09.11 Скласти та розв‘язати дві задачі (перша – по курсу «Теорія ймовірностей», друга – по курсу «Математична статистика») за тематикою спеціальності з використанням теоретичного матеріалу (формул), який не зустрічався в завданнях варіанту.

*Задача №1:* Щоб захистити користувача від спаму в електронній пошті (масової розсилки кореспонденції рекламного чи іншого характеру людям, які не висловили бажання її одержувати) використовується спеціальна програма (спам-фільтр). При «навчанні» фільтра для кожного слова у листах розраховується і зберігається його «вага» - оцінка ймовірності того, що листи з цим словом – спам.

Користувач отримав повідомлення на пошту, що містить слово «безкоштовно». Яка ймовірність того, що це повідомлення – спам? Врахувати, що:

* Ймовірність того, що довільне повідомлення – спамове – 0.75;
* Ймовірність того, що слово «безкоштовно» з’являється у повідомленнях, якщо вони є спамом – 0.2; якщо вони безпечні – 0.5;

*Розв’язок:*

Нехай подія А – при перевірці повідомлення в ньому виявлено слово «безкоштовно».  *–* гіпотеза, повідомлення є спамом; – гіпотеза*,* повідомлення безпечне. Отже, ці дві події утворюють повну групу (вони є несумісними та єдиноможливими).

Тоді з умови :

Подія А відбувається за умови виконання однієї з гіпотез:

Можемо застосувати формулу Байєса, оскільки необхідно переоцінити ймовірність гіпотези, за умови, що відбулась подія А:

, де – ймовірності гіпотез; умовні ймовірності події А;

Оскільки нас цікавить випадок, коли повідомлення зі словом «безкоштовно» - спам, то ми маємо знайти *.* Маємо*:*

Відповідь*: 0.545;*

*Задача №2*: У статистичному дослідженні швидкості завантаження файлів розважального характеру з серверу протягом двох діб (1-го та 2-го січня в однаковому тарифі) з інтервалом в одну годину були отримані наступні величини (Мбіт/сек ):

*3.48 3.40 3.67 3.88 4.13 4.28 4.59 4.99 4.87 4.60 4.34 4.33*

*4.12 4.01 3.83 3.56 3.40 3.41 3.45 3.54 3.78 3.45 3.67 3.89*

*4.03 4.14 4.35 4.35 4.35 4.43 4.32 4.56 4.76 4.99 5.03 5.34*

*5.78 5.99 6.30 6.19 5.67 5.78 6.01 6.01 6.23 6.58 7.01 7.15*

1. Побудуйте таблицю частоти, відносної частоти, щільності частоти, щільності відносної частоти, накопиченої частоти, накопиченої відносної частоти.
2. Знайдіть медіану та моду.
3. Побудуйте гістограму щільності відносної частоти.
4. Побудуйте полігон щільності частоти.

*Розв’язок:*

Варіційний ряд (впорядкована вибірка):

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 3.4 | 3.4 | 3.41 | 3.45 | 3.45 | 3.45 | 3.48 | 3.54 | 3.56 | 3.67 | 3.67 | 3.78 |
| 3.83 | 3.88 | 3.89 | 4.01 | 4.03 | 4.12 | 4.13 | 4.14 | 4.28 | 4.32 | 4.33 | 4.34 |
| 4.35 | 4.35 | 4.35 | 4.43 | 4.56 | 4.59 | 4.6 | 4.76 | 4.87 | 4.99 | 4.99 | 5.03 |
| 5.34 | 5.67 | 5.78 | 5.99 | 6.01 | 6.01 | 6.19 | 6.23 | 6.3 | 6.58 | 7.01 | 7.15 |

Найбільше значення , а найменше – . Розмах . Тоді виділяємо класів із шириною .

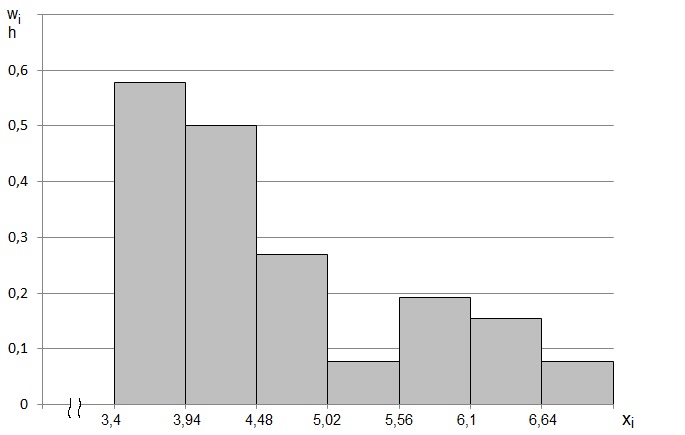
Згрупована об’єднана таблиця частоти , відносної частоти , щільності частоти , щільності відносної частоти , накопиченої частоти , накопиченої відносної частоти :

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |
| 3.4 - 3.94 | 15 | 0.3125 | 27.7778 | 0.5787 | 15 | 0.3125 |
| 3.94 - 4.48 | 13 | 0.2708 | 24.074 | 0.5015 | 28 | 0.5833 |
| * 1. 5.02 | 7 | 0.1458 | 12.963 | 0.27 | 35 | 0.7292 |
| 5.02 - 5.56 | 2 | 0.0417 | 3.7037 | 0.0772 | 37 | 0.7708 |
| 5.56 - 6.1 | 5 | 0.1042 | 9.2593 | 0.193 | 42 | 0.875 |
| 6.1 - 6.64 | 4 | 0.0833 | 7.4074 | 0.1543 | 46 | 0.9583 |
| 6.64 - 7.18 | 2 | 0.0417 | 3.7037 | 0.0772 | 48 | 1 |
|  | 48 | 1 |  |  |  |  |

*Медіана* - середнє арифметичне двох центральних елементів варіаційного ряду (об’єм всієї вибірки *N* = 48 – парне число):

*Мода* – варіанта,що зустрічається найчастіше у варіаційному ряді. В даному випадку ми маємо полі модальну вибірку:

За даними таблиці вище побудуємо шукану гістограму щільності відносної частоти:



Полігон щільності частоти:

