Теория вероятностей и математическая статистика

Громова Е.В., Тур А.В.

Санкт-Петербург, 2020

Благодарим за помощь в компьютерной верстке конспекта Голокоза А., Портнова В.
Иллюстрации: студенты 1 курса псих.факультета СПбГУ Потёмкина Д., Митрофанова Н.

ОБНОВЛЁННАЯ ВЕРСИЯ С НАЧАЛЬНЫМИ ГЛАВАМИ

Оглавление

Часть I

Теория вероятностей

Введение

Теория вероятностей — наука, изучающая закономерности случайных явлений.

Глава 1

Случайные события. Дискретные вероятностные пространства

1.1 Основные понятия теории вероятностей. Пространство элементарных исходов. Случайное событие

Испытание(опыт, эксперимент) — фиксация какого-либо явления или факта при соблюдении определенной (воспроизводимой) совокупности условий.

Событие — факт, появление которого регистрируется в результате испытания.

Событие может не произойти в результате опыта!

Опыт не обязательно ставится человеком.

Идеальный опыт – опыт, который можно воспроизвести бесконечное число раз.

Примеры

1. Опыт – подбрасывание монеты. Событие $A =$ "Г".
2. Опыт – подбрасывание 3 монет. Событие $B =$ "ГГГ".
3. Опыт – подбрасываем кость. Событие $C =$ "3".
4. Опыт – подбрасываем кость. Событие $D =$ "четное число".
5. Опыт – выстрел. Событие $E =$ "попадание".
6. Опыт – рождение ребёнка. Событие $F =$ "рождение мальчика".

События бывают двух типов:

1. *Составные.*
2. *Элементарные.* Обозначаются ω_i.

   $\Omega = \{\omega_i\}_{i=1}^n$ – *множество (пространство) элементарных событий (элементарных исходов).*

Случайное событие A – любое подмножество пространства элементарных исходов $A \subseteq \Omega$

Примеры пространств элементарных исходов $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$:

1. $\Omega = \{$Г, Р$\}$
2. $\Omega = \{$ГГГ, РРР, ГГР, ГРГ, РГГ, ГРР, РГР, РРГ$\}$
3. $\Omega = \{1, 2, 3, 4, 5, 6\}$
4. $\Omega = \{1, 2, 3, 4, 5, 6\}$
5. $\Omega = \{$попал, не попал$\}$
6. $\Omega = \{$мальчик, девочка$\}$

1.2 Операции над событиями. Сумма и произведение событий. Полная группа событий

Если в результате испытания событие обязательно произойдёт, то оно называется *достоверным* событием ($A = \Omega$).

Если в результате испытания событие заведомо не произойдёт, то оно называется *невозможным* ($A = \varnothing$).

- $A \subseteq \Omega$ — случайное событие.
- $A = \varnothing$ — невозможное событие.
- $A = \Omega$ — достоверное событие.
- $\bar{A}$ — противоположное событие. Событие $\bar{A}$ называется противоположным к A

событием, если оно происходит тогда и только тогда, когда не происходит событие A.

- Если $A \cdot B = \varnothing$, то A, B — несовместные события (появление одного события исключает появление другого).

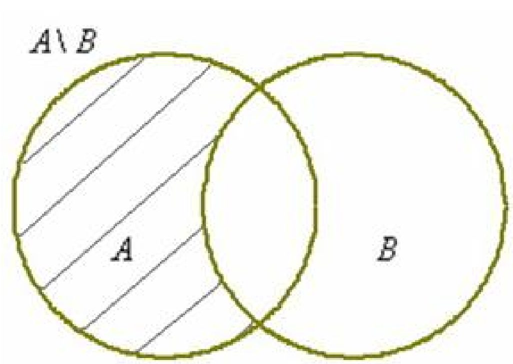
Произведение событий $A \cap B$ $(A \cdot B)$ – событие C, состоящее в появлении и A, и B.
Сумма событий $A \cup B$ $(A + B)$ – событие C, состоящее в появлении или A или B, или и A, и B.

Полная группа событий – группа событий $A_1, \ldots, A_n$, таких что:

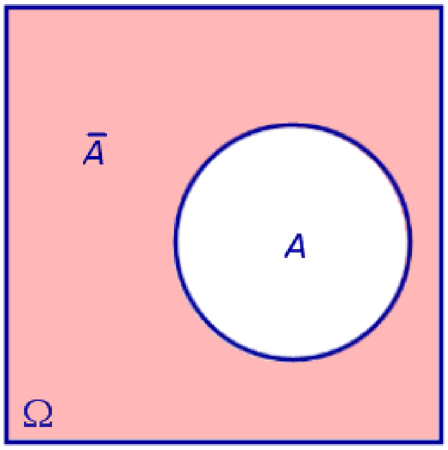
1. $A_i \cdot A_j = \varnothing, \ \forall i \neq j$
2. $A_1 + \ldots + A_n = \Omega$

Мощность конечного множества — количество его элементов.

Разность множеств A и B — множество, состоящее из элементов, принадлежащих множеству A, но не принадлежащих множеству B. Обозначается как $A \backslash B$.



Заметим, что противоположное к A событие $\bar{A} = \Omega \backslash A$. Тогда



$$|\bar{A}| = |\Omega| - |A|. \tag{1.1}$$

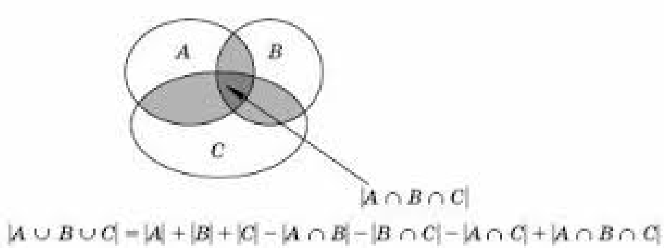
Аналогично получаем, что

$$|\overline{A \cup B}| = |\Omega| - |A \cup B|;$$
$$|\overline{A \cup B \cup C}| = |\Omega| - |A \cup B \cup C|. \tag{1.2}$$

Формула включений – исключений:

$$|A \cup B| = |A| + |B| - |A \cap B|. \tag{1.3}$$

$$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |B \cap C| - |A \cap C| + |A \cap B \cap C|. \tag{1.4}$$



Задача 1. В компании, занимающейся переводами, каждый из сотрудников владеет как минимум одним их трех иностранных языков: английским, немецким, французским. 20

сотрудников знают английский язык, 15 – немецкий, 10 –французский. 5 человек владеют и английским, и немецким. 4 человека – немецким и французским. 3 – французским и английским, двое знают все три языка. Сколько сотрудников в компании?

Задача 2. В трёх группах учатся 70 первокурсников. Из них 27 увлекаются общей психологией, 32 – социальной психологией, 22 – инженерной психологией. И общей, и социальной психологией интересуются 10 человек, социальной и инженерной – 6, общей и инженерной – 8, а всеми тремя дисциплинами – 3 человека. Сколько студентов не увлекается ни одной из дисциплин?

Общая формула:

$$|A_1 \cup \ldots \cup A_n| = \sum_{k=1}^{n} \sum_{1 \le i_1 \le \ldots \le i_k \le n} (-1)^{k-1} |A_{i1} \cap \ldots \cap A_{ik}|.$$

1.3 Вероятности событий. Классическая формула вычисления вероятностей. Примеры

Вероятность события A — число, характеризующее степень возможности этого события.

Обозначается как $P(A)$ (“p” — probability).

1. $P(\varnothing) = 0$,
2. $P(\Omega) = 1$,
3. $0 \le P(A) \le 1$.

Пусть выполнены следующие предпосылки:

1. $|\Omega| = n$ — пространство элементарных исходов конечно.
2. Все элементарные исходы равновозможны (*симметрия шансов*).

Тогда

$$P(A) = \frac{m}{n}, \tag{1.5}$$

где m — число благоприятствующих событию A элементарных исходов,

n — общее число элементарных исходов.

Используем обозначение *мощности* множества:

$$|A| = m; \quad |\Omega| = n.$$

$$P(A) = \frac{|A|}{|\Omega|}, \tag{1.6}$$

Пример. Опыт — вытаскиваем из колоды в 52 карты одну карту. Событие A — вытащили красную карту.

$$P(A) = \frac{26}{52} = \frac{1}{2}.$$

B — туз.

$$P(B) = \frac{4}{52} = \frac{1}{13}.$$

1.4 Статистическая вероятность

Используется, если для Ω выполнено условие конечности пространства, а второе условие (равновозможность элементарных событий) нарушено:

1. $|\Omega| = n$ — пространство элементарных исходов конечно.

~~2. Все элементарные исходы равновозможны (*симметрия шансов*).~~

$$P(A) = \frac{m_A}{n_A},$$

где m_A — число появлений события A,

n_A — общее число испытаний.

КАКОВА ВЕРОЯТНОСТЬ ВСТРЕТИТЬ ДИНОЗАВРА НА НЕВСКОМ?

1.5 Геометрическая вероятность. Задача о встрече

Используется, если для Ω нарушено условие конечности пространства, а второе условие (равновозможность элементарных событий) выполнено:

~~1. $|\Omega| = n$ — пространство элементарных исходов конечно.~~

2. Все элементарные исходы равновозможны (*симметрия шансов*).

Пример 1. Имеется отрезок длины L. На него наудачу равновероятно для любой точки отрезка ставится точка. Найти вероятность того, что точка попадёт на отрезок длины l. Условие 1 нарушено, условие 2 выполено, вероятность

$$P(A) = \frac{l}{L}.$$

Для задачи на плоскости — отношение площадей, для задачи в объеме — отношение объемов.

Геометрическая вероятность в общем случае (задается и для пространств другой размерности):

$$P(A) = \frac{m(A)}{m(\Omega)},$$

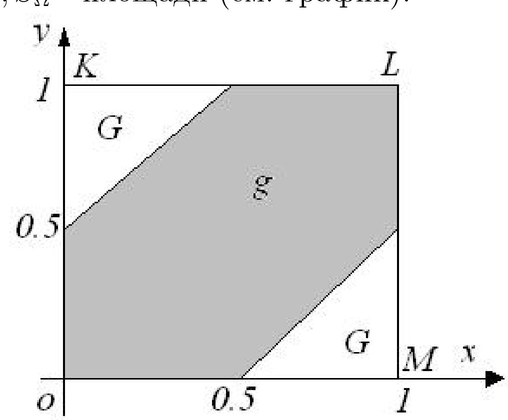
где $m(A), m(\Omega)$ – меры соответствующих множеств.

Пример 2. *Задача о встрече.* 2 студента договорились о встрече в промежуток 13:00-13:30. Каждый из них приходит равновозможно в этот интервал, ждёт другого 15 минут, потом уходит. Найти вероятность того, что студенты встретятся.

Решение. Пусть x — время прихода студента A, y — время прихода студента B.
$x \in [0; 30]$, $y \in [0; 30]$, $|x - y| \leq 15$

$$P(A) = \frac{S_A}{S_\Omega} = \frac{30^2 - 2 \cdot \frac{1}{2} \cdot 15^2}{30^2} = \frac{3}{4}.$$

S_A, S_Ω - площади (см. график):



Пример 3. *Игла Бюффона.* Плоскость разлинована параллельными прямыми, расстояние между которыми равно L. Бросают иглу длины l ($l \leq L$). Найти вероятность,

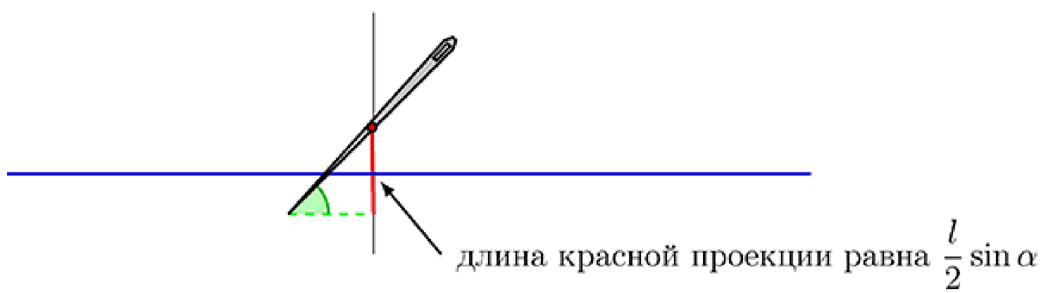
что игла пересечёт прямую.

Решение. Пусть α — угол между иглой и ближайшей прямой, x — расстояние от середины иглы до ближайшей прямой.
Имеем:

$$x \in [0; \frac{L}{2}], \qquad \alpha \in [0, \pi].$$

Условие того, что игла пересечет прямую, имеет вид:

$$x \leq \frac{l}{2} \sin \alpha.$$



Тогда

$$P(A) = \frac{S_A}{\pi \frac{L}{2}} = \frac{2l}{\pi L},$$

поскольку

$$S_A = \int_0^{\pi} \frac{l}{2} \sin \alpha \ d\alpha = l.$$

См. также парадокс Бертрана.

1.6 Аксиоматическое определение вероятностей А.Н. Колмогорова

~~1. $|\Omega| = n$ — пространство элементарных исходов конечно.~~

~~2. Все элементарные исходы равновозможны (*симметрия шансов*).~~

В таком случае вероятность может быть введена аксиоматически. Продолжение следует...

1.7 Основы комбинаторики. Неупорядоченный и упорядоченный выбор с возвращением и без возвращения. Перестановки, сочетания, размещения. Формула включений-исключений

Перестановки — комбинации из n элементов, отличающиеся только порядком расположения элементов.

Число всевозможных перестановок:

$$P_n = n!$$

Замечание.

$$0! = 1$$

Задача. Каким числом способов можно посадить в ряд 5 человек?

Ответ:

$$5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120.$$

Размещения — комбинации из n элементов по m элементов, отличающиеся либо составом, либо порядком элементов.

Число всевозможных размещений:

$$A_n^m = n(n-1)\dots(n-m+1) = \frac{n!}{(n-m)!}$$

Задача. Каким числом способов можно выбрать победителей турнира (1, 2, 3 место) из 25 человек?

Ответ:

$$A_{25}^3 = 25 \cdot 24 \cdot 23 = 13800.$$

Сочетания — комбинации из n элементов по k элементов, отличающаяся только составом элементов. Число всевозможных сочетаний:

$$C_n^k = \frac{n!}{k!(n-k)!} = C_n^{n-k}.$$

Задача. Каким числом способов можно выбрать троих из 25 человек?

Ответ:

$$C_{25}^3 = \frac{25!}{3!22!} = \frac{23 \cdot 24 \cdot 25}{2 \cdot 3} = 2300.$$

Заметим, что

$$13800 = A_{25}^3 = C_{25}^3 \cdot P_3 = 2300 \cdot 6.$$

Имеем:

$$A_n^k = C_n^k \cdot P_k.$$

Правило произведения. Пусть имеется m элементов: $a_1, \ldots, a_m$ и n элементов: $b_1, \ldots, b_n$, тогда пару (a_i, b_j) можно выбрать $(n \cdot m)$ способами.

Правило сложения. Если объект A можно выбрать n способами, а объект B можно выбрать m способами, тогда выбрать либо A, либо B можно $(n + m)$ способами.

Выборки. Пусть имеется урна, содержащая n пронумерованных объектов (шаров). Мы выбираем из этой урны m шаров. Результатом выбора является набор из m шаров. Нас интересует, сколькими способами можно выбрать m шаров из n, или сколько различных результатов может получиться. На этот вопрос нельзя дать однозначный ответ, пока мы не определимся с тем, как организован выбор (можно ли шары возвращать в урну), и с тем, что понимается под различными результатами выбора. Рассмотрим 4 возможных случая:

1. Упорядоченная выборка с возвращением. (Шары вынимаются по одному, записывается номер и шар убирается обратно в урну). Количество вариантов выбора m шаров из n:
$$|\Omega| = n^m.$$
2. Упорядоченная выборка без возвращения $(a_{i_1}, a_{i_2}, \ldots, a_{i_m})$, $i_1 \neq i_2 \neq \ldots \neq i_m$ (Шары последовательно извлекают, записывают номер и не возвращают обратно):
$$|\Omega| = A_n^m = n(n-1) \cdot \ldots \cdot (n-m+1) = \frac{n!}{(n-m)!}$$
3. Неупорядоченная выборка без возвращений $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_m} \rangle, i_1 \neq i_2 \neq \ldots \neq i_m$ (Выбираем m шаров из n, порядок не важен, шары не возвращаем):
$$|\Omega| = C_n^m = \frac{A_n^m}{m!} = \frac{n!}{m!(n-m)!}.$$

4. Неупорядоченная выборка с возвращением $\langle a_{i_1}, a_{i_2}, \ldots, a_{i_m}\rangle$ (Выбираем m шаров из n, порядок не важен, шары возвращаем):

$$|\Omega| = C_{n+m-1}^{m} = C_{n+m-1}^{n-1}$$

1.7.1 Примеры непосредственного вычисления вероятности по классической формуле

Задача 1. Из группы 25 человек, в которой имеется 16 мужчин, наугад выбирают двоих. Найти вероятность того, что это люди разного пола.

Решение:

$$P(A) = \frac{16 \cdot 9}{C_{25}^2} = 12/25 = 0,48.$$

Задача 2. Из колоды в 52 карты наугад выбирают три. Найти вероятность того, что это тройка, семерка, туз без учета порядка.

Решение:

$$P(A) = \frac{4 \cdot 4 \cdot 4}{C_{52}^3} \approx 0,003.$$

Задача 3

Набирая номер телефона, абонент забыл последние три цифры, но помнит, что они различны. Он набрал номер наудачу. Найти вероятность того, что был набран нужный номер.

Решение: Элементарным событием в данной задаче является набор из трёх цифр, причем важен порядок, в котором они расположены (изменения порядка цифр приводит к изменению номера телефона). Повторения исключены. Значит, число возможных наборов будем искать как число размещений без повторений из 10 (всего цифр) по 3: $|\Omega| = A_{10}^3$. Пусть событие А заключается в том, что абонент дозвонился, тогда $|A| = 1$, т.к. только один набор соответствует нужному номеру телефона. Значит, $P(A) = \frac{|A|}{|\Omega|} = \frac{1}{A_{10}^3} = \frac{1}{720}$.

Задача 4

Числа 1, 2, 3, 4, 5 написаны на пяти карточках. Наугад последовательно выбираются три карточки, и вытянутые таким образом цифры ставятся слева направо. Найти вероятность того, что полученное при этом трехзначное число будет четным.

Решение: Элементарным событием в данной задаче является набор из трёх цифр от 1 до 5, причем важен порядок, в котором они расположены, и цифры не повторяются. Значит, число возможных полученных чисел будем искать как число размещений без повторений из 5 по 3: $|\Omega| = A_5^3$.

Пусть событие A – выбранное число четное. Чтобы число было четным, оно должно заканчиваться цифрой 2, либо 4. В каждом из этих случаев, оставшиеся 2 цифры можно выбрать A_4^2 способами, то есть $|A| = A_4^2 + A_4^2$ (используем принцип сложения). Получаем, $P(A) = \frac{|A|}{|\Omega|} = \frac{2A_4^2}{A_5^3} = 0,4$.

Задача 5

В урне находится 10 белых и 15 черных шаров. Наудачу из урны извлекают 5 шаров. Найти вероятности событий: A={все выбранные шары белые}, B={среди выбранных шаров два белых и три черных}, C={среди выбранных шаров хотя бы 1 белый}.

Решение: Элементарным событием в данной задаче является набор из пяти шаров, причем важен только состав набора (порядок не важен). Значит, число возможных наборов будем искать как число сочетаний из 25 (всего 10+15 шаров) по 5: $|\Omega| = C_{25}^5$. Очевидно, что $|A| = C_{10}^5$, тогда $P(A) = \frac{|A|}{|\Omega|} = \frac{C_{10}^5}{C_{25}^5} \approx 0,005$.

Выбрать два белых шара из 10 можно C_{10}^2 способами. Выбрать три черных шара из 15 можно C_{15}^3 способами. Используя правило произведения, получаем $|B| = C_{10}^2 \cdot C_{15}^3$. Тогда $P(B) = \frac{C_{10}^2 \cdot C_{15}^3}{C_{25}^5} \approx 0,39$.

Для вычисления вероятности события C перейдём к противоположному событию $\overline{C}$ ={все выбранные шары — черные}. Выбрать пять черных шаров из 15 можно C_{15}^5 способами. $|\overline{C}| = C_{15}^5$, $P(|\overline{C}|) = \frac{C_{15}^5}{C_{25}^5}$. Учитывая, что $P(C) = 1 - P(\overline{C})$, получаем $P(C) = 1 - \frac{C_{15}^5}{C_{25}^5} \approx 0,94$.

1.8 Основные теоремы теории вероятностей. Условная вероятность. Независимость событий

Теорема 1. О сложении для несовместных событий.

$\forall A, B, \quad$ т.ч. $\; A \cdot B = \varnothing \Rightarrow P(A+B) = P(A) + P(B)$

(Для двух несовместных событий A и B вероятность появления либо A, либо B равна сумме вероятностей событий A и B.)

Следствие.

$$P(\bar{A}) = 1 - P(A).$$

Задача. В корзине находится 10 шаров: 5 белых, 3 синих, 2 желтых. Один шар извлекается наудачу. Найти вероятность того, что он цветной (либо синий, либо желтый).

Решение: Пусть событие A – вытянут шар желтого цвета, событие B — шар синего цвета. Тогда

$$P(A+B) = P(A) + P(B) = 3/10 + 2/10 = 0,5.$$

Другой способ решения. Пусть C — событие, состоящее в том, что извлечен белый шар. Тогда событие C являетя противоположным к событию $A+B$ (вытянут цветной шар).

Тогда

$$P(A+B) = 1 - P(\overline{A+B}) = 1 - P(C) = 1 - 5/10 = 5/10 = 0,5.$$

Теорема 2. О сложении для произвольных событий.

$\forall A, B \;\; P(A+B) = P(A) + P(B) - P(A \cdot B)$

Следствие.

$$P(A+B+C) = P(A) + P(B) + P(C) - P(A \cdot B) - P(B \cdot C) - P(A \cdot C) + P(A \cdot B \cdot C).$$

Задача

Среди целых чисел от 1 до 30 наудачу извлекают одно. Найти вероятность того, что это число делится на два или на три.

Решение:

Рассмотрим события A={Число делится на два}, B={Число делится на три.}

Имеем:

$$\Omega = \{1, \dots, 30\}, \qquad |\Omega| = 30$$

$$A = \{2, 4, 6, 8, 10, \dots, 30\}, \qquad |A| = 15$$

$$B = \{3, 6, 9, 12, \dots, 30\}, \qquad |B| = 10$$

Необходимо найти $P(A+B)$. По формуле сложения вероятностей:

$$P(A+B) = P(A) + P(B) - P(A \cdot B).$$

$$A \cap B = \{6, 12, 18, 24, 30\}, \qquad |A \cap B| = 5$$

$P(A) = \frac{1}{2}$, $P(B) = \frac{1}{3}$. Событие $A \cdot B$ включает в себя все числа от 1 до 30 кратные шести, значит $P(A \cdot B) = \frac{5}{30} = 1/6$. Тогда

$$P(A+B) = \frac{1}{2} + \frac{1}{3} - \frac{1}{6} = \frac{2}{3}.$$

Задача

Когда 3 друга уходили из гостей, неожиданно погас свет, и им пришлось надевать шапки наугад в темноте. Найти вероятность того, что хотя бы один из них ушел в своей шапке.

Решение:

Зададим события A_i=\{i-й друг ушёл в своей шапке\}.

Необходимо найти $P(A_1 + A_2 + A_3)$. По формуле сложения вероятностей:

$$P(A_1 + A_2 + A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cdot A_2) - P(A_1 \cdot A_3) - P(A_2 \cdot A_3) +$$

$$+ P(A_1 \cdot A_2 \cdot A_3).$$

Мощность множества Ω равна числу перестановок трёх элементов. $|\Omega| = 3! = 6$. Мощность множества A_1 равна числу перестановок двух элементов (варианты для 2-го и 3-го друга). $|A_1| = 2!$. Тогда

$$P(A_1) = \frac{2!}{|\Omega|} = \frac{1}{3} = P(A_2) = P(A_3).$$

Аналогично

$$P(A_1 \cdot A_2) = \frac{|A_1 \cdot A_2|}{|\Omega|} = \frac{1}{6} = P(A_1 \cdot A_3) = P(A_2 \cdot A_3),$$

$$P(A_1 \cdot A_2 \cdot A_3) = \frac{|A_1 \cdot A_2 \cdot A_3|}{|\Omega|} = \frac{1}{6}.$$

Значит,

$$P(A_1 + A_2 + A_3) = \frac{1}{3} + \frac{1}{3} + \frac{1}{3} - \frac{1}{6} - \frac{1}{6} - \frac{1}{6} + \frac{1}{6} = \frac{2}{3}.$$

Условная вероятность. $P(A/B)$ — вероятность события A при условии, что событие B произошло:

$$P(A/B) = \frac{P(A \cdot B)}{P(B)}.$$

Аксиоматика вероятностной меры для условной вероятности выполнена. $P(B)$ — безусловная вероятность, $P(B) > 0$.

Теорема 3. Об умножении вероятностей. Вероятность того, что произойдёт и событие A, и событие B равна

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

Следствие. $P(A \cdot B \cdot C) = P(C) \cdot P(B/C) \cdot P(A/(B \cdot C))$

Задача

В урне M белых и N-M черных шаров. По схеме "выборка без возвращения"последовательно выбирают 2 шара. Найти вероятность того, что оба окажутся белыми.

Решение: Пусть A=$\{$1-й вытянутый шар белый$\}$, B=$\{$2-й вытянутый шар белый$\}$. Необходимо найти $P(AB)$.

$$P(AB) = P(A)P(B/A),$$

где $P(A) = \frac{M}{N}$, $P(B/A) = \frac{M-1}{N-1}$ (учитываем, что один белый шар из урны уже достали).Тогда

$$P(AB) = \frac{M(M-1)}{N(N-1)}.$$

Задача. Из колоды в 52 карты наугад выбирают три. Найти вероятность того, что это тройка, семерка, туз в указанном порядке.

Решение: ЭТО НЕ ЕДИНСТВЕННЫЙ СПОСОБ РЕШЕНИЯ

$$P(A \cdot B \cdot C) = 4/52 \cdot 4/51 \cdot 4/50.$$

Определение. Два события A и B называются *независимыми*, если вероятность их совместного появления равна произведению их вероятностей:

$$P(A \cdot B) = P(A) \cdot P(B).$$

Для независимых событий $P(A/B) = P(A)$, при условии $P(B) > 0$.

Замечание. Пусть события A и B несовместны. Из этого не следует их независимость.

Доказательство. Самостоятельно $\square$

Множество событий $A_1, \dots, A_n$ называются *попарно независимыми*, если
$\forall A_i, A_j \;\; P(A_i \cdot A_j) = P(A_i) \cdot P(A_j)$

Множество событий $A_1, \dots, A_n$ называются *независимыми в совокупности*, если
$\forall A_{i_1}, \dots, A_{i_m} \;\; P(\cap_{k=1}^{m} A_{i_k}) = \prod_{k=1}^{m} P(A_{i_k})$

Пример – "Пирамида Бернштейна". Правильная четырёхгранная пирамида с одинаковой вероятностью 1/4 падает на одну из своих граней. Три любые грани раскрашиваются в три цвета: красный, зелёный и синий. Четвёртая грань делится на три части, каждая из которых раскрашивается одним из указанных цветов.

Рассмотрим 4 события:

1. $A = \{\text{красный}\}$ – пирамида упала гранью, содержащей красный цвет.
2. $B = \{\text{синий}\}$ – пирамида упала гранью, содержащей синий цвет.
3. $C = \{\text{зелёный}\}$ – пирамида упала гранью, содержащей зелёный цвет.

A, B, C содержат два элементарных исхода, следовательно

$$P(A) = P(B) = P(C) = 2/4 = 1/2.$$

Если эти события независимые, то согласно формуле вероятности совместного появления независимых событий

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C) = (1/2)^3 = 1/8$$

Но в нашем случае одновременное осуществление всех трёх событий возможно лишь при выпадении грани, содержащей все три цвета, поэтому $P(A \cdot B \cdot C) = 1/4 \neq 1/8$. Тождество не выполняется, а значит события A, B, C — не независимы в совокупности.

С другой стороны, события A, B, C попарно независимы. В самом деле,

$$P(A \cdot B) = 1/4 = P(A) \cdot P(B)$$

Точно такие же равенства справедливы для остальных пар.

1.9 Формула полной вероятности

Предположим, что событие A может произойти только одновременно с одним из событий из системы гипотез:

$$H_1, H_2, \ldots, H_n$$

где $H_1, H_2, \ldots, H_n$ — полная группа событий, то есть

$$H_1 + H_2 + \ldots + H_n = \Omega, \quad H_i \cdot H_j = \varnothing, \quad \forall i, j, \quad i \neq j.$$

Тогда

$$P(A) = P(A \cdot \Omega) = P(A \cdot (H_1 + \ldots + H_n)) = \sum_{i=1}^{m} P(A \cdot H_i) = \sum_{i=1}^{n} P(H_i) \cdot P(A/H_i)$$

Последнее равенство — *формула полной вероятности*:

$$P(A) = P(H_1) \cdot P(A/H_1) + \ldots + P(H_n) \cdot P(A/H_n) = \sum_{i=1}^{n} P(H_i) \cdot P(A/H_i). \quad (1.7)$$

Пример 1. В тире 5 ружей. Вероятность попаданий из этих ружей в мишень разная: $0.8, 0.7, 0.6, 0.4, 0.3$. Человек берёт одно ружье наудачу и производит выстрел. Какова вероятность попадания в мишень?

Решение. Вероятность попаданий из этих ружей в мишень разная: $0.8, 0.7, 0.6, 0.4, 0.3$. Имеем 5 гипотез — ружье с каким номером было выбрано: $H_1, \ldots, H_5$ такие, что $P(H_i) = \frac{1}{5}$, $i = 1..5$ - вероятность, что будет выбрано i-е оружие. Тогда

$$P(A) = \frac{1}{5}0.8 + \frac{1}{5}0.7 + \frac{1}{5}0.6 + \frac{1}{5}0.4 + \frac{1}{5}0.3 = 0.56.$$

Пример 2. Студент N знает m билетов из n. Первым или последним следует брать билет студенту N?

Решение.

Событие A — студент N вытянул билет, который он знает. Пусть он берёт билет первым, тогда $P(A) = m/n$.

Пусть студент N тянет билет вторым: H_1 — первый (предыдущий) студент взял билет из m известных студенту с именем N билетов.

$$P(H_1) = \frac{m}{n}, \quad P(A/H_1) = \frac{m-1}{n-1}$$

Другая гипотеза H_2 — первый студент взял не из m известных студенту с именем N билетов. Тогда

$$P(H_2) = \frac{n-m}{n}, \quad P(A/H_2) = \frac{m}{n-1}$$

в итоге:

$$P(A) = \frac{m}{n} \cdot \frac{m-1}{n-1} + \frac{n-m}{n} \cdot \frac{m}{n-1} = \frac{m^2 - m + mn - m^2}{n(n-1)} = \frac{m}{n}$$

И т.д. Так как вероятность с каждым разом остаётся постоянной (доказывается по индукции), то в итоге получим, что для студента неважен порядок выбора билета.

Задача 3

Два автомата производят детали, которые сбрасываются на общий конвеер. Производительность первого автомата вдвое больше второго. 60% деталей, производимых первым автоматом, отличного качества, и 84% деталей, производимых вторым – отличного качества.

1. Найти вероятность, что наудачу взятая деталь – отличного качетсва;
2. Наудачу взятая с конвеера деталь имеет отличное качество. Найти вероятность того, что деталь произведена первым автоматом.

Решение:

1. Введём события:

   A={деталь отличного качества}.

   B_1={деталь произведена первым автоматом}.

   B_2={деталь произведена вторым автоматом}.

   $P(B_1) = \frac{2}{3}$, $P(B_2) = \frac{1}{3}$, $P(A/B_2) = 0,6$, $P(A/B_1) = 0,84$.

$$P(A) = P(B_1)P(A/B_1) + P(B_2)P(A/B_2) = \frac{2}{3} \cdot 0,6 + \frac{1}{3} \cdot 0,84 = 0,68$$

2.
$$P(B_1/A) = \frac{P(A/B_1)P(B_1)}{P(A)} = \frac{\frac{2}{3} \cdot 0,6}{0,68} = \frac{10}{17}.$$

Задача 4

Имеется две партии приборов: первая состоит из 100 приборов, среди которых 4 дефектных, вторая состоит из 80 приборов, среди которых 2 дефектных. Из первой партии берется 10 приборов, а из второй - 15 приборов. Из новой партии берут один прибор. Какова вероятность того, что он будет дефектным?

Решение:

Пусть событие A={из новой партии взяли дефектный прибор}. Выскажем следующие гипотезы: B_1 ={взятый из новой партии прибор, первоначально находился в первой партии}; B_2 ={взятый из новой партии прибор, первоначально находился во второй партии}. Очевидно, что эти гипотезы образуют полную группу. Вычислим вероятности гипотез: $P(B_1) = 0,4$ и $P(B_2) = 0,6$ (в новой партии 25 приборов, из них 10 из первой и 15 из второй партии). Проверим, что $P(B_1) + P(B_2) = 1$. Вычислим условные вероятности события А: $P(A/B_1) = 0,04$ (так как мы вычисляем вероятность события А, считая, что прибор взят из первой партии, а в этой партии всего 100 приборов и среди них 4 дефектных), аналогично $P(A/B_2) = 0,025$. По формуле полной вероятности найдем: $P(A) = 0,4 \cdot 0,04 + 0,6 \cdot 0,025 = 0,031$.

Задача 5

В ящике лежат 20 теннисных мячей, в том числе 15 новых и 5 игранных. Для игры наудачу выбираются два мяча и после игры возвращаются обратно. Какова вероятность того, что на вторую игру выберут наудачу новые мячи?

Решение:

Пусть событие A={на вторую игру взяли два новых мяча}. Выскажем следующие гипотезы: B_1 ={на первую игру взяли два новых мяча}; B_2={на первую игру взяли 1 новый и 1 игранный мяч}; B_3 ={на первую игру взяли два игранных мяча}. Очевидно, что события B_1, B_2, B_3 попарно несовместны и, в результате опыта, одно из них должно произойти. Вычислим вероятности гипотез. Два мяча из 20 можно выбрать C_{20}^2 способами, два новых мяча из 5 – C_{15}^2 способами и выбрать 1 новый и 1 играный мяч – $C_{15}^1 C_5^1$ способами.

$$P(B_1) = \frac{C_{15}^2}{C_{20}^2} = \frac{21}{38},$$

$$P(B_2) = \frac{C_{15}^1 C_5^1}{C_{20}^2} = \frac{15}{38},$$

$$P(B_3) = \frac{C_5^2}{C_{20}^2} = \frac{1}{19}.$$

Проверим, что $P(B_1) + P(B_2) + P(B_3) = 1$. Вычислим условные вероятности события А, также с использованием классического определения вероятности: $P(A/B_1) = \frac{C_{13}^2}{C_{20}^2} = \frac{39}{95}$, $P(A/B_2) = \frac{C_{14}^2}{C_{20}^2} = \frac{91}{190}$, $P(A/B_3) = \frac{C_{15}^2}{C_{20}^2} = \frac{21}{38}$. . По формуле полной вероятности найдем: $P(A) = \frac{1}{19} \cdot \frac{39}{95} + \frac{15}{38} \cdot \frac{91}{190} + \frac{21}{38} \cdot \frac{21}{38} = 0,516$.

1.9.1 Переоценка гипотез. Формулы Байеса

Пусть дан набор гипотез

$$H_1, \ldots, H_n$$

Тогда $P(H_1), \ldots, P(H_n)$ — *априорные* вероятности гипотез.

Нужно оценить вероятность гипотез: $P(H_i/A)$ — *апостериорная* вероятность гипотезы.

$$\sum_{i=1}^{n} p(H_i) = 1$$

$$P(H_i/A) = \frac{P(H_i \cdot A)}{P(A)} = \frac{P(H_i) \cdot P(A/H_i)}{\sum_{i=1}^{n} P(H_i) \cdot P(A/H_i)}, \qquad i = 1, \ldots, n. \tag{1.8}$$

Задача 1

Три студента печатают некоторый текст. Владимир напечатал 40% общего объема, Иван – 50%, Мария – 10% (объем считается по количеству слов). Известно, что Владимир делает ошибки в 5% слов, Иван – в 6%, Мария – в 10% слов. Случайно выбранное слово оказалось напечатанным с ошибкой. Найдите вероятность того, что это слова напечатано Иваном.

Решение: Введём события:

V={слово напечатано Владимиром},

I={слово напечатано Иваном},

M={слово напечатано Марией},

E={слово напечатано с ошибкой}.

$$P(E) = P(V)P(E/V) + P(I)P(E/I) + P(M)P(E/M) =$$

$$= 0,4 \cdot 0,05 + 0,5 \cdot 0,06 + 0,10 \cdot 0,10 = 0,06.$$

$$P(I/E) = \frac{P(I)P(E/I)}{P(E)} = \frac{0,03}{0,06} = 0,5$$

Задача 2

Для снижения количества краж на производстве руководство компании принимает решение подвергнуть всех работников испытанию на детекторе лжи. Детектор в 90%

случаев делает правильный вывод (как для виновных, так и для невиновных). Компания собирается уволить всех работников, не прошедших испытание. Предположим, что 5% работников время от времени занимаются кражей на производстве.

1. Найти вероятность, что уволят невиновного работника;
2. Найти вероятность того, что случайный работник, которого не уволили, виновен.

Решение:

Введём события:

I={работник не виновен},

S={работник прошел испытание}.

Тогда $P(I) = 0,95$, $P(S/I) = P(\overline{S}/\overline{I}) = 0,9$ и, значит, $P(\overline{S}/I) = P(S/\overline{I}) = 0,1$.

По формуле Байеса получаем:

1. $P(I/\overline{S}) = \frac{P(\overline{S}/I)P(I)}{P(\overline{S}/I)P(I)+P(\overline{S}/\overline{I})P(\overline{I})} = \frac{0,1\cdot 0,95}{0,1\cdot 0,95+0,9\cdot 0,05} = 0.6786;$
2. $P(\overline{I}/S) = \frac{P(S/\overline{I})P(\overline{I})}{P(S/I)P(I)+P(S/\overline{I})P(\overline{I})} = \frac{0,1\cdot 0,05}{0,1\cdot 0,05+0,9\cdot 0,95} = 0.0058;$

To be continued

Парадок Монти Холла

1.10 Повторные испытания Якоба Бернулли

Схема Бернулли (повторные испытания Бернулли) — последовательность из n независимых испытаний, в каждом из которых вероятность появления события A одинакова и равна p ($p(A) = p$ в каждом испытании). Если событие A произошло, то говорят, что произошел "успех", иначе ($\bar{A}$) — "неудача".

Очевидно, что в каждом испытании вероятность успеха равна p. Вероятность неудачи принято обозначать как $q = 1 - p$.

В таких задачах нас часто интересует вероятность того, что из n испытаний будет ровно k успешных:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}, \qquad k = 0, \dots, n, \qquad q = 1 - p. \tag{1.9}$$

Формула (1.9) — *формула Бернулли.*

Вычислим вероятности следующих событий:

1. Событие A произошло менее k раз: $\sum_{i=0}^{k-1} P_n(i)$
2. Событие произошло более k раз: $\sum_{i=k+1}^{n} P_n(i)$

3. Событие произошло не менее k раз: $\sum_{i=k}^{n} P_n(i)$
4. Событие произошло не более k раз: $\sum_{i=0}^{k} P_n(i)$

События, рассмотренные в пунктах 1, 3 и 2, 4 являются противоположными.

Пусть событие B состоит в том, что в схеме Бернулли из n независимых испытаний событие A появилось хотя бы один раз. Тогда событие $\bar{B}$ состоит в том, что событие A не появилось ни одного раза. Имеем:

$$P(B) = 1 - P(\bar{B}) = 1 - P_n(0) = 1 - q^n.$$

Наиболее вероятное число успехов в схеме Бернулли

Пусть m_0 — число успехов, при котором $P_n(k)$ принимает наибольшее значение.

- Справедливо следующее неравенство:

  $$np - q \leqslant m_0 < np + p.$$

  Заметим, что $np + p = np - q + 1$. Тогда неравенство можно записать так:

  $$np - q \leqslant m_0 < np - q + 1.$$

  Имеем:

  – Если np — целое, то $m_0 = np$. Иначе см. далее.
  – Если $(np - q)$ — целое, тогда существует два наиболее вероятных числа успехов: $m_0 = np - q$ и $m_0 = np + p$. Иначе см. далее.
  – Если $(np - q)$ — дробное число, то $m_0 = [np + p]$ (целая часть числа обозначена квадратными скобками).

Пример. Симметричную монету подбрасывают 3 раза. Пусть событие A (успех) — выпадение герба. Имеем $n = 3$, $p = \frac{1}{2}$. Тогда
вероятность того, что 2 раза выпадет герб: $P_3(2) = C_3^2 \cdot (\frac{1}{2})^2 \cdot \frac{1}{2} = \frac{3}{8}$;
вероятность того, что 1 раз выпадет герб: $P_3(1) = C_3^1 \cdot (\frac{1}{2})^3 = \frac{3}{8}$;
вероятность того, что ни разу не выпадет герб: $P_3(0) = q^3 = (\frac{1}{2})^3 = \frac{1}{8}$;
вероятность того, что хотя бы один раз выпадет герб: $1 - P_3(0) = \frac{7}{8}$;
вероятность того, что хотя бы два раза выпадет герб: $P_3(2) + P_3(3) = \frac{3}{8} + (\frac{1}{2})^3 = 1 - P_3(0) - P_3(1) = 1 - \frac{4}{8} = \frac{4}{8} = \frac{1}{2}$.

Задача 1

В квадрат со стороной $2a$ вписан круг. В квадрат наудачу бросают 4 точки. Найти вероятность следующих событий: A={две точки попали в круг}, B={хотя бы три точки попали в круг}.

Решение:

Производится 4 независимых испытания (бросают 4 точки), $n = 4$.

Вероятность успеха найдем, используя формулу геометрической вероятности, как отношение площади круга к площади квадрата:

$$p = \frac{\pi a^2}{4a^2} = \frac{\pi}{4}.$$

Тогда

$$P(A) = C_4^2 p^2 q^2 = \frac{4!}{2!2!} \cdot \frac{\pi^2}{16} \left(1 - \frac{\pi}{4}\right)^2 \approx 0,17.$$

Найдём вероятность события B, которое заключается в том, что в круг попали 3 или 4 точки. Заметим, что $P(B) = C_4^3 p^3 q^1 + C_4^4 p^4 = 4p^3(1-p) + p^4 \approx 0,76$.

Задача 2

Контрольное задание состоит из шести вопросов. На каждый вопрос представлены 4 ответа, среди которых необходимо выбрать один правильный. Найдите вероятность события: A={методом простого угадывание удастся ответить, по крайней мере, на 5 вопросов}.

Решение:

Производится 6 независимых испытаний (6 вопросов), $n = 6$. Поскольку ответ выбирается угадыванием, вероятность успеха в каждом испытании равна $p = \frac{1}{4}$.

Событие А заключается в том, что отвечено на 5 или 6 вопросов. Значит,

$$P(A) = C_6^5 p^5 q + C_6^6 p^6 = 6 \cdot (0,25)^5 \cdot 0,75 + (0,25)^6 = 0,0046.$$

Задача 3

Показать, что более вероятно: при одновременном бросании четырёх костей получить хотя бы одну единицу или при 24-х бросаниях двух костей получить хотя бы один раз две единицы. Ответ известен как парадокс де Мере. Игрок шевалье де Мере считал эти вероятности равными и обвинял математиков в своих проигрышах.

Решение:

Рассмотрим первый опыт, который заключается в одновременном бросании четырёх костей. Можно сказать, что производится 4 независимых испытания (подбрасывание одной монеты – испытание), $n_1 = 4$. Успех – выпадение единицы, $p_1 = \frac{1}{6}, q_1 = \frac{5}{6}$. Пусть A=$\{$при одновременном бросании четырёх костей получи хотя бы одну единицу$\}$. Рассмотрим событие $\overline{A}$, которое заключается в том, что при одновременном бросании четырёх костей единиц не получили.

$$P(\overline{A}) = C_4^0 p_1^0 q_1^4 = q_1^4 = \left(\frac{5}{6}\right)^4,$$

$$P(A) = 1 - P(\overline{A}) = 1 - \left(\frac{5}{6}\right)^4 \approx 0,518.$$

Рассмотрим теперь событие B=$\{$при 24-х бросаниях двух костей получили хотя бы один раз две единицы$\}$. В этом опыте производится $n_2 = 24$ независимых испытания. Вероятность выпадения двух единиц в одном испытании равна $p_2 = \frac{1}{36}$.

$$P(\overline{B}) = C_{24}^0 p_2^0 q_2^{24} = q_2^{24} = \left(\frac{35}{36}\right)^{24},$$

$$P(B) = 1 - P(\overline{B}) = 1 - \left(\frac{35}{36}\right)^{24} \approx 0,491.$$

Получили, что событие A более вероятно.

Задача 4

В результате каждого визита страхового агента договор заключается с вероятностью 0,1. Найти наивероятнейшее число заключенных договоров после 25 визитов.

Решение:

По условию задачи $n = 25$, $p = 0,1$, $q = 0,9$.

Наивероятнейшее число успехов m^* в серии из n испытаний удовлетворяет условию $np - q \leq m^* \leq np + p$.

$$np - q = 25 \cdot 0,1 - 0,9 = 1,6,$$

$$np + p = 25 \cdot 0,1 + 0,1 = 2,6.$$

Получаем, что $1,6 \leq m^* \leq 2,6$. Значит, $m^* = 2$. Наивероятнейшее число заключенных договоров после 25 визитов равно двум.

Задача 5 Проверяется качествов партии из n товаров. Пусть вероятность того, что товар имеет хорошее качество, равна p. Найти наиболее вероятное количество качественных товаров в партии, если

a) $n = 15$; $p = 0,9$;

b) $n = 24$, $p = 0,6$;

c) $n = 12$, $p = 0,5$.

a) $q = 1 - p = 0,1$;

$np - q = 15 \cdot 0,9 - 0,1 = 13,4$ — дробное число;

$np + p = 14,4$;

$13,4 \leqslant m_0 < 14,4$, поскольку m_0 целое, то $m_0 = 14$;

b) $q = 1 - p = 0,4$;

$np - q = 24 \cdot 0,6 - 0,4 = 14$ — целое; два наиболее вероятных числа $m_0 = 14$ и $m_0 = 15$;

c) $np = 12 \cdot 0,5 = 6$ целое; $m_0 = 6$.

1.11 Формула Пуассона

Напомним формуллу Бернулли для вычисления вероятности появления ровно k успехов в схеме Бернулли, состоящей из n независимых испытаний.

Имеем:

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k}, \qquad q = 1 - p, \qquad k = 0, \ldots, n.$$

Теорема Пуассона. Пусть имеется последовательность серий испытания Бернулли. Предположим, что

1. $n \to \infty$ (n достаточно велико);
2. Вероятность успеха $p \to 0$ (p достаточно мало);
3. $np \to \lambda > 0$,

тогда

$$P_n(k) \approx \frac{\lambda^k}{k!}e^{-\lambda}, \qquad np = \lambda.$$

формула Пуассона

Закон редких явлений — см. далее в разделе распределение Пуассона. Сорняки в урожае, изюминки в булочках, количество разбитых бутылок при перевозке, звезды в пространстве, количество несчастных случаев, число левшей и пр.

Доказательство.

$$P_n(k) = C_n^k \cdot p^k \cdot q^{n-k} = \frac{n!}{k!(n-k)!} \cdot \left(\frac{\lambda}{n}\right)^k \cdot \left(1 - \frac{\lambda}{n}\right)^{n-k} =$$

$$= \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n^k} \cdot \frac{\lambda^k}{k!} \cdot \left(1 - \frac{\lambda}{n}\right)^n \cdot \left(1 - \frac{\lambda}{n}\right)^{-k}$$

$$\lim_{n\to\infty} P_n(k) = \frac{\lambda^k}{k!} \cdot \lim_{n\to\infty} \left(1 - \frac{\lambda}{n}\right)^n = \frac{\lambda^k}{k!}e^{-\lambda}.$$

Задача 1

8 процентов населения составляют левши. Найти вероятность того, что среди 200 студентов ровно 4 левши.

Решение: $n = 200,\ p = 0,08,\ \lambda = 200 * 0,08 = 16$ (в среднем 16 левшей на 200 человек).

$$P_200(4) \approx \frac{16^4}{4!}\,e^{-16} \approx 0,00034.$$

Задача 2

В городке 1 000 домов, каждый из которых застрахован от пожара в некоторой страховой компании на сумму 1 000 000 рублей. Страховой взнос за год составляет 2 000 рублей. Для данного городка вероятность пожара в доме в течение года оценивается как 0,003. Какова вероятность того, что в течение года страховая компания потерпит убытки?

Решение:

Страховая компания собирает в городке со всех домов общий взнос в 2 000 000 рублей. Следовательно, если сгорят за год не менее, чем три дома, то компания потерпит убытки. Найдём вероятность события A={за год сгорят не менее, чем три дома}.

В данной задаче производится $n = 1000$ независимых испытаний, с вероятностью успеха в каждом $p = 0,003$. Можем считать пожар редким событием и воспользоваться для вычисления приближенного значения вероятности теоремой Пуассона.

Вычислим параметр $\lambda = 1000 \cdot 0,0003 = 3$. В данной задаче удобнее перейти к событию $\overline{A}$={сгорит меньше чем три дома}. Вычислим предварительно вероятности событий: «не сгорит ни один дом», «сгорит один дом», «сгорят два дома».

$$P_{1000}(0) \approx \frac{3^0 e^{-3}}{0!} \approx 0,0498,$$

$$P_{1000}(1) \approx \frac{3^1 e{-3}}{1!} \approx 0,1494,$$

$$P_{1000}(2) \approx \frac{3^2 e{-3}}{2!} \approx 0,2241.$$

Тогда $P(\overline{A}) = 0,0498 + 0,1494 + 0,2241 = 0,4233$.

Найдём теперь $P(A) = 1 - P(\overline{A}) = 1 - 0,4233 = 0,5767$. Таким образом, вероятность того, что страховая компания в течение года потерпит убытки, составляет 0, 5767.

1.11.1 Потоки событий. Пуассоновский поток

Поток событий — последовательность событий, которые наступают в случайные моменты времени.

Пример. Поток посетителей в супермаркете. Покупатель зашёл — событие произошло.

Поток событий называется *пуассоновским (простейшим)*, если выполнены 3 свойства:

1. *Стационарность*: вероятность того, что произошло k событий за время t зависит только от k и t и не завсисит от начала отсчёта времени.
2. *Отсутствие последействия*: вероятность появления k событий за время t не зависит от предыстории процесса.
3. *Ординарность*: вероятность наступления двух и более событий за малый промежуток времени Δt близка к нулю.

Тогда вероятность наступления k событий за промежуток времени t равна

$$P_t(k) = \frac{e^{-\lambda t}(\lambda t)^k}{k!},$$

где λ — интенсивность потока (среднее число событий, которое происходит в единицу времени).

Задача

Среднее число заказов такси за 1 минуту равно 3. Найти вероятность того, что за 2 минуты поступит 4 вызова. Поток считать простейшим.

Решение

Имеем $\lambda = 3$. Тогда
$$P_2(4) = \frac{e^{-3\cdot 2}(3\cdot 2)^4}{4!} \approx 0,135.$$

Глава 2

Случайные величины

2.1 Случайные величины

Случайная величина — числовая измеримая функция $\xi : \Omega \mapsto R^1$.

Функция называется *измеримой*, если прообраз борелевского множества измерим, то есть $\xi^{(-1)}(B) \in \mathfrak{F}$.

Если Ω конечно, то *случайная величина* — любая числовая функция $\xi : \Omega \mapsto R^1$.

Неформальное определение: случайной величиной называется функция, которая в результате испытания примет одно и только одно из (известных) возможных значений, но какое — наперед неизвестно.

Случайные величины:

- *Дискретные случайные величины* — случайные величины, множество возможных значений которых не более чем счётно (имеют конечный либо счетный набор отдельных изолированных значений).

  Пример 1. Число родившихся мальчиков (или девочек) среди ста новорожденных.

  Пример 2. Число студентов, пришедших на пару, из списка группы.

- *Непрерывные величины* — определим позже (см. раздел Функция распределения). Неформально: непрерывные случайные величины характеризуются тем, что множество возможных значений представляет собой не набор отдельных изолированных значений, а некий интервал, отрезок, полуинтервал или их счетное объединение.

Пример 3. Расстояние, которое пролетит снаряд, при стрельбе на дальность.

- *Сингулярные (смешанные) величины*

Проходить не будем.

2.2 Дискретные случайные величины

Закон распределения дискретной случайной величины X — перечень всех её возможных значений $\{x_k\}$ и соответствующих им вероятностей $\{p_k\}$. Заметим, что $\sum_k p_k = 1$.

Две формы задания закона распределения дискретной случайной величины: табличная и при помощи формулы.

X	x_1	x_2	$\dots$	x_k	$\dots$	x_n	$\dots$
$\sum_k p_k = 1$	p_1	p_2	$\dots$	p_k	$\dots$	p_n	$\dots$

$$p_k = P\{X = x_k\}, \qquad k = 1, 2, \dots$$

Графическое представление закона распределения дискретной случайной величины называется *многоугольником распределения.*

Пример 1. Пусть имеется 1000 лотерейных билетов. Из них определяются

- 1 выигрышный на 10 000 рублей,
- 2 выигрышных на 5 000 рублей,
- 4 выигрышных на 500 рублей,
- 10 выигрышных на 100 рублей,

Случайная величина Y — величина выигрыша для владельца лотерейного билета.

Пример 2. По данной американской статистики 25-летний человек достигает 26-летия без серьезного ущерба для здоровья с вероятностью 0.992. Предположим, что 25-летний

Y	10000	5000	500	100	0
	$0,001$	$0,002$	$0,004$	$0,01$	$1-(0,001+0,002+0,004+0,01)=0,983$

молодой человек страхует свою жизнь на год, страховой взнос составляет 10 долларов. При наступлении страхового случая страховая фирма выплачивает страховую выплату в размере 1000 долларов.

Пусть случайная величина X — величина прибыли с одного клиента страховой фирмы.

X	10	-990
	$0,992$	$0,008$

2.3 Основные распределения дискретных случайных величин

1. *Биномиальное.*

   Пусть S_n — число успехов в схеме Бернулли, включающей n независимых испытаний, в каждом из которых вероятность успеха равна p.

   S_n принимает целочисленные значения $0,\dots,n$.

   Закон распределения биномиальной случайной величины:

$$p_k = P\{S_n = k\} = P_n(k) = C_n^k p^k q^{n-k}, \qquad k=\overline{0,n},\ \ q = 1-p.$$

S_n	0	1	2	$\dots$	k	$\dots$	$n-1$	n
$\sum_{k=0}^n p_k = 1$	q^n	npq^{n-1}	$\frac{n(n-1)}{2}p^2q^{n-2}$	$\dots$	$C_n^k p^k q^{n-k}$	$\dots$	$np^{n-1}q$	p^n

Короткая запись:

$$S_n \sim B(n;p).$$

Сумма вероятностей равна 1 из бинома Ньютона:

$$1 = 1^n = (p+q)^n = \sum_{k=0}^{n} C_n^k p^k q^{n-k}.$$

Пример. В семье пятеро детей. Считая вероятность рождения мальчика равной 0.514 записать закон распределения для случайной величины S_n — числа мальчиков в семье.

S_n	0	1	2	3	4	5
	$0,486^5$	$5 \cdot 0,512 \cdot 0,486^4$	$10 \cdot 0,512^2 \cdot 0,486^3$	$10 \cdot 0,512^3 \cdot 0,486^2$	$5 \cdot 0,512^4 \cdot 0,486$	$0,512^5$

2. *Распределение Пуассона (закон редких явлений).*

$$p_k = p\{X = k\} = \frac{e^{-\lambda} \cdot \lambda^k}{k!}$$

$$\sum_{k=0}^{\infty} p_k = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} \cdot e^{\lambda} = 1$$

Короткая запись:

$$X \sim P(\lambda)$$

(или $X \sim \Pi(\lambda)$).

Смысл параметра λ будет ясен чуть позже.

3. *Геометрическое.*

Пусть вероятность наступления события A одинакова в каждом опыте и равна p. Опыты производят до первого наступления события A. Пусть X — число проведенных опытов (до первого появления события A).

Тогда

$$p_k = P\{X = k\} = q^{k-1} \cdot p, \qquad k = 1, 2, \ldots$$

$$\sum_{k=1}^{\infty} p_k = p + p \cdot q + p \cdot q^2 + \ldots = \frac{p}{1-q} = 1$$

Короткая запись:

$$X \sim G(p).$$

Пример. Игральный кубик подбрасывается до первого появления 6 очков. Пусть X — число проведенных испытаний до первого появления 6 очков.

Имеем:

$$p = \frac{1}{6}, \qquad q = \frac{5}{6}.$$

X	1	2	3	4	...
	$\frac{1}{6}$	$\frac{5}{6}\frac{1}{6}$	$(\frac{5}{6})^2\frac{1}{6}$	$(\frac{5}{6})^3\frac{1}{6}$	...

4. *Гипергеометрическое.*

Пусть имется N деталей, M из них — стандартные, остальные — бракованные. Наудачу выбираем n деталей из N. Найдем вероятность того, что среди n деталей будет ровно k стандартных.

Случайная величина X — число стандартных деталей среди n отобранных. Эта величина принимает целочисленные значения $0, \ldots, min\{M, n\}$

Закон распределения с.в. X:

$$p_k = P\{X = k\} = \frac{C_M^k \cdot C_{N-M}^{n-k}}{C_N^n}.$$

Короткая запись:

$$X \sim HG(N, M, n).$$

2.4 Числовые характеристики дискретных случайных величин

2.4.1 Математическое ожидание

Пусть дана дискретная случайная величина X с известным законом распределения

Математическое ожидание — среднее значение случайной величины X:

X	x_1	x_2	$\ldots$	x_k	$\ldots$	x_n	$\ldots$
	p_1	p_2	$\ldots$	p_k	$\ldots$	p_n	$\ldots$

$$M(X) = \sum_i x_i \cdot p_i.$$

Иногда вместо $M(X)$ пишут $E(X)$ (от английского “expectation’)..

Пример 1. В задаче с данными о страховых выплатах 25-летнему молодому человеку имеем:

X	10	-990
	$0,992$	$0,008$

Тогда средняя прибыль страховой фирмы с одного клиента равна

$$M(X) = 10 \cdot 0,992 + (-990) \cdot 0,008 = 2 > 0.$$

Пример 2. Игра ”Однорукий бандит”. Автомат имеет два окошка, в каждом из которых независимо друг от друга показывается одна из трёх картинок: колокольчик, вишенка, яблоко. Вероятность того, что в окошке в данный момент показывается колокольчик равна 0.5, вишня — 0.4, яблоко — 0.1.

Участие в игре стоит 5 центов. При выпадении двух одинаковых картинок игрок выигрывает определённую сумму.

- “Колокольчик, колокольчик” $\to$ выигрыш 5 центов;
- “Вишня, вишня” $\to$ выигрыш 10 центов;
- “Яблоко, яблоко” $\to$ выигрыш 50 центов.

Случайная величина Y — чистый выигрыш игрока (с учетом платы за участие в игре). Тогда

Тогда

$$M(Y) = 0 \cdot 0,25 + 5 \cdot 0,16 + 45 \cdot 0,01 + (-5) \cdot 0,58 = -1,65 < 0.$$

Y	0	5	45	-5
	$0.5 \cdot 0.5 = 0.25$	$0.4 \cdot 0.4 = 0.16$	$0.1 \cdot 0.1 = 0.01$	$1 - (0,25 + 0,16 + 0,01) = 0,58$

2.4.2 Свойства математического ожидания

Для дискретных с.в. несложно доказать выполнение следующих свойств. Отметим, что те же свойства справедливы и для абсолютно непрерывных случайных величин.

1. $M(const) = const$
2. $M(const \cdot X) = const \cdot M(X)$
3. $M(X+Y) = M(X) + M(Y)$
4. Если X, Y — независимые случайные величины $\Rightarrow M(X \cdot Y) = M(X) \cdot M(Y)$

Санкт-Петербургский парадокс (игра). Игрок играет против казино. Подбрасывается монета. Если при i-м бросании монета упала гербом, то игрок получает 2^i долларов. Вопрос: какой размер вступительного взноса делает такую игру справедливой? Рассчитаем математическое ожидание выигрыша игрока (равного проигрышу казино):

$$M(Y) = 2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{4} + \ldots + 2^k \cdot \frac{1}{2^k} + \ldots = \infty$$

Парадокс: для входа в игру оптимальная сумма — бесконечность.

2.4.3 Математическое ожидание для основных дискретных случайных величин

1. Биномиальное распределение.
   $M(S_n) = ?,\quad S_n = \sum_{i=1}^n I_i \Rightarrow M(S_n) = M(\sum_{i=1}^n I_i) = \sum_{i=1}^n M(I_i) = np$

2. Распределение Пуассона.
   $p_k = \frac{e^{-\lambda} \cdot \lambda^k}{k!},\quad k = 0, 1, 2, \ldots$

$$M(X) = \sum_{k=0}^{\infty} x_k \cdot p_k = \sum_{k=0}^{\infty} k \cdot p_k = \sum_{k=0}^{\infty} k \cdot \frac{e^{-\lambda} \cdot \lambda^k}{k!} = \lambda \cdot \sum_{m=0}^{\infty} \frac{e^{-\lambda} \cdot \lambda^m}{m!} = \lambda$$

3. Геометрическое распределение.

   $X \sim G(p)\ \ p_k = p\{x = k\} = q^{k-1}p, \ \ k = 1, 2, \dots$

$$M(X) = \sum_k x_k p_k = \sum_{k=1}^{\infty} k p_k = p \sum_{k=1}^{\infty} k q^{k-1} = p \sum_{i=1}^{\infty} \frac{dq^k}{dq} = p \frac{d}{dq} \sum_{i=1}^{\infty} q^k = p \frac{d}{dq}\left(\frac{q}{1-q}\right) =$$

$$= p \frac{1}{(1-q)^2} = \frac{1}{p}$$

4. Гипергеометрическое распределение.

$$M(X) = M \cdot \frac{n}{N}$$

2.4.4 Дисперсия случайной величины

Дисперсия характеризует разброс значений случайной величины вокруг её математического ожидания.

Пусть имеются две случайные величины:

Y_1	$-0,001$	$0,001$
	0.5	0.5

Y_2	-1000	1000
	0.5	0.5

Очевидно, что

$$M(Y_1) = M(Y_2) = 0,$$

однако значения случайной величины совершенно по-разному разбросаны вокруг 0.

Дисперсией случайной величины X называется

$$D(X) = M\Big((X - M(X))^2\Big). \tag{2.1}$$

Дисперсия с.в. иногда также обозначается как $V(X)$, $Var(X)$ (от английского “variance”).

Справедлива также вторая формула для вычисления дисперсии:

$$D(X) = M(X^2) - (M(X))^2. \tag{2.2}$$

Доказательство. $D(X) = M(X - M(X))^2 = M(X^2) + M(-2 \cdot XM(X)) + M^2(X) = M(X^2) - M^2(X)$ □

Вычисление дисперсии для дискретных случайных величин.

Пусть X — дискретная случайная величина:

X	x_1	x_2	$\dots$	x_n
	p_1	p_2	$\dots$	p_n

Рассмотрим случайную величину $Y = (X - M(X))^2$. Тогда она имеет следующий закон распределения:

$Y = (X - M(X))^2$	$(x_1 - M(X))^2$	$(x_2 - M(X))^2$	$\dots$	$(x_n - M(X))^2$
	p_1	p_2	$\dots$	p_n

$$D(X) = M(Y) = \sum_i y_i \cdot p_i = \sum_i (x_i - M(X))^2 \cdot p_i = \sum_i x_i^2 \cdot p_i - \left(\sum_i x_i \cdot p_i \right)^2$$

Свойства дисперсии:

1. $D(X) \geq 0$
2. $D(C) = 0$
3. $D(C \cdot X) = C^2 \cdot D(X)$
4. Если X, Y — независимые случайные величины, тогда $D(X \pm Y) = D(X) + D(Y)$
5. $D(X + Y) = D(X) + D(Y) + 2cov(X, Y),\ cov(X, Y) = M[(X - M(X))(Y - M(Y))]$

Пример. Имеются две игральные кости. X — сумма очков на двух костях, Y — произведение.

$X = X_1 + X_2$. Вероятность выпадения числа очков i равна $\frac{1}{6}$, $i = 1, \dots, 6$.
$M(X) = M(X_1) + M(X_2) = \frac{7}{2} \cdot 2 = 7$.

$Y = X_1 \cdot X_2 \Rightarrow M(Y) = M(X_1) \cdot M(X_2) = \frac{7}{2} \cdot \frac{7}{2} = 12.25.$

2.4.5 Среднее квадратическое отклонение

Имеет смысл ввести характеристику рассеивания значений случайной величины, которая имеет те же единицы измерения, что и сама случайная величина. Например, пусть X — рост. Тогда

X — [м], $M(X)$ — [м], $D(X)$ — [м^2] (здесь [м] — метры).

Среднее квадратическое отклонение (с.к.о.)

$$\sigma(X) = \sqrt{D(X)}$$

Очевидно, что $\sigma(X)$ имеет те же единицы измерения, что и X.

2.4.6 Начальные и центральные моменты случайных величин

Начальный момент порядка k:

$$\mu_k = M(X^k).$$

Центральный момент порядка k:

$$\nu_k = M((X - M(X))^k).$$

2.4.7 Дисперсия для основных распределений дискретных случайных величин

1. Биномиальное распределение.

   S_n - число успехов. $S_n = I_1 + \ldots + I_n$, I_i – индикаторная случайная величина:

   $$DS_n = D\left(\sum_i I_i\right); \qquad D(I_i) = \sum (x_i - M(X))^2 \cdot p_i = (1-p)^2 \cdot p + (0-p)^2 \cdot q = pq$$

   $$DS_n = \sum_{i=1}^{n} D(I_i) = \sum_{i=1}^{n} p \cdot q = npq$$

2. Распределение Пуассона.

   $p_k = \frac{e^{-\lambda} \cdot \lambda^k}{k!}, k = 0, 1, \ldots,\ M(X) = \lambda,$

$$D(X) = \sum_{i=0}^{\infty} x_i^2 p_i - (M(X))^2 = \sum_{k=0}^{\infty} k^2 \frac{e^{-\lambda}\lambda^k}{k!} - \lambda^2 = \sum_{k=0}^{\infty} \frac{e^{-\lambda} k k \lambda^k}{k!} - \lambda^2 =$$

$$= \sum_{k=0}^{\infty} \frac{e^{-\lambda}(k-1+1)\lambda\lambda^{k-1}}{(k-1)!} - \lambda^2 = \lambda \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^{k-1}}{(k-1)!} + \lambda^2 \sum_{k=0}^{\infty} \frac{e^{-\lambda}\lambda^{k-2}}{(k-2)!} - \lambda^2 = \lambda + \lambda^2 - \lambda^2 = \lambda$$

3. Геометрическое распределение.

$$D(X) = \frac{q}{p^2}$$

4. Гипергеометрическое распределение.

$$D(X) = M\frac{n}{N} \cdot \left(1 - \frac{n}{N}\right) \cdot \frac{N-M}{N-1}$$