Министерство науки и высшего образования Российской Федерации

федеральное государственное автономное образовательное учреждение высшего образования

«НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ ИТМО»

Теория вероятностей и математическая статистика

Отчет по лабораторной работе №1

Авторы:

Юманов М.А.

Алексеева В.В.

Преподаватель:

Лимар И.А.

Санкт-Петербург, 2024

Задание 1

Условие:

В городе с населением $(n+1)$ человек некто узнает новость. Он передает ее первому встречному, тот еще одному и т.д. На каждом шаге услышавший новость может сообщить её любому из n человек с одинаковыми вероятностями. Найти вероятности того, что в течение r единиц времени новость не возвратиться к впервые узнавшему ее горожанину и новость не будет никем повторена. Решить ту же задачу в предположении, что на каждом шаге новость сообщается группе из m человек.

Решение:

Вероятность, что новость не вернется к впервые узнавшему ее горожанину через r шагов обозначим $P(r)$. Обозначим $(n+1)$ как человека, который первым все знал.

Вероятность того, что новость не вернется к первоначальному человеку после первого шага и далее:

$$P(1) = \frac{n}{n} \quad P(2) = \frac{n-1}{n} \quad P(3) = \frac{n-2}{n}$$

На следующем шаге каждый новый слушатель может сообщить новость любому из n человек. Тогда вероятность того, что в течение r единиц времени новость не возвратиться к впервые узнавшему ее горожанину, равна

$$P(r) = \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdots \frac{n-r+1}{n} = \frac{n!}{n^r \cdot (n-r)!}$$

Запись подразумевает, что новость не была никем повторена.

Далее рассмотрим случай, когда новость передается группе из m человек на каждом шаге. Возьмем количество сочетаний по m как число людей, которые не знают новость. Кроме того, в знаменателе были описаны не все люди, а лишь их выборка из m человек.

Итого вероятность имеет вид

$$P(r) = \frac{C_n^m \cdot C_{n-m}^m \cdot \ldots \cdot C_{n-mr}^m}{(C_n^m)^r}$$

Задание 2

Условие:

Уходя из детского сада, каждый из $\boldsymbol{n}$ случайно берет один левый и один правый ботинок. Найти вероятность того, что каждый их них возьмет не свой левый и не свой правый ботинок.

Решение:

Вначале предлагается попытаться заставить малыша взять свой левый ботинок.

Пусть A_i – событие, при котором i малышей взяли свой ботинок. Тогда по формуле включения-исключения число комбинаций равно

$$|A_1 \cup \dots \cup A_n| = \bigcup_{i=1}^{n} |A_i| = \sum_{i=1}^{n} A_i - \sum_{i<j}^{n} |A_i \cap A_j| \dots (-1)^{n+1} |A_1 \cap \dots \cap A_n|$$

Сумма i объединений может быть записана как

$$|A_1 \cup \dots \cup A_n| = C_n^1 \cdot (n-1)! - C_n^2 \cdot (n-2)! \dots (-1)^{n+1}$$

$$= n! - \frac{n!}{2!} + \cdots + (-1)^{n+1}$$

Правые ботинки описываются аналогично.

Тогда вероятность того, что оба ботинка были взяты правильно равна произведению вероятностей того, что ребенок взял свой левый и правый ботинок, так как эти события независимы.

Иначе описать это можно словами: каждый из $\boldsymbol{n}$ детей взял свой левый и свой правый ботинок.

Итого запишем в формуле

$$\left(\frac{n! - \frac{n!}{2!} + \cdots + (-1)^{n+1}}{n!}\right)^2$$

При этом важно учесть, что вопрос поставлен противоположно. Чтобы найти вероятность того, что каждый малыш возьмет не свой левый и не свой правый ботинок, необходимо вычесть из полной вероятности рассмотренный выше случай.

В итоге искомый ответ

$$P(n) = \left(1 - \frac{n! - \frac{n!}{2!} + \cdots + (-1)^{n+1}}{n!}\right)^2$$

Задание 3

Условие:

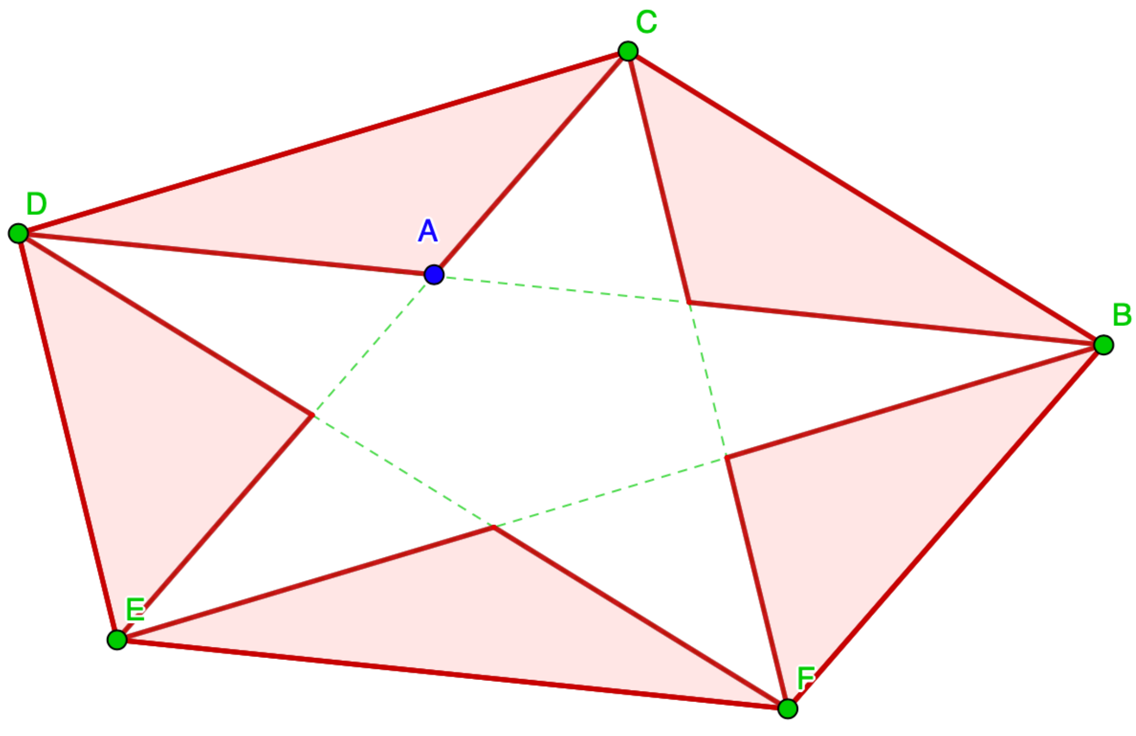
Случайная точка A имеет равномерное распределение в правильном $n-$ угольнике. Найти вероятность $P(n)$, что точка A находится ближе к границе многоугольника, чем к его диагоналям.

Найти числа C, x, что

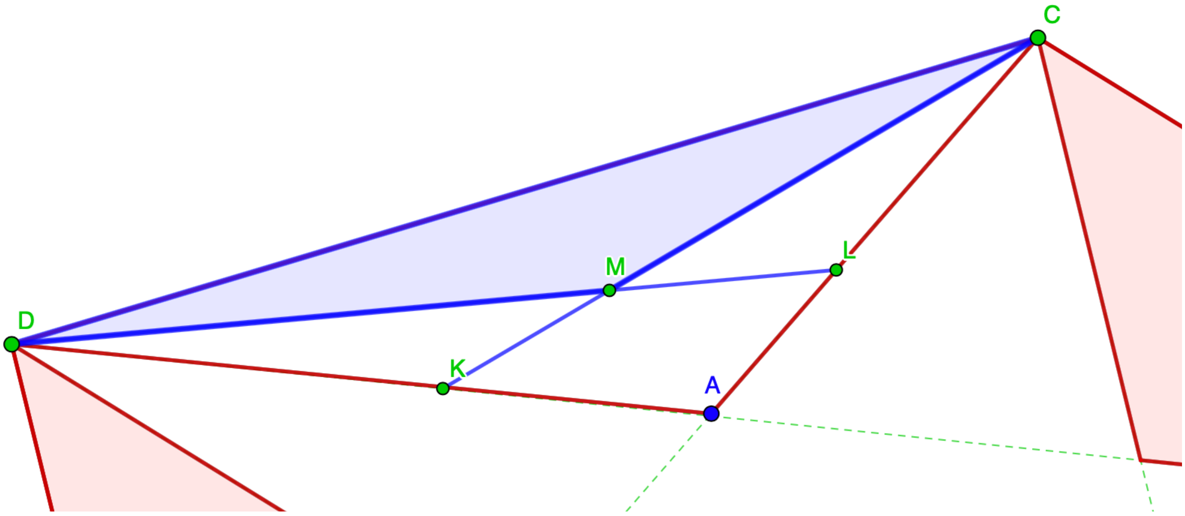
$$P(n) = Cn^x\big(1 + o(1)\big), \qquad n \to \infty$$

Решение:

Осознание задачи начинается с поиска места, в котором точка может находится ближе к границе многоугольника, чем к его диагоналям. Для этого необходимо, чтобы она находилась внутри определенного сектора, ограниченного границей и диагональю. Нам надо провести линии, которые бы проходили через две вершины. На рисунке обозначим эти диагонали зеленым цветом.



Далее рассмотрим один из секторов в виде треугольника, который образовался посредством соединения нашей точки A и диагоналей. Проведем две крыски-биссектриски CC_1 и DD_1 внутри треугольника.



Произведенные манипуляции привели местонахождению искомой области, а именно синего треугольника.

Перейдем к его рассмотрению и найдем все углы.

Возьмем $\angle CDA = \delta \Rightarrow \angle CAD = 180° - 2\delta$

Угол в правильном n *−угольнике равен по формуле* $180°(n-2)/_n$

В таком случае $\angle FDA = 180° - \frac{360°}{n} - 2\delta$ *и далее*

$$\angle DAE = \frac{180° - \angle FDA}{2} = \frac{180°}{n} + \delta$$

Учитывая, что $\angle DAE + \angle DAC = 180° \Rightarrow \delta = \frac{180°}{n}$

Следом после углов рассмотрим $\triangle CDM$ *и найдем его площадь.*

Пусть $CD = x$ *и* $CM = y$. *Тогда по теореме косинусов запишем:*

$$x^2 = 2y^2 - 2y^2 \cos\left(180° - \frac{180°}{n}\right)$$

$$y^2 = \frac{x^2}{2 + 2\cos\left(\frac{180°}{n}\right)}$$

Треугольник $\triangle CDM$ *равнобедренный, поэтому его площадь*

$$S = \frac{1}{2} y^2 \sin\left(180^\circ - \frac{180^\circ}{n}\right) = \frac{1}{2} y^2 \sin\left(\frac{180^\circ}{n}\right)$$

Выразим площадь через x:

$$S = \frac{x^2 \sin\left(\frac{180^\circ}{n}\right)}{4 + 4\cos\left(\frac{180^\circ}{n}\right)}$$

Перейдем искомой вероятности $P(n)$.

Итак, она равна отношению площади треугольника, образованного вершинами $n-$*угольника и точкой* A, *к площади всего* $n-$*угольника.*

Площадь нашего треугольника может быть найдена как

$$S_\triangle = n \cdot \frac{x^2 \sin\left(\frac{\pi}{n}\right)}{4 + 4\cos\left(\frac{\pi}{n}\right)}$$

Площадь всего $n-$*угольника:*

$$S_n = n \cdot \frac{x^2}{4\tan\left(\frac{\pi}{n}\right)}$$

Итого объединим полученное ранее и запишем

$$P(n) = \frac{S_\triangle}{S_n} = \frac{\sin\left(\frac{\pi}{n}\right)\tan\left(\frac{\pi}{n}\right)}{1 + \cos\left(\frac{\pi}{n}\right)} = \tan\left(\frac{\pi}{2n}\right)\tan\left(\frac{\pi}{n}\right)$$

Далее рассмотрим предел при $n \to \infty$

$$\lim_{n\to\infty}\left[\tan\left(\frac{\pi}{2n}\right)\tan\left(\frac{\pi}{n}\right)\right] = \lim_{n\to\infty}\left[\tan\left(\frac{\pi}{2n}\right)\right]\lim_{n\to\infty}\left[\tan\left(\frac{\pi}{n}\right)\right] = \left.\frac{\pi^2}{2n^2}\right|_{n\to\infty}$$

С учетом того, что

$$P(n) = Cn^x\big(1 + o(1)\big), \qquad n \to \infty$$

Искомые числа

$$C = {\pi^2}/{2} \text{ и } x = -2$$

Задание 4

Условие:

Введем события $A_i = \{X = i\}$, $B_i = \{Y = i\}$, $i \geq 0$. *Известно, что для любых* $i \geq 0$ *и* $j \geq 0$ *события* A_i *и* B_j *– независимы, при этом*

$$P(X = i) = P(Y = i) = (1 - p)^i p \quad i \geq 0,$$

где $p \in (0,1)$

Найти $P(X = i|X + Y = j)$.

Решение:

Воспользуемся теоремой Байеса:

$$P(X = i|X + Y = j) = \frac{P(X + Y = j|X = i) \cdot P(X = i)}{P(X + Y = j)}$$

Заметим далее, что $P(X + Y = j|X = i)$ *есть то же самое, что и* $P(Y = j - i)$.

Далее подставим полученное:

$$= \frac{P(Y = j - i) \cdot P(X = i)}{P(X + Y = j)}$$

Кроме того, по формуле полной вероятности, запишем

$$P(X + Y = j) = \sum_{k=0}^{j} P\big((X = k) \cdot (Y = j - k)\big)$$

Независимость событий подразумевает, что $P(AB) = P(A) \cdot P(B)$. *Иначе это преобразование можно описать как* $P\big((X = k) \cdot (Y = j - k)\big) = P(X = k) \cdot P(Y = j - k)$

Итого мы пришли к вероятности:

$$= \frac{P(Y = j - i) \cdot P(X = i)}{\sum_{k=0}^{j} P(X = k) \cdot P(Y = j - k)} = \frac{(1 - p)^{j-i}p \cdot (1 - p)^i p}{\sum_{k=0}^{j}(1 - p)^k p \cdot (1 - p)^{j-k} p}$$

$$= \frac{(1 - p)^j}{\sum_{k=0}^{j}(1 - p)^j} = \frac{(1 - p)^j}{j \cdot (1 - p)^j} = \frac{1}{j}$$

Задание 5

Условие:

В заключительной задаче нам предстоит рассмотреть схемы Бернулли при $n \in \{100, 1000, 10000\}$ *и* $p \in \{0.001, 0.01, 0.1, 0.25, 0.5\}$. *Для каждой пары* (n, p) *рассчитаем вероятности* $P(S_n \leq 5)$, $P(S_n = k_*)$, $P\left(S_n \in \left[n/2 - \sqrt{npq}, n/2 + \sqrt{npq}\right]\right)$ *по точным формулам (если это возможно), где* S_n *— количество успехов в* n *испытаниях,* k_* *— наиболее вероятное количество успехов. Кроме того, каждую вероятность вычислим приближенно с помощью теоремы Пуассона, локальной и интегральной теоремы Муавра-Лапласа. Сравним результаты, сделаем выводы.*

Решение:

Работа с кодом была проведена к Google Colab. Именно в нем содержатся результаты для различных заданных условий и различных теорем. ***Торжественно*** *предоставляем Вам наш блокнот:*

https://colab.research.google.com/drive/1PXw3rPRg9dlgdz5pqIvOiNsXuFJub9kd?usp=sharing

Формула Бернулли:

$$P(k) = C_n^k p^k (1-p)^{n-k}$$

Теорема Пуассона:

$$P(k) = {\lambda^k e^{-\lambda}}/{k!}, \qquad \lambda = np$$

Локальная теорема Муавра-Лапласа:

$$P(k) = \frac{\Phi(x)}{\sqrt{npq}}, \qquad \Phi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}}, \qquad x = \frac{k - np}{\sqrt{npq}}$$

Интегральная теорема Муавра-Лапласа:

$$P(a \leq k \leq b) \approx \Phi(x_b) - \Phi(x_a)$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{\frac{-t^2}{2}} dt, \qquad x_a = \frac{a - np}{\sqrt{npq}}, \qquad x_b = \frac{b - np}{\sqrt{npq}}$$

Какие же выводЫ?

Можно заметить, что все методы показывают схожие результаты (сходятся к нулю). При достаточно больших значениях n и k в численных методах получается близкий ответ, поэтому их можно смело использовать при сложных вычислениях для ускорения процесса (единственное, нужно следить за тем, чтобы число «влезло» в тип данных).