1. Basics

1.1. **Discrete Probability Preliminaries.** A *finite probability space* (P, Ω) consists of a finite ground set Ω and a probability measure P on Ω. I.e. $\forall x \in \Omega$, $P(x)$ is nonnegative and $\Sigma_{x\in\Omega}P(x) = 1$.

The *uniform* distribution over Ω satisfies $\forall x \in \Omega, P(x) = \frac{1}{|\Omega|}$.

An *event* E is a subset of Ω. Its probability,$P(E)$, is $\Sigma_{x\in E}P(x)$.

Subadditivity of Probabilities $P(X_1 \cup X_2) \leqslant P(X_1) + P(X_2)$.

Two events A and B are *independent* if $P(A \cap B) = P(A)P(B)$.

If $P(B) > 0$ then the *conditional probability of A given B*, $P(A|B) = \frac{P(A\cap B)}{P(B)}$.

If A and B are independent and $P(B) > 0$ then $P(A|B) = P(A)$.

A *Random real-valued variable* X is a function $X : \Omega - > \mathcal{R}$. Its *expected value*, $E(X)$ is $\Sigma_{y\in\Omega}X(y)P(y)$.

Linearity of Expectation(LoE): $E(X_1 + X_2) = E(X_1) + E(X_2)$.

The *product* of probability spaces (P_1, Ω_1) and (P_2, Ω_2) has ground set $\{(x, y)|x \in \Omega_1, y \in \Omega_2\}$ and probability distribution P where $P((x, y)) = P_1(x)P_2(y)$.

For such a product space, the product E_1xE_2 of events $E_1 \subseteq \Omega_1$ and $E_2 \subset \Omega_2$ is $\{(x, y)|x \in E_1, y \in E_2)\}$. Clearly $P(E_1xE_2) = P_1(E_1)P_2(E_2)$.

A Bernoulli Trial is a probability space whose ground set consists of two elements: success or failure. $Bin(n, p)$ is the random variable which counts the number of succeses in the probability space which is the product of n Bernoulli trials in each of which the probability of success is p. By LoE, $E(Bin(n, p)) = pn$. This is often expressed as the number of successes in a sequence of n independent trials. This sequence is simply a choice of an order in which to expose the outcome in a product space. We will often refer to such sequences.

$P(Bin(n, p) = k) = \binom{n}{k}p^k(1 - p)^{n-k}$. Using Sterling's formula we see that this is $\theta(\sqrt{\frac{1}{n}})$ for p a constant and $k \in \{\lfloor pn \rfloor, \lceil pn \rceil\}$.

Also:

$$\frac{P(Bin(n, p) = k)}{P(Bin(n, p) = k + 1)} = \frac{\binom{n}{k}p^k(1 - p)^{n-k}}{\binom{n}{k+1}p^{k+1}(1 - p)^{n-k-1}} = \frac{(k + 1)(1 - p)}{p(n - k)}.$$

This is greater than 1 if $k > pn$ and less than 1 if $k + 1 < pn$. it follows that $Prob(Bin(n, p) = k)$ is maximized either at $k = \lceil pn \rceil$ or $k = \lfloor pn \rfloor$.

Exercise 1: Show that for any fixed p, $C > 0$, sufficiently large n, and $k = \lceil \frac{n}{2} + C\sqrt{n\ log\ n} \rceil$ we have:

$$Prob(Bin(n, p) = k) = \theta(\frac{\sqrt{logn}Prob(Bin(n, P) \geqslant k)}{\sqrt{n}}).$$

Exercise 2: Show that for any fixed p, for any n we can choose k with $k = \frac{n}{2} + \theta(\sqrt{n\ log\ n})$ such that:

$$Prob(Bin(n,p) \geqslant k) = (1 + o(1))n^{\frac{-3}{4}}.$$

1.2. **Graph Theory Preliminaries.** A *graph* $G = (V, E)$ consists of a set $V = V(G)$ of vertices and a set $E = E(G)$ of edges each of which is an unordered pair of distinct vertices, its *endpoints*. A *multigraph* $G = (V, E)$ consists of a set V of vertices and a multiset E of edges each of which is an unordered pair of vertices. A *digraph* $D = (V, A)$ consists of a set V of vertices and a set A of arcs each of which is an ordered pair of distinct vertices. The arc goes from its *tail* to its *head*. A *multidigraph* $D = (V, A)$ consists of a set V of vertices and a multiset A of arcs each of which is an ordered pair of vertices.

Two vertices are *adjacent* if they are joined by an edge. An edge e and a vertex v are *incident* if v is an endpoint of G. The *neighbourhood* of v, $N(v)$ is the set of vertices adjacent to v. The *degree* of v is the number of edges it is incident to.

The complement of a graph G, denoted $\overline{G}$, has vertex set $V(G)$ and its edge set conists of those pairs of distinct vertices of G which do not form edges in G.

In a digraph, the *outneighbourhood* of v, denoted $N^+(v)$ is the set of vertices which are heads of arcs of which v is a tail. The *outdegree* of G, denoted $d^+(v)$ is the number of arcs of which v is a tail. *Inneighbourhood* and *indegree* are defined symmetrically.

A *subgraph* H of G is a graph with $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. It is *induced* if $E(H) = \{(x, y)|(x, y) \in E(G) s.t. x, y \in V(H)\}$. We have similar definitions for subgraphs and induced subgraphs of multigraphs, digraphs and, multidigraphs.

The union of two graphs G and H has vertex set $V(H) \cup V(G)$ and edge set $E(H) \cup E(G)$. For any set X of vertices of a graph G, $G - X$ is the subgraph of G induced by $V - X$.

A *path of length* $k - 1$, or P_k, is a graph with $k - 1$ edges and k distinct vertices which can be enumerated so each edge joins vertices which are consecutive in the order. Its *endpoints* are the first and last vertex in this order, i.e. the vertices which have degree 1 in the path. For $k > 2$, *a cycle of length* k, or C_k, is obtained from a P_k by adding an edge between the first and last vertex.

A *walk of length* k from x to y is a multigraph with k edges for which we can enumerate the multiset of endpoints of these edges as $x = x_1, x_2, ..., y = x_{k+1}$ so that the edges of the walk are $\{(x_i, x_{i+1})|1 \leqslant i \leqslant k\}$. A closed walk of length k is a walk of length k from x to x for some vertex x.

A graph is *connected* if for every pair of vertices (x, y) it contains a path with endpoints x and y as a subgraph.

Observation: A path is a walk. If a graph contains a walk from x to y, it contains a path with endpoints x and y.

Corollary: A graph is connected precisely if for every pair (x, y) of its vertices, there is a walk from x to y.

Corollary: The union of two connected graphs G and H which intersect is connected.

Proof. For any vertex z in $V(H) \cap V(G)$ and pair of vertices (x, y) of $V(G) \cup V(H)$. Both x and z lie in one of the connected subgraphs H or G of $H \cup G$. Hence there is a path from x to z in $H \cup G$. Symmetrically there is a path from z to y in $H \cup G$. Concatenating these two walks yields a walk from x to y in $H \cup G$. □

The *components* of a graph are its maximal connected subgraphs.

Corollary: The components of a graph are disjoint.

A digraph is *strongly connected* if for every two vertices x and y there is a path from x to y. The *strong components* of a digraph are its maximal strongly connected subgraphs. Again these are disjoint.

A *clique of size* l or K_l is a graph with k vertices, every two of which are joined by an edge.

A *stable set of size* l or S_l in a graph G is a set of k vertices no two of which are joined by an edge.

A *matching* of G is a set of disjoint edges.

A k-*colouring* of G is a function f mapping V to $\{1, ..., k\}$ such that no two adjacent vertices are incident,

A *tree* is a connected graph with $|V| - 1$ edges, or equivalently a connected graph with no cycles. Every tree contains a *leaf*, that is a vertex of degree one. Deleting a leaf yields a new tree.

Exercise 3: Show that a tree with at most one vertex of degree 2 has at least $\frac{|V(T)|}{2}$ leaves.

Exercise 4 Show that for every edge e of a tree $T - e$ is disconnected.

A *spanning tree* for a graph G is a subgraph which is a tree and has vertex set $V(G)$.

We root a tree to obtain a *rooted tree* by choosing a *root* r, and constructing a digraph by orienting each edge e of T towards the component of $T - e$ not containing r. The parent $p(v)$ of a non-root node v in the tree, is the unique node u such that uv is an edge of this digraph. We say w is a child of v if $p(w) = v$.

2. Further Fundamental Background

2.1. **A First Random Model.** $G_{n,p}$ is a random graph with vertex set $V_n = \{1, ..., n\}$ where for each graph H on V_n, $Prob(G_{n,p} = H) = p^{|E(H)|}(1-p)^{\binom{n}{2}-|E(H)|}$. Equivalently $G_{n,p}$ is a product space which is the product of $(P_{i,j}, \Omega_{i,j})$ for $1 \leqslant i < j \leqslant n$ where $\Omega_{i,j} = \{ij \text{ is an edge}, ij \text{ is not an edge}\}$ and $P_{i,j}(ij \text{ is an edge}) = p$. Note that $G_{n,1/2}$ is a uniformily chosen graph from amongst the $2^{\binom{n}{2}}$ graphs on V_n. $|E(G_{n,p})|$ is $Bin(\binom{n}{2}, p)$ and for every vertex v of $G_{n,p}$, $d(v)$ is $Bin(n-1, p)$.

2.2. **A Second Random Model (The Configuration Model).** Given an enumeration of the vertices of a (multi)graph G as $\{v_1, ..., v_n\}$, the degree sequence of G is $\{d_1, ..., d_n\}$ where d_i is the degree of v_i.

For a sequence $\mathcal{D} = \{d_1, ..., d_n\}$ of nonnegative integers whose sum D is even, we construct a random multigraph $G_{\mathcal{D}}$ with degree sequence $\mathcal{D}$ as follows. We generate d_i copies of each vertex v_i and consider a uniformly random matching $M_{\mathcal{D}}$ on the resultant set of D vertices. We then merge the copies of each v_i into one vertex to get $G_{\mathcal{D}}$. Thus, for $i \neq j$, the number of edges between v_i and v_j in $G_{\mathcal{D}}$ is the number of edges of $M_{\mathcal{D}}$ joining copies of v_i to copies of v_j and the number of loops at v_i is the number of edges of $M_{\mathcal{D}}$ joining two copies of v_i.

As a simple example consider $\mathcal{D} = \{2, 2, 2\}$. Then there are 15 matchings between the six vertex copies. Exactly one matching yields the multigraph with three loops, so the probability that $G_{\mathcal{D}}$ is this graph is $\frac{1}{15}$. For each of the three multigraphs consisting of a loop at one vertex and two parallel edges between the other two vertices, there are two matching which yield the multigraph, so the probability that $G_{\mathcal{D}}$ is a specific such multigraph is $\frac{2}{15}$. Finally, there are eight matchings corresponding to the triangle on these three vertices, so the probability that $G_{\mathcal{D}}$ is a triangle is $\frac{8}{15}$.

Given a list of the edges of a multigraph G with degree sequence $\mathcal{D}$, we can specify all of the matchings which correspond to G by specifying a bijection between the copies of v_i and its appearances on the edge list. If G is simple all of these matchings are distinct, so the number of such matchings is $\Pi_{i-1}^{n} d_i!$. If G has loops or multiple edges, the matchings created will not be distinct. We see that all simple graphs with degree sequence $\mathcal{D}$ are equally likely to be $G_{\mathcal{D}}$.

$M_{\mathcal{D}}$ can be generated by choosing a random permutation of the vertex copies as $s_1, .., s_D$ and then using the edges $s_{2i-1}s_{2i}$ for $i \in \{1, ..., \frac{D}{2}\}$. We note that every matching corresponds to $2^{\frac{D}{2}}\frac{D}{2}!$ permutations, as we can list the edges in any order, and put either vertex of each edge first. We can also generate the matching, one edge at a time, as having exposed s one of its edges, for any unmatched vertex copy, the vertex copy it is matched to is equally likely to be any other vertex copy (verify this if you like).

Exercise 5: Show that the expected number of loops in $\mathcal{G}_D$ is $\theta(\frac{\Sigma_{i=1}^n d_i(d_i-1)}{2(D-1)})$ and the expected number of pairs of nonloop edges of $\mathcal{G}_D$ with the same endpoints is $\theta(\frac{\Sigma_{i=1}^n \Sigma_{j\neq i} d_i(d_i-1)d_j(d_j-1)}{4(D-1)(D-3)})$.

2.3. **Concentration Inequalities.** Linearity of Expectation often allows us to compute easily the expected value of a random variable. If we can show that it is concentrated around its expected value, i.e the probability it is far from its expected value is small, then we have valuable information about its distribution.

We begin with a one-sided inequality:

Markov's Inequality: If X is a nonnegative variable and a is a nonnegative real then $P(X \geqslant aE(X)) \leqslant \frac{1}{a}$.

Proof.

$$E(X) = \Sigma_{y\in\Omega} X(y)P(y) \geqslant \Sigma_{y\in\Omega, X(y)\geqslant aE(X)} X(y)P(y)$$
$$\geqslant \Sigma_{y\in\Omega, X(y)\geqslant aE(X)} aE(X)P(y) = aE(X)P(X \geqslant aE(X)).$$

□

Corollary: If X is a nonnegative integer random variable then $Prob(X > 0) = Prob(X \geqslant 1) \leqslant E(X)$.

Exercise 6: Show that for $s_n = \lceil 2log\ n \rceil$, the probablity that $G_{n,1/2}$ contains a stable set of size s_n goes to zero as n goes to infinity.

Exercise 7: Show that if every element of $\mathcal{D}$ is 4 then the probability $G_{\mathcal{D}}$ is connected is $1 - o(1)$.

Applying Markov's inequality to $(X - E(X))^2$ we obtain:

Chebyshev's Inequality: $P(|X - E(X)| \geqslant k) \leqslant \frac{(E(X^2)-E(X)^2)}{k^2}$.

Proof. Letting $Y = (X - E(X))^2$, we have $|X - E(X)| \geqslant k$ precisely if $Y \geqslant k^2$, and $E(Y) = E(X^2 - 2E(X)X + E(X)^2) = E(X^2) - E(X)^2$. □

We often apply this to $X = \Sigma_{i=1}^N X_i$ where X_i is a 0-1 variable.

In this case $E(X^2) = \Sigma_{i,j} P(X_i = 1\ AND\ X_j = 1)$.

Exercise 8: Show that for any p, n and the k guaranteed to exist by Exercise 2, we have (i) with probability $1 - o(1)$, $G(n,p)$ has at least $\frac{n^{1/4}}{2}$ vertices of degree at least k, and (ii) The expected number of vertices of $G_{n,p}$ of degree at least k wich have the same degree as another vertex is $o(n^{1/4})$. Deduce that with probability $1 - o(1)$, the set of vertivecs $\{v|\ \nexists u \neq v\ s.t\ d(u) = d(v)\}$ of $G_{n,p}$ contains more than $\frac{n^{1/4}}{3}$ vertices.

We can obtain much better bounds for sums of independent variables.We begin with the case of $Bin(n,p)$.

Chernoff Bound: $P(|Bin(n,p) - E(Bin(n,p)| \geqslant t) \leqslant 2e^{\frac{-2t^2}{3pn}}$.

We can obtain a similar bound for sums of independent variable which lie between 0 and 1, which are ot identically dstributed.

Hoeffding's Inequality Suppose X is the sum of n independent variables each lying between 0 and 1, then $P(|X - E(X)|) \geqslant t) \leqslant 2e^{\frac{-2t^2}{n}}$.

We can also obtain bounds for random variables define by a sequence of independent variables which are not just sums, provided each choice in the sequence can only affect the random variable by a limited amount. To take a concrete example suppose we generate a sequence of n independent random variables $T_1, .., T_n$ and let X be the longest monotone subsequence. Then, changing T_i can affect X by at most one, since X is at most one more than the length of the longest monotone subsequence in $T_1, .., T_{i-1}, T_{i+1}, ..., T_n$.

McDiarmid's Inequality Suppose X is a random variable determined by a sequence of n independent variables $T_1, ..., T_n$ such that changing the outcome of T_i can change the value of X by at most c_i. Then $P(|X - E(X)|) \geqslant t) \leqslant 2e^{\frac{-2t^2}{\Sigma_{i=1}^n c_i^2}}$.

We remark that we obtain Hoeffding's inequality as a corollary by setting each $c_i = 1$.

We can actually strengthen this result by weakening the condition and only insisting that each outcome in the sequence changes the conditional expected value of X by a bounded amount. To give a concrete example let $T_2, ..., T_n$ be uniformly chosen elements of $\{1, 2.3\}$, let Z_1 be a uniformly chosen element of $\{H, T\}$, and then for $i = 2, ..., n$ we choose Z_i in $\{H, T\}$ so that $Z_i = Z_{i-1}$ precisely if $T_i \in \{1, 2\}$. I.e. we perform a sequence of random coin flips, where in the first step we flip a fair coin, and in the remaining steps we flip a biased coin, which yields the previous result with probability $\frac{2}{3}$. We let X be the number of heads obtained. Symmetry tells us that $E(X) = \frac{n}{2}$. We cannot use McDiarmid's Inequality to show X is concentrated, because changing the value of T_i can change the value of X by $n - i$ (if every T_j with $j > i$ is 1). However:

Exercise 9: For any choice , $t_2, ..t_{i-1}$ each in $\{H, T\}$, we have $|E(X|T_1 = t_1, .., T_{i-1} = t_{i-1}, T_i = H) - E(X|T_1 = t_1, .., T_{i-1} = t_{i-1}, T_i = T)| \leqslant 2$.

Which allows us to apply the

Simplified Azuma's Inequality; Suppose X is a random variable determined by a sequence of n independent trials such that for any sequenced of possible outcomes for the first $i-1$ trials $t_1, ..., t_{i-1}$ and two possible outcomes t_i and t_i' for T_i we have:

$$|E(X|T_1 = t_1, .., T_{i-1} = t_{i-1}, T_i = t_i) - E(X|T_1 = t_1, .., T_{i-1} = t_{i-1}, T_i = t_i')| \leqslant c_i.$$

Then: $P(|X - E(X)|) \geqslant t) \leqslant 2^{\frac{-2t^2}{\Sigma_{i=1}^n c_i^2}}$.

3. Further Background

3.1. **Counting.** We can specify a tree on a set S of n (labelled) vertices by specifying a leaf l, its neighbour, and the tree formed by $S - v$. It follows, by induction, that there are at fewer than n^{2n} trees on S. Now, there are $\frac{n!}{2}$ distinct paths on S, each one corresponding to a permutation on S and its reverse. Hence there are at least this many trees. In fact, an elegant argument using Pruefer Sequences yields:

There are n^{n-2} trees on a set of n labelled vertices

The Bell number $\mathcal{B}_n$ is the number of (unlabelled) partitions of $1, ..., n$. We can specify an ordered partition into k parts by specifying the index of the element containing each i, which yields at most k^n choices. Of course this is a gross overcount, If there are k elements in an unordered partition of $\{1, .., n\}$ then there are $k!$ ordered partitions corresponding to it. So, an upper bound on B_n is $\Sigma_{k=1}^{n} \frac{k^n}{k!}$. This is an upper bound since we only want to count the fraction of the k^n assignments where each of the elements is non-empty. It is easy to see that the k giving the maximum term in this sum satisfies $k = (1 + o(1))\frac{n}{ln\ n}$. This shows that $B_n = O(n(\frac{(1+o(1))n}{ln\ n})^n)$. The following exercvise shows that this is not too far from the truth.

Exercise 10: Show that for any $\epsilon > 0$ and $k < (1 - \epsilon)\frac{n}{log\ n}$, there are $(1 + o(1))\frac{k^n}{k!}$ partitions of $\{1, .., n\}$ into k non-empty parts.

There are tighter estimates of B_n, which we shall use. We also need:

Exercise 11: Show that there are postive c_1 and c_2 such that for sufficiently large n, $\frac{c_1 n}{logn} \leqslant \frac{\mathcal{B}_n}{\mathcal{B}_{n-1}} \leqslant \frac{c_2 n}{logn}$.

3.2. **Ramsey. Exercise 12:** Show that every graph G contains a stable set S and a clique C such that $|S| + |C| \geqslant log\ |V(G)|$.

3.3. **Regular Pairs.** Our discussion of the structure of a random graph without H as an induced subgraph will require the use of Szemeredi's celebrated Regularity Lemma. We present some relevant definitions.

For two disjoint subsets X and Y of $V(G)$, $E(X, Y)$ is the set of edges between A and B, and the *density* between A and B denoted $d(A, B)$, is $\frac{|E(A,B)|}{|A||B|}$.

For $\epsilon > 0$. We say a pair of disjoint subsets A and B of the vertex set of a graph G are ϵ-*regular* if for every $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geqslant |A|$ and $|B'| \geqslant |B|$ we have $|d(A', B') - d(A, B)| \leqslant \epsilon$.

Exercise 13 Show that for $\epsilon < 1/2$ if A and B are ϵ^2 regular and $C \subseteq A$, $D \subseteq B$ with $|C| \geqslant \epsilon|A|$ and $|D| \geqslant \epsilon|B|$ then (C, D) is ϵ-regular.

Exercise 14: Show that if A and B are ϵ-regular then:

$$|\{x \in\ A\ s.t.(d(A, B) - \epsilon)|B| \leqslant (N(X) \cap B) \leqslant (d(A, B) + \epsilon)|B|\}| \geqslant (1 - 2\epsilon)|A|.$$

Exercise 15: Show that if (A, B), (A, C), and (B, C) are ϵ-regular and $d(A, B), d(B, C), d(A, C) > 3\epsilon$ then G contains moe than $(1-2\epsilon)(d(A, B)-\epsilon)(d(B, C)-\epsilon)(d(A, C)-\epsilon)|A||B||C|$ triangles.

3.4. **Graph Minors.** We *contract* an edge xy in a graph G by deleting x and y and adding a new vertex z adjacent to $N(X) \cup N(y) - x - y$. We say H is a *minor* of G, and write $H <_M G$ if we can obtain H from G by a sequence of edge deletions, vertex deletions, and edge contractions. A minor of G is *proper* if it is not G itself.

Obviously every subgraph of G is a minor of G. Furthermore $<_M$ is clearly transitive. In particular if $H \not<_M G$ then $K_{|V(H)|} \not<_M G$.

The average degree of G is $\frac{2|E(G)|}{|V(G)|}$.

Exercise 16: Show that if for some integer a G has average degree at least a but no proper minor does then for every edge xy of G, $|N(x) \cap N(y)| > \frac{a-2}{2}$.

Exercise 17: Deduce that if G has average degree at least 2^{l-1} then it contains K_l as a minor.

3.5. **H-free Graphs.** A graph is H-free if it contains no induced subgraph isomorphic to H.

Exercise 18 : Show that every component of a P_3-free graph is a clique. Deduce that there are $\mathcal{B}_n$ P_3-free graphs on n vertices.

Exercise 19: Show that for every P_4-free graph either G or $\overline{G}$ is disconnected.

P**Exercise 20:** Deduce that the number of connected P_4 free graphs on $n > 1$ vertices is the same of the number of rooted trees in which every non-leaf has two children, there are n leaves, and the root is not a leaf. Deduce that the number of P_4-free graphs on n vertices is at most $(2n)^{2n}$.