Листок 2.

Задача 1. В семье два ребенка, причем один из них мальчик. Какова вероятность того, что в семье два мальчика?

Задача 2. В коробке a белых и b черных шариков. Случайным образом извлекается шар. Этот шар возвращается обратно, затем добавляется еще c шаров одного с ним цвета.

(a) Найдите вероятность того, что при $n+m$ извлечениях появилось n белых и m черных шаров.

(b) Докажите, что вероятность извлечения белого шара на k-м шаге равна $a/(a+b)$.

Задача 3. Из множества $\{1, 2, \ldots, 100\}$ без возвращения по очереди выбирают три различных числа. Найдите условную вероятность того, что третье число лежит между первым и вторым, при условии, что первое число меньше второго.

Задача 4. Случайным образом выбираем из $\{1, 2, \ldots, n\}$ одно число. Событие A – выбранное число делится на 2, событие B – выбранное число делится на 5. (a) Выяснить, независимы ли события A и B при $n = 99, 100, 101$. (b) Найдите все n такие, что события A и B независимы.

Задача 5. Новичок играет три партии в теннис против двух противников – слабого и сильного. Он должен победить в двух партиях подряд. Порядок партий может быть следующий: слабый – сильный – слабый или сильный – слабый – сильный. Вероятность победить слабого p, вероятность победить сильного q, $q < p$. Результаты партий независимы в совокупности. Какой вариант предпочтительней для новичка и какова вероятность выиграть?

Задача 6. (a) Пусть события A и B независимы. Докажите, что события A и $\Omega \setminus B$ независимы.

(b) События A, B, C попарно независимы и равновероятны, $A \cap B \cap C = \emptyset$. Найти максимально возможное значение $P(A)$.

Задача 7. Пространство элементарных исходов Ω состоит из n элементов. Выяснить, при каких k на Ω можно определить вероятностную меру P и выделить события $A_1, A_2, \ldots, A_k$ так, что эти события независимы в совокупности и $0 < P(A_i) < 1$ для всех $i = 1, 2, \ldots, k$.

Задача 8. Пусть $q > 1/2$. Про три события A, B, C известно, что $P(A \cap B) \geq q$ и $P(A \cap C) \geq q$. Докажите, что $P(A|B \cap C) \geq q$.

*Задача 9.** (Лемма Ловаса) Пусть события $A_1, \ldots, A_n$ и число $d \in \{1, 2, \ldots, n\}$ удовлетворяют следующим двум условиям: 1) для каждого события A_k найдутся события $A_{i_1}, \ldots, A_{i_s}$ с $s \geq n-d$ такие, что A_k независимо со всеми пересечениями этих событий, 2) для каждого k верно неравенство $P(A_k) \geq 1 - \frac{1}{e(d+1)}$. Докажите, что $P(\cap_k A_k) > 0$.

Задача 10. В научном центре работают специалисты по различным направлениям естественных наук (число направлений мы не знаем). По каждому направлению в центре работает ровно 7 ученых, причем вполне может быть, что один ученый является специалистом сразу по нескольким направлениям, но не больше трех. Докажите, что всегда можно так распределить всех ученых центра по двум одновременно проходящим конференциям, что на каждой будут присутствовать специалисты по всем направлениям (и каждый будет участвовать лишь в одной конференции).

*Задача 11.** По кругу стоят 1600 студентов из 100 групп по 16 человек. Докажите, что в каждой группе можно выбрать старосту так, что рядом не окажется двух старост.

*Задача 12.** Докажите, что для всяких 25 чисел A_1, A_2, $\ldots$, A_{25} можно раскрасить все вещественные числа в три цвета так, что для каждого x среди чисел x, $x + A_1$, $\ldots$, $x + A_n$ встретятся все три цвета.

Remarks on the problem set 2

Ex 1. There are four options $\{MM, MF, FM, FF\}$, a priori equiprobable. As we know that one is male, we are left with three, *still equiprobable* possibilities $\{MM, MF, FM\}$, which implies that the answer is $1/3$. Notice that, had we known that say the eldest (or the tallest, etc) sibling is a male, then the probability that the family has to males is $1/2$. No tricky math here, just tricky language.

Ex 2. (a) This is an interesting exercise that can be generalized in several ways (e.g. several colors). First suppose that we want to find the probability of a given sequence of colors, say BBWB (B for black, W for white): this is easily calculated as $\frac{b}{a+b}\frac{b+c}{a+b+c}\frac{a}{a+b+2c}\frac{b+2c}{a+b+3c}$.

If we similarly calculate the probability of a sequence of length $n+m$, it is then clear that, no matter the colors in the sequence, we have at the denominator a product $(a+b)(a+b+c)\cdots(a+b+(n+m-1)c)$. Similarly, the numerator will only depend on how many white balls (say n) and black balls (say m) are extracted (and does not depend on the actual order of appearance of such colors), and is given by$a(a+c)\cdots(a+(n-1)c)b\cdots(b+(m-1)c)$. If one introduces the notation

$$w^{(u,v)} := \prod_{i=0}^{v-1}(w+i\,u)$$

then the probability of a given sequence with m white and n black is $a^{(c,m)}b^{(c,n)}/(a+b)^{(c,n+m)}$. Since there are $\binom{n+m}{n}$ such (ordered) sequences, the wanted probability is

$$\binom{n+m}{n}\frac{a^{(c,m)}b^{(c,n)}}{(a+b)^{(c,n+m)}}$$

(b) From the discussion above, it is clear that the probability is invariant under permutations of the order of sampled colors, so the probability to get a white at the first try should equal the probability to get a white at a later extraction. As an exercise, let's make this statement formal. Let k be the total number of extractions, and consider the probability space $S := \{B,W\}^k$. For $x \in S$ and π a k-permutation, let x^π be defined by $x^\pi_i = x_{\pi(i)}$. From the calculations in point (a), we know that $\mathbb{P}(\{x^\pi\}) = \mathbb{P}(\{x\})$. Therefore, setting E_k the event "*a white ball is extracted at the k-th extraction*", and fixing any permutation π with $\pi(k)=1$, we have

$$\begin{aligned}\mathbb{P}(E_k) &= \mathbb{P}(\{x \in S\,:\,x_k = W\}) = \mathbb{P}(\{x^\pi \in S\,:\,x^\pi_i = W\})\\ &= \mathbb{P}(\{x \in S\,:\,x^\pi_i = W\}) = \mathbb{P}(E_1) = \tfrac{a}{a+b}\end{aligned}$$

Further remarks on Ex. 2 Notice that for $c=-1$, the formulas in (a) gives the probability of the same event in the *sampling without replacement.* For $c=0$ it gives the probability in the *sampling with replacement* (independent extractions).

A remarkable case is obtained when $a=b=c>0$. In such a case, the probability of extracting k white balls in N extractions is $1/(N+1)$, namely it is independent of $k \in \{0,1,\dots,N\}$.

Ex 5. Let A_i be the event "*the player wins the i-th game*". Then $\mathbb{P}((A_1\cap A_2)\cup(A_2\cap A_3)) = \mathbb{P}(A_1\cap A_2)+\mathbb{P}(A_2\cap A_3)-\mathbb{P}(A_1\cap A_2\cap A_3) = pq+qp-pqp$ (exchange the role of p and q in the case of strong-weak-strong).

Ex 6. For point (b), let $p = \mathbb{P}(A) = \mathbb{P}(B) = \mathbb{P}(C)$. Then

$$(1-p)^2 = \mathbb{P}(\bar{B} \cap \bar{C}) \geq \mathbb{P}(\bar{A} \cap \bar{B} \cap \bar{C}) = 1 - \mathbb{P}(A \cup B \cup C) = 1 - (3p - 3p^2 + 0)$$

As a consequence, $p \leq 1/2$. It is easy to check that $p = 1/2$ is possible. E.g. $\Omega = \{1,2,3,4\}$ with uniform probability, and $A = \{1,2\}$, $B = \{1,3\}$, $C = \{2,3\}$.

Ex 7. k has to be such that $2^k \leq n$.

First we prove that this inequality is necessary for such a probability space to exist. Recall that, as each B_i runs in $\{A_i, \bar{A}_i\}$, the sets $B_1, \ldots, B_k$ are also independent. Thus $\mathbb{P}(B_1 \cap \cdots, \cap B_k) = \mathbb{P}(B_1) \cdots \mathbb{P}(B_k) \in (0,1)$. Therefore the 2^k sets of the partition of Ω given by $\{ \cap_{i=1}^k B_i,\, B_i \in \{A_i, \bar{A}_i\}\}$ are all distinct and nonempty. This yields $2^k \leq n$.

Next, we prove that the inequality is sufficient. Indeed, up to adding up some points with vanishing probability to Ω, we can assume $n = 2^k$. Then take $\Omega = \{0,1\}^k$, and $A_i = \{\omega \in \Omega : \omega_i = 0\}$ and $\mathbb{P}$ uniform.

Ex 8. There are various ways to prove this statement. For instance

$$\begin{aligned} &2\mathbb{P}(A \cap B \cap C) \geq \mathbb{P}(A \cap B) + \mathbb{P}(A \cap C) - \mathbb{P}(B \Delta C) \\ &\mathbb{P}(B \cap C) \geq 1 - \mathbb{P}(B \Delta C) \end{aligned}$$

so that (for $p = \mathbb{P}(B\Delta C)$) $\mathbb{P}(A|B \cap C) \geq \frac{2q-p}{2-2p} \geq q$, for $q \geq 1/2$.

Ex 9. We will prove the following fact:

For each $A \in \{A_1, \ldots, A_n\}$ and each subset $\mathcal{E} \subset \{A_1, \ldots, A_n\} \setminus A$, it holds

$$\mathbb{P}(\bar{A}| \cap_{B \in \mathcal{E}} B) \leq \tfrac{1}{d+1} \tag{1}$$

Once (1) is proved, the wanted statement follows from

$$\mathbb{P}(\cap_{k=1}^n A_k) = \mathbb{P}(A_1 | \cap_{k=2}^n A_k)\mathbb{P}(A_2 | \cap_{k=3}^n A_k) \cdots \mathbb{P}(A_n) \geq (1 - \tfrac{1}{d+1})^n$$

As for the proof of (1), we proceed by induction on the cardinality of $\mathcal{E}$. The base is trivial, since if $\mathcal{E}$ is empty then $\mathbb{P}(\bar{A}) \leq \frac{1}{e(d+1)} < \frac{1}{d+1}$. Let's then fix a nonempty $\mathcal{E}$, and assume that (1) holds for any $A \in \{A_1, \ldots, A_n\}$, whenever $\mathcal{E}$ is replaced with a subset of $\{A_1, \ldots, A_n\} \setminus A$ with cardinality strictly smaller than $\mathcal{E}$ itself. Let $\mathcal{F} \subset \{A_1, \ldots, A_n\} \setminus A$ be the collection of sets on which A is *not* independent; by hypothesis $|\mathcal{F}| \leq d$. Let $\mathcal{E}_1 = \mathcal{E} \cap \mathcal{F}$, $\mathcal{E}_2 = \mathcal{E} \setminus \mathcal{E}_1$. By some elementary algebra

$$\mathbb{P}(\bar{A}| \cap_{B \in \mathcal{E}} B) = \frac{\mathbb{P}(\bar{A} \cap_{B \in \mathcal{E}_1} B | \cap_{B \in \mathcal{E}_2} B)}{\mathbb{P}(\cap_{B \in \mathcal{E}_1} B | \cap_{B \in \mathcal{E}_2} B)} \tag{2}$$

Now the numerator is bounded as

$$\mathbb{P}(\bar{A} \cap_{B \in \mathcal{E}_1} B | \cap_{B \in \mathcal{E}_2} B) \leq \mathbb{P}(\bar{A} | \cap_{B \in \mathcal{E}_2} B) = \mathbb{P}(\bar{A}) \leq 1/(e(d+1)) \tag{3}$$

where we used that A is independent of the B's in $\mathcal{E}_2$.

As for the denominator in (2), write $\mathcal{E}_2 = \{B_1, \ldots, B_q\}$ and

$$\begin{aligned} \mathbb{P}(\cap_{B \in \mathcal{E}_1} B | \cap_{B \in \mathcal{E}_2} B) &= \mathbb{P}(B_1 | \cap_{j=2}^q B_j \cap_{B \in \mathcal{E}_2} B)\mathbb{P}(B_2 | \cap_{j=3}^q B_j \cap_{B \in \mathcal{E}_2} B) \cdots \mathbb{P}(B_n | \cap_{B \in \mathcal{E}_2} B) \\ &\geq (1 - \tfrac{1}{d+1})^{|\mathcal{E}_1|} \geq (1 - \tfrac{1}{d+1})^d \geq \tfrac{1}{e} \end{aligned} \tag{4}$$

From (2), (3) and (4) we obtain (1).

Further remarks on Ex. 9 There is a sharper version of this statement: if we let $p := \sup_k \mathbb{P}(\bar{A}_k)$ and assume

$$p < \begin{cases} 1 & \text{if } d = 0 \\ 1/2 & \text{if } d = 1 \\ \frac{(d-1)^{d-1}}{d^d} & \text{if } d \geq 2 \end{cases}$$

then $\mathbb{P}(\cap_{k=1}^n A_k) > 0$.

Ex 10. We assign each scientist to one of the two conferences randomly, with equal probability and independently one of the other. Say that there are n subjects of expertise, and for $i = 1, \dots, n$ let A_i be the event "*there is an expert in the i-th subject in both conferences*". We want to prove that $\mathbb{P}(\cap_{i=1}^n A_i) > 0$. We have $\mathbb{P}(\bar{A}_i) = 2 * 2^{-7} = 2^{-6}$. On the other hand, A_i only depend on the assignment of the 7 scientists that are expert in i, and those scientists are (all together) at most expert in other $2 * 7 = 14$ subjects. We can then apply the statement in Ex. 9 with $d = 14$ since $e\, 2^{-6}(14 + 1) < 1$.

Using the sharper statement at the end of Ex. 9, it is actually enough to know that for each subject there are (at least) 6 experts.

Ex. 11 We chose randomly the head of each team, giving the same probability 1/16 to each member of each team. We then let A_i be the event "*the students at positions $i, i+1$ are not both heads of their groups*". We want to apply Ex. 9 to show that $\mathbb{P}(\cap_{i=1}^{1600} A_i) > 0$. If the students at $i, i+1$ are in the same team, then $\mathbb{P}(\bar{A}_i) = 0$; otherwise $\mathbb{P}(\bar{A}_i) = 1/16^2$. In any case $\mathbb{P}(\bar{A}_i) \leq 2^{-8}$. On the other hand, A_i is determined only by the choice of the heads in the teams of the students at positions $i, i+1$, and is independent on the choice of heads of the remaining 98 (or 99) teams. Since those teams are relevant for at most other $(2*(16-1)+1)*2 = 62$ adjacent positions $j, j+1$, we can take $d = 62$ in Ex. 9. And since $e * 2^{-8} * 63 < 1$ we can apply the result in that exercise.

Notice that the same calculation holds if we had N teams with $k \geq 11$ students in each.