**Урок 3. EDA (exploratory data analysis) или Разведочный анализ**

**Задача 1**

Даны значения зарплат из выборки выпускников: 100, 80, 75, 77, 89, 33, 45, 25, 65, 17, 30, 24, 57, 55, 70, 75, 65, 84, 90, 150. Посчитать (желательно без использования статистических методов наподобие std, var, mean) среднее арифметическое, среднее квадратичное отклонение, смещенную и несмещенную оценки дисперсий для данной выборки.

17, 24, 25, 30, 33, 45, 55, 57, 65, 65, 70, 75, 75, 77, 80, 84, 89, 90, 100, 150

среднее арифметическое = (17+24+25+30+33+45+55+57+65+65+70+75+75+77+80+84+89+90+100+150)/20 = 65.3

дисперсия = = = 950,11

среднее квадратичное отклонение = = = 30,82

дисперсия несмещенная= = 1000.1

**Задача 2**

В первом ящике находится 8 мячей, из которых 5 - белые. Во втором ящике - 12 мячей, из которых 5 белых. Из первого ящика вытаскивают случайным образом два мяча, из второго - 4. Какова вероятность того, что 3 мяча белые?

**Задача 3**

На соревновании по биатлону один из трех спортсменов стреляет и попадает в мишень. Вероятность попадания для первого спортсмена равна 0.9, для второго — 0.8, для третьего — 0.6. Найти вероятность того, что выстрел произведен: a). первым спортсменом б). вторым спортсменом в). третьим спортсменом.

Ф-ла Байеса

Событие А – попадание

Событие B – выстрел каждого спортсмена

P попадание = P(cпортсмен1|попал)\*P(cпортсмен1) + P(cпортсмен2|попал)\*P(cпортсмен2) + P(cпортсмен3|попал)\*P(cпортсмен3) = 0,9\*1/3 + 0,8\* 1/3 + 0,6\* 1/3 = 0,766

a). первым спортсменом

P(cпортсмен1|попал)=(0,9\*1/3) / 0,766 = 0,3911

б). вторым спортсменом

P(cпортсмен2|попал)=(0,8\*1/3) / 0,766 = 0,3476

в). третьим спортсменом.

P(cпортсмен3|попал)=(0,6\*1/3) / 0,766 = 0,2607

**Задача 4**

В университет на факультеты A и B поступило равное количество студентов, а на факультет C студентов поступило столько же, сколько на A и B вместе. Вероятность того, что студент факультета A сдаст первую сессию, равна 0.8. Для студента факультета B эта вероятность равна 0.7, а для студента факультета C - 0.9. Студент сдал первую сессию. Какова вероятность, что он учится: a). на факультете A б). на факультете B в). на факультете C?

P сдача сессии = P(факА|сдачаА)\*P(cтудент с факА) + P(факB|сдачаB)\*P(cтудент с факB) + P(факC|сдачаC)\*P(cтудент с факC) = 0,8\*0.25 + 0,7\* 0.25 + 0,9\* 0.5 = 0,825

Ф-ла Байеса

Событие А – сдача сессии

Событие B – факультет студента

a). на факультете A

P(cтудент с факА|сдачаА) = (P(сдача|студА) \* P(студА)) / P(сдача) = 0,25\*0,8/0,825 = 0,24

б). на факультете B

P(cтудент с факB|сдачаB) = (P(сдача|студB) \* P(студB)) / P(сдача) = 0,25\*0,7/0,825 = 0,21

в). на факультете C?

P(cтудент с факC|сдачаC) = (P(сдача|студC) \* P(студC)) / P(сдача) = 0,5\*0,9/0,825 = 0,54

**Задача 5**

Устройство состоит из трех деталей. Для первой детали вероятность выйти из строя в первый месяц равна 0.1, для второй - 0.2, для третьей - 0.25. Какова вероятность того, что в первый месяц выйдут из строя: а). все детали б). только две детали в). хотя бы одна деталь г). от одной до двух деталей?

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Деталь 1 | Х | 0 | 0 | Х | Х | Х | 0 | 0 |
| Деталь 2 | Х | 0 | Х | 0 | Х | 0 | Х | 0 |
| Деталь 3 | Х | Х | 0 | 0 | 0 | Х | Х | 0 |

а). все детали

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Деталь 1 | Х | 0 | 0 | Х | Х | Х | 0 | 0 |
| Деталь 2 | Х | 0 | Х | 0 | Х | 0 | Х | 0 |
| Деталь 3 | Х | Х | 0 | 0 | 0 | Х | Х | 0 |

Pбрак1\* Pбрак2\* Pбрак3 = 0,1\*0,2\*0,25 = 0,005

б). только две детали

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Деталь 1 | Х | 0 | 0 | Х | Х | Х | 0 | 0 |
| Деталь 2 | Х | 0 | Х | 0 | Х | 0 | Х | 0 |
| Деталь 3 | Х | Х | 0 | 0 | 0 | Х | Х | 0 |

Pбрак1\* Pбрак2\*(1- Pбрак3) + Pбрак1\*(1- Pбрак2)\* Pбрак3+ (1- Pбрак1)\* Pбрак2\* Pбрак3= 0.1\*0.2\*(1-0.25) + 0.1\*(1-0.2)\*0.25+ (1-0.1)\*0.2\*0.25 = 0.08

в). хотя бы одна деталь

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Деталь 1 | Х | 0 | 0 | Х | Х | Х | 0 | 0 |
| Деталь 2 | Х | 0 | Х | 0 | Х | 0 | Х | 0 |
| Деталь 3 | Х | Х | 0 | 0 | 0 | Х | Х | 0 |

1 – P(ни одна деталь не сломалась) = 1 – ((1- Pбрак1) \* (1- Pбрак2) \* (1- Pбрак3)) = 1-0,54 = 0,46

г). от одной до двух деталей?

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Деталь 1 | Х | 0 | 0 | Х | Х | Х | 0 | 0 |
| Деталь 2 | Х | 0 | Х | 0 | Х | 0 | Х | 0 |
| Деталь 3 | Х | Х | 0 | 0 | 0 | Х | Х | 0 |

1 – P(ни одна деталь не сломалась) - P(все детали сломаны) = 1 – ((1- Pбрак1) \* (1- Pбрак2) \* (1- Pбрак3)) – (Pбрак1\* Pбрак2\* Pбрак3) = 1-0,54 -0,005 = 0,535