1. Постановка задачи

Определим процесс бернулиевского града, падающего на горячую поверхность (Bernoulli Hail on a Hot Ground). Пусть время дискретно, и поверхность представляет из себя целочисленную прямую $\mathbb{Z}$. Предположим, что в каждый момент времени $n \in \mathbb{Z}_+$ в каждую точку $i \in \mathbb{Z}$ поверхности, с вероятностью p, независмо от других параметров системы, приходит клиент, обозначаемый $\{i, i+1\}$ (клиент типа i) и ждет обслуживания. Каждому клиенту необходимо h единиц времени ($h = 1$), чтобы завершить его. Обслуживание клиентов происходит в порядке FIFO (first in first out). Однако, в один момент времени, не может происходить обслуживание клиентов, стоящих рядом. В случае, если в один момент времени в систему пришли клиенты типа i и $i+1$, то их порядок определяется случайным образом, а именно с вероятностью 1/2 кто-то из них окажется первее. Построим граф $G(p)$, определяющий порядок обслуживания клиентов. Сперва рассмотрим случай $p = 1$ и затем, на его основе, опишем граф для произвольного значения p. Узлами данного графа являются элементы множества $\{(i, n) : i \in \mathbb{Z},\, n \in \mathbb{Z}_+\}$. Ребро вида $(i_1, n_1) \to (i_2, n_2)$ означает, что клиент типа i_1, пришедший в момент времени n_1, не начнет обслуживание, пока клиент типа i_2, пришедший в момент времени n_2, не завершит его. Таким образом, учитывая ограничения модели, в каждой точке (i, n), будем строить граф следующим образом:

1. Строим ребро $(i+1, n) \to (i, n)$ с вероятностью 1/2, либо ребро $(i, n) \to (i+1, n)$ иначе.
2. Строим ребра $(i, n) \to (i-1, n-1)$, $(i, n) \to (i, n-1)$, $(i, n) \to (i+1, n-1)$ для всех $n \geq 2$.

Теперь опишем граф $G(p)$, при $0 \leq p < 1$.

1. Закрасим каждый узел графа $G(1)$ с вероятностью $1 - p$ в белый цвет и с вероятностью p в черный
2. Если узел (i, n) белый, то удалим все смежные с ним ребра, кроме ребра $(i, n) \to (i, n-1)$.

Заметим, что, для данного графа справедливо **свойство монотонности**, то есть, при $p \leq q$, верно включение $G(p) \subseteq G(q)$.

Пусть (i_1, n_1), (i_2, n_2),..., (i_m, n_m) — последовательно соединенные узлы графа $G(p)$ (все ребра направлены вправо). Назовем данную последовательность **путем** длины m, начинающийся в узле (i_1, n_1). Определим **высоту** пути как суммарное время обслуживания вдоль данного пути. Пусть теперь $H_n^i = H_n^i(p)$ — максимальная высота среди всех путей, начинающихся в узле (i, n) в графе $G(p)$. Покажем, что $H_n^i(p)$ ограничена. Учитывая свойство монотонности модели, а также однородность по пространству, достаточно это показать для $H_n^0(1)$. Пусть

$$t_{n,n}^{+} = \min\{i \geq 1 : (i, n) \to (i-1, n)\},$$

и для $m = n-1, n-2, ..., 1$ определим

$$t_{m,n}^{+} = \min\{i > t_{m+1,n}^{+} + 1 : (i, n) \to (i-1, n)\}.$$

Аналогично, пусть

$$t_{n,n}^{-} = \max\{i \leq -1 : (i, n) \to (i+1, n)\},$$

и для $m = n-1, n-2, ..., 1$

$$t^-_{m,n} = \max\{i < t^-_{m,n} - 1 : (i,n) \to (i+1,n)\}.$$

Тогда получим следующую оценку

$$H^0_n \le \sum_{i=1}^{n} (t^+_{i,n} - t^-_{i,n}) + n \tag{1}$$

Теперь, с помощью полученной оценки, оценим математическое ожидание H^i_n. Сперва найдем распределения случайных величин $\{t^+_{m,n}\}$, при $n \ge 1$ и $1 \le m \le n$ (для $\{t^-_{m,n}\}$ рассуждение аналогично). Сперва заметим, что

$$t^+_{m,n} - (t^+_{m+1,n} + 1) \overset{d}{=} t^+_{m,m}.$$

Тогда ясно, что

$$t^+_{m,n} \overset{d}{=} \sum_{k=0}^{m-n} t^+_{m+k,m+k} + (n-m).$$

Найдем распределение $t^+_{m,m}$. Так как

$$\{t^+_{m,m} = k\} = \{(0,m) \to (1,m), (1,m) \to (2,m), ..., (k-2,m) \to (k-1,m), (k-1,m) \leftarrow (k,m)\},$$

где $k \ge 1$, получим следующее распределение случайной величины $t^+_{m,m}$

$$\mathbb{P}(t^+_{m,m} = k) = 2^{-k}.$$

Тогда, для математического ожидания случайной величины $t^+_{m,m}$ получим

$$\mathbb{E}\, t^+_{m,m} = \sum_{k=1}^{\infty} \mathbb{P}(t^+_{m,m} \ge k) = \sum_{k=1}^{\infty} 2^{-k+1} = \sum_{k=0}^{\infty} 2^{-k} = 2 \tag{2}$$

Итого, учитывая формулы (1) и (2), для математического ожидания H^0_n получим

$$\mathbb{E}\, H^0_n \le \sum_{i=1}^{n} 6(n-i+1) + n = O(n^2) \tag{3}$$

2. Вопросы

1. Оценка производных для приближения по формуле Тейлора?
2. Возможно ли построить каплинг для данной модели со свойствами суб/супер аддитивности?
3. Можно ли получить оценки для предельных констант в случае, если вместо исходной поверхности будем рассматривать окружность с тремя точками? С произвольным количеством точек?
4. Можно ли комбинаторными методами найти распределение H^i_n, при некоторых допущениях для модели?
5. Провести компьютерную симуляцию модели.

3.1. Рассмотрим сперва случай $p = 1$. В силу правил построения ребер на одном временном уровне $(t - Fix)$, а также определения величины $H_n^i(p)$ ясно, что максимальная длина очереди на каждом временном уровне равна 3, а также, так как от каждого черного узла будет выходить три ребра на нижестоящий уровень, то ни один клиент, пришедший в момент времени n не будет обслужен, пока не закончат обслуживания клиенты на уровне $n - 1$. Таким образом, легко выписать распределение для $H_n^i(1)$, где $i = 1, 2, 3$ и $n \geq 0$.

$$H_n^i = \begin{cases} 3n-2, & \text{prob} = 1/3 \\ 3n-1, & \text{prob} = 1/3 \\ 3n, & \text{prob} = 1/3 \end{cases} \tag{4}$$

Пусть теперь $0 < p < 1$. Аналогично предыдущему случаю, можно заметить, что клиенты, пришедшие в момент времени n не будут обслужены, пока не завершится обслуживание клиентов, пришедших на уровень $n - 1$. Таким образом, искомую случайную величину H_n^i можно выразить в виде суммы двух независимых случайных величин ξ_n^i – длина очереди на уровне n и $B(3n-3, p)$ – биномиальная случайная величина с параметрами $3n - 3$ и p. Итого получим

$$H_n^i = \xi_n^i + B(3n-3, p). \tag{5}$$

Распределение ξ_n^i легко находится обычным перебором, поэтому приведем только результат вычислений

$$\xi_n^i = \begin{cases} 0, & \text{prob} = 1-p \\ 1, & \text{prob} = p - p^2 + \frac{1}{4}p^3 \\ 2, & \text{prob} = p^2(1-p/2) \\ 3, & \text{prob} = \frac{1}{4}p^3 \end{cases} \tag{6}$$

Распределение H_n^i несложно найти, например, по формуле полной вероятности. Коэффициент линейного роста можно найти из формулы 5, с помощью закона больших чисел. Итого получим

$$H_n^i/n \to 3p \quad i = 1, 2, 3. \tag{7}$$

4.1 Найдем оценку для распределения H_n^0 "напрямую". Для этого воспользуемся следующими рассуждениями

$$\begin{aligned} \mathbb{P}(H_n^0 = k) &= \mathbb{P}(\exists \text{ путь длины } k \text{ и } \neg\exists \text{ пути длины } k+1) \\ &= 1 - \mathbb{P}(\neg\exists \text{ пути длины } k \text{ или } \exists \text{ путь длины } k+1) \\ &\geq 1 - \mathbb{P}(\neg\exists \text{ пути длины } k) - \mathbb{P}(\exists \text{ путь длины } k+1) \\ &= \mathbb{P}(\exists \text{ путь длины } k) - \mathbb{P}(\exists \text{ путь длины } k+1). \end{aligned}$$

Теперь найдем вероятность событий $\{\exists \text{ путь длины } k\}$ для всех $k \geq 0$. Сперва заметим, что если мы зафиксируем точку откуда путь будет "стартовать"(пусть $(0, m)$), а также длину пути k, то количество различных конфигураций пути с заданными параметрами зависит от количества вариантов разложения числа k в сумму из m слагаемых $n_1, n_2, ..., n_m$, где n_i можно интерпретировать как длину части исходного пути на i-ом временном уровне, и количества вариантов разложить заданную конфигурацию частей пути. Обозначим далее через

l_i – части исходного пути, то есть $l_1 + l_2 + ... + l_m = H_m^0$. Найдем вероятность события $\{l_1 = n_1, l_2 = n_2, ..., l_m = n_m\}$

$$\begin{aligned}&\mathbb{P}(l_1 = n_1, l_2 = n_2, ..., l_m = n_m)\\ &= p2^{I(n_m>1)}\,\mathbb{P}(l_m = n_m)\cdot 3^{I(n_{m-1}>0)}2^{I(n_{m-1}>1)}\,\mathbb{P}(l_{m-1} = n_{m-1})\cdot ...\\ &\cdot 3^{I(n_1>0)}2^{I(n_1>1)}\,\mathbb{P}(l_1 = n_1).\end{aligned}$$

Заметим, что $l_1, l_2, ..., l_m$ – независимые одинаково распределенные случайные величины. Найдем распределение l_1

$$\mathbb{P}(l_1 = n_1) = p^{n_1}(\frac{1}{2})^{n_1-1}((1-p)^2 + 2\frac{1}{2}(1-p) + (\frac{1}{2})^2) = \frac{1}{2}(\frac{p}{2})^{n_1}(3-2p).$$

Таким образом, получим

$$\begin{aligned}&\mathbb{P}(l_1 = n_1, l_2 = n_2, ..., l_m = n_m)\\ &= p3^{\sum_{i=1}^m I(n_i>0)}2^{\sum_{i=1}^m I(n_i>1)}2^m(\frac{p}{2})^{\sum_{i=1}^m n_i}(1-\frac{p}{2})^{2\sum_{i=1}^m I(n_i>0)}(1-p)^{\sum_{i=1}^m I(n_i=0)}.\end{aligned} \tag{8}$$

Для события $\{\exists$ путь длины $k\}$ имеем

$$\mathbb{P}(\exists \text{ путь длины } k) = \sum_{n_1+n_2+...+n_m=k}\mathbb{P}(l_1 = n_1, l_2 = n_2, ..., l_m = n_m).$$

Разобъем множество $A = \{n_1, n_2, ..., n_m \ : \ n_1 + n_2 + ...n_m = k\}$ на такие подмножества, что на каждом из них индикаторные функции в выражении (8) принимают фиксированные значения. Обозначим через $Z_i^m = \{n_1, n_2, ..., n_m$: i слагаемых равны нулю и $n_1 + n_2 + ...n_m = k\} = \{n_1, n_2, ..., n_{m-i} \ : \ n_1 + n_2 + ... + n_{m-i} = k\}$, а также $O_i^m = \{n_1, n_2, ..., n_m$: j слагаемых равны 1 и $n_1 + n_2 + ... + n_m = k\} = \{n_1, n_2, ..., n_{m-i} \ : \ n_1 + n_2 + ... + n_{m-i} = m - i\}$. Тогда, при условии, что значение k и m достаточно велики и $k > m$, для события A имеем следующее разложение

$$A = \bigsqcup_{i=0}^{m-1}\bigsqcup_{j=0}^{m-i-1} A\cap Z_i^m \cap O_j^{m-i-1}.$$

Найдем мощность каждого подмножества в разложении. Для этого сперва заметим, что задача нахождения мощности множества $A\cap Z_i^m\cap O_j^{m-i-1}$ для некоторых i, j сводится к комбинаторной задаче о количестве вариантов распределить $k-2(m-1-i-j)-j$ шаров по $m-1-i-j$ ящикам. Таким образом получаем

$$\begin{aligned}|A\cap Z_i^m \cap O_j^{m-i-1}| &= C_{m-1}^i C_{m-1-i}^j C_{k-2(m-1-i-j)-j+m-1-i-j-1}^{m-1-i-j-1}\\ &= C_{m-1}^i C_{m-1-i}^j C_{k-m+i}^{m-i-j-2}.\end{aligned}$$

Теперь заметим, что на множестве $A\cap Z_i^m\cap O_j^{m-i-1}$ для некоторых i, j имеем

$$\mathbb{P}(l_1 = n_1, l_2 = n_2, ..., l_m = n_m)$$
$$= p3^{m-i}2^{m-i-j}2^m(\frac{p}{2})^k(1-\frac{p}{2})^{2(m-i)}(1-p)^i.$$

Итого для события $\{\exists$ путь длины $k\}$ получим

$$\mathbb{P}(\exists \text{ путь длины } k) = \sum_{n_1+n_2+...+n_m=k} \mathbb{P}(l_1 = n_1, l_2 = n_2, ..., l_m = n_m)$$
$$= \sum_{i=0}^{m-1}\sum_{j=0}^{m-i-1} C_{m-1}^i C_{m-1-i}^j C_{k-m+i}^{m-i-j-2} p3^{m-i}2^{m-i-j}2^m(\frac{p}{2})^k(1-\frac{p}{2})^{2(m-i)}(1-p)^i.$$

4.2 Пусть теперь в системе всего $k+1$ сервер. Обозначим h_t – максимальная длина пути на временном уровне t. Тогда ясно, что справедлива следующая оценка

$$H_n^i \leq \sum_{t=1}^n h_t \quad i = 0, 1, ..., k. \tag{9}$$

Легко понять, что $\{h_t\}_{t=0}^\infty$ – независимые одинаково распределенные случайные величины. Таким образом, чтобы получить оценку сверху для искомого коэффициента линейного роста, с помощью закона больших чисел, достаточно найти математическое ожидание h_1. Для этого воспользуемся следующей формулой

$$\mathbb{E}\, h_1 = \sum_{j=0}^k \mathbb{P}(h_1 \geq j).$$

Можно заметить, что

$$\mathbb{P}(h_1 \geq j) = \mathbb{P}(\{\exists \text{ путь длины } j\}).$$

Найдем вероятность события $\{\exists$ путь длины $j\}$.

$$\mathbb{P}(\{\exists \text{ путь длины } j\}) = 2(p/2)^{k+1} + 4(p/2)^k(1-p/2) + 6(p/2)^{k-1}(1-p/2)^2$$
$$+8(p/2)^{k-2}(1-p/2)^2 + ... + 2(k-j+2)(p/2)^j(1-p/2)^2$$
$$= 2(p/2)^{k+1} + 4(p/2)^k(1-p/2) + \sum_{i=2}^{k-j+1} 2(i+1)(p/2)^{k-i+1}(1-p/2)^2$$
$$= 2(p/2)^{k+1} + 4(p/2)^k(1-p/2) + 2(p/2)^{k+1}(1-p/2)^2 \sum_{i=2}^{k-j+1}(i+1)(p/2)^{-i}.$$

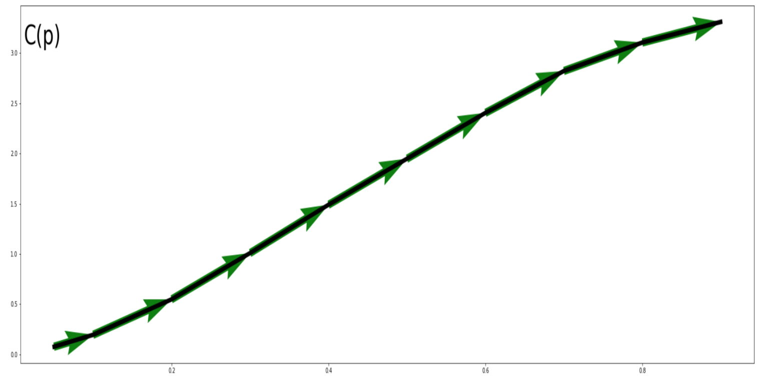
Посчитаем последнюю сумму отдельно.

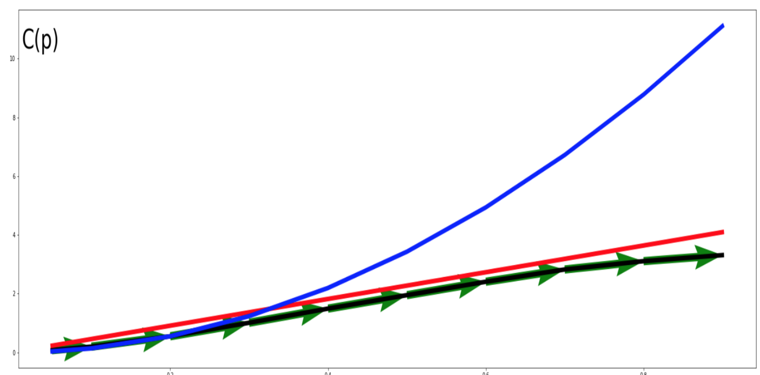
$$\sum_{i=2}^{k-j+1}(i+1)(p/2)^{-i} = (p/2)^{-1}\Big(\sum_{i=1}^{k-j} i(p/2)^{-i} + 2\sum_{i=1}^{k-j}(p/2)^{-i}\Big)$$
$$= (p/2)^{-1}\Big(\sum_{i=1}^{k-j}\sum_{s=i}^{k-j}(p/2)^{-s} + 2\sum_{i=1}^{k-j}(p/2)^{-i}\Big)$$
$$= (p/2)^{-1}\Big(\sum_{i=1}^{k-j}(p/2)^{-i}\frac{(1-(p/2)^{-(k-j+1-i)})}{(1-(p/2)^{-1})} + 2(p/2)^{-1}\frac{(1-(p/2)^{-(k-j)})}{(1-(p/2)^{-1})}\Big)$$
$$= \frac{(p/2)^{-1}}{(1-(p/2)^{-1})}\Big(\sum_{i=1}^{k-j}(p/2)^{-j} - \sum_{i=1}^{k-j}(p/2)^{-(k-j+1)} + 2(p/2)^{-1}(1-(p/2)^{-(k-j)})\Big)$$
$$= \frac{(p/2)^{-2}}{(1-(p/2)^{-1})}\Big(\frac{1-(p/2)^{-(k-j)}}{1-(p/2)^{-1}} - (k-j)(p/2)^{-(k-j)} + 2(1-(p/2)^{-(k-j)})\Big).$$

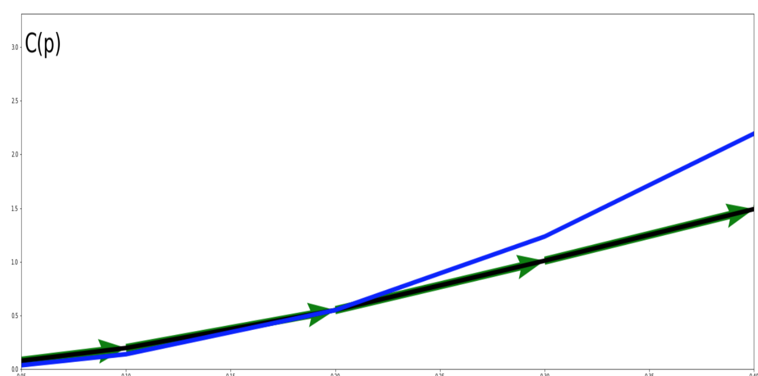
Таким образом, итого для распределения имеем

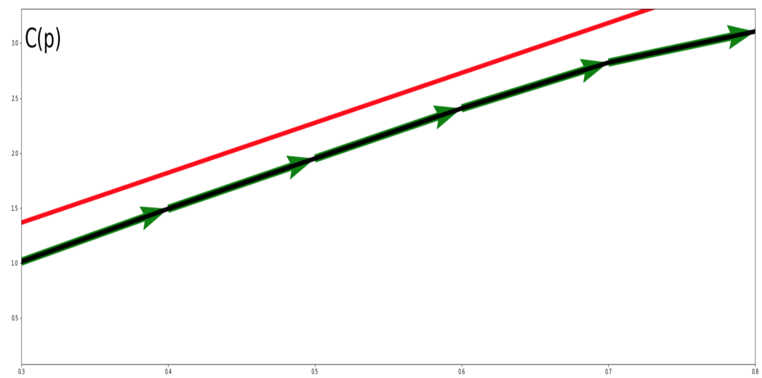
$$\mathbb{P}(\{\exists \text{ путь длины } j\}) = 2(p/2)^{k+1} + 4(p/2)^k(1-p/2) + 2(p/2)^{k+1}(1-p/2)^2\frac{(p/2)^{-2}}{(1-(p/2)^{-1})}$$
$$\cdot\Big(\frac{1-(p/2)^{-(k-j)}}{1-(p/2)^{-1}} - (k-j)(p/2)^{-(k-j)} + 2(1-(p/2)^{-(k-j)})\Big).$$

5. Для компьютерной симуляции использовались языки программирования python и c++. Задача состояла в реализации алгоритма вычисления H_n^0. Однако в силу ограничения компьютерных мощностей, вместо инициализации множества узлов $\{(i,t) : i \in \mathbb{Z},\ 0 \le t \le m\}$ для некоторого $m > 0$, производилась инициализация множества $\{(i,t) : i \in [-k,k],\ 0 \le t \le m\}$ для некоторого $k > 0$. В данной задаче Python используется для визуализации данных, однако его производительность сравнительно низкая, поэтому для вычисления алгоритма при больших значениях k и m используется c++, а python только вызывает скомпилированный код. Сперва рассмотрим графики средних значений на 100 реализациях функции $H_n^0(p)/n$ при фиксированном значении $n = 10^4$ и значениях $p = 0.05, 0.1, 0.2, ..., 0.9$. На графиках можно наблюдать синию параболу с коэффициентом ≈ 13.7, красную прямую с углом наклона ≈ 4.5 и упомянутые средние значения, соединенные зелеными стрелками. Можно заметить, что в начале график напоминает параболу, но затем становится больше похож на прямую. При том такую же зависимость можно наблюдать при фиксированных значениях $n = 5000, 2000$.

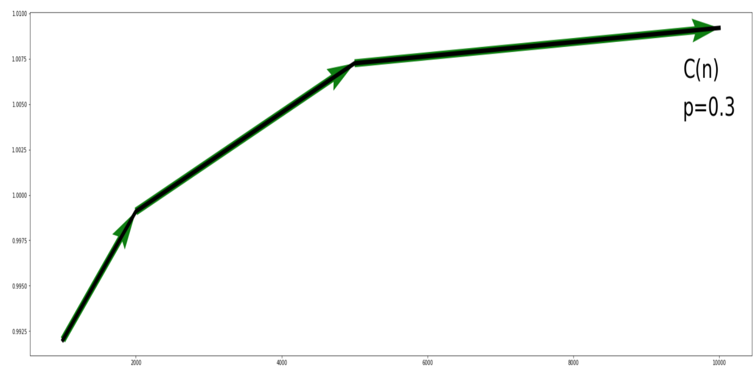

C(p)

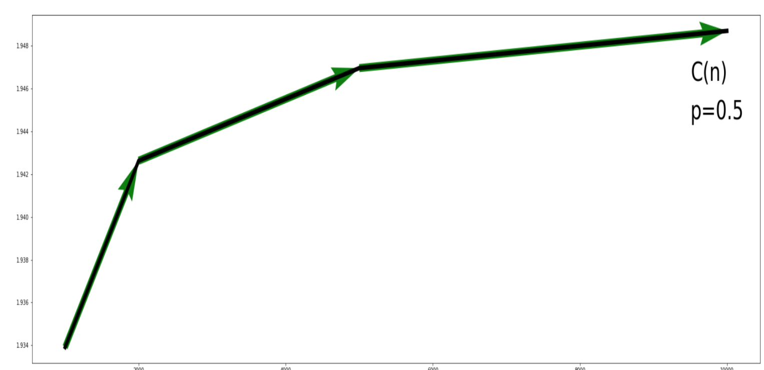

C(p)

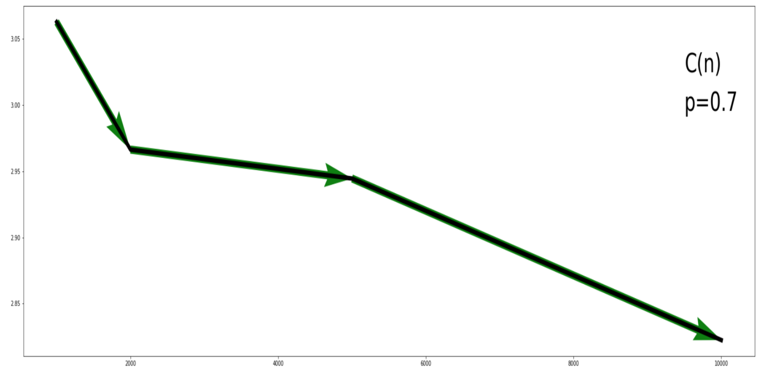

C(p)

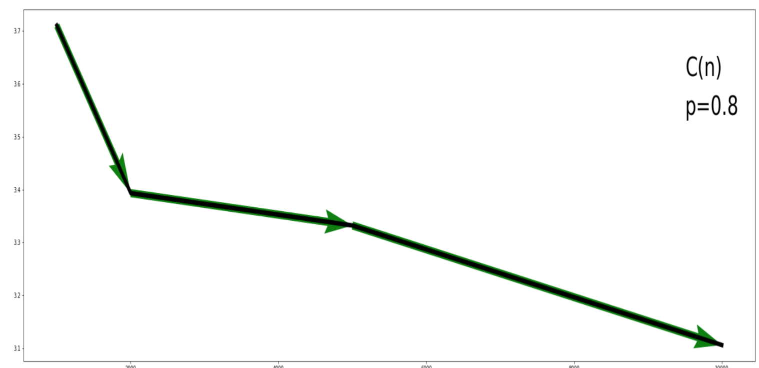


Теперь рассмотрим графики средних значений на 100 реализациях функции H_n^0/n при различных фиксированных значениях $p = 0.3, 0.5, 0.7, 0.8$ и переменных значениях $n = 1000, 2000, 5000, 10000$.






C(n)
p=0.7
3.05
3.00
2.95
2.90
2.85
2000
4000
6000
8000
10000


C(n)
p=0.8
3.7
3.6
3.5
3.4
3.3
3.2
3.1
2000
4000
6000
8000
10000

ALEXANDR VADIMOVICH REZLER
NOVOSIBIRSK STATE UNIVERSITY
2, PIROGOVA STR.
Email address: rezlers123@gmail.com

MIKHAIL GEORGIEVICH CHEBUNIN
KARLSRUHE INSTITUTE OF TECHNOLOGY,
INSTITUTE OF STOCHASTICS,
KARLSRUHE, 76131, GERMANY,
NOVOSIBIRSK STATE UNIVERSITY,
2, PIROGOVA STR.,
NOVOSIBIRSK, 630090, RUSSIA.
Email address: chebuninmikhail@gmail.com