|  |  |  |
| --- | --- | --- |
|  | | |
| Федеральное государственное бюджетное  образовательное учреждение высшего образования «Новосибирский государственный технический университет» | | |
|  | | |
| Кафедра теоретической и прикладной информатики | | |
| Лабораторная работа № 1 | | |
| по дисциплине «Статистический анализ нечисловых данных» | | |
|  | | |
|  | | |
|  | Группа ПМ-13 |  |
| Бригада 5 | зиянуров артём |
| Вариант 5 | вострецова екатерина |
|  |  |
|  |  |
| Преподаватель | Тимофеева анастасия юрьевна |
|  |  |
|  |  |
| Новосибирск, 2024 | | |

# Данные

Attribute Information:

A1: b, a.

A2: continuous.

A3: continuous.

A4: u, y, l, t.

A5: g, p, gg.

A6: c, d, cc, i, j, k, m, r, q, w, x, e, aa, ff.

A7: v, h, bb, j, n, z, dd, ff, o.

A8: continuous.

A9: t, f.

A10: t, f.

A11: continuous.

A12: t, f.

A13: g, p, s.

A14: continuous.

A15: continuous.

A16: +,- (class attribute)

# Ход работы

1. **Из набора данных Вашего варианта сформируйте два массива данных:**

**- массив количественных данных (все количественные данные оставить без изменений, для всех качественных данных кроме переменной класса1 произвести калибровку с учетом априорного шанса с поправкой Лапласа, значения переменной класса задать как 1, если положительный класс, 0 иначе);**

**- массив качественных данных (для всех качественных данных кроме переменной класса произвести калибровку с учетом априорного шанса с поправкой Лапласа, значения переменной класса задать как 1, если положительный класс, 0 иначе; для всех количественных данных произвести дискретизацию с равной частотой, в качестве границ интервалов взять выборочные квантили порядка 0, 0.2, 0.4, 0.6, 0.8, 1 2 ).**

В данном варианте в качественными данными являются:

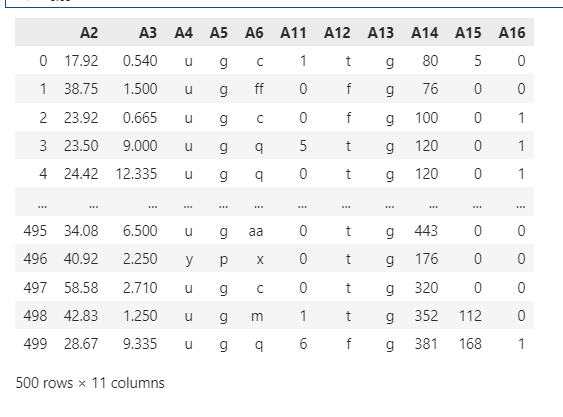
A4, A5, A6, A12, A13

Количественными:

A2, A3, A11, A14, A15

A16 является переменной класса

**Данные ДО калибровки**

****

**Значения после калибровки**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Столбец А4** | **Всего** | **1** | **0** | **с** | **Ответ** |
| u | **385** | **192** | **193** | **0,8115942** | **0,5507216410** |
| y | **114** | **31** | **83** | **0,3194444444** |
| l | **1** | **1** | **0** | **0,7113402062** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Столбец А5** | **Всего** | **1** | **0** | **с** | **Ответ** |
| **g** | **385** | **192** | **193** | **0,8115942** | **0,5507216410** |
| **p** | **114** | **31** | **83** | **0,3194444444** |
| **gg** | **1** | **1** | **0** | **0,7113402062** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Столбец А6** | **Всего** | **1** | **0** | **с** | **Ответ** |
| **C** | **105** | **46** | **59** | **0,8115942** | **0,491140391** |
| **D** | **19** | **5** | **14** | **0,330143541** |
| **cc** | **26** | **18** | **8** | **0,72231405** |
| **I** | **43** | **8** | **35** | **0,235494881** |
| **J** | **10** | **3** | **7** | **0,38121547** |
| **K** | **36** | **11** | **25** | **0,362521891** |
| **M** | **29** | **11** | **18** | **0,437632135** |
| **R** | **3** | **2** | **1** | **0,648902821** |
| **Q** | **59** | **41** | **18** | **0,731448763** |
| **W** | **51** | **28** | **23** | **0,598206278** |
| **X** | **30** | **25** | **5** | **0,842253521** |
| **E** | **14** | **9** | **5** | **0,67251462** |
| **Aa** | **37** | **14** | **23** | **0,435056747** |
| **ff** | **38** | **3** | **35** | **0,120418848** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Столбец А12** | **Всего** | **1** | **0** | **с** | **Ответ** |
| **T** | **231** | **108** | **123** | **0,8115942** | **0,519944694** |
| **F** | **269** | **116** | **153** | **0,48350003** |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Столбец А13** | **Всего** | **1** | **0** | **с** | **Ответ** |
| **G** | **459** | **213** | **246** | **0,8115942** | **0,516329813** |
| **P** | **1** | **0** | **1** | **0,38121547** |
| **s** | **40** | **11** | **29** | **0,330143541** |

**Данные после калибровки с учетом априорного шанса с поправкой Лапласа:**

****

**2**

**К количественным данным относятся столбцы**

A2, A3, A11, A14, A15

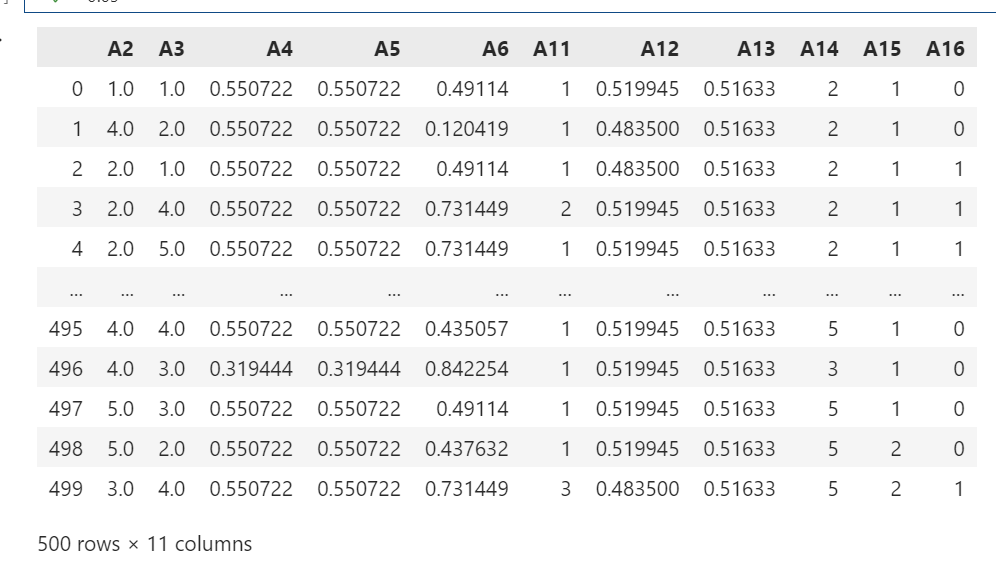
Данные ДО дискретизации, НО уже с калибровкой с учетом априорного шанса с поправкой Лапласа:

****

Получились следующие границы интервалов:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **A2** | **A3** | **A11** | **A14** | **A15** |
| [15.17; 21.814]  (21.814; 25.468]  (25.468;32.282]  (32.282; 40.92]  (40.92; 74.83] | [0.0; 0.822]  (0.822; 2.0]  (2.0;4.415]  (4.415; 9.5]  (9.5; 28.0] | [0.0; 1]  (1; 5.0]  (5.0;67.0] | [0.0; 21.6]  (21.6; 120.0]  (120.0;181.0]  (181.0; 290.0]  (290.0; 980.0] | [0.0; 50]  (50; 587.8]  (587.8;51100.0] |

Данные после дискретизации выглядят так:

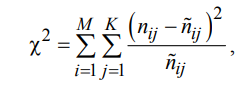
****

1. **Рассчитайте все парные показатели взаимосвязи между переменными из набора данных, соответствующего Вашему варианту.**

**Нашему варианту соответствует m=0**

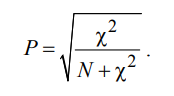
**m=0 – статистика хи-квадрат, коэффициент сопряженности Пирсона, коэффициент Крамера**

Парные показатели взаимосвязи между переменными в массиве с качественными данными, **статистика хи-квадрат**

****

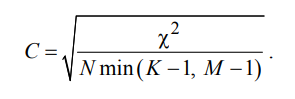
|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A2** | **A3** | **A4** | **A5** | **A6** | **A11** | **A12** | **A13** | **A14** | **A15** | **A16** |
| **A2** |  | 72,77135 | 7,501664 | 7,501664 | 63,84842 | 22,48606 | 12,10203 | 4,923498 | 81,6324 | 10,48556 | 19,0395 |
| **A3** | 72,77135 |  | 12,69411 | 12,69411 | 79,65124 | 29,51267 | 5,203919 | 13,21982 | 67,38893 | 12,85672 | 35,44714 |
| **A4** | 7,501664 | 12,69411 |  | 1000 | 41,12267 | 11,17258 | 1,463126 | 12,39473 | 24,59755 | 6,675919 | 19,52508 |
| **A5** | 7,501664 | 12,69411 | 1000 |  | 41,12267 | 11,17258 | 1,463126 | 12,39473 | 24,59755 | 6,675919 | 19,52508 |
| **A6** | 63,84842 | 79,65124 | 41,12267 | 41,12267 |  | 49,90691 | 25,13915 | 50,71118 | 98,36848 | 23,07514 | 84,31273 |
| **A11** | 22,48606 | 29,51267 | 11,17258 | 11,17258 | 49,90691 |  | 2,614751 | 17,44896 | 17,73217 | 82,69857 | 120,9278 |
| **A12** | 12,10203 | 5,203919 | 1,463126 | 1,463126 | 25,13915 | 2,614751 |  | 1,10093 | 21,25118 | 2,045121 | 0,523734 |
| **A13** | 4,923498 | 13,21982 | 12,39473 | 12,39473 | 50,71118 | 17,44896 | 1,10093 |  | 8,956503 | 29,28191 | 6,13086 |
| **A14** | 81,6324 | 67,38893 | 24,59755 | 24,59755 | 98,36848 | 17,73217 | 21,25118 | 8,956503 |  | 10,32786 | 32,91096 |
| **A15** | 10,48556 | 12,85672 | 6,675919 | 6,675919 | 23,07514 | 82,69857 | 2,045121 | 29,28191 | 10,32786 |  | 79,48565 |
| **A16** | 19,0395 | 35,44714 | 19,52508 | 19,52508 | 84,31273 | 120,9278 | 0,523734 | 6,13086 | 32,91096 | 79,48565 |  |

**Коэффициент сопряженности Пирсона для массива количественных данных**

****

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A2** | **A3** | **A4** | **A5** | **A6** | **A11** | **A12** | **A13** | **A14** | **A15** | **A16** |
| **A2** |  | 0,356443 | 0,121579 | 0,121579 | 0,336507 | 0,207453 | 0,153727 | 0,098747 | 0,374634 | 0,143319 | 0,191526 |
| **A3** | 0,356443 |  | 0,157352 | 0,157352 | 0,370692 | 0,236084 | 0,101492 | 0,160495 | 0,344631 | 0,158331 | 0,257296 |
| **A4** | 0,121579 | 0,157352 |  | 0,816497 | 0,275672 | 0,14784 | 0,054016 | 0,155531 | 0,216537 | 0,114786 | 0,193862 |
| **A5** | 0,121579 | 0,157352 | 0,816497 |  | 0,275672 | 0,14784 | 0,054016 | 0,155531 | 0,216537 | 0,114786 | 0,193862 |
| **A6** | 0,336507 | 0,370692 | 0,275672 | 0,275672 |  | 0,301256 | 0,218795 | 0,303452 | 0,405456 | 0,210034 | 0,37986 |
| **A11** | 0,207453 | 0,236084 | 0,14784 | 0,14784 | 0,301256 |  | 0,072127 | 0,183633 | 0,185067 | 0,376727 | 0,441309 |
| **A12** | 0,153727 | 0,101492 | 0,054016 | 0,054016 | 0,218795 | 0,072127 |  | 0,046872 | 0,201915 | 0,063825 | 0,032348 |
| **A13** | 0,098747 | 0,160495 | 0,155531 | 0,155531 | 0,303452 | 0,183633 | 0,046872 |  | 0,132657 | 0,23521 | 0,11006 |
| **A14** | 0,374634 | 0,344631 | 0,216537 | 0,216537 | 0,405456 | 0,185067 | 0,201915 | 0,132657 |  | 0,142259 | 0,248509 |
| **A15** | 0,143319 | 0,158331 | 0,114786 | 0,114786 | 0,210034 | 0,376727 | 0,063825 | 0,23521 | 0,142259 |  | 0,370359 |
| **A16** | 0,191526 | 0,257296 | 0,193862 | 0,193862 | 0,37986 | 0,441309 | 0,032348 | 0,11006 | 0,248509 | 0,370359 |  |

**Коэффициент Крамера**

****

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A2** | **A3** | **A4** | **A5** | **A6** | **A11** | **A12** | **A13** | **A14** | **A15** | **A16** |
| **A2** |  | 0,19075 | 0,086612 | 0,086612 | 0,178673 | 0,149954 | 0,155577 | 0,070168 | 0,20203 | 0,102399 | 0,195138 |
| **A3** | 0,19075 |  | 0,112668 | 0,112668 | 0,199564 | 0,171793 | 0,102019 | 0,114977 | 0,183561 | 0,113387 | 0,26626 |
| **A4** | 0,086612 | 0,112668 |  | 1 | 0,202787 | 0,1057 | 0,054095 | 0,111332 | 0,156836 | 0,081706 | 0,197611 |
| **A5** | 0,086612 | 0,112668 | 1 |  | 0,202787 | 0,1057 | 0,054095 | 0,111332 | 0,156836 | 0,081706 | 0,197611 |
| **A6** | 0,178673 | 0,199564 | 0,202787 | 0,202787 |  | 0,223399 | 0,224228 | 0,225191 | 0,221775 | 0,151905 | 0,41064 |
| **A11** | 0,149954 | 0,171793 | 0,1057 | 0,1057 | 0,223399 |  | 0,072315 | 0,132095 | 0,133162 | 0,287574 | 0,491788 |
| **A12** | 0,155577 | 0,102019 | 0,054095 | 0,054095 | 0,224228 | 0,072315 |  | 0,046924 | 0,206161 | 0,063955 | 0,032365 |
| **A13** | 0,070168 | 0,114977 | 0,111332 | 0,111332 | 0,225191 | 0,132095 | 0,046924 |  | 0,094639 | 0,17112 | 0,110733 |
| **A14** | 0,20203 | 0,183561 | 0,156836 | 0,156836 | 0,221775 | 0,133162 | 0,206161 | 0,094639 |  | 0,101626 | 0,256558 |
| **A15** | 0,102399 | 0,113387 | 0,081706 | 0,081706 | 0,151905 | 0,287574 | 0,063955 | 0,17112 | 0,101626 |  | 0,398712 |
| **A16** | 0,195138 | 0,26626 | 0,197611 | 0,197611 | 0,41064 | 0,491788 | 0,032365 | 0,110733 | 0,256558 | 0,398712 |  |

1. **Проверьте гипотезы о значимости взаимосвязей между переменными:**

**m=0 – по критерию хи-квадрат и с помощью перестановочного критерия;**

**Для массива качественных данных:**

1. **Значения статистики хи-квадрат**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A2** | **A3** | **A4** | **A5** | **A6** | **A11** | **A12** | **A13** | **A14** | **A15** | **A16** |
| **A2** |  | 72,77135 | 7,501664 | 7,501664 | 63,84842 | 22,48606 | 12,10203 | 4,923498 | 81,6324 | 10,48556 | 19,0395 |
| **A3** | 72,77135 |  | 12,69411 | 12,69411 | 79,65124 | 29,51267 | 5,203919 | 13,21982 | 67,38893 | 12,85672 | 35,44714 |
| **A4** | 7,501664 | 12,69411 |  | 1000 | 41,12267 | 11,17258 | 1,463126 | 12,39473 | 24,59755 | 6,675919 | 19,52508 |
| **A5** | 7,501664 | 12,69411 | 1000 |  | 41,12267 | 11,17258 | 1,463126 | 12,39473 | 24,59755 | 6,675919 | 19,52508 |
| **A6** | 63,84842 | 79,65124 | 41,12267 | 41,12267 |  | 49,90691 | 25,13915 | 50,71118 | 98,36848 | 23,07514 | 84,31273 |
| **A11** | 22,48606 | 29,51267 | 11,17258 | 11,17258 | 49,90691 |  | 2,614751 | 17,44896 | 17,73217 | 82,69857 | 120,9278 |
| **A12** | 12,10203 | 5,203919 | 1,463126 | 1,463126 | 25,13915 | 2,614751 |  | 1,10093 | 21,25118 | 2,045121 | 0,523734 |
| **A13** | 4,923498 | 13,21982 | 12,39473 | 12,39473 | 50,71118 | 17,44896 | 1,10093 |  | 8,956503 | 29,28191 | 6,13086 |
| **A14** | 81,6324 | 67,38893 | 24,59755 | 24,59755 | 98,36848 | 17,73217 | 21,25118 | 8,956503 |  | 10,32786 | 32,91096 |
| **A15** | 10,48556 | 12,85672 | 6,675919 | 6,675919 | 23,07514 | 82,69857 | 2,045121 | 29,28191 | 10,32786 |  | 79,48565 |
| **A16** | 19,0395 | 35,44714 | 19,52508 | 19,52508 | 84,31273 | 120,9278 | 0,523734 | 6,13086 | 32,91096 | 79,48565 |  |

1. **P-value**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A2** | **A3** | **A4** | **A5** | **A6** | **A11** | **A12** | **A13** | **A14** | **A15** | **A16** |
| **A2** |  | 3,25E-09 | 0,483595 | 0,483595 | 0,125483 | 0,004091 | 0,016608 | 0,765721 | 8,44E-11 | 0,232585 | 0,000772 |
| **A3** | 3,25E-09 |  | 0,122817 | 0,122817 | 0,008122 | 0,000258 | 0,267007 | 0,104507 | 2,85E-08 | 0,116875 | 3,76E-07 |
| **A4** | 0,483595 | 0,122817 |  | 3,6E-215 | 0,030163 | 0,024691 | 0,481156 | 0,014645 | 0,001818 | 0,154038 | 5,76E-05 |
| **A5** | 0,483595 | 0,122817 | 3,6E-215 |  | 0,030163 | 0,024691 | 0,481156 | 0,014645 | 0,001818 | 0,154038 | 5,76E-05 |
| **A6** | 0,125483 | 0,008122 | 0,030163 | 0,030163 |  | 0,003226 | 0,022131 | 0,002583 | 0,000109 | 0,628693 | 1,69E-12 |
| **A11** | 0,004091 | 0,000258 | 0,024691 | 0,024691 | 0,003226 |  | 0,270529 | 0,001581 | 0,023327 | 4,67E-17 | 5,51E-27 |
| **A12** | 0,016608 | 0,267007 | 0,481156 | 0,481156 | 0,022131 | 0,270529 |  | 0,576682 | 0,000282 | 0,359673 | 0,469253 |
| **A13** | 0,765721 | 0,104507 | 0,014645 | 0,014645 | 0,002583 | 0,001581 | 0,576682 |  | 0,345979 | 6,85E-06 | 0,046634 |
| **A14** | 8,44E-11 | 2,85E-08 | 0,001818 | 0,001818 | 0,000109 | 0,023327 | 0,000282 | 0,345979 |  | 0,242764 | 1,25E-06 |
| **A15** | 0,232585 | 0,116875 | 0,154038 | 0,154038 | 0,628693 | 4,67E-17 | 0,359673 | 6,85E-06 | 0,242764 |  | 5,49E-18 |
| **A16** | 0,000772 | 3,76E-07 | 5,76E-05 | 5,76E-05 | 1,69E-12 | 5,51E-27 | 0,469253 | 0,046634 | 1,25E-06 | 5,49E-18 |  |

1. **Если, значение p-value меньше, чем 0.05, то гипотеза о независимости отвергается**

**1 – правда(отвергается)**

**0 – ложь(НЕ отвергается)**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A2** | **A3** | **A4** | **A5** | **A6** | **A11** | **A12** | **A13** | **A14** | **A15** | **A16** |
| **A2** |  | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| **A3** | 1 |  | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| **A4** | 0 | 0 |  | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| **A5** | 0 | 0 | 1 |  | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| **A6** | 0 | 1 | 1 | 1 |  | 1 | 1 | 1 | 1 | 0 | 1 |
| **A11** | 1 | 1 | 1 | 1 | 1 |  | 0 | 1 | 1 | 1 | 1 |
| **A12** | 1 | 0 | 0 | 0 | 1 | 0 |  | 0 | 1 | 0 | 0 |
| **A13** | 0 | 0 | 1 | 1 | 1 | 1 | 0 |  | 0 | 1 | 1 |
| **A14** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |  | 0 | 1 |
| **A15** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |  | 1 |
| **A16** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |  |

**Перестановочный критерий**

1. **Значения перестановочного критерия**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A2** | **A3** | **A4** | **A5** | **A6** | **A11** | **A12** | **A13** | **A14** | **A15** | **A16** |
| **A2** |  | 25,56298 | 13,13001 | 13,00106 | 68,1557 | 16,02579 | 9,336138 | 14,14336 | 27,64771 | 13,68261 | 9,666088 |
| **A3** | 25,56298 |  | 13,11263 | 12,9169 | 70,54453 | 15,34992 | 10,0453 | 13,51594 | 27,22226 | 15,35706 | 9,195691 |
| **A4** | 13,13001 | 13,11263 |  | 7,690864 | 45,48136 | 9,863391 | 4,506817 | 11,99784 | 13,62581 | 8,298668 | 4,913912 |
| **A5** | 13,00106 | 12,9169 | 7,690864 |  | 46,72065 | 9,574142 | 6,16662 | 12,39473 | 13,63315 | 9,048063 | 4,913912 |
| **A6** | 68,1557 | 70,54453 | 45,48136 | 46,72065 |  | 39,38986 | 22,85649 | 45,2841 | 71,20903 | 38,90172 | 21,29312 |
| **A11** | 16,02579 | 15,34992 | 9,863391 | 9,574142 | 39,38986 |  | 5,413225 | 9,276592 | 14,5601 | 10,18706 | 6,284621 |
| **A12** | 9,336138 | 10,0453 | 4,506817 | 6,16662 | 22,85649 | 5,413225 |  | 5,456151 | 9,350645 | 5,887463 | 3,261597 |
| **A13** | 14,14336 | 13,51594 | 11,99784 | 12,39473 | 45,2841 | 9,276592 | 5,456151 |  | 13,71441 | 9,115573 | 4,828099 |
| **A14** | 27,64771 | 27,22226 | 13,62581 | 13,63315 | 71,20903 | 14,5601 | 9,350645 | 13,71441 |  | 15,63672 | 9,52722 |
| **A15** | 13,68261 | 15,35706 | 8,298668 | 9,048063 | 38,90172 | 10,18706 | 5,887463 | 9,115573 | 15,63672 |  | 5,38884 |
| **A16** | 9,666088 | 9,195691 | 4,913912 | 4,913912 | 21,29312 | 6,284621 | 3,261597 | 4,828099 | 9,52722 | 5,38884 |  |

1. **Если значение в столбце перестановочного критерия меньше, чем значение хи-квадрат, то гипотеза о независимости отвергается**

**1-отвергается**

**0-НЕ отвергается**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **A2** | **A3** | **A4** | **A5** | **A6** | **A11** | **A12** | **A13** | **A14** | **A15** | **A16** |
| **A2** |  | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| **A3** | 1 |  | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 1 |
| **A4** | 0 | 0 |  | 1 | 0 | 1 | 0 | 1 | 1 | 0 | 1 |
| **A5** | 0 | 0 | 1 |  | 0 | 1 | 0 | 0 | 1 | 0 | 1 |
| **A6** | 0 | 1 | 0 | 0 |  | 1 | 1 | 1 | 1 | 0 | 1 |
| **A11** | 1 | 1 | 1 | 1 | 1 |  | 0 | 1 | 1 | 1 | 1 |
| **A12** | 1 | 0 | 0 | 0 | 1 | 0 |  | 0 | 1 | 0 | 0 |
| **A13** | 0 | 0 | 1 | 0 | 1 | 1 | 0 |  | 0 | 1 | 1 |
| **A14** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 |  | 0 | 1 |
| **A15** | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |  | 1 |
| **A16** | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 |  |

**3.Из подмножества признаков, значимо влияющих на переменную класса, выберите пару признаков, наиболее сильно взаимосвязанных друг с другом. Обозначим их X1 и X2. Проверьте наличие частной корреляции между признаком X1 (X2) и откликом при исключении влияния X2 (X1) с помощью**

Взаимная информация и частная взаимная информация для качественных данных (Х1 – это признак А4, Х2 – это признак А5, Y – это А16)

MI(X1,Y) = 0.02059015902799573

MI(X1, Y| X2) = 0

MI(X2,Y) = 0.02059015902799573

MI(X2, Y| X1) = 0

Так как взаимная информация не нулевая, а частная нулевая, то делаем вывод о том, что при исключении одного из признаков информация, которая передаётся в отклик равна нулю. Принимая ко вниманию также таблицу сопряжённости мы замечаем, что данные одинаковые. И признаки A4 и A5 несут одинаковую информацию. Значит один из них можно убрать.

**4.Дайте интерпретацию полученным результатам исходя из практических соображений. Уделите особое внимание интерпретации взаимосвязей между объясняющими переменными и переменной класса.**

Ранжирование данных по коэффициенту Крамера от наиболее значимого к наименее:

A11, A6, A15, A3, A14, A4, A5, A2, A13, A12

Учитывая, что А4 и А5 имеют одинаковую информацию, то в итоге мы получаем следующие значимые признаки для отклика:

A11, A6, A15, A3, A14, A4 , A2, A13, A12.