Problem of the Month

Problem 7: April 2021

For an integer $n \geq 3$, n^2 points form an $n \times n$ square grid.

Define $P(n)$ to be the probability that three distinct points randomly selected from the grid are the vertices of a triangle with positive area. Also define $f(n)$ to be the number of sets of three distinct points from the grid that lie on a common line. We can think of $f(n)$ as the number of sets of three distinct points from the grid that are the vertices of a triangle with area 0.

For instance, with $n = 3$, it can be shown that there are 84 possible ways to select three distinct points, that 8 of the sets of three points lie on a line, and that 76 of the sets of three points form the vertices of a triangle with positive area. Thus, $f(3) = 8$ and $P(3) = \dfrac{76}{84} = \dfrac{19}{21}$.

The goal of this problem is to estimate $P(n)$ for large n. The approach outlined will be to estimate $f(n)$ and use it to estimate $P(n)$.

(a) When $n = 3$, $f(n) = 8$ and $P(3) = \dfrac{19}{21}$. Compute $f(n)$ and $P(n)$ for $n = 4$ and $n = 5$.

(b) For $n \geq 3$, prove that $f(n+1) < f(n)+5n^4+5n^3+5n^2+5n$. This will allow us to understand how quickly $f(n)$ grows which will help to estimate $P(n)$.

(c) Using part (b), prove that $f(n) < n^5$ for all $n \geq 3$.

(d) Prove that there is a constant c with the property that $P(n) > 1 - \dfrac{c}{n}$ for all $n \geq 3$. Use this to explain why the following statement makes sense: "For very large n, it is nearly certain that three points selected randomly from an $n \times n$ grid will be the vertices of a triangle with positive area."

As indicated in part (d), this problem is meant to examine what happens to $P(n)$ as n gets large. Since it seems very difficult to calculate $f(n)$ (and hence, $P(n)$) directly for large n, we instead estimate its value. As long as we carefully keep track of how good/bad the estimates can be, we can say something meaningful about $P(n)$ for large n without actually computing it directly. Very frequently, mathematicians use estimates like these when exact answers are difficult or impossible to obtain. These estimates are often as useful as exact answers.