1. А. И. Маркушевич Замечательные кривые М., 1978 г., 48 стр. с ил.
2. Вольф Д. OpenGL 4. Язык шейдеров. Книга рецептов/ пер. с англ. А. Н. Киселева – М.: ДМК Пресс, 2015. – 368 стр. с ил.
3. Выгодский М.Я. Справочник по высшей математике. Изд-во «Наука». М. 1977 г.
4. Штейнгауз Г. Математический калейдоскоп: Пер. с польского. – М.: Наука. Главная редакция физико-математической литературы, 1981 г.

УДК 615-073

Артемьева А.

ИССЛЕДОВАНИЕ ПОВЕДЕНИЯ ТЕСТОВ НА МНОЖЕСТВЕННЫЕ СРАВНЕНИЯ, ПРИМЕНЁННЫХ К НЕНОРМАЛЬНЫМ НЕОДНОРОДНЫМ РАСПРЕДЕЛЕНИЯМ, МЕТОДОМ МОНТЕ-КАРЛО

INVESTIGATION OF THE BEHAVIOR OF MULTIPLE COMPARISON TESTS APPLIED TO ABNORMAL NON-UNIFORM DISTRIBUTIONS USING THE MONTE-CARLO METHOD

Филиал «Протвино» государственного университета «Дубна»
Секция «Естественные и инженерные науки»

Автор: Артемьева Анастасия, студентка 4 курса направления «Физика» филиала «Протвино» государственного университета «Дубна».

Научный руководитель: Масликов Александр Альбертович, кандидат физико-математических наук, доцент кафедры технической физики филиала «Протвино» государственного университета «Дубна».

Author: Artemeva Anastasia, 4d year student of the direction "Physics" of the branch "Protvino" state University "Dubna".

Scientific adviser: Maslikov Alexander Albertovich, candidate of physical and mathematical Sciences, associate professor of the department of technical physics of the branch "Protvino" state University "Dubna".

Аннотация

Методом Монте-Карло симулируется использование тестов выполняемых при множественных сравнениях на выборках малого объёма из искаженных (по сравнению с нормальным) распределений. Вычисляются эффективные вероятности ошибок 1-го рода и делаются выводы о влиянии негомогенности дисперсий, «ненормальности» эксцесса и асимметрии на эффективность исследуемых критериев.

Abstract

The Monte Carlo method simulates the use of tests performed in multiple comparisons on small-volume samples from distorted (compared to normal) distributions. The effective probabilities of errors of the 1st kind are calculated and conclusions are drawn about the influence of inhomogeneity of variances,"abnormality" of excesses and asymmetry on the effectiveness of the studied criteria.

Ключевые слова: математическая статистика, множественные сравнения, ANOVA.

Keywords: Mathematical statistics, multiple comparisons, ANOVA.

Отличительной чертой современной фармакологии является использование продвинутых статистических методов анализа данных. После проведения биохимического анализа формируется массив данных, который требует обработки. Выбор параметров описательной статистики основывается на типе данных и форме их распределения. Характеристику типа распределения данных дают с использованием аналитических статистических критериев.

Мы проанализируем возможности использования классических критериев Шапиро-Уилка, Левене, дисперсионного анализа (ANOVA), а также непараметрического критерия Краскела—Уоллиса для выборок с ненормальным распределением. Нас будет интересовать – как ведут себя эти критерии на ненормальных и/или неоднородных по дисперсии выборках малого объёма.

Для исследования мы использовали программы, написанные в пакете WolframMathematica (WM). Идея в том, чтобы генерировать выборки из заданных распределений и проверять, насколько частоты попадания результатов в критическую область обсуждаемого критерия близки к теоретически ожидаемым частотам.

Пакет аналитических вычислений WM позволяет формировать компактные программы для генерации огромного числа случайных выборок (мы использовали 10^5—10^6) из большого набора библиотечных распределений. Далее эти выборки можно прогонять через интересующие тесты и методом Монте-Карло получать статистически достоверные результаты. В свете формата фармакокинетических исследований нас будет интересовать случай малых выборок объёмом 5—10 элементов. Отклонения исследуемых распределений от нормального Гауссового распределения можно описывать величинами эксцесс (E) и асимметрия (A). Для нормального распределения коэффициент островершинности E =3, а A=0.

Корректность работы программ устанавливается путём применения критериев к выборкам из нормальных распределений. Так Shapiro-Wilk, ANOVA, Kraskal-Wallis выдают значения для вероятности ошибки 1-го рода при стандартном квантиле уровня значимости $\alpha = 0.05$: в диапазоне 0.050—0.051, что соответствует ожидаемым из теории значениям.

В классической формулировке тест ANOVA (множественные сравнения выборок на основе дисперсионного анализа) требует нормальности и однородности по дисперсиям выборок. Представляется интересным исследовать, насколько катастрофичным для ANOVA будет утрата выборками нормальности и/или однородности. Рассмотрим случай 4-х выборок объема 5 с кардинально нарушенной гомогенностью. Будем рассматривать выборки из нормальных распределений с геометрически растущим среднеквадратическим отклонением (СКО): σ, 2σ, 4σ, 8σ и одинаковым мат-ожиданием (МО). Метод Монте-Карло, состоящий в, данном случае, в применении теста Левене к 10^5 четвёрок таких выборок позволяет сделать вывод, что ~73% этих четвёрок не проходят тест на гомогенность на уровне значимости $\alpha = 0.05$. Теперь на этом же материале проведём дисперсионный анализ в стандартной форме с $\alpha = 0.05$. В результате получаем, что неоднородность приводит к эффективному увеличению вероятности ошибки 1-го рода всего до 0.10 против 0.05 в случае гомогенности дисперсий.

Далее проведём исследования для критериев ANOVA и Краскела—Уоллиса с контролем нормальности по Шапиро-Вилка. Будем брать выборки с отклонениями по эксцессу и асимметрии по сравнению с нормальным распределением. Для этого в качестве тестовых используем по 4 выборки различного объёма из распределений Стьюдента, хи-квадрат и логнормального.

Множественные сравнения (ANOVA, Краскел—Уоллис) для выборок из асимметричных распределений (логнормальное, хи-квадрат) мы исследуем в 2-х вариантах: 1) все 4 выборки асимметричны в одну сторону; 2) асимметрия 2-х выборок противоположна 2-м другим (в этом случае 2 выборки берутся отражёнными относительно МО (или медианы), чтобы мат-ожидания (или медианы для непараметрического критерия) оставались неизменными. Итак, мы прогоняем через указанные тесты большое число ($>10^5$) четвёрок выборок из известных распределений (используя стандартные квантили для $\alpha=0.05$) и, в соответствии с методом Монте-Карло, получаем эмпирическую вероятность ошибки 1-го рода. Результаты представлены в таблицах.

Для логнормального распределения мы выберем параметры $\mu=0$, $\sigma = 1$, тогда $E = 114$, асимметрия $A = 6.18$, мат-ожидание $M=\sqrt{e}$, медиана $m=1$. Получим следующие значения:

Таблица 1. Наблюдаемая вероятность ошибки 1-го рода для логнормального распределения.

Выборки\Критерии	Shapiro-Wilk	ANOVA	Kruskal-Wallis
4 одинаковых логнормальных распределения	0.24	0.033	0.055
2 логнормальных распределения 2 отражённых логнормальных	0.24	0.100	0.09

Критерий Shapiro-Wilk уверенно указывает на ненормальность распределений, однако, в случае сонаправленности асимметрий это не сказывается критическим образом на работе тестов ANOVA и Краскел—Уоллис. Только при разнонаправленных асимметриях начинаются некоторые проблемы, возрастает вероятность ошибки 1-го рода.

Теперь в качестве тестового будем использовать распределение хи-квадрат (χ^2). Во второй таблице представлены результаты теста Shapiro-Wilk проведённого для выборок разного объёма из распределений (χ^2) с различным числом степеней свободы (df).

Таблица 2. Тест Shapiro-Wilk (контроль нормальности) для выборок из распределений хи-квадрат (χ^2). 10^6 испытаний.

Эксцесс (E), асимметрия (A)	df \ объём выборки	5	6	7	8	9	10
E=6; A=1.41;	df=4	0.097	0.125	0.154	0.185	0.214	0.244
E=5; A=1.15;	df=6	0.078	0.096	0.117	0.137	0.155	0.174
E=4.5; A=1;	df=8	0.070	0.083	0.099	0.114	0.127	0.142

Видно, что тест Shapiro-Wilk уверенно обнаруживает ненормальность распределений.

Теперь исследуем – как это скажется на работе ANOVA. Следующие две таблицы представляют собой тест ANOVA (множественное сравнение средних значений) проведённый для выборок из распределений хи-квадрат (χ^2):

Таблица 3. Тест ANOVA для выборок из распределений хи-квадрат. 4 одинаковых χ^2 распределения.10^6 испытаний.

Эксцесс (E), ас-ия (A)	df \ объём выборки	5	6	7	8	9	10
E=6; A=1.41;	df=4	0.0449	0.045	0.0454	0.0457	0.0462	0.0460
E=5; A=1.15;	df=6	0.0465	0.0465	0.0468	0.0467	0.0475	0.0474
E=4.5; A=1;	df=8	0.0474	0.0470	0.0480	0.0474	0.0477	0.048

Таблица 4. Тест ANOVA для выборок из распределений хи-квадрат (χ^2). 2 нормальных 2 отраженных χ^2 распределения.10^6 испытаний.

Эксцесс (E), ас-ия (A)	df \ объём выборки	5	6	7	8	9	10
E=6; A=1.41;	df=4	0.0580	0.0572	0.0571	0.0561	0.0557	0.0553
E=5; A=1.15;	df=6	0.0552	0.0548	0.055	0.0540	0.0538	0.0538
E=4.5; A=1;	df=8	0.0535	0.0539	0.053	0.0528	0.053	0.053

Видим, что ненормальность распределений не сказалась критическим образом на работоспособности ANOVA – ошибка 1-го рода мало изменилась. Но в случае разнонаправленных асимметрий ситуация несколько хуже, как было и в предыдущем примере.

Результаты в таблицах 5 и 6 представлены для проведённого теста Kruskal-Wallis (множественное сравнение медиан) для выборок из распределений хи-квадрат (χ^2).

Таблица 5. Тест Kruskal-Wallis (множественное сравнение медиан) для выборок из распределений хи-квадрат (χ^2). 4 одинаковых χ^2 распределения. 10^6 испытаний.

Эксцесс (E), ас-ия (A)	df \ объём выборки	5	6	7	8	9	10
E=6; A=1.41;	df=4	0.056	0.054	0.053	0.053	0.052	0.051
E=5; A=1.15;	df=6	0.055	0.053	0.053	0.052	0.0519	0.052
E=4.5; A=1;	df=8	0.0553	0.054	0.0529	0.0528	0.052	0.0517

Таблица 6. Тест Kruskal-Wallis для выборок из распределений хи-квадрат (χ^2). 2 нормальных и 2 отражённых χ^2 распределения.10^6 испытаний.

Эксцесс (E), ас-ия (A)	Медиан a(m)	df \ объём выборки	5	6	7	8	9	10
E=6; A=1.41;	3.36	df=4	0.068	0.071	0.073	0.077	0.08	0.083
E=5; A=1.15;	5.34	df=6	0.063	0.064	0.066	0.068	0.07	0.071

E=4.5; A=1;	7.34	df=8	0.0614	0.0616	0.0625	0.0634	0,065	0,067

Видим, что опять нет критического роста ошибки 1-го рода, но в целом тест Kruskal-Wallis ведёт себя менее устойчиво, чем ANOVA.

Далее мы провели аналогичные тесты для выборок из распределений Стьюдента. Результаты можно увидеть в таблицах 7 - 9:

Таблица 7. Тест Shapiro-Wilk (проверка гипотезы нормальности) для выборок из распределений Стьюдента. 10^6 испытаний.

Эксцесс (E), ас-ия (A)	df \ объём выборки	5	6	7	8	9	10
E=9; A=0;	df=5	0.066	0.077	0.087	0.097	0.106	0.114
E=6; A=0;	df=6	0.061	0.07	0.078	0.087	0.092	0.098
E=5; A=0;	df=7	0.059	0.065	0.072	0.079	0.084	0.089

Таблица 8. Тест ANOVA (множественное сравнение средних значений) для выборок из распределений Стьюдента. 10^6 испытаний.

Эксцесс (E), ас-ия (A)	df \ объём выборки	5	6	7	8	9	10
E=9; A=0;	df=5	0.045	0.045	0.046	0.0462	0.0465	0.047
E=6; A=0;	df=6	0.046	0.0462	0.047	0.0475	0.047	0.045
E=5; A=0;	df=7	0.0467	0.047	0.0471	0.0477	0.048	0.048

Таблица 9. Тест Kruskal-Wallis (множественное сравнение медиан) для выборок из распределений Стьюдента. 10^6 испытаний.

Эксцесс (E), ас-ия (A)	df \ объём выборки	5	6	7	8	9	10
E=9; A=0;	df=5	0.0554	0.0535	0.053	0.052	0.052	0.052
E=6; A=0;	df=6	0.0555	0.0531	0.053	0.0526	0.0519	0.0518
E=5; A=0;	df=7	0.0552	0.0537	0.0528	0.0525	0.052	0.0516

Теоретические эксцессы рассматриваемых распределений Стьюдента заметно выше нормального. Видно, что наши выборки из распределений Стьюдента тест Shapiro-Wilk вполне уверенно определяет как ненормальные, в то же время ANOVA и тест Kruskal-Wallis продолжают надёжно работать, почти не искажая вероятность ошибки 1-го рода. Следуя ANOVA мы излишне рискуем посчитать выборки различающимися, если, например пронаблюдается $\alpha = 0.049$. А непараметрический его аналог тест Kruskal-Wallis наоборот более консервативен. Т.е. есть риск не увидеть различий там, где при выполнении условия нормальности они были бы обнаружены.

В целом же, изменение эксцесса (у нас в 3 раза) некатастрофично для ANOVA и KruskalWallis. В этом случае можно рекомендовать их совместное применение. При согласии их выводов заключение можно считать достоверным.

Заключение

Итак, мы может сделать следующие выводы:

1) негомогенность дисперсий (даже сильная) не оказывает катастрофического влияния на критерий ANOVA. В нашем примере с геометрическим ростом СКО наблюдалось возрастание эффективной вероятности ошибки 1-го рода всего с 5% до 10%;

2) тесты множественных сравнений (ANOVA, Kruskal-Wallis) практически нечувствительны к увеличению эксцесса по сравнению с нормальным распределением при условии сонаправленности асимметрий;

3) наблюдается некоторое ухудшение ситуации в случае с распределениями, имеющими разнонаправленные асимметрии.

Представляется полезным итерационным методом найти уровни значимости, соответствующие статистически наблюдаемым ошибкам 1-го рода для используемых «ненормальных» распределений. Имеет смысл провести аналогичные исследования для других распределений с малыми объёмами выборок. Также мы планируем аналогичным методом исследовать влияние «ненормальности» на мощность критериев.

Библиографический список

1. Гмурман В.Е. Теория вероятностей и математическая статистика – М., Высшая школа, 2003.
2. Эверитт, Б.С. Большой словарь по статистике / Б.С. Эверитт. - М.: Проспект, 2012. - 736 с.
3. Upton, GrahamCook, Ian. Dictionaryofstatistics. - OxfordUniversityPress, Оксфорд, Великобритания, 2006. – 464 с.
4. Лапач С.Н., Чубенко А.В., Бабич П.Н., Статистические методы в медико-биологических исследованиях с использованием Excel. – 2-е изд. перераб. и доп. – К.: МОРИОН, 2001. – 408 с.
5. С. Гланц. Медико-биологическая статистика. Пер. с англ. — М., Практика, 1998. — 459 с.
6. Кобзарь А. И. «Прикладная математическая статистика», для инженеров и научных работников, 2008. Размер 8,1 Мб, 816 с.

УДК 20.51.01

Ахтырский А.И.

РАЗВИТИЕ ИГРОВОГО ПРОЕКТА“SPACE STATION 13” НА ПЛАТФОРМЕ BYOND ЯЗЫКА DREAM MAKER

DEVELOPMENT OF THE “SPACE STATION 13” GAME PROJECT ON THE BYOND PLATFORM OF THE DREAM MAKER LANGUAGE

Филиал «Протвино» государственного университета «Дубна»
Секция «Информационные технологии»