[7] B. Bollobás. The chromatic number of random graphs. *Combinatorica* 8:49-55, 1988

[8] P. Erdős. Some remarks on the theory of graphs. *Bulletin of the Amer. Math. Soc.* 53:292-294, 1947

[9] P. Erdős and H. Hanani, On a limit theorem in combinatorial analysis, *Publ. Math. Debrecen*, 10 (1963), 10-13.

[10] P. Erdős and A. Rényi. On the evolution of random graphs. *Magyar Tud. Akad. Mat. Kut. Int. Közl* 5:17-61, 1960

[11] D. J. Kleitman. On a combinatorial problem of Erdős. *J. of Combinatorial Theory* 1:209-214, 1966

[12] P. Erdős and L. Lovász. Problems and results on 3-chromatic hypergraphs and some related questions, in *Infinite and Finite Sets*, A. Hajnal et. al. eds, North Holland 1975, 609-628

[13] S. Janson, T. Luczak and A. Rucinski. An exponential bound for the probability of nonexistence of specified subgraphs of a random graph, in *Procedings of Random Graphs '87*, M. Karonski et. al. eds, J. Wiley 1990, 73-87

[14] S. Janson. Poisson approximation for large deviations. *Random Stuctures & Algorithms* 1:221-230, 1990

[15] J. Matoušek, Tight upper bounds for the discrepancy of halfspaces. KAM Series (Tech. Report), Charles University, Prague 1994

[16] J. Matoušek and J. Spencer. Discrepancy in Arithmetic Progressions (to appear)

[17] V. Rödl, On a packing and covering problem, *European Journal of Combinatorics*, 5 (1985), 69-78.

[18] K. F. Roth. Remark concerning integer sequences. *Acta Arithmetica* 9:257–260, 1964

[19] J. Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289:679–706, 1985

[20] M. Talagrand, A new isoperimetric inequality for product measure and the tails of sums of independent random variables, *Geometric and Functional Analysis* 1:211-223, 1991

To determine if E survives at time c we create a tree with root E. If $A \supset E$ and $x_A < c$ we consider the $Q-1$ l-sets $E' \subset A$, $E' \neq E$, as a brood of children of E, born at time x. If all these E' survive at time x then either A is placed in F at time x or some $A' \supset E$ had already been placed in F. Either way E does not survive at time c.

In Asymptopia this becomes a continuous time birth process. E, now Eve, has birthdate c. Time goes backwards. Eve gives birth to broods of size $Q-1$ by a Poisson process with unit density. Children with birthdate x in turn have broods in $[0,x)$ by the same process. With probability one a finite tree T is produced. Survival is defined inductively. Childless E' survive and E' does not survive if and only if she has a brood all of whom survive. Let $f(c)$ be the probability Eve survives. Some technical work gives $\lim_{n\to\infty} f_n(c) = f(c)$.

In Asymptopia we estimate $f(c) - f(c+\Delta c)$. The difference for Eve is if she has no surviving broods born in $[0,c)$, a brood born in $[c, c+\Delta c)$, and that brood all survive. For Δc small

$$f(c) - f(c+\Delta c) \sim f(c)(\Delta c) f(c)^{Q-1}$$

Here we bring out the most powerful tool of all, Calculus! In the limit the derivative $f'(c) = -f(c)^Q$. Eve born at $c=0$ is always childless so $f(0)=1$. We solve the differential equation

$$f(c) = [1 + (Q-1)c]^{-1/(Q-1)}$$

For any $\epsilon > 0$ we find c and then n so that on average fewer than $\epsilon\binom{n}{l}$ E survive at time c. Thus there *exists* an outcome for which fewer than $\epsilon\binom{n}{l}$ E survive. Then the $A \in F_c$ must cover $(1-\epsilon)\binom{n}{l}$ sets E and

$$|F| \geq |F_c| \geq (1-\epsilon)\binom{n}{l} / \binom{k}{l}$$

as desired.

References

[1] N. Alon. A parallel algorithmic version of the local lemma. *Random Stuctures & Algorithms* 2:367-378, 1991

[2] N. Alon and J. Spencer. *The probabilistic method*, J. Wiley & Sons, 1993

[3] J. Beck. Roth's estimate on the discrepancy of integer sequences is nearly sharp. *Combinatorica* 1(4):319–325, 1981

[4] J. Beck. An algorithmic approach to the Lovász Local Lemma I. *Random Stuctures & Algorithms* 2:343-365, 1991

[5] J. Beck and T. Fiala. Integer-making Theorems. *Discrete Applied Math.* 3:1-8, 1981

[6] B. Bollobás. *Random Graphs*, Academic Press, 1985

that the proper magnification with which to slow down the double jump is

$$p = \frac{1}{n} + \frac{\lambda}{n^{4/3}}$$

This narrower range of p is called the Phase Transition. When $\lambda = \lambda(n) \to -\infty$ the largest components are all of size $o(n^{2/3})$, they are all almost the same size and they are all trees. The Phase Transition has not started. By the time $\lambda = \lambda(n) \to +\infty$ there is a dominant component whose size is $\gg n^{2/3}$ while all other components have size $o(n^{2/3})$. Moreover the complexity (defined as edges minus vertices) of the dominant component goes to infinity. In Asymptopia the situation at λ constant is given by an infinite sequence $c_1 > c_2 > \ldots$, representing components of sizes $c_1 n^{2/3}, c_2 n^{2/3}, \ldots$ in $G(n,p)$. We think of this as an infinite asteroid belt with asteroids of these sizes. The distribution of these sequences is complex. But the dynamic situation, moving from time λ to time $\lambda + d\lambda$ is easy to describe. Given components of sizes $c_i n^{2/3}, c_j n^{2/3}$ there are $c_i c_j n^{4/3}$ potential edges between them and $n^{2/3} d\lambda/2$ random edges are being selected so they are joined with probability $\sim c_1 c_2 d\lambda$. In Asymptopia we have a peculiar physics in which with probability $c_1 c_2 d\lambda$ asteroids of sizes c_1, c_2 merge to form a new asteroid of size $c_1 + c_2$. Each asteroid further has a complexity x_i, the complexity of the component. For λ large negative most of the components will be trees so $x_i = -1$. When asteroids of complexities x_i, x_j merge the merged asteroid has complexity $x_i + x_j - 1$. With λ large negative the asteroids are all tiny but as λ increases moderate size asteroids are created. This physics favors the rich, a larger asteroid is more likely to merge with others and so become still larger. Computer experiments reveal the process quite strikingly, when $\lambda = -4$ the sizes are small while by $\lambda = +4$ in over 90% of the cases a clear dominant component has emerged.

4.2 Asymptotic Packing

For $2 \leq l < k < n$ let $m(n,k,l)$ denote the maximal size of a family F of k-element subsets of $\{1, \ldots, n\}$ so that no l-set E is contained in more than one $A \in F$. We set $Q = \binom{k}{l}$ for notational convenience. Elementary counting gives $m(n,k,l) \leq \binom{n}{l}/Q$, with equality holding if and only if there is an appropriate tactical configuration. (For $l = 2, k = 3$ these are the Steiner Triple Systems.) In 1963 Paul Erdős and Haim Hanani [9] conjectured that for all $2 \leq l < k$

$$\lim_{n \to \infty} m(n,k,l) Q / \binom{n}{l} = 1$$

This was first proven by Vojtech Rödl [17]. Here we outline a new proof. Indeed, we show that a random greedy algorithm gives F of desired size.

We describe a greedy algorithm with a handy parametrization. Assign to each k-set A a random real $x_A \in [0, \binom{n-l}{k-l}]$. This orders the k-sets. Consider them in order accepting A if no B with $|A \cap B| \geq l$ has already been accepted. Let F_c be the family of A accepted with $x_A < c$. An l-set E is said to survive at "time" c if no $A \in F_c$ contains E.

ask a precise question. Fix $k = 10$. Given $S_1, \ldots, S_n \subseteq [n]$ as above, can the desired χ be found with a polynomial (in n) time algorithm? Even allowing randomized algorithms the answer is not clear. Though LLL guarantees $\Pr[\wedge_{i \in I} \overline{B}_i] \neq 0$ it will be exponentially small in n so checking random χ would take expected exponential time. As stated the problem remains open. But a recent breakthrough by J. Beck[4] gives an algorithm when k is somewhat larger.

We outline Beck's idea as a randomized algorithm though it can be, and originally was, expressed in deterministic fashion. Fix $k = 100$ for definiteness. First $[n]$ is colored randomly. Any S_i with more than 80 (say) points in one color is considered dangerous. All points in dangerous sets are uncolored. If S_i still has red and blue colors, fine. Otherwise we say S_i survives and let S_i^* be the set of uncolored points. Then $|S_i^*| \geq 20$ for otherwise it had had more than 80 points all one color, so it was dangerous and all points were uncolored. Let $\mathcal{F}^*$ be the family of S_i^*. We want a 2-coloring χ of $\mathcal{F}^*$ with no S_i^* monochromatic. Having picked $100, 80, 20$ appropriately LLL applies and χ exists. But isn't this begging the question. Surprisingly, no. The family $\mathcal{F}^*$ has, almost surely, a quite simple structure. Make a graph G with vertices the indices $1 \leq i \leq n$ and adjacency $i \sim j$ if $S_i \cap S_j \neq \emptyset$. Each i has at most 10^4 neighbors. For S_i to survive one of its neighbors must be dangerous, and this occurs with probability at most a very small constant ϵ. Let G^* be the restriction of G to the surviving i. Imagine that each i survived with independent probability ϵ. When i survived it would have in G^* on average $\gamma = 10^4\epsilon$ surviving neighbors who would have on average γ^2 further neighbors, etc. With $\gamma < 1$ the neighborhood of i looks locally like a birth process which will almost surely die. An even better analogy is to components of the random graph $G(n, \frac{\gamma}{n})$ with $\gamma < 1$. There, as discussed in §4.1, all components are of size $O(\ln n)$ Of course, the i do not survive independently, when $i \sim j$ the dependence can be quite strong. Nonetheless Beck showed that G^* almost surely has all components of size $O(\ln n)$. The coloring of $\mathcal{F}^*$ then breaks into coloring the at most n components separately. Each component has $O(\ln n)$ sets hence $O(\ln n)$ vertices. On each component a coloring χ exists. Beck finds it by using exhaustive search! This takes exponential time but the problem has only logarithmic size so the time is polynomial in n. Alon[1] has given an alternate, parallelizable, version of this algorithm and many applications. Still, the general, if ill-formed, question of whether LLL always admits an algorithmic implementation remains open. More likely the opposite is true. A class of problems may well be found where the existence of solutions are guaranteed by LLL but a polynomial time algorithm to find them would violate usual assumptions in complexity theory.

4 Adventures in Asymptopia

4.1 Inside the Double Jump

In their original [10] Paul Erdős and Alfred Rényi discovered what they called the "double jump" in the evolution of the random graph $G(n, p)$ around $p = n^{-1}$. When $p = \gamma n^{-1}$, $\gamma < 1$, all components of G are small, the largest of size $\Theta(\ln n)$, but when $\gamma > 1$ a giant component of size $\Theta(n)$ has been created. We now know

With care we can ensure the entropy requirement and that $2\sum f(s) = O(n^{1/4})$. This gives a substantial partial coloring of $[n]$ with $|\chi(A)| = O(n^{1/4})$ for all $A \in \mathcal{F}$. The iteration of this method to get a full coloring χ (without losing a logarithmic factor!) uses interesting but noncombinatorial ideas.

Matoušek[15] applied entropy to discrepancy of halfplanes. Let P be a set of n points in the plane and $\mathcal{F}$ the family of $H \cap P$, H a halfplane. Here the decomposition is more difficult, the end result again being a family $\mathcal{G}$ so that all $A \in \mathcal{F}$ are expressible in terms of $B \in \mathcal{G}$ of distinct cardinalities 2^j. Again $\mathcal{G}$ has $\sim n^2 s^{-2}$ sets of size s and the entropy argument gives a partial coloring χ – which again can be extended to a full coloring χ – with $|\chi(A)| = O(n^{1/4})$ for all $A \in \mathcal{F}$. This result is best possible up to constants and the method works for halfspaces in R^d for any constant d. Indeed discrepancy of halfplanes came first and motivated the reinvestigation of Roth's result.

Let $\vec{v}_i = (a_{i1}, \ldots, a_{in}) \in R^n$, $1 \leq i \leq n$. For $\chi : [n] \to \{-1, +1\}$ set

$$\vec{S} = \sum_{i=1}^{n} \chi(i)\vec{v}_i = (L_1, \ldots, L_n)$$

with $L_j = \sum_i \chi(i)a_{ij}$. Entropy methods give that if $|\vec{v}_i|_\infty \leq 1$ there exists χ with $|\vec{S}|_\infty \leq cn^{1/2}$. (When $a_{ij} \in \{0, 1\}$ this reduces to n sets on n points and the same proof applies.) Linear algebra methods[5] give that if $|\vec{v}_i|_1 \leq 1$ there exists χ with $|\vec{S}|_\infty \leq 2$. Assume now $|\vec{v}_i|_2 \leq 1$. Set $\sigma_j^2 = \sum_i a_{ij}^2$ so $\sum \sigma_j^2 = \sum \sum a_{ij}^2 \leq n$. Let χ be random, L_i acts like $\sigma_i N$. For k large $ENT(\sigma N, k) < \epsilon$ when $\sigma \sim 1$. Further $ENT(\sigma N, k) < \epsilon\sigma^2$ for *all* σ. One calculates $\sum ENT(L_i, k) < \epsilon n$ so there exists $\chi : [n] \to \{-1, 0, +1\}$ with many $\chi(i) \neq 0$ and $|\vec{S}|_\infty \leq K$. Here iteration fails! More precisely, one may [19] iterate the process $O(\ln n)$ times to give χ with all $\chi(i) = \pm 1$ and $|\vec{S}|_\infty = O(\ln n)$. Still open is a challenging conjecture of J. Komlós that such χ exists with $|\vec{S}|_\infty \leq K$.

3 Algorithmic Sieve

Let $B_i, i \in I$ be events, I finite. Let $\sim$ be a symmetric relation on I so that B_i is mutually independent of all B_j with $i \not\sim j$. This includes the Janson scenario of §1 but is far more general.

Lovász Local Lemma[12] (symmetric case). If all $\Pr[B_i] \leq p$ and, for each $i \in I$, $i \sim j$ for at most d $j \in I$ and if $p < d^d(d+1)^{-(d+1)}$ then $\wedge_{i \in I} \overline{B}_i \neq \emptyset$.

The strength of LLL is that I may be of arbitrary size. With B_i as bad events it sieves out a good outcome. We'll concentrate on one example. Let $S_1, \ldots, S_n \subseteq [n]$ with all $|S_i| = k$ and all j in precisely $k+1$ sets S_i. We want a coloring $\chi : [n] \to \{Red, Blue\}$ so that no S_i is monochromatic. Let χ be random and let B_i be the event that S_i is monochromatic. We naturally define $i \sim i'$ when $S_i \cap S_{i'} \neq \emptyset$. Then $p = 2^{1-k}$ and $d = k^2$. For k large ($k = 10$ suffices) the LLL conditions hold and χ exists.

The probabilistic method has always had a magical quality - just where is the coloring, graph, tournament or whatever that we have proved exists? Here we can

and

$$\sum_{j=0}^{\gamma n} \binom{n}{j} < 2^{n(1-\epsilon)}$$

Then there is a partial coloring χ of Ω with

$$|\chi(S_i)| \leq b_i \text{ for all } i$$

and more than $2\gamma n$ points $x \in \Omega$ colored.
Proof. Let $\chi : \Omega \to \{-1, +1\}$ and define

$$L(\chi) = (R_{b_1}(\chi(S_1)), \ldots, R_{b_v}(\chi(S_v))$$

Entropy, critically, is subadditive so L has entropy at most ϵn. Therefore some value of L obtained with probability at least $2^{-\epsilon n}$, and some $2^{(1-\epsilon)n}$ colorings χ have the same L-value. Colorings χ can be considered points on the Hamming Cube $\{-1, +1\}^n$. A classic result of D. Kleitman[11] gives that some two χ_1, χ_2 of these must differ in at least $2\gamma n$ coordinates. Then $\chi = (\chi_1 - \chi_2)/2$ gives the desired partial coloring. □

Its best to consider $ENT(n, b)$ under the parametrization $b = \lambda n^{1/2}$. Then $R_b(S_n)$ is roughly $R_\lambda(N)$, with N standard Gaussian. For λ large $ENT(n, b) < e^{-c\lambda^2}$, the terms $R = 0, \pm 1$ dominating. In particular, for λ a large constant $ENT < \epsilon$. For λ small $ENT(n, b) < c \ln(\lambda^{-1})$, the dominating factor being that R is roughly uniform for $|i| = O(\lambda^{-1})$.

Suppose $\mathcal{F}$ consists of n sets on an n-set Ω, so all sets have size at most n. For λ a large constant (six will suffice) the Theorem gives a coloring with only a small (but fixed) fraction of the points uncolored and all $|\chi(A)| \leq \lambda n^{1/2}$. Appropriately iterating this author[19] showed that for suitable constant λ one can find χ as above with *no* points uncolored.

Let $\Omega = [n]$ and $\mathcal{F}$ be the arithmetic progressions on $[n]$. The discrepency $disc(\mathcal{F})$ is the least $g(n)$ for which there is a $\chi : \Omega \to \{-1, +1\}$ with $|\chi(A)| \leq g(n)$ for all $A \in \mathcal{F}$. In 1964 K. F. Roth[18] used analytic methods to show $disc(\mathcal{F}) > cn^{1/4}$. The upper bound has been lowered from $n^{.5+o(1)}$ to $n^{1/3+o(1)}$ to $n^{1/4} \ln^c n$ [3] over the decades and just recently to $c'n^{1/4}$ by Jiri Matoušek and this author[16]. Beck[3] provided a key decomposition. For each $d \leq n$, $0 \leq i < d$ and $j \geq 0$ with $2^j \leq n$ split $\{x \in [n] : x \equiv i \bmod d\}$ into consecutive intervals of length 2^j, leaving out the excess. Let $\mathcal{G}$ be the family of sets obtained. Any $A \in \mathcal{F}$ can be written $A = B - C$ with $C \subset B$ and both B, C the disjoint union of $S \in \mathcal{G}$ of distinct cardinalities. Thus a coloring χ for which all $S \in \mathcal{G}$ with $|S| = 2^j$ have $|\chi(S)| \leq f(2^j)$ would have the property that $|\chi(A)| \leq 2 \sum_j f(2^j)$ for all $A \in \mathcal{F}$. Calculation gives that $\mathcal{G}$ has roughly $n^2 s^{-2}$ sets of size $s = 2^j$. To get a substantial partial coloring with $|\chi(A)| \leq f(|A|)$ for $A \in \mathcal{G}$ the entropy requirement becomes

$$\sum n^2 s^{-2} ENT(s, f(s)) \leq \epsilon n$$

When $s \sim n^{1/2}$ we may take $f(s) = kn^{1/4}$. For larger s the savings in s^{-2} allows for a smaller $f(s)$ and for smaller s the savings in $s^{1/2}$ also allows for a smaller $f(s)$.

coordinates outside of I. Let y' agree with y on I and agree with z outside of I. By the certification $h(y') \geq b$. Now y', z differ in at most $t\sqrt{f(b)}$ coordinates and so, by Lipschitz,

$$h(z) > h(y') - t\sqrt{f(b)} \geq b - t\sqrt{f(b)}$$

but then $z \notin A$, a contradiction. So $\Pr[X > b] \leq 1 - \Pr[A_t]$ so

$$\Pr[X < b - t\sqrt{f(b)}] \Pr[X \geq b] \leq e^{-t^2/4}$$

As the right hand side is continuous in t we may replace $<$ by $\leq$ giving the Corollary. □

Letting b (or $b - t\sqrt{f(b)}$) be the median of X the Corollary gives a sharp concentration result. For example, let $\Omega = [0,1]^n$ with uniform distribution and let $X(x_1, \ldots, x_n)$ be the length of the longest monotone subsequence of $x_1, \ldots, x_n$. X is Lipschitz and f-certifiable with $f(s) = s$ as a monotone subsequence certifies itself. It is known that $X \sim 2\sqrt{n}$ almost surely. Therefore X almost surely lies within $n^{1/4}\omega(n)$ ($\omega(n) \to \infty$) of its median.

In $G(n, .5)$ let X be, as before, the maximal number of edge disjoint k-cliques. X is Lipschitz and f-certifiable with $f(s) = \binom{k}{2}s$ as the s k-cliques certify themselves. While medians are notoriously difficult to calculate tight concentration yields that the median $b \sim \mu > cn^2k^{-4}$ as previously discussed. Setting $t = bf(b)^{-1/2}$

$$\Pr[\omega(G) < k] = \Pr[X = 0] = \Pr[X \leq b - t\sqrt{f(b)}] < 2e^{-t^2/4} < ce^{-c'n^2 \ln^{-6} n}$$

2 Entropy

Let $\mathcal{F}$ be a family of subsets of Ω. A two-coloring is a map $\chi : \Omega \to \{-1, +1\}$. For $A \subseteq \Omega$ define $\chi(A) = \sum_{a \in A} \chi(a)$ so that $|\chi(A)|$ is small if the coloring is "nearly balanced" on A. An object of discrepancy theory is to find χ so all $|\chi(A)|$, $A \in \mathcal{F}$, are small. Its convenient to also define partial colorations as maps $\chi : \Omega \to \{-1, 0, 1\}$, a is called colored when $\chi(a) \neq 0$, $\chi(A)$ is as before.

Under random coloring of an n-set A, $\chi(A)$ has distribution S_n, roughly Gaussian with zero mean and standard deviation $n^{1/2}$. Chernoff bounds give $\Pr[|\chi(A)| > \lambda n^{-1/2}] < 2e^{-\lambda^2/2}$. When $\mathcal{F}$ consists of m sets, each of size n, one sets $\lambda = (2\ln(2m))^{1/2}$ so these "failure events" each have probability less than $\frac{1}{m}$ and thus there exists χ with all $|\chi(A)| \leq \lambda\sqrt{n}$. With entropy we can sometimes do better.

Define the roundoff function $R_b(x)$ as that integer i with $2bi$ closest to x. Note $R_b(S_n) = 0$ when $|S_n| < b$. Define $ENT(n, b)$ to be the entropy of the random variable $R_b(S_n)$.

Theorem: Let $\mathcal{F} = \{S_1, \ldots, S_v\}$ with $|\Omega| = n$ and $|S_i| = n_i$. Suppose b_i, ϵ and $\gamma < \frac{1}{2}$ are such that

$$\sum_{i=1}^{v} ENT(n_i, b_i) \leq \epsilon n$$

B_i. Let ϵ be an upper bound for all $\Pr[B_i]$. Set

$$M = \prod \Pr[\overline{B}_i] \text{ and } \Delta = \sum_{i \sim j} \Pr[B_i \wedge B_j]$$

Janson's Inequality:

$$M \leq \Pr[\wedge \overline{B}_i] \leq M e^{\frac{1}{1-\epsilon}\frac{\Delta}{2}}$$

Generalized Janson Inequality: If $\Delta \geq \mu(1-\epsilon)$ then

$$\Pr[\wedge \overline{B}_i] \leq e^{-\mu^2(1-\epsilon)/\Delta}$$

In many cases $\epsilon \to 0$, $\Delta \to 0$ and $M \sim e^{-\mu}$ so that Janson's Inequality gives $\Pr[X = 0] \sim e^{-\mu}$. In this sense Janson's Inequality acts as a Poisson approximation for X, though with particular emphasis at $X = 0$. For example, when $p = c/n$ and $A_{ijk} = \{\{i,j\},\{i,k\},\{j,k\}\}$ range over all triangles these conditions hold and $G(n,p)$ is trianglefree with probability $\sim \exp(-c^3/6)$, as known to Erdős and Rényi. Sweeping generalizations of this are given in [13] where the first proof of Janson's Inequality may be found. Other proofs and generalizations are given in [14][2].

Applying Janson to $\Pr[\omega(G(n,.5)) < k]$ we let $A_S = [S]^2$, S ranging over the k-sets of vertices. Then $\epsilon \to 0$, $\mu = f(k)$. Δ is the expected number of edge overlapping k-cliques, calculation gives domination by cliques overlapping in a single edge and $\Delta \sim \mu^2(2k^4n^{-2})$. The Poisson approximation does *not* apply but the Extended Janson Inequality gives

$$\Pr[\omega(G(n,.5)) < k] < e^{-c\mu^2/\Delta} = e^{-c'n^2 \ln^{-4} n}$$

The newest result, Talagrand's Inequality, has a similar framework to Azuma. Let $\Omega = \prod_1^m \Omega_i$ be a product probability space. For $A \subseteq \Omega$, $x = (x_1, \ldots, x_t) \in \Omega$ define a "distance" $\rho(A,x)$ as the least t so that for any real $\alpha_1, \ldots, \alpha_m$ with $\sum \alpha_i^2 = 1$ there exists $y = (y_1, \ldots, y_t) \in A$ with $\sum_{x_i \neq y_i} \alpha_i \leq t$. Note critically that y may depend on $\alpha_1, \ldots, \alpha_m$. Set A_t equal the set of all $x \in \Omega$ with $\rho(A,x) \leq t$. Talagrand's Inequality[20]:

$$\Pr[A] \Pr[\overline{A}_t] \leq e^{-t^2/4}$$

Call $X : \Omega \to R$ f-certifiable ($f : N \to N$) if whenever $X(x) \geq s$, $x = (x_1, \ldots, x_m)$, there is a set of at most $f(s)$ indices I that certify $X \geq s$ in that if $y = (y_1, \ldots, y_m)$ has $y_i = x_i$ for $i \in I$ then $X(y) \geq s$.
Corollary: If X is Lipschitz and f-certifiable then for all $t \geq 0$, b

$$\Pr[X \leq b - t\sqrt{f(b)}] \Pr[X \geq b] \leq e^{-t^2/4}$$

Proof. Set $A = \{x : h(x) < b - t\sqrt{f(b)}\}$. Now suppose $h(y) \geq b$. We claim $y \notin A_t$. Let I be a set of indices of size at most $f(b)$ that certifies $h(y) \geq b$ as given above. Define $\alpha_i = 0$ when $i \notin I$, $\alpha_i = |I|^{-1/2}$ when $i \in I$. If $y \in A_t$ there exists a $z \in A$ that differs from y in at most $t\sqrt{f(b)}$ coordinates of I though at arbitrary

be more accurate attributions) bound the "large deviation"

$$\Pr[X > a] < e^{-\lambda a} E[e^{\lambda X}] = e^{-\lambda a} \prod_i E[e^{\lambda X_i}]$$

(See, e.g., the appendix of [2].) The power in the inequality is that it holds for all $\lambda > 0$ and one chooses $\lambda = \lambda(a)$ for optimal results. Suppose, for example, that $|X_i| \leq 1$. One can show $E[e^{\lambda X_i}] \leq \cosh(\lambda) \leq \exp(\lambda^2/2)$, the extreme case when $X_i = \pm 1$ uniformly. Then $\Pr[X > a] < \exp(-\lambda a + \lambda^2 m/2) = \exp(-a^2/2m)$ by the optimal choice $\lambda = a/m$. Intuition is guided by comparison to the Gaussian, in the above example $Var(X_i) \leq 1$ so $Var(X) \leq m$ and the probability of being more than $a = \sigma\sqrt{m}$ of the mean should, and here does, drop like the chance of being σ standard deviations off the mean, like $\exp(-\sigma^2/2)$.

The new inequalities are used when the X_i exhibit slight depencies. To illustrate them, let $G \sim G(n, .5)$. Let $f(x) = \binom{n}{x} 2^{-\binom{x}{2}}$ be the expected number of x-cliques and let $k_0 = k_0(n)$ satisfy $f(k_0) > 1 > f(k_0 + 1)$. Calculation gives $k_0 \sim 2\log_2 n$ and its long been known that $\omega(G)$ is almost surely very close to k_0. Now set $k = k_0 - 4$ so that $f(k) > n^{3+o(1)}$ is large. We show thrice that

$$\Pr[\omega(G) < k] < 2^{-n^2 \ln^{-c} n}$$

(As G may be empty the probability is at least 2^{-cn^2}.) The proof via Azuma's Inequality, given below, was given by Béla Bollobás[7] and was essential to his discovery that the chromatic number $\chi(G) \sim n/(2\log_2 n)$ almost surely.
Azuma's Inequality: Let $\mu = X_0, X_1, \ldots, X_m = X$ be a martingale in which $|X_{i+1} - X_i| \leq 1$. Then $\Pr[X > \mu + a] < \exp(-a^2/2m)$.

In application we use an isoperimetric version. Let $\Omega = \prod_{i=1}^m \Omega_i$ be a product probability space and X a random variable on it. Call X *Lipschitz* if whenever $\omega, \omega' \in \Omega$ differ on only one coordinate $|X(\omega) - X(\omega')| \leq 1$. Set $\mu = E[X]$.
Azuma's Perimetric Inequality: $\Pr[X \geq \mu + a] < e^{-a^2/2m}$.

The connection is via the Doob Martingale, $X_i(\omega)$ being the conditional expectation of X given the first i coordinates of ω. The same inequality holds for $\Pr[X \leq \mu - a]$. The random graph $G(n, .5)$ can be viewed as the product of its $m = \binom{n}{2}$ coin flips. Bollobás set X equal the maximal number of edge disjoint k-cliques. From probabilistic methods he showed $E[X] > cn^2 k^{-4}$. (One may conjecture that the true value is $\Theta(n^2 k^{-2})$.) Then $\omega(G) < k$ if and only if $X = 0$ and

$$\Pr[X = 0] = \Pr[X \leq \mu - \mu] < e^{-\mu^2/2m} = e^{-\Theta(n^2 \ln^{-8} n)}$$

For Janson's Inequalities let Ω be a fixed set and $Y \subseteq \Omega$ a random subset (so, formally, 2^Ω is the probability space) where the events $y \in Y$ are mutually independent over $y \in \Omega$. $G(n,p)$ fits this perfectly with $\Omega = [n]^2$ the set of potential edges and $\Pr[y \in G(n,p)] = p$ for every $y \in \Omega$. Let $A_1, \ldots, A_m \subseteq \Omega$. Let B_i be the event $Y \supseteq A_i$, I_i its characteristic function, $X = \sum I_i$ and $\mu = E[X]$. Write $i \sim j$ if $i \neq j$ and $A_i \cap A_j \neq \emptyset$. Roughly $\sim$ represents dependence of the corresponding

Probabilistic Methods in Combinatorics

Joel Spencer

In 1947 Paul Erdős[8] began what is now called the probabilistic method. He showed that if $\binom{n}{k}2^{1-\binom{k}{2}} < 1$ then there *exists* a graph G on n vertices with clique number $\omega(G) < k$ and independence number $\alpha(G) < k$. (In terms of the Ramsey function, $R(k,k) > n$.) In modern language he considered the random graph $G(n,.5)$ as described below. For each k-set S let B_S denote the "bad" event that S is either a clique or independent set. Then $\Pr[B_S] = 2^{1-\binom{k}{2}}$ so that $\sum \Pr[B_S] < 1$ hence $\wedge \overline{B}_S \neq \emptyset$ and a graph satisfying $\wedge \overline{B}_S$ must exist.

In 1961 Erdős with Alfred Rényi[10] began the systematic study of Random Graphs. Formally $G(n,p)$ is a probability space whose points are graphs on a fixed labelled set of n vertices and where every pair of vertices is adjacent with independent probability p. A graph theoretic property A becomes an event. While in the probabilistic method one generally requires only $\Pr[A] > 0$ from which one deduces the existence of the desired object, in Random Graphs estimate of $\Pr[A]$ is the object itself. Let A denote connectedness. In their most celebrated result Erdős and Rényi showed that if $p = p(n) = \frac{\ln n}{n} + \frac{c}{n}$ then $\Pr[A] \to \exp(-e^{-c})$. We give[2][6] as general references for these topics.

While pure probability underlies these fields most of the basic results use fairly straightforward methods. The past ten years (our emphasis here) have seen the use of a number of more sophisticated probability results. The Chernoff bounds have been enhanced by inequalities of Janson and Talagrand and new appreciation of an inequality of Azuma. Entropy is used in new ways. In its early days the probabilistic method had a magical quality – where is the graph that Erdős in 1947 proved existed. With the rise of Theoretical Computer Science these questions take on an algorithmic tone, having proven the existence of a graph or other structure can it be constructed in polynomial time. A recent success of J. Beck allows the Lovász Local Lemma to be derandomized. Sometimes. We close with two forays into a land dubbed Asymptopia by David Aldous. There the asymptotic behavior of random objects are given by an infinite object, allowing powerful noncombinatorial tools to be used.

1 Chernoff, Azuma, Janson, Talagrand

Let $X = X_1 + \ldots + X_m$ with the X_i mutually independent and normalized so that $E[X] = E[X_i] = 0$. The so-called Chernoff bounds (Bernstein or antiquity might