МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. Ломоносова
МЕХАНИКО-МАТЕМАТИЧЕСКИЙ ФАКУЛЬТЕТ
КАФЕДРА ТЕОРИИ ВЕРОЯТНОСТЕЙ

КУРСОВАЯ РАБОТА

«Сравнение мощности распространенных критериев нормальности распределения»

Студентки 3-го курса 308-ой группы
кафедры теории вероятностей
Дьячковой Екатерины

Научный руководитель:
доктор физико-математических наук, профессор
Яровая Елена Борисовна

МОСКВА 2020

Содержание

Введение

Целью курсовой работы явилось сравнение мощности таких распространённых критериев нормальности распределения как критерий χ^2 Пирсона, критерий Харке-Бера, критерий Лиллиефорса, а также критерий Шапиро — Уилка. Большое количество статистических методов исходит из предположения нормальности распределения изучаемых данных. Поэтому на начальном этапе исследования бывает очень важным проверить, подчиняется ли выборка нормальному закону. В разделе 1 мы напомним результаты по аналитическому и эмпирическому исследованию мощности. В разделе 2 мы приведём теоретическое описание наиболее распространённых критериев нормальности. Третий раздел будет посвящён сравнению их мощности.
Для лучшего восприятия материала приведём вступительную теорию из [1].
Пусть нам дана выборка $\mathcal{X} = (X_1, ..., X_n)$ из распределения P, которое нам неизвестно, но мы знаем, что оно принадлежит некоторому параметризованному семейству распределений: $P \in \mathcal{P} = \{P_\theta, \text{где } \theta \in \Theta\}$. Для проверки предположений о виде такого распределения используют статистические критерии.
Параметрические статистические гипотезы — это пара из предположения H_0 о неизвестном параметре и альтернативы H_1:

$$\begin{cases} H_0 : \theta \in \Theta_0 \\ H_1 : \theta \in \Theta_1 = \Theta \setminus \Theta_0 \end{cases} \tag{1}$$

Если множество Θ_0 (Θ_1) состоит из одной точки, то гипотезу H_0 (альтернативу H_1) называют *простой*; в противном случае гипотезу (или альтернативу) называют *сложной*. *Статистическим критерием* называется правило, по которому принимают или отклоняют гипотезу H_0. Это правило строится следующим образом: выбирается *критериальная статистика* $S = S(\mathcal{X})$ — функция, зависящая от выборки, но не от параметра, и *критическая область* G. Если критериальная статистика попала в критическое множество: $S \in G$, то мы отклоняем гипотезу H_0. При тестировании возможны два вида ошибок: *ошибка I рода* — принимаем H_1, когда на самом деле верна H_0, и, наоборот, *ошибка II рода* — принимаем H_0, когда на самом деле верна H_1. Введём обозначение: $P_{\theta'}(X) = P(X|\theta = \theta')$, где $\theta' \in \Theta$ — некоторое фиксированное значение. Тогда *мощностью* статистического критерия называют вероятность отвергнуть гипотезу H_0 при значении параметра θ:

$$w(\theta) = P_\theta(S \in G) \tag{2}$$

Задачей отбора подходящего теста является минимизация ошибок и, соответственно, максимизация мощности статистического критерия. Вероятности ошибок I, II рода — $\alpha(\theta)$, $\beta(\theta)$, уровень значимости — α и мощность связаны следующим образом:

$$\begin{cases} w(\theta) = P_\theta(S \in G) = P(H_1|H_0) = \alpha(\theta) \le \alpha \ \forall \theta \in \Theta_0 \\ w(\theta) = P_\theta(S \in G) = P(H_1|H_1) = 1 - P(H_0|H_1) = 1 - \beta(\theta) \ \forall \theta \in \Theta_1 \end{cases} \tag{3}$$

Если зафиксированы уровень значимости α, критериальная статистика S и простые гипотезы:

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 \end{cases} \tag{4}$$

задача нахождения наиболее мощного критерия разрешима т.к., зависит только от вида критической области G. Перебирая различные множества, выбираем тот тест, мощность которого будет наибольшей:

$$\begin{cases} P_{\theta_0}(S \in G) = P_0(G) \leq \alpha \\ P_{\theta_1}(S \in G) = P_1(G) = 1 - \beta \to max_G \end{cases} \tag{5}$$

В случае, когда альтернатива не является простой, а полученный тест оказывается наиболее мощным для каждого фиксированного $\theta_1 \in \Theta_1$, такой критерий называют *равномерно наиболее мощным.*

1 Аналитическое и эмпирическое исследование мощности критериев

1.1 Теорема Неймана-Пирсона

На основании материала, изложенного в [2], рассмотрим базовый пример: тест Стьюдента. Пусть дана выборка из нормального распределения: $\mathcal{X} = (X_1, ..., X_n)$, где $X_i \sim N(\theta, \sigma^2)$ с неизвестным средним и известной дисперсией. Фиксируем уровень значимости α. Рассмотрим гипотезу с правосторонней критической областью, вида:

$$\begin{cases} H_0 : \theta = \theta_0 \\ H_1 : \theta = \theta_1 (\theta_0 < \theta_1) \end{cases} \tag{6}$$

В качестве критериальной статистики рассмотрим выборочное среднее $S = \bar{X}$: $\bar{X} \sim_{H_0} N(\theta_0, \frac{\sigma^2}{n})$, $\bar{X} \sim_{H_1} N(\theta_1, \frac{\sigma^2}{n})$.

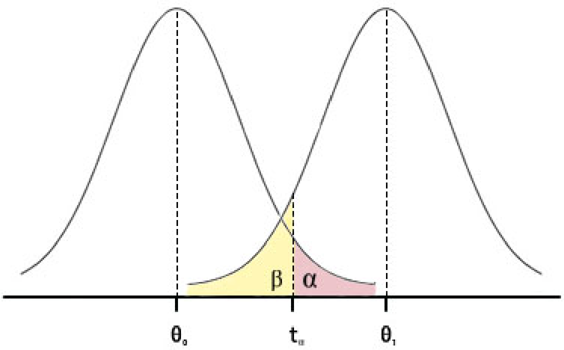

Рис.1: Распределение статистики S при гипотезах H_0 и H_1.

Найдём квантиль t_α для критической области. Т.к., $\sqrt{n} \cdot \frac{\bar{X} - \theta_0}{\sigma} \sim_{H_0} N(0, 1) \Rightarrow$ $1 - \alpha = P_0\left(\bar{X} \leq t_\alpha\right) = P_0\left(\sqrt{n} \cdot \frac{\bar{X} - \theta_0}{\sigma} \leq \sqrt{n} \cdot \frac{t_\alpha - \theta_0}{\sigma}\right) = \Phi\left(\sqrt{n} \cdot \frac{t_\alpha - \theta_0}{\sigma}\right) = \Phi\left(q_{1-\alpha}\right)$, где Φ — функция стандартного нормального распределения, $q_{1-\alpha}$ — квантиль уровня $1 - \alpha$.
Следовательно, искомый квантиль $t_\alpha = \theta_0 + \frac{(\sigma \cdot q_{1-\alpha})}{\sqrt{n}}$.
Получили критерий: выборочное среднее $\bar{X} \geqslant t_\alpha \Rightarrow$ отклоняем H_0.

Вернёмся к задаче вычисления мощности при фиксированной критериальной статистике. Аналогично примерам из [2] рассмотрим одновыборочный тест Стьюдента с правосторонней альтернативой. Пусть дана выборка размера $n = 100$ из нормального распределения с неизвестным средним θ и дисперсией $\sigma^2 = 4$, $\alpha = 0.05$. Сформулируем гипотезы:

$$\begin{cases} H_0 : \theta = 0 \\ H_1 : \theta > 0 \end{cases} \tag{7}$$

В отличие от уже известных моделей, рассмотрим следующие нестандартные критические области:

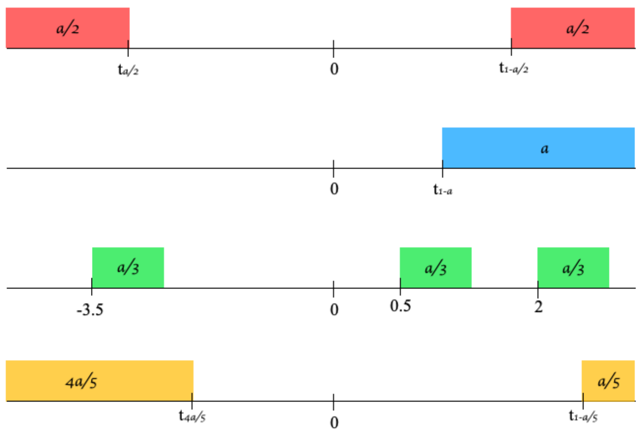

Рис.2 Рассматриваемые критические области для теста Стьюдента.

Зафиксируем альтернативу в гипотезах (7), получим тест Стьюдента с простыми гипотезами:

$$\begin{cases} H_0 : \theta = 0 \\ H_1 : \theta = \theta_1 \end{cases} \tag{8}$$

Рассмотрим график зависимости мощности тестов от односторонних альтернатив:

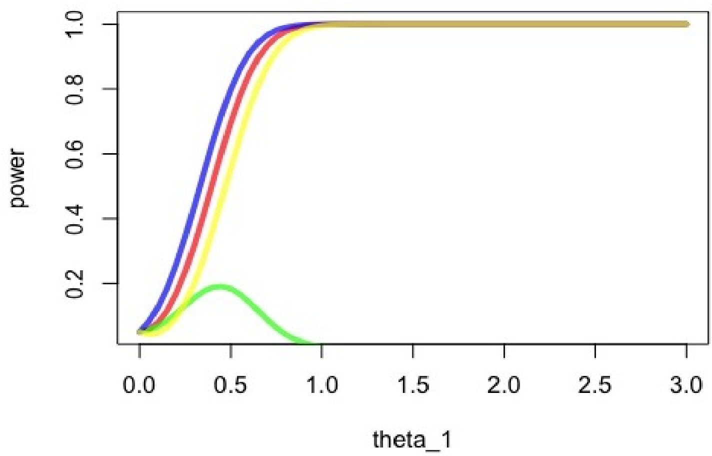

Рис.3 График зависимости мощности теста Стьюдента от односторонней альтернативы.

Цвет каждой линии графика на рис.3 соответствует цвету критической области на рис.2. Как видно из рис.3, для произвольной фиксированной альтернативы тест с

голубой критической областью будет иметь наибольшую мощность среди четырёх рассмотренных. Покажем, что произвольный статистический тест с такой критической областью и односторонними гипотезами вида (7) будет иметь наибольшую мощность среди тестов со всеми возможными критическими областями.
Напомним теорему Неймана-Пирсона, которую мы цитируем по учебнику [3].

Теорема 1 (Нейман-Пирсон) *Дана выборка рамера n. Вводим систему вложенных множеств $G_c := \{x \in \mathbb{R}^n : \frac{p_1(x)}{p_0(x)} \geq c\}$ и функцию $\varphi(c) := P_0(G_c)$, требуем выполнение двух условий:*
1.) плотности выборки $p_0(x)$ и $p_1(x)$, при H_0 и H_1 соответственно, положительны при всех $x \in \mathbb{R}^n$;
2.) для заданного уровня $\alpha \in (0;1)$ существует $c = c_\alpha : \varphi(c_\alpha) = \alpha$ (всегда выполнено при непрерывной φ).
Тогда при заданных условиях (1) и (2) наиболее мощный критерий уровня $\alpha \in (0;1)$ задаётся критическим множеством $G^* := G_{c_\alpha} = \{x \in \mathbb{R}^n : \frac{p_1(x)}{p_0(x)} \geq c_\alpha\}$.

В примере с односторонней альтернативой и тестом Стьюдента голубая критическая область задавалась следующим образом:

$$\{X \in \mathbb{R}^n : S = \sqrt{n} \cdot \frac{\bar{X} - \theta_0}{\sigma} \geq t_{1-\alpha}\} = \{X \in \mathbb{R}^n : \bar{X} \geq \frac{\sigma \cdot t_{1-\alpha}}{\sqrt{n}} + \theta_0\} = \tag{9}$$

$$= \{X \in \mathbb{R}^n : \bar{X} \geq c, \text{ где } c \in \mathbb{R}\} \tag{10}$$

По теореме Неймана-Пирсона критическая область наиболее мощного теста в условиях этой задачи задаётся множеством:

$$\{X \in \mathbb{R}^n : \frac{p_1(X)}{p_0(X)} = \exp^{-\frac{\sum_{i=0}^n (X_i - \theta_1)^2}{2 \cdot \sigma^2} + \frac{\sum_{i=0}^n (X_i - \theta_0)^2}{2 \cdot \sigma^2}} \geq c_\alpha\} = \tag{11}$$

$$= \{X \in \mathbb{R}^n : \bar{X} \geq \frac{2 \cdot \sigma^2 \cdot \ln c_\alpha + n \cdot (\theta_1^2 - \theta_0^2)}{2 \cdot n \cdot (\theta_1 -_0)}\} = \tag{12}$$

$$= \{X \in \mathbb{R}^n : \bar{X} \geq c_\alpha^{'}, \text{ где } c_\alpha^{'} \in \mathbb{R}\} \tag{13}$$

Видим, что множества (10) и (13) являются односторонними для некоторых констант c и c'_α. Для фиксированного уровня значимости α эти константы совпадут. Условия теоремы Неймана-Пирсона выполнены. Критическая область наиболее мощного теста с гипотезами (8) имеет вид (10) и не зависит от альтернативы, т.е., величины $\theta_1 > 0$. Поэтому тест с гипотезами (7) и голубой критической областью является *равномерно наиболее мощным.*
Как мы показали в предыдущем абзаце, теорема Неймана-Пирсона явным образом доказывает, что равномерно наиболее мощный критерий в случае односторонней альтернативы существует. Рассмотрим пример, демонстрирующий, что в случае двусторонней альтернативы наиболее мощный тест для каждой фиксированной θ_1 существует, но равномерно наиболее мощного — нет.

Пусть дана выборка размера $n = 100$ из нормального распределения с неизвестным средним θ и дисперсией $\sigma^2 = 4$. Используем одновыборочный тест Стьюдента, $\alpha = 0.05$ и двусторонние гипотезы:

$$\begin{cases} H_0 : \theta = 0 \\ H_1 : \theta \neq 0 \end{cases} \tag{14}$$

Аналогично примеру теста Стьюдента с односторонней альтернативой проводим тесты для каждой из четырёх критических областей на рис.2 с фиксированными альтернативами:

$$\begin{cases} H_0 : \theta = 0 \\ H_1 : \theta = \theta_1 \end{cases} \tag{15}$$

Строим графики зависимости мощности тестов от альтернативы:

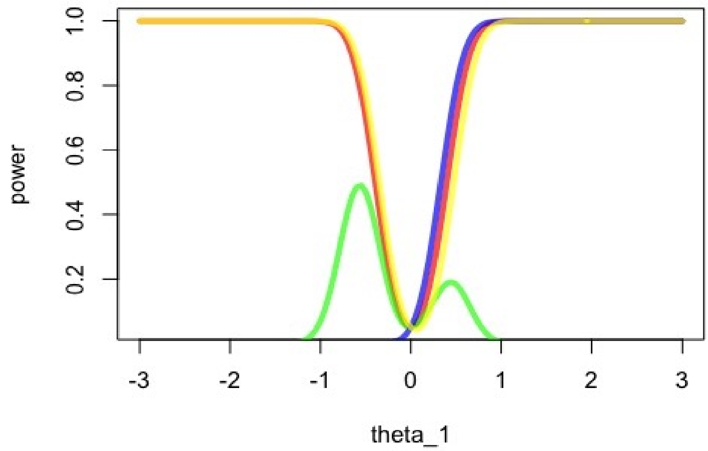

Рис.4: Графики зависимости мощности тестов Стьюдента от двусторонней альтернативы.

На правой полуоси $\theta_1 > 0$ наибольшая мощность у теста с правой критической областью, это вытекает из теоремы Неймана-Пирсона. Мощнее него при $\theta_1 > 0$ — нет. Тем не менее, на левой полуоси $\theta_1 < 0$ наибольшая мощность у теста с жёлтой критической областью. Получается, что на разных полуосях наибольшую мощность имеют разные тесты и в этом примере теста с наибольшей мощностью для любой θ_1 нет.
То есть, как и утверждалось, для теста с двусторонней альтернативой равномерно наиболее мощного теста не существует.

1.2 Метод Монте-Карло

В предыдущем подразделе примеры были связаны с тестом Стьюдента, а оценки мощности критериев вычислялись аналитически. На практике гипотеза может иметь сложную критическую область из-за чего аналитическое вычисление может быть трудоёмким. Поэтому для вычисления оценки мощности статистических тестов можно использовать метод Монте-Карло, описанный в [4] и [7]. Этот метод при помощи закона больших чисел позволяет вычислять оценку мощности произвольного статистического критерия при фиксированной альтернативе и заданной критической области. Приведём алгоритм метода Монте-Карло:
1.) Пусть заданы гипотеза $H_0 : \theta \in \Theta_0$ и альтернатива $H_1 : \theta \in \Theta_1$.
2.) Фиксируем простую альтернативу $H_1 : \theta = \theta_1$;
3.) Выбираем достаточно большое натуральное число K и генерируем K раз выборку $\mathcal{X}$ при условиях гипотезы H_1;
4.) Для каждой выборки вычисляем статистику t_{H_1};
5.) Выясняем, попала ли статистика в критическую область G. Если да, то $m_i = 1$;
6.) Вычисляем количество отвержений гипотезы H_0: $M = \sum_{i=1}^{K} m_i$;

7.) Вычисляем оценку мощности теста: $W = \frac{M}{K}$.
Для лучшего восприятия приведём схему метода Монте-Карло:

$$\forall i = 1, \ldots, K : \mathcal{X}_{H_1} \longrightarrow t_{H_1} \longrightarrow m_i = I(t_{H_1} \in G^*) \longrightarrow M = \sum_{i=1}^{K} m_i \longrightarrow W = \frac{M}{K} \quad (16)$$

Покажем преимущества этого метода на примере теста Стьюдента с двусторонней альтернативой и зелёной критической областью. Для сравнения аналитического вычисления и метода Монте-Карло нами были написаны две программы в среде статистического анализа R 3.5.1. Коды программ представлены на рис.5: при аналитическом вычислении легко ошибиться, требуется подробное описание критической области при помощи формул, которые для каждой области свои; метод Монте-Карло — более универсальный:

```
myp=seq(1,121,by=1)
pw=function(i){
  m=pt(q1, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
    pt(-3.5, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)+
    pt(q2, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
    pt(0.5, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)+
    pt(q3, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
    pt(2, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)
  return(m)
}

for (j in myp) {
  myp[j] = pw((j-1)/20 -3)
}
```

Рис.5: а) Аналитическое вычисление

```
K=1000 #количество итераций
s=2 #sd выборки
mc_pwr=function(i)
{
  res_plus=0
  for( j in 1:K)
  {
    X=rnorm(n = N, mean = i, sd = s)
    test=t.test(X)
    win=(test$statistic>-3.5 & test$statistic<q1) |
      (test$statistic>0.5 & test$statistic<q2) |
      (test$statistic>2 & test$statistic<q3)
    res_plus=res_plus+win
  }
  return(res_plus/K)
}
myc=seq(1,121,by=1)
for (j in myc) {
  myc[j] = mc_pwr((j-1)/20 -3)
}
```

Рис.5: б) Метод Монте-Карло

Также можно увидеть, как быстро этот метод достигает необходимой точности. При $K = 1000$ оценка мощности методом Монте-Карло близка к оценке мощности, вычисленной аналитически: в 60% точек абсолютная ошибка не превышает 5%. При $K = 10000$ графики уже почти неотличимы: в 80% точек абсолютная ошибка не превышает 5%

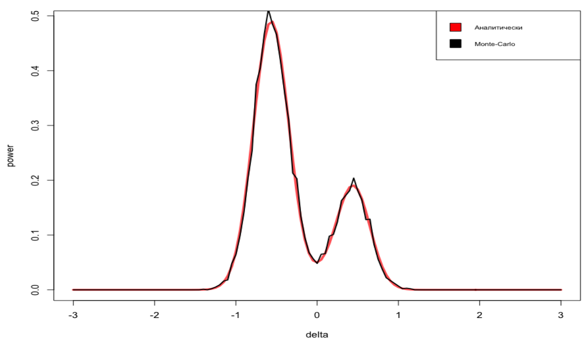

Рис.6: а) $K = 1000$

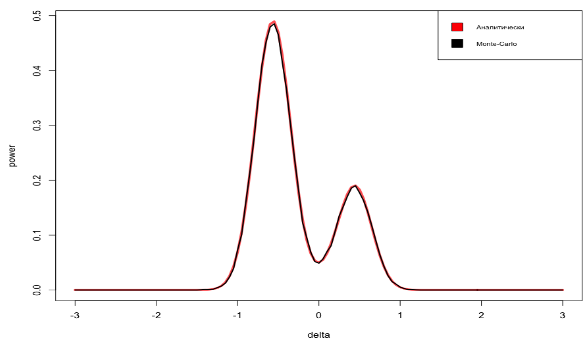

Рис.6: б) $K = 10000$

2 Критерии нормальности

2.1 Критерий χ^2 Пирсона

Пусть дана выборка: $\mathcal{X} = (X_1, \ldots, X_n)$ с неизвестной функцией распределения $F(x)$. Зафиксируем уровень значимости α, сформулируем гипотезы:

$$\begin{cases} H_0 : F(x) \in \mathcal{F} = \{\Phi_{\mu,\sigma^2}(x) : (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\} \\ H_1 : F(x) \notin \mathcal{F} \end{cases} \tag{17}$$

Разбиваем область значений X_1 на N промежутков: $\Delta_j = (a_i; b_j]$, $j = 1, ..., N$. Пусть ν_j — количество X_i, попавших в Δ_j, $p_j(\mu, \sigma^2) = P(X_1 \in \Delta_j)$. Вычисляем оценку максимального правдоподобия по сгруппированным данным:

$$(\hat{\mu}, \hat{\sigma^2}) = argmax_{(\mu,\sigma^2)} P\left(\nu_1 = l_1, ..., \nu_N = l_N\right) = \tag{18}$$

$$= argmax_{(\mu,\sigma^2)} \frac{n!}{l_1! \cdot ... \cdot l_N!} \cdot [p_1(\mu, \sigma^2)]^{l_1} \cdot ... \cdot [p_N(\mu, \sigma^2)]^{l_N} = \tag{19}$$

$$= argmax_{(\mu,\sigma^2)} \sum_{j=1}^{N} l_j \cdot \ln p_j(\mu, \sigma^2) \tag{20}$$

Тогда критериальная статистика выглядит следущим образом:

$$\chi_n^2 = \sum_{j=1}^{N} \frac{\left(\nu_j - n \cdot p_j(\hat{\mu}, \hat{\sigma^2})\right)^2}{n \cdot p_j(\hat{\mu}, \hat{\sigma^2})} \tag{21}$$

Если H_0 верна, критериальная статистика сходится по распределению к χ^2 распределению с $N - 3$ степенями свободы: $\chi_n^2 \xrightarrow[n\to\infty]{d} \chi^2(N-3)$. Поэтому, если критериальная статистика после вычисления оказывается больше $q_{1-\alpha}$ — квантиля распеделения $\chi^2(N-3)$ уровня $1 - \alpha$, то мы отклоняем H_0.

2.2 Критерий Харке - Бера

Этот тест основан на поиске отклонений выборочного распредления от нормального при помощи коэффициентов асимметрии и эксцесса. У нормального распределения они принимают значения 0 и 3 соответственно. Пусть дана выборка: $\mathcal{X} = (X_1, \ldots, X_n)$. Обозначим за S выборочную асимметрию, а за K — выборочный эксцесс, тогда:

$$S = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^3}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right)^{3/2}} \tag{22}$$

$$K = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^4}{\left(\frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2\right)^{2}} \tag{23}$$

Фиксируем уровень значимости α, формулируем гипотезы:

$$\begin{cases} H_0 : S = 0, K = 3 \\ H_1 : S \neq 0 \text{ и(или) } K \neq 3 \end{cases} \tag{24}$$

Критериальная статистика вычисляется следущим образом:

$$JB = \frac{n}{6}\left(S^2 + \frac{(K-3)^2}{4}\right) \tag{25}$$

Коэффициенты S и K являются асимптотически нормальными. Если H_0 выполнена, критериальная статистика сходится по распредлению к χ^2 с 2 степенями свободы: $JB \xrightarrow[n\to\infty]{d} \chi^2(2)$. Аналогично критерию χ^2 Пирсона, если критериальная статистика JB после вычисления оказывается больше $q_{1-\alpha}$ — квантиля распеделения $\chi^2(2)$ уровня $1-\alpha$, то мы отклоняем H_0.

2.3 Критерий Лиллиефорса

Этот критерий является модификацией теста Колмогорова-Смирнова. Пусть дана выборка $\mathcal{X} = (X_1, \ldots, X_n)$ с неизвестной функцией распределения $F(x)$. Вычисляем оценки: $\mu = \frac{1}{n} \cdot \sum_{i=1}^n X_i = \bar{X}$, $s^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$. Формулируем гипотезы:

$$\begin{cases} H_0 : F(x) = \Phi_{\mu,s^2}(x) \\ H_1 : F(x) \neq \Phi_{\mu,s^2}(x) \end{cases} \tag{26}$$

где $\Phi_{\mu,s^2}(x)$ — функция нормального распределения со средним μ и дисперсией s^2. Статистика теста вычисляется по формуле:

$$D_n = sup_{x\in\mathbb{R}}|F_n(x) - \Phi_{\mu,s^2}(x)| \tag{27}$$

где $F_n(x) = \frac{1}{n}\sum_{i=1}^n I(X_i \leq x)$ — эмпирическая функция распределения выборки $\mathcal{X}$. Если H_0 верна, критериальная статистика сходится по распределению к распределению Лиллиефорса. Критические значения находятся при помощи таблиц или метода Монте-Карло.

2.4 Критерий Шапиро — Уилка

Пусть дана выборка: $\mathcal{X} = (X_1, \ldots, X_n)$ с неизвестной функцией распределения $F(x)$. Сформулируем гипотезы:

$$\begin{cases} H_0 : F(x) \in \mathcal{F} = \{\Phi_{\mu,\sigma^2}(x) : (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+\} \\ H_1 : F(x) \notin \mathcal{F} \end{cases} \tag{28}$$

Критериальная статистика вычисляется следующим образом:

$$W = \frac{\left(\sum_{i=1}^n a_i \cdot X_{(i)}\right)^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \tag{29}$$

где $X_{(1)} \leq \ldots \leq X_{(n)}$ — вариационный ряд порядковых статистик, а коэффициенты a_i вычисляются так:

$$(a_1, \ldots, a_n) = \frac{m^T \cdot V^{-1}}{\sqrt{m^T \cdot V^{-1} \cdot V^{-1} \cdot m}} \tag{30}$$

где $m = (m_1, ..., m_n)$ и V — это вектор математических ожиданий и ковариационная матрица порядковых статистик из стандартного нормального распределения соответственно. Критические значения для заданного уровня значимости α также находятся из таблиц.

3 Исследование мощности наиболее распространенных критериев

Для исследование мощности тестов на нормальность, рассмотренных в разделе 2 нами был использован метод Монте-Карло, о котором говорилось в разделе 1. Уровень значимости зафиксировали $\alpha = 0.05$, далее генерировали $K = 10000$ раз выборки размера $n = 10, \ldots, 2000$ из четырёх распределений:
1.) Бета с параметрами $\alpha = 2$ и $\beta = 2$, его плотность:

$$p(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)} \tag{31}$$

2.) Гамма с параметрами $k = 4$ и $\theta = 5$, его плотность:

$$p(x) = \frac{1}{\Gamma(k)\theta^k} x^{k-1} \exp -\frac{x}{\theta} \tag{32}$$

3.) Гамма с параметрами $k = 1$, $\theta = 5$;
4.) Распределение Лапласа с параметрами $\alpha = 1$ и $\beta = 0$, его плотность:

$$p(x) = \frac{\alpha}{2} \exp -\alpha|x-\beta| \tag{33}$$

Далее мы применяли метод Монте-Карло и вычисляли оценки мощности тестов. Результаты представлены на рисунках 7-10.

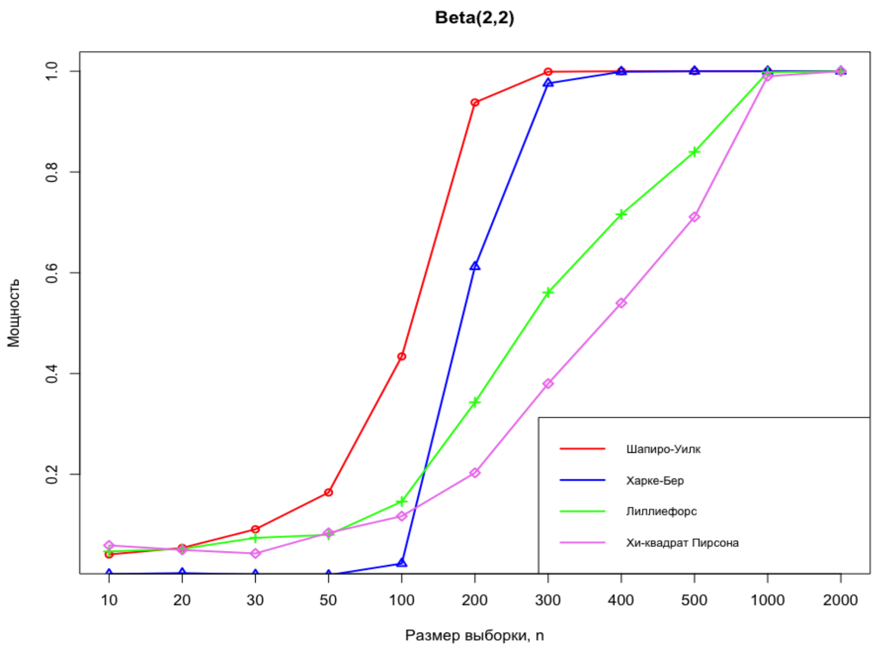

Рис.7 Сравнение мощности тестов на нормальность на распределении Бета(2,2).

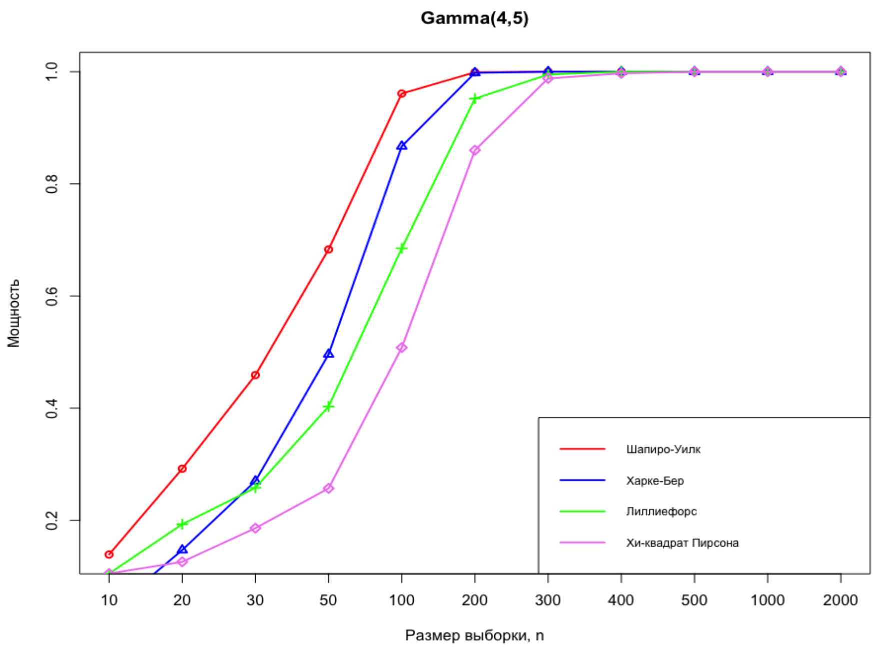

Рис.8 Сравнение мощности тестов на нормальность на распределении Гамма(4,5).

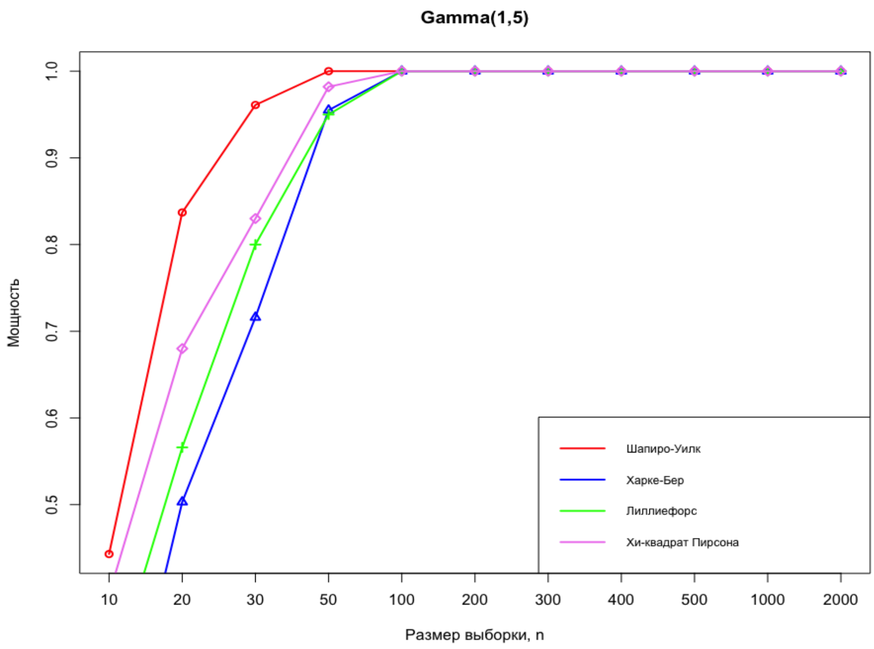

Рис.9 Сравнение мощности тестов на нормальность на распределении Гамма(1,5).

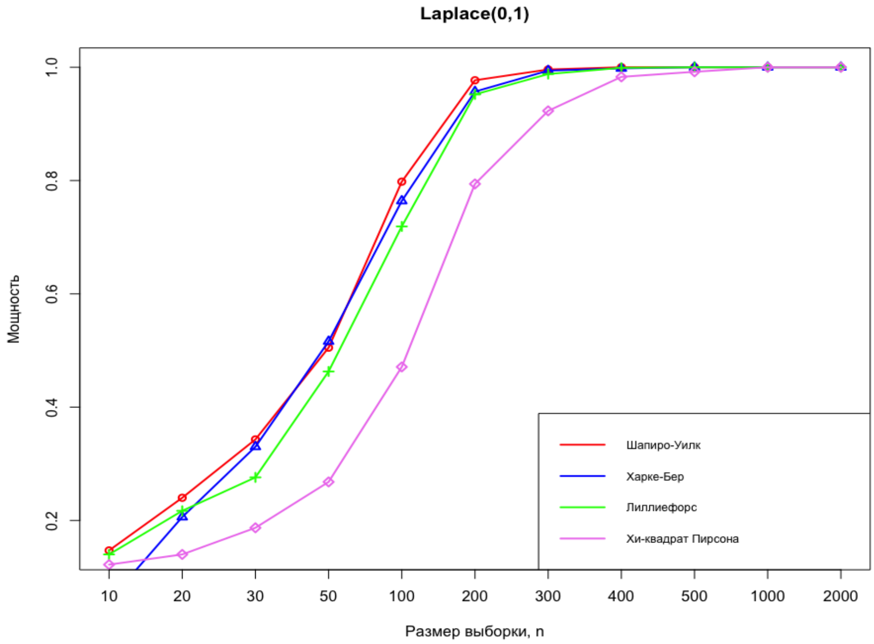

Рис.10 Сравнение мощности тестов на нормальность на распределении Лапласа(0,1).

Заключение

Из рассмотренных в работе методов наибольшей мощностью на исследуемых распределениях обладал тест Шапиро-Уилка. Эти результаты также подтверждаются в работах [5] и [6]. Таким образом, наше исследование подтверждает, что критерий Шапиро-Уилка является наилучшим среди наиболее распространённых критериев для проверки гипотезы о нормальном распределении. Также мы показали, что метод Монте-Карло может быть использован для оценки мощности критериев, когда аналитические подходы сложны или недоступны.

Приложение

```
N=100
alpha=0.05
nu=seq(0,3,by=0.05)

qw1=qt(alpha/2, df=N-1, ncp=0, lower.tail = T, log.p = FALSE)
qw2=qt(1-alpha/2, df=N-1, ncp=0, lower.tail = T, log.p = FALSE)
mypow_t=seq(1,61,by=1)
powf_t=function(i){
  return(1-pt(qw2, df=N-1, ncp=((i) * sqrt(N)/2), lower.tail = T, log.p = FALSE)+
  pt(qw1, df=N-1, ncp=((i)* sqrt(N)/2), lower.tail = T, log.p = FALSE))
}
for (j in mypow_t){
  mypow_t[j] = powf_t((j-1)/20)
}

q=qt(1-alpha, df=N-1, ncp=0, lower.tail = T, log.p = FALSE)
mypow_o=seq(1,61,by=1)
powf_o=function(i){
  return(1-pt(q, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE))
}
for (j in mypow_o){
  mypow_o[j] = powf_o((j-1)/20)
}

f1=pt(-3.5, df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
q1=qt(f1+(alpha/3), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
f2=pt(0.5, df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
q2=qt(f2+(alpha/3), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
f3=pt(2, df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
q3=qt(f3+(alpha/3), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
myp=seq(1,61,by=1)
pw=function(i){
  m=pt(q1, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
  pt(-3.5, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)+
  pt(q2, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
  pt(0.5, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)+
  pt(q3, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
  pt(2, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)
  return(m)
}
for (j in myp){
  myp[j] = pw((j-1)/20)
}

qw041=qt((4*alpha/5), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
```

```
qw042=qt((1-alpha/5), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
myp04=seq(1,61,by=1)
pw04=function(i){
  return(return(1-pt(qw042, df=N-1, ncp=((i) * sqrt(N)/2), lower.tail = T, log.p = FA
  pt(qw041, df=N-1, ncp=((i) * sqrt(N)/2), lower.tail = T, log.p = FALSE)))
}
for (j in myp04){
  myp04[j] = pw04((j-1)/20)
}

plot(nu,mypow_t,xlab="theta_1",ylab="power",col=rgb(1,0,0,0.7),type = "l",lwd=4)
lines(nu,myp,type="l",col=rgb(0,1,0,0.7),lwd=4)
lines(nu,mypow_o,col=rgb(0,0,1,0.7),type = "l",lwd=4)
lines(nu,myp04,type="l",col=rgb(1,1,0,0.7),lwd=4)

nu=seq(-3,3,by=0.05)

qw1=qt(alpha/2, df=N-1, ncp=0, lower.tail = T, log.p = FALSE)
qw2=qt(1-alpha/2, df=N-1, ncp=0, lower.tail = T, log.p = FALSE)
mypow_t=seq(1,121,by=1)
powf_t=function(i){
  return(1-pt(qw2, df=N-1, ncp=((i) * sqrt(N)/2), lower.tail = T, log.p = FALSE)+
  pt(qw1, df=N-1, ncp=((i) * sqrt(N)/2), lower.tail = T, log.p = FALSE))
}
for (j in mypow_t){
  mypow_t[j] = powf_t((j-1)/20 -3)
}

q=qt(1-alpha, df=N-1, ncp=0, lower.tail = T, log.p = FALSE)
mypow_o=seq(1,121,by=1)
powf_o=function(i){
  return(1-pt(q, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE))
}
for (j in mypow_o){
  mypow_o[j] = powf_o((j-1)/20 -3)
}

f1=pt(-3.5, df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
q1=qt(f1+(alpha/3), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
f2=pt(0.5, df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
q2=qt(f2+(alpha/3), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
f3=pt(2, df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
q3=qt(f3+(alpha/3), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
myp=seq(1,121,by=1)
pw=function(i){
  m=pt(q1, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
  pt(-3.5, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)+
```

```
  pt(q2, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
  pt(0.5, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)+
  pt(q3, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)-
  pt(2, df=N-1, ncp=(i) * sqrt(N)/2, lower.tail = T, log.p = FALSE)
  return(m)
}

for (j in myp){
  myp[j] = pw((j-1)/20 -3)
}
qw041=qt((4*alpha/5), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
qw042=qt((1-alpha/5), df=N-1, ncp=0, lower.tail = TRUE, log.p = FALSE)
myp04=seq(1,121,by=1)
pw04=function(i){
  return(return(1-pt(qw042, df=N-1, ncp=((i) * sqrt(N)/2), lower.tail = T, log.p = FA
  pt(qw041, df=N-1, ncp=((i) * sqrt(N)/2), lower.tail = T, log.p = FALSE)))
}
for (j in myp04){
  myp04[j] = pw04((j-1)/20 -3)
}

plot(nu,mypow_t,xlab="theta_1",ylab="power",col=rgb(1,0,0,0.7),type = "l",lwd=4)
lines(nu,myp,type="l",col=rgb(0,1,0,0.7),lwd=4)
lines(nu,mypow_o,col=rgb(0,0,1,0.7),type = "l",lwd=4)
lines(nu,myp04,type="l",col=rgb(1,1,0,0.7),lwd=4)

library(moments)
library(nortest)

K = 10000

gen_sample = function(name, N) #function for quick generation{
  if (name == "beta_22"){
    return(rbeta(shape1 = 2, shape2 = 2, n = N))
  }
  if (name == "gamma45"){
    return(rgamma(shape = 4, scale = 5, n = N))
  }
  if (name == "gamma15"){
    return(rgamma(shape = 1, scale = 5, n = N))
  }
  if (name == "laplace01"){
    return(rlaplace(m=0, s=1, n = N))
  }
}

quick_pearson = function(x, alpha = 0.05)
```

```
  return(pearson.test(x)$p.value < 0.05)

pearson_power = function (K = 1000, type = "beta_22", N = 100){
  sample_matrix = replicate(n = K, expr =  gen_sample(name = type, N=N))
  sample_matrix = as.data.frame(sample_matrix)
  p_vals = lapply(sample_matrix, quick_pearson)
  p_vals = as.numeric(p_vals)
  power = sum(p_vals)/length(p_vals)
  return(power)
}

quick_jarque = function(x, alpha = 0.05)
  return(jarque.test(x)$p.value < 0.05)

jarque_power = function (K = 1000, type = "beta_22", N = 100){
  sample_matrix = replicate(n = K, expr =  gen_sample(name = type, N=N))
  sample_matrix = as.data.frame(sample_matrix)
  p_vals = lapply(sample_matrix, quick_jarque)
  p_vals = as.numeric(p_vals)
  power = sum(p_vals)/length(p_vals)
  return(power)
}

quick_slillie = function(x, alpha = 0.05)
  return(lillie.test(x)$p.value < 0.05)

lillie_power = function (K = 1000, type = "beta_22", N = 100){
  sample_matrix = replicate(n = K, expr =  gen_sample(name = type, N=N))
  sample_matrix = as.data.frame(sample_matrix)
  p_vals = lapply(sample_matrix, quick_slillie)
  p_vals = as.numeric(p_vals)
  power = sum(p_vals)/length(p_vals)
  return(power)
}

quick_shapiro = function(x, alpha = 0.05)
  return(shapiro.test(x)$p.value < 0.05)

Shapiro_wilk_power = function (K = 1000, type = "beta_22", N = 100){
  sample_matrix = replicate(n = K, expr =  gen_sample(name = type, N=N))
  sample_matrix = as.data.frame(sample_matrix)
  p_vals = lapply(sample_matrix, quick_shapiro)
  p_vals = as.numeric(p_vals)
  power = sum(p_vals)/length(p_vals)
  return(power)
}
```

```
N_vector = list(10, 20, 30, 50, 100, 200, 300, 400, 500, 1000,2000)
X=seq(1,length(N_vector),by=1)
p_b22=sapply(N_vector, function(n) pearson_power(type = "beta_22", N = n))
j_b22=sapply(N_vector, function(n) jarque_power(type = "beta_22", N = n))
l_b22=sapply(N_vector, function(n) lillie_power(type = "beta_22", N = n))
s_b22=sapply(N_vector, function(n) Shapiro_wilk_power(type = "beta_22", N = n))
plot(x=X,y=s_b22,type = "o",main="Beta(2,2)",ylab = "Мощность",
xlab="Размер выборки, n",col="red",pch=1,lwd=2,xaxt="n")
lines(x=X,y=j_b22,type = "o",col="blue",pch=2,lwd=2,xaxt="n")
lines(x=X,y=l_b22,type = "o",col="green",pch=3,lwd=2,xaxt="n")
lines(x=X,y=p_b22,type = "o",col="violet",pch=5,lwd=2,xaxt="n")
legend("bottomright",legend=c("Шапиро-Уилк","Харке-Бер","Лиллиефорс",
"Хи-квадрат Пирсона"), col=c("red","blue","green","violet"),lwd=2,cex = 0.75)
axis(1, at=X,labels=N_vector, las=1)

p_l01=sapply(N_vector, function(n) pearson_power(type = "laplace01", N = n))
j_l01=sapply(N_vector, function(n) jarque_power(type = "laplace01", N = n))
l_l01=sapply(N_vector, function(n) lillie_power(type = "laplace01", N = n))
s_l01=sapply(N_vector, function(n) Shapiro_wilk_power(type = "laplace01", N = n))
plot(x=X,y=s_l01,type = "o",main="Laplace(0,1)",ylab = "Мощность",
xlab="Размер выборки, n",col="red",pch=1,lwd=2,xaxt="n")
lines(x=X,y=j_l01,type = "o",col="blue",pch=2,lwd=2,xaxt="n")
lines(x=X,y=l_l01,type = "o",col="green",pch=3,lwd=2,xaxt="n")
lines(x=X,y=p_l01,type = "o",col="violet",pch=5,lwd=2,xaxt="n")
legend("bottomright",legend=c("Шапиро-Уилк","Харке-Бер","Лиллиефорс",
"Хи-квадрат Пирсона"),col=c("red","blue","green","violet"),lwd=2,cex = 0.75)
axis(1, at=X,labels=N_vector, las=1)

p_g45=sapply(N_vector, function(n) pearson_power(type = "gamma45", N = n))
j_g45=sapply(N_vector, function(n) jarque_power(type = "gamma45", N = n))
l_g45=sapply(N_vector, function(n) lillie_power(type = "gamma45", N = n))
s_g45=sapply(N_vector, function(n) Shapiro_wilk_power(type = "gamma45", N = n))
plot(x=X,y=s_g45,type = "o",main="Gamma(4,5)",ylab = "Мощность",
xlab="Размер выборки, n",col="red",pch=1,lwd=2,xaxt="n")
lines(x=X,y=j_g45,type = "o",col="blue",pch=2,lwd=2,xaxt="n")
lines(x=X,y=l_g45,type = "o",col="green",pch=3,lwd=2,xaxt="n")
lines(x=X,y=p_g45,type = "o",col="violet",pch=5,lwd=2,xaxt="n")
legend("bottomright",legend=c("Шапиро-Уилк","Харке-Бер","Лиллиефорс",
"Хи-квадрат Пирсона"),col=c("red","blue","green","violet"),lwd=2,cex = 0.75)
axis(1, at=X,labels=N_vector, las=1)

p_g15=sapply(N_vector, function(n) pearson_power(type = "gamma15", N = n))
j_g15=sapply(N_vector, function(n) jarque_power(type = "gamma15", N = n))
l_g15=sapply(N_vector, function(n) lillie_power(type = "gamma15", N = n))
s_g15=sapply(N_vector, function(n) Shapiro_wilk_power(type = "gamma15", N = n))
plot(x=X,y=s_g15,type = "o",main="Gamma(1,5)",ylab = "Мощность",
xlab="Размер выборки, n",col="red",pch=1,lwd=2,xaxt="n")
```

```
lines(x=X,y=j_g15,type = "o",col="blue",pch=2,lwd=2,xaxt="n")
lines(x=X,y=l_g15,type = "o",col="green",pch=3,lwd=2,xaxt="n")
lines(x=X,y=p_g15,type = "o",col="violet",pch=5,lwd=2,xaxt="n")
legend("bottomright",legend=c("Шапиро-Уилк","Харке-Бер","Лиллиефорс",
"Хи-квадрат Пирсона"),col=c("red","blue","green","violet"),lwd=2,cex = 0.75)
axis(1, at=X,labels=N_vector, las=1)
```

Список литературы

[1] Ивченко Г. И., Медведев Ю.И. Математическая статистика: Учеб. пособие для втузов.— М.: Высш.шк., 1984. — 248 с.

[2] Zar, Jerrold H. Biostatistical analysis: Pearson new international edition. Pearson Higher Ed, 2013.

[3] М.Б. Лагутин. Наглядная математическая статистика: учебное пособие - 2-е изд., испр.— М.: БИНОМ. Лаборатория знаний, 2009. — 472 с.

[4] Rizzo M. L. Statistical computing with R. – CRC Press, 2019.

[5] Razali, Nornadiah Mohd, and Yap Bee Wah. "Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests."Journal of statistical modeling and analytics 2.1 (2011): 21-33.

[6] Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников. – 2012.

[7] Постовалов С. Н. Применение компьютерного моделирования для расширения прикладных возможностей классических методов проверки статистических гипотез //Дисс. на соискание уч. степенид. т. н., НГТУ, 2013г.–298с. – 2013.

[8] Боровков А. А. Математическая статистика. – 2007.