

Школа лингвистики, 2019-20 уч. год**Теория вероятностей****Доверительные интервалы (20.04.2020)***И. В. Щуров, Д. А. Филимонов, Р. Я. Будылин*

Задачи основаны на книге Statistics, David Freedman, Robert Pisani, Roger Purves (fourth edition).

Задача 1. Частота встречаемости фиксированного слова в кусках текста известной длины распределена по нормальному закону. Для нахождения параметров этого распределения были подсчитаны частоты для нескольких случайных текстов и получены следующие значения: 0.42; 0.62; 0.42; 0.65. Вы работаете с корпусом **хорошо** изученного языка и известно, что для достаточно употребимых слов $\sigma^2 = 0.2$

- Построить доверительный интервал для параметра μ на уровне доверия 95%.
- То же на уровне доверия 99%.
- Сколько элементов нужно выбрать, чтобы уменьшить длину доверительного интервала в 3 раза?

Задача 2. Частота встречаемости фиксированного слова в кусках текста известной длины распределена по нормальному закону. Для нахождения параметров этого распределения были подсчитаны частоты для нескольких случайных текстов и получены следующие значения: 0.42; 0.62; 0.42; 0.65. Вы работаете с корпусом **мало** изученного языка и про σ ничего не известно.

- Построить доверительный интервал для параметра μ на уровне доверия 95%.
- То же на уровне доверия 99%.
- Сколько элементов нужно выбрать, чтобы уменьшить длину доверительного интервала в 3 раза?

Задача 3. Из большого корпуса выбирается 100 элементов. Среди них 30 обладает свойством X.

- Построить доверительный интервал для доли всех элементов в корпусе, обладающих свойством X, на уровне доверия 95%.
- То же на уровне доверия 99%.
- Сколько элементов нужно выбрать, чтобы уменьшить длину доверительного интервала в 3 раза?

Задача 4. Монетку подкинули 50 раз. Из них 25 раз она выпала орлом вверх. Построить доверительный интервал для вероятности выпадения орла в одном бросании на уровне доверия 95%.

Задача 5. В некотором городе проживает 10 000 человек в возрасте от 18 до 24 лет. Из них случайным образом выбрали и опросили 500 человек. Среди опрошенных 194 человека сообщили, что учатся в вузах.

- Построить доверительный интервал для процента студентов в городе среди всех людей в возрасте от 18 до 24 лет на уровне доверия 95%.
- То же, что в пункте (а), но на уровне доверия 99%.
- То же, что в пункте (б), но в предположении, что в опросе участвовало 2,000 человек, из которых 776 учится в вузах.
- Сколько человек нужно опросить, чтобы длина доверительного интервала на уровне доверия 95% составляла не больше 2 процентных пунктов.