciones es correcta. Finalmente, como el valor esperado de $e_p(W, \Gamma_u)$ es menor o igual que $p|\Gamma_u|s \leq 2^{k/2}t^2 ns$, podemos aplicar el Lema ... para inferir que

$$\begin{aligned}
\mathbb{P}\left(e_p(W, \Gamma_u) = \Omega(2^k t^2 ns)\right) &\leq \exp\left(-\Omega(2^k t^2 ns \times k)\right) \\
&= \exp(-\Omega(b)).
\end{aligned}$$

Combinando esto con la desigualdad anterior, hemos mostrado (4.17) como deseábamos.

4.4. Desvío positivo de triángulos en $\mathbf{G}(n, m)$: límite superior

En esta sección mostraremos de una forma más sistemática el resultado principal de este capítulo (para ser más exactos, mostraremos el límite superior del Teorema 4.1). En otras palabras, nuestro objetivo es mostrar que

$$D_{\Delta}(G_m) \lesssim^{*b} M(b, t) \tag{4.18}$$

siempre que $t = \Omega(n^{-1/2}(\ln n)^{1/2})$ y $3 \ln n \leq b \leq tn^2 \ell$.

Como ya fue descrito en la sección 4.2, usaremos la descomposición martingala de

$D_\Delta(G_m)$ junto con la estrategia “Divide-and-Conquer” para obtener la descomposición

$$D_\Delta(G_m) = D'_\Delta(G_m) + N^*_\Delta(G_m),$$

donde $D'_\Delta(G_m)$ es la parte truncada de la representación martingala de $D_\Delta(G_m)$ y $N^*_\Delta(G_m)$ es la parte no truncada respectiva. Así, como ya fue comentado en la sección 4.2, nuestra tarea se reduce a mostrar que

$$\begin{aligned} D'_\Delta(G_m) &\lesssim^{*b} M \text{ y} \\ N^*_\Delta(G_m) &\lesssim^{*b} M. \end{aligned}$$

En las siguientes subsecciones mostraremos cada una de estas desigualdades probabilísticas respectivamente, lo que concluirá la prueba de nuestro resultado principal.

4.4.1. Controlando $D'_\Delta(G_m)$

Para controlar la parte truncada $D'_\Delta(G_m)$ usaremos la desigualdad de Freedman. Para esto, debemos justificar algunas hipótesis junto con algunas estimativas de segundo momento (condicionado).

Primeramente el lector debe notar que $D'_\Delta(G_m)$ está expresado como un supermartingala. Esto es inmediato del hecho que $D_\Delta(G_m)$ es un martingala y que el truncamiento se dá en valores positivos (grandes). Así, la esperanza de los incrementos del truncamiento no serán superiores a la esperanza de los incrementos de la martingala original (igual a 0). Por comodidad, denotaremos el incremento i -ésimo de $D'_\Delta(G_m)$ por

$$\mathbb{X}'_i := 3 \frac{(N-m)_2 (m-i)}{(N-i)_3} X'_\wedge(\sigma_i) + \frac{(N-m)_3}{(N-i)_3} X'_\Delta(\sigma_i)$$

Así, necesitamos cierto control sobre los incrementos \mathbb{X}'_i para aplicar Freedman’s Inequality. Note que del truncamiento aplicado, obtenemos directamente que

$$|\mathbb{X}'_i| \lesssim \frac{M(b, t)}{b}$$

para todo $i \in [m]$.

Por otro lado, podemos limitar el cuadrado de \mathbb{X}'_i trivialmente por

$$\begin{aligned} |\mathbb{X}'_i|^2 &\lesssim t^2 |X'_\wedge(\sigma_i)|^2 + |X'_\Delta(\sigma_i)|^2 \\ &\leq t^2 |X_\wedge(\sigma_i)|^2 + |X_\Delta(\sigma_i)|^2 \end{aligned} \quad (4.19)$$

para todo $i \in [m]$. Así, podemos limitar el segundo momento de \mathbb{X}'_i vía el siguiente lema.

Lema 4.4 (Varianza Condicionada de $X_\wedge(\sigma_i)$ y $X_\Delta(\sigma_i)$: orden de magnitud - caso disperso). *Sea $\Omega(n^{-1/2}(\ln n)^{1/2}) = t \leq 1/2$. Para todo $b = \Omega(tn)$ y para todo $i \in [m]$, se cumplen las siguientes desigualdades no-deterministas,*

$$\mathbb{E} [X_\wedge(\sigma_i)^2 | G_{i-1}] \lesssim^{*b} \frac{\kappa(b, t)}{n}.$$

Análogamente, para $X_\Delta(\sigma_i)$, tenemos que para $n \leq b \leq t^2 n^2$, se satisface la siguiente desigualdad,

$$\mathbb{E} [X_\Delta(\sigma_i)^2 | G_{i-1}] \lesssim^{*b} bt^2,$$

y para $b > t^2 n^2$, se cumple que

$$\mathbb{E} [X_\Delta(\sigma_i)^2 | G_{i-1}] \lesssim^{*b} \frac{b^2}{n^2}.$$

Demostración del Lema 4.4:

Usando las estimativas del capítulo anterior y los nuevos resultados sobre grados y cogrados, la prueba del lema resulta directa. Como ilustración mostraremos la primera desigualdad no-determinista.

Note que la ecuación (3.30) se aplica a nuestra situación pues $m = tN \leq N/2$. Así, tenemos que

$$\mathbb{E} [X_\wedge(\sigma_i)^2 | G_{i-1}] = O\left(\frac{\sum_{u \in [n]} D_u(G_{i-1})^2}{n}\right)$$

Ahora bien, por la Proposición 4.1 junto con la desigualdad anterior, obtenemos el control deseado,

$$\begin{aligned} \mathbb{E} [X_{\wedge}(\sigma_i)^2 | G_{i-1}] &= O \left(\frac{\sum_{u \in [n]} D_u(G_{i-1})^2}{n} \right) \\ &\lesssim^{*b} \frac{\kappa(b, t)}{n}. \end{aligned}$$

■

Combinando el Lema 4.4 con la ecuación (4.19), obtenemos el siguiente corolario.

Corolario 4.2. *Sea $\Omega(n^{-1/2}(\ln n)^{1/2}) = t \leq 1/2$. Para todo $n \leq b \leq t^2 n^2$ y para todo $i \in [m]$, se satisface la siguiente desigualdad,*

$$\mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] \lesssim^{*b} bt^2,$$

y para $b > t^2 n^2$, se cumple que

$$\mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] \lesssim^{*b} \frac{b^2}{n^2}.$$

Más aún, para todo $b \geq n$ tenemos que

$$\sum_{i=1}^m \mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] \lesssim^{*b} bt^3 n^2 + b^2 t.$$

Demostración del Corolario 4.2:

Hacemos notar al lector que la tercera desigualdad (probabilística) es consecuencia de las dos primeras (junto con Union Bound). En efecto, asumiendo cierta estas dos primeras desigualdades, es simple deducir que

$$\mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] \lesssim^{*b} bt^2 + \frac{b^2}{n^2}$$

para todo $b \geq n$ y para todo $i \in [m]$.

Así pues, sumando sobre todos los valores posibles de i (y considerando que $b \gg \ln n$), obtenemos la tercera desigualdad deseada, esto es,

$$\begin{aligned} \sum_{i=1}^m \mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] &\lesssim^{*b} m \left(bt^2 + \frac{b^2}{n^2} \right) \\ &\leq tn^2 \left(bt^2 + \frac{b^2}{n^2} \right) \\ &= bt^3 n^2 + b^2 t. \end{aligned}$$

Así pues, nuestra tarea se reduce a probar las dos primeras desigualdades. Como las cuentas son similares, solamente mostraremos como obtener la primera.

De la ecuación (4.19), aplicando esperanza condicionada, resulta que (determinísticamente)

$$\mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] \lesssim t^2 \mathbb{E} [X_{\wedge}(\sigma_i)^2 | G_{i-1}] + \mathbb{E} [X_{\Delta}(\sigma_i)^2 | G_{i-1}].$$

Por otro lado, del Lema 4.4 aplicado a

$$n \leq b \leq t^2 n^2 \text{ y } , i \in [m],$$

sabemos que

$$\mathbb{E} [X_{\wedge}(\sigma_i)^2 | G_{i-1}] \lesssim^{*b} \frac{\kappa(b, t)}{n}.$$

y

$$\mathbb{E} [X_{\Delta}(\sigma_i)^2 | G_{i-1}] \lesssim^{*b} bt^2.$$

Luego, juntando estas tres desigualdades, deducimos que

$$\mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] \lesssim^{*b} \frac{t^2 \kappa(b, t)}{n} + bt^2.$$

Así, todo se reduce a mostrar que

$$\kappa(b, t) \leq bn$$

para todo $b \geq n$, lo que es un simple ejercicio dada la definición de $\kappa(b, t)$ (y usando que $t \ll 1$, por simplicidad). ■

Con esto, ya tenemos los ingredientes necesarios para controlar $D'_\Delta(G_m)$. En efecto, vamos a aplicar la desigualdad de Freedman para supermartingalas 2.2 con valor final $D'_\Delta(G_m)$. Primeramente, suponga que $b \geq n$. En este caso, haciendo $R \asymp \frac{M(b,t)}{b}$ y $\beta \asymp \frac{M^2(b,t)}{b}$, tenemos que

$$\begin{aligned}
& \mathbb{P} \left(D'_\Delta(G_m) = \Omega(M(b,t)) \right) \\
& \leq \mathbb{P} \left(D'_\Delta(G_m) = \Omega(M(b,t)) \text{ y } \sum_{i=1}^m \mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] \leq \beta \right) \\
& \quad + \mathbb{P} \left(\sum_{i=1}^m \mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] > \beta \right) \\
& \leq \exp \left(-\Omega \left(\frac{M^2(b,t)}{M^2(b,t)/b + M(b,t) \times M(b,t)/b} \right) \right) + \exp(-\Omega(b)) \\
& \leq \exp(-b)
\end{aligned}$$

donde hemos utilizado el Corolario 4.2 junto con la estimativa $\beta \geq bt^3n^2 + b^2t$, lo que es fácilmente comprobable.

Para finalizar, indicaremos como proceder para el caso $b < n$. Como podemos ver, nuestros resultados sobre varianzas (condicionadas) no abordan explícitamente el régimen de b entre $\ln n$ y n (o entorno de n). Esto se debe a que el desvío estándar produce un incremento natural en la varianza total (de orden t^3n^3). Esto sugiere que no podemos esperar un mejor resultado que en $b = n$. Así, nuestra estrategia en estos regímenes menores será usar el caso $b = n$ junto con ciertas estimativas y concluir. Afortunadamente este abordaje funciona. En efecto, sea $\beta^* \asymp t^3n^3$. Aplicando el Teorema 2.2 como en el caso anterior para β^* en lugar de β , resulta que

$$\begin{aligned}
& \mathbb{P} \left(D'_\Delta(G_m) = \Omega(M(b,t)) \right) \\
& \leq \mathbb{P} \left(D'_\Delta(G_m) = \Omega(M(b,t)) \text{ y } \sum_{i=1}^m \mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] \leq \beta^* \right) \\
& \quad + \mathbb{P} \left(\sum_{i=1}^m \mathbb{E} [(\mathbb{X}'_i)^2 | G_{i-1}] > \beta^* \right) \\
& \leq \exp \left(-\Omega \left(\frac{M^2(b,t)}{t^3n^3 + M(b,t) \times M(b,t)/b} \right) \right) + \exp(-\Omega(n)) \\
& \leq \exp(-b)
\end{aligned}$$

donde hemos utilizado el Corolario 4.2 para $b = n^{10}$, además de la estimativa $t^3 n^3 \leq \frac{M^2(b,t)}{b}$ (para $b < n$). Esta última estimativa es verificable usando el hecho que $M(b,t) \geq N(b,t)$. Con lo que concluimos esta primera parte de la prueba. ■

4.4.2. Controlando $N_{\Delta}^*(G_m)$

Recordamos al lector que esta sección contiene las nuevas ideas que introducimos al inicio de este capítulo. Así, recomendamos una lectura a esa sección inicial caso haya sido ignorada. Aquí será mostrado con mayor formalismo lo anterior expuesto, sin repetir demasiado notaciones y/o terminologías.

Brevemente, denotamos la parte no truncada de los incrementos por $Z_{\wedge}(\sigma_i)$ y $Z_{\Delta}(\sigma_i)$. De esta forma tenemos la desigualdad determinista

$$N_{\Delta}^*(G_m) \leq 3t \sum_{i=1}^m Z_{\wedge}(\sigma_i) + \sum_{i=1}^m Z_{\Delta}(\sigma_i).$$

Por lo tanto, el problema se reduce a controlar (estocásticamente) cada una de las contribuciones de la suma anterior. Esto es,

$$t \sum_{i=1}^m Z_{\wedge}(\sigma_i) \lesssim^{*b} M(b,t) \text{ y} \tag{4.20}$$

$$\sum_{i=1}^m Z_{\Delta}(\sigma_i) \lesssim^{*b} M(b,t). \tag{4.21}$$

Nuevamente para facilitar la exposición, dividiremos esta sección en dos partes, donde cada parte se corresponde con la prueba de (4.20) y (4.21).

Controlando “ $t \sum_{i=1}^m Z_{\wedge}(\sigma_i)$ ”.

Esta parte de la prueba está vinculada con los nuevos resultados de grados que obtuvimos y es relativamente sencilla de mostrar usando estos resultados como ‘black box’. La parte crucial es que podemos limitar la variable aleatoria $\sum_{i=1}^m Z_{\wedge}(\sigma_i)$ por $\sum_u d_u^2(G_m)$ donde u varía entre los vértices con grado alto, así pues con alta predictibilidad sabemos controlar este último sumatorio por $\kappa^+(b,t)^{11}$.

¹⁰Note que $t^3 n^3 \geq n^2 t$ pues estamos asumiendo que $t \gg n^{-1/2} \ln^{1/2} n$.

¹¹O “cero” en el caso que b es menor a $\Theta(tn)$

Finalmente basta observar que $t\kappa^+(b, t)$ es inferior a $M(b, t)$. A seguir escribiremos este esbozo con un poco más de detalles.

Afirmamos que

$$\sum_{i=1}^m Z_{\wedge}(\sigma_i) \leq 4 \sum_u d_u^2(G_m) \mathbf{1}_{\{d_u(G_m) > K_{\wedge}/4\}} \quad (4.22)$$

Antes de mostrar esta desigualdad, veremos como deducir el control deseado sobre

$$t \sum_{i=1}^m Z_{\wedge}(\sigma_i).$$

Primero suponga que $b \geq 32tn$. Por la Proposición 4.1, sabemos que

$$\sum_u d_u^2(G_m) \mathbf{1}_{\{d_u(G_m) \geq 32tn\}} \lesssim^{*b} \kappa^+(b, t).$$

En particular, como $K_{\wedge}/4 > 32tn$, tenemos de (4.22)

$$\begin{aligned} t \sum_{i=1}^m Z_{\wedge}(\sigma_i) &\lesssim^{*b} t\kappa^+(b, t) \\ &\leq M(b, t) \end{aligned}$$

donde la última desigualdad se cumple para $b \geq 32tn$ y es un simple ejercicio.

Para el caso en que $b < 32tn$ no podemos usar el parámetro κ en lugar de κ^+ , pues no siempre es cierto que $t\kappa(b, t)$ es menor que $M(b, t)$. Afortunadamente si $b < 32tn$, tenemos que

$$\Delta(G_m) \lesssim^{*b} tn$$

Por tanto, de (4.22) nuevamente, tenemos que

$$t \sum_{i=1}^m Z_{\wedge}(\sigma_i) = 0 \leq M(b, t)$$

excepto con probabilidad de $e^{-\Omega(b)}$, como deseábamos.

Para finalizar esta subsección mostraremos la desigualdad (4.22). La idea de esta desigualdad es recordar que los incrementos X_\wedge 's pueden ser vistos como (suma de) grados.

Esto junto con la monotonía (para pasar todo al modelo G_m) nos permite limitar la suma total por la suma de cerezas de vértices con grado alto, esto es (esencialmente) como suma de los cuadrados grados altos. Así, recalamos que

$$X_\wedge(\sigma_i) \leq A_\wedge(\sigma_i) = 2 \sum_{v \in e_i} d_v(G_{i-1})$$

Luego, el evento $X_\wedge(\sigma_i) > K_\wedge$ implica que alguno de los vértices $v \in e_i$ satisface

$$d_v(G_{i-1}) > K_\wedge/4.$$

Por lo tanto, tenemos que

$$\begin{aligned} Z_\wedge(\sigma_i) &= X_\wedge(\sigma_i) \mathbf{1}_{\{X_\wedge(\sigma_i) > K_\wedge\}} \\ &\leq 4 \sum_{v \in e_i} d_v(G_{i-1}) \mathbf{1}_{\{d_v(G_{i-1}) > K_\wedge/4\}} \\ &\leq 4 \sum_{v \in e_i} d_v(G_m) \mathbf{1}_{\{d_v(G_m) > K_\wedge/4\}} \end{aligned}$$

donde la segunda desigualdad proviene de la monotonía del proceso Erdős-Rényi (y que $i \leq m$). Ahora bien, sumando sobre todos los $i \in [m]$, se tiene

$$\begin{aligned} \sum_{i=1}^m Z_\wedge(\sigma_i) &\leq 4 \sum_{i=1}^m \sum_{v \in e_i} d_v(G_m) \mathbf{1}_{\{d_v(G_m) > K_\wedge/4\}} \\ &= 4 \sum_{e \in E(G_m)} \sum_{v \in e} d_v(G_m) \mathbf{1}_{\{d_v(G_m) > K_\wedge/4\}} \\ &= 4 \sum_v \sum_{e: v \in e} d_v(G_m) \mathbf{1}_{\{d_v(G_m) > K_\wedge/4\}} \\ &= 4 \sum_v d_v^2(G_m) \mathbf{1}_{\{d_v(G_m) > K_\wedge/4\}} \end{aligned}$$

donde hemos usado un argumento de doble conteo en la penúltima igualdad. Hemos pues probado la afirmación deseada.

Controlando “ $\sum_{i=1}^m Z_{\Delta}(\sigma_i)$ ”.

Análogamente al caso anterior, aquí comenzamos estimando nuestro proceso aleatorio por variables combinatorias (cogrados) que son más fáciles de entender. Así, tenemos que

$$\sum_{i=1}^m Z_{\Delta}(\sigma_i) \leq \sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{d_{e_i}(G_{i-1}) > K_{\Delta}/6\}} \quad (4.23)$$

Del mismo modo que en nuestro análisis de cogrados, va ser muy útil dividir el lado derecho de (4.23) en dos sumatorios dependiendo de si el alto cogrado es explicado por el grado (o no). Recordemos que un par (de vértices) f es j -good ($j \in [m]$) si

$$d_f(G_j) \lesssim t \left(\sum_{v \in f} d_v(G_j) \right).$$

Así, decimos que “ e_i es good” si se cumple el evento “ $d_{e_i}(G_{i-1}) > K_{\Delta}/6$ y $\exists j \in [i-1, m]$ tal que e_i es j -good”. Complementariamente, si $d_{e_i}(G_{i-1}) > K_{\Delta}/6$ pero e_i no es good, diremos que “ e_i es bad”. Así, el lado derecho de (4.23) puede ser descompuesto por la suma

$$N_{\text{good}}^*(m) + N_{\text{bad}}^*(m)$$

donde

$$N_{\text{good}}^*(m) := \sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{e_i \text{ es good}\}}, \text{ y}$$

$$N_{\text{bad}}^*(m) := \sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{e_i \text{ es bad}\}}.$$

Como el lector atento podrá notar, sólo estamos reescribiendo (sistemáticamente) el panorama general dado al inicio de este capítulo. Ahora bien, la variable aleatoria $N_{\text{good}}^*(m)$ puede ser controlada por estimativas similares a las mostradas en la parte de cogrados

(usando los grados, por definición)¹². Así, nos ahorraremos este trabajo asumiendo que

$$N_{\text{good}}^*(m) \lesssim^{*b} M(b, t).$$

Nuestra tarea central será pues mostrar el siguiente lema, que finaliza la prueba.

Lema 4.5. *Sea $n^{-1/2} \ln^{1/2}(n) \lesssim t \ll 1$ y $\ln(n) \lesssim b \leq tn^2\ell$. Entonces, la siguiente desigualdad estocástica es válida,*

$$N_{\text{bad}}^*(m) \lesssim^{*b} M(b, t)$$

Para lo que resta, nos dedicaremos a la prueba de este lema. Nuevamente, el lector puede ver en nuestro panorama (dado al inicio del capítulo) las ideas envueltas en la prueba que sigue, pues aquí seremos más técnicos.

Demostración del Lemma 4.5:

Como fue mencionado anteriormente, nuestro primer paso para estimar la variable $N_{\text{bad}}^*(m)$ dividiremos la contribución de los cogrados (altos) involucrados en dos tipos : (altos-) gigantes y (altos-) no gigante. Nuestra punto de inflexión sera " $K_{\Delta}\ell$ ", así tenemos la siguiente descomposición:

$$\begin{aligned} N_{\text{bad}}^*(m) &= \sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{e_i \text{ es bad}, d_{e_i}(G_{i-1}) > K_{\Delta}\ell\}} \\ &\quad + \sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{e_i \text{ es bad}, d_{e_i}(G_{i-1}) \leq K_{\Delta}\ell\}}. \end{aligned}$$

De está forma, la tarea que nos resta será mostrar que cada sumando satisface la desigualdad (probabilística) deseada. Específicamente,

$$\sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{e_i \text{ es bad y } d_{e_i}(G_{i-1}) > K_{\Delta}\ell\}} \lesssim^{*b} M(b, t) \quad (4.24)$$

$$\sum_{i=1}^m d_{e_i}(G_{i-1}) \mathbf{1}_{\{e_i \text{ es bad y } d_{e_i}(G_{i-1}) \leq K_{\Delta}\ell\}} \lesssim^{*b} M(b, t) \quad (4.25)$$

¹²Un esbozo es dado al inicio del capítulo para beneficio del lector más curioso.

El análisis de (4.24) será la parte relativamente sencilla. Esto se debe a que podremos aplicar directamente nuestras herramientas de cogrados (vía argumento diádica) en el limitante superior

$$\sum_{f \in E(K_n)} d_f(G_m) \mathbf{1}_{\{f \text{ es } m\text{-bad y } d_f(G_m) > K_\Delta \ell\}}$$

Este limitante es consecuencia de la definición de “bad” de las aristas e_i , además que es una variable mas simple (y cruda) de analizar pues no ignora la información de seleccionar aristas (al sustituir e_i por f).

Por ende, nuestros esfuerzos estarán enfocados sólo en mostrar (4.25). Aquí, será crucial considerar el costo de escoger muchas aristas dentro de un número limitado posible. Para este fin, definimos los siguientes objetos aleatorios:

$$\begin{aligned} H^*(G_i) &:= \bigcup_{k=k_0}^{k_1} H_k(G_i), \\ \mathbf{a}^*(\sigma_m) &:= \sum_{i=1}^m \mathbf{1}_{\{e_i \in H^*(G_{i-1})\}} \text{ y,} \\ \mathbf{a}^{**}(\sigma_m) &:= \sum_{i=1}^m \mathbf{1}_{\{e_i \in H^*(G_{i-1}) \text{ y } |H^*(G_{i-1})| \lesssim b^4 / (\ell^2 M^2(b,t))\}}, \end{aligned}$$

donde k_0 es el último entero tal que $2^{k_0} t^2 n \geq K_\Delta$ y k_1 es el último entero que cumple que $2^{k_1-1} t^2 n \leq K_\Delta \ell$.

Crucialmente, observamos que el evento

$$“e_i \text{ es bad y } d_{e_i}(G_{i-1}) \leq K_\Delta \ell”$$

implica el evento

$$“e_i \in H^*(G_{i-1})”.$$

Note que esto es resultado inmediato de las definiciones. Así, usando esta implicación junto con una cota cruda de cogrados, tenemos que (4.25) es limitado superiormente por la variable aleatoria

$$K_\Delta \ell \mathbf{a}^*(\sigma_m) \asymp \frac{M(b,t)\ell}{b} \mathbf{a}^*(\sigma_m).$$

Luego, es suficiente mostrar que

$$\mathbf{a}^* \lesssim^{*b} \frac{b}{\ell}. \quad (4.26)$$

En lugar de estudiar la variable \mathbf{a}^* , estimaremos la variable \mathbf{a}^{**} que es estocásticamente el mismo objeto salvo un evento muy raro (que podemos ‘ignorar’). La razón de este cambio se debe, heurísticamente, a que dado G_{i-1} , la probabilidad del evento $e_i \in H^*(G_{i-1})$ es $O(|H^*(G_{i-1})|/N)$. Esta probabilidad es típicamente baja, sin embargo en el peor de los casos podría ser constante, lo que arruinaría nuestro análisis. Afortunadamente, tal posibilidad es extremadamente improbable y puede ser ‘ignorada’. Formalmente tenemos las siguientes desigualdades:

$$\mathbb{P}(\mathbf{a}^* \neq \mathbf{a}^{**}) = \exp(-\Omega(b)), \text{ y} \quad (4.27)$$

$$\mathbb{P}(\mathbf{a}^{**} \gtrsim b/\ell) = \exp(-\Omega(b)). \quad (4.28)$$

En efecto esto muestra (4.26), como podemos conferir en las siguientes desigualdades,

$$\begin{aligned} \mathbb{P}(\mathbf{a}^* \gtrsim b/\ell) &\leq \mathbb{P}(\mathbf{a}^* \gtrsim b/\ell \text{ y } \mathbf{a}^* = \mathbf{a}^{**}) + \mathbb{P}(\mathbf{a}^* \neq \mathbf{a}^{**}) \\ &\leq \mathbb{P}(\mathbf{a}^{**} \gtrsim b/\ell) + \mathbb{P}(\mathbf{a}^* \neq \mathbf{a}^{**}) \\ &= \exp(-\Omega(b)). \end{aligned}$$

De esta manera, nuestra tarea finalizará una vez mostremos las desigualdades (4.27) y (4.28). Mostrar (4.27) es relativamente sencillo. En efecto, note que si ocurre el evento “ $\mathbf{a}^* \neq \mathbf{a}^{**}$ ” entonces “para algún $i \in [m]$ tenemos que $H^*(G_{i-1}) \gtrsim b^4/(\ell^2 M^2(b, t))$ ”. Así,

$$\mathbb{P}(\mathbf{a}^* \neq \mathbf{a}^{**}) \leq \mathbb{P}(\exists i \in [m], H^*(G_{i-1}) \gtrsim b^4/(\ell^2 M^2(b, t))).$$

Mas, como veremos a continuación, el término de la derecha es limitado por $\exp(-\Omega(b))$. En virtud de la Proposición 4.2 sabemos que, excepto por una probabilidad $\exp(-\Omega(b))$, para todo $k \in [k_0, k_1]$ y todo $i \in [m]$, se cumple que

$$h_k(G_{i-1}) = |H_k(G_{i-1})| \lesssim \frac{b^2}{4^k t^4 n^2 k^2}.$$

Sumando sobre todos los $k \in [k_0, k_1]$, resulta que

$$\begin{aligned}
|H^*(G_{i-1})| &\leq \sum_{k=k_0}^{k_1} \frac{b^2}{4^k t^4 n^2 k^2} \\
&\lesssim \frac{b^2}{4^{k_0} t^4 n^2 k_0^2} \\
&\lesssim \frac{b^2}{K_\Delta^2 \ell^2} \\
&\lesssim \frac{b^4}{M^2(b, t) \ell^2},
\end{aligned}$$

donde usamos que los sumandos $g(k) = \frac{b^2}{4^k t^4 n^2 k^2}$ decaen geoméricamente, junto que $K_\Delta \gtrsim t^{3/2} n$. Esto termina la verificación de (4.27). Ahora pasaremos a mostrar (4.28). Este resultado ya no es tan directo. Por conveniencia, denotemos por

$$X_i := \mathbf{1}_{\{e_i \in H^*(G_{i-1}) \text{ y } |H^*(G_{i-1})| \lesssim b^4 / (\ell^2 M^2(b, t))\}},$$

para cada $i \in [m]$. Note que X_i es una variable aleatoria σ_i -medible, en particular el vector aleatorio (X_1, \dots, X_{i-1}) es σ_{i-1} -medible. Con esto, podemos ver la variable aleatoria $\alpha^{**} = \alpha^{**}(\sigma_m)$ como un proceso estocástico cuyos incrementos X_i son variables de Bernoulli con filtración dada por σ_i . Ya que estamos interesados en estimar su desvío, emplearemos el Lema 2.2 para este fin. Así, nuestra primera tarea será encontrar p_i tal que

$$\mathbb{P}(X_i = 1 \mid X_1, \dots, X_{i-1}) \leq p_i.$$

Por lo ya observado sobre el proceso α^{**} , será suficiente computar, para cada $i \in [m]$ la cantidad

$$\mathbb{P}(X_i = 1 \mid G_{i-1}) = \mathbb{P}(e_i \in H^*(G_{i-1}) \text{ y } |H^*(G_{i-1})| \lesssim b^4 / (\ell^2 M^2(b, t)) \mid G_{i-1}). \quad (4.29)$$

Así pues, si la realización G_{i-1} es tal que $|H^*(G_{i-1})| = \Omega(b^4 / (\ell^2 M^2(b, t)))$, entonces (4.29) será 0, por obvias razones. Por el contrario, si la realización G_{i-1} satisface $|H^*(G_{i-1})| \lesssim b^4 / (\ell^2 M^2(b, t))$, entonces la arista aleatoria e_i tiene distribución uniforme en $E(K_n) \setminus$

$E(G_{i-1})$. En particular, (4.29) estará limitado por

$$\begin{aligned} \frac{|H^*(G_{i-1})|}{N-i+1} &\lesssim \frac{b^4}{\ell^2 M^2(b, t) n^2} \\ &\lesssim \frac{b\ell}{n^2}, \end{aligned}$$

donde hemos usado que $M(b, t) \geq \text{Clique}(b, t)$. Denote por $q := \Theta(b\ell/n^2)$ y observe que q no depende del tiempo i . Aplicando el Lema 2.2 (con $p_i = q$) inferimos que $\mathbf{a}^{**} =$

$\sum_{i=1}^m X_i$ es dominada estocásticamente por $\text{Bin}(m, q)$.

En particular, tenemos que

$$\begin{aligned} \mathbb{P}(\mathbf{a}^{**} \gtrsim b/\ell) &\leq \mathbb{P}(\text{Bin}(m, q) \gtrsim b/\ell) \\ &= \exp\left(-\Omega\left(\frac{b}{\ell} \ln\left(e \frac{b/\ell}{qm}\right)\right)\right) \\ &= \exp\left(-\Omega\left(\frac{b}{\ell} \ln\left(\frac{1}{t\ell^2}\right)\right)\right) \\ &= \exp(-\Omega(b/\ell \times \ell)) = \exp(-\Omega(b)), \end{aligned}$$

donde hemos usado que $t \ll 1$. Lo que completa la prueba central de este trabajo. ■

Capítulo 5

Conclusiones

En este trabajo hemos querido mostrar los avances recientes en el t3pico de “desv3os en conteo de subgrafos” (para tri3ngulos y cerezas, con especial 3nfasis) de una forma amigable al lector.

Para esto, inicialmente presentamos de forma sucinta el manuscrito [21] donde se resuelve el problema general para el caso denso (p fijo) y donde se sientan las bases te3ricas.

Posteriormente, presentamos el subsecuente art3culo [1] donde estas t3cnicas son explotadas a la par que otras estrategias son incorporadas, con el objetivo de tratar el caso disperso ($p = o(1)$).

Tanto en el art3culo [21] como en el m3s reciente [1] son obtenidos resultados adicionales. Ya que estos resultados est3n fuera del objetivo de este trabajo, no los hemos incorporado aqu3. Sin embargo, el lector curioso puede dirigirse a las fuentes para entrar en detalles. S3lo aprovecharemos este par3grafo para describir sucintamente tales resultados del manuscrito [1]. Usando la teor3a de martingalas (junto con fen3menos de localizaci3n), tambi3n son obtenidos ‘l3mites inferiores’ para los desv3os de conteo de tri3ngulos y cerezas, donde los resultados son justos salvo constantes (en el exponente). Tambi3n recalcamos que los detalles para el modelo binomial $\mathbf{G}(n, p)$ se encuentran en dicho manuscrito. Estos dos resultados juntos cubren todo en rango de desv3o para ambos modelos siempre que la densidad sea superior a $n^{-1/2}\sqrt{\log n}$.

El segundo art3culo que estudiamos aqu3, surgi3 motivado por entender el problema de desv3os para modelos dispersos. Estos “desv3os” pueden ser entendidos y estudiados a varios niveles de dificultad.

En un trabajo paralelo (a3n no publicado) hemos abordado el problema de entender el desv3o asint3ticamente (en el exponente) para una gran familia de subgrafos (que abarca tri3ngulos y cerezas) con el costo de que el desv3o est3 limitado (lo que corresponde a

$b \approx tn$ en $\mathbf{G}(n, m)$ para triángulos, por ejemplo).

Otro problema de interés es comprender los desvíos negativos. En [21] son dadas cotas superiores que son óptimas para desvíos moderados. Posteriormente, para grandes desvíos, en el artículo [35] los autores determinan el desvío negativo para el conteo de triángulos ¹ usando métodos espectrales en su análisis que contrastan grandemente con los métodos martingales usados aquí. Esto sugiere que la causa de desvíos negativos grandes se debe a razones sutiles vinculadas con los autovalores de los grafos. En ese mismo artículo los autores proponen una conjetura sobre el limitante inferior de los desvíos (para triángulos en el modelo denso). En un trabajo subsecuente (junto con Simon Griffiths y Gabriel D. do Couto) conseguimos responder afirmativamente esta conjetura, completando la estimación de los desvíos negativos para triángulos en el modelo binomial denso (salvo constante en el exponente). Las ideas para esa prueba son totalmente disjuntos con lo discutido en este trabajo y se basan en argumentos estadísticos.

Claramente la historia no termina aquí. Nuestros resultados sólo incluyen triángulos (y cerezas) para el modelo disperso. Así, queda en abierto el problema de determinar los desvíos para grafos como cliques, o más general, para cualquier subgrafo fijo H . Por otro lado, mismo en el caso de triángulos (y cerezas), una cuestión importante en la teoría de los grandes desvíos es entender la “causa” de los desvíos. En este sentido, nos gustaría mostrar que en el caso de fenómenos de localización, la probabilidad de que el grafo aleatorio contenga una cierta estructura sutil, dado que este presenta un cierto desvío, tiende a 1. El problema de los desvíos negativos en modelos dispersos también se encuentra poco explorado.

Para cerrar este trabajo, sólo resta decir que si bien hemos presentado algunos resultados interesantes, sólo el problema de desvíos para conteo de subgrafos aún es un gran área a explorar, donde la matemática abre sus puertas a nuevas mentes y deseamos que esta pequeña puerta haya despertado el interés de alguna mente curiosa.

¹De hecho, posteriormente los autores extienden sus métodos para ciclos impares.

Bibliografía

- [1] J. ALVARADO, L. DE OLIVEIRA AND S. GRIFFITHS, *Moderate deviations of triangle counts in sparse Erdős-Rényi random graphs $G(n, m)$ and $G(n, p)$* , arXiv e-prints, <https://arxiv.org/pdf/2305.04326.pdf>, (2023).
- [2] A. BARABÁSI AND R. ALBERT, *Emergence of scaling in random networks*, *Science* 286 (5439) (1999), 509–512.
- [3] A. BARBOUR, M. KAROŃSKI AND A. RUCIŃSKI, *A central limit theorem for decomposable random variables with applications to random graphs*, *J. Combin. Theory Ser. B* 47 (2) (1989), 125–145.
- [4] B. BOLLOBÁS, *Modern Graph Theory*, Graduate Texts in Mathematics 184, Springer-Verlag New York, 1998.
- [5] B. BOLLOBÁS, *Random graphs*, Cambridge University Press, 2001.
- [6] S. CHATTERJEE, *The missing log in large deviations for triangle counts*, *Random Structures and Algorithms* 40 (4) (2011), 437–451.
- [7] S. CHATERJEE, *An Introduction to large deviations for random graphs*, *Amer. Math. Soc.* 53 (4) (2016), 617–642.
- [8] S. CHATERJEE AND A. DEMBO, *Nonlinear large deviations*, *Advances in Mathematics*, 299 (20) (2016), 396–450.
- [9] S. CHATERJEE AND S. VARADHAN, *The large deviation principle for the Erdős-Rényi random graph*, *European J. Combin.* 32 (7) (2011), 1000–1017.
- [10] N. COOK AND A. DEMBO, *Large deviations of subgraph counts for sparse Erdős-Rényi graphs* *Advances in Mathematics*, 373 (2020), 107–289.
- [11] B. DEMARCO AND J. KAHN, *Upper tails for triangles* *Random Structures Algorithms*, 40 (4) (2012), 452–459.

- [12] R. DIESTEL, *Graph Theory*, Graduate Texts in Mathematics, Springer, 2005.
- [13] H. DÖRING AND P. EICHELSBACHER, *Moderate deviations in a random graph and for the spectrum of Bernoulli random matrices*, Electron. J. Probab. 14 (92) (2009), 2636–2656.
- [14] H. DÖRING AND P. EICHELSBACHER, *Moderate deviations via cumulants*, J. Theoret. Probab. 26 (2) (2013), 360–385.
- [15] R. ELKAN, *Gaussian-width gradient complexity, reverse log-Sobolev inequalities and non-linear large deviations*, Geom. Funct. Anal. 28 (2018), 1548–1596.
- [16] P. ERDŐS AND A. RÉNYI, *On random graphs, I*, Publ. Math. Debrecen 6 (1959), 290–297.
- [17] P. ERDŐS AND A. RÉNYI, *The Evolution of Random Graphs*, Magyar Tud. Akad. Mat. Kutató Int. Közl 5 (1960), 17–61.
- [18] V. FÉRAY, P. MÉLIOT, AND A. NIKEGHBALI, *Mod- ϕ Convergence. Normality Zones and Precise Deviations*, SpringerBriefs in Probability and Mathematical Statistics, Springer Cham, 2016.
- [19] D. FREEDMAN, *On tail probabilities for martingales*, Ann. Probab. 3 (1975), 100–118.
- [20] E. N. GILBERT, *Random Graphs*, Ann. Math. Statist. 30 (4) (1959), 1141–1144.
- [21] C. GOLDSCHMIDT, S. GRIFFITHS AND A. SCOTT, *Moderate Deviations of Subgraph Counts in the Erdős-Rényi Random Graphs $G(n, m)$ and $G(n, p)$* , Transactions of the American Mathematical Society 373 (2020), 5517–5585.
- [22] M. HAREL, F. MOUSSET AND W. SAMOTIJ, *Upper tails via high moments and entropic stability*, arXiv e-prints (2019), <https://arxiv.org/pdf/1904.08212.pdf>.
- [23] S. JANSON, *A functional limit theorem for random graphs with applications to subgraph count statistics*, Random Structures Algorithms 1 (1) (1990), 15–37.
- [24] S. JANSON AND K. NOWICKI, *The asymptotic distributions of generalized U-statistics with applications to random graphs*, Probab. Theory Related Fields 90 (3) (1991), 341–375.
- [25] S. JANSON, *Orthogonal decompositions and functional limit theorems for random graph statistics*, Mem. Amer. Math. Soc. 111 (534), 1994, vi+78.
- [26] S. JANSON, K. OLESZKIEWICZ AND A. RUCIŃSKI, *Upper tails for subgraph counts in random graphs*, Isr. J. Math. 142 (2004), 61–92.
- [27] O. KALLENBERG, *Foundations of Modern Probability*, Probability and Its Applications, Springer, 2002.

- [28] J. KIM AND V. VU, *Divide and conquer martingales and the number of triangles in a random graph*, Random Structures Algorithms, 24 (2) (2004), 166–174.
- [29] K. KROKOWSKI, A. REICHENBACHS AND C. THÄLE, *Discrete Malliavin-Stein method: Berry-Esseen bounds for random graphs and percolation*, Ann. Probab. 45 (2) (2017), 1071–1109.
- [30] J. LESKOVEC, A. RAJARAMAN AND J. ULLMAN, *Mining of Massive Datasets*, Cambridge University Press, 2014.
- [31] A. LIEBENAU AND N. WORMALD, *Asymptotic enumeration of graphs by degree sequence, and the degree sequence of a random graph*, arXiv e-prints (2017), <https://arxiv.org/pdf/1702.08373.pdf>.
- [32] E. LUBETZKY E Y. ZHAO, *On the variational problem for upper tails in sparse random graphs*, Random Structures Algorithms 50 (3) (2016), 420–436.
- [33] C. MCDIARMID, *On the method of bounded differences*, In Surveys in Combinatorics, (1989), 148–188, Cambridge Univ. Press.
- [34] B. MCKAY AND N. WORMALD, *Asymptotic enumeration by degree sequence of graphs of high degree*, Eur. J. Combin 6 (1990), 565–580.
- [35] J. NEEMAN, C. RADIN, AND L. SADUN, *Moderate Deviations in Cycle Count*. arXiv e-prints (2022), <https://arxiv.org/pdf/2101.08249.pdf>.
- [36] G. REINERT AND A. RÖLLIN, *Random subgraph counts and U-statistics: multivariate normal approximation via exchangeable pairs and embedding*, J. Appl. Probab. 47 (2) (2010), 378–393.
- [37] A. RÖLLIN, *Kolmogorov Bound for the Normal Approximation of the number of triangles in the Erdős-Rényi Random Graph*, Probability in the Engineering and Informational Sciences (2021), 1–27.
- [38] A. RUCIŃSKI, *When are small subgraphs of a random graph normally distributed?*, Probability Theory and Related Fields 78 (1988), 1–10.
- [39] V. VU, *A Large Deviation Result on the Number of Small Subgraphs of a Random Graph*, Combinatorics, Probability and Computing 10 (1) (2001), 79–94.
- [40] D. WATTS AND S. STROGATZ, *Collective dynamics of ‘small-world’ networks*, Nature 393 (1998), 440–442.