Assignment 3

Random Graphs, Technion, 2017

December 20, 2017

1. **Generalized Bonferroni inequalities.** Let $E_1, \ldots, E_n$ be dependent events, and let X be the number of events among them that occur. The inclusion-exclusion principle states that

$$\Pr[X = 0] = \sum_{t=0}^{n} (-1)^t \sum_{i_1 < \cdots < i_t} \Pr[E_{i_1} \wedge \cdots \wedge E_{i_t}],$$

and the Bonferroni inequalities state that if we stop the sum at $t = s$ then we get an upper bound (if s is even) or a lower bound (if s is odd), matching the sign of the following term. The goal of this exercise is to obtain similar formulas for $\Pr[X = r]$.

   (a) Clearly $X = r$ if there exist $j_1, \ldots, j_r$ such that $E_{j_1}, \ldots, E_{j_r}$ occur but no other events occur. Using the inclusion-exclusion principle, obtain a similar principle of the form

$$\Pr[X = r] = \sum_{t=r}^{n} c_{r,t} \sum_{i_1 < \cdots < i_t} \Pr[E_{i_1} \wedge \cdots \wedge E_{i_t}],$$

   for appropriate constants $c_{r,t}$.

   (b) Use Bonferroni's inequalities to deduce similar inequalities for $\Pr[X = r]$.

   (c) Let X_n denote the number of isolated vertices in $G(n, \frac{\log n + c}{n})$, where $c > 0$ is a constant, and let P be a Poisson random variable with expectation e^{-c}. Show that for every r, $\Pr[X_n = r]$ tends to $\Pr[P = r]$ as $n \to \infty$.

   (d) Show that moreover, $\sum_r |\Pr[X_n = r] - \Pr[P = r]| \to 0$. (In other words, the total variation distance between X_n and P tends to zero.)

2. **Google interview question.** The goal of this exercise is to solve the following question: *There are n whole pills in a bag. At each step, you pick up a random pill (whole or half). If it is whole, you break it into two halves, and consume one of them. If it is a half pill, you just consume it. How many steps does it take before the number of half pills exceeds the number of whole pills?*

   (a) Consider the experiment of throwing m balls into n bins. Denote the number of bins containing exactly k balls after m steps by $X_k(m)$. Suppose that there exist constants $\epsilon, C > 0$ such that $\epsilon n \le m \le Cn$. Show that $\mathbb{E}[X_k(m)] = e^{-m/n} \frac{(m/n)^k}{k!} n + O(1)$ and $\mathbb{V}[X_k(m)] = O(n)$, where the hidden constants can depend on ϵ, C, k.

   (b) Fix $\epsilon, C > 0$ and $k \ge 0$. Show that there exists a function $\delta(n) = o(n)$ such that with high probability, $|X_k(m) - \mathbb{E}[X_k(m)]| \le \delta$ for all m in the range $\epsilon n \le m \le Cn$. (Hint: Consider first only $o(n)$ evenly spread values of m.)

   (c) Let m^* be the first point at which $X_1(m^*) \ge X_0(m^*)$. Show that there exist constants α, β and a function $\eta(n) = o(n)$ such that with high probability, $|m^* - \alpha n| < \eta(n)$ and $|X_k(m^*) - \beta n| < \eta(n)$ for $k = 0, 1$.

   (d) **The interview question.** Let $W(t), H(t)$ be the number of whole pills and half pills (respectively) after t steps. Let t^* be the first time that $H(t) \ge W(t)$. Show that there exist constants a, b and a function $\zeta(n) = o(n)$ such that with high probability, $|t^* - an| < \zeta(n)$, $|W(t^*) - bn| < \zeta(n)$, and $|H(t^*) - bn| < \zeta(n)$. (Hint: Analyze first $H(t^*), W(t^*)$, and deduce the bounds on t^*.)