«КАЗАНСКИЙ (ПРИВОЛЖСКИЙ) ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ»

Лабораторная работа №7 по теме: «Построение и исследование доверительных интервалов для математического ожидания нормальной с.в.

Выполнил

Студент 1 курса

группы 09-115(3)

Зиновьев Е. А.

Преподаватель:

Шустова Е.П

Казань 2021

План работы:

1. Придумайте с.в., распределение которой можно считать близким к нормальному (например, объём продаж продукции по месяцам, вес детали, ошибка измерения некоторой величины).
2. Получите выборку из распределения выбранной с.в. (объём выборки и параметры распределения задайте сами).
3. Постройте 3 различных доверительных интервала для математического ожидания с.в.; выясните, как соотносятся полученные интервалы, объясните различия;
4. Последовательно полагая объём выборки n равным 100, 500, 1000 и 10000, сгенерируйте выборки соответствующих объёмов (не меняя значения параметров распределения), постройте 95% доверительный интервал для математического ожидания с.в.;выясните, как соотносятся полученные интервалы, объясните различия.

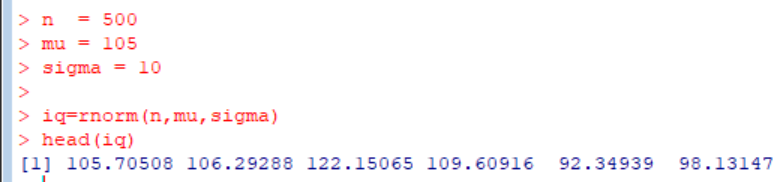
Ход работы:

* 1. Рассмотрим случайную величину, распределение которой можно считать близким к нормальному: количество баллов теста на IQ у людей:

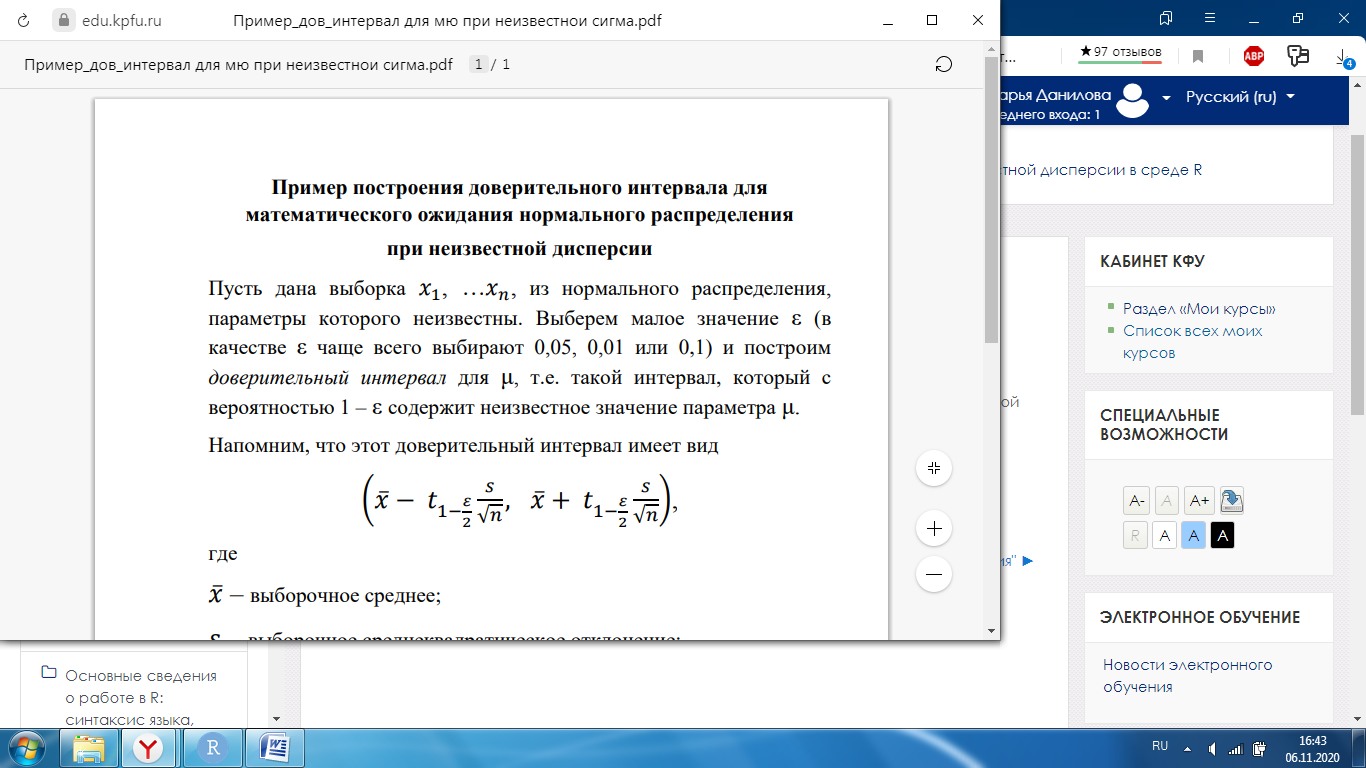
<https://brainapps.ru/blog/2015/08/urovni-znacheniy-iq-i-ikh-rasshifrovka/>

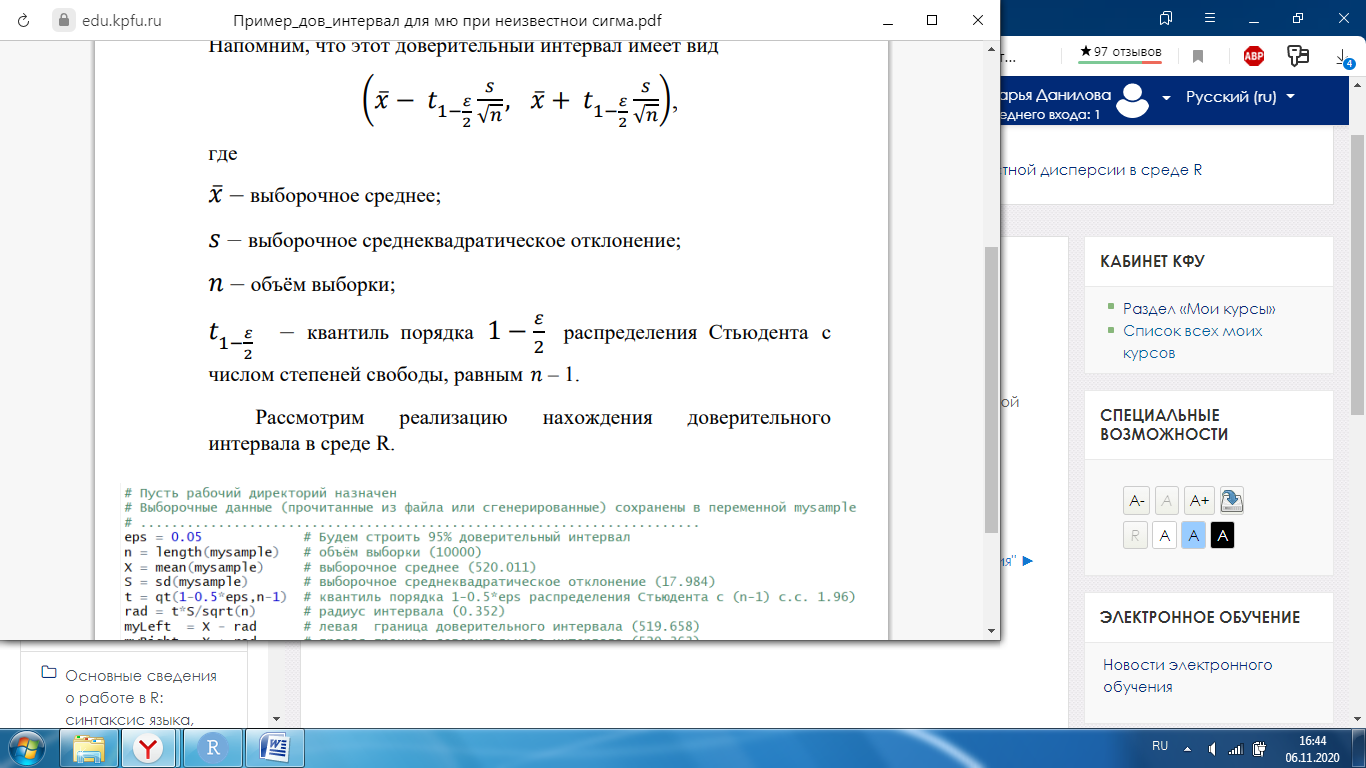
* 1. Объем выборки: 500 человек, µ (среднее IQ человека) = 105, σ (среднеквадратическое отклонение IQ человека) = 10.

Смоделируем нашу выборку:



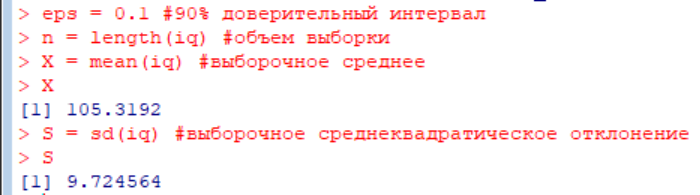
Для построения доверительного интервала нам необходимо знать:





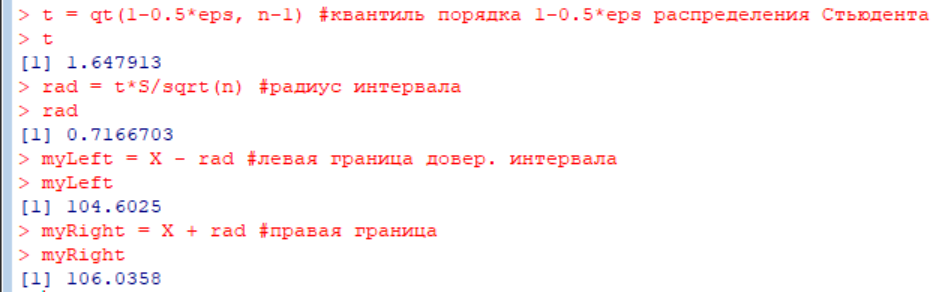
3.1 Построим 90% доверительный интервал для нашей выборки.

Найдем среднее и среднеквадратическое отклонение нашей выборки:



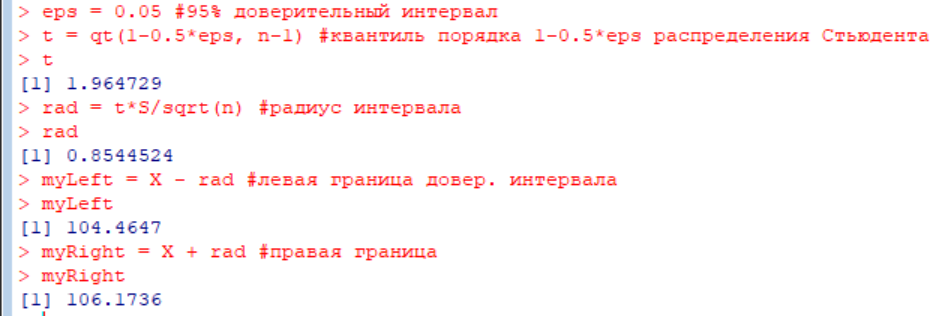
При генерации выборки среднее отклонилось меньше, чем на 0.3% от заданного, а среднеквадратическое отклонение почти на 3%.

Найдем квантиль порядка 1-eps распределения Стьюдента, из него вычислим радиус доверительного интервала и значения правой и левой границ:



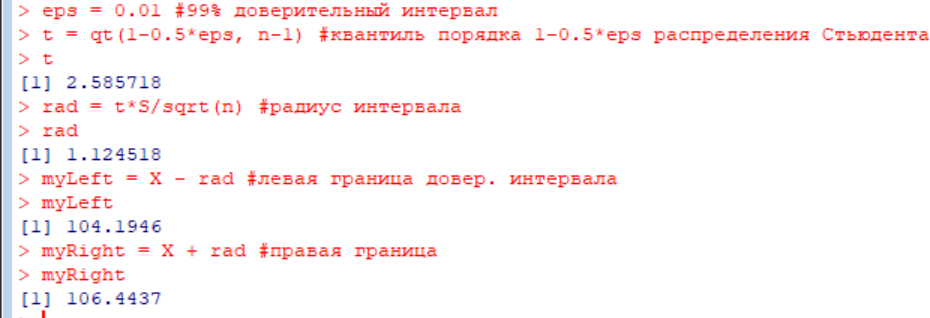
Доверительный интервал составил 15% от среднеквадратического отклонения и равен [104.6025, 106.0358]

3.2 Построим 95% доверительный интервал для нашей выборки:



Доверительный интервал повысился на 2% относительно 90% интервала и составил 17% от среднеквадратического отклонения и равен [104.4647, 106.1736]

3.3 Построим 99% доверительный интервал для нашей выборки:



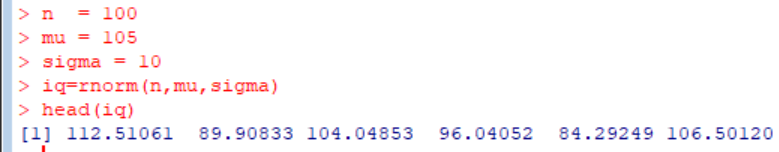
Доверительный интервал повысился на 6% относительно 95% интервала и составил 23% от среднеквадратического отклонения и равен [104.1946, 106.4437]

Взглянем на полученные результаты:

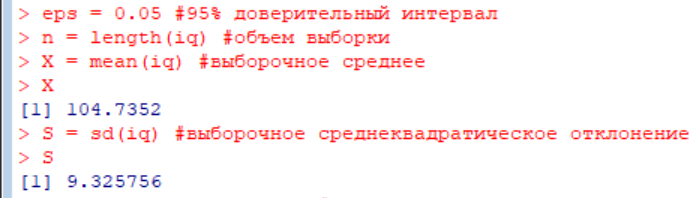
|  |  |  |
| --- | --- | --- |
|  | Левая граница | Правая граница: |
| 90%-ый доверительный интервал | 104.6025 | 106.0358 |
| 95%-ый доверительный интервал | 104.4647 | 106.1737 |
| 99%-ый доверительный интервал | 104.1946 | 106.4437 |

Видим, что в зависимости от уровня доверия (при его увеличении), наши границы смещаются, становятся «шире». Доверительный интервал занял 15,17,23% значений одного среднеквадратического отклонения для 90,95,99% уровня доверия соответственно. Это говорит о том, что для большей уверенности включения среднего значения генеральной совокупности необходимо увеличить доверительный интервал.

4.1 Полагаем объём выборки n равным 100, сгенерируем выборку соответствующего объёма (не меняя значения параметров распределения)

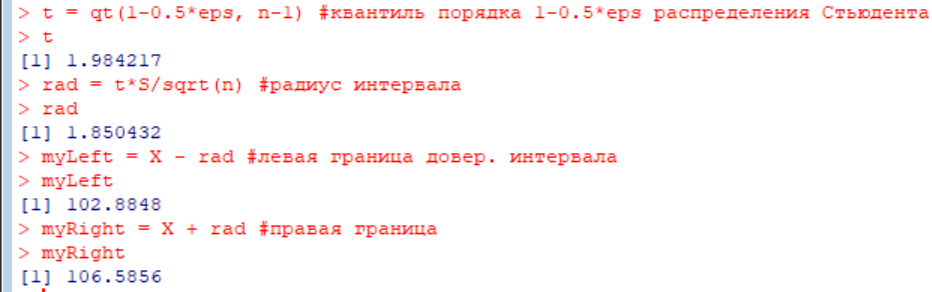


Найдем среднее и среднеквадратическое отклонение нашей выборки:



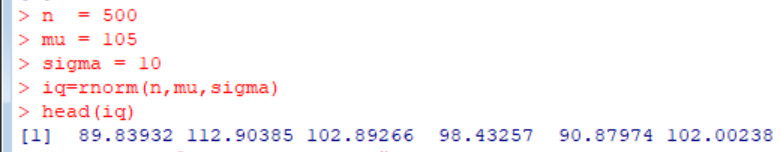
При генерации выборки среднее отклонилось меньше, чем на 0.25% от заданного, а среднеквадратическое отклонение на 7%.

Построим 95% доверительный интервал для математического ожидания с.в.:

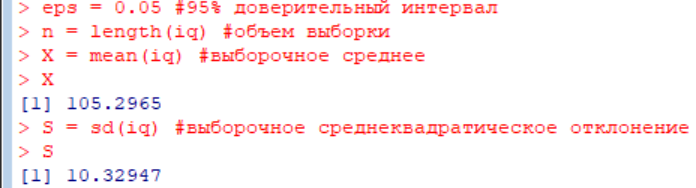


Доверительный интервал составил 37% от среднеквадратического отклонения и равен [102.8848, 106.5856]

4.2 Полагаем объём выборки n равным 500, сгенерируем выборку соответствующего объёма (не меняя значения параметров распределения)

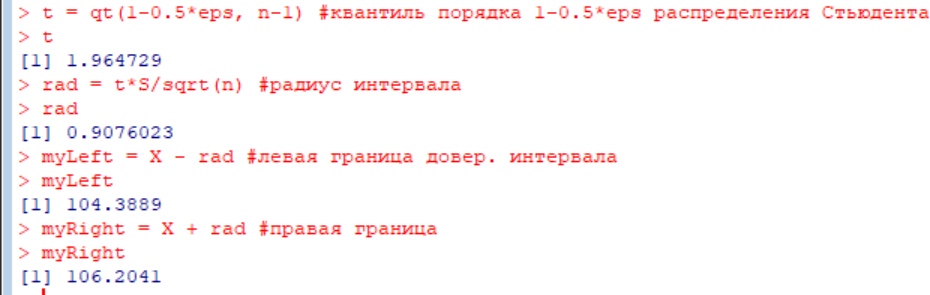


Найдем среднее и среднеквадратическое отклонение нашей выборки:



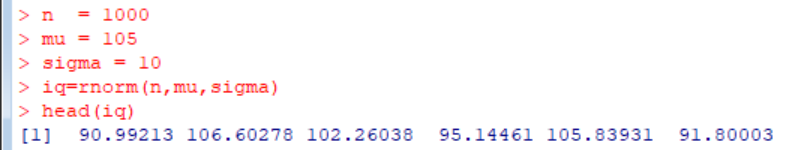
При генерации выборки среднее отклонилось меньше, чем на 0.3% от заданного, а среднеквадратическое отклонение на 3.3%, то есть вдвое уменьшилось относительно прошлой выборки.

Построим 95% доверительный интервал для математического ожидания с.в.:

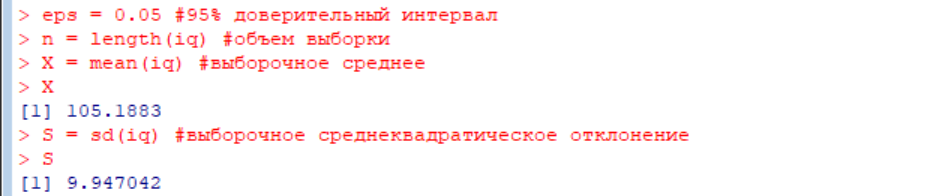


Доверительный интервал составил 18% от среднеквадратического отклонения и равен [104.3889, 106.2041], то есть также уменьшился вдвое относительно прошлой выборки.

4.3 Полагаем объём выборки n равным 1000, сгенерируем выборку соответствующего объёма (не меняя значения параметров распределения)

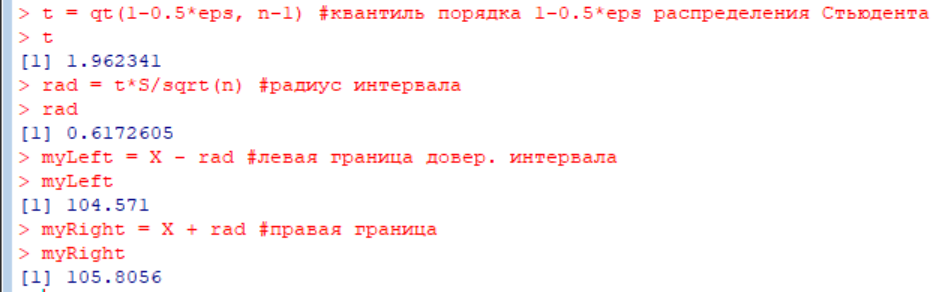


Найдем среднее и среднеквадратическое отклонение нашей выборки:



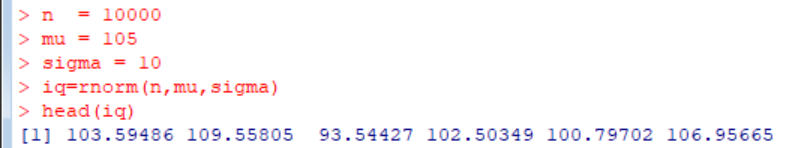
При генерации выборки среднее отклонилось меньше, чем на 0.2% от заданного, а среднеквадратическое отклонение на 0.6%, то есть в 5.5 раз уменьшилось относительно прошлой выборки.

Построим 95% доверительный интервал для математического ожидания с.в.:

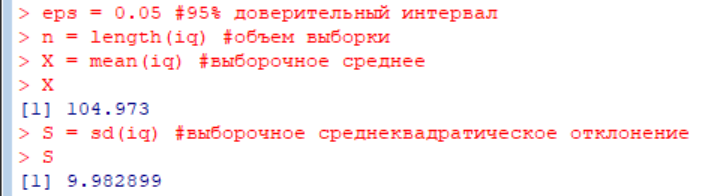


Доверительный интервал составил 12.4% от среднеквадратического отклонения и равен [104.571, 105.8056], то есть уменьшился в 1.5 раза относительно прошлой выборки.

4.4 Полагаем объём выборки n равным 10000, сгенерируем выборку соответствующего объёма (не меняя значения параметров распределения)

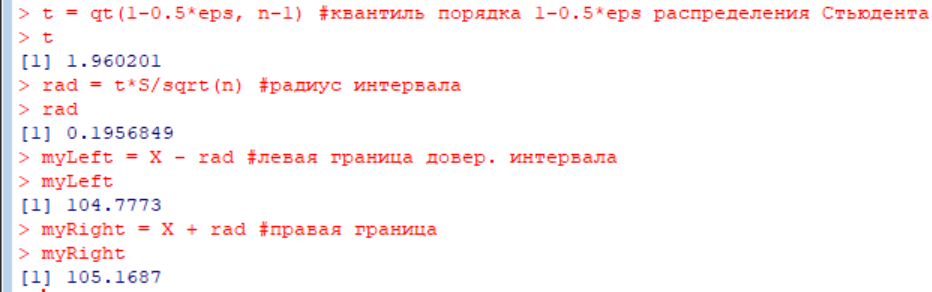


Найдем среднее и среднеквадратическое отклонение нашей выборки:



При генерации выборки среднее отклонилось меньше, чем на 0.03% от заданного, а среднеквадратическое отклонение на 0.2%, то есть в 3 раза уменьшилось относительно прошлой выборки.

Построим 95% доверительный интервал для математического ожидания с.в.:



Доверительный интервал составил 2% от среднеквадратического отклонения и равен [104.7773, 105.1687], то есть уменьшился в 6 раз относительно прошлой выборки.

Посмотрим табличку с полученными результатами:

|  |  |  |
| --- | --- | --- |
|  | Левая граница | Правая граница |
| Объем выборки = 100 | 102.8848 | 106.5856 |
| Объем выборки = 500 | 104.3889 | 106.2041 |
| Объем выборки = 1000 | 104.571 | 105.8056 |
| Объем выборки = 10000 | 104.7773 | 105.1687 |

Видим, что при увеличении объема выборки границы смещаются, становятся более «узкими» при одном и том же уровне доверия 95%. То есть с увеличением выборки вероятность попадания истинного среднего значения генеральной совокупности в интервал повышается (если, например, закрепить один доверительный интервал для 4 ранее рассмотренных случаев, то вероятность попадания среднего значения генеральной совокупности будет выше при выборке в 10 000 элементов, чем при 1000 элементов).