Random Graphs G22.3033-007
Assignment 8. Due Monday, Apr 3,2006

1. In $G(n, p)$ with $p = c/n$ let X be the number of *isolated* triangles. Let $\mu = E[X]$. In Assignment 4 you calculated the limiting value of μ. Now for each r calculate $E[\binom{X}{r}]$. Use Brun's Sieve to deduce the limiting value of $\Pr[X = 0]$. (Hint: There is only one "picture" for r isolated triangles!)

2. Set $m = n \ln n + cn$ where c is a constant. (Don't worry about integrality.) Let f be a random function from $\{1, \ldots, m\}$ to $\{1, \ldots, n\}$. Call $j \in \{1, \ldots, n\}$ *missed* if there is no $i \in \{1, \ldots, m\}$ with $f(i) = j$. Let X be the number of missed $j \in \{1, \ldots, n\}$.

   (a) Find $E[X]$ precisely.

   (b) Find the limiting value of $E[X]$.

   (c) For $r \geq 2$ find $E[\binom{X}{r}]$ precisely.

   (d) For $r \geq 2$ find the limiting value of $E[\binom{X}{r}]$.

   (e) Apply Brun's Sieve to find the limiting value of $\Pr[X = 0]$.

3. In $G \sim G(n, p)$ let X denote the number of isolated edges – i.e., the number of v, w adjacent to each other and no other vertices.

   (a) Find $E[X]$ precisely.

   (b) Give an explicit parameterization $p = f_1(n) + cf_2(n)$ so that $E[X] \to g(c)$ where $g(c)$ will be an explicit continuous function with $\lim_{c \to -\infty} g(c) = 0$ and $\lim_{c \to +\infty} g(c) = +\infty$. (When X is the number of isolated vertices the parametrization $p = \frac{\ln n}{n} + \frac{c}{n}$ was given in class. This is similar, though the answers are not the same.)

   (c) With the above parametrization set $\mu := E[X] \sim g(c)$. Use the Brun's Sieve method to show that X approaches a Poisson Distribution with mean μ.

   (d) Put everything together to make a statement analogous to the isolated vertices statement of the form: If $p =$ blah blah blah then the probability that G has no isolated edges is yadda yadda yadda.

4. For T_n a tournament on n players $1, \dots, n$ and σ a permutation of the players. We call a pair $\{i, j\}$ of players an upset if the higher ranked player won the game between them, i.e. $(\sigma(i) < \sigma(j)$ and $(j, i) \in T_n)$ or $(\sigma(j) < \sigma(i)$ and $(i, j) \in T_n)$. Otherwise call $\{i, j\}$ a nonupset and define the *fit* $fit(T_n, \sigma)$ as the number of nonupsets minus the number of upsets. Here we show that a random tournament T_n has $fit(T_n, \sigma) < Cn^{3/2}$ for all $\sigma \in S_n$, where C is a computable constant. (This is a long and not easy question. Good understanding will be achieved if you work out the $i = 1$ case below finding β_1 and c_1.) We set $n = 2^t$ and assume (avoiding some technical stuff) that t is a positive integer. For $1 \leq i \leq t$ let $fit_i(T_n, \sigma)$ be the number of nonupsets minus the number of upsets in the games between $\sigma(j), \sigma(k)$ where

$$(2u-2)n2^{-i} < j \leq (2u-1)n2^{-i} < k \leq 2un2^{-i}$$

and $1 \leq u \leq 2^{i-1}$. (The dominating cases turn out to be when i is small. Plugging in $i = 1$ and $i = 2$ both above and below will be helpful in understanding the problem.) Call σ_1, σ_2 i-similar if the pairs $\sigma(j), \sigma(k)$ above are the same for σ_1, σ_2. Note that when this holds $fit_i(T, \sigma_1) = fit_i(T, \sigma_2)$ for any tournament T on n players. This splits S_n into equivalence classes.

(a) Give a precise formula for the number $A_i(n)$ of equivalence classes under the i-similarity. (There is no probability in this part but the counting is a little tricky. Let me just say that if you count the number of ways of partitioning an n-set into 2^i sets, each of size $n2^{-i}$, you going in the right direction but you aren't done.)

(b) Give precisely the distribution of $fit_i(T_n, \sigma)$ for σ fixed and T_n the random tournament. (It will be a distribution we have seen often.)

(c) Let i be fixed, with $n \to \infty$. Find the best constant β_i so that $A_i(n) \leq \beta_i^n$. For this β_i show that $A_i(n)\beta_i^{-n} \to 0$.

(d) Let i be fixed, with $n \to \infty$. Find a constant c_i (the best you can find) so that $\Pr[FAIL_i] \to 0$ where $FAIL_i$ is the event that some $fit_i(T_n, \sigma) > c_i n^{3/2}$.

(e) (*) Now take the formula for c_i you just derived and apply it for all $1 \leq i \leq t$. Show that $\sum_{i=1}^{t} \Pr[FAIL_i] \to 0$. (One approach: Show that each $\Pr[FAIL_i] = o(t^{-1})$. Indeed, the failure probabilities will be much smaller than this.)

(f) Give $C := \sum_{i=1}^{\infty} c_i$ explicitly. Deduce that there exists a tournament T_n on n players with $fit(T_n, \sigma) < Cn^{3/2}$ for all $\sigma \in S_n$. This proof was given by F. de la Vega. The original proof of this result, due to your instructor, has been relegated to the antiBook.

5. Here we give a Ramsey result: There exists a triangle free graph on n vertices with no independent set of size $m = Kn^{1/2} \ln n$, K a large constant. This was first shown by Erdős in 1961, the proof we give was part of the doctoral dissertation of Michael Krivelevich (now at Tel Aviv University) in the 1990s. Later Jeong Han Kim (now at Microsoft) showed that this holds when $m = K_1 n^{1/2}\sqrt{\ln n}$ which matched (except for the constant, which remains open) a bound in the other direction due to Ajtai, Komlós and Szemerédi. We let $G \sim G(n,p)$ with $p = cn^{-1/2}$. We think of c, K fixed and n approaching infinity.

   (a) For a given set S of m vertices let X be the number of edges $\{i,j\}$ of G with both vertices in S. Set $\mu_1 = E[X]$. Find μ_1 asymptotically.

   (b) Use large deviation results (I suggest A.1.13) to bound $\Pr[X < \frac{1}{2}\mu_1]$.

   (c) Find conditions on K, c so that with probability approaching one *every* set S of m vertices has at least $\frac{1}{2}\mu_1$ edges. [It will be helpful to bound $\binom{n}{m} \le (\frac{ne}{m})^m \le n^{m/2}$.]

   (d) For a given set S of m vertices let Y be the number of triangles $\{i,j,k\}$ of G with two or three vertices in S. Set $\mu_2 = E[Y]$. Find μ_2 asymptotically.

   (e) Use Lemma 8.4.1 to bound that probability that, for a given set S of m vertices, there exists a family of $3\mu_2$ edge disjoint triangles, each of which has two or three vertices in S.

   (f) Find conditions on K, c so that with probability approaching one for *every* set S of m vertices there does *not* exist a family of $3\mu_2$ edge disjoint edges, each of which has two or three vertices in S.

   (g) Find conditions on K, c so that $9\mu_2 < \frac{1}{2}\mu_1$.

   (h) Show that there exist K, c satifying simultaneously the conditions of (5c,5f,5g).

   (i) Let K, c be as in (5h). Use Erdős Magic to show that there exists a graph G on n vertices (n sufficiently large) satisfying the conditions (5c,5f).

(j) Let G be as in (5i). Take any maximal family of edge disjoint triangles of G. Delete all these triangles from G (i.e., delete the three edges – do *not* delete the vertices) giving a graph G^*. Argue that G^* is trianglefree. Argue that for every set S of m vertices there is at least one edge $\{i, j\}$ of G^* with both vertices $i, j \in S$.

If you take a number and double it and double it again and then double it a few more times, the number gets bigger and bigger and goes higher and higher and only arithmetic can tell you what the number is when you quit doubling.
from *Arithmetic* by Carl Sandburg