1. Введение

Имеется гипотеза:

$$H_0 : \text{Выборка из класса распределения N},$$

а также альтернативная гипотеза:

$$H_1 : H_0 \text{ не верна}.$$

Необходимо рассмотреть различные критерии и сравнить их по мощности для различных альтернатив.

2. Распределение χ^2, критерий согласия Пирсона

Для проверки того, соответсвует ли выборка объема n некоторому дискретному распределению, можно воспользоваться критерием согласия Пуассона. Критерий подразумевает работу с группированными данными, поэтому разделим область на k интервалов:

$$x_0,\, x_1,\, x_2, \ldots, x_k$$

После разбиения необходимо подсчитать, сколько в каждый из интервалов попало наблюдений n_i.

Поскольку сам критерий используется для проверки дискретных распределений, а задача стоит в проверке непрерывного распределения, введем теоретическую вероятность попадания в каждый из интервалов при известных значениях функции предполагаемого распределения $F(x, \hat{\theta})$, где $\hat{\theta}$ зависит от выборки:

$$P_i(\hat{\theta}) = F(x_i, \hat{\theta}) - F(x_{i-1}, \hat{\theta})$$

На основе введенных обозначений статистика критерия имеет вид:

$$t = n \sum_{i=1}^{k} \frac{\left(\frac{n_i}{n} - P_i(\hat{\theta})\right)^2}{P_i(\hat{\theta})}$$

В пределе статистика имеет распределение χ^2 с $k-p-1$ степенями свободы, где p — число параметров в $\hat{\theta}$.

В нашем случае в качестве $\hat{\theta}$ будет выступать $(\bar{x}, s^2)$ — параметры полученные с помощью метода моментов, F — функция нормального распределния.

3. Критерий согласия Колмогорова

Критерий согласия Колмогорова также служит для проверки того, соответсвует ли выборка объема n некоторому распределению. В качестве статистики служит следующая функция:

$$t = \sqrt{n} D_n, \qquad D_n = \sup_{x \in \mathbf{R}} \left| F_n(x) - F(x) \right|,$$

где $F_n(x)$ — эмпирическая функция распределения и имеет вид:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{X_i < x}, \qquad I_{X_i < x} = \begin{cases} 1, \ X_i \leq x \\ 0, \ X_i > x \end{cases}.$$

Кроме того, в силу возрастания $F(x)$ и кусочного постоянства $F_n(x)$ можно переписать D_n:

$$D_n = \max_{i=2,\ldots,n} \left\{ \left| \frac{i-1}{n} - F(X_i) \right|, \left| \frac{i}{n} - F(X_i) \right| \right\}$$

В пределе такая статистика имеет распределение Колмогорова $K(x)$:

$$K(x) = \sum_{j=-\infty}^{\infty} (-1)^j e^{-2j^2 x^2}$$

Также часто используют статистику с правками Большева:

$$t = \sqrt{n} D_n + \frac{1}{6\sqrt{n}}$$

4. Критерий Лиллиефорса

Критерий Лиллиефорса является модификацией критерия согласия Колмогорова специально для проверки соответствия выборки объема n нормальному закону. Статистика в этом случае имеет аналогичный вид D_n. Результат проверяется по таблице:

Размер выборки n	$\alpha = 0,20$	$\alpha = 0,15$	$\alpha = 0,10$	$\alpha = 0,05$	$\alpha = 0,01$
5	0,285	0,299	0,315	0,337	0,405
10	0,215	0,224	0,239	0,258	0,294
20	0,160	0,166	0,174	0,190	0,231
25	0,149	0,153	0,165	0,180	0,203
30	0,131	0,136	0,144	0,161	0,187
>30	$\frac{0.736}{\sqrt{n}}$	$\frac{0.768}{\sqrt{n}}$	$\frac{0.805}{\sqrt{n}}$	$\frac{0.886}{\sqrt{n}}$	$\frac{1.031}{\sqrt{n}}$

В случае, если посчитанный D_n меньше соответсвующего табличного значения, то нулевая гипотеза H_0 принимается, иначе отвергается.

5. Критерий Шапиро-Уилка

Критерий Шапиро-Уилка также как и предыдущий предназначен специально для проверки соответствия выборки объема n нормальному закону. Расчетное значение критерия получают по формуле:

$$W_{\text{расч}} = \frac{B^2}{s^2}, \qquad B^2 = \left| \sum_{i=1}^{[n/2]} a_i (X_{n-i+1} - X_i) \right|^2, \quad s^2 = \sum_{i=1}^{n} (X_i - \bar{X})^2.$$

Значения a_i берутся из таблиц. Однако для программной реализации удобнее использовать приближенный формулы:

$$a_i = a_0 \left(z + \frac{1483}{(3-z)^{10.845}} + \frac{71.6 \cdot 10^{-10}}{(1.1-z)^{8.26}} \right), \qquad a_0 = \frac{0.899}{(n-2.4)^{0.4162}} - 0.02,$$

$$z = \frac{n - 2i + 1}{n - 0.5}.$$

В зависимости от α выбирается соответсвующее $W_{\text{табл}}$:

α	$W_{\text{табл}}$
0.01	$(-0,0148n^4 + 2,1875n^3 - 122,61n^2 + 3257,3n + 55585)/100000$
0,05	$(-0,0113n^4 + 1,656n^3 - 91,88n^2 + 2408,6n + 67608)/100000$
0,1	$(-0,0084n^4 + 1,2513n^3 - 70,724n^2 + 1890n + 73840)/100000$

Если $W_{\text{расч}} < W_{\text{табл}}$, то нулевая гипотеза H_0 отвергается, иначе принимается.

6. Реализация

Для реализации каждого из критериев используется библиотека scipy. В качестве функции, отвечающей за критерий согласия Пирсона, используется рукописная функция. Для Шапиро-Уилка — scipy.stats.shapiro. Поскольку критерий Лиллиефорса является улучшенной версией критерия согласия Колмогорова в рамках проверки на нормальность, то реализовываться будет именно первый. Для него используется функция scipy.stats.kstest. Каждая из функций генерирует пару значение критерия и p-value. Для каждого набора результатов для различных выборок будем сопоставлять полученное p-value с частотой получения значений меньших текущего. Синяя кривая показывает критерий хи-квадрат, оранжевая — критерий Шапиро-Уилка, зеленая — критерий Лиллиефорса.

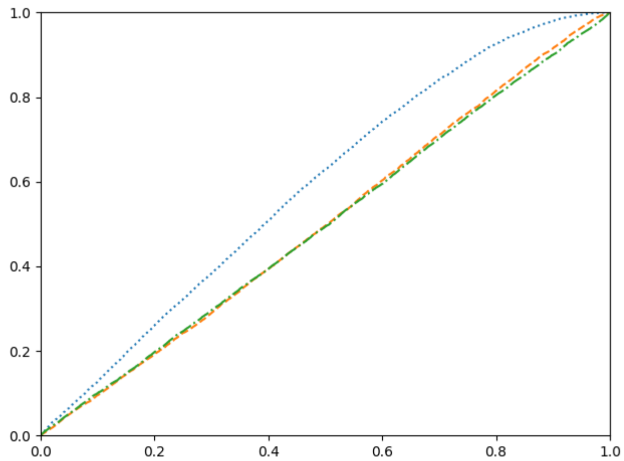

Рис. 6.1.: Проверка нулевой гипотезы Python.

В первую очередь проверим работоспособность критериев. Для этого будем генерировать выборки объема 100 с нормальным распределением (0, 1), с помощью встроенной в scipy функции scipy.stats.norm.rvs(size=100), и проверять их критериями. Критерий проверяется выборками 1000 раз. В результате должен образоваться рисунок близкий к прямой $y = x$

Как видно на рисунке 6.1, результаты, действительно, совпадают с предполагаемыми, однако график хи-квадрат очень сильно отклоняется от нее, а значит все критерии, кроме хи-квадрат, исправно работают для нормального распределения.

Из-за невозможности полноценно оценить каждый критерий, будем рассматривать реализацию на языке R. Принцип работы аналогичен предложенному ранее. Для критерия согласия Пирсона используется функция PearsonTest, для Шапиро-Уилка — ShapiroFranciaTest, для Лиллиефорса — LillieTest. Проведем проверку, аналогичную ранее проведенной.

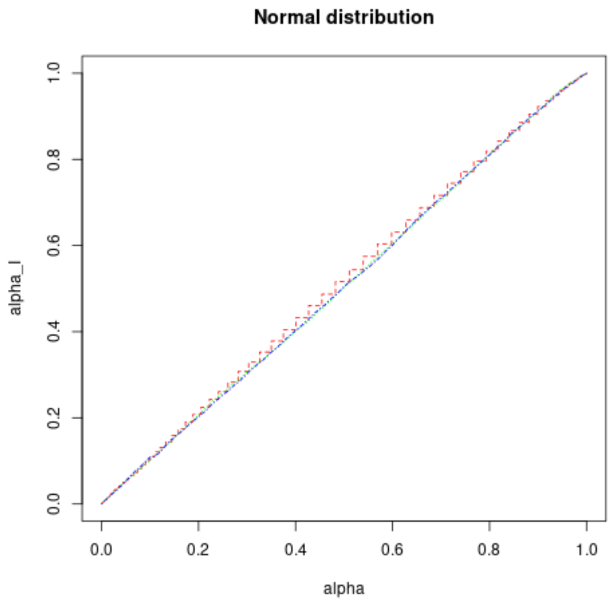

Рис. 6.2.: Проверка нулевой гипотезы R

Из рисунка 6.2 видно, что результаты по проверке нулевой гипотезы полностью совпадают с их теоретическим поведением.

Ниже представлены результаты расчетов для различных альтернатив. Для каждой выборки будет представлен коэфициент эксцесса, а также какой из критериев лучше работает.

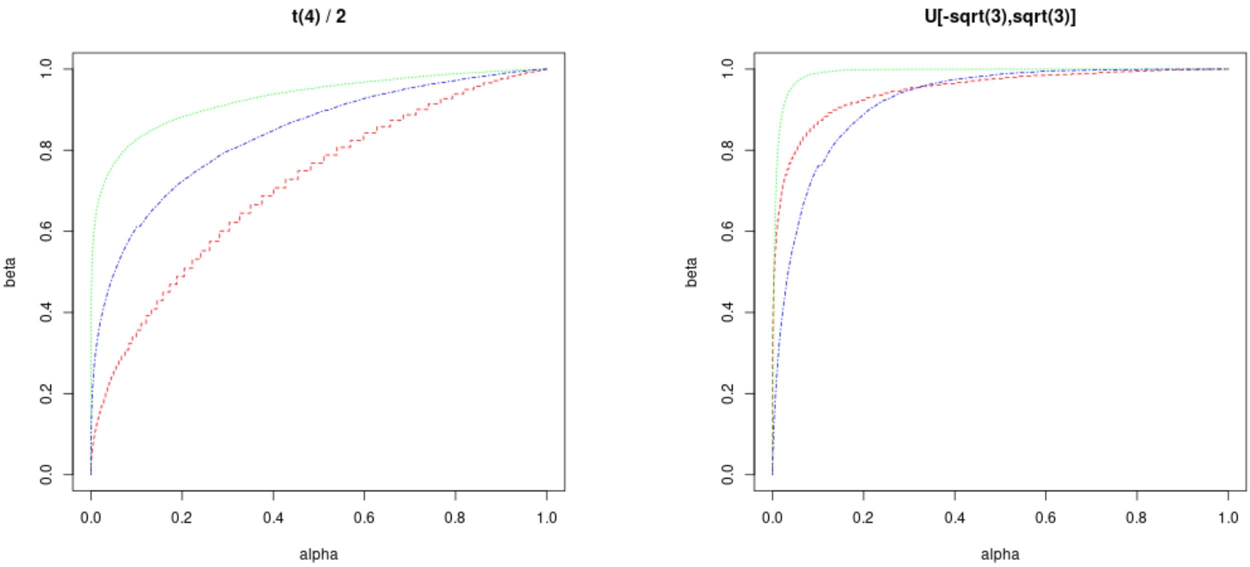

Рис. 6.3.: Зависимость мощностей от α для разных альтернатив

Красная кривая указывает на критерий Пирсона, синяя — на критерий Лиллиефорса, зеленая — на критерий Шапиро-Уилка. Теперь, опираясь на рисунок 6.3, определим какой критерий лучше в отдельной ситуации и занесем все данные в таблицу, будем рассматривать выборки с мат. ожиданием 0 и единичной дисперсией.

Распределение	Генерация выборки в языке R	Коэфициент эксцесса γ_2	Критерий Пирсона	Критерий Шапиро-Уилка	Критерий Лиллиефорса
$t(4)/2$	rt(100, 4) / 2	$\gamma_2 > 0$	3	1	2
$U\left[-\sqrt{3}, \sqrt{3}\right]$	runif(100, -sqrt(3), sqrt(3))	$\gamma_2 < 0$	2	1	3

Таблица 6.1.: Результаты сравнения мощностей для различных альтернатив

Основываясь на результатах сравнения, можем заключить, что для $\gamma_2 > 0$ ориентируясь по сумме мест, лучшим оказался критерий Шапиро-Уилка, следующим — критерий Лиллиефорса, а худшим — критерий Пирсона. В предложенной к рассмотрению статье для $\gamma_2 > 0$, напротив, худшим из трех предложенных оказался критерий Лиллиефорса. Также там было указано, что для $\gamma_2 < 0$ лучшим так и остался критерий Шапиро-Уилка, а худшим теперь является критерий Пирсона. Однако, исходя из проведенных расчетов, худшим является критерий Лиллиефорса, а критерий Шапиро-Уилка так и остался лучшим.

Вполне возможно, что в рамках этой работы было рассмотрено недостаточное количество различных распределений.

7. Вывод

В ходе проделанной работы были рассмотрены три критерия для проверки поставленной ранее гипотезы. Помимо этого для каждого из них было рассмотрено шесть различных альтернатив. И на основе сравнения по каждому было выявлено, что для проверки гипотезы нормальности наилучший критерием оказался критерий Шапиро-Уилка, а критерии Пирсона и Лиллиефорса показали себя примерно одинаково, однако, хуже первого.

A. Реализация зависимости β от α на языке R

```
library(DescTools)

N = 10000
x = seq(0, 1, length.out = N)
chi2_results = c(0, 1)
shapiro_results = c(0, 1)
kstest_resultsf = c(0, 1)
for(i in 3:N){
        sample = rnorm(100, 0, 1)
        chi2_results[i] <- PearsonTest(sample, n.classes=21)['p.value']
        shapiro_results [i] <- ShapiroFranciaTest(sample  )['p.value']
        kstest_resultsf [i] <- LillieTest(sample )['p.value']
}
chi2_results= sort(unlist(chi2_results), decreasing = FALSE)
shapiro_results = sort(unlist(shapiro_results ), decreasing = FALSE)
kstest_resultsf  = sort(unlist(kstest_resultsf  ), decreasing = FALSE)

plot(chi2_results, x, pch=100, type="l",
xlim = c(0,1), ylim = c(0,1),
main="Sample Name", xlab="alpha", ylab="beta", lty = 2, col="red")
lines(shapiro_results, x, pch=100, lty=3, col="green")
lines(kstest_resultsf, x, pch=100, lty=4, col="blue")
```