Санкт-Петербургский
государственный
университет

Бросание шарика на доску Гальтона

Татьяненко Алексей, учащийся 11 М класса
Академической гимназии СПбГУ им. Д.К. Фаддеева

Научный руководитель: **Звонарев Никита Константинович**, к. ф.-м. н.,
старший преподаватель кафедры статистического моделирования
математико-механического факультета СПбГУ

XXX Всероссийская научно-практическая конференция «Университетская гимназия»
Санкт-Петербург, 18 марта 2021 г.

Что такое доска Гальтона?

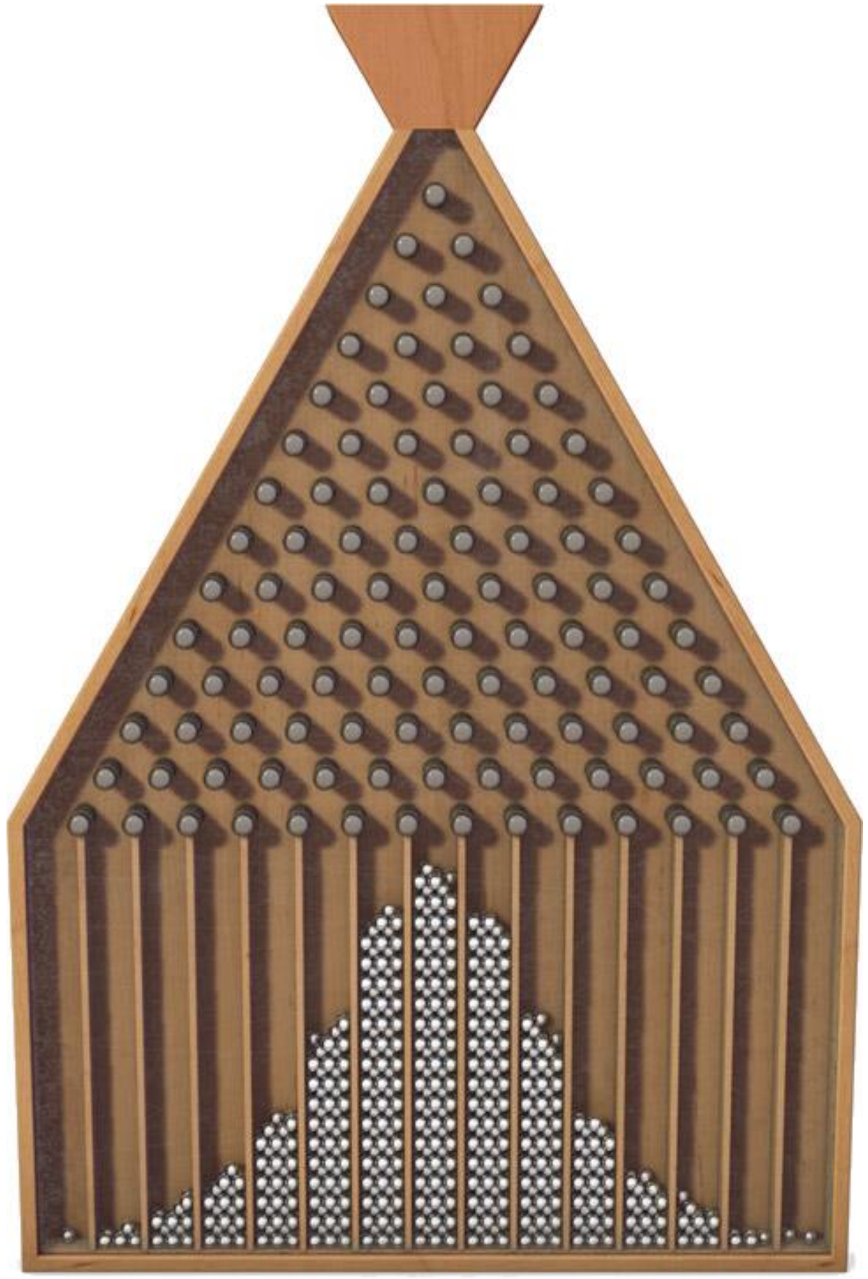
Доска Гальтона — устройство, изобретённое английским учёным Фрэнсисом Гальтоном для демонстрации нормального закона распределения.

В рассмативаемом случае представляет из себя треугольник, образуемый N рядами гвоздиков, вбитых в доску, (количество гвоздиков в ряде равно номеру ряда). В классической доске Гальтона вероятности поворота направо и налево равны $1/2$.

В $N+1$ ряду находится $N+1$ лунка.

Сверху на гвоздики забрасывается шарик, после чего он спускается по рядам до лунок.

Если шариков много, то моделируется распределение.



Постановка задачи

Рассмотрим обобщённую модель доски Гальтона:

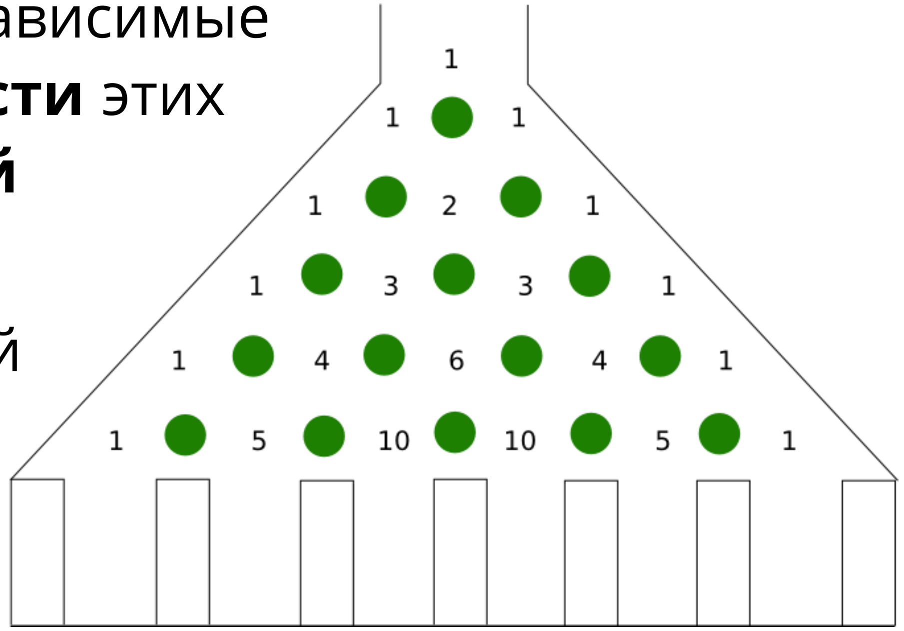
Обозначим через (i, j) гвоздик в i–ом ряду $(1 \leq i \leq N)$ в j-ой позиции $(1 \leq j \leq i)$.

1) Изначально шарик бросается на гвоздик $(1,1)$.
2) В каждом следующем ряду шарик
   - с вероятностью $P_{i,j}$ падает на гвоздик $(i + 1, j + 1)$ либо
   - с вероятностью $(1 - P_{i,j})$ на гвоздик $(i + 1, j)$.

- 1. При каком наборе вероятностей $P_{i,j}$ шарик с одинаковой вероятностью $1/(N + 1)$ попадает в каждую из лунок?
  2. С вероятностями $P_1, P_2, \ldots, P_{N+1}$ в соответствующие лунки?

Вероятности в доске Гальтона

Т.к. поворот на гвоздиках i-го и $(i+1)$-го рядов — независимые события, **вероятность** некоторой **последовательности** этих **двух поворотов** будет ***произведением*** **вероятностей** осуществлённых на каждом из них **поворотов**.



При отсутствии нулевых или единичных вероятностей поворота число маршрутов, ведущих к гвоздику (i, j), равно C_{i-1}^{j-1}. Это видно при наложении треугольника Паскаля (с перенумерацией) на доску Гальтона.

Тогда, если предположить равенство всех $P_{i,j} = p$, мы сможем использовать формулу испытаний Бернулли для определения вероятности попадания в лунку $(N+1, j)$:

$$P_{N+1,j} = C_N^{j-1} p^{j-1} (1-p)^{N-j}.$$

Однако эта формула *не позволяет* получить более двух одинаковых вероятностей попадания в лунку, поэтому предположение неверно, и **необходимо искать другие подходы** к решению задачи.

Актуальность задачи

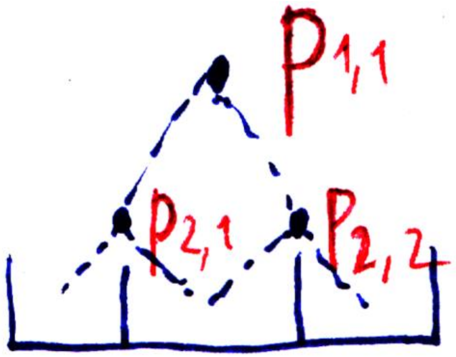
Модифицированная таким образом доска Гальтона позволяет моделировать равномерное либо произвольное дискретное распределение с $N+1$ исходами, используя только последовательность из распределений Бернулли с вычисленными мной вероятностями успеха.

Это может использоваться, например, для генерации случайных чисел.

Случай $N = 2$

Найдём $P_{1,1}, P_{2,1}, P_{2,2}$ в общем виде. Записывая вероятности попадания в каждую из лунок как сумму вероятностей попадания в неё каждым из возможных маршрутов (несовместных исходов), получим и решим систему уравнений:

$$\begin{cases} P_{1,1}P_{2,2} = \frac{1}{3} \\ (1 - P_{1,1})(1 - P_{2,1}) = \frac{1}{3} \\ (1 - P_{1,1})P_{2,1} + P_{1,1}(1 - P_{2,2}) = \frac{1}{3} \end{cases} \Leftrightarrow \begin{cases} P_{2,1} = \frac{3P_{1,1} - 2}{3P_{1,1} - 3} \\ P_{2,2} = \frac{1}{3P_{1,1}} \\ P_{1,1} \in [1/3\,; 2/3] \end{cases}$$



При $P_{1,1} \in (2/3\,; 1)$ $P_{2,1} < 0$, при $P_{1,1} \in (0; 1/3)$ $P_{2,2} > 1$, а при $P_{1,1} = 0$, $P_{1,1} = 1$ пропадают маршруты, ведущие в крайние лунки.

Выражение двух неизвестных через третью указывает на то, что исходные уравнения не были независимыми.

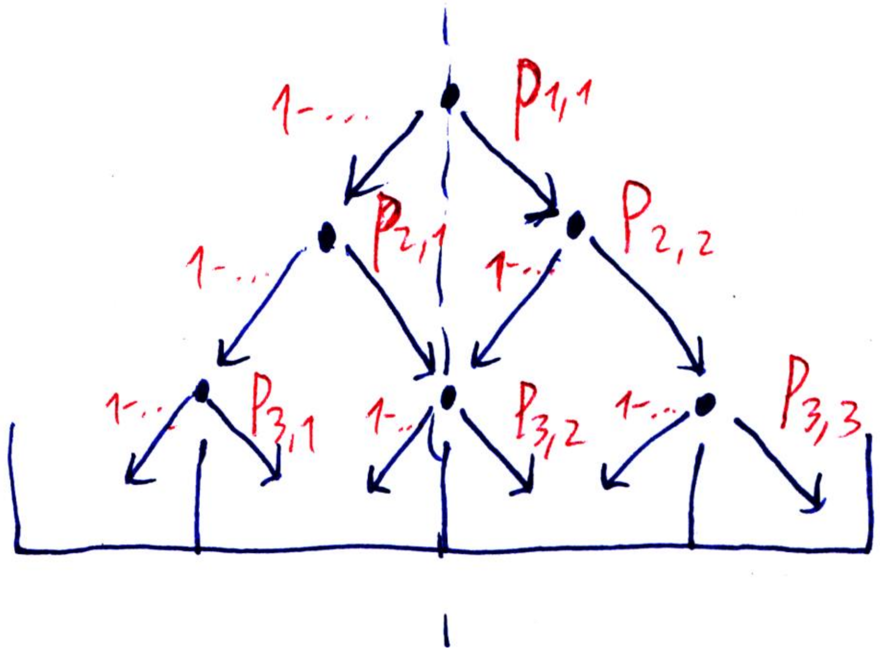
Случай $N = 3$.
Представление в виде взвешенного орграфа

Для случая $N = 3$ решение в общем виде представляется заметно более трудоёмким, так как теперь система включает в себя 6 неизвестных.

Поэтому сменим подход.

Представим доску Гальтона как взвешенный ориентированный граф:

- гвоздики и лунки соответствуют вершинам,
- рёбра соединяют гвоздик с двумя гвоздиками на ряд ниже, куда может упасть шарик,
- рёбрам приписан вес, равный вероятности прохождения по ним (при условии, что шарик уже находится на этом гвоздике).



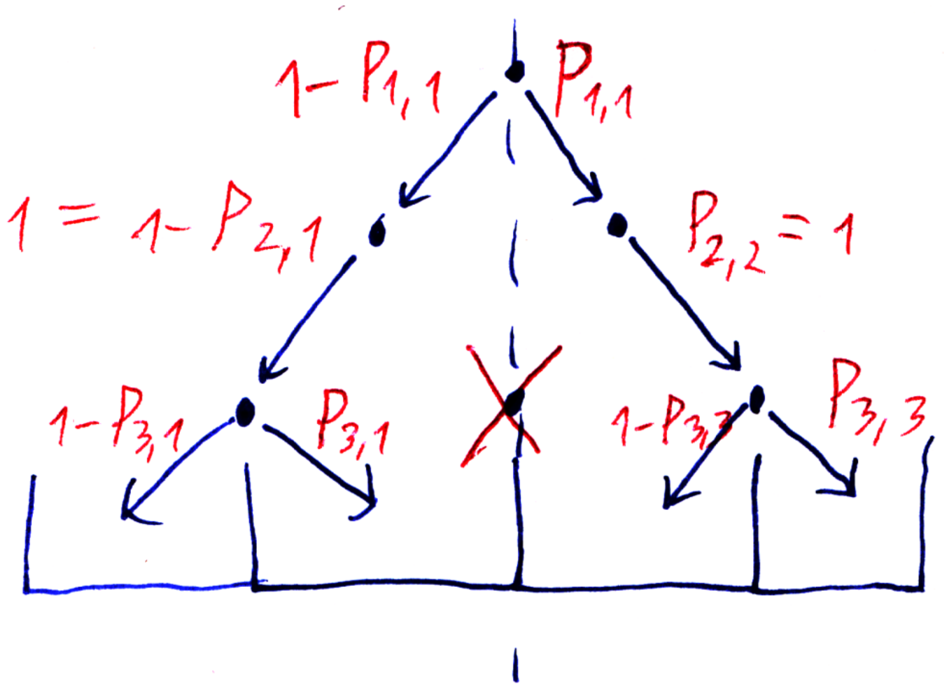
Случай $N = 3$. Решение

Через вершины $(1,1)$ и $(3,2)$ проходит ось симметрии доски. Исключив вершину $(3,2)$ и ведущие из неё рёбра, получим дерево, симметричное относительно этой оси.

Для этого приписываем рёбрам из $(2,1)$ в $(3,1)$ и из $(2,2)$ в $(3,3)$ веса 1.

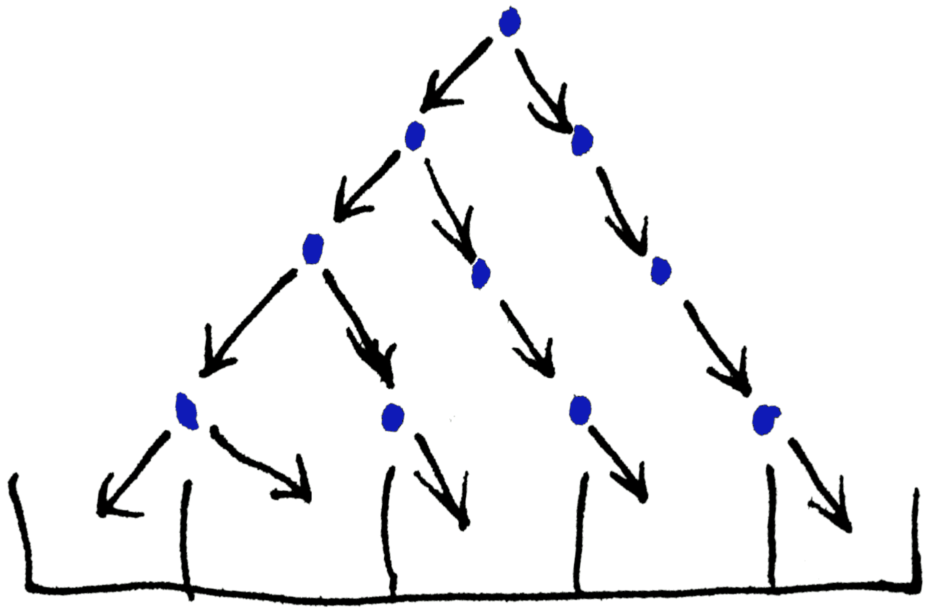
Тогда ответом будет:

$P_{1,1} = 1/2, \qquad P_{2,1} = 0, \qquad P_{2,2} = 1, \qquad P_{3,1} = 1/2,$
$P_{3,2} \in [0;1], \qquad P_{3,3} = 1/2$



Случай произвольного N, вероятность зависит от N. Выбор подхода

- Для случая произвольного N с зависимостью $P_{i,j}$ от N ограничим маршруты иначе: **однажды** повернув направо, шарик больше **никогда** не поворачивает налево.
- Тогда $P_{i,j} = 1$ для $i \in [2; N]$, $j \in [2;\ i]$



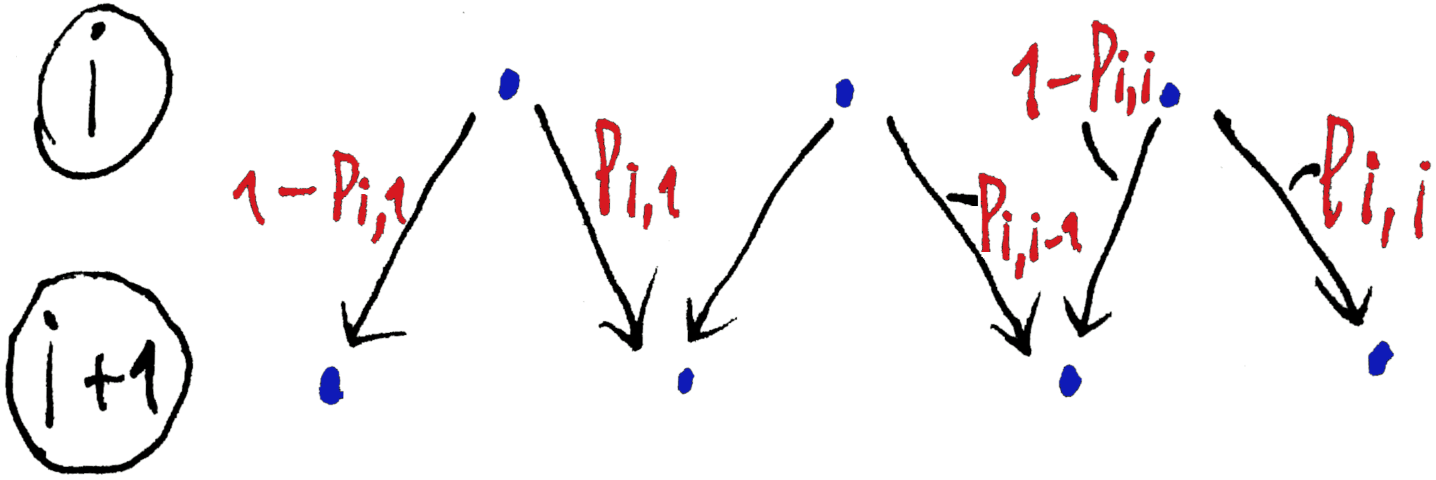
Случай произвольного N, вероятность зависит от N. Решение

- Выразим $P_{1,1}$, $P_{2,1}$, $P_{3,1}$ и подметим закономерность.
- Далее предположим, что $P_{i,1} = 1/(N + 2 - i)$, где $i \in [1; N]$, и докажем это с помощью ММИ.
- После доказательства этой формулы получаем систему, определяющую все возможные вероятности $P_{i,j}$:

$$\begin{cases} P_{i,j} = 1 \text{ для } i \in [2; N], j \in [2;\ i], \\ P_{i,1} = 1/(N + 2 - i)\text{, где } i \in [1; N]. \end{cases}$$

Случай произвольного N, вероятность не зависит от N . Выбор подхода

- Поскольку вероятности $P_{i,j}$ не зависят от N, для досок с разным количеством рядов гвоздики (i, j) неразличимы.
- Выдвинем и докажем предположение, что на **всех** рядах **любой** такой доски достигается дискретное равномерное распределение

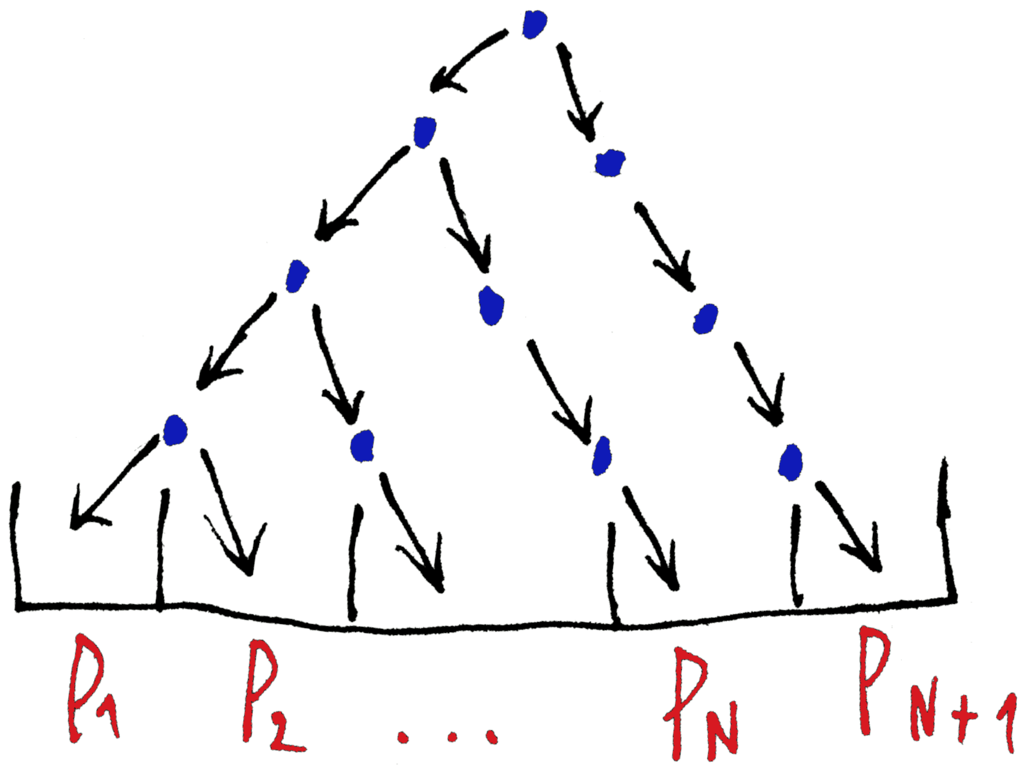


Случай произвольного N, вероятность не зависит от N . Решение

- Выразим $P_{i,1}$, $P_{i,i-1}$, $P_{i,i}$ и подметим закономерность.
- Предположим, что $P_{i,1} = j/(i+1)$, где $i \in [1; N]$, j $\in [1; i-1]$.
- Запишем вероятность попадания шарика в узел $(i+1, j+1)$ и покажем, что формула даёт тождество.
- Проверим узел $(i+1, i+1)$.
- Получим ответ: $P_{i,j} = j/(i+1)$, где $i \in [1; N], j \in [1;\ i]$.
- Важно заметить, что при таком алгоритме моделирования вероятность $P_{i,j}$ можно вычислить при сколь угодно большом и заведомо неизвестном N.

Произвольное дискретное распределение с зависимостью от N. Выбор подхода

- Воспользуемся ограничением маршрутов по методу, применявшемуся в случае с моделированием дискретного равномерного распределения.
- Из разбиения следует $P_{i,j} = 1$ для $i \in [2; N]$, $j \in [2;\ i]$.



Произвольное дискретное распределение с зависимостью от N. Решение

- Запишем формулу для $P_{i,1}$ и, сократив дроби, получим:

$$P_{i,1} = \frac{P_{N+2-i}}{1 - \sum_{k=N+3-i}^{N+1} P_k}.$$

- После проверки соблюдения условий, получаем ответ

$$\begin{cases} P_{i,j} = 1 \text{ для } i \in [2; N], j \in [2;\ i], \\ P_{1,1} = \ P_{N+1}, \\ P_{i,1} = \dfrac{P_{N+2-i}}{1 - \sum_{k=N+3-i}^{N+1} P_k}, \text{где } i \in [2; N]. \end{cases}$$

Заключение

В ходе работы были получены:

- дискретное равномерное распределение, где $P_{i,j}$ зависит от N,
- дискретное равномерное распределение, где $P_{i,j}$ не зависит от N,
- некоторое дискретное распределение ($P_{i,j}$ зависит от N), соответствующее произвольному набору вероятностей попадания в лунку $P_1, P_2, \ldots, P_{N+1}$.

Используемая литература

- Лютикас В.С., Школьнику о теории вероятностей, 2-е изд., М.: Просвещение, 1983, гл. I.
- Колмогоров А.Н., Журбенко И.Г., Прохоров А.В., Введение в теорию вероятно-стей, М: Наука, 1982, гл. 1, §5.
- Ширяев А.Н., Вероятность — 1, 3-е изд., М: МЦНМО, 2004, гл. 1, §1–3.
- Карпов Д.В., Теория графов. [Рукопись] https://logic.pdmi.ras.ru/~dvk/graphs_dk.pdf