

# **Конспект лекций по теории вероятностей и математической статистике.**

## **4 семестр**

**Лектор – Лохвицкий Михаил Сергеевич**

## **Часть 2**

### **1.6. Формулы полной вероятности и Байеса**

**Пример 1.12.** «Типичная» задача для формулы полной вероятности : На складе изделия с трёх заводов: с первого завода 50% всех изделий, со второго 40% и с третьего -10%. Вероятность того, что деталь высшего качества на первом заводе, равна 0,9, для второго - 0,7, для третьего -0,5. Найти вероятность того, что наудачу взятая со склада деталь высшего качества.

Обозначим события  $H_i$  – деталь сделана на  $i$ -ом заводе,  $A$  –наудачу взятая деталь высшего качества. Тогда  $P(H_1) = 50\%/100\% = 0,5$ ;

$$P(H_2) = 40\%/100\% = 0,4; P(H_3) = 10\%/100\% = 0,1; P(A/H_1) = 0,9;$$

$$P(A/H_2) = 0,7; P(A/H_3) = 0,5.$$

(Решение этой задачи будет дано после вывода формулы.)

**Формула полной вероятности.** Пусть события  $H_1, H_2, \dots, H_n$  образуют полную группу несовместных событий (будем их называть гипотезами). Событие  $A$  может произойти только с одним из  $H_i$ ,

$$P(A) = \sum_{i=1}^n P(H_i)P(A/H_i). \quad (1.15)$$

*Доказательство:*

Поскольку  $H_1 + H_2 + \dots + H_n = \Omega$ , то умножив обе части последнего равенства на событие  $A$ , получим  $A(H_1 + H_2 + \dots + H_n) = A\Omega = A$ . Из того, что  $H_i H_j = \emptyset$  ( $i \neq j$ ), следует, что  $(AH_i) \cdot (AH_j) = \emptyset$  и  $AH_1 + AH_2 + \dots + AH_n = A$ . По свойству 4 после определения 1.18

$$P(A) = \sum_{i=1}^n P(H_i A), \quad \text{а из свойства 1 после определения 1.19}$$

следует, что  $P(H_i A) = P(H_i)P(A/H_i)$ , так что  $P(A) = \sum_{i=1}^n P(H_i)P(A/H_i)$ .

**Решение** примера 1.12.

$$P(A) = P(H_1)P(A/H_1) + P(H_2)P(A/H_2) + P(H_3)P(A/H_3) = 0,5 \cdot 0,9 + 0,4 \cdot 0,7 + 0,1 \cdot 0,5 = 0,78$$

**Формула Байеса.** В условиях формулы полной вероятности требуется пересчитать вероятности гипотез при условии, что событие произошло:

$$P(H_k / A) = \frac{P(H_k)P(A/H_k)}{P(A)} \quad (1.16)$$

Здесь  $P(A)$  вычисляется по формуле полной вероятности.

Вероятности  $P(H_i)$  называются априорными (до опыта), а вероятности  $P(H_k / A)$  апостериорными (после опыта).

Вывод формулы Байеса: Запишем вероятность произведения двух событий  $A$  и  $H_k$  двумя способами:

$$P(A \cdot H_k) = P(A) \cdot P(H_k / A) = P(H_k) \cdot P(A / H_k).$$

Из последнего равенства найдем выражение для вероятности  $P(H_k / A)$  и в итоге и получим формулу Байеса (1.9).

**Пример 1. 13.** «Типичная» задача для формулы Байеса:

В условиях «Типичной» задачи для формулы полной вероятности (см. выше пример 1.12). Требуется определить вероятности того, что деталь сделана на 1, 2, 3 заводах при условии, того, что она высшего качества.

Решение.  $P(H_1 / A) = \frac{P(H_1)P(A/H_1)}{P(A)} = \frac{0,5 \cdot 0,9}{0,78} = 0,577;$

$$P(H_2 / A) = \frac{P(H_2)P(A/H_2)}{P(A)} = \frac{0,4 \cdot 0,7}{0,78} = 0,359;$$

$$P(H_3 / A) = \frac{P(H_3)P(A/H_3)}{P(A)} = \frac{0,1 \cdot 0,5}{0,78} = 0,064;$$

**Задача 1.8.** Экзамен в группе принимают профессор и доцент. Доцент принимает 70% всех студентов, остальных - профессор. Вероятность сдать экзамен с высокой оценкой профессору равна 0,9, а доценту 0,6. Какова вероятность получить высокую оценку на экзамене.

**Задача 1.9..** В условиях предыдущей задачи определить вероятность того, что студент сдавал экзамен профессору (доценту) при условии, что оценка высокая.

### **1.7.Схема и формула Бернулли.**

Проводится серия из  $n$  независимых испытаний, в каждом из которых может произойти событие  $A$  с одной и той же вероятностью  $p$ . Вероятность того, что событие произойдёт ровно  $k$  раз в  $n$  испытаниях равна:

$$P_n(k) = C_n^k p^k q^{n-k} \quad (1.17)$$

Здесь  $q = 1 - p$ .

Вывод формулы Бернулли. Перебираются все возможные варианты того, что событие произойдёт ровно  $k$  раз в  $n$  испытаниях (т.е. искомое событие раскладывается на сумму  $C_n^k$  событий). Эти события несовместны, и имеют одну и ту же вероятность  $p^k q^{n-k}$ .

**Пример 1.14.** Стрелок делает 4 независимых выстрелов по мишени. Вероятность попадания при каждом выстреле равна 0,8. Найти вероятность того, что будет: а) 3 попадания, б) ни одного промаха, в) не менее 3-х попаданий, г) не более 2 -х попаданий.

**Решение:** 1.14.а)  $P_4(3) = C_4^3 0,8^3 0,2^1 = 4 \cdot 0,384 \cdot 0,2 = 0,3072$ ; 1.14.б) т.е. 4 попадания  $P_4(4) = C_4^4 0,8^4 0,2^0 = 1 \cdot 0,4096 \cdot 1 = 0,4096$ ;

1.14.в)  $P_4(k \geq 3) = P_4(k=3 \text{ или } k=4) = P_4(k=3) + P_4(k=4) = 0,3072 + 0,4096 = 0,7168$ ; 1.14.г)  $P_4(k \leq 2) = P_4(k=0 \text{ или } k=1 \text{ или } k=2) = 1 - P_4(k \geq 3) = 1 - 0,7168 = 0,2832$ .

**Задача 1.10.** Студент сдаёт в сессию 3 экзамена. Вероятность сдать каждый экзамен с оценкой отлично равна 0,9. Найти вероятность того, что он сдаст с оценкой отлично: а) все три экзамена, б) два из трёх, в) ни один экзамен.

#### **Обобщение формулы Бернулли.**

Пусть в последовательности  $n$  независимых однородных испытаний в каждом из исходов могут наблюдаться события  $A_1, A_2, \dots, A_r$ ; с вероятностями  $p_1, p_2, \dots, p_r$  соответственно и  $\sum_{i=1}^r p_i = 1$ . Требуется определить вероятность того, что событие  $A_1$  появится  $k_1$  раз, событие  $A_2$  появится  $k_2$  раза, ..., событие  $A_r$  появится  $k_r$  раз,

Для дальнейшего рассмотрения перепишем формулу (1.17) в другой форме: введем обозначения  $A_1 = A$ ;  $A_2 = \bar{A}$ ;  $k_1 = k$ ;  $k_2 = n - k$ ;  $p_1 = p$ ;  $p_2 = q$ . При этом  $A_1 + A_2 = \Omega$ ;  $k_1 + k_2 = n$ ;  $p_1 + p_2 = 1$ . В новых обозначениях формула Бернулли будет иметь такой вид:

$$p_n(k_1; k_2) = \frac{n!}{k_1! k_2!} p_1^{k_1} p_2^{k_2}$$

Теперь эту формулу обобщим на  $r$  событий  $A_1, A_2, \dots, A_r$ ;

$$p_n(k_1; k_2; \dots; k_r) = \frac{n!}{k_1! k_2! \dots k_r!} p_1^{k_1} p_2^{k_2} \cdot \dots \cdot p_r^{k_r} \quad (1.18)$$

где  $k_1 + k_2 + \dots + k_r = n, k_r \geq 0$ .

**Пример 1.15.** По статистике в Принстоне (штат Нью Джерси, США) 60% обращающихся по телефону спасения 911 нуждаются в помощи полиции, 25% - медиков и 15% - пожарных. Наудачу набираются 10 звонков. Найдём вероятность того, что в 6 случаях обращение будет переадресовано в полицию, в 3 - в скорую помощь и в одном - в пожарную часть.

Вводя события  $A_1 = \{\text{нужна полиция}\}$ ,  $A_2 = \{\text{нужен врач}\}$ ,  $A_3 = \{\text{нужны пожарные}\}$ , получаем:

$$p_1 = 0,6; p_2 = 0,25; p_3 = 0,15.$$

$$P(k_1 = 6, k_2 = 3, k_3 = 1) = \frac{10!}{6! 3! 1!} (0,6)^6 \cdot (0,25)^3 \cdot (0,15)^1 \approx 0,092.$$

**Замечание.** При вычислении сочетаний нужно вычислять факториалы. Например,  $20! = 2432902 \cdot 10^{12}$ . Т.е. при больших  $n$  это очень большие числа, а вероятности  $p$  и  $q$  в больших степенях числа очень маленькие. Поэтому использование формулы (1.10) при больших  $n, k$  и малых  $p$  и (или)  $q$  вызывает определенные вычислительные проблемы. Существуют предельные теоремы для формулы Бернулли, которые позволяют вычислить соответствующие вероятности приближенно.

## 1.8. Предельные теоремы в схеме Бернулли

**Локальная теорема Муавра-Лапласа.** Если  $p \neq 0$  и  $p \neq 1$ , то при  $n \rightarrow \infty$

$$P_n(k) \approx \frac{1}{\sqrt{npq}} \varphi(x), \quad (1.19)$$

$$\text{где } x = \frac{k-np}{\sqrt{npq}}, \varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad (1.20)$$

**Замечание.** Функция  $\varphi(x)$  табулирована и таблица с её значениями есть в любом учебнике по теории вероятностей. Далее мы увидим, что это плотность распределения вероятностей нормального закона с параметрами  $N(0, 1)$ .

**Интегральная теорема Муавра-Лапласа.** Если  $p \neq 0$  и  $p \neq 1$ , то при  $n \rightarrow \infty$

$$P_n(k_1 \leq k \leq k_2) \approx \Phi(x_2) - \Phi(x_1), \quad (1.21)$$

$$\text{Где } x_1 = \frac{k_1-np}{\sqrt{npq}}, x_2 = \frac{k_2-np}{\sqrt{npq}},$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-x^2/2} dx \quad (1.22)$$

**Важное замечание 1.** Функция  $\Phi(x)$  называется функцией Лапласа. Она табулирована и таблица с её значениями есть в любом учебнике по теории вероятностей. Однако, в разных учебниках выражение, а следовательно и таблица, для этой функции может отличаться от формулы (1.22). В этом случае и формула для вероятности (1.21) может **быть другой**.

Функция Лапласа – нечетная функция. При  $x \geq 5$  функция  $\Phi(x) \approx 0,5$ .

Доказательство теорем Муавра-Лапласа следует из Центральной предельной теоремы, которая будет в дальнейшем курсе.

**Важное замечание 2.** При малых значениях вероятности  $p$ , когда  $np \leq 9$ , лучшие приближения, чем формула локальной теоремы Муавра-Лапласа, дает следующая теорема.

**Теорема Пуассона.** Если  $p \rightarrow 0, n \rightarrow \infty$ , но так, что  $np = \lambda$ , то

$$P_n(k) \rightarrow P(k) = P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad (1.23)$$

**Пример 1.16.** По каналу связи передается сообщение из  $n=10^5$  двоичных символов. Вероятность искажения каждого символа равна  $p=5 \cdot 10^{-6}$ . Вероятность того, что в сообщении будет искажено а) три символа; б) 0 символов; в) не более трёх символов.

Решение. Так как  $\lambda = np = 0,5 \leq 9$ , используем теорему Пуассона: а)  $P(3) = e^{-\lambda} \cdot \lambda^3 / 3! = 0,013$ ; б)  $P(0) = e^{-\lambda} = 0,624$ ; в)  $P(k \leq 3) = P(0) + P(1) + P(2) + P(3) = e^{-\lambda} (1 + \lambda + \lambda^2 / 2! + \lambda^3 / 3!) = 0,998$ .

**Пример 1.17.** Вероятность приёма каждого из 100 передаваемых сигналов равна 0,8. Найти вероятность того, что будет принято а) 75 сигналов; б) не менее 75, но не более 90 сигналов; в) не менее 75; г) не более 74 сигналов.

Решение. Этот пример решается по теоремам Муавра-Лапласа, т.к.  $np = 100 \cdot 0,8 = 80 \geq 10$ . а) в соответствии с локальной теоремой (1.18)  $P_{100}(75) \approx \frac{1}{\sqrt{100 \cdot 0,8 \cdot 0,2}} \varphi\left(\frac{75 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}}\right) = \frac{1}{4} \varphi(-1,25) = 0,25 \cdot 0,1827 = 0,046$ . б) по интегральной теореме  $P_{100}(75 \leq k \leq 90) \approx \Phi\left(\frac{90 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}}\right) - \Phi\left(\frac{75 - 100 \cdot 0,8}{\sqrt{100 \cdot 0,8 \cdot 0,2}}\right) = \Phi(2,5) + \Phi(1,25) = 0,4938 + 0,3944 = 0,8882$ ; (было

использовано свойство нечетности функции Лапласа); в)  $P_{100}(75 \leq k \leq 100) \approx$

$$\Phi\left(\frac{100-100 \cdot 0.8}{\sqrt{100 \cdot 0.8 \cdot 0.2}}\right) - \Phi\left(\frac{75-100 \cdot 0.8}{\sqrt{100 \cdot 0.8 \cdot 0.2}}\right) = \Phi(5) + \Phi(1,25) = 0,5 + 0,3944 = 0,8944;$$

$$з) P_{100}(0 \leq k \leq 74) \approx 1 - P_{100}(75 \leq k \leq 100) = 1 - 0,8944 = 0,1056.$$