|  |  |  |
| --- | --- | --- |
|  | | |
| Федеральное государственное бюджетное  образовательное учреждение высшего образования «Новосибирский государственный технический университет» | | |
|  | | |
| Кафедра прикладной математики | | |
| Практическое задание № 2 | | |
| по дисциплине «Методы принятия оптимальных решений» | | |
| **Оценивание параметров и проверка адекватности****построенных моделей** | | |
|  | | |
|  | Бригада | ПМ-13 Буданцев дмитрий |
| . | ПМ-13 Форкин Кирилл |
|  |  |
|  |  |
|  |  |
|  |  |
| Преподаватели | Лемешко Борис Юрьевич |
|  |  |
| Новосибирск,2024 | | |

1. **Цель**

**Цель работы – исследовать:**

* как с увеличением объемов выборок, сгенерированных в соответствии с законом с заданными параметрами, эмпирическое распределение приближается к соответствующему теоретическому;
* как это отражается на достигаемых уровнях значимости применяемых критериев согласия при проверке простых гипотез;
* как это отражается на достигаемых уровнях значимости применяемых критериев согласия при проверке сложных гипотез;
* как меняется вид закона (функции распределения и функции плотности) с изменением его параметров;
* возможность приближения эмпирических распределений другими моделями.

1. **Проведённые исследования**
2. **Проведите простую гипотезу о согласии с законом (и параметрами), указанным в первой строке файла с выборкой, зафиксируйте p-value по всем критериям, обратите внимание на плотность и гистограмму.**

Результаты проверки простых гипотез на основании предложенных файлов с выборками объемом 50, 100, 500 наблюдений, сгенерированных в соответсвии с распределением Вейбулла с параметром формы, меняющимся от 0.5 до 5 с одним и тем же параметром масштаба, опираясь на критерии согласия Колмагорова (см. таблица 1), Крамера-Мизиса-Смирнова (см. таблица 2), Андесона-Дарлинга (см. таблица 3) и Хи-квадрат Пирсона (см. таблица 4).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Критерий Колмогорова | | | | | | |
|  | 50 | | 100 | | 500 | |
|  |  |  |  |  |  |  |
| 0.5 | 1.095 | 0.1817 | 0.6754 | 0.7517 | 0.7194 | 0.6787 |
| 1 | 0.6709 | 0.7589 | 0.6322 | 0.8191 | 0.9219 | 0.3632 |
| 2 | 0.5727 | 0.8983 | 0.5627 | 0.9095 | 0.8959 | 0.3984 |
| 3 | 1.0093 | 0.2602 | 0.7543 | 0.62 | 0.8196 | 0.5127 |
| 4 | 1.1078 | 0.1717 | 0.8556 | 0.4569 | 0.703 | 0.7062 |
| 5 | 0.8366 | 0.4859 | 0.6664 | 0.7662 | 0.9189 | 0.3672 |

*Таблица 1.* Результаты проверки простых гипотез критерием Колмогорова

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Критерий Крамера-Мизиса-Смирнова | | | | | | |
|  | 50 | | 100 | | 500 | |
|  |  |  |  |  |  |  |
| 0.5 | 0.2579 | 0.1786 | 0.0632 | 0.794 | 0.0633 | 0.7935 |
| 1 | 0.0396 | 0.9354 | 0.0517 | 0.8658 | 0.1555 | 0.3731 |
| 2 | 0.0384 | 0.9414 | 0.0441 | 0.9108 | 0.114 | 0.5205 |
| 3 | 0.1749 | 0.3218 | 0.0815 | 0.683 | 0.172 | 0.329 |
| 4 | 0.2186 | 0.2341 | 0.1616 | 0.356 | 0.1606 | 0.3588 |
| 5 | 0.0904 | 0.6338 | 0.0626 | 0.7975 | 0.1919 | 0.2838 |

*Таблица 2.* Результаты проверки простых гипотез критерием Крамера-Мизиса-Смирнова

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Критерий Андерсона-Дарлинга | | | | | | |
|  | 50 | | 100 | | 500 | |
|  |  |  |  |  |  |  |
| 0.5 | 1.586 | 0.1572 | 0.4513 | 0.7968 | 0.6773 | 0.5781 |
| 1 | 0.2903 | 0.9453 | 0.4368 | 0.8116 | 0.7612 | 0.5099 |
| 2 | 0.2979 | 0.9397 | 0.2879 | 0.947 | 0.7743 | 0.5 |
| 3 | 1.1579 | 0.2841 | 0.4433 | 0.805 | 0.8881 | 0.4217 |
| 4 | 1.0514 | 0.3314 | 0.9323 | 0.3948 | 1.0415 | 0.3362 |
| 5 | 0.5489 | 0.6975 | 0.384 | 0.8645 | 1.235 | 0.2546 |

*Таблица 3.* Результаты проверки простых гипотез критерием Андерсона-Дарлинга

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Критерий Хи-квадрат | | | | | | |
|  | 50 | | 100 | | 500 | |
|  |  |  |  |  |  |  |
| 0.5 | 4.67 | 0.5868 | 4.313 | 0.6344 | 4.2625 | 0.6412 |
| 1 | 3.6806 | 0.7198 | 9.5515 | 0.1449 | 1.9902 | 0.9206 |
| 2 | 3.1698 | 0.7873 | 2.8026 | 0.8332 | 7.3829 | 0.2869 |
| 3 | 4.6025 | 0.5957 | 1.7962 | 0.9375 | 1.8621 | 0.9319 |
| 4 | 2.9047 | 0.8207 | 5.417 | 0.4915 | 3.0206 | 0.8063 |
| 5 | 3.9715 | 0.6805 | 3.8861 | 0.6921 | 4.0006 | 0.6766 |

*Таблица 4.* Результаты проверки простых гипотез критерием Хи-квадрат Пирсона

Из результатов (*см. таблица 1-4*) видно, что при проверке простых гипотез и с ростом объёма выборки применяемые критерии Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и Хи-квадрат Пирсона, не дают никаких оснований для отклонения справедливой гипотезы , что и следовало ожидать.

1. **Проверьте сложную гипотезу о согласии с этим же законом при оценивании параметров методом максимального правдоподобия, зафиксируйте p-value по всем критериям; посмотрите на плотность и гистограмму**

Результаты проверки сложных гипотез на основании предложенных файлов с выборками объемом 50, 100, 500 наблюдений, сгенерированных в соответсвии с распределением Вейбулла с параметром формы, меняющимся от 0.5 до 5 с одним и тем же параметром масштаба, опираясь на критерии согласия Колмагорова (см. таблица 5), Крамера-Мизиса-Смирнова (см. таблица 6), Андесона-Дарлинга (см. таблица 7) и Хи-квадрат Пирсона (см. таблица 8).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Критерий Колмогорова | | | | | | |
|  | 50 | | 100 | | 500 | |
|  |  |  |  |  |  |  |
| 0.5 | 0.6014 | 0.5329 | 0.477 | 0.871 | 0.4308 | 0.9484 |
| 1 | 0.568 | 0.6309 | 0.6313 | 0.4486 | 0.5932 | 0.5567 |
| 2 | 0.6437 | 0.4154 | 0.4119 | 0.9686 | 0.5341 | 0.7288 |
| 3 | 0.8038 | 0.1208 | 0.5581 | 0.66 | 0.3838 | 0.9877 |
| 4 | 0.618 | 0.4853 | 0.5084 | 0.7979 | 0.3715 | 0.9926 |
| 5 | 0.5904 | 0.5651 | 0.5792 | 0.5979 | 0.7062 | 0.2696 |

*Таблица 5.* Результаты проверки сложных гипотез критерием Колмогорова

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Критерий Крамера-Мизиса-Смирнова | | | | | | |
|  | 50 | | 100 | | 500 | |
|  |  |  |  |  |  |  |
| 0.5 | 0.0612 | 0.365 | 0.0248 | 0.9166 | 0.0265 | 0.8923 |
| 1 | 0.0293 | 0.8514 | 0.042 | 0.6363 | 0.0495 | 0.5177 |
| 2 | 0.0473 | 0.5509 | 0.0255 | 0.9073 | 0.0368 | 0.724 |
| 3 | 0.1047 | 0.0892 | 0.0615 | 0.3624 | 0.0188 | 0.9769 |
| 4 | 0.0636 | 0.3392 | 0.0319 | 0.8083 | 0.0239 | 0.9273 |
| 5 | 0.0435 | 0.6124 | 0.0524 | 0.4764 | 0.0934 | 0.1301 |

*Таблица 6.* Результаты проверки сложных гипотез критерием Крамера-Мизиса-Смирнова

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Критерий Андерсона-Дарлинга | | | | | | |
|  | 50 | | 100 | | 500 | |
|  |  |  |  |  |  |  |
| 0.5 | 0.3996 | 0.376 | 0.1623 | 0.959 | 0.2091 | 0.863 |
| 1 | 0.1744 | 0.9392 | 0.3192 | 0.5606 | 0.3095 | 0.5861 |
| 2 | 0.3017 | 0.6073 | 0.1885 | 0.9112 | 0.2345 | 0.7958 |
| 3 | 0.5785 | 0.1407 | 0.3499 | 0.4837 | 0.1562 | 0.9676 |
| 4 | 0.4129 | 0.3507 | 0.2782 | 0.6725 | 0.1686 | 0.9492 |
| 5 | 0.2479 | 0.7583 | 0.3457 | 0.4939 | 0.6312 | 0.104 |

*Таблица 7.* Результаты проверки сложных гипотез критерием Андерсона-Дарлинга

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Критерий Хи-квадрат Пирсона | | | | | | |
|  | 50 | | 100 | | 500 | |
|  |  |  |  |  |  |  |
| 0.5 | 3.5034 | 0.4774 | 5.1032 | 0.2769 | 12.8431 | 0.0121 |
| 1 | 3.0656 | 0.5469 | 6.8646 | 0.1432 | 3.9318 | 0.4153 |
| 2 | 3.102 | 0.5409 | 1.3656 | 0.8501 | 4.466 | 0.3466 |
| 3 | 3.4604 | 0.4839 | 0.9007 | 0.9245 | 0.7465 | 0.9455 |
| 4 | 1.1406 | 0.8878 | 3.8088 | 0.4325 | 1.6868 | 0.7931 |
| 5 | 0.7241 | 0.9483 | 3.5698 | 0.4673 | 2.4446 | 0.6546 |

*Таблица 8.* Результаты проверки сложных гипотез критерием Хи-квадрат Пирсона

Из результатов (*см. таблица 5-8*) видно, что при проверке сложных гипотез при оценивании двух параметров (формы и масштаба) и с ростом объёма выборки применяемые критерии Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и Хи-квадрат Пирсона, не дают никаких оснований для отклонения справедливой гипотезы , что и следовало ожидать.

1. **Проанализируйте результаты по значениям p-value и сформулируйте выводы по 1-му и 2-му пунктам.**

Исходя из полученных результатов видно, что для применяемых критериев согласия при проверке как сложных, так и простых гипотез не наблюдается отклонение справедливой гипотезы  при больших объёмах выборок.

Так как, в применяемых выборках нет ошибок округления. Критерий Колмогорова используются с поправкой Большева, что помогает пренебречь зависимость распределения статистики от объема выборки. Применение зарекомендовавшего метода оценивания параметров “Метод максимального правдоподобия” при данных объемах выборок сходится достаточно близко. Таким образом, полученный результат был предсказуемым и даёт основания сделать вывод, что критерии зависят от объёма выборки незначительно.

1. **Визуально посмотрите, как меняется распределение Вейбулла (интегральные функции распределения и функции плотности) с изменением параметров. Сформулируйте выводы по 2-му пункту.**

Результаты построения графиков интегральных функций распределения (см. график 1-2) и функции плотности (см. график 3-4)



*График 1.* Интегральных функций распределения при проверке сложных гипотез



*График 2.* Интегральных функций распределения при проверке простых гипотез



*График 3.* Функции плотности при проверке сложных гипотез



*График 4.* Функции плотности при проверке простых гипотез

Попарно сравнивая графики (см. графики 1-4), видно подтверждение, того что было написано ранее. Метод оценки параметров “Метод максимального правдоподобия” при данных объемах выборок сходится достаточно близко к истинным параметрам. Сравнив графики тяжело найти разницу между ними.

1. **Выборки объёмом 500 наблюдений попробуйте сгладить другой параметрической моделью, например, обобщенным распределением Вейбулла и гамма-распределением В последнем случае проверьте эффект от применения “группирования при оценивании”.**

Результаты проверки сложных гипотез используя параметрические модели: обобщенное распределение Вейбулла (см. таблица 9) и гамма-распределение (см. таблица 10), на основании предложенных файлов с выборками объемом 500 наблюдений, сгенерированных в соответсвии с распределением Вейбулла с параметром формы, меняющимся от 0.5 до 5 с одним и тем же параметром масштаба, опираясь на критерии согласия Колмагорова, Крамера-Мизиса-Смирнова, Андесона-Дарлинга и Хи-квадрат Пирсона.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Критерий  Колмогорова | | Критерий  КМС | | Критерий  АД | | Критерий  Хи-квадрат | |
|  | 500 | | | | | | | |
|  |  |  |  |  |  |  |  |  |
| 0.5 | 0.9014 | 0 | 0.2662 | 0 | 1.91 | 0 | 5.49E+06 | 0 |
| 1 | 1.1566 | 0 | 0.3464 | 0 | 2.0988 | 0 | 0.4637 | 0.9268 |
| 2 | 0.6127 | 0 | 0.1078 | 0 | 0.7076 | 0 | 3.2394 | 0.3562 |
| 3 | 0.3733 | 0 | 0.0186 | 0 | 0.1558 | 0 | 1.0849 | 0.7807 |
| 4 | 0.3767 | 0 | 0.0196 | 0 | 0.1409 | 0 | 13.0433 | 0.0045 |
| 5 | 0.5113 | 0 | 0.0417 | 0 | 0.3932 | 0 | 386.978 | 0 |

*Таблица 9.* Результаты проверки сложных гипотез, используя обобщенное распределение Вейбулла

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Критерий  Колмогорова | | Критерий  КМС | | Критерий  АД | | Критерий  Хи-квадрат | |
|  | 500 | | | | | | | |
|  |  |  |  |  |  |  |  |  |
| 0.5 | 3.2397 | 0 | 3.4511 | 0 | 19.1759 | 0 | 475.9408 | 0 |
| 1 | 0.7004 | 0.3377 | 0.0777 | 0.2593 | 0.5391 | 0.1906 | 466.8004 | 0 |
| 2 | 1.5212 | 0 | 0.6181 | 0 | 3.5326 | 0 | 691.2111 | 0 |
| 3 | 2.0082 | 0 | 1.3095 | 0 | 7.6246 | 0 | 929.3105 | 0 |
| 4 | 2.0205 | 0 | 1.2961 | 0 | 8.0472 | 0 | 964.1838 | 0 |
| 5 | 1.5658 | 0 | 0.5641 | 0 | 5.5002 | 0 | 612.5873 | 0 |

*Таблица 10.* Результаты проверки сложных гипотез, используя гамма-распределение

1. **Выборки объёмом 500 наблюдений при параметре формы 4 и 5 попробуйте описать:**
   * + **нормальным законом**
     + **обобщенным нормальным законом**
     + **Сформулируйте выводы по 5-му и 6-му пункту.**

Результаты проверки сложных гипотез используя параметрические модели: **нормальным законом** (см. таблица 11) и **распределением Лапласа** (см. таблица 12), на основании предложенных файлов с выборками объемом 500 наблюдений, сгенерированных в соответсвии с распределением Вейбулла с параметром формы, меняющимся от 4 до 5 с одним и тем же параметром масштаба, опираясь на критерии согласия Колмагорова, Крамера-Мизиса-Смирнова, Андесона-Дарлинга и Хи-квадрат Пирсона.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Критерий  Колмогорова | | Критерий  КМС | | Критерий  АД | | Критерий  Хи-квадрат | |
|  | 500 | | | | | | | |
|  |  |  |  |  |  |  |  |  |
| 4 | 0.4642 | 0.9015 | 0.0338 | 0.7819 | 0.2278 | 0.8141 | 2.8036 | 0.5912 |
| 5 | 0.5888 | 0.5839 | 0.0395 | 0.6827 | 0.2614 | 0.7181 | 5.0708 | 0.2801 |

*Таблица 11.* Результаты проверки сложных гипотез, используя нормальный закон

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Критерий  Колмогорова | | Критерий  КМС | | Критерий  АД | | Критерий  Хи-квадрат | |
|  | 500 | | | | | | | |
|  |  |  |  |  |  |  |  |  |
| 4 | 1.294 | 0.0013 | 0.5223 | 0 | 3.3512 | 0 | 45.36 | 0 |
| 5 | 1.756 | 0 | 0.5217 | 0 | 3.1548 | 0 | 38.0608 | 0 |

*Таблица 12.* Результаты проверки сложных гипотез, используя распределением Лапласа

Из результатов (*см. таблица 11*) видно, что при проверке сложных гипотез применяемые критерии Колмогорова, Крамера-Мизеса-Смирнова, Андерсона-Дарлинга и Хи-квадрат Пирсона, не дают никаких оснований для отклонения не верной гипотезы. Дело в том, что при данных параметрах a(1) ложная гипотеза  близка c  (см. графики 5-6). Для проверки таких близких конкурирующих гипотез необходимо иметь больший объём выборки.

В то же время, обратив внимание на Таблицу 12, можем заметить, что критерии ожидаемо, отвергают ложную гипотезу . Поскольку гипотеза , в данном случае, в отличие от гипотезы , менее близока с гипотезой . И данного объема выборки достаточно для того, чтобы проверить это.



*График 5.* Интегральные функции распределения конкурирующих гипотез



*График 5.* Функции плотности при проверке конкурирующих гипотез