Цель исследования

Оценить возможности некоторых критериев, предназначенных для проверки статистических гипотез о случайности данных и отсутствии в анализируемых выборках признаков **тренда**. Проследить, как в зависимости от объемов выборок меняются результаты проверки (*p-value*). По результатам экспериментов оценить, примерно какой объем выборок потребуется, чтобы принять верное решение и отклонить «несправедливую» проверяемую гипотезу.

Этапы исследования

1. В предложенных 5 выборках объемом $n = 200, 400, 600, 800, 1000$ *первая половина* моделировалась в соответствии с **нормальным законом** с параметрами *сдвига* и *масштаба* $(0, 1)$, *вторая половина* - с параметрами $(0.1, 1)$, то есть математическое ожидание для второй части *(скачком)* увеличилось на 10% от стандартного отклонения. Это достаточно *малое* отклонение, мало заметное при $\sigma = 1$.
    1. Последовательно, используя критерии **автокорреляции**, **Кокса-Стьюарта** и **Бартелса**, проверьте гипотезу о случайности и отсутствии тренда в 5 предложенных выборках. Проследите, как меняется **достигаемый уровень значимости** при проверке гипотезы по соответствующим критериям.
    2. *Зафиксируйте* результаты проверок в таблице.
    3. *Оцените*, примерно какой **объем** выборок потребуется, чтобы принять верное решение и *отклонить* **«несправедливую»** проверяемую гипотезу при задании вероятности ошибки 1-*го* рода $\alpha = 0.1, 0.05, 0.01$.
2. Создайте *(смоделируйте)* аналогичные выборки объемом $n = 200, 400, 600, 800, 1000$, в которых в середине математическое ожидание изменяется на величину 30% от стандартного отклонения. Такую выборку также можно сформировать объединением 2-*х* выборок, смоделированных с различными параметрами сдвига.
    1. Как и в **п.1**, используя критерии **автокорреляции**, **Кокса-Стьюарта** и **Бартелса**, проверьте гипотезу о случайности и отсутствии тренда в полученных выборках. Проследите, как меняется **достигаемый уровень значимости** при проверке гипотезы в данном случае. Насколько уменьшатся объемы выборок с обнаруженным трендом *(с отклонением гипотезы о случайности)*.
3. Смоделируйте *(аналогично п. 2)* 5 выборок $n = 200, 400, 600, 800, 1000$ с наличием *«сдвига в дисперсии»*, так чтобы в *первой* половине элементы моделировались в соответствии с нормальным законом с параметрами $(0, 1)$, а во *второй* - с параметрами $(0, 1.1)$, то есть стандартное отклонение второй части отличается на 10% от первой в большую сторону.

    1. Последовательно, используя критерий **Хсу** с h*-статистикой*, критерий **Кокса-Стьюарта для дисперсий**, ранговые критерии с метками **Клотца** и **Сэвиджа**, проверьте гипотезу о случайности и отсутствии тренда в 5 построенных выборках. Проследите, как меняется **достигаемый уровень значимости** при проверке гипотезы.
    2. *Зафиксируйте* результаты проверок в таблице.
    3. *Оцените*, примерно какой **объем** выборок потребуется, чтобы принять верное решение и *отклонить* **«несправедливую»** проверяемую гипотезу при задании вероятности ошибки 1*-го* рода $\alpha = 0.1, 0.05, 0.01$.
4. Аналогично **п. 2** создайте *(смоделируйте)* аналогичные выборки, в которых в середине стандартное отклонение (параметр масштаба) изменяется в большую сторону на величину 30%. Как и ранее это можно сделать объединением 2*-х* выборок, смоделированных с различными параметрами масштаба, но проще воспользоваться преобразованием выборки, задав параметр t[8] (сдвиг масштаба) равным +0.3.
    1. Используя те же критерии, что и в **п. 3** (критерий **Хсу** с h*-статистикой*, критерий **Кокса-Стьюарта для дисперсий**, ранговые критерии с метками **Клотца** и **Сэвиджа**), проверьте гипотезу о случайности и отсутствии тренда в 5 построенных выборках. Проследите, как меняется **достигаемый уровень значимости** при проверке гипотезы.
    2. *Зафиксируйте* результаты проверок в таблице.
    3. *Оцените*, насколько раньше *(при каком* ***объеме*** *выборок)* будет принято *верное* решение об *отклонении* **«несправедливой»** проверяемой гипотезы при задании вероятности ошибки 1*-го* рода $\alpha = 0.1, 0.05, 0.01$.
5. **Кратко сформулируйте для себя выводы, вытекающие из Ваших результатов.**

Выполнение пунктов

Проверка гипотезы о случайности и отсутствия тренда на основе предложенных выборок

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **автокорреляции** на основе предложенных выборок.

n	Значение S	Значение P
200	1.35607	0.17
400	1.45033	0.136
600	2.63469	0.012

n	**Значение** S	**Значение** P
800	2.30747	0.026
1000	2.329	0.01986

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Кокса-Стьюарта** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	0.08487	0.904
400	-0.01837	0.99
600	-2.124	0.038
800	-1.43436	0.16
1000	-2.27236	0.02306

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Бартелса** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	-1.188	0.228
400	-1.26198	0.204
600	-2.99919	0.002
800	-2.67884	0.008
1000	-2.58607	0.00971

Из полученных результатов, можно сделать вывод о том, что среди рассматриваемых критериев, при наименьших рассматриваемых значений n, лучше всего справился критерий **автокорреляции**, поскольку на малых значения n показал самое наименьшее значение **достигаемого уровня значимости**, при ложной H_1 гипотезе.

Для данной ситуации, критерий **Бартелса**, среди остальных критериев, оказался **самым мощным**, поскольку для большинства значений n, он показал *наименьшее* значение **достигаемого уровня значимости**. Можно отметить, что среди других критериев, значения на порядок меньше других, что однозначно говорит о **большей мощности**.

Среди трёх рассмотренных критериев, хуже всего себя показал критерий **Кокса-Стьюарта**. Как можно заметить для наименьших среди рассмотренных значений n, критерий не отвергает ложную гипотезу H_1. но достаточно увеличив n значения достигнутого уровня значимости становятся близкими к значениям критерия **автокорреляции**.

Можно сделать вывод, что для того чтобы отклонить **"несправедливую"** гипотезу при ошибки первого рода $\alpha = 0.1$, для параметрических и непараметрических критериев потребуется $400 < n^* \leq 600$ наблюдений. При ошибки первого рода $\alpha = 0.05$, для параметрических и непараметрических критериев потребуется $400 < n^* \leq 600$ наблюдений. При ошибки первого рода $\alpha = 0.01$, для критериев **автокорреляции** и **Кокса-Стьюарта** потребуется $n^* > 1000$ наблюдений, для критериев **Бартелса** потребуется $800 < n^* \leq 1000$ наблюдений.

Моделирование 5 выборок со скачком математического ожидания в 30% от стандартного отклонения

Выборки получаем путём объединения двух выборок. Одну выборку моделируем по закону $N(0, 1)$ с объёмом в два раза меньшим от размера результирующей выборки. Вторую выборку моделируем аналогично первой, только с изменённым параметром математического ожидания на 30% относительно стандартного отклонения т.е. с законом распределения $N(0.3, 1)$.

Эмпирическая функция распределения *(см. график 1)* выборки *model_1* с объёмом $n = 200$

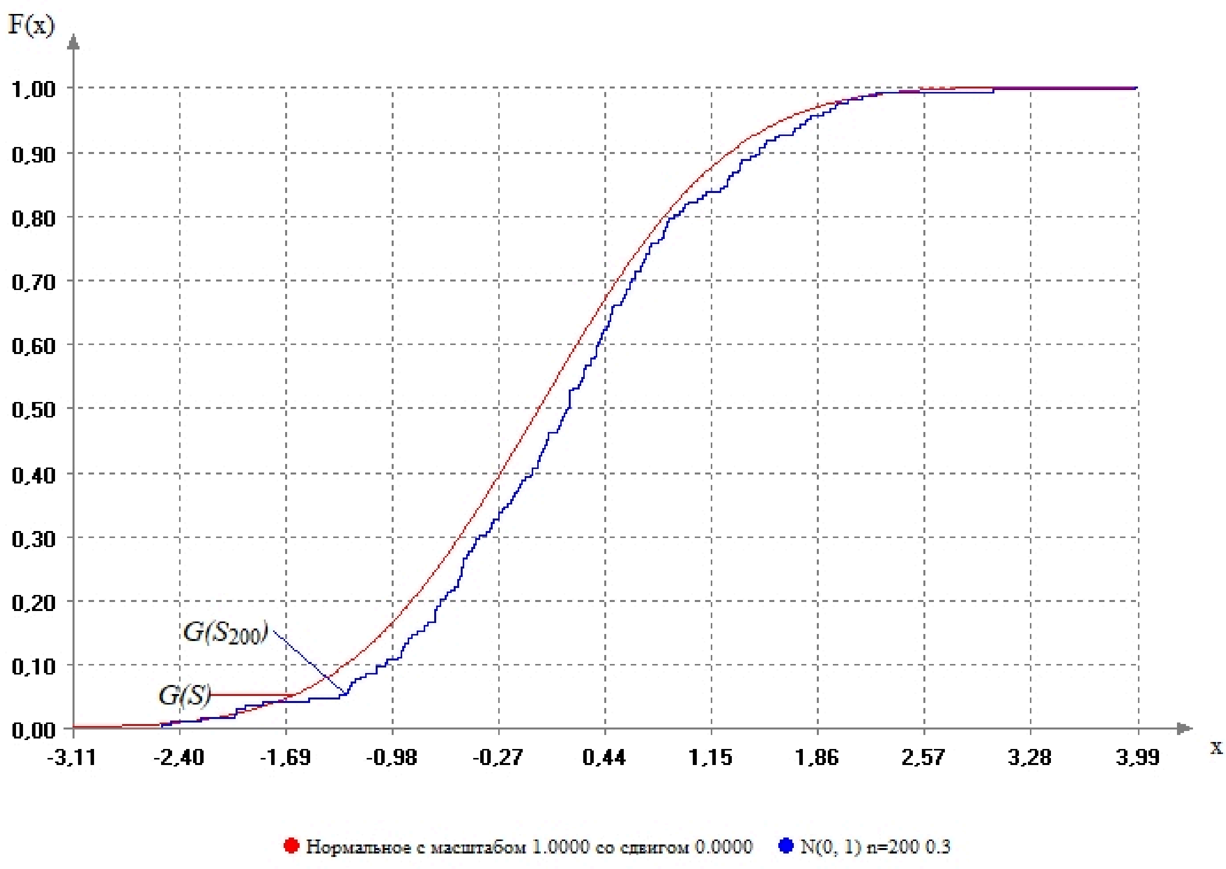


График 1 Эмпирическая функция распределения выборки *model_1*

Функция плотности распределения *(см. график 2)* выборки *model_1* с объёмом $n = 200$

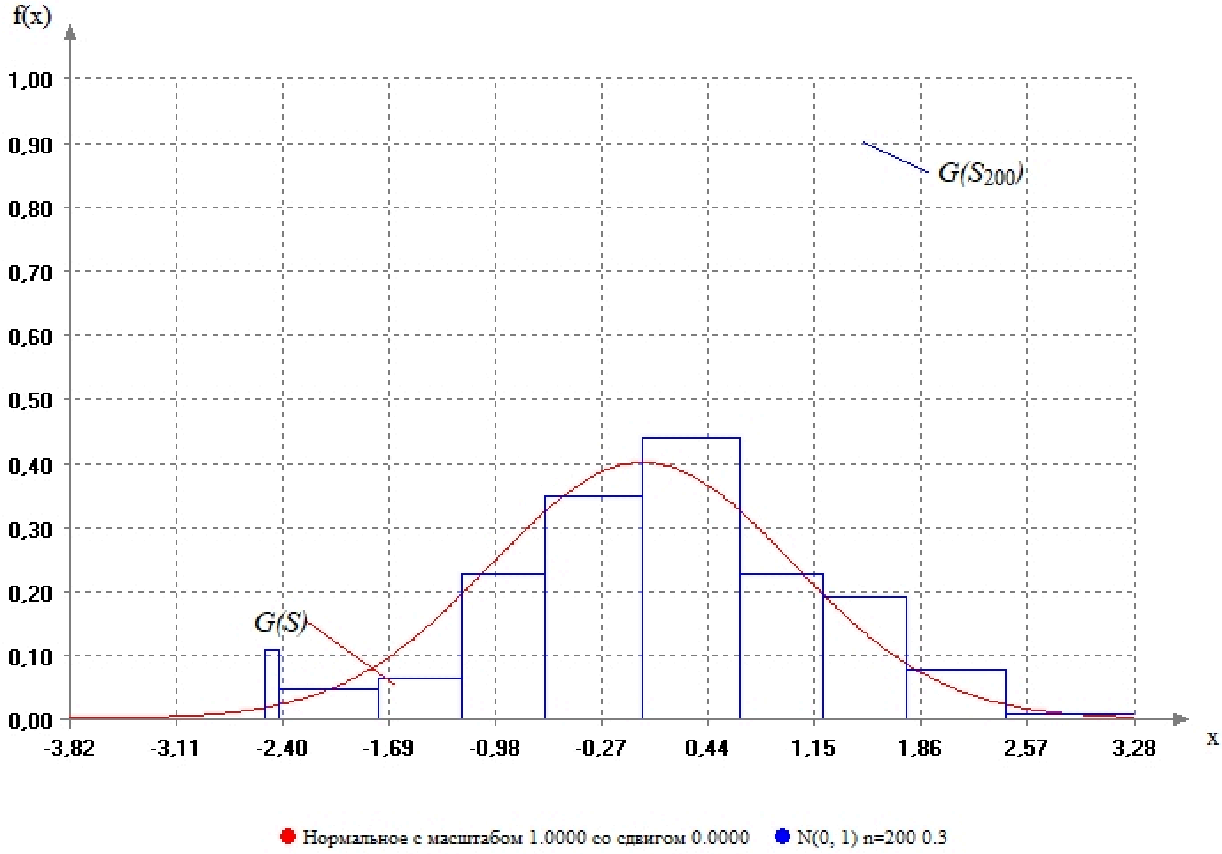


График 2 Функция плотности распределения выборки *model_1*

Для остальных выборок действуем аналогично. Для примера приведу график эмпирической функции распределения *(см. график 3)* и график функции плотности *(см. график 4)* выборки *model_5* с объёмом $n = 1000$.

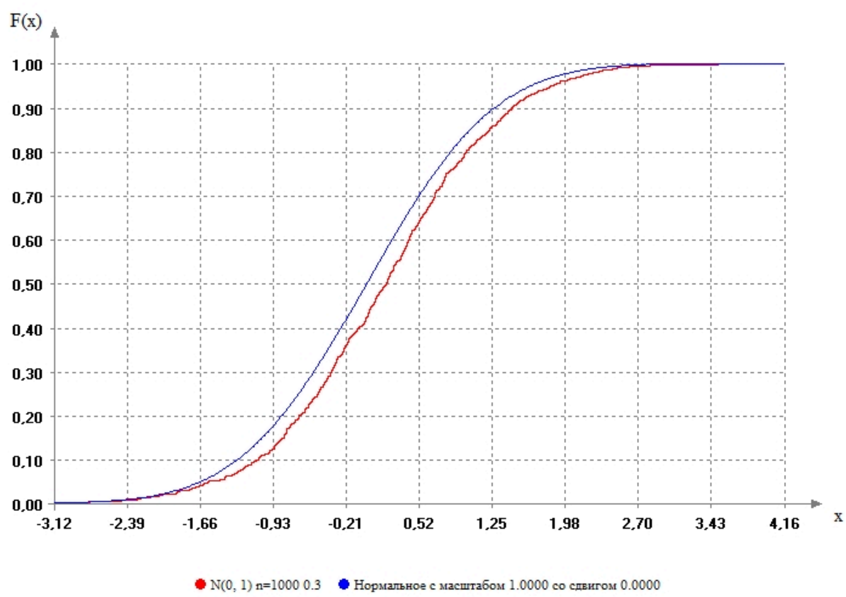


График 3 Эмпирическая функция распределения выборки *model_5*

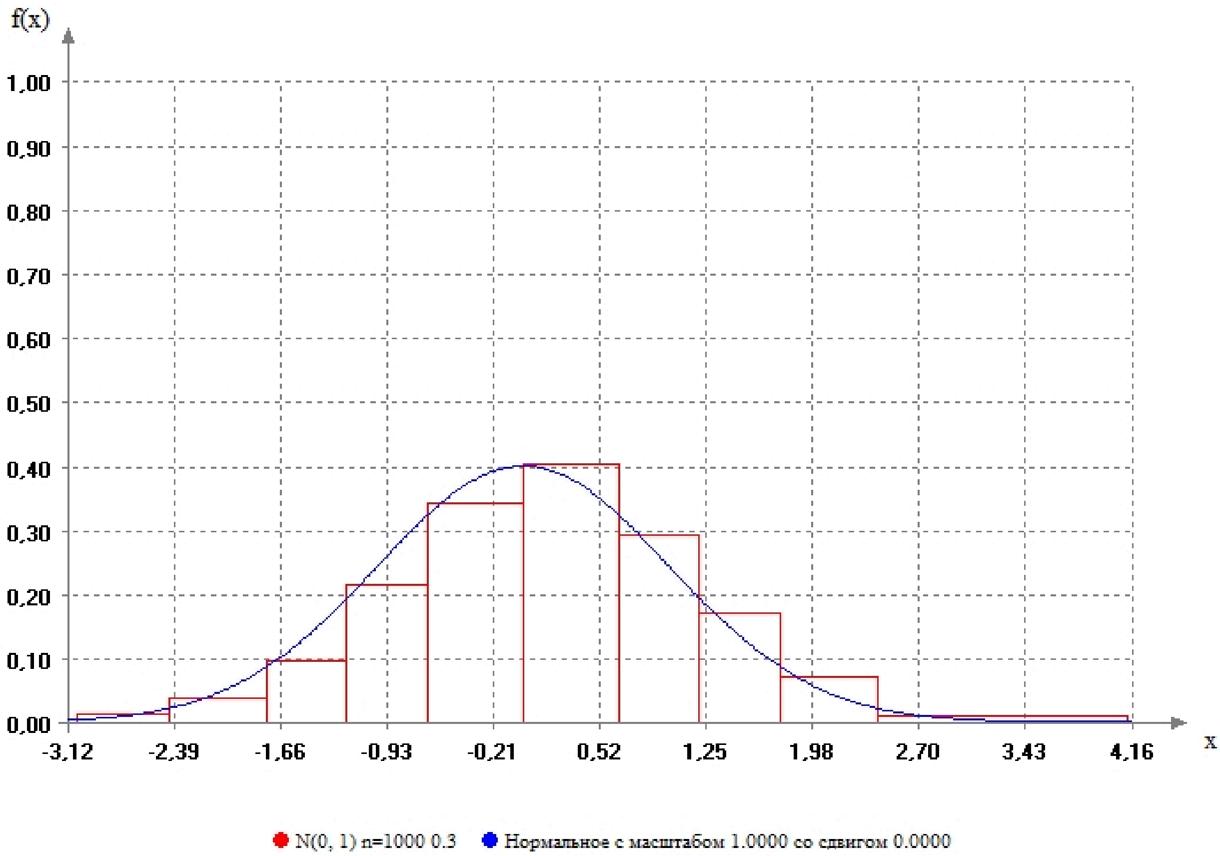


График 4 Функция плотности распределения выборки *model_5*

Проверка гипотезы о случайности и отсутствия тренда на основе смоделированных выборок со скачком по параметрам μ на 30% от стандартного отклонения

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **автокорреляции** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	1.18616	0.232
400	2.36933	0.018
600	1.18616	0.25
800	0.31085	0.784
1000	1.15031	0.28

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Кокса-Стьюарта** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	-0.73094	0.492
400	-1.67546	0.1
600	-0.73094	0.536
800	-0.87339	0.394
1000	-4.41923	0

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Бартелса** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	-1.31908	0.168
400	-2.21534	0.026
600	-1.31908	0.194
800	-0.07521	0.97
1000	-0.94085	0.374

Полученные результаты по критериям **Бартелса** и **автокорреляции** не однозначны. Легко можно увидеть, что в данном случае критерий **Кокса-Стьюарта** показал наилучший результат среди других. В отличие от других критериев имеет наименьшие скачки в значениях **достигаемого уровня значимости**, но, как и в первом пункте, при наименьшем из рассматриваемых значений n, имеет наибольшее среди других значений *достигнутого уровня значимости*.

Критерии **автокорреляции** и **Бартелса** имеют схожие значения *достигаемого уровня значимости*, имеют одно и то же значение n при котором происходит скачок *достигнутого уровня значимости*. Среди этих двух критериев, критерий **Бартелса** при наименьшем количестве наблюдений имеет наименьшее значение *достигнутого уровня значимости*, но при этом значения *достигнутого уровня значимости*, у обоих критериев, по какой-то причине **растут**, а не **уменьшается**.

Если сравнивать с предыдущим пунктом, то полученные результаты **парадоксальны**. Учитывая, что скачок по **математическому ожиданию** в 3 раза **больше**, чем в прошлом пункте. Ожидалось, что критерии более уверенно отвергнут гипотезу. Полученные результаты проверились на нескольких генерациях файлах. И во всех попытках были получены похожие результаты.

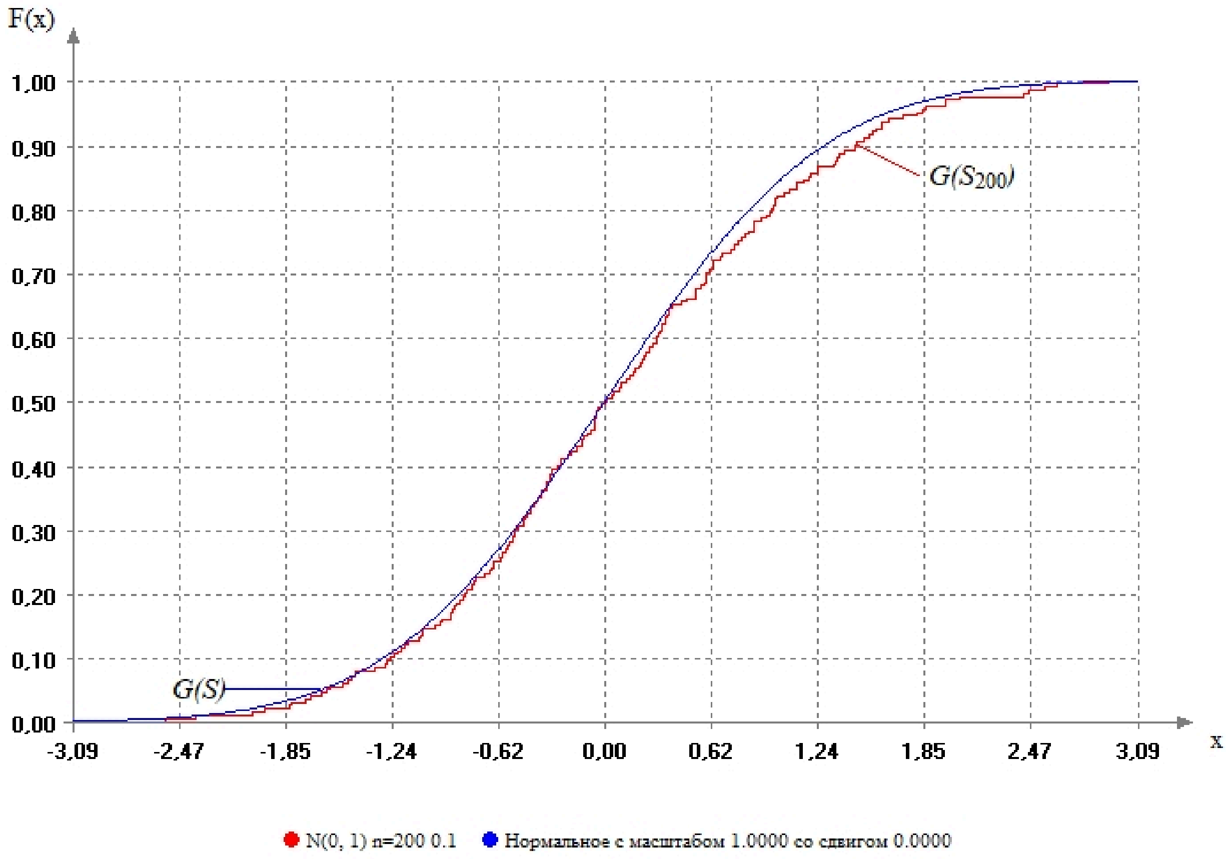
Можно сделать вывод, что для того, чтобы *отклонить* **"несправедливую"** гипотезу при ошибке первого рода $\alpha = 0.1$, для критерия **Кокса-Стьюарта** потребуется $800 < n^* \leq 1000$ наблюдений. Для остальных двух критериев потребуется $n^* > 1000$

наблюдений. Для всех критериев при ошибке первого рода меньших $\alpha = 0.1$ потребуется $n^* > 1000$.

Моделирование 5 выборок со скачком параметра σ на 10% от первого в большую сторону

Выборки получаем путём объединения двух выборок. Одну выборку моделируем по закону $N(0,1)$ с объёмом в два раза меньшим от размера результирующей выборки. Вторую выборку моделируем аналогично первой, только с изменённым параметром среднеквадратичного отклонения на 10% т.е. с законом распределения $N(0,1.1)$.

Эмпирическая функция распределения *(см. график 5)* выборки *model_6* с объёмом $n = 200$



<u>*График 5*</u> Эмпирическая функция распределения выборки <u>*model_6*</u>

Функция плотности распределения *(см. график 6)* выборки *model_6* с объёмом $n = 200$

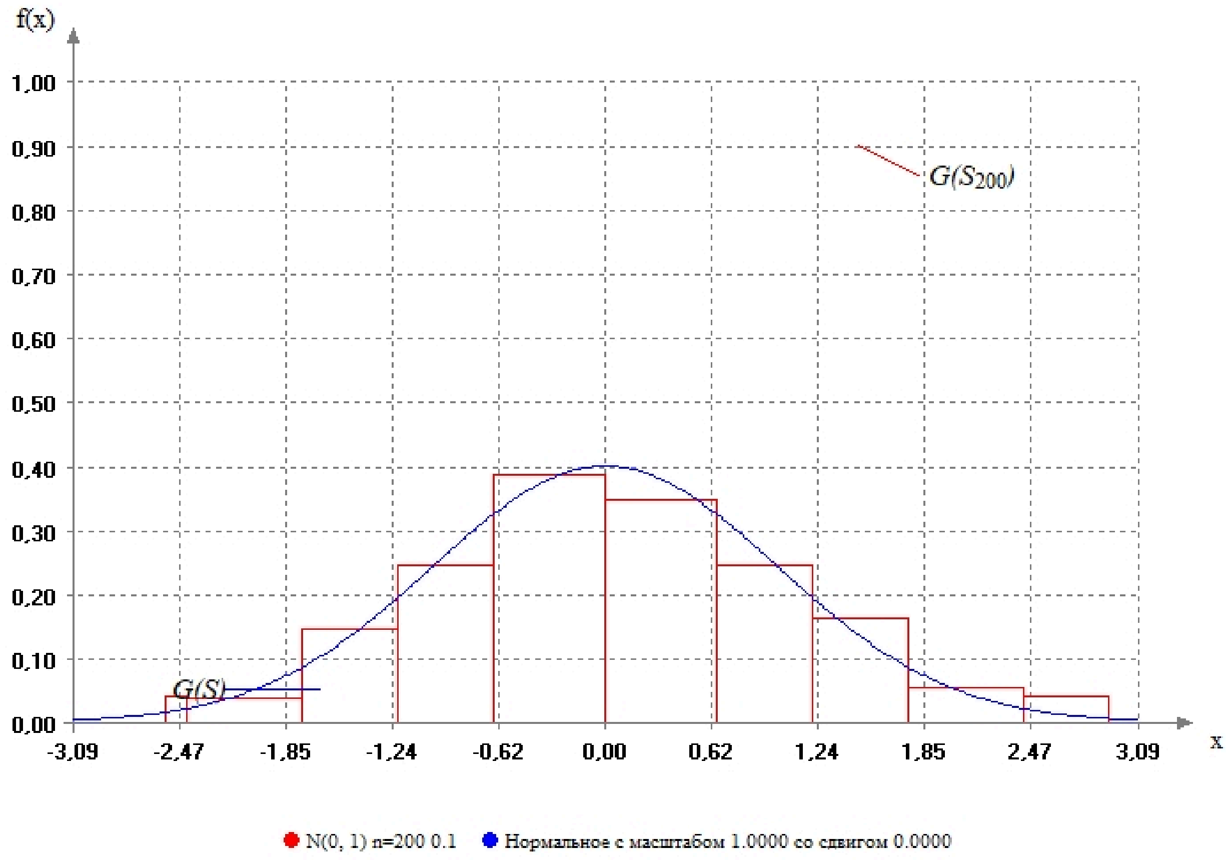


График 6 Функция плотности распределения выборки *model_6*

Для остальных выборок действуем аналогично. Для примера приведу график эмпирической функции распределения *(см. график 7)* и график функции плотности *(см. график 8)* выборки model_10 с объёмом $n = 1000$.

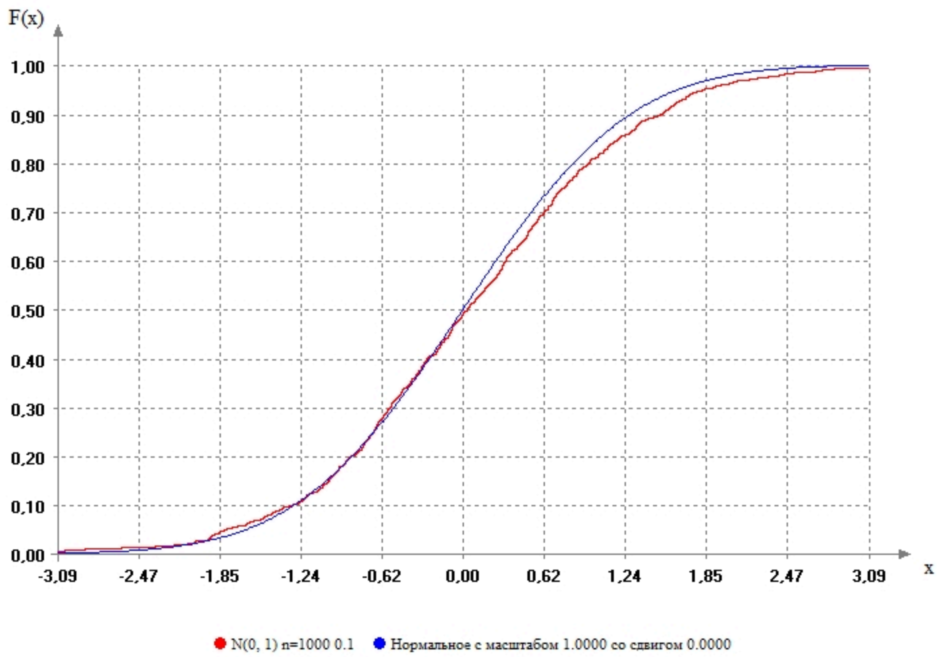


График 7 Эмпирическая функция распределения выборки *model_10*

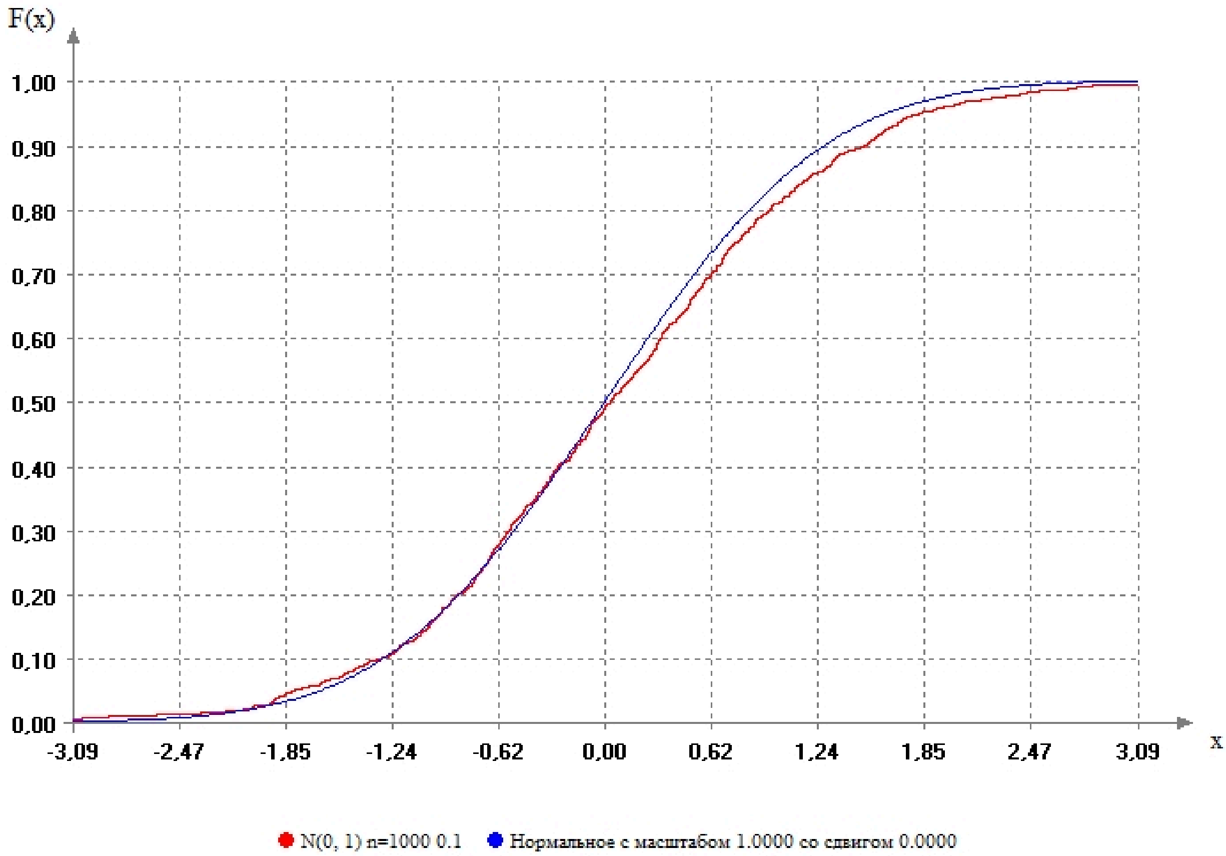


График 8 Функция плотности распределения выборки *model_10*

Проверка гипотезы о случайности и отсутствия тренда на основе смоделированных выборок со скачком по параметру σ на 0.1

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Хсу** на h-*статистики* на основе предложенных выборок.

n	Значение S	Значение P
200	-0.57520	0.722
400	1.51224	0.053
600	1.60022	0.054
800	1.83672	0.0318
1000	3.28065	0

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Кокса-Стьюарта для дисперсий** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	-0.44553	0.666
400	-1.49951	0.156
600	0.09690	0.908
800	-2.25606	0.0222
1000	-2.23957	0.02

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Клотца** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	-0.58922	0.524
400	1.28957	0.192
600	1.52799	0.128
800	1.99013	0.0436
1000	3.13799	0

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Сэвиджа** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	0.49177	0.644
400	1.44736	0.164
600	1.46772	0.172
800	1.83604	0.061
1000	1.78235	0.088

Из полученных данных, можно сделать вывод, что значения достигнутого уровня значимости ниже у критерия **Хсу**. Можно отметить так же, что в отличие от других критериев, уже при $n = 400$ данный критерий отвергает ложную гипотезу при ошибке 1-го рода $\alpha = 0.1$. Можно предположить, что, среди других критериев, данный имеет наибольшую мощность.

Ранговые критерии показали одинаковые результаты. Можно отметить, что критерий **Клотца** имеет наименьший **достигнутый уровень значимости** среди другого рангового критерия. Можно предположить, что данные критерии имеют похожую мощность.

В отличие от всех остальных критериев, критерий **Кокса-Стьюарта** имеет скачок при $n = 600$, можно предположить, что среди остальных критериев, возможно он имеет наименьшую мощность.

Можно сделать вывод, что для того, чтобы отклонить **"несправедливую"** гипотезу при ошибки первого рода $\alpha = 0.1$, для параметрического критерия **Хсу** потребуется $200 < n^* \leq 400$ наблюдений, для непараметрических критериев потребуется $600 < n^* \leq 800$ наблюдений. При ошибки первого рода $\alpha = 0.05$, для параметрических и непараметрических критериев потребуется $600 < n^* \leq 800$ наблюдений. При ошибки первого рода $\alpha = 0.01$, для критериев **Хсу** и **Клотца** потребуется $800 < n^* \leq 1000$ наблюдений. Для остальных критериев понадобится $n^* > 1000$ наблюдений.

Моделирование 5 выборок со скачком параметра σ на 10% от первого в большую сторону

Выборки получаем путём объединения двух выборок. Одну выборку моделируем по закону $N(0, 1)$ с объёмом в два раза меньшим от размера результирующей выборки. Вторую выборку моделируем аналогично первой, только с изменённым параметром среднеквадратичного отклонения на 30% т.е. с законом распределения $N(0, 1.3)$.

Эмпирическая функция распределения *(см. график 9)* выборки *model_11* с объёмом $n = 200$

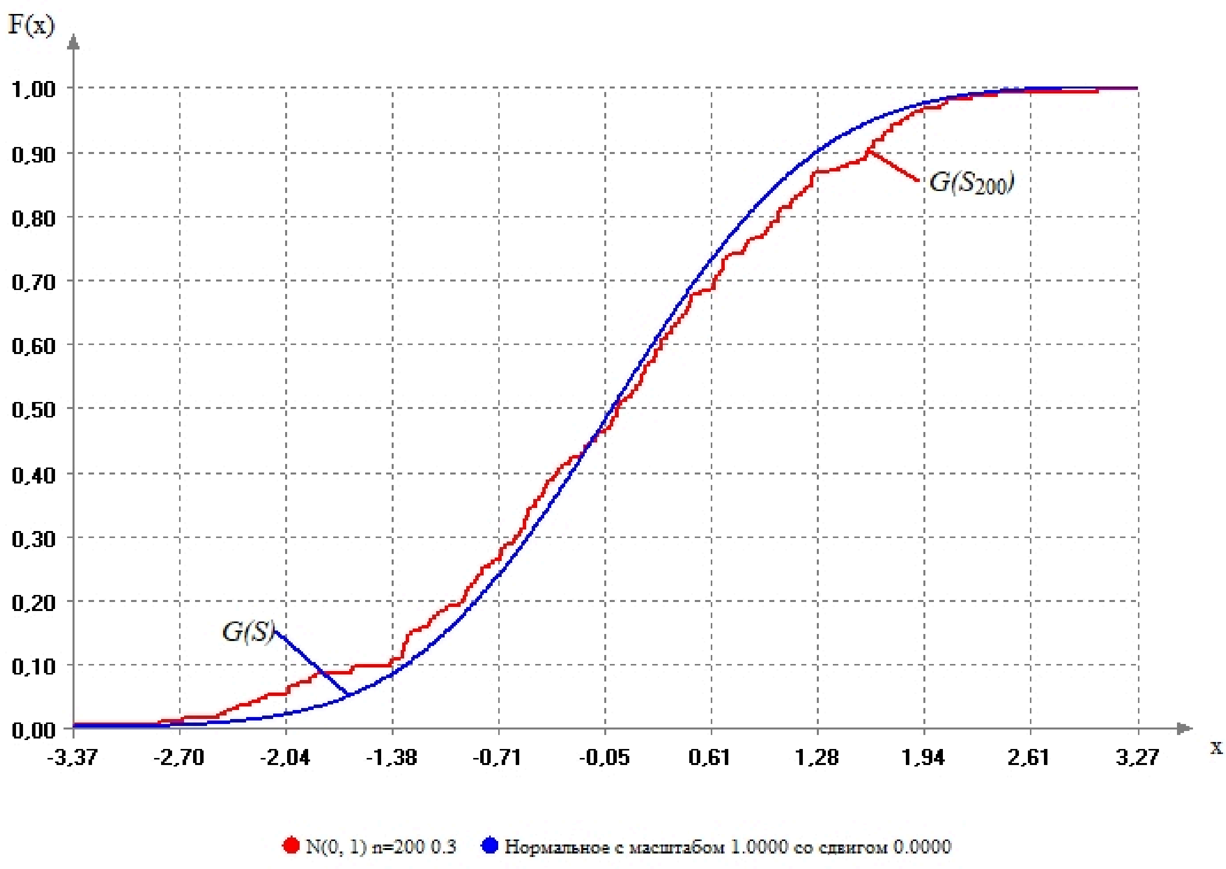
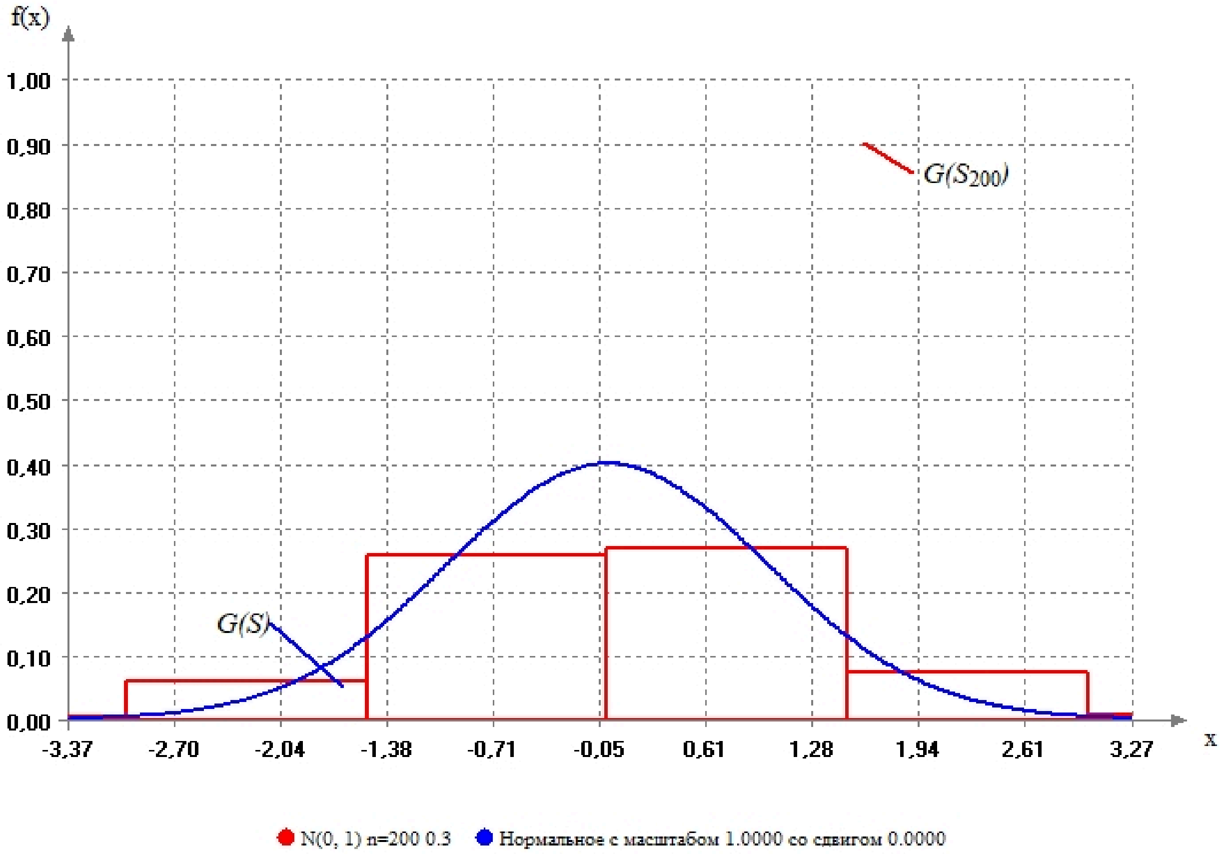


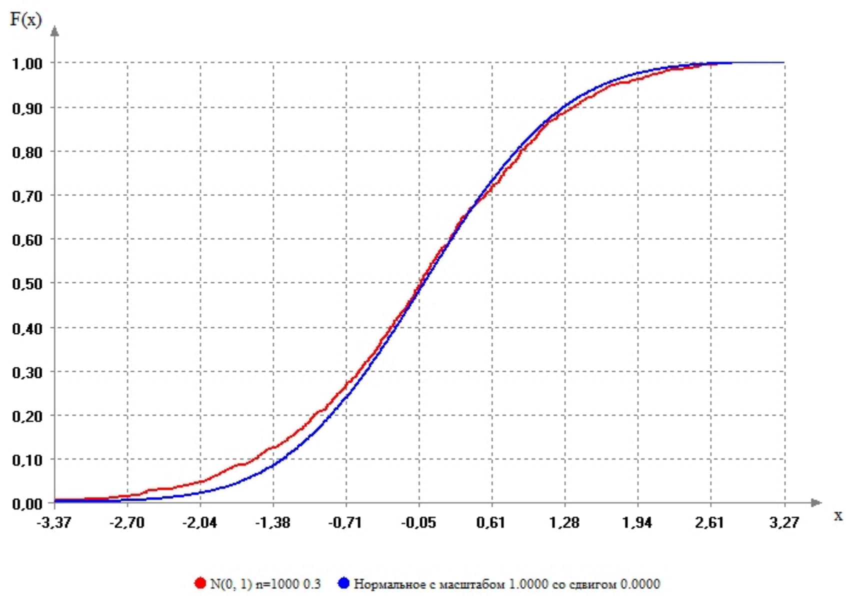
График 9 Эмпирическая функция распределения выборки *model_11*

Функция плотности распределения *(см. график 10)* выборки *model_11* с объёмом $n = 200$



<u>График 10</u> Функция плотности распределения выборки *<u>model_11</u>*

Для остальных выборок действуем аналогично. Для примера приведу график эмпирической функции распределения *(см. график 11)* и график функции плотности *(см. график 12)* выборки model_15 с объёмом $n = 1000$.



<u>График 11</u> Эмпирическая функция распределения выборки *<u>model_15</u>*

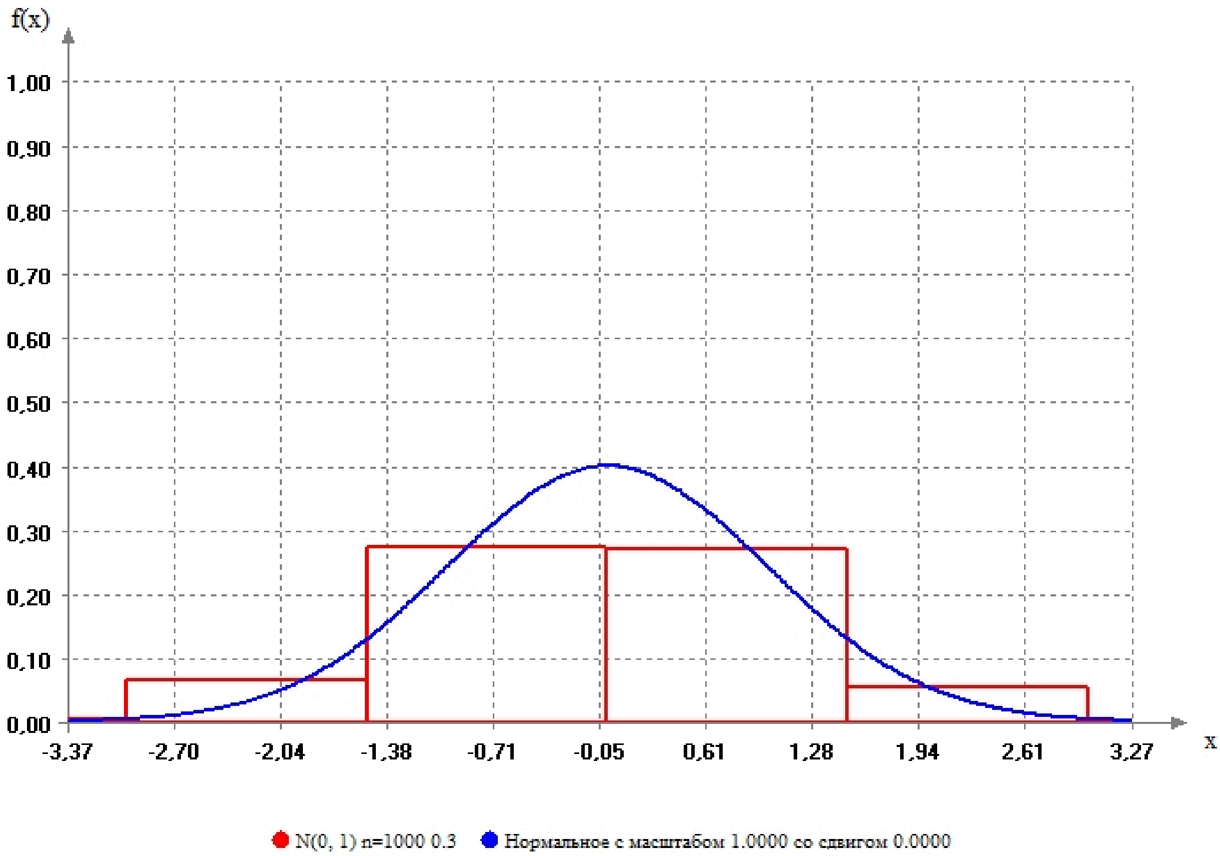


График 12 Функция плотности распределения выборки *model_15*

Проверка гипотезы о случайности и отсутствия тренда на основе смоделированных выборок со скачком по параметру σ на 0.3

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Хсу** на h-*статистики* на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	3.16968	0.002
400	1.61839	0.039
600	3.62731	0
800	4.79944	0
1000	4.66488	0

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Кокса-Стьюарта для дисперсий** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	-2.59571	0.014
400	-1.38995	0.192
600	-2.1467	0.034
800	-3.11781	0.002
1000	-3.23205	0

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Клотца** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	3.51916	0
400	1.2229	0.218
600	3.54452	0
800	4.75807	0
1000	4.77519	0

Результаты проверки гипотезы о **случайности и отсутствия тренда** по критерию **Сэвиджа** на основе предложенных выборок.

n	**Значение** S	**Значение** P
200	1.63717	0.116
400	1.38426	0.182
600	0.71217	0.538
800	1.23035	0.24
1000	2.98392	0.002

Полученные результаты, вполне ожидаемы. В отличие от прошлого пункта заметно то, что критерии показывают более низкие **достигнутые уровни значимости**, по сравнению с прошлым пунктом. Понятно, что увеличив *скачок*, критерием будет легче его обнаружить, что наглядно демонстрируется из полученных *результатов*.

Среди данных критериев, несложно заметить *выделяющийся* критерий **Сэвиджа**, на фоне других критериев он показывает значения **достигнутого уровня значимости** превышающее **сумму из достигнутых уровней значимости** остальных критериев. Однозначно из данной ситуации можно сказать, что критерий имеет мощность **меньше** остальных.

Как и в прошлом пункте критерий **Хсу** оказался **мощнее** остальных по той причине, что значения **достигнутого уровня значимости** *ниже* остальных.

Вторым по мощности, возможно, является критерий **Клотца**, основываясь на полученных значений **достигнутого уровня значимости**.

Вторым по мощности, возможно, является критерий **Кокса-Стьюарта**, основываясь на полученных значений **достигнутого уровня значимости**.

Можно сделать вывод, что для того, чтобы отклонить **"несправедливую"** гипотезу при ошибки первого рода $\alpha = 0.1$, для критериев **Хсу**, **Клотца** и **Кокса-Стьюарта** потребуется $n^* \leq 200$ наблюдений, для критерия **Сэвиджа** потребуется $800 < n^* \leq 1000$ наблюдений. При ошибки первого рода $\alpha = 0.05$, для критериев **Хсу**, **Клотца** и **Кокса-Стьюарта** потребуется $n^* \leq 200$ наблюдений, для критерия **Сэвиджа** потребуется $800 < n^* \leq 1000$ наблюдений. При ошибки первого рода $\alpha = 0.01$, для критериев **Хсу**, **Клотца** потребуется $n^* \leq 200$ наблюдений, для критерия **Кокса-Стьюарта** потребуется $600 < n^* \leq 800$ наблюдений, для критерия **Сэвиджа** потребуется $800 < n^* \leq 1000$ наблюдений.

Выводы

Из проведённого исследования можно сделать вывод, что для проверки гипотезы об **отсутствии тренда** при скачке **метаматематического ожидания** при малых количествах наблюдений можно использовать критерий **автокорреляции**, при больших количествах наблюдений можно использовать, как метод **автокорреляции**, так и метод **Бартелса**. На основе проделанного исследования можно сделать вывод, что мощность критерия **Бартелса** при больших количествах наблюдений больше, чем у критерия **автокорреляции**.

На основе исследования мощности критериев для проверки гипотезы об **отсутствии тренда** при скачке **среднеквадратичного отклонения** располагаются в следующем порядке: критерий **Хуса** с h-*статистикой* $\succ$ ранговый критерий **Клотца** $\succ$ непараметрический критерий **Кокса-Стьюарта** $\succ$ ранговый критерий **Сэвиджа**.