*Цель занятия* – освоить вычисление оптимальных *L*-оценок по выборочным квантилям, проследить, как влияет на точность оценок выбор числа интервалов.

# Этапы исследования

1. Смоделировать выборку в соответствии с нормальным законом объемом $n = 1000$. Внести её в таблицу **Excel**, отсортировать по возрастанию. Далее опираться на рекомендации
2. Предполагая, что выборка принадлежит нормальному закону, найти оптимальные $L$*-оценки* (обоих) параметров закона. Для этого выбрать из соответствующей таблицы **АОГ** оптимальные вероятности при необходимом числе интервалов $k$. В соответствии с этими вероятностями найти оценки выборочных квантилей, разбивающие выборку на части, пропорциональные данным вероятностям. Выбрать из соответствующей таблицы коэффициенты, необходимые для вычисления оптимальных $L$*-оценок*. Вычислить оптимальные $L$*-оценки* как соответствующие линейные комбинации.
   1. Найти оценки при $k = 4, 5, 8, 10$.
   2. Сравнить полученные оценки с ОМП (при вычислении в ISW).
   3. Предполагая, что Вы нашли оценки по некоторой другой выборке, проверьте простую гипотезу о согласии с нормальным законом со значениями параметров, полученными при $k = 10$.
3. Предполагая, что выборка принадлежит логистическому закону[1], выполнить ту же последовательность действий при вычислении оптимальных $L$*-оценок* параметров этого закона, ограничившись $k = 10$.
4. Смоделировать выборку в соответствии с распределением Коши объемом $n = 1000$. Вычислить оптимальные $L$*-оценки* параметров этого закона при $k = 10$. Сравнить с ОМП. Проверить “**простую**” гипотезу о согласии с данным распределением *Коши*.
5. Предполагая, что выборка принадлежит нормальному закону, выполнить ту же последовательность действий при вычислении оптимальных $L$*-оценок* параметров нормального закона, так же ограничившись $k = 10$.
6. Кратко сформулируйте для себя выводы, вытекающие из ваших результатов.

# Выполненные исследования

## Моделирование выборки нормального распределения

Эмпирическая функция распределения *(см. график 1)* выборки *model_1*, смоделированная в соответствии $N(\sigma = 4, \mu = 0)$ со с объёмом $n = 1000$

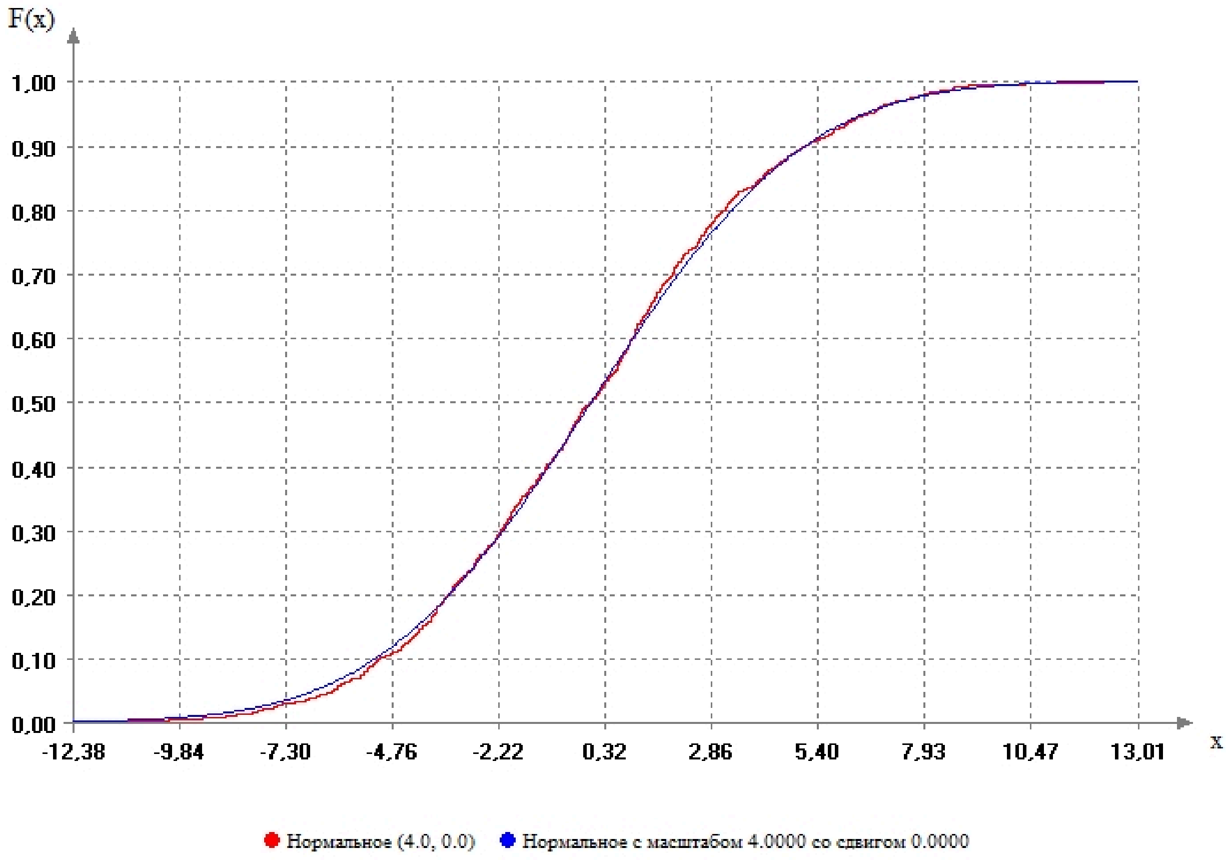



*График 1* Эмпирическая функция распределения выборки *model_1*

Функция плотности распределения *(см. график 2)* выборки *model_1*, смоделированная в соответствии $N(\sigma = 4, \mu = 0)$ со с объёмом $n = 1000$

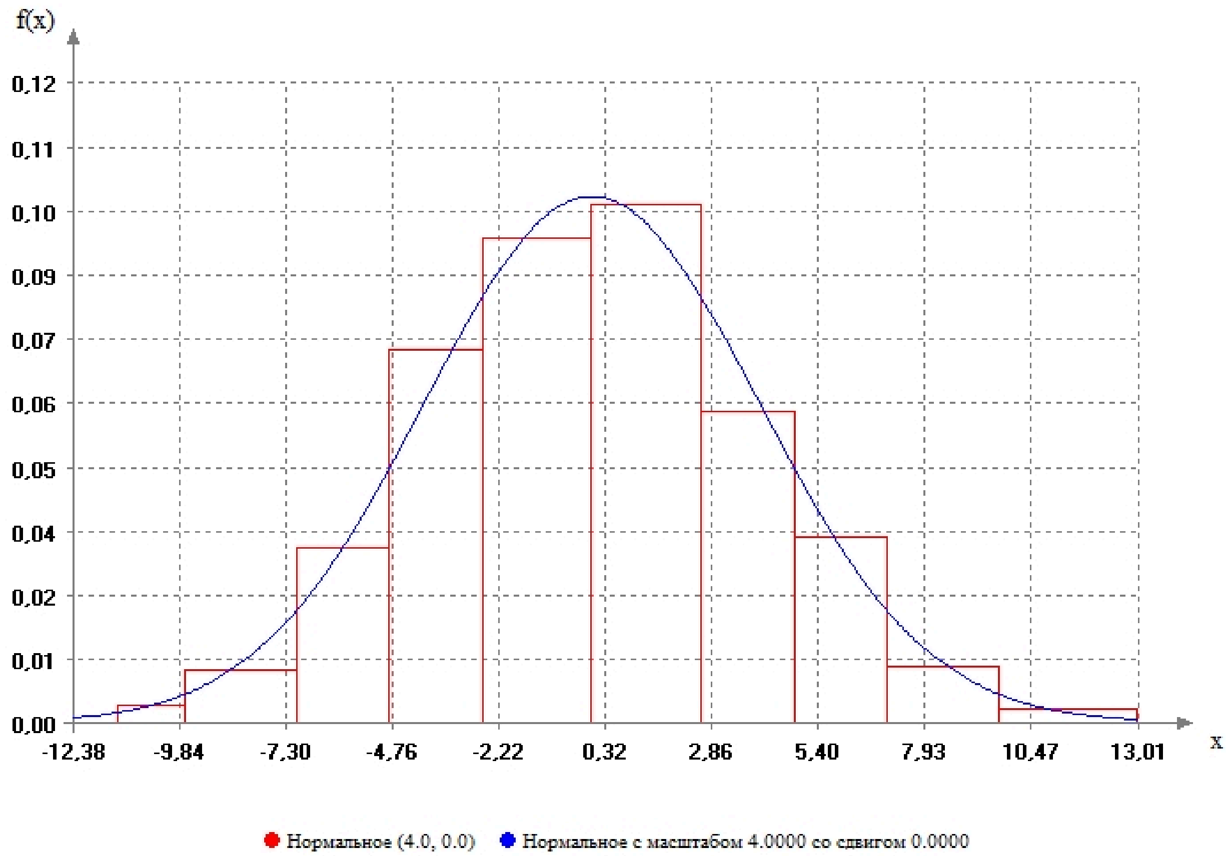


*График 2* Функция плотности распределения выборки *model_1*

# Получение L-оценки для нормального закона

## Получение оптимальных вероятностей для $k = 4, 5, 8, 10$

Для получения оптимальных интервалов, получим оптимальные вероятности при необходимом числе интервалов $k$ при оценивании двух параметров. Необходимые вероятности можно получить из таблицы А.29.

| $k$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ | $P_7$ | $P_8$ | $P_9$ | $P_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| 4 | 0.0833 | 0.4167 | 0.4167 | 0.0833 | - | - | - | - | - | - |
| 5 | 0.0449 | 0.2004 | 0.5094 | 0.2004 | 0.0449 | - | - | - | - | - |
| 8 | 0.0141 | 0.0587 | 0.1431 | 0.2841 | 0.2841 | 0.1431 | 0.0587 | 0.0141 | - | - |
| 10 | 0.0077 | 0.0317 | 0.0748 | 0.1438 | 0.2420 | 0.2420 | 0.1438 | 0.0748 | 0.0317 | 0.0077 |

Получим выборочные квантили выборки **model_1**, сформированного в прошлом пункте по оптимальным вероятностям. Сформируем таблицу значений границ интервалов.

| $k$ | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | $I$ |
|---|---|---|---|---|---|---|---|---|---|
| 4 | $-11.3023$ | $-5.3176$ | 0.0005 | 5.6716 | 12.9968 | - | - | - | |
| 5 | $-11.3023$ | $-6.2126$ | $-2.7955$ | 2.5971 | 6.7996 | 12.9968 | - | - | |
| 8 | $-11.3023$ | $-8.0621$ | $-5.4886$ | $-3.2206$ | 0.0005 | 2.9323 | 5.9502 | 8.4333 | 12.9 |
| 10 | $-11.3023$ | $-8.7633$ | $-6.5387$ | $-4.5537$ | $-2.6720$ | 0.0005 | 2.4607 | 4.8169 | 6.9 |

## Сравнение полученных результатов с ISW

Получим значения границ интервалов используя средства ISW.

| $k$ | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | $I$ |
|---|---|---|---|---|---|---|---|---|---|
| 4 | $-11.3023$ | $-5.5336$ | 0.0000 | 5.5336 | 12.9968 | - | - | - | |
| 5 | $-11.3023$ | $-6.7844$ | $-2.7576$ | 2.7576 | 6.7844 | 12.9968 | - | - | |
| 8 | $-11.3023$ | $-8.7816$ | $-5.8208$ | $-3.1452$ | 0.0000 | 3.1452 | 5.8208 | 8.7816 | 12.9 |
| 10 | $-11.3023$ | $-9.6900$ | $-7.0312$ | $-4.8184$ | $-2.5988$ | 0.0000 | 2.5988 | 4.8184 | 7.0 |

Из полученных результатов можно сделать вывод, что значения похожи, но отличны друг от друга. Дело в том, что в ISW квантили находятся по плотности модели $N(\sigma = 4, \mu = 0)$, а в нашем случае квантили выборочные т.е. получены из выборки *model_1*. Если же получить значения границ интервалов по квантилям модели, то можно заметить схожесть.

Убедится в том, что ISW использует квантили модели можно исходя из значений. Как видно при чётном количестве интервалов можно увидеть математическое ожидание, равное во всех случаях 0, что нельзя сказать к значениям полученным по выборочным квантилям.

## Проверка простой гипотезы

Результаты проверки **простой гипотезы** о согласии, используя критерий $\chi^2$ *Пирсона* для параметрической модели $N(\sigma = 4, \mu = 0)$, на группированной выборке $k = 10$ , граничные точки которой были получены ранее.

| **Критерий** | **Значение** $S$ | **Значение** $P$ |
|---|---|---|
| $\chi^2$ Пирсона | $8.4459$ | $0.4899$ |

# Получение L-оценки для логистичекого закона $k = 10$

Получим оценки для Логистического распределения с параметрами $\sigma = 4/\sqrt{3} \approx 2.309491, \mu = 0$. Для этого воспользуемся соответствующей [таблицей А45](#) для получение оптимальных вероятностей при $k = 10$.

| $k$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ | $P_7$ | $P_8$ | $P_9$ | $P_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $10$ | $0.0153$ | $0.0510$ | $0.0946$ | $0.1441$ | $0.1950$ | $0.1950$ | $0.1441$ | $0.0946$ | $0.0510$ | $0.015$ |

Получим выборочные квантили выборки **model_1**, сформированного в пункте №1 по оптимальным вероятностям. Сформируем таблицу значений границ интервалов.

| $k$ | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | $I_9$ |
|---|---|---|---|---|---|---|---|---|---|
| $10$ | $-11.3023$ | $-8.0327$ | $-5.7182$ | $-3.8195$ | $-2.0748$ | $0.0005$ | $1.8819$ | $3.9081$ | $6.10$ |

## Сравнение полученных результатов с ISW

Получим значения границ интервалов используя средства ISW.

| $k$ | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | $I_9$ |
|---|---|---|---|---|---|---|---|---|---|
| $10$ | $-11.3023$ | $-9.6225$ | $-6.1068$ | $-3.8134$ | $-1.9019$ | $0.0000$ | $1.9019$ | $3.8134$ | $6.10$ |

Аналогично прошлому пункту, можем наблюдать различия между полученными значениями границ интервалов.

# Проверка простой гипотезы

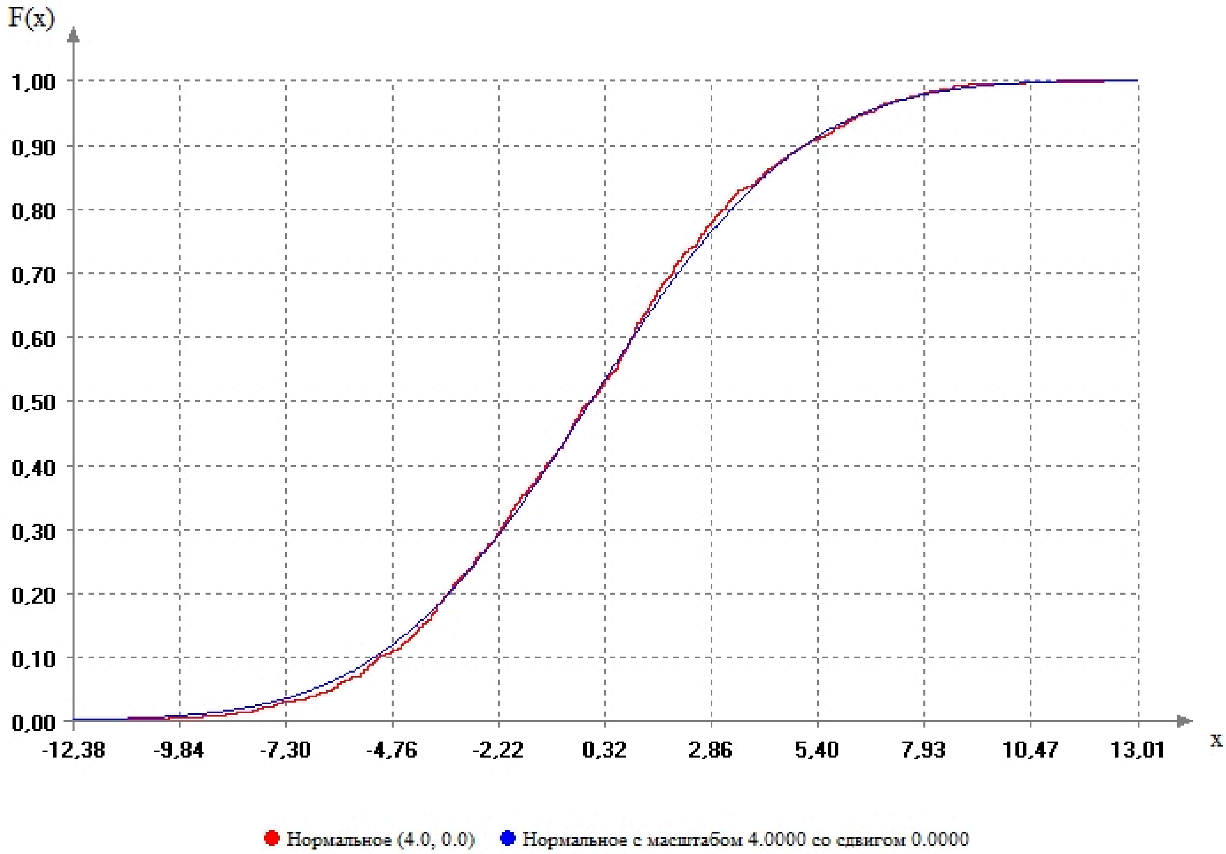
Результаты проверки **простой гипотезы** о согласии, используя критерий $\chi^2$ *Пирсона* для параметрической модели $Logist(\sigma \approx 2.309491, \mu = 0)$, на группированной выборке $k = 10$ , граничные точки которой были получены ранее.

| Критерий | Значение $S$ | Значение $P$ |
|---|---|---|
| $\chi^2$ Пирсона | $19.1281$ | $0.02412$ |

Из результата, видно, что значение p-value критерия $\chi^2$ Пирсона при проверке простой гипотезы $H_1$ значительно ниже значения p-value полученного при проверке гипотезы $H_0$ .

# Моделирование выборки Коши

Эмпирическая функция распределения *(см. график 2)* выборки _model_2, смоделированная в соответствии $C(x_0 = 0, \gamma = 3)$ со с объёмом $n = 1000$




*<u>График 3</u>* Эмпирическая функция распределения выборки *<u>model_2</u>*

Функция плотности распределения *(см. график 4)* выборки _model_2, смоделированная в соответствии $C(x_0 = 0, \gamma = 3)$ со с объёмом $n = 1000$

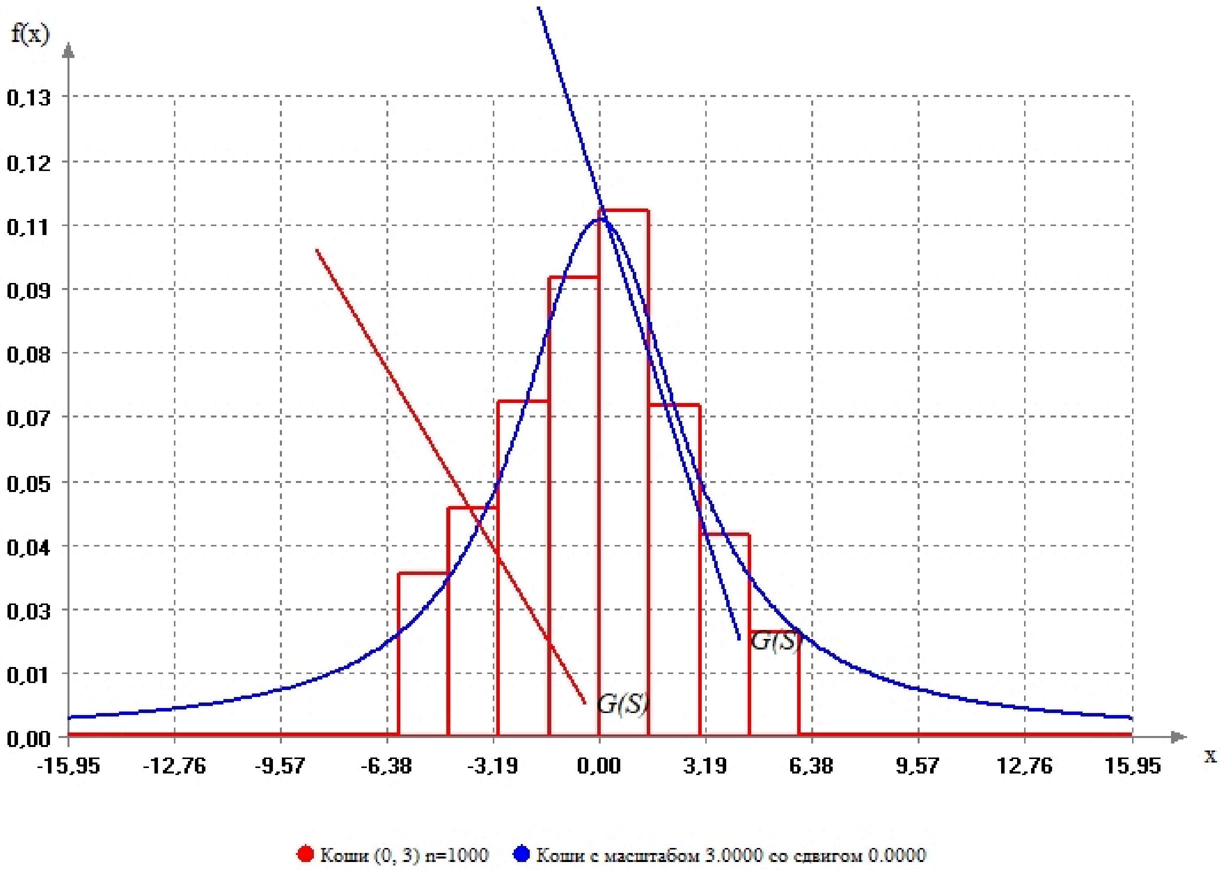



*График 4* Функция плотности распределения выборки *model_2*

# Вычисление L-оценок для закона Коши

## Получение оптимальных вероятностей для $k = 10$

Получим оценки для распределения Коши с параметрами $x_0 = 0.0, \gamma = 3$. Исходя из [источника](#) для получения $L$-оценки требуется использовать равновероятный способ группирования. Для этого воспользуемся ISW. Сформируем таблицу значений границ интервалов.

| $k$ | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | $I_9$ |
|---|---|---|---|---|---|---|---|---|---|
| 10 | $-404.6613$ | $-9.2331$ | $-4.1291$ | $-2.1796$ | $-0.9747$ | $0.0000$ | $0.9747$ | $2.1796$ | $4.12$ |

## Сравнение полученных результатов с ISW

Получим значения границ интервалов используя средства ISW.

| $k$ | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | $I_9$ |
|---|---|---|---|---|---|---|---|---|---|
| 10 | $-404.6613$ | $-9.2334$ | $-4.1292$ | $-2.1798$ | $-0.9747$ | $0.0000$ | $0.9747$ | $2.1798$ | $4.12$ |

При сравнении можно заметить, что значения крайне похожи друг на друга. Разница заметна в 4 знаке после запятой.

## Проверка простой гипотезы

Результаты проверки **простой гипотезы** о согласии, используя критерий $\chi^2$ *Пирсона* для параметрической модели $C(x_0 = 0, \gamma = 3)$, на группированной выборке $k = 10$ , граничные точки которой были получены ранее.

| Критерий | Значение $S$ | Значение $P$ |
|---|---|---|
| $\chi^2$ Пирсона | $5.9200$ | $0.7478$ |

# Получение L-оценки для нормального закона

## Получение оптимальных вероятностей для $k = 10$

Получим оценки для нормального закона распределения с параметрами $\mu = 0.0, \sigma = 5.53445$. Используем уже ранее показанную таблицу для получение оптимальных вероятностей при $k = 10$

| $k$ | $P_1$ | $P_2$ | $P_3$ | $P_4$ | $P_5$ | $P_6$ | $P_7$ | $P_8$ | $P_9$ | $P_{10}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $10$ | $0.0077$ | $0.0317$ | $0.0748$ | $0.1438$ | $0.2420$ | $0.2420$ | $0.1438$ | $0.0748$ | $0.0317$ | $0.0077$ |

Получим выборочные квантили выборки **model_2**, сформированного в прошлом пункте по оптимальным вероятностям. Сформируем таблицу значений границ интервалов.

| $k$ | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | |
|---|---|---|---|---|---|---|---|---|---|
| $10$ | $-404.6613$ | $-67.1396$ | $-18.5429$ | $-7.0866$ | $-2.9612$ | $0.0005$ | $2.6579$ | $7.9962$ | 2 |

## Сравнение полученных результатов с ISW

Получим значения границ интервалов используя средства ISW.

| $k$ | $I_1$ | $I_2$ | $I_3$ | $I_4$ | $I_5$ | $I_6$ | $I_7$ | $I_8$ | |
|---|---|---|---|---|---|---|---|---|---|
| $10$ | $-404.6613$ | $-13.4072$ | $-9.7285$ | $-6.6668$ | $-3.5957$ | $0.0000$ | $3.5957$ | $6.6668$ | 9.7 |

Можно заметить существенное отличие, пояснение было ранее. Здесь аналогично.

## Проверка простой гипотезы

Результаты проверки **простой гипотезы** о согласии, используя критерий $\chi^2$ *Пирсона* для параметрической модели $Logist(\sigma \approx 2.309491, \mu = 0)$, на группированной выборке $k = 10$ , граничные точки которой были получены ранее.

| Критерий | Значение $S$ | Значение $P$ |
|---|---|---|
| $\chi^2$ Пирсона | $64633.501$ | $0$ |

Вполне ожидаемый результат. Нормальный закон распределения не похож на закон Коши. И в данном случае это показывается.

# Вывод

По итогу работы удалось освоить вычисление оптимальных *L*-оценок по выборочным квантилям. По итогу работы удалось показать, что при данном способе, критерий $\chi^2$ Пирсона способен отличать близкие конкурирующие гипотезы. Удалось показать, отличие $L$-оценок и ОМП.

---

1. Указание: Логистический закон в **ISW** представлен несколько в другом виде, чем рассматриваемый при построении $L$-*оценок* этого закона. Поэтому полученную $L$-*оценку* параметра масштаба необходимо разделить на $\sqrt{3}$ ↩