ДЕВЕТНАДЕСЕТА УЧЕНИЧЕСКА КОНФЕРЕНЦИЯ НА УЧЕНИЧЕСКИЯ ИНСТИТУТ ПО МАТЕМАТИКА И ИНФОРМАТИКА

УС'19

ПРИЛОЖЕНИЯ НА ВЕРОЯТНОСТНИЯ МЕТОД

Автор:

Анна Михалкова

(СМГ „Паисий Хилендарски“, град София)

Научен ръководител:

Димитър Чакъров

(МГ „Акад.К.Попов“, град Пловдив)

Резюме

Разработеният реферат е в областта на комбинаториката и съставлява едно широко обобщение на вероятностния метод. Въведени са основната терминология и теоритичния апарат, нужен за прилагане на метода. Към тях са предоставени примерни задачи с директно приложение. Главният фокус е около състезателни задачи от национални и международни форуми, в чиито решения метода може да се използва и допринася за олекотяване на решенията. Една от основните ни цели е показване широката приложимост на вероятностния метод не само в теоритични клонове на математиката, но и в задачи със състезателен характер. Разгледаните задачи са с повишена трудност и смятаме, че биха били от полза на ученици при подготовката им за олимпиади и състезания по математика.

Summary

The following project is in the field of combinatorics. It is a broad summary of the probabilistic method. Basic definitions and theorems, needed for the proper use of the probabilistic method, are introduced. Simple examples are provided in order to illustrate the theory in practice and to introduce the method itself. The project focuses on competitive problems from major national and international competitions, in whose solutions could be used probabilistic method, contributing to their briefness. One of our main aims is demonstrating the broad application of the probabilistic method not only in some theoretical fields of mathematics, but also in competitive problems. The problems in the abstract are with increased difficulty and we believe, they can serve as a preparation for mathematical olympiads and competititons.

1 Въведение

През последните години вероятностният метод се развива интензивно и става едно от най-използваните и широкоразпространени средства за решаване на комбинаторни проблеми. Една от главните причини за голямото развитие в областта е съществената роля на теорията на вероятности в компютърните науки - област, източник на множество интригуващи комбинаторни задачи.

Началото на метода поставя Пол Ердьош (един от най-плодовитите математици на 20. век) в края на четиридесетте години на миналия век . Не е изненада и че той има значителен принос за развитието на областта и вида, в който я познаваме.

Основната идея на вероятностния метод е да докаже съществуването на математическа структура с определени свойства, от които ние се интересуваме. Разглеждаме произволно състояние на конструкцията с дадени параметри и задаваме вероятност на елементарните събития, изграждащи я. Използвайки това доказваме, че желаните свойства съществуват с положителна вероятност, тоест има състояние, в което търсеното от нас е изпълнено. Истинността на получените чрез метода резултати се предполага от формалната дефиниция за вероятност. Ето защо и той намира широко приложение в теория на числата, алгебра, анализ, теория на информация, комбинаторика, а и косвено в области като алгоритмичното програмиране.

В следващата секция на този проект са въведени важни факти, дефиниции и използваната терминология. В третата секция са дадени задачи, за които е нужно единствено директното приложение на основните понятия. Четвъртата секция представя някои от необходимите теореми за прилагане на вероятностния метод. Петата секция предлага още задачи, които са от значително по-голяма трудност и изискват немалка съобразителност по отношение на мястото на метода в цялостното решение. Последните две секции са съответно заключение, бъдещо развитие и благодарности.

2 Дефиниции и означения

В тази секция излагаме важни дефиниции и нотацията, която ще използваме в реферата.

Дефиниция 1. *Случайна величина*

Случайна величина е такава променлива, която мени стойността си произволно. (Например, ако се разгледат възможните точки при хвърляне на зарче със 6 страни (D_6), събитието да бъдат получени определен брой точки е случайна величина.)

Дефиниция 2. *Вероятност*

Вероятността да се случи някакво елементарно събитие (или още отношението на броя на благоприятните изходи към общия им брой) ще бележим с P.

Пример:

В случая със зарчето

$P(D_6 = 1) = P(D_6 = 2) = P(D_6 = 3) = P(D_6 = 4) = P(D_6 = 5) = P(D_6 = 6) = \dfrac{1}{6}$;

$P(D_6 = 0) = 0$;

$P(D_6 \geq 4) = \dfrac{1}{2}$.

Дефиниция 3. *Математическо очакване*

Математическо очакване представлява характеристична стойност на вероятностното разпределение на случайна величина. Математическо очакване за случайна величина X може да се интерпретира като нейна средна стойност, въпреки че тази стойност може да не бъде възможен неин изход. Математическото очакване не трябва да се бърка с най-вероятния изход!

Формалната дефиниция е

$$E[X] = \sum_x P(X = x) \cdot x$$

Но нека се върнем на примера със зарчето:

$$E[D_6] = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = 3.5$$

Накратко математическото очакване е сумата от възможните изходи, умножени съответно с вероятността да се случат.

Теорема 1 (Линейност на математическото очакване)**.** *При дадени случайни величини* $X_1, X_2, \dots, X_n$ *винаги важи*

$$E[X_1 + X_2 + \cdots + X_n] = E[X_1] + E[X_2] + \cdots + E[X_n]$$

Теоремата е очевидна, ако приемем, че сбитията са независими едно от другоако хвърлим зарче 100 *пъти очакваме да получим средно* 350 *точки. По силното твърдение е, че теоремата важи и за събития, които зависят едно от друго. В много от случаите* $E[X]$ *може да се пресметне при разбиване на главното събитие на елементарни такива и прилагане на теоремата за линейност.*

3 Задачи за загрявка

В тази секция даваме лесноусвоими примери за директно приложение на вероятностния метод.

Задача 1 (AHSME 1989)**.** 7 момичета и 13 момчета са се наредили в редица. Нека означим с S колко пъти момче и момиче в редицата стоят едно до друго.(Например в редицата 12211121211121211211, където на мястото на 1 има момиче, а на 2-момче, $S = 12$). Намерете математическото очакване за стойността на S.

Решение. Трябва да получим колко общо измежду деветнадесетте двойки са с момче и момиче – това е същото,като да разгледаме всяка двойка поотделно и да съберем стойностите. Ще означаваме двойка момче-момче или двойка момиче-момиче с 0, а двойка момче-момиче или момиче-момче с 1. С S означаваме общия брой на двойките момче-момиче и момиче-момче, а $S_i (i = 1, 2, \dots, 19)$ приема стойност 0 или 1 в съвпадаща с тази на вида на двойката. Знаем, че:

$$S = S_1 + S_2 + \cdots + S_{19}$$

Всичко това ни навежда на мисълта да приложим теоремата за линейност на математическото очакване. Преди това обаче трябва да изчислим вероятността една двойка да е момче-момиче или момиче-момче:

$$P = \frac{7 \cdot 13}{\frac{20 \cdot 19}{2}} = \frac{7 \cdot 13}{10 \cdot 19} = \frac{91}{190}$$

$$E[S] = E[S_1] + E[S_2] + \cdots + E[S_{19}] = 19 \cdot \frac{91}{190} = 9.1$$

Задача 2 (NIMO 5.6)**.** Том има научен калкулатор. За нещастие, всички копчета са счупени освен 1,2,3,$+$ и $-$. Той въвежда поредица от пет знака, като всеки от тях има еднаква вероятност да бъде избран. След това калкулаторът пресмята израза и извежда резултат . Намерете математическото очакване за стойността на .

Решение. Ако заменим плюс с минус и обратното, средната стойност на всички числа, с които сме извършили действието, остава числото,образувано от цифрите преди поява на знак. Затова можем да означим плюсовете и минусите със сигнал СПРИ. За всеки израз се зачита стойността му до сигнал СПРИ. Ако разгледаме един израз и заменим 1 с 3 и обратно, средната стойност на първоначалния израз и новополучения се записва само с двойки на мястото на цифрите 1, 2 и 3. Следователно можем да заменим всяка цифра с 2. Извършваме определената размяна.

Вероятността един знак да е 2 е $\frac{3}{5}$. Означаваме $p = \frac{3}{5}$. Вероятността първият знак да бъде 2 е p, а за всеки следващ става $10^k p$ за $k = 1, 2, 3, 4$. Прилагаме теоремата за линейността и получаваме израза:

$$2(p + 10p^2 + 10^2 p^3 + 10^3 p^4 + 10^4 p^5) = 2 \cdot p \cdot \frac{(10p)^5 - 1}{10p - 1} = 1866$$

Задача 3 (AIME 2006/6)**.** Нека S е множеството от реалните числа, които могат да бъдат представени като десетична дроб $0.abc$, където a, b и c са различни цифри. Намерете сумата на елементите на S.

Решение. В числото abc a,b,c за различни, следователно броят на елементите на S е $10 \cdot 9 \cdot 8 = 720$. Вероятността едно число да принадлежи на S е $\frac{1}{2}$(всяко от числата или удовлетворява условието, или не). Чрез теоремата за линейността, сумата на дробите става $720 \cdot \frac{1}{2} = 360$.

Задача 4 (NIMO 4.3)**.** Един ден кон и офицер са на квадратчета на един и същи ред от безкрайна шахматна дъска. Когато изведнъж започва голяма метеоритна буря, пада метеор във всяко квадратче независимо и случайно с вероятност p. Нито конят, нито офицерът са оцелени, но тяхното движение може да бъде възпрепятствано. За каква стойност на p очакваният брой на квадратчетата, до които може да стигне коня за един ход, е равен на очаквания брой квадратчета, до които може да стигне офицерът за един ход?

Решение. Всеки кон може да стигне до 8 полета с един ход, а вероятността едно поле да е без метеорит е $1-p$. Един офицер може да стигне до 4 полета на разстояние n по диагонал от неговото поле(n варира от 1 до безкрайност). За да може да стигне до поле на разстояние n, трябва и всички полета преди избраното, и самото то да не съдържа метеорит. Следователно вероятността да може да стигне до поле на разстояние n от сегашното е $(1-p)^n$.

За да намерим стойността, искана в условието, изравняваме двете стойности за математическото очакване за достъпните полета, използвайки теоремата за линейността. Така достигаме до уравнението:

$$8(1-p) = 4((1-p) + (1-p)^2 + (1-p)^3 + \dots)$$
$$2(1-p) = ((1-p) + (1-p)^2 + (1-p)^3 + \dots)$$

$p = 1$ е решение както на уравнението, така и на задачата. Ако p е различно от 1:

$$1 = ((1-p) + (1-p)^2 + \dots)$$
$$1 = \frac{(1-p)(1-(1-p)^n)}{p}$$

При нарастване на n до безкрайност $(1-p)^n$ клони към 0, а съответно $1-(1-p)^n$ към 1. Оттук

$$lim_{n\to\infty}((1-p) + (1-p)^2 + (1-p)^3 + \dots) = \frac{1-p}{p}$$
$$1 = \frac{1-p}{p}$$

Следователно окончателно p е равно на 1 или $\frac{1}{2}$.

4 Полезна теория

В тази секция предоставяме на читателя няколко теореми, които могат да бъдат използвани в по-сложни многостъпкови задачи от състезания и олимпиади.

Лема 1. *За положителни n и k*

$$\binom{n}{k} = \frac{1}{e}\left(\frac{en}{k}\right)^k,$$

където $e \approx 2.718\ldots$ е Неперовото число.

Теорема 2 (Неравенство на Джордж Бул)**.** *Нека $A_1, A_2, \ldots, A_k$ са елементарни събития. Ако*

$$P(A_1) + P(A_2) + \cdots + P(A_k) < 1,$$

то има положителна вероятност никое от събитията да не се случи.

Теорема 3 (Неравенство на Марков)**.** *Нека X е произволна променлива, която приема само неотрицателни стойности, а $E[X] = c$. Следователно*

$$P(X \geq rc) \leq \frac{1}{r}.$$

Теорема 4 (Лема на László Lovász)**.** *Съществуват няколко елементарни събития, чиято вероятност е не по-голяма от p, такива, че всяко от тях е зависимо от най-много d от останалите събития. Тогава ако*

$$epd \leq 1,$$

вероятността никое от събитията да не се случи е положителна.

5 Състезателни задачи

Задача 5 (Russia 1996)**.** В парламента има 1600 народни представители, които формират 16000 комитети с по 80 човека всеки. Докажете, че съществуват два комитета, които имат поне 4 общи члена.

Решение. Означаваме с X броя на народните представители, които са членове на два комитета (произволни), а с $X_1, X_2, \ldots, X_{1600}$ дали всеки представител поотделно е член и на двата комитета. X_i приема стойност 1, ако човекът е член, и стойност 0, ако не е. Забелязваме, че

$$X = X_1 + X_2 + \cdots + X_{1600}$$

От това следва, че можем да приложим теорема за линейността:

$$E[X] = \sum_{i=1}^{i=1600} X_i 5$$

С n_i означаваме броя на комитетите, в които i-тият представител членува. От условието намираме, че $\sum_{i=1}^{i=1600} n_i = 80.16000 = 1280000$, а средната стойност на $n_i = \dfrac{16000.80}{1600} = 800$. Също така знаем, че

$$E[X_i] = \frac{\binom{n_i}{2}}{\binom{16000}{2}}.$$

От горните факти можем да направим следната оценка

$$E[X] \geq 1600 \cdot \binom{800}{2} / \binom{16000}{2} = 1600.\frac{800 \cdot 799}{16000 \cdot 15999} = 3.995$$

и тъй като $E[X]$ по условие трябва да бъде цяло, доказателството е извършено.

Задача 6. Покажете, че може да се организира кръгов турнир (всеки играе с всеки по веднъж) с повече от 1000 човека, така че във всяка група от 1000 човека може да се намери такъв, който е победил всички от групата си.

Решение. Нека n е общият брой на участниците, а една група се състои от k участника (в случая $k = 1000$). Тогава $\binom{n}{k}$ е броят на всички такива групи. Вероятността да няма участник, който побеждава всички останали е $(1 - \dfrac{1}{2^k})$. Прилагайки Теорема 4, за да съществува турнир с исканите свойства, трябва да е изпълнено:

$$\binom{n}{k}\left(1 - \frac{1}{2^k}\right)^{n-k} < 1$$

Пресмятаме за $k = 1000$ и използвайки лемата, достигаме до:

$$\binom{n}{1000}\left(\frac{2^{1000} - 1}{2^{1000}}\right)^{n-1000} < 1$$

$$\binom{n}{1000}\Big(\frac{2^{1000}-1}{2^{1000}}\Big)^{n-1000} < \frac{1}{e}\Big(\frac{en}{1000}\Big)^{1000}\Big(\frac{2^{1000}-1}{2^{1000}}\Big) < 1,$$

което е вярно за достатъчно голямо n.

Задача 7 (BAMO 2004)**.** Дадени са n реални числа, не всички равни на нула, но със сума равна на нула. Докажете, че числата могат да бъдат означени $a_1, a_2, \dots a_n$, така че

$$a_1a_2 + a_2a_3 + \cdots + a_na_1 < 0.$$

Решение. Да допуснем, че всички суми са по-големи или равни на нула. Произведението a_ia_j (i е различно от j) се появява $n(n-2)!$ пъти (тук стойностите на a_i и a_j не се променят-фиксирали сме ги).

$$n(n-2)!\sum a_ia_j \geq 0$$

Следователно

$$2\sum a_ia_j \geq 0$$

и

$$\sum_{i=1}^{i=n} a_i^2 + 2\sum a_ia_j > 0$$

От друга страна горепосоченият израз е точно равен на сумата на елементите на квадрат, което е 0. Така достигаме до противоречие с допуснатото и твърдението е доказано.

Задача 8 (Russia 1999)**.** В училище всяко момче харесва поне едно момиче. Докажете, че съществува подмножество S от поне половината ученици, така че всяко момче харесва нечетен брой момичета от S.

Решение. Ще покажем как такова множество може да бъде построено. Първо ще разпределим момичетата – за всяко от тях има вероятност да бъде избрано $\frac{1}{2}$ (или ще бъде, или няма да бъде в подмножеството). Така избрани броят на харесваните от определено момче момичета в групата вече може да се намери – ако е нечетен, включваме момчето в подмножеството, а ако е четен – не. Следователно вероятността да бъде избрано едно момче също е $\frac{1}{2}$. Така от теоремата за линейността получаваме, че полученото множество се състои от поне половината ученици, с което доказателството е завършено.

Задача 9 (IMC 2002)**.** На олимпиада иа 6 задачи и 200 участника, които са много умни и затова всяка задача е решена от поне 120 от тях. Докажете, че съществуват двама участника, за които всяка задача е решена от поне един от двамата.

Решение. Съществуват $\binom{200}{2} = 19900$ различни двойки участници. Всяка задача не е решена от максимум $\binom{80}{2} = 3160$ двойки участници. Следователно вероятността да не бъде решена в определена двойка е $\frac{3160}{19900}$. Изчисляваме математическото очакване за броя на двойките, в които има нерешена задача. Означаваме със Z броя на нерешените задачи, а $Z_1, Z_2, \ldots, Z_6$ са променливи, които показват дали всяка задача е решена или не от поне един от двойката (приемат стойност 1 при нерешена задача и 0 при решена). Знаем, че:

$$Z = Z_1 + Z_2 + \cdots + Z_6$$

$$E[Z] = E[Z_1] + \cdots + E[Z_6] = 6 \cdot \frac{3160}{19900} = \frac{18960}{19900} < 1$$

Тъй като математическото очакване за нерешените задачи е нецяло число по-малко от едно, а в действителност броят на задачите е цяло число, то със сигурност съществува двойка, която разполага с 0 нерешени задачи, тоест всяка задача е решена от поне един от двамата.

Задача 10 (Shortlist 1999 C4)**.** Нека е множество от N остатъка по модул N^2. Покажете, че съществува множество B от N остатъка по модул N^2, така че $A + B = \{a + b | a \in A, b \in B\}$ съдържа поне половината остатъци по модул N^2.

Решение. Означаваме с X броя на всички възможни остатъци, които $A+B$ съдържа. Нека остатъкът $i \in \{A+B\}$. Тъй като $|A| = n$, съществуват n на брой остатъка $b \in B$, които удовлетворяват $A+b \equiv i(\mod N^2)$. Вероятността остатъкът i да принадлежи на $A+B$ е $1-\left(1-\frac{n}{n^2}\right)^n$ От всички възможни изходи изваждаме неблагоприятните изходи-всеки от възможните n остатъци не принадлежи на множеството B. Така определяме и математическото очакване за броя на остатъците:

$$E[X]=n^2\cdot\left(1-\left(1-\frac{n}{n^2}\right)^n\right)$$

За да се съдържат поне половината остатъци, трябва единствено да докажем:

$$1-\left(1-\frac{n}{n^2}\right)^n\geq\frac{1}{2}$$

$$\frac{1}{2}\geq\left(1-\frac{n}{n^2}\right)^n.$$

Последното неравенство е тривиално и може да бъде доказано по индукция.

Задача 11 (Iran TST 2008/6)**.** Нека 799 отбора играят в турнир всеки срещу всеки по веднъж. Докажете, че съществуват две групи A и B, всяка с по 7 отбора, които нямат общ отбор, така че всички отбори от едната група са победили всички от другата.

Решение. Означаваме с произволна група от 7 отбора. Нека X е броят на отборите, които са победени и от седемте отбора в група . Построяваме насочен граф с върхове отборите и ребра изиграните мачове. Ребрата са насочени от победител към победен. С d_v означаваме ребрата, които завършват във връх v. От условието знаем, че $\sum_v d_v = \binom{799}{2}$, от което заключаваме, че средната стойност на d_v е 399. Изчисляваме $E[X]$:

$$E[X]=\sum_v\binom{d_v}{7}/\binom{799}{7}\geq 799\cdot\binom{399}{7}/\binom{799}{7}\approx 800\cdot\left(\frac{1}{2}\right)^7\approx 6,25$$

Тъй като математическото очакване за победените отбори е нецяло число, а в действителност те са цяло число, оценката е достатъчна и със сигурност съществуват 7 победени отбора от отборите в група (ще ги наречем група B).

Задача 12 (Romania 2004)**.** Докажете, че за които и да било комплексни числа $z_1, z_2, \ldots, z_n$, удовлетворяващи условието

$$|z_1|^2+|z_2|^2+\cdots+|z_n|^2=1,$$

могат да се изберат $\epsilon_1, \epsilon_2, \ldots, \epsilon_n \in \{-1, 1\}$, така че

$$\left| \sum_{k=1}^{k=n} \epsilon_k z_k \right| \leq 1.$$

Решение. Избираме ϵ_i произволно, а лявата страна на исканото означаваме с LHS. За всяко z важи $|z|^2 = z\overline{z}$. Знаем, че:

$$LHS^2 = \sum_k |z_k|^2 + 2 \sum_{i<j} \epsilon_i \epsilon_j (z_i \overline{z_j} + z_j \overline{z_i})$$

Избираме ϵ_i и ϵ_j произволно и равнопоставено, значи в два от случаите произведението им е -1, а в другите два е 1, което ни води до $E(\epsilon_i \epsilon_j) = 0$. Следователно след като съберем почленно всички леви страни, сумите с множител 2 пред тях се съкращават (има равен брой с минус и с плюс пред тях). Получаваме:

$$E[LHS^2] = \sum_k |z_k|^2 = 1.$$

Това означава, че понякога можем да изберем ϵ_i, така че $LHS^2 \leq 1$, към което се целим.

Задача 13 (USAMO 2012/6)**.** Нека $x_1, x_2, \ldots, x_n$ са реални числа за $n \geq 2$ (n е цяло), удовлетворяващи равенствата $x_1 + x_2 + \cdots + x_n = 0$ и $x_1^2 + x_2^2 + \cdots + x_n^2 = 1$. За всяко подмножество $A \subseteq \{1, 2, \ldots, n\}$ дефинираме

$$S_A = \sum_{i \in A} x_i.$$

(Ако A е празно множество, $S_A = 0$). Докажете, че за положително число λ, броят на множествата A, за които $S_A \geq \lambda$, е най-много $2^{n-3}/\lambda^2$. Кога се достига равенство?

Решение. Знаем, че

$$2 \sum_{1 \leq i \leq j \leq n} x_i x_j = \left(\sum_{i=1}^{n} x_i \right)^2 - \sum_{i=1}^{n} {x_i}^2 = -1$$

Следователно

$$\sum_{1 \leq i \leq j \leq n} x_i x_j = -\frac{1}{2}$$

Оттук заключаваме, че

$$\sum_{A\subseteq N} S_A^2 = 2^{n-1}\sum_{i=1}^{n} {x_i}^2 + 2^{n-1}\sum_{i=1}^{n} x_i x_j = 2^{n-1} - 2^{n-2} = 2^{n-2}$$

$S_A^2 \geq 0$. Забелязваме, че $S_A + S_{N|A} = 0$ и оттук най-много половината от $|S_A| \geq \lambda$.

От гореспоменатото директно следва(от неравенство на Марков), че максималният брой на търсените множества е $2^{n-3}/\lambda^2$. За да се достигне максималната стойност, трябва да се достига равенство и при $|S_A| \geq \lambda$. Следователно $S_A \in \{-\lambda, 0, \lambda\}$. Измежду множествата може да съществува най-много едно със стойност λ и едно със стойност $-\lambda$(в противен случай ще бъде нарушено условието за сбора). Така за останалите $n-2$ множества остава стойност 0 и за тях важи следното:

$$\sum_{i=1}^{n} x_i^2 = 2\lambda^2 = 1$$

От условието λ е положително число и заради това $\lambda = \dfrac{1}{\sqrt{2}}$.

Задача 14 (Online Math Olimpiad, Ray Li)**.** Кевин има $2^n - 1$ бисквитки, всяка от които се отличава от другите чрез различно непразно подмножество от $\{1, \ldots, n\}$. Всеки ден той избира на случаен принцип една от все още неизядените бисквитки и изяжда нея и всички други, които са с подмножество на нейното множество. Пресметнете математическото очакване за броя дни, през които Марк е ял бисквитки, докато накрая всички свършат.

Решение. Броят на дните е равен на броя на случаите, в които една бисквитка е елиминирана или избрана. Нека C е множеството на бисквитките, които са избрани по време на процеса, A – множеството, характеризиращо определена бисквитка, D – броят на дните, а $S = \{1, \ldots, n\}$. X определяме като индикатор(0 или 1) за това дали $A \in C$. Ясно е, че

$$E[D] = E[|C|] = \sum_{A\subseteq S} E[X(A \in C)] = \sum_{A\subseteq S} P(A \in C)$$

За всяко $A \in C$, е елиминирана в момента, когато надмножество A' на A е избрано. Следователно

$$P(A \in C) = P(A = A') = \frac{1}{K} = \frac{1}{2^{n-|A|}}$$

K-брой на надмножествата на A

Заместваме в първото равенство:

$$\sum_{A\subseteq S} P(A\in C) = \sum_{A\subseteq S} 2^{|A|-n} = \sum_{1\leq k\leq n} 2^k \cdot 2^{k-n} = \frac{4^{n+1}-4}{3\cdot 2^n}$$

Задача 15. Нека n е положително цяло число. С a_k означаваме броя на пермутациите на n елемента, от които k са фиксирани. Пресметнете:

$$a_1 + 4a_2 + 9a_3 + \cdots + n^2 a_n$$

Решение. За произволна пермутация X е броят на фиксираните елементи, а търсеният израз е еквивалентен на $n!\, E[X^2]$.

$$E[X^2] = 1\cdot \frac{a_1}{n!} + 2^2 \cdot \frac{a_2}{n!} + \cdots + n^2 \cdot \frac{a_n}{n!} = \frac{a_1 + 4a_2 + \cdots + n^2 a_n}{n!}$$

Лесно се вижда, че

$$X^2 = X + 2\cdot \binom{X}{2}$$

$$E[X^2] = E[X] + E\left[2\cdot\binom{X}{2}\right] = E[X] + 2\cdot E\left[\binom{X}{2}\right]$$

Ще изчислим $E[X]$ и $E[\binom{X}{2}]$:

$$E[X] = \frac{1}{n!}\cdot n\cdot (n-1)! = 1$$

$$E\left[\binom{X}{2}\right] = \sum_{ij} P(i,j = const) = \binom{n}{2}\cdot\frac{1}{n}\cdot\frac{1}{n-1} = \frac{1}{2}$$

$P(i,j = const)$-вероятността да съществуват две фиксирани числа i и j сред всички n Следователно

$$E[X] + E[2\cdot\binom{X}{2}] = n!\, E[X^2] = n!\cdot(1 + 2\cdot\frac{1}{2}) = 2n!$$

Задача 16 (Ердьош)**.** Докажете, че във всяко множество S от n различни положителни цели числа съществува подмножество T от $\left[\frac{1}{3}\right] n$ или повече елемента със свойството $a + b \neq c$ за всеки $a, b, c \in T$ (не непременно различни).

Решение. Нека $p = 3k+2$ е просто, което удовлетворява неравенството $p > 2.\max_{1\leq i\leq n} |b_i|$, а C е такова множество: $C = \{k+1, k+2, \ldots, 2k+1\}$. В C няма два елемента със сбор, който принадлежи на множеството. Забелязваме, че $\dfrac{|C|}{p-1} = \dfrac{k+1}{3k+1} > \dfrac{1}{3}$. Избираме произволно число x, такова че $1 \leq x < p$ и дефинираме $d_1, d_2, \dots, d_n$, така че $d_i \equiv xb_i(\mod p)$, $0 \leq d_i < p$. За всяко фиксирано $i(1 \leq i \leq n)$ x варира измежду $1, 2, \dots, p-1$, а d_i варира измежду ненулевите елементи на множеството S. Изчисляваме вероятността d_i да принадлежи на C: $P(d_i \in C) = \dfrac{|C|}{p-1} > \dfrac{1}{3}$. Следователно математическото очакване за броя на елементите b_i, за които $d_i \in C$, е по-голямо от $n/3$. Това означава, че съществува x, $1 \leq x < p$ и подредица , принадлежаща на B с дължина $|A| > \dfrac{n}{3}$, така че $xa(\mod p) \in C$, за всяко $a \in C$. В със сигурност съществуват някакви a_1, a_2, a_3: $a_1 + a_2 = a_3$ и $xa_1 + xa_2 \equiv xa_3(\mod p)$. От последното равенство следва, че в C има два елемента със сбор, който принадлежи на множеството-противоречие с твърдението в началото на решението.

Задача 17 (Sperner)**.** Дадено е множеството A от N различни подмножества $S_1, S_2, \dots, S_N$ от $\{1, 2, \dots, n\}$, така че никое S_i не е подмножество на S_j. Докажете, че

$$N \leq \binom{n}{\left[\frac{1}{2}n\right]}$$

Решение. Означаваме b като произволна пермутация, а X дефинираме като $\{i : b(1), \dots, b(i) \in A\}$. Разглеждаме математическото очакване за X. От дефиницията на A по условие променливата X е не по-голяма от 1, а събитията $\{b(1), \dots, b(k)\}$ са независими за всяко k. Нека N_k е броят на подмножествата с k елемента в .

$$E[X] = \sum_{k=1}^{n} P\left[\{b(1), \dots, b(k) \in A\}\right] = \sum_{k=1}^{n} \frac{N_k}{\binom{n}{k}} \geq \sum_{k=1}^{n} \frac{N_k}{\binom{n}{[n/2]}}$$

Тъй като $E[X] \leq 1$, то $\sum_{k=1}^{k=n} N_k = N \leq \binom{n}{[n/2]}$, което трябваше да докажем.

Задача 18. $11n$ точки са разположени по окръжност и са оцветени в n различни цвята, като има по 11 точки от всеки цвят. Докажете, че могат да се изберат n точки по една от всеки цвят, така че никои две не са съседни.

Решение. Означаваме точките по окръжността с $a_1, \dots, a_{11n}$. Вземаме по една точка от всеки цвяти конструираме произволно множество B от n точки. За всяко $1 \leq i \leq$

$11n$ дефинираме събитието X_i, когато a_i и a_{i+1} принадлежат на B. $X = X_1 \cup X_2 \cup \ldots \cup X_{11n}$ е обединението на всички възможни неблагоприятни изходи, които искаме да избегнем. Да отбележим, че $P(X_i) = \frac{1}{11} \cdot \frac{1}{11} = \frac{1}{121}$, ако a_i и a_{i+1} са разноцветни, а $P(X_i) = 0$ при едноцветни точки. Прилагаме Теорема 3:

$$P(X) \leq \sum_{i=1}^{11n} P(X_i) < \frac{11n}{11^2} = \frac{n}{11}$$

Теоремата обаче не ни дава $P(X) < 1$ при $n > 11$. Забелязваме, че X_i е независимо от X_j за всички j, освен ако a_i или a_{i+1} не е в същия цвят като a_j или a_{j+1}. 10 други точки освен a_i са оцветени в същия цвят и всяка от тях е част от две последователни двойки от точки. Следователно има $2 \cdot 10 + 1 = 21$ двойки (a_j, a_{j+1}), различни от(a_i, a_{i+1}), които имат един и същ цвят с a_i. Аналогично има 21 двойки, които имат общ цвят с a_{i+1}. Така X_i и X_j са независими за всички без най-много 42 стойности на j. Време е да използваме Теорема 4 с вероятност $p = \frac{1}{11^2}$ и $d = 42$:

$$ep(d+1) = e \cdot \frac{1}{11^2} \cdot 43 < \frac{28}{10} \cdot \frac{43}{121} < 1$$

Резултатът след прилагането на теоремата е, че вероятността никой от неблагоприятните изходи да не възникне е положителна. Това означава, че със сигурност има множество от n несъседни две по две точки, които са с различни цветове.

Задача 19 (Алон и Спенсър, глава 1)**.** Да се докаже, че за всеки две независими случайни реални величини X и Y

$$P(|X - Y| \leq 2) \leq 3P(|X - Y| \leq 1)$$

Решение. Нека $\{x_i\}_{i \leq n}$ са n на брой реални числа, а с $\#$ означаваме брой. Тогава първо ще докажем

$$\#\{(i,j) : |x_i - x_j| \leq 2\} \leq 3\#\{(i,j) : |x_i - x_j| \leq 1\}$$

Ще докажем твърдението по индукция. Базата е тривиална. Допускаме, че твърдението е вярно за всяко множество с по-малко или равно на n точки. Прекарваме синьо ребро между две точки, ако те са на разстояние 1 или по-малко, и с червено ребро, ако те са на разстояние между 2 включително и 1. За целите на броенето ще свържем x_i със себе си с половин синьо ребро, тъй като (i, i) се брои като една двойка, а двойката (i, j) се брои и като (j, i). Целта е да покажем, че червените ребра са най-много два пъти повече от сините ребра.

Нека съществуват $n + 1$ реални числа $\{x_i\}$ и нека $y_1 < y_2 < \cdots < y_k$ е максимално дългото подмножество, за което всяка двойка (y_i, y_{i+1}) е свързана със червено ребро. Ние твърдим, че при премахване на цялата редица със съответстващите и ребра ще сме премахнали най-много два пъти повече червени от сини ребра. Това е вярно поради следното:

Съществуват $k - 1$ червени ребра между (y_i, y_{i+1}) и съответно $k/2$ сини ребра между (y_i, y_i). За всяко червено ребро от x_i до y_j разглеждаме два случая: $x_i < y_j$ и $x_i > y_j$. Ако $x_i < y_j$, тъй като редицата е максимална, $j > 1$, трябва $|x_i - y_{j-1}| \leq 1$ и затова съпоставяме червено ребро (x_i, y_j) на синьо ребро (x_i, y_{j-1}). Аналогично ако $x_i > y_j$, следователно $j < k$ и червеното ребро (y_j, x_i) съответства на синьо ребро (x_i, y_{j+1}). Всяко премахнато синьо ребро може да бъде съпоставено на най-много едно червено ребро, идващо отдясно, и на най-много едно червено ребро, идващо отляво. Така се оказва, че броят на сините ребра, принадлежащи на нашата редица $y_1, y_2, \ldots, y_k$, е най-много половината от броят на червените ребра. И така сме готови с нашето доказателство по индукция.

Нека $X_1, X_2, \ldots, X_n$ са n независими променливи от условието. Съчетавайки горното твърдение и теоремата за линейността, получаваме, че

$$E[\#\{(i, j) : |X_i - X_j| \leq 2\}] \leq 3E[\#\{(i, j) : |X_i - X_j| \leq 1\}]$$

$$n + 2\binom{n}{2} P[|X - Y| \leq 2] \leq 3n + 6\binom{n}{2} P[|X - Y| \leq 1]$$

Границата се достига, когато n клони към безкрайност.

Задача 20 (Алон и Спенсър, глава 1)**.** Нека $\{(A_i, B_i), 1 \leq i \leq h\}$ е семейство от двойки подмножества на множеството на целите числа, така че $|A_i| = k$ за всяко i,

$|B_i|= l$ за всяко i, $A_i \cap B_i = \emptyset$ и $(A_i \cap B_j) \cup (A_j \cap B_i) \neq \emptyset$ за всяко $i \neq j$. Докажете, че $h \leq \dfrac{(k+l)^{k+l}}{k^k l^l}$.

Решение. Нека S е множеството от целите числа, които принадлежат на някое от A_i и B_i. Приемаме, че f е произволна функция, която изпраща S във $[1, k+l]$, а X_i е събитието, което изпраща A_i във $[1, k]$, а B_i-в [k+1, k+l]. Следователно

$$P[X_i] = \frac{k^k l^l}{(k+l)^{k+l}}$$

Нека $h > \dfrac{(k+l)^{k+l}}{k^k l^l}$. Тогава математическото очакване за X_i е по-голямо от 1 и така поне две от събитията се случват с положителна вероятност. Ако случаят е за X_i и X_j, тогава $A_i \cap B_j$ и $A_j \cap B_i$ са празни, противоречие с условието.

Задача 21 (Алон и Спенсър, глава 1)**.** Нека $n \geq 4$ и H е n-хиперграф с най-много $\dfrac{4^{n-1}}{3^n}$ ръба. Докажете, че съществува оцветяване на върховете на H в 4 цвята, така че във всеки ръб се съдържат и четирите цвята. (В един n-хиперграф всяко ребро, с други думи множество от върхове,съдържа точно n върха.)

Решение. Произволно оцветяваме всеки връх в един от четирите цвята. Вероятността всеки връх да съдържа най-много 3 цвята е по-малка или равна на $4 \cdot \dfrac{3^n}{4^n} = \dfrac{3^n}{4^{n-1}}$. Така математическото очакване за върховете, които съдържат по-малко от 4 цвята, е по-малко от 1. От последното следва, че съществува някакво оцветяване без ръбове с по-малко от 4 цвята, което и трябваше да докажем.

Задача 22 (Алон и Спенсър, глава 1)**.** За всяко естествено число k съществува граф турнир $T_k = (V, E)$ с $|V|> k$, така че за всяко множество U с най-много k върха от T_k съществува връх v, за който всички насочени ребра $\{(v, u) : u \in U\}$ принадлежат на E. Покажете, че всеки такъв граф турнир съдържа най-малко $\Omega(k2^k)$ върха.

Решение. Нека T_k е граф турнир със свойството S_k (свойството на графа турнир в условието) и с n върха. Избираме произволен връх v_k, математическото очакване за чиято степен е по-малко от $n/2$. Фиксираме връх със степен по-малка от $n/2$. Нека T_{k-1} е подтурнирът от върхове, които са победили v_k. Тогава $|T_{k-1}|< n/2$ и

T_{k-1} има свойство S_{k-1} (Във всяко множество от T_{k-1} с $k-1$ върха, заедно с v_k, съществува връх u, който побеждава всички останали, включително и v_k, и следователно принадлежи на множеството T_{k-1}). Продължавайки аналогично, достигаме до $T_k \supset T_{k-1} \supset T_{k-2} \supset \ldots \supset T_2 \supset T_1$, като за всяко i T_i е множество от върхове, които побеждават $v_k, v_{k+1}, \ldots, v_i, v_{i+1}$, а v_i е избран връх, така че $2|T_i| < |T_{i+1}|$. Оказва се, че T_1 е непразното множество от всички върхове, които побеждават $v_k, \ldots, v_2$. Нека l е някой връх извън T_k. Заедно $v_k, \ldots, v_2, l$ образуват k-елементно подмножество на T_k и трябва да има $u \in T_1$, който побеждава всички останали. По принцип обаче няма l, който да побеждава всички в T_1. И тъй като T_1 принадлежи на турнир T_k със свойство k, $|T_1| \geq k+1$.Така

$$|T_k| \geq 2^{k-1}|T_1| \geq 2^{k-1}(k+1),$$

което и трябваше да докажем.

Задача 23 (Алон и Спенсър, глава 2)**.** Нека $n \geq 2$ и $H = (V, E)$ е n-хиперграф с $|E| = 4^{n-1}$ ребра. Докажете, че съществува оцветяване на V с 4 цвята, така че никое ребро не е едноцветно.(В един n-хиперграф всяко ребро, с други думи множество от върхове,съдържа точно n върха.)

Решение. Оцветяваме върховете на H произволно в четирите цвята. За всяко ребро има вероятност 4^{-n+1} да бъде едноцветен. Следователно математическото очакване за едноцветните ребра е 1. Знаем, че има начин да получим повече от едно едноцветно ребро(всичко е в един цвят) и така заключаваме, че трябва да съществува начин, по който се достига до 0.

Задача 24 (Алон и Спенсър, глава 2)**.** Нека $p > n > 10m^2$, p е просто, а $0 < a_1 < a_2 < \cdots < a_m < p$ са цели числа. Докажете, че съществува цяло число x, $0 < x < p$, за което m-те на брой числа

$$((xa_i)(modp))modn, (1 \leq i \leq m)$$

са две по две различни.

Решение. Избираме за x някое от $1, 2, \dots, p-1$ и дефинираме индикатора X_{ij} като 1, ако $(xa_i modp) modn = (xa_j modp) modn$. Съществуват най-много $2(p-1)/n$ кратни на n, чиято стойност може да приеме $(xa_i modp) - (xa_j modp)$ и всяко от тях е с вероятност $1/p-1$. Следователно математическото очакване за броя на двойките от този тип е най-много

$$\frac{2}{n}\binom{m}{2} < 1,$$

тъй като $10m^2 < n$. От последното заключаваме, че за някоя стойност на x няма две равни числа.

Задача 25 (Алон и Спенсър, глава 2)**.** Нека F е семейство от подмножествва на множеството $N = \{1, 2, \dots, n\}$ и няма $A, B \in F$, за които $A \subset B$. Тогава с $\sigma \in S_n$ означаваме произволна пермутация на елементите на N, а с X случайната величина

$$X = |\{i : \{\sigma(1), \sigma(2), \dots, \sigma(i)\} \in F\}|$$

Пресмятайки математическото очакване за X, докажете, че $|F| \leq \binom{n}{\lfloor n/2 \rfloor}$.

Решение. От условието, че няма $A, B \in F$, за които $A \subset B$, знаем, че $X \leq 1$. От друга страна всеки елемент $S \in F$ има най-малко вероятност $1/\binom{n}{\lfloor n/2 \rfloor}$ да се появи в X. От това и линейността на математическото очакване следва, че $|F| \leq \binom{n}{\lfloor n/2 \rfloor}$.

Задача 26 (Алон и Спенсър, глава 2)**.** Нека $G = (V, E)$ е двуделен граф с n върха и списък $S(v)$ с повече от $\log_2 n$ цвята за всеки връх $v \in V$. Докажете, че съществува подходящо оцветяване на G, така че всеки връх v е в цвят от своя списък $S(v)$.

Решение. Нека C е множеството, образувано от всички цветове, които се срещат във някой от списъците. За всеки цвят имаме вероятност $\frac{1}{2}$ да се срещне в някой от дяловете на графа. Ако всеки връх е в един и същи дял с поне един от цветовете си, твърдението е доказано, като използваме някой от тези цветове. Вероятността за определен връх v всеки от цветовете във списъка му да се среща в другия дял е $\frac{1}{2^{|S(v)|}} < \frac{1}{n}$. Така матемическото очакване за върховете, които са с несъответстващ цвят, е по-малко от 1 от линейността на математическото очакване, от което следва и нашият резултат.

Задача 27 (USAMO 2012/2)**.** Окръжност е разделена на 432 равни дъги от 432 точки. Точките са оцветени в 4 цвята, така че 108 от тях са оцветени в червено, 108–в зелено, 108–в синьо, а останалите 108 са оцветени в жълто. Докажете, че могат да бъдат избрани по три точки от всеки цвят, така че триъгълниците, определени от едноцветните точки, са еднакви.

Решение. Ако завъртим червените точки 431 пъти, те ще се препокрият със зелените точки 108^2 пъти, което прави средно по $\dfrac{108^2}{431} > 27$. Следователно в някакъв момент поне 28 червени точки препокриват зелените точки и съществуват два еднакви (червен и зелен) 28-ъгълника.

Ако завъртим тези 28 червени точки 431 пъти, те ще се препокрият със сините точки $108 \cdot 28$ пъти, което прави средно по $\dfrac{108 \cdot 28}{431} > 7$. Следователно в някакъв момент поне 8 червени точки препокриват сините точки и съществуват три еднакви (червен, зелен и син) 8-ъгълника.

Ако завъртим тези 8 червени точки 431 пъти, те ще се препокрият с жълтите точки $108 \cdot 8$ пъти, което прави средно по $\dfrac{108 \cdot 8}{431} > 2$. Следователно в някакъв момент поне 3 червени точки препокриват жълтите точки и съществуват четири еднакви (червен, зелен, син, жълт) триъгълника.

Задача 28 (IMO 2006 C3 Shortlist)**.** Нека S е множество от точки в равнината, така че никои три от тях да не лежат на една права. За всеки изпъкнал многоъгълник P, означаваме с $a(P)$ броят на върховете на P, а с $b(P)$ броят на точките от S извън P. Докажете, че за всяко реално число x

$$\sum_P x^{a(P)}(1-x)^{b(P)} = 1,$$

където сумата е изчислена за всички изпъкнали многоъгълници с върхове във S.(Смята се, че отсечка, точка и празното множество са съответно многоъгълници с 2, 1 и 0 върха.)

Решение. Броят на точките в многоъгълника P е $c(P) = |S| - a(P) - b(P)$. Ако положим $y = 1 - x$ и заместим в лявата част на исканото равенство получаваме

$$\sum_P x^{a(P)}(1-x)^{b(P)} = \sum_P x^{a(P)} y^{b(P)} (x+y)^{c(P)} =$$

$$= \sum_P \sum_{i=0}^{c(P)} \binom{c(P)}{i} x^{a(P)+i} y^{b(P)+c(P)-i} =$$

$$= \sum_P \sum_{k=a(P)}^{a(P)+c(P)} \binom{c(P)}{k-a(P)} x^k y^{|S|-k} =$$

$$= \binom{|S|}{k} x^k y^{|S|-k} = 1$$

Коефициентът пред $x^k y^{|S|-k}$ е сумата $\sum_P \binom{c(P)}{k-a(P)}$, която е равна на броя на двойките (P, H) от изпъкнал многоъгълник P и k-елементно подмножество от S, чиято изпъкнала обвивка е P, което пък от своя страна е равно на $\binom{|S|}{k}$. Оттук доказателството следва директно.

6 Заключение и бъдещо развитие

В настоящата разработка въведохме един от най-широкоразпространените и полезни комбинаторни методи – вероятностния метод, и доказахме голямото му приложение. Разгледахме няколко помощни теореми. Бяха дадени примери за основни задачи и бяха обяснени ясно и разбираемо принципите на вероятностния метод. Като извод забелязахме, че наличието на метода в решението на определена задача спомага за простотата му и улеснява решаващия значително.
Като бъдещо развитие авторът си поставя за цел да изследва по-конкретни и/или отворени проблеми и варианти за решаването им чрез вероятностен метод.

7 Благодарности

Искам да благодаря на научния си ръководител Димитър Чакъров за това, че ми възложи тази тема и ми предостави нужната литература. Изказвам своите благодарности към него и за конструктивната критика и всички полезни съвети, които ми даде. Благодаря на Никола Стайков за помощта относно текста на разработката.

Литература

[1] Noga Alon, Joel H. Spencer, **The probabilistic method**, John Wiley and Sons, 2008

[2] Evan Chen, **Expected uses of probability**, August 2014

[3] Pranav A. Sriram, **Olympiad Combinatorics**, August 2014