Homework 9

Problem 1. *Find an example to verify the claim that '(pairwise) independence does not imply mutual independence'. Pls give a detailed proof.*

Solution. (by S. Bernstein)

Suppose X and Y are two independent tosses of a fair coin, where we designate 1 for heads and 0 for tails. Let the third random variable $Z = (X + Y) \bmod 2$.

Then jointly the triple $\langle X, Y, Z\rangle$ has the following probability distribution:

$$\langle X, Y, Z\rangle = \begin{cases} \langle 0,0,0\rangle & \text{with probability } 1/4 \\ \langle 0,1,1\rangle & \text{with probability } 1/4 \\ \langle 1,0,1\rangle & \text{with probability } 1/4 \\ \langle 1,1,0\rangle & \text{with probability } 1/4 \end{cases}$$

$i, j, k \in \{0, 1\}$.

It is easy to verify that $Pr(X = i) = Pr(Y = j) = Pr(Z = k) = 1/2$ and $Pr(X = i, Y = j) = Pr(X = i, Z = k) = Pr(Y = j, Z = k) = 1/4$. i.e., X, Y, Z are pairwise independent.

However, $Pr(X = i, Y = j, Z = k) \neq Pr(X = i) \cdot Pr(Y = j) \cdot Pr(Z = k)$. For example, the left side equals $1/4$ for $\langle x, y, z\rangle = \langle 0, 0, 0\rangle$ while the right side equals $1/8$.

In fact, any of $\langle X, Y, Z\rangle$ is completely determined by the first two components. That is as far from independence as random variables can get.

□

Problem 2. *Show that, if $E_1, E_2, \ldots, E_n$ are mutually independent, then so are $\overline{E_1}, \overline{E_2}, \ldots, \overline{E_n}$.*

Solution. (sketch) It will be enough to prove that for any $2 \le k \le n$, and $\{F_1, F_2, \ldots, F_k\} \subseteq \{E_1, E_2, \ldots, E_n\}$

$$Pr\left(\bigcap_{i=1}^{k} \overline{F_i}\right) = \prod_{i=1}^{k} Pr(\overline{F_i})$$

Let $Pr(F_i) = f_i$, then

$$Pr\left(\bigcap_{i=1}^{k} \overline{F_i}\right) = 1 - Pr\left(\bigcup_{i=1}^{k} F_i\right) = 1 - \sum_{i=1}^{k} f_i + \sum_{1\le i<j\le k} f_i f_j - \sum_{1\le i<j<l\le k} f_i f_j f_l + \cdots$$

The right hand side of the above equation is $(1 - f_1)(1 - f_2)\cdots(1 - f_k) = \prod_{i=1}^{k} Pr(\overline{F_i})$. □

Problem 3. *A monkey types on a 26 -letter keyboard that has lowercase letters only. Each letter is chosen independently and uniformly at random from the alphabet. If the monkey types 1,000,000 letters. what is the expected number of times the sequence "proof" appears?*

Solution. By the linearity of expectation:

$$E[X] = (1/26)^5 \times (1000000 - 4)$$

□

Problem 4. *We have 27 fair coins and one counterfeit coin (28 coins in all), which looks like a fair coin but is a bit heavier. Show that one needs at least 4 weighings to determine the counterfeit coin. We have no calibrated weights, and in one weighing we can only find out which of two groups of some k coins each is heavier, assuming that if both groups consist of fair coins only the result is an equilibrium.*

Solution. Each weighting has 3 possible outcomes, and hence 3 weightings can only distinguish one among 3^3 possibilities.

□

Problem 5.

1. *Prove that, for every integer n, there exists a coloring of the edges of the complete graph K_n by two colors so that the total number of monochromatic copies of K_4 is at most $\binom{n}{4}2^{-5}$.*
2. *Give a randomized algorithm for finding a coloring with at most $\binom{n}{4}2^{-5}$ monochromatic (i.e. single-color) copies of K_4 that runs in expected time polynomial in n.*

Solution.

1. Coloring every edge in K_4 by red or blue with probability $1/2$. The expected value of the total number of monochromatic copies of K_4 is then $2 \times \binom{n}{4} \times \left(\frac{1}{2}\right)^6$. Then there must exist some coloring scheme where the total number of monochromatic copies of K_4 is less or equal to $\binom{n}{4}2^{-5}$ (otherwise the expectation would be strictly larger than $\binom{n}{4}2^{-5}$.

2. Color each edge independently and uniformly. Let $p = Pr\left(X \leq \binom{n}{4}2^{-5}\right)$ where X is the number of chromatic K_4.

$$\begin{aligned}\binom{n}{4}2^{-5} &= \mathbf{E}(X) \\ &= \sum_{i \leq \binom{n}{4}2^{-5}} i \cdot Pr(X = i) + \sum_{i > \binom{n}{4}2^{-5}} i \cdot Pr(X = i) \\ &\geq p + (1-p)\left(\binom{n}{4}2^{-5} + 1\right)\end{aligned}$$

which implies $p \geq \frac{32}{\binom{n}{4}}$. The expected number of sampling before finding a suitable coloring is $1/p = \frac{\binom{n}{4}}{32}$. For each sampling, the time needs to count the number of chromatic K_4 is bounded by $\binom{n}{4}$ which is also polynomial. Thus the expected running time of this algorithm is polynomial.

□

Problem 6. *Use the Lovasz local lemma to show that if*

$$4\binom{k}{2}\binom{n}{k-2}2^{1-\binom{k}{2}} \leq 1$$

then it is possible to color the edges of K_n with two colors so that it has no monochromatic (i.e. single color) K_k subgraph.

Solution. E_i: the i–th K_k is monochromatic. $Pr(E_i) = 2^{1-\binom{k}{2}}$. Consider the dependency graph, for any different E_i and E_j, they are adjacent if the corresponding K_k share at least one edge. Thus the degree of the dependency graph is bounded by $\binom{k}{2}\binom{n}{k-2}$.

According to the Lovasz local lemma, it is possible that none of the E_i happens under the given inequality. □

Problem 7. *What is the expected number of trees with k vertices in $G \in \mathcal{G}(n, p)$?*

Solution. By *Cayley's formula* and the linearity of expectation, it is $\binom{n}{k}k^{k-2}p^{k-1}$ □

Problem 8. *Show that if almost all $G \in \mathcal{G}(n, p)$ have a graph property $\mathcal{P}_1$ and almost all $G \in \mathcal{G}(n, p)$ have a graph property $\mathcal{P}_2$, then almost all $G \in \mathcal{G}(n, p)$ have both properties.*

Solution. The portion of the graphs have both properties equals 1 minus the portion of the graphs which does not have property $\mathcal{P}_1$ or $\mathcal{P}_2$. However the portion of the graph does not have property $\mathcal{P}_1$ or $\mathcal{P}_2$ is bounded by the sum of the portion of the graphs does not have property $\mathcal{P}_1$ and the the portion of the graphs does not have property $\mathcal{P}_2$, which both tend to 0 as n approaches ∞. The claim in the question then follows. $\square$