1. How many distinct subsets does a finite set with n elements have? (The empty set and the set itself are should be included in the count.) (7)

Let A be a set with $|A| = n$, with its elements labeled $1, \dots, n$. Then a subset of A is uniquely characterized by a binary sequence of length n, where a 1 in position i denotes that element i is included in the subset, and a 0 denotes otherwise.

Thus, the number of subsets of A is the number of distinct binary sequences of length n, which is 2^n.

2. Three dice are tossed.

(a) What is the probability of obtaining at least one 6?

(b) What is the probability of obtaining at most one 6?

(5+5)

(a) $P(\text{no } 6) = \frac{5^3}{6^3}$

$P(\text{at least one } 6) = 1 - P(\text{no } 6) = 1 - \frac{5^3}{6^3} \left(= \frac{91}{216} \right)$

(b) $P(\text{exactly one } 6) = 3 \cdot P(\text{exactly one 6 in first die})$

$= 3 \cdot \frac{5^2}{6^3} = \frac{25}{72}$

$P(\text{at most one } 6) = P(\text{no } 6) + P(\text{exactly one } 6)$

$= \frac{5^3}{6^3} + \frac{25}{72} \left(= \frac{25}{27} \right)$

3. Use the method of generating functions to find a closed form expression for the members of the sequence

$$a_n = \tfrac{5}{2}\, a_{n-1} - a_{n-2}$$

where $a_0 = 2$ and $a_1 = \frac{5}{2}$. (10)

Use a generating function

$$\phi(x) = \sum_{n=0}^{\infty} a_n x^n$$

$$\Rightarrow \left(1 - \tfrac{5}{2}x + x^2\right)\phi(x) = a_0 + a_1 x + a_2 x^2 + \ldots$$

$$- \tfrac{5}{2}a_0 x - \tfrac{5}{2}a_1 x^2 - \ldots$$

$$+ \underbrace{a_0 x^2}_{=0} + \underbrace{\ldots}_{=0} \ldots$$

$$= 2 + \tfrac{5}{2}x - 5x = 2 - \tfrac{5}{2}x$$

$$\Rightarrow \phi(x) = \frac{2 - \frac{5}{2}x}{1 - \frac{5}{2}x + x^2} = \frac{2 - \frac{5}{2}x}{(1-2x)(1-\frac{1}{2}x)} = \frac{A}{1-2x} + \frac{B}{1-\frac{1}{2}x}$$

where $(1-\frac{1}{2}x)A + (1-2x)B = 2 - \frac{5}{2}x$

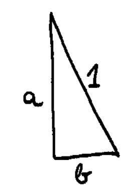
$\Rightarrow A + B = 2$ and $-\frac{1}{2}A - 2B = -\frac{5}{2}$

$\Rightarrow A = B = 1$

$$\Rightarrow \phi(x) = \frac{1}{1-2x} + \frac{1}{1-\frac{1}{2}x} = \left(1 + 2x + (2x)^2 + \ldots\right) + \left(1 + \tfrac{x}{2} + \left(\tfrac{x}{2}\right)^2 + \ldots\right)$$

$$\Rightarrow a_n = 2^n + \frac{1}{2^n}$$

4. Show that, among all right triangles with hypotenuse of length 1, the isosceles triangle has the largest area. (8)



$A = \frac{1}{2}ab \leq \frac{1}{4}(a^2 + b^2)$ by AM-GM. inequality

$= \frac{1}{4}$ by Pythagoras : $a^2 + b^2 = 1$

where equality holds when $a = b$, i.e. the triangle is isosceles.

5. Show that, for arbitrary numbers $a_1, \ldots, a_n$,

$$\sum_{k=1}^{n} a_k \le n^{\frac{1}{2}} \left(\sum_{k=1}^{n} a_k^2 \right)^{\frac{1}{2}} .$$

(7)

$$\sum_{k=1}^{n} a_k \le \sum_{k=1}^{n} 1 \cdot |a_k| \underset{\substack{\uparrow \\ \text{Cauchy}}}{\le} \Bigg(\underbrace{\sum_{k=1}^{n} 1^2}_{=n} \Bigg)^{\frac{1}{2}} \left(\sum_{k=1}^{n} a_k^2 \right)^{\frac{1}{2}}$$

$\square$

6. In class, we computed the probability P that among a group of k people at least two are born on the same day of the year. In the derivation, we assumed that that each day has equal probability of being the birthday of someone. In reality, however, there are weekly and seasonal variations of the distribution of births. Would, under more realistic assumptions on the distribution of births, the quantity P increase or decrease?

Note: No credits without explanation. Full credit for a convincing qualitative argument. Up to five extra credits for a complete proof. (8+5)

1. Qualitative argument:

Consider an extreme case of an uneven distribution of births: One (or more) days(s) is excluded as a possible birthday, with no further bias between the remaining days. Thus, we have a birthday problem with a shorter year, so clearly P goes up.

We conjecture that this tendency holds for all uneven distributions of birth.

2. Proof for the special case $k=2$:

Recall that in the standard birthday problem for $k=2$ and a year with n days,

$$P(\text{no two born on the sames day}) = \frac{n(n-1)}{n^2} = \frac{n-1}{n} = 1 - \frac{1}{n}$$

There is an alternative way of writing this:

$$P(\text{no two born on the same day})$$
$$= \sum_{i=1}^{n} \underbrace{P(\text{person 1 born on day } i)}_{\equiv p_i} \underbrace{P(\text{person 2 not born on day } i)}_{\equiv (1-p_i)}$$

where p_i is the probability that a randomly selected person is born on day i. In the standard case, $p_i = \frac{1}{365}$, but in general, we only require that $0 \leq p_i \leq 1$ with $\sum_{i=1}^{n} p_i = 1$.

$$\Rightarrow P(\text{no two born on same day}) = \sum_{i=1}^{n} p_i - p_i^2 = 1 - \underbrace{\sum_{i=1}^{n} p_i^2}_{\geq \frac{1}{n} \ (*)} \leq 1 - \frac{1}{n}$$

(*) by Cauchy's inequality, see question 5, where equality holds only if the p_i are all the same.

Thus, $P(\text{at least two on same day}) = 1 - P(\text{no two born on same day})$ is <u>minimal</u> in the standard case where $p_i = \frac{1}{n}$.

3. General case

The general case is more difficult, and we start with proving a Lemma.

<u>Lemma:</u> Let $k \leq n$ and $0 \leq p_i \leq 1$ with $\sum_{i=1}^{n} p_i = 1$. Then

$$\sum_{i_1, \dots, i_k \text{ distinct}} p_{i_1} p_{i_2} \cdots p_{i_k} \leq \frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n}$$

with equality if and only if $p_i = \frac{1}{n}$ for $i = 1, \dots, n$.

<u>Proof:</u> We proceed as in the second proof of the AM-GM inequality (Ivanov, p. 51). First, note that, when $p_i = \frac{1}{n}$, the LHS of the inequality can be written

$$\underbrace{\sum_{i_1=1}^{n} \frac{1}{n}}_{n \text{ terms}} \underbrace{\sum_{\substack{i_2=1 \\ i_2 \neq i_1}}^{n} \frac{1}{n}}_{n-1 \text{ terms}} \cdots \underbrace{\sum_{\substack{i_k=1 \\ i_k \neq i_1, \dots, i_{k-1}}}^{n} \frac{1}{n}}_{n-k+1 \text{ terms}} = \frac{n}{n} \frac{n-1}{n} \cdots \frac{n-k+1}{n}.$$

We claim that this is the maximum of the LHS of the inequality (which exists as the constraint set $\sum_{i=1}^{n} p_i = 1$ is compact, cf. the remark in Ivanov).

We prove this indirectly, by showing that if $p_i \neq p_j$ for some $i \neq j$, then this cannot be the maximum. Assume, therefore, that, possibly after permuting the indices, that $p_1 \neq p_2$. Note, further, that any permutation of the indices $i_1, \dots, i_k$ contributes the same product $p_{i_1} \cdots p_{i_k}$ to the sum, so that it is sufficient to consider a particular, say increasing, ordering of indices. More precisely

$$\sum_{i_1,\dots,i_k \text{ distinct}} p_{i_1} \cdots p_{i_k} = k! \sum_{1 \le i_1 < \dots < i_k \le n} p_{i_1} \cdots p_{i_k} .$$

Let $\varphi^* = \dfrac{p_1 + p_2}{2}$, so that $\varphi^* + \varphi^* + \sum_{i=3}^{n} p_i = 1$.

Then,

$$\sum_{1 \le i_1 < \dots < i_k \le n} p_{i_1} \cdots p_{i_k} = p_1 \sum_{2 \le i_2 < \dots < i_k \le n} p_{i_2} \cdots p_{i_k} + \sum_{2 \le i_1 < \dots < i_k \le n} p_{i_1} \cdots p_{i_k}$$

$$= p_1 p_2 \sum_{3 \le i_3 < \dots < i_k \le n} p_{i_3} \cdots p_{i_k} + p_1 \sum_{3 \le i_2 < \dots < i_k \le n} p_{i_2} \cdots p_{i_k}$$

$$+ p_2 \sum_{3 \le i_2 < \dots < i_k \le n} p_{i_2} \cdots p_{i_k} + \sum_{3 \le i_1 < \dots < i_k \le n} p_{i_1} \cdots p_{i_k}$$

$$< \varphi^* \varphi^* \sum_{3 \le i_3 < \dots < i_k \le n} p_{i_3} \cdots p_{i_k} + (\varphi^* + \varphi^*) \sum_{3 \le i_2 < \dots < i_k \le n} p_{i_2} \cdots p_{i_k} + \sum_{3 \le i_1 < \dots < i_k \le n} p_{i_1} \cdots p_{i_k}$$

$$= \sum_{1 \le i_1 < \dots < i_k \le n} p^*_{i_1} \cdots p^*_{i_k} \qquad \text{where} \quad \begin{array}{l} p^*_1 = p^*_2 = \varphi^* \\ p^*_i = p_i \ \text{ for } i = 3, \dots, n \end{array}$$

since, by the two-dimensional AM-GM inequality $p_1 p_2 < (\varphi^*)^2$ and $p_1 + p_2 = 2\varphi^*$.

Thus, the maximum cannot be attained if $p_i \ne p_j$ for some $i \ne j$.

This completes the proof. □

The connection to the birthday problem is the observation that

P(no two born on same day)

$$= \sum_{i_1=1}^{n} \underbrace{P(\text{person 1 born on day } i_1)}_{\equiv p_{i_1}} \underbrace{P(\text{no other person born on day } i_1)}$$

$$= \sum_{\substack{i_2=1 \\ i_2 \ne i_1}}^{n} P(\text{person 2 born on day } i_2) \cdot P(\text{no other person born on } i_1 \text{ or } i$$

$$= \dots = \sum_{i_1=1}^{n} p_{i_1} \sum_{\substack{i_2=1 \\ i_2 \ne i_1}}^{n} p_{i_2} \cdots \sum_{\substack{i_k=1 \\ i_k \ne i_1,\dots,i_{k-1}}}^{n} p_{i_k}$$

$$= \sum_{i_1,\dots,i_k \text{ distinct}} p_{i_1} \cdots p_{i_k} .$$