

Проект по случайным графам

Чегодаева Таисия и Купряков Дмитрий, ПАДПИ, 2 курс

18 мая 2025 г.

Часть I

Исследование свойств
характеристики.

Глава 1

Исследовать, как ведет себя числовая характеристика τ в зависимости от параметров распределений θ и ν , зафиксировав размер выборки и параметр процедуры построения графа.

1.1 Характеристика τ^{KNN} .

1.1.1 Распределение SkewNormal с параметром α .

Зафиксируем размер выборки $n = 100$ и количество соседей $k = 5$. Число итераций для метода Монте-Карло равно 1000.

Будем перебирать $\theta = \{0.001, 0.01, 0.1, 0.5, 0.75, 1, 3, 5, 10, 15, 20, 50, 100, 500, 1000\}$.

Результаты

Усредненная характеристика τ^{KNN} при любых значениях параметра приближенно равна 9, но при больших значениях это приближение становится

более заметным.

1.1.2 Распределение Normal с параметром-дисперсией σ и матожиданием 0.

Зафиксируем размер выборки $n = 100$ и количество соседей $k = 5$. Число итераций для метода Монте-Карло равно 1000.

Будем перебирать $\nu = \{0.001, 0.01, 0.1, 0.5, 0.75, 1, 3, 5, 10, 15, 20, 50, 100, 500, 1000\}$.

Результаты

Усредненная характеристика τ^{KNN} принимает значения в окрестности числа 9 независимо от параметра ν . Но при больших значениях параметра можно заметить здесь, что характеристика τ^{KNN} начинает отклоняться от своего среднего значения.

1.2 Характеристика τ^{dist} .

1.2.1 Распределение SkewNormal с параметром α .

Зафиксируем размер выборки $n = 100$ и расстояние $dist = 1$. Число итераций для метода Монте-Карло равно 1000.

Будем перебирать

$\theta = \{0.001, 0.01, 0.1, 0.5, 0.75, 1, 3, 5, 10, 15, 20, 50, 100, 500, 1000, 10000, 150000, 300000, 500000, 15000000\}$.

Результаты.

Характеристика τ^{dist} при $\theta \in (0, 1)$ принимает в среднем значение 5, а при больших θ принимает значения, близкие к 3. Это хорошо видно на графике.

1.2.2 Распределение Normal с параметром-дисперсией σ и матожиданием 0.

Зафиксируем размер выборки $n = 100$ и расстояние $dist = 5$. Число итераций для метода Монте-Карло равно 1000.

Будем перебирать

$\nu = \{0.001, 0.01, 0.1, 0.5, 0.75, 1, 3, 5, 10, 15, 20, 50, 100, 500, 1000, 10000, 150000, 300000, 500000, 15000000\}$

Результаты

Характеристика τ^{dist} при $\nu \in (0, 0.5)$ принимает значение 1 (т.е. при таких ν граф – полный). С увеличением параметра растет среднее значение

характеристики (можно посмотреть [здесь](#)).

Глава 2

Исследовать, как ведет себя числовая характеристика τ в зависимости от параметров процедуры построения графа и размера выборки при фиксированных значениях $\theta = \theta_0$ и $\nu = \nu_0$.

2.1 Характеристика τ^{KNN} .

2.1.1 Распределение SkewNormal с параметром $\alpha_0 = 1$.

Будем перебирать параметры с 1000 итерациями метода Монте-Карло:

1. $n_samples = \{[1, 5, 10, 25, 50, 100, 300]\}$
2. $k_neighbours = \{1, 3, 5, 7, 9, 15, 20\}$

Результаты

Можно заметить, что средняя величина характеристики τ^{KNN} увеличивается, по мере роста перебираемых параметров. Но также часто встречаются ситуация, когда среднее значение совпадает с реальным.

2.1.2 Распределение Normal с параметром-дисперсией $\sigma_0 = 1$ и матожиданием 0.

Будем перебирать параметры с 1000 итерациями метода Монтэ-Карло:

1. $n_samples = \{[1, 5, 10, 25, 50, 100, 300]\}$
2. $k_neighbours = \{1, 3, 5, 7, 9, 15, 20\}$

Результаты

Можем наблюдать такую же тенденцию – с ростом параметров растет среднее значение характеристики, даже значения принимаются такие же со сдвигом на небольшой ϵ .

2.2 Характеристика τ^{dist} .

2.2.1 Распределение SkewNormal с параметром $\alpha_0 = 1$.

Будем перебирать параметры с 1000 итерациями метода Монтэ-Карло:

1. $n_samples = \{[1, 5, 10, 25, 50, 100, 300]\}$
2. $dist = \{0.001, 0.01, 0.1, 0.5, 1, 3, 5\}$

Результаты

Можно заметить, что больше всего на значение характеристики τ^{dist} влияет параметр $n_samples$, а с увеличением параметра $dist$ увеличивается количество ребер из-за этого уменьшается количество независимых вершин.

2.2.2 Распределение Normal с параметром-дисперсией $\sigma_0 = 1$ и матожиданием 0.

Будем перебирать параметры с 1000 итерациями метода Монтэ-Карло:

1. $n_samples = \{[1, 5, 10, 25, 50, 100, 300]\}$
2. $k_neighbours = \{1, 3, 5, 7, 9, 15, 20\}$

Результаты

Для каждого значения параметра $n_samples$ можем заметить довольно плотное распределение среднего значения характеристики τ^{dist} , но с ростом этого параметра растет количество выбросов и колебания.