Министерство образования и науки РФ

Новосибирский государственный технический университет

Кафедра ТПИ

**Курсовой проект**

**По дисциплине:** «Компьютерные технологии анализа данных и исследования статистических закономерностей»

**Тема:** «Исследование мощности критериев равномерности Морана, Ченга-Спринга, Шермана для проверки статистических гипотез»

Студент: Кочнев А.В.

Факультет: ПМИ

Группа: ПММ-71

Преподаватель: Постовалов С.Н.

Новосибирск 2019

Оглавление

[ВВЕДЕНИЕ 3](#_Toc533546570)

[1. ПОСТАНОВКА ЗАДАЧИ 4](#_Toc533546571)

[2. КРИТЕРИЙ МОРАНА 4](#_Toc533546572)

[3. КРИТЕРИЙ ЧЕНГА-СПРИНГА 5](#_Toc533546573)

[4. КРИТЕРИЙ ШЕРМАНА 6](#_Toc533546574)

[5. ИССЛЕОВАНИЯ 8](#_Toc533546575)

[ВЫВОДЫ 15](#_Toc533546576)

[СПИСОК ЛИТЕРАТУРЫ 16](#_Toc533546577)

[ПРИЛОЖЕНИЕ. ТЕКСТ ПРОГРАММЫ 17](#_Toc533546578)

## ВВЕДЕНИЕ

В прикладной математической статистике равномерный закон распределения вероятностей занимает видное место. В определенной степени обилие критериев обусловлено тем интересом, который проявляется к использованию модели равномерного закона в различных приложениях Частота применения модели равномерного закона в задачах статистического анализа в приложениях не в последнюю очередь определяется тем, что использование такой простой модели во многих ситуациях позволяет найти решение задачи с опорой только на аналитические методы. Равномерный закон, например, зачастую используется для описания ошибок измерений некоторых приборов или измерительных систем.

Критериев, ориентированных на проверку гипотезы о принадлежности выборки равномерному закону, достаточно много. Это, например, такие критерии, как критерии: Шермана, Кимбелла, Морана, Ченга–Спиринга, Хегази–Грина, Янга, Фросини, Гринвуда, Неймана–Бартона и др. [1]

В данной работе будут рассматриваться критерии Морана, Ченга-Спринга и критерий Шермана.

Целью данной работы является знакомство с современными тенденциями развития аппарата прикладной математической статистики и состоянием программного обеспечения задач статистического анализа. Освоение методов статистического моделирования как средства исследования и развития аппарата прикладной математической статистики. Исследование особенностей методов оценивания и критериев проверки статистических гипотез. Исследование статистических свойств оценок. Исследование мощности критериев относительно различных конкурирующих гипотез. Сравнительный анализ мощности групп критериев, ориентированных на проверку аналогичных гипотез. Закрепление навыков проведения самостоятельных исследований.

## ПОСТАНОВКА ЗАДАЧИ

Пусть имеется равномерно распределенная случайная величина X на интервале [0,1] Функция распределения вероятностей равномерного закона имеет вид  при . При проверке гипотезы о принадлежности наблюдаемой случайной величины равномерному закону простая проверяемая гипотеза имеет вид:  .

Пусть  – выборка независимых наблюдений случайной величины . По элементам выборки  строится вариационный ряд: . В статистиках критериев используются порядковые статистики вида , , , .

## КРИТЕРИЙ МОРАНА

Статистика критерия Морана разбивается на два вида статистик: Морана 1 и Морана 2. Статистика Морана 1 предложена в [3] очень близка к статистике критерия Кимбелла и имеет вид:

, (1)

где , .

Критерий является правосторонним, то есть гипотеза  отклоняется при больших значениях статистики .

Процентные точки распределения статистики (1), полученные при статистическом моделировании представлены в работе [1]. Распределение статистики критерия имеет зависимость от объёма выборок n при справедливости проверяемой гипотезы , что является существенным недостатком данного критерия, также, как и необходимость использования таблиц процентных точек.

Статистика критерия Морана 2, предложенная в [4] имеет вид:

 (2)

где , .

При заданном уровне значимости α проверяемая гипотеза о равномерности не отклоняется при . Критические значения статистики (2), полученные в результате статистического моделирования представлены в [1].

Также в работе [1] говориться о том, что оба критерия обладают не очень высокой мощностью и является не предпочтительным для применения.

## КРИТЕРИЙ ЧЕНГА-СПРИНГА

Данный критерий предложен в работе [5]. Статистика критерия имеет вид:

 (2)

где разность  представляет собой выборочный размах, а

.

Критерий является двусторонним, то есть проверяемая гипотеза не отклоняется, если . Критические значения представлены в работе [1]. В работе также говорится, что данный критерий более мощный, чем предыдущий, однако некоторые альтернативы может не различить.

К недостаткам использования критерия относится зависимость распределения статистики от объема выборки n и необходимость использования таблицы процентных точек.

## КРИТЕРИЙ ШЕРМАНА

Статистика критерия Шермана для проверки равномерности распределения имеет вид [6]:

. (3)

Критерий является правосторонним, то есть гипотеза  отклоняется при больших значениях статистики .

При справедливости  и больших объёмах выборок  статистика  приближенно подчиняется нормальному распределению [1]. Данный критерий не будет исследоваться в данной работе в силу того, что нам неизвестны параметры предельного нормального распределения, к которому сходится распределение статистики Шермана. Поэтому рассмотрим модификации критерия Шермана.

Нормализованная статистика Шермана имеет вид:

, (4)

где математическое ожидание и дисперсия статистики  определяются выражениями:

 . (5)

При  нормализованная статистика Шермана сходится к стандартному нормальному закону.

Также в [6] приводиться еще одна модификация статистики Шермана, вида:

, (6)

. (7)

Распределение модифицированной статистики  ещё быстрее сходится к стандартному нормальному закону.

В формуле 7 константа 0.3679 получена из:

.

А выражение  было найдено путем следующих манипуляций:



.

Из-за таких грубых округлений точность падает больше чем в два раза и скорость сходимость увеличивается в разы. Поэтому в рамках данной работы было принято решение о том, чтобы не использовать приблизительные значения формулы 7, а вычислять математическое ожидание и дисперсию пользуясь формулой 5.

## ИССЛЕОВАНИЯ

В данной работе предполагается исследовать мощность критериев. Мощностью критерия называется величина , где вероятность ошибки второго рода, т.е. мощность – вероятность того, что при справедливой конкурирующей гипотезе  будет отклонена гипотеза . Очевидно, что чем выше мощность используемого критерия при заданном значении , тем лучше он различает гипотезы  и .

В качестве меры близости конкурирующих гипотез  к гипотезе  будет использоваться расстояние:

Количество повторений моделирования статистик в методе Монте-Карло .

Для исследования мощность основная гипотеза имеет вид: о принадлежности наблюдаемой случайной величины равномерному закону, конкурирующие гипотезы подбиралось, основываясь на Бета-распределение 1-го рода. совпадает с равномерным распределением, поэтому параметры варьировались не отходя далеко от изначальных, чтобы расстояние между и конкурирующей гипотезой было небольшим. А также на обобщенно нормальном распределении .

|  |  |  |
| --- | --- | --- |
| Гипотеза | Закон распределения | Расстояние |
|  |  | 0.1056659 |
|  |  | 0.1054688 |
|  |  | 0.1051284 |
|  |  | 0.0193088 |
|  |  | 0.0693088 |
|  |  | 0.0985638 |

Рассмотрим исследования мощности при разных объемах выборок, выбрав в качестве конкурирующих гипотез те, которые основаны на бета распределении.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Объем выборок | | | | |
|  |  |  |  |  |
| Моран 1 | 0.15 | 0.185 | 0.227 | 0.21 |
| 0.1 | 0.129 | 0.163 | 0.149 |
| 0.05 | 0.069 | 0.092 | 0.082 |
| 0.025 | 0.036 | 0.052 | 0.045 |
| 0.01 | 0.015 | 0.023 | 0.019 |
| Моран 2 | 0.15 | 0.179 | 0.247 | 0.216 |
| 0.1 | 0.122 | 0.178 | 0.153 |
| 0.05 | 0.065 | 0.104 | 0.088 |
| 0.025 | 0.035 | 0.061 | 0.049 |
| 0.01 | 0.014 | 0.029 | 0.022 |
| Ченга-Спринг | 0.15 | 0.129 | 0.155 | 0.19 |
| 0.1 | 0.082 | 0.105 | 0.132 |
| 0.05 | 0.039 | 0.054 | 0.072 |
| 0.025 | 0.019 | 0.028 | 0.041 |
| 0.01 | 0.007 | 0.012 | 0.019 |
| Шерман 1 | 0.15 | 0.19 | 0.234 | 0.182 |
| 0.1 | 0.131 | 0.167 | 0.127 |
| 0.05 | 0.071 | 0.095 | 0.070 |
| 0.025 | 0.037 | 0.053 | 0.038 |
| 0.01 | 0.016 | 0.024 | 0.016 |
| Шерман 2 | 0.15 | 0.191 | 0.236 | 0.182 |
| 0.1 | 0.133 | 0.171 | 0.128 |
| 0.05 | 0.071 | 0.096 | 0.069 |
| 0.025 | 0.038 | 0.055 | 0.037 |
| 0.01 | 0.016 | 0.026 | 0.016 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Объем выборок | | | | |
|  |  |  |  |  |
| Моран 1 | 0.15 | 0.211 | 0.248 | 0.233 |
| 0.1 | 0.15 | 0.18 | 0.168 |
| 0.05 | 0.081 | 0.101 | 0.093 |
| 0.025 | 0.045 | 0.057 | 0.053 |
| 0.01 | 0.021 | 0.027 | 0.024 |
| Моран 2 | 0.15 | 0.197 | 0.263 | 0.235 |
| 0.1 | 0.136 | 0.192 | 0.168 |
| 0.05 | 0.073 | 0.112 | 0.095 |
| 0.025 | 0.039 | 0.065 | 0.054 |
| 0.01 | 0.017 | 0.031 | 0.024 |
| Ченга-Спринг | 0.15 | 0.135 | 0.177 | 0.183 |
| 0.1 | 0.087 | 0.123 | 0.13 |
| 0.05 | 0.0412 | 0.066 | 0.071 |
| 0.025 | 0.02 | 0.036 | 0.039 |
| 0.01 | 0.007 | 0.016 | 0.18 |
| Шерман 1 | 0.15 | 0.212 | 0.255 | 0.207 |
| 0.1 | 0.147 | 0.185 | 0.145 |
| 0.05 | 0.08 | 0.107 | 0.08 |
| 0.025 | 0.044 | 0.061 | 0.043 |
| 0.01 | 0.019 | 0.028 | 0.02 |
| Шерман 2 | 0.15 | 0.212 | 0.253 | 0.205 |
| 0.1 | 0.151 | 0.186 | 0.146 |
| 0.05 | 0.082 | 0.108 | 0.081 |
| 0.025 | 0.045 | 0.062 | 0.046 |
| 0.01 | 0.02 | 0.029 | 0.021 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Объем выборок | | | | |
|  |  |  |  |  |
| Моран 1 | 0.15 | 0.271 | 0.294 | 0.282 |
| 0.1 | 0.2 | 0.219 | 0.208 |
| 0.05 | 0.117 | 0.128 | 0.123 |
| 0.025 | 0.069 | 0.074 | 0.071 |
| 0.01 | 0.032 | 0.035 | 0.033 |
| Моран 2 | 0.15 | 0.234 | 0.306 | 0.273 |
| 0.1 | 0.167 | 0.228 | 0.2 |
| 0.05 | 0.092 | 0.0134 | 0.116 |
| 0.025 | 0.05 | 0.079 | 0.067 |
| 0.01 | 0.021 | 0.037 | 0.03 |
| Ченга-Спринг | 0.15 | 0.156 | 0.227 | 0.178 |
| 0.1 | 0.106 | 0.167 | 0.125 |
| 0.05 | 0.053 | 0.096 | 0.066 |
| 0.025 | 0.026 | 0.056 | 0.035 |
| 0.01 | 0.011 | 0.028 | 0.015 |
| Шерман 1 | 0.15 | 0.257 | 0.303 | 0.253 |
| 0.1 | 0.184 | 0.225 | 0.185 |
| 0.05 | 0.108 | 0.136 | 0.107 |
| 0.025 | 0.031 | 0.08 | 0.061 |
| 0.01 | 0.029 | 0.04 | 0.029 |
| Шерман 2 | 0.15 | 0.257 | 0.302 | 0.251 |
| 0.1 | 0.186 | 0.226 | 0.184 |
| 0.05 | 0.107 | 0.138 | 0.107 |
| 0.025 | 0.060 | 0.083 | 0.061 |
| 0.01 | 0.029 | 0.041 | 0.029 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Объем выборок | | | | |
|  |  |  |  |  |
| Моран 1 | 0.15 | 0.355 | 0.356 | 0.345 |
| 0.1 | 0.274 | 0.273 | 0.262 |
| 0.05 | 0.171 | 0.168 | 0.162 |
| 0.025 | 0.107 | 0.101 | 0.098 |
| 0.01 | 0.057 | 0.052 | 0.051 |
| Моран 2 | 0.15 | 0.277 | 0.353 | 0.315 |
| 0.1 | 0.202 | 0.27 | 0.236 |
| 0.05 | 0.114 | 0.165 | 0.140 |
| 0.025 | 0.064 | 0.102 | 0.084 |
| 0.01 | 0.029 | 0.053 | 0.041 |
| Ченга-Спринг | 0.15 | 0.188 | 0.278 | 0.178 |
| 0.1 | 0.131 | 0.211 | 0.122 |
| 0.05 | 0.071 | 0.133 | 0.065 |
| 0.025 | 0.039 | 0.081 | 0.035 |
| 0.01 | 0.017 | 0.043 | 0.015 |
| Шерман 1 | 0.15 | 0.313 | 0.364 | 0.312 |
| 0.1 | 0.233 | 0.281 | 0.235 |
| 0.05 | 0.14 | 0.176 | 0.143 |
| 0.025 | 0.083 | 0.111 | 0.086 |
| 0.01 | 0.042 | 0.059 | 0.044 |
| Шерман 2 | 0.15 | 0.312 | 0.362 | 0.314 |
| 0.1 | 0.232 | 0.276 | 0.235 |
| 0.05 | 0.141 | 0.173 | 0.143 |
| 0.025 | 0.083 | 0.106 | 0.085 |
| 0.01 | 0.041 | 0.055 | 0.042 |

По таблицам видно, критерий Ченга-Спринга является смещенным при гипотезе и маленьких объемах выборок N=10, 20. В большинстве можно отдать предпочтение критериям Морана, однако стоит отметить, что при разных альтернативных гипотезах наибольшая мощность у разных критериев.

Рассмотрим исследования мощности при разных объемах выборок, выбрав в качестве конкурирующих гипотез те, которые основаны на обобщенно нормальном распределении.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Объем выборок | | | | |
|  |  |  |  |  |
| Морана 1 | 0.15 | 0.137 | 0.159 | 0.085 |
| 0.1 | 0.089 | 0.106 | 0.05 |
| 0.05 | 0.044 | 0.052 | 0.021 |
| 0.025 | 0.021 | 0.025 | 0.008 |
| 0.01 | 0.007 | 0.009 | 0.002 |
| Морана 2 | 0.15 | 0.123 | 0.124 | 0.129 |
| 0.1 | 0.079 | 0.079 | 0.083 |
| 0.05 | 0.039 | 0.039 | 0.040 |
| 0.025 | 0.019 | 0.019 | 0.02 |
| 0.01 | 0.007 | 0.007 | 0.007 |
| Ченга-Спринга | 0.15 | 0.143 | 0.143 | 0.143 |
| 0.1 | 0.093 | 0.094 | 0.093 |
| 0.05 | 0.046 | 0.047 | 0.045 |
| 0.025 | 0.023 | 0.023 | 0.022 |
| 0.01 | 0.009 | 0.009 | 0.009 |
| Шерман 1 | 0.15 | 0.150 | 0.18 | 0.072 |
| 0.1 | 0.1 | 0.128 | 0.042 |
| 0.05 | 0.052 | 0.073 | 0.018 |
| 0.025 | 0.027 | 0.040 | 0.007 |
| 0.01 | 0.011 | 0.019 | 0.002 |
| Шерман 2 | 0.15 | 0.151 | 0.182 | 0.072 |
| 0.1 | 0.102 | 0.131 | 0.043 |
| 0.05 | 0.051 | 0.073 | 0.017 |
| 0.025 | 0.027 | 0.043 | 0.007 |
| 0.01 | 0.012 | 0.02 | 0.002 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Объем выборок | | | | |
|  |  |  |  |  |
| Моран 1 | 0.15 | 0.164 | 0.216 | 0.196 |
| 0.1 | 0.108 | 0.148 | 0.129 |
| 0.05 | 0.052 | 0.073 | 0.59 |
| 0.025 | 0.025 | 0.036 | 0.026 |
| 0.01 | 0.01 | 0.014 | 0.008 |
| Моран 2 | 0.15 | 0.118 | 0.114 | 0.208 |
| 0.1 | 0.078 | 0.075 | 0.142 |
| 0.05 | 0.038 | 0.037 | 0.076 |
| 0.025 | 0.019 | 0.018 | 0.039 |
| 0.01 | 0.007 | 0.007 | 0.017 |
| Ченга-Спринг | 0.15 | 0.184 | 0.184 | 0.184 |
| 0.1 | 0.128 | 0.127 | 0.127 |
| 0.05 | 0.067 | 0.068 | 0.068 |
| 0.025 | 0.034 | 0.034 | 0.034 |
| 0.01 | 0.014 | 0.013 | 0.014 |
| Шерман 1 | 0.15 | 0.207 | 0.273 | 0.1 |
| 0.1 | 0.149 | 0.208 | 0.061 |
| 0.05 | 0.087 | 0.132 | 0.027 |
| 0.025 | 0.051 | 0.086 | 0.012 |
| 0.01 | 0.024 | 0.048 | 0.004 |
| Шерман 2 | 0.15 | 0.206 | 0.271 | 0.097 |
| 0.1 | 0.151 | 0.21 | 0.061 |
| 0.05 | 0.0287 | 0.134 | 0.027 |
| 0.025 | 0.051 | 0.087 | 0.012 |
| 0.01 | 0.026 | 0.048 | 0.004 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Объем выборок | | | | |
|  |  |  |  |  |
| Моран 1 | 0.15 | 0.317 | 0.486 | 0.7 |
| 0.1 | 0.234 | 0.389 | 0.608 |
| 0.05 | 0.135 | 0.253 | 0.119 |
| 0.025 | 0.074 | 0.156 | 0.315 |
| 0.01 | 0.031 | 0.076 | 0.178 |
| Моран 2 | 0.15 | 0.1 | 0.094 | 0.421 |
| 0.1 | 0.064 | 0.061 | 0.323 |
| 0.05 | 0.03 | 0.029 | 0.198 |
| 0.025 | 0.014 | 0.014 | 0.118 |
| 0.01 | 0.005 | 0.005 | 0.055 |
| Ченга-Спринг | 0.15 | 0.411 | 0.412 | 0.406 |
| 0.1 | 0.333 | 0.335 | 0.33 |
| 0.05 | 0.223 | 0.224 | 0223 |
| 0.025 | 0.148 | 0.15 | 0.148 |
| 0.01 | 0.083 | 0.083 | 0.084 |
| Шерман 1 | 0.15 | 0.413 | 0.597 | 0.174 |
| 0.1 | 0.341 | 0.498 | 0.117 |
| 0.05 | 0.25 | 0.197 | 0.06 |
| 0.025 | 0.181 | 0.315 | 0.03 |
| 0.01 | 0.12 | 0.231 | 0.013 |
| Шерман 2 | 0.15 | 0.409 | 0.565 | 0.178 |
| 0.1 | 0.338 | 0.497 | 0.121 |
| 0.05 | 0.246 | 0.395 | 0.062 |
| 0.025 | 0.18 | 0.312 | 0.032 |
| 0.01 | 0.119 | 0.229 | 0.013 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Объем выборок | | | | |
|  |  |  |  |  |
| Моран 1 | 0.15 | 0.639 | 0.827 | 0.978 |
| 0.1 | 0.56 | 0.765 | 0.962 |
| 0.05 | 0.44 | 0.65 | 0.925 |
| 0.025 | 0.335 | 0.536 | 0.872 |
| 0.01 | 0.231 | 0.397 | 0.781 |
| Моран 2 | 0.15 | 0.07 | 0.064 | 0.674 |
| 0.1 | 0.044 | 0.041 | 0.577 |
| 0.05 | 0.02 | 0.019 | 0.421 |
| 0.025 | 0.01 | 0.009 | 0.298 |
| 0.01 | 0.004 | 0.003 | 0.177 |
| Ченга-Спринг | 0.15 | 0.76 | 0.76 | 0.76 |
| 0.1 | 0.694 | 0.695 | 0.697 |
| 0.05 | 0.582 | 0.585 | 0.585 |
| 0.025 | 0.481 | 0.484 | 0.482 |
| 0.01 | 0.35 | 0.355 | 0.35 |
| Шерман 1 | 0.15 | 0.697 | 0.861 | 0.304 |
| 0.1 | 0.636 | 0.823 | 0.226 |
| 0.05 | 0.543 | 0.758 | 0.135 |
| 0.025 | 0.463 | 0.693 | 0.082 |
| 0.01 | 0.373 | 0.611 | 0.04 |
| Шерман 2 | 0.15 | 0.695 | 0.860 | 0.302 |
| 0.1 | 0.634 | 0.823 | 0.223 |
| 0.05 | 0.541 | 0.757 | 0.133 |
| 0.025 | 0.46 | 0.69 | 0.078 |
| 0.01 | 0.371 | 0.604 | 0.038 |

По таблицам видно, что при маленьких объемах выборок все критерии являются смещенными. Отдельно стоит отметить критерий Морана 2, он является смещенным почти при всех объемах выборок. Наибольшую мощность тут показал критерий Морана 1.

Определим наиболее предпочтительный критерий по методам Вальда и Сэвиджа при .

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Объем выборок | | | | | | | Вальд |
|  |  |  |  |  |  |  | MIN |
| М1 | 0.129 | 0.163 | 0.149 | 0.089 | 0.106 | 0.05 | 0.05 |
| М2 | 0.122 | 0.178 | 0.153 | 0.079 | 0.079 | 0.083 | 0.079 |
| ЧС | 0.082 | 0.105 | 0.132 | 0.093 | 0.094 | 0.093 | 0.093 |
| Ш1 | 0.131 | 0.167 | 0.127 | 0.1 | 0.128 | 0.042 | 0.042 |
| Ш2 | 0.133 | 0.171 | 0.128 | 0.102 | 0.131 | 0.043 | 0.043 |
| Объем выборок | | | | | | | ЧС |
| М1 | 0.15 | 0.18 | 0.168 | 0.108 | 0.148 | 0.129 | 0.108 |
| М2 | 0.136 | 0.192 | 0.168 | 0.078 | 0.075 | 0.142 | 0.075 |
| ЧС | 0.0871 | 0.123 | 0.13 | 0.128 | 0.127 | 0.127 | 0.087 |
| Ш1 | 0.147 | 0.185 | 0.145 | 0.149 | 0.208 | 0.061 | 0.061 |
| Ш2 | 0.151 | 0.186 | 0.146 | 0.151 | 0.21 | 0.061 | 0.061 |
| Объем выборок | | | | | | | М1 |
| М1 | 0.2 | 0.219 | 0.208 | 0.234 | 0.389 | 0.608 | 0.2 |
| М2 | 0.167 | 0.228 | 0.2 | 0.064 | 0.061 | 0.323 | 0.064 |
| ЧС | 0.106 | 0.167 | 0.125 | 0.333 | 0.335 | 0.33 | 0.125 |
| Ш1 | 0.184 | 0.225 | 0.185 | 0.341 | 0.498 | 0.117 | 0.117 |
| Ш2 | 0.186 | 0.226 | 0.184 | 0.338 | 0.497 | 0.121 | 0.121 |
| Объем выборок | | | | | | | М1 |
| М1 | 0.274 | 0.273 | 0.262 | 0.56 | 0.765 | 0.962 | 0.262 |
| М2 | 0.202 | 0.27 | 0.236 | 0.044 | 0.041 | 0.577 | 0.041 |
| ЧС | 0.131 | 0.211 | 0.122 | 0.694 | 0.695 | 0.697 | 0.122 |
| Ш1 | 0.233 | 0.281 | 0.235 | 0.636 | 0.823 | 0.226 | 0.226 |
| Ш2 | 0.232 | 0.276 | 0.235 | 0.634 | 0.823 | 0.223 | 0.223 |
|  | | | | | | | M1 |

По критерию Вальда можно сказать, что наиболее предпочтительным при маленьких объемах выборок является критерий Ченга-Спринга, а при больших Морана 1.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Объем выборок | | | | | | | Сэвидж |
|  |  |  |  |  |  |  | MAX |
| М1 | 0.004 | 0.015 | 0.004 | 0.011 | 0.025 | 0.043 | 0.025 |
| М2 | 0.011 | 0 | 0 | 0.021 | 0.052 | 0.01 | 0.052 |
| ЧС | 0.051 | 0.073 | 0.021 | 0.007 | 0.037 | 0 | 0.073 |
| Ш1 | 0.002 | 0.011 | 0.026 | 0.002 | 0.005 | 0.051 | 0.051 |
| Ш2 | 0 | 0.007 | 0.025 | 0 | 0 | 0.050 | 0.05 |
| Объем выборок | | | | | | | М1 |
| М1 | 0.001 | 0.012 | 0 | 0.043 | 0.062 | 0.013 | 0.062 |
| М2 | 0.014 | 0 | 0 | 0.073 | 0.135 | 0 | 0.135 |
| ЧС | 0.064 | 0.069 | 0.038 | 0.023 | 0.083 | 0.015 | 0.083 |
| Ш1 | 0.004 | 0.007 | 0.023 | 0.002 | 0.003 | 0.081 | 0.081 |
| Ш2 | 0 | 0.006 | 0.022 | 0 | 0 | 0.081 | 0.081 |
| Объем выборок | | | | | | | М1 |
| М1 | 0 | 0.009 | 0 | 0.099 | 0.109 | 0 | 0.109 |
| М2 | 0.033 | 0 | 0.008 | 0.277 | 0.175 | 0.285 | 0.285 |
| ЧС | 0.094 | 0.061 | 0.083 | 0.008 | 0.163 | 0.278 | 0.278 |
| Ш1 | 0.016 | 0.003 | 0.023 | 0 | 0 | 0.491 | 0.491 |
| Ш2 | 0.014 | 0.002 | 0.024 | 0.003 | 0.001 | 0.487 | 0.487 |
| Объем выборок | | | | | | | М1 |
| М1 | 0 | 0.008 | 0 | 0.134 | 0.058 | 0 | 0.134 |
| М2 | 0.072 | 0.011 | 0.026 | 0.65 | 0.782 | 0.385 | 0.782 |
| ЧС | 0.143 | 0.07 | 0.14 | 0 | 0.128 | 0.265 | 0.265 |
| Ш1 | 0.041 | 0 | 0.027 | 0.058 | 0 | 0.736 | 0.736 |
| Ш2 | 0.042 | 0.005 | 0.027 | 0.06 | 0 | 0.739 | 0.739 |
|  | | | | | | | М1 |

По критерию Сэвиджа при всех объемах выборок наиболее предпочтительным является критерий Морана 1. Результаты, полученные по критерию Сэвиджа совпадают с результатами исследований предыдущих лет.

## ВЫВОДЫ

В данной работе была исследована мощность таких критериев как, критерии Морана 1, Морана 2, критерий Ченга-Спринга и Шермана. В качестве альтернативных гипотез были выбраны гипотезы, основанные на бета распределении и на обобщенно нормальном распределении, в силу того, что они могу быть очень близки к равномерному распределению при определенных параметрах.

В результате данных исследований было выявлено, что наиболее мощным является критерий Морана 1.

## СПИСОК ЛИТЕРАТУРЫ

1. П. Ю. Блинов, Б. Ю. Лемешко // Критерии проверки отклонения распределения от равномерного закона:. Новосибирск : Изд-во НГТУ, 2015. – УДК 519.23 – С. 182
2. Кобзарь А. И. Прикладная математическая статистика. Для инженеров и научных работников / А. И. Кобзарь. – M : ФИЗМАТЛИТ, 2006. – 816 с
3. Moran P. A. P. The random division of an intervals / P. A. P. Moran // J. R. Statist. Soc. – 1947. – Ser. B. V.9. No. 1. – P. 92-98.
4. Moran P. A. P. The random division of an intervals. II / P. A. P. Moran // J. R. Statist. Soc. – 1951. – Ser. B. V.13. No. 2. – P. 147-150.
5. Cheng S. W. A test to Identify the uniform distribution with applications / S. W. Cheng, F. A. Spiring // IEEE Trans. Reliability. – 1987. – V. R-36. – P. 98- 105
6. Sherman B. A random variable related to the spacing of sample values / B. Sherman // The Annals of Mathematical Statistics. – 1950. – V.21, №3. – P. 339-361.

## ПРИЛОЖЕНИЕ. ТЕКСТ ПРОГРАММЫ

Весь исходный код программы можно найти на https://github.com/andymost/CTDA\_3\_course\_work

monteCarlo.py

import numpy as np

import scipy.stats as st

import stats.chengSpring as chengSpring

import stats.moran as moran

import stats.sherman as sherman

from monteCarlo import monteCarlo

#количество наблюдений в выборке

n = 100

#количество выборок

N = 1000

def generateFunc(case = 0):

if case is 0:

return lambda:np.random.uniform(0, 1, n);

if case is 1:

return lambda:st.beta.rvs(1.22, 0.9, size=n)

if case is 2:

return lambda:st.beta.rvs(1.0, 0.75, size=n)

if case is 3:

return lambda:st.beta.rvs(0.8, 1.08, size=n)

if case is 4:

return lambda:st.gennorm.rvs(beta=8, loc=0.5, scale=0.5, size=n)

if case is 5:

return lambda:st.gennorm.rvs(beta=8, loc=0.45, scale=0.5, size=n)

return lambda:st.gennorm.rvs(beta = 8, loc = 0.5, scale = 0.4, size=n)

def statFunc(case = 'chengSpring'):

if case is 'chengSpring':

return chengSpring.stat

if case is 'moran' or case is 'moran1':

return moran.stat1

if case is 'moran2':

return moran.stat2

if case is 'sherman':

return sherman.stat

if case is 'shermanNormalized':

return sherman.statNormalized

return sherman.statModified

statNames = [

'chengSpring',

'moran',

'moran2',

'sherman',

'shermanNormalized',

'shermanModified',

]

hypotises = range(0, 7)

for statName in statNames:

for hypotise in hypotises:

print(statName, 'for hypotise', hypotise)

statF = statFunc(statName)

generateF = generateFunc(hypotise)

statSample = monteCarlo(N, n, statF, generateF, statName, hypotise is 0)

f = open(f'./output/H{hypotise}\_{statName}\_n={n}\_N={N}.dat', 'w')

f.write(f'H{hypotise}\_{statName}\_{n}\n')

f.write(f'0 {n}\n')

for item in statSample:

f.write(f'{item}\n')

f.close()

monteCarlo.py

def monteCarlo(N, n, statFunc, generateFunc, generateName, isH0):

statistica = []

for i in range(N):

x = list(generateFunc())

statistica.append(statFunc(x))

if (i%100 == 0):

print(f'Done {i/N \* 100}% MonteCarlo {generateName} {isH0}')

return statistica