# AIM Qualifying Review Exam: Probability and Discrete Mathematics

*January 8, 2024*

There are five (5) problems in this examination.

There should be sufficient room in this booklet for all your work. But if you use other sheets of paper, be sure to mark them clearly and staple them to the booklet.

**Problem 1**

Let $M, B, K$ denote the the outputs of rolling a maize, blue, and khaki die, respectively. (A die has six sides, labeled $1, 2, \ldots, 6$, and each side comes up with equal probability $1/6$ when the die is rolled.)

(a) What is the probability that no two dice land on the same number?

(b) Given that no two of the dice land on the same number, what is the conditional probability that $M < B < K$?

(c) What is $P(M < B < K)$?

**Solution**

Ross 3.22, p. 103.

(a) The blue outcome differs from the maize (whatever the latter) with probability $\frac{5}{6}$. Given that, the khaki outcome is different again with probability $\frac{4}{6}$. So all three are different with probability $\frac{6}{6} \cdot \frac{5}{6}\frac{4}{6} = \frac{20}{36} = \frac{5}{9}$.

(b) By symmetry over permuting the colors, $\frac{1}{3!} = \frac{1}{6}$.

(c) The product, $\frac{5}{9} \cdot \frac{1}{6} = \frac{5}{54}$.

**Mathematical concepts:** conditional discrete probability

**Problem 2** Let $Y$ be a continuous random variable without point masses and with density $f_Y$ (so $\int_{-\infty}^{\infty} f_Y(y)dy =$

1). Show that the expectation $\mathbb{E}[Y]$ is

$$\mathbb{E}[Y] = \int_0^\infty P\{Y > y\}dy - \int_0^\infty P\{Y < -y\}dy$$

.

**Solution**

We have $\int_0^\infty P\{Y > y\}dy = \int_{y=0}^\infty \int_{x\geq y} f_Y(x)dx\,dy$. Exchanging the order of integration, we get

$$\int_0^\infty P\{Y > y\}dy = \int_{x=0}^\infty \int_{0\leq y\leq x} f_Y(x)dy\,dx.$$

Now the inner integrand $\int_{0\leq y\leq x} f_Y(x)dy = f_Y(x)\int_{0\leq y\leq x} dy$ is constant, so we can replace $\int_{0\leq y\leq x} dy$ with $x$, getting $\int_0^\infty P\{Y > y\}dy = \int_{x=0}^\infty x f_Y(x)dx$, which is the expectation conditioned on $Y \geq 0$.

Similarly, $\int_0^\infty P\{Y < -y\}dy = \int_{y=-\infty}^0 P\{Y < y\}dy$, changing $y$ to $-y$ in the integrand, limits, and $dy$, then multiplying by $-1$ to swap the limits of integration. This is

$$\int_{y=-\infty}^0 \int_{x\leq y} f_Y(x)dx\,dy = \int_{x=-\infty}^0 \int_{x\leq y\leq 0} f_Y(x)dy\,dx = -\int_{x=-\infty}^0 x f_Y(x)dx.$$

**<u>Problem 3</u>**

The $n$'th triangular number, $A_n$, is defined as $A_n = 1 + 2 + 3 + \cdots + n$. The $n$'th tetrahedral number, $V_n$, is defined as $V_n = A_1 + A_2 + \cdots + A_n$. (Since triangle and tetrahedron start with the same letter, use $A$ as in "area" for triangles and $V$ as in "volume" for tetrahedra.)

(a) Find a formula (without $\cdots$) for $A_n$.

(b) Show by induction that $V_n = \binom{n+2}{3}$, the binomial symbol $\binom{n+2}{3} = \frac{(n+2)(n+1)n}{1\cdot 2\cdot 3}$.

(c) Show by induction that $V_{2n} = 2^2 + 4^2 + 6^2 + \cdots + (2n)^2$. (Note even index in $V_{2n}$.)

Best wishes for $2024 = V_{22} = 1 + (1+2) + (1+2+3) + \cdots + (1+2+\cdots+22) = 2^2 + 4^2 + \cdots + 22^2 = \binom{22+2}{3}$.

**Solution**

(a) We can use the sum of an arithmetic sequence: $A_n = (1+n)+(2+(n-1))+\cdots = \frac{(n+1)n}{2}$, whether $n$ is even or odd. Alternatively, $2A_n - n = n^2$, whence $A_n = \frac{n^2+n}{2}$. E.g., for $n = 3$, we have $2A_3 = 2\cdot 6 = 3^2 + 3$:

```
* ###       *##       #
** ##   =   **#   +   #
*** #       ***       #
```

In general, $\sum_{j=1}^n \binom{j}{r} = \binom{n+1}{r+1}$. So $\sum_{j=1}^n \binom{j}{1} = \binom{n+1}{2}$.

(b) (See also general formula, above.) The case $n = 0$ depends on definitions and special cases, so start with $n = 1$, giving $1 = 1$. Inducting from $n-1$ to $n$ seems to be a bit cleaner than $n$ to $n+1$. Inductively, supposing $V_{n-1} = \binom{n+1}{3}$, then

$$\begin{aligned} V_n &= V_{n-1} + A_n \\ &= \binom{n+1}{3} + \binom{n+1}{2} \\ &= \frac{(n+1)n(n-1)}{6} + \frac{(n+1)n}{2} \\ &= \frac{(n+1)n(n-1) + 3(n+1)n}{6} \\ &= \frac{(n+2)(n+1)n}{6} \\ &= \binom{n+2}{3}. \end{aligned}$$

Note also that $\binom{n+1}{3} + \binom{n+1}{2} = \binom{n+2}{3}$ is the 2-3 diagonal case of the familiar Pascal's triangle relation.

(c) Similarly, we have

$$\begin{aligned} V_{2n} &= V_{2(n-1)} + A_{2n-1} + A_{2n} \\ &= (2^2 + 4^2 + \cdots 4(n-1)^2) + \binom{2n}{2} + \binom{2n+1}{2} \\ &= (2^2 + 4^2 + \cdots 4(n-1)^2) + \frac{2n(2n-1) + (2n+1)(2n)}{2} \\ &= 2^2 + 4^2 + \cdots (2n)^2. \end{aligned}$$

Note the sum of consecutive triangles is a square:

```
***           ***
**   +   #  = **#
*       ##    *##
```

**Mathematical concepts:** induction, binomial symbols

**Problem 4**

(a) Show that if 51 distinct numbers are chosen from $1, 2, \cdots, 100$, then there is a pair that differs by 1.

(b) Show that if 5 points are chosen in the unit square, some pair of points is at distance at most $\frac{\sqrt{2}}{2}$.

**Solution**

Brualdi, 2, p. 83

(a) Consider the 50-partition $\{1, 2\} \cup \{3, 4\} \cup \cdots \cup \{99, 100\}$. One of these partitions must have two of the 51 chosen numbers.

(b) Partition the unit square into four quadrants. Each has diameter $\frac{\sqrt{2}}{2}$.

**Mathematical concepts:** pigeonhole principle

**Problem 5** The Karatsuba algorithm multiplies two $n$-term polynomials, $p$ and $q$, by divide-and-conquer,

more efficiently than the gradeschool algorithm, that looks at all $n^2$ pairs of a term in $p$ and a term in $q$. As a base case, $(a_0 + a_1x)(b_0 + b_1x)$ is done by computing $a_0b_0, (a_0 + a_1)(b_0 + b_1)$, and $a_1b_1$. To multiply $(a_0 + a_1x + a_2x^2 + a_3x^3)(b_0 + b_1x + b_2x^2 + b_3x^3)$, put

$$\begin{aligned} u &= x^2 \\ A_0 &= a_0 + a_1x \\ A_1 &= a_2 + a_3x \\ B_0 &= b_0 + b_1x \\ B_1 &= b_2 + b_3x, \end{aligned}$$

then do $(A_0 + A_1u)(B_0 + B_1u)$, using the base case algorithm on capital letters and on recursive calls: for example, $A_1B_1$ is unwound to $(a_2 + a_3x)(b_2 + b_3x)$. The general case for larger $n$ is similar.

(a) This algorithm can be described as taking a problem of size $n$ and reducing it to $s$ problems of size at most $n/t$. What are $s$ and $t$?

(b) Suppose the original coefficients $a_0, a_1, \dots$ are reals. How many multiplications of real numbers are incurred by the algorithm when $n = 4$? (Do not simplify in case of zero coefficients, etc. And do **not** count $A_1 B_1$ as a multiplication, since it is a virtual multiplication of polynomials, not a multiplication of real numbers.)

(c) How many multiplications of real numbers are incurred for general $n$? Answer in the form $O(n^r)$, i.e., give the exponent $r$, in worst case, without need to calculate a constant factor $c$ in $cn^r$.

**Solution**

Kleinberg and Tardos, pp. 215–216.

(a) The algorithm reduces a problem of size $n$ to 3 problems of size $n/2$.

(b) When $n = 4$ (above), we get exactly 9 multiplications.

(c) In general, we get exactly $3^{\log_2 n} = n^{\log_2 3} = n^{1.58\dots}$ multiplications in worst case when $n$ is a power of 2. If $n$ is not a power of 2, round $n$ up to $m$, which is a power of 2 at most $2n$. Then we get at most $m^{1.58\dots} \le (2n)^{1.58\dots} = 3n^{1.58\dots}$

**Mathematical concepts:** divide and conquer