Дискретная теория вероятностей.

Лектор — Юрий Александрович Давыдов
Создатель конспекта — Глеб Минаев *

TODOs

Содержание

*Оригинал конспекта расположен на GitHub. Также на GitHub доступен репозиторий с другими конспектами.

Литература:

- А.Н. Ширяев, “Вероятность”.
- М.А. Лифшиц, “Лекции”
- Ю.Г. Борисович, Н.М. Близняков, Я.А. Израилевич, Т.Н. Фоменко, “Введение в топологию”, М.:Наука. Физматлит, 1995.
- James Munkres, Topology.
- Биография Сергея Натановича Бернштейна.

1 Вероятностные пространства и стандартные следствия

1.1 Вероятностное пространство

Определение 1. *Вероятностное пространство* — это тройка $(\Omega, \mathcal{F}, \mathbb{P})$, где

- $\Omega \neq \varnothing$ — множество объектов случайной природы, называемых *элементарными событиями (исходами)*,
- $\mathcal{F}$ — сигма-алгебра над множеством Ω (т.е. такое подмножество $\mathcal{P}(\Omega)$, что
  1. $\mathcal{F}$ содержит Ω,
  2. для всякого $A \in \mathcal{F}$ множество $\Omega \setminus A$ содержится в $\mathcal{F}$,
  3. для всякого не более чем счётного семейства $\{A_i\}_{i \in I}$ множеств из $\mathcal{F}$ множества
  $$\bigcup_{i \in I} A_i \qquad \text{и} \qquad \bigcap_{i \in I} A_i$$
  содержатся в $\mathcal{F}$),

  которая называется множеством *(случайных) событий*,
- $\mathbb{P}$ — счётно-аддитивная мера, что $\mathbb{P}(\Omega) = 1$ (т.е. функция из $\mathcal{F}$ в $[0; 1]$, что
  1. $\mathbb{P}(\varnothing) = 0$, $\mathbb{P}(\Omega) = 1$,
  2. для всякого не более чем счётного семейства $\{A_i\}_{i \in I}$ дизъюнктных множеств из $\mathcal{F}$
  $$\mathbb{P}\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} \mathbb{P}(A_i);$$

  значение $\mathbb{P}(A)$ называется *вероятностью события* A).

Пример 1. Пусть мы бросаем монетку (один раз).

1. Тогда множество исходов будет состоять из элементарных событий “выпал орёл” и “выпала решка”:
$$\Omega := \{\text{Орёл}; \text{Решка}\}$$

2. Множество событий будет состоять из событий:

   (a) $\varnothing$ — ничего не выпало, т.е. ничего не произошло,

   (b) $\{\text{Орёл}\}$ — выпал орёл,

   (c) $\{\text{Решка}\}$ — выпала решка,

   (d) $\{\text{Орёл}; \text{Решка}\}$ — выпал орёл или решка, т.е. что-то произошло.

   Т.е. в данном случае $\mathcal{F} = \mathcal{P}(\Omega)$.

3. Понятно, что
$$\mathbb{P}(\varnothing) = 0, \quad \text{а} \quad \mathbb{P}(\{\text{Орёл}; \text{Решка}\}) = 1.$$
   При этом для всякой величины $p \in [0; 1]$ может быть, что
$$\mathbb{P}(\{\text{Орёл}\}) = p, \quad \text{а} \quad \mathbb{P}(\{\text{Решка}\}) = 1 - p.$$
   В случае $p = \frac{1}{2}$ монетку называют *симметричной* (иначе *несимметричной*).

Замечание (“стабилизация частот”). Пусть мы проводим один и тот же эксперимент n раз и смотрим, сколько раз реализовалось событие A. Обозначим это количество реализаций за $\nu_n(A)$. При этом если эксперимент брать “реальным”, например, взятым из физики (подбрасывание монеты, бросание игрального кубика, etc.), то можно заметить следующие явления.

1. (эмпирический факт) Для всякого события A имеет место сходимость
$$\lim_{n\to\infty} \frac{\nu_n(A)}{n} = P(A).$$
   Это и хочется назвать вероятностью.

2. Если $A = \Omega$, то $\nu_n(A)$ должно ровняться n, а значит
$$P(A) = \lim_{n\to\infty} \frac{\nu_n(A)}{n} = 1.$$
   А если $A = \varnothing$, то $P(A) = 0$.

3. Для всякого события A верно, что $\nu_n(A) \in [0; n]$. Следовательно $P(A) \in [0; 1]$.

4. Если события A и B дизъюнктны, то $\nu_n(A) + \nu_n(B) = \nu_n(A \cup B)$. Отсюда следует аддитивность вероятности; и по аналогии получается счётная аддитивность.

Лемма 1.

1. *Для всяких событий A и B*
$$A \subseteq B \quad \Longrightarrow \quad \mathbb{P}(A) \leqslant \mathbb{P}(B).$$

2. *Для всякого события A*

$$\mathbb{P}(A) + \mathbb{P}(A^C) = 1,$$

где $A^C := \Omega \setminus A$.

3. *Для всяких событий A и B*

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

4. *Для всякого не более чем счётного семейства событий $\{A_i\}_{i\in I}$*

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) \leqslant \sum_{i\in I} \mathbb{P}(A_i),$$

и равенство достигается только если $\mathbb{P}(A_i \cap A_j) = 0$ для всех $i \neq j$.

5. *Для всякой последовательности вложенных событий $\{A_n\}_{n=0}^{\infty}$ ($A_{n+1} \supseteq A_n$)*

$$\mathbb{P}\left(\bigcap_{n=0}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

6. *Для всякой последовательности вложенных событий $\{A_n\}_{n=0}^{\infty}$ ($A_{n+1} \subseteq A_n$)*

$$\mathbb{P}\left(\bigcap_{n=0}^{\infty} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

Определение 2. Вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$ называется *дискретным*, если Ω не более чем счётно, а $\mathcal{F} = \mathcal{P}(\Omega)$.

Теорема 2. *Пусть даны не более чем счётное Ω и $\mathcal{F} = \mathcal{P}(\Omega)$.*

1. *Пусть $(\Omega, \mathcal{F}, \mathbb{P})$ — дискретное вероятностное пространство. Для всякого $\omega \in \Omega$ можно обозначить*

$$p_\omega := \mathbb{P}(\omega).$$

Тогда

(a) *каждое $p_\omega \geqslant 0$,*

(b)

$$\sum_{\omega\in\Omega} p_\omega = 1.$$

И при этом $\mathbb{P}$ можно задать условием

$$\mathbb{P}(A) = \sum_{\omega\in A} p_\omega$$

2. *Пусть для всякого $\omega \in \Omega$ определено вещественное p_ω, что*

(a) *каждое $p_\omega \geqslant 0$,*

(b)

$$\sum_{\omega\in\Omega} p_\omega = 1.$$

Тогда можно задать функцию $\mathbb{P} : \mathcal{F} \to [0;1]$ *условием*

$$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega,$$

и тогда $(\Omega, \mathcal{F}, \mathbb{P})$ *будет дискретным вероятностным пространством.*

Доказательство.

1. Действительно:

   (a) Каждое
   $$p_\omega = \mathbb{P}(\omega) \geqslant 0.$$

   (b) Поскольку $\Omega = \bigsqcup_{\omega \in \Omega} \{\omega\}$, то
   $$\sum_{\omega \in \Omega} p_\omega = \sum_{\omega \in \Omega} \mathbb{P}(\omega) = \mathbb{P}(\Omega) = 1.$$

   И аналогично $A = \bigsqcup_{\omega \in A} \{\omega\}$, а значит
   $$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) = \sum_{\omega \in A} p_\omega.$$

2. Действительно, если $A = \bigsqcup_{i \in I} A_i$ (I не более чем счётно), то (так как каждое A_i не более чем счётно)
   $$\mathbb{P}(A) = \sum_{\omega \in A} p_\omega = \sum_{i \in I} \sum_{\omega \in A_i} p_\omega = \sum_{i \in I} \mathbb{P}(A_i)$$
   (так как мы рассуждаем в рамках абсолютно сходящегося ряда). Ну и, конечно,
   $$\mathbb{P}(\varnothing) = \sum_{\omega \in \varnothing} p_\omega = 0 \qquad \text{и} \qquad \mathbb{P}(\Omega) = \sum_{\omega \in \Omega} p_\omega = 1$$

$\square$

Определение 3. $(\Omega, \mathcal{F}, \mathbb{P})$ называется *пространством классического типа*, если Ω кончено, $\mathcal{F} = \mathcal{P}(\Omega)$ и для всякого $\omega \in \Omega$
$$\mathbb{P}(\omega) = \frac{1}{|\Omega|}.$$

Замечание 1. В классическом пространстве соответственно имеем, что
$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}$$

1.2 Условная вероятность

Определение 4. *Вероятность события A при условии события B* (где $\mathbb{P}(B) \neq 0$) есть
$$\mathbb{P}_B(A) = \mathbb{P}(A \mid B) := \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Теорема 3. *Пусть даны вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$ и $B \in \mathcal{F}$, что $\mathbb{P}(B) \neq 0$. Тогда тройки $(\Omega, \mathcal{F}, P)$ и $(B, \mathcal{F}_B, P_B)$, где*

$$\mathcal{F}_B := \{S \in \mathcal{F} \mid S \subseteq B\},$$

а

$$P : \mathcal{F} \to [0; 1], A \mapsto \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

и

$$P : \mathcal{F}_B \to [0; 1], A \mapsto \frac{\mathbb{P}(A)}{\mathbb{P}(B)},$$

являются вероятностными пространствами.

Доказательство. Понятно, что

$$P(\varnothing) = 0 \qquad \text{и} \qquad P(\Omega) = 1.$$

Также если $A = \bigsqcup_{i \in I} A_i$ (где I не более чем счётно), то

$$\left(\bigcup_{i \in I} A_i\right) \cap B = \bigcup_{i \in I} A_i \cap B \quad \Longrightarrow \quad \mathbb{P}\left(\left(\bigcup_{i \in I} A_i\right) \cap B\right) = \sum_{i \in I} \mathbb{P}(A_i \cap B).$$

Значит первая тройка является вероятностным пространством.

Заметим, что отношение $\sim$ на $\mathcal{F}$, заданное условием

$$S \sim T \quad \Longleftrightarrow \quad S \cap B = T \cap B$$

определяет классы эквивалентности, минимальные по включению предстаители которых (представителем $[A]$ будет $A \cap B$), образуют $\mathcal{F}_B$. При этом для всяких S и T из $S \sim T$ следует, что

$$\mathbb{P}(S \cap B) = \mathbb{P}(T \cap B).$$

Также несложно понять, что $\mathcal{F}_B$ будет сигма-алгеброй. Значит P_B сужением P на $\mathcal{F}$ с тем же множеством значений. Таким образом вторая тройка тоже будет вероятностным пространством. □

Лемма 4 (формула полной вероятности)**.** *Пусть дано не более чем счётное разбиение $\{B_i\}_{i \in I}$ множества Ω на множества из $\mathcal{F}$. Тогда для всякого события A*

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(B_i)\mathbb{P}(A \mid B_i).$$

Доказательство. Поскольку $\{B_i \cap A\}_{i \in I}$ есть разбиение A, значит

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \cap B_i) = \sum_{i \in I} \mathbb{P}(B_i)\mathbb{P}(A \mid B_i)$$

□

Лемма 5 (формула Байеса)**.** *Для всяких событий A и B*

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A)\mathbb{P}(B \mid A)}{\mathbb{P}(B)}$$

Доказательство.

$$\mathbb{P}(A \mid B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A)\mathbb{P}(B \mid A)}{\mathbb{P}(B)}.$$

$\square$

Следствие 5.1. *Пусть дано не более чем счётное разбиение $\{B_i\}_{i\in I}$ множества Ω на множества из $\mathcal{F}$. Тогда для всякого события A и индекса $j \in I$*

$$\mathbb{P}(B_j \mid A) = \frac{\mathbb{P}(B_j) \mid (A \mid B_j)}{\sum_{i\in I} \mathbb{P}(B_i)\mathbb{P}(A \mid B_i)}.$$

Лемма 6 (формула умножения). *Для всяких событий $\{A_k\}_{i=1}^{n}$. Тогда*

$$\mathbb{P}\left(\bigcap_{k=1}^{n} A_i\right) = \prod_{k=1}^{n} \mathbb{P}\left(A_k \mid \bigcap_{i=1}^{k-1} A_i\right) = \mathbb{P}(A_1) \cdot \mathbb{P}(A_2 \mid A_1) \cdot \dots \cdot \mathbb{P}\left(A_n \mid \bigcap_{i=1}^{n-1} A_i\right)$$

1.3 Независимые события

Определение 5. События A и B называются *независимыми*, если

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$$

Лемма 7. *Для любых двух событий A и B TFAE*

1. *A и B независимы и их вероятности > 0,*
2. *$\mathbb{P}(A) = \mathbb{P}(A \mid B)$ (а $\mathbb{P}(B) = \mathbb{P}(B \mid A)$).*

Определение 6. Семейство событий $\{A_i\}_{i\in I}$ называется *независимым* (или также *"независимым в совокупности"* или *"совместно независимым"*), если для всякого конечного $S \subseteq I$ верно равенство

$$\mathbb{P}\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} \mathbb{P}(A_i).$$

Замечание 2. Независимость (в совокупности) есть частный случай попарной независимости (что понятно, из определения), но не является равносильным ему свойством.

Пример 2 (пирамида Бернштейна). Рассмотрим пространство классического типа с $\Omega = \{1; 2; 3; 4\}$. Пусть

$$A_1 := \{1; 4\}, \qquad A_2 := \{2; 4\} \qquad \text{и} \qquad A_3 := \{3; 4\}.$$

Тогда

$$\mathbb{P}(A_i) = \frac{1}{2}, \qquad \mathbb{P}(A_i \cap A_j) = \frac{1}{4} = \mathbb{P}(A_i)\cdot\mathbb{P}(A_j) \qquad \text{и} \qquad \mathbb{P}(A_1 \cap A_2 \cap A_3) = \frac{1}{4} = \mathbb{P}(A_1)\cdot\mathbb{P}(A_2)\cdot\mathbb{P}(A_3).$$

Отсюда, например, следует, что

$$\mathbb{P}(A_3 \mid A_1 \cap A_2) \neq \mathbb{P}(A_3)$$

Теорема 8. *Пусть даны некоторые натуральные $\{m_i\}_{i=1}^{n}$ и семейство независимых событий*

$$\{A_{i,j}\}_{\substack{i\in\{1;\dots;n\} \\ j\in\{1;\dots;m_i\}}}.$$

Тогда семейство событий

$$\left\{\bigcup_{j=1}^{m_i} A_i\right\}_{i=1}^{n}$$

независимо.

Доказательство.

Лемма 9. *Пусть дано семейство независимых событий* $\{A_i\}_{i\in I} \cup \{B; C\}$. *Тогда семейства*

$$\{A_i\}_{i\in I} \cup \{B \cap C\} \qquad \text{и} \qquad \{A_i\}_{i\in I} \cup \{B \cup C\}$$

являются независимыми (сами для себя, а не друг для друга).

Доказательство.

1. Покажем для пересечения. Пусть $S \subseteq I$ — конечное подмножество. Тогда

$$\mathbb{P}\left((B\cap C) \cap \bigcap_{i\in S} A_i\right) = \mathbb{P}(B)\cdot\mathbb{P}(C)\cdot\prod_{i\in S}\mathbb{P}(A_i) = \mathbb{P}(B\cap C)\cdot\prod_{i\in S}\mathbb{P}(A_i).$$

   Значит

$$\{A_i\}_{i\in I} \cup \{B \cap C\},$$

   действительно, независим.

2. Покажем для объединения. Пусть $S \subseteq I$ — конечное подмножество. Тогда

$$\begin{aligned}
\mathbb{P}\left((B\cup C) \cap \bigcap_{i\in S} A_i\right) &= \mathbb{P}\left(B \cap \bigcap_{i\in S} A_i\right) + \mathbb{P}\left(C \cap \bigcap_{i\in S} A_i\right) - \mathbb{P}\left((B\cap C) \cap \bigcap_{i\in S} A_i\right)\\
&= \mathbb{P}(B)\cdot\prod_{i\in S}\mathbb{P}(A_i) + \mathbb{P}(C)\cdot\prod_{i\in S}\mathbb{P}(A_i) - \mathbb{P}(B)\cdot\mathbb{P}(C)\cdot\prod_{i\in S}\mathbb{P}(A_i)\\
&= (\mathbb{P}(B)+\mathbb{P}(C)-\mathbb{P}(B)\cdot\mathbb{P}(C))\cdot\prod_{i\in S}\mathbb{P}(A_i)\\
&= \mathbb{P}(B\cup C)\cdot\prod_{i\in S}\mathbb{P}(A_i).
\end{aligned}$$

   Значит

$$\{A_i\}_{i\in I} \cup \{B \cup C\},$$

   действительно, независим.

□

Несложно понять, что операциями из леммы выше из семейства

$$\{A_{i,j}\}_{\substack{i\in\{1;\dots;n\}\\ j\in\{1;\dots;m_i\}}}$$

можно получить семейство

$$\left\{\bigcup_{j=1}^{m_i} A_i\right\}_{i=1}^{n},$$

и при этом семейство будет оставаться независимым после каждой операции. Следовательно конечное семейство будет независимым. □

1.4 Случайные величины

Определение 7. *Случайная величина* X в вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ на измеримое пространство $(R, \mathcal{G})$ — $\mathcal{F}/\mathcal{G}$-измеримая функция из Ω в R.

Определение 8. Случайная величина X называется *дискртеной*, если множество её значений $X(\Omega)$ не более чем счётно.

Замечание. Мы будем рассматривать в качестве R множество $\mathbb{R}$, а в качестве $\mathcal{G}$ — $\mathcal{P}(\mathbb{R})$. При этом поскольку мы рассматриваем дискретные вероятностные пространства, то всякое отображение из Ω в $\mathbb{R}$ будет измеримым, а всякая случайная величина (на дискретном пространстве) будет дискретной.

Или же всё-таки $\mathcal{G}$ будет борелевской сигма-алгеброй?

Определение 9. *Распределение случанйной величины* X в вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ на измеримое пространство $(R, \mathcal{G})$ — функция

$$\mathbb{P}_X : \mathcal{G} \to [0; 1], S \mapsto \mathbb{P}(X^{-1}(S)).$$

Замечание 3. $(R, \mathcal{G}, \mathbb{P}_X)$ — вероятностное пространство.

Замечание 4. Для дискретной случайной величины можно считать, что $X(\Omega) = \{a_i\}_{i \in I}$, где I не более чем счётно. Значит можно определить

$$A_i := \{\omega \in \Omega \mid X(\omega) = a_i\} \qquad \text{и} \qquad p_i := \mathbb{P}(A_i).$$

Тогда

1. $p_i \geqslant 0$,
2. $\sum_{i \in I} p_i = 1$.

Поэтому в качестве распределения случайной величины можно также рассматривать сужение $\mathbb{P}_X$ на множество синглтонов $\{\{a_i\}\}_{i \in I}$.

Определение 10. Пусть X и Y — случайные в (возможно, разных) вероятностных пространствах на измеримое пространство $(R, \mathcal{G})$. Тогда говорят, что X *имеет распределение* Y или Y *имеет распределение* X, и пишут $X \overset{\mathcal{D}}{=} Y$, если их функции распределения совпадают.

Пример 3.

1. **Вырожденное распределение.** Пусть $X(\Omega) = a$. Тогда распределение X будет состоять только из сопоставления
$$a \mapsto 1.$$
2. **Распределение Бернулли:** $B(1, p)$**.** Для всякой величины $p \in [0; 1]$ можно рассмотреть случайную величину $X \overset{\mathcal{D}}{=} B(1, p)$ с распределением
$$X = \begin{cases} 0 & \text{с вероятностью } 1 - p, \\ 1 & \text{с вероятностью } p. \end{cases}$$
3. **Биномиальное распределение:** $B(n, p)$**.** Для всякой величины $p \in [0; 1]$ можно рассмотреть случайную величину $X \overset{\mathcal{D}}{=} B(n, p)$ с множеством значений $\{0; \ldots; n\}$ и распределением
$$\mathbb{P}\{X = k\} = \binom{n}{k} p^k (1 - p)^{n-k}.$$

4. **Геометрическое распределение.** Для всякой величины $p \in [0;1]$ можно рассмотреть случайную величину X с множеством значений $\mathbb{N} \cup \{0\}$ и распределением
$$\mathbb{P}\{X = k\} = p^k(1-p).$$

5. **Распределение Пуассона: $\mathcal{P}(\alpha)$.** Для всякой величины $\alpha > 0$ можно рассмотреть случайную величину $X \overset{\mathcal{D}}{=} \mathcal{P}(\alpha)$ с множеством значений $\mathbb{N} \cup \{0\}$ и распределением
$$\mathbb{P}\{X = k\} = \frac{\alpha^k}{k!}e^{-\alpha}.$$

Определение 11. Семейство случайных величин $\{X_i\}_{i \in I}$ в вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ на измеримые пространства $\{(R_i, \mathcal{G}_i)\}_{i \in I}$ называются *независимым*, если для всякого семейства $\{S_i\}_{i \in I}$, что $S_i \in \mathcal{G}_i$, семейство событий
$$\{X_i^{-1}(S_i)\}_{i \in I}$$
независимо.

Говоря проще, распределение вероятностей всякого X_i не зависит от конечного количества условий на другие случайные величины.

Теорема 10. *Пусть дано семейство случайных величин $\{X_i\}_{i=1}^n$. Величина X_i имеет распределение $\{a_{i,j} \mapsto p_{i,j}\}_{j \in I_i}$. Тогда семейство $\{X_i\}_{i=1}^n$ независимо тогда и только тогда, когда для всяких $\{j_i\}_{i=1}^n$, что $j_i \in I_i$, верно, что*
$$\mathbb{P}\left\{\bigwedge_{i=1}^n X_i = a_{i,j_i}\right\} = \prod_{i=1}^n p_{i,j_i}.$$

Доказательство. Понятно, что равенство в условии равносильно части условия независимости
$$\{X_i^{-1}(\{a_{i,j_i}\})\}_{i=1}^n,$$
где выбираемое множество индексов $S = \{1; \dots; n\}$. Таким образом из независимости $\{X_i\}_{i=1}^n$, очевидно, следует предыдущее утверждение; покажем теперь следование в обратную сторону.

Покажем, что из утверждения выше следует полная независимость семейства событий
$$\{X_i^{-1}(\{a_{i,j_i}\})\}_{i=1}^n.$$
Понятно, что для всякого i семейство множеств
$$\{X_i^{-1}(a_{i,k_i})\}_{k_i \in I_i}$$
есть разбиение Ω. Значит для всякого $S \subseteq \{1; \dots n\}$ мы имеем, что
$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i \in S} X_i^{-1}(a_{i,j_i})\right) &= \sum_{\substack{\{k_i\}_{i \notin S} \\ k_i \in I_i}} \mathbb{P}\left(\bigcap_{i \in S} X_i^{-1}(a_{i,j_i}) \cap \bigcap_{i \notin S} X_i^{-1}(a_{i,k_i})\right) \\
&= \sum_{\substack{\{k_i\}_{i \notin S} \\ k_i \in I_i}} \prod_{i \in S} \mathbb{P}(X_i^{-1}(a_{i,j_i})) \cdot \prod_{i \notin S} \mathbb{P}(X_i^{-1}(a_{i,k_i})) \\
&= \prod_{i \in S} \mathbb{P}(X_i^{-1}(a_{i,j_i})) \cdot \sum_{\substack{\{k_i\}_{i \notin S} \\ k_i \in I_i}} \prod_{i \notin S} \mathbb{P}(X_i^{-1}(a_{i,k_i})) \\
&= \prod_{i \in S} \mathbb{P}(X_i^{-1}(a_{i,j_i})) \cdot \prod_{i \notin S} \sum_{k_i \in I_i} \mathbb{P}(X_i^{-1}(a_{i,k_i})) \\
&= \prod_{i \in S} \mathbb{P}(X_i^{-1}(a_{i,j_i}))
\end{aligned}$$

Теперь покажем, что из независимости прообразов одноэлементных множеств следует независимость прообразов любых множеств. Действительно, пусть дано семейство множеств $\{B_i\}_{i=1}^n$, что $B_i \subseteq \Omega(X_i)$ (следовательно B_i не более чем счётно). Тогда для всякого конечного $S \subseteq \{1; \ldots; n\}$

$$\begin{aligned}
\mathbb{P}\left(\bigcap_{i\in S} X_i^{-1}(B_i)\right) &= \mathbb{P}\left(\bigsqcup_{\substack{\{a_i\}_{i\in S}\\ a_i\in B_i}} \bigcap_{i\in S} X_i^{-1}(a_i)\right) &&= \sum_{\substack{\{a_i\}_{i\in S}\\ a_i\in B_i}} \mathbb{P}\left(\bigcap_{i\in S} X_i^{-1}(a_i)\right) \\
&= \sum_{\substack{\{a_i\}_{i\in S}\\ a_i\in B_i}} \prod_{i\in S} \mathbb{P}(X_i^{-1}(a_i)) &&= \prod_{i\in S} \sum_{a_i\in B_i} \mathbb{P}(X_i^{-1}(a_i)) \\
&= \prod_{i\in S} \mathbb{P}(X_i^{-1}(B_i))
\end{aligned}$$

□

Пример 4. Пусть $X \overset{\mathcal{D}}{=} \mathcal{P}(\alpha)$, $Y \overset{\mathcal{D}}{=} \mathcal{P}(\beta)$ и X и Y независимы. Значит $X+Y$ имеет множество значений $\mathbb{N} \cup \{0\}$, а её распределение

$$\begin{aligned}
\mathbb{P}\{X+Y=n\} &= \mathbb{P}\left(\bigsqcup_{k=0}^{n} \{X=k \wedge Y=n-k\}\right) &&= \sum_{k=0}^{n} \mathbb{P}\{X=k \wedge Y=n-k\} \\
&= \sum_{k=0}^{n} \mathbb{P}\{X=k\} \cdot \mathbb{P}\{Y=n-k\} &&= \sum_{k=0}^{n} \frac{\alpha^k}{k!} e^{-\alpha} \cdot \frac{\beta^{n-k}}{(n-k)!} e^{-\beta} \\
&= \sum_{k=0}^{n} \frac{\alpha^k \beta^{n-k} \binom{n}{k}}{n!} e^{-(\alpha+\beta)} &&= \frac{(\alpha+\beta)^n}{n!} e^{-(\alpha+\beta)},
\end{aligned}$$

т.е. $X+Y \overset{\mathcal{D}}{=} \mathcal{P}(\alpha+\beta)$.

Пример 5 (испытания Бернулли). Пусть $\{\varepsilon_i\}_{i=1}^n$ — независимые случайные величины с распределением Бернулли $B(1,p)$. Тогда $\sum_{i=1}^n \varepsilon_i$ имеет множество значений $\{0; \ldots; n\}$, а её распределение

$$\begin{aligned}
\mathbb{P}\left\{\sum_{i=1}^{n}\varepsilon_i = k\right\} &= \mathbb{P}\left(\bigsqcup_{\substack{S\subseteq\{1;\ldots;n\}\\|S|=k}}\left\{\bigwedge_{i\in S}X_i = 1 \wedge \bigwedge_{j\notin S}X_j = 0\right\}\right)\\
&= \sum_{\substack{S\subseteq\{1;\ldots;n\}\\|S|=k}}\mathbb{P}\left\{\bigwedge_{i\in S}X_i = 1 \wedge \bigwedge_{j\notin S}X_j = 0\right\}\\
&= \sum_{\substack{S\subseteq\{1;\ldots;n\}\\|S|=k}}\prod_{i\in S}\mathbb{P}\{X_i = 1\}\cdot\prod_{j\notin S}\mathbb{P}\{X_j = 0\}\\
&= \sum_{\substack{S\subseteq\{1;\ldots;n\}\\|S|=k}}\prod_{i\in S}\mathbb{P}\{X_i = 1\}\cdot\prod_{j\notin S}\mathbb{P}\{X_j = 0\}\\
&= \sum_{\substack{S\subseteq\{1;\ldots;n\}\\|S|=k}} p^{|S|}\cdot(1-p)^{n-|S|}\\
&= p^k(1-p)^{n-k}\sum_{\substack{S\subseteq\{1;\ldots;n\}\\|S|=k}} 1\\
&= \binom{n}{k}p^k(1-p)^{n-k},
\end{aligned}$$

т.е.

$$\sum_{i=1}^{n}\varepsilon_i \overset{\mathcal{D}}{=} B(n,p).$$

1.5 Построение сложных вероятностных пространств

Теорема 11.

1. *Пусть даны вероятностные пространства* $(\Omega_1,\mathcal{F}_1,\mathbb{P}_1)$ *и* $(\Omega_2,\mathcal{F}_2,\mathbb{P}_2)$. *Обозначим*

   - $\Omega := \Omega_1\times\Omega_2$,
   - $\mathcal{F}$ *— минимальная сигма-алгебра, содержащая как подмножество*

     $$\{S_1\times S_2 \mid S_1\in\mathcal{F}_1\wedge S_2\in\mathcal{F}_2\},$$

   - $\mathbb{P}$ *— счётно-аддитивная функция на* $(\Omega,\mathcal{F})$, *что для всяких* $S_1\in\mathcal{F}_1$ *и* $S_2\in\mathcal{F}_2$

     $$\mathbb{P}(S_1\times S_2) = \mathbb{P}_1(S_1)\cdot\mathbb{P}_2(S_2)$$

     (такая функция существует и единственна).

   Тогда $(\Omega,\mathcal{F},\mathbb{P})$ *— вероятностное пространство.*

2. *Пусть даны события* A_1 *и* A_2 *в вероятностных пространствах* $(\Omega_1,\mathcal{F}_1,\mathbb{P}_1)$ *и* $(\Omega_2,\mathcal{F}_2,\mathbb{P}_2)$. *Тогда множества*

   $$B_1 := A_1\times\Omega_2 \qquad \text{и} \qquad B_2 := \Omega_1\times A_2$$

   являются событиями в вероятностном пространстве $(\Omega,\mathcal{F},\mathbb{P})$, *причём*

   $$\mathbb{P}_1(A_1) = \mathbb{P}(B_1) \qquad \text{и} \qquad \mathbb{P}_2(A_2) = \mathbb{P}(B_2).$$

3. *Пусть даны независимые семейства события $\{A_{1,i}\}_{i\in I_1}$ и $\{A_{2,i}\}_{i\in I_2}$ в вероятностных пространствах $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ и $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$. Тогда объединение семейств $\{B_{1,i}\}_{i\in I_1}$ и $\{B_{2,i}\}_{i\in I_2}$, где*
$$B_{1,i} := A_{1,i} \times \Omega_2 \qquad \text{и} \qquad B_{2,i} := \Omega_1 \times A_{2,i},$$
является независимым в вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$.

4. *Пусть даны случайные величины X_1 и X_2 в вероятностных пространствах $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ и $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ на измеримые пространства $(R_1, \mathcal{G}_1)$ и $(R_2, \mathcal{G}_2)$ соответственно. Тогда функции*
$$Y_1 : \Omega \to R_1, (\omega_1, \omega_2) \mapsto X_1(\omega_1) \qquad \text{и} \qquad Y_2 : \Omega \to R_2, (\omega_1, \omega_2) \mapsto X_2(\omega_2)$$
являются случайными величинами в вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$ на те же измеримые пространства, причём
$$X_1 \overset{\mathcal{D}}{=} Y_1 \qquad \text{и} \qquad X_2 \overset{\mathcal{D}}{=} Y_2.$$

5. *Пусть даны независимые семейства случайных величин $\{X_{1,i}\}_{i\in I_1}$ и $\{X_{2,i}\}_{i\in I_2}$ в вероятностных пространствах $(\Omega_1, \mathcal{F}_1, \mathbb{P}_1)$ и $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$. Тогда объединение семейств $\{Y_{1,i}\}_{i\in I_1}$ и $\{Y_{2,i}\}_{i\in I_2}$, где*
$$Y_{1,i} : \Omega \to R_{1,i}, (\omega_1, \omega_2) \mapsto X_{1,i}(\omega_1) \qquad \text{и} \qquad Y_{2,i} : \Omega \to R_{2,i}, (\omega_1, \omega_2) \mapsto X_{2,i}(\omega_2),$$
является независимым в вероятностном пространстве $(\Omega, \mathcal{F}, \mathbb{P})$.

Доказательство.

1. Пока непростое утверждение. Можно посмотреть, например, здесь.

2. Поскольку $\Omega_2 \in \mathcal{F}_2$, то B_1 по определению лежит в $\mathcal{F}$, т.е. B_1 является событием. И по тому же определению
$$\mathbb{P}(B_1) = \mathbb{P}(A_1 \times \Omega_2) = \mathbb{P}_1(A_1) \cdot \mathbb{P}_2(\Omega_2) = \mathbb{P}_1(A_1).$$
Аналогично будет и для A_2.

3. Заметим, что всякий конечное подсемейство нового семейства $\{B_{1,i}\}_{i\in I_1} \cup \{B_{2,i}\}_{i\in I_2}$ можно получить, рассмотрев для некоторых конечных $S_1 \subseteq I_1$ и $S_2 \subseteq I_2$ семейство $\{B_{1,i}\}_{i\in S_1} \cup \{B_{2,i}\}_{i\in S_2}$. Тогда
$$\begin{aligned}\mathbb{P}\left(\bigcap_{i\in S_1} B_{1,i} \cap \bigcap_{i\in S_2} B_{2,i}\right) &= \mathbb{P}\left(\left(\bigcap_{i\in S_1} A_{1,i}\right) \times \left(\bigcap_{i\in S_2} A_{2,i}\right)\right) \\ &= \mathbb{P}_1\left(\bigcap_{i\in S_1} A_{1,i}\right) \cdot \mathbb{P}_2\left(\bigcap_{i\in S_2} A_{2,i}\right) \\ &= \prod_{i\in S_1} \mathbb{P}_1(A_{1,i}) \cdot \prod_{i\in S_2} \mathbb{P}_2(A_{2,i}) \\ &= \prod_{i\in S_1} \mathbb{P}(B_{1,i}) \cdot \prod_{i\in S_2} \mathbb{P}(B_{2,i}).\end{aligned}$$
Таким образом, рассматривая все конечные подсемейства нашего семейства, получаем условие независимости.

4. Заметим, что для всякого $B \in \mathcal{G}_1$

$$Y_1^{-1}(B) = X_1^{-1}(B) \times \Omega_2 \in \mathcal{F},$$

т.е. Y_1 — случайная величина. При этом

$$\mathbb{P}(Y_1^{-1}(B)) = \mathbb{P}(X_1^{-1}(B) \times \Omega_2) = \mathbb{P}_1(X_1^{-1}(B)),$$

т.е. X_1 и Y_1 порождают одно распределение, т.е. $X_1 \overset{\mathcal{D}}{=} Y_1$. Аналогично и для X_2.

5. Пусть $\{B_{1,i}\}_{i \in I_1}$ и $\{B_{2,i}\}_{i \in I_2}$ — всякие семейства множеств, где $B_{1,i} \in \mathcal{G}_{1,i}$ и $B_{1,i} \in \mathcal{G}_{2,i}$. Таким образом мы получаем, что семейства

$$\{X_{1,i}^{-1}(B_{1,i})\}_{i \in I_1} \qquad \text{и} \qquad \{X_{2,i}^{-1}(B_{2,i})\}_{i \in I_2}$$

будут независимыми, а значит семейство их вложений в новое вероятностное пространство будет независимым по предыдущим пунктам. При этом их вложения в новое вероятностное пространство есть прообразы соответствующих множеств по соответствующим Y-величинам. Перебирая все возможные $B_{1,i}$ и $B_{2,i}$ получаем условие независимости всех Y-величин (в совокупности).

□

Теорема 12.

То же самое для взятия по условию события.

Замечание. Таким образом мы теперь умеем склеивать два вероятностных пространства в новое большое вероятностное пространство и сжимать вероятностное пространство по модулю всякого его события.

Теорема 13.

1. *Распределение Бернулли $B(1,p)$ реализуемо.*
2. *Биномиальное $B(n,p)$ распределение реализуемо.*
3. *Распределение Пуассона $\mathcal{P}(\alpha)$ реализуемо.*
4. *Всякое не более чем счётное распределение реализуемо.*
5. *Всякое распределение на всяком измеримом пространстве $(R, \mathcal{G})$ реализуемо.*

Доказательство.

1. Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$, где
   - $\Omega = \{0; 1\}$,
   - $\mathcal{F} = \mathcal{P}(\Omega)$,
   - $\mathbb{P}(\varnothing) = 0$, $\mathbb{P}(\{0\}) = 1 - p$, $\mathbb{P}(\{1\}) = p$, $\mathbb{P}(\{0; 1\}) = 1$.

   В таком пространстве случайная величина $X(\omega) = \omega$ задаёт распределение Бернулли $B(1,p)$.

2. Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$, где

- $\Omega = \{0;1\}^n$,
- $\mathcal{F} = \mathcal{P}(\Omega)$,
- $\mathbb{P}((a_1, \dots, a_n)) = p^{\sum a_i} \cdot (1-p)^{n-\sum a_i}$.

В таком пространстве случайная величина $X((a_1, \dots, a_n)) = \sum_{i=1}^n a_i$ задаёт биномиальное распределение $B(n,p)$.

3. Рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$, где

   - $\Omega = \mathbb{N} \cup \{0\}$,
   - $\mathcal{F} = \mathcal{P}(\Omega)$,
   - $\mathbb{P}(S) = \sum_{k \in S} \frac{\alpha^k}{k!} e^{-\alpha}$.

   В таком пространстве случайная величина $X(\omega) = \omega$ задаёт распределение Пуассона $B(n,p)$.

4. Пусть даны множество значений $\{a_i\}_{i \in I}$ случайной величины и соответствующие им вероятности $\{p_i\}_{i \in I}$ ($p_i \geqslant 0$ и $\sum_{i \in I} p_i = 1$), где I не более чем счётно. Тогда рассмотрим вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$, где

   - $\Omega = \{a_i\}_{i \in I}$,
   - $\mathcal{F} = \mathcal{P}(\Omega)$,
   - $\mathbb{P}(\{a_i\}_{i \in S}) = \sum_{i \in S} p_i$ для всякого $S \subseteq I$.

   В таком пространстве случайная величина $X(\omega) = \omega$ задаёт распределение Пуассона $B(n,p)$.

5. Пусть дана функция распределения P, т.е. такая функция $P : \mathcal{G} \to [0;1]$, что $(R, \mathcal{G}, P)$ будет вероятностным пространством (напомним, что распределение всякой случайной величины индуцирует вероятностную меру на измеримом пространстве-образе). Поэтому рассмотрим вероятностное пространство $(R, \mathcal{G}, P)$. В таком пространстве случайная величина $X(\omega) = \omega$ задаёт распределение P.

Заметим, что во всех случаях несложно проверить, что вероятностная мера в рассматриваемых вероятностных пространств, действительно, удовлетворяет всем условиям на вероятностную меру. $\square$

1.6 Пуассоновская аппроксимация

Теорема 14. *Пусть дана последовательность чисел $\{p_n\}_{n=0}^\infty$ из отрезка $[0;1]$, что*

$$\lim_{n \to \infty} np_n \to \alpha$$

для некоторой константы α, и последовательность случайных величин $\{X_n\}_{n=0}^\infty$, что

$$X_n \overset{\mathcal{D}}{=} B(n, p_n).$$

Тогда для всякого $k \in \mathbb{N} \cup \{0\}$

$$\lim_{n \to \infty} \mathbb{P}\{X_n = k\} = \frac{\alpha^k}{k!} e^{-\alpha}$$

Для всякого $n \in \mathbb{N} \cup \{0\}$

$$\sum_{k=0}^{\infty} |\mathbb{P}\{X_n = k\} - \frac{(np_n)^k}{k!} e^{-np_n}| \leqslant 2np^2$$

Доказательство.

$$\mathbb{P}\{X_n = k\} = \frac{n!}{k!(n-k)!} p_n^k (1-p_n)^{n-k} = \frac{n \cdot \dots \cdot (n-k+1)}{k! n^k} (np_n)^k (1-p_n)^{n-k}.$$

При этом

$$\frac{n \cdot \dots \cdot (n-k+1)}{k! n^k} \to \frac{1}{k!}, \qquad (np_n)^k \to \alpha^k, \qquad (1-p_n)^{n-k} = \left(1 - \frac{\alpha + o(1)}{n}\right)^{n(1+o(1))} \to e^{-\alpha}.$$

Отсюда и получается требуемое утверждение. □

Теорема 15. *Пусть даны некоторое подмножество $A \subseteq \mathbb{R}$ и случайные виличины $S_n \stackrel{\mathcal{D}}{=} B(n,p)$ и $S \stackrel{\mathcal{D}}{=} \mathcal{P}(np)$. Тогда*

$$|\mathbb{P}\{S_n \in A\} - \mathbb{P}\{S \in A\}| \leqslant np^2.$$

Доказательство. Рассмотрим следующее вероятностное пространство. Пусть

$$\Omega_1 := \mathbb{N} \cup \{0; -1\}, \qquad \mathcal{F}_1 := \mathcal{P}(\Omega) \qquad \text{и} \qquad \mathbb{P}_1(n) := \begin{cases} 1-p & \text{если } n = -1, \\ \frac{p^0}{0!} e^{-p} - (1-p) & \text{если } n = 0, \\ \frac{p^n}{n!} e^{-p} & \text{иначе.} \end{cases}$$

Определим на нём случайные величины

$$\varepsilon_1 := \begin{cases} 0 & \text{если } n = -1, \\ 1 & \text{иначе,} \end{cases} \qquad \text{и} \qquad \eta_1 := \begin{cases} 0 & \text{если } n = -1, \\ n & \text{иначе.} \end{cases}$$

Несложно видеть, что

$$\varepsilon_1 \stackrel{\mathcal{D}}{=} B(1,p), \qquad \text{а} \qquad \eta_1 \stackrel{\mathcal{D}}{=} \mathcal{P}(p).$$

При этом

$$\mathbb{P}_1\{\varepsilon_1 \neq \eta_1\} = 1 - \mathbb{P}_1\{\varepsilon_1 = \eta_1\} = 1 - \mathbb{P}_1(\{-1; 1\}) = 1 - ((1-p) + pe^{-p}) = p(1 - e^{-p}) \leqslant p^2.$$

Теперь сделаем ещё $n-1$ дубликатов нашего пространства и построенных случайных величин и перемножим их как в теореме 11. Получим пространство $(\Omega, \mathcal{F}, \mathbb{P})$ в котором выбраны случайные величины $\{\varepsilon_i\}_{i=1}^n$ и $\{\eta_i\}_{i=1}^n$, где

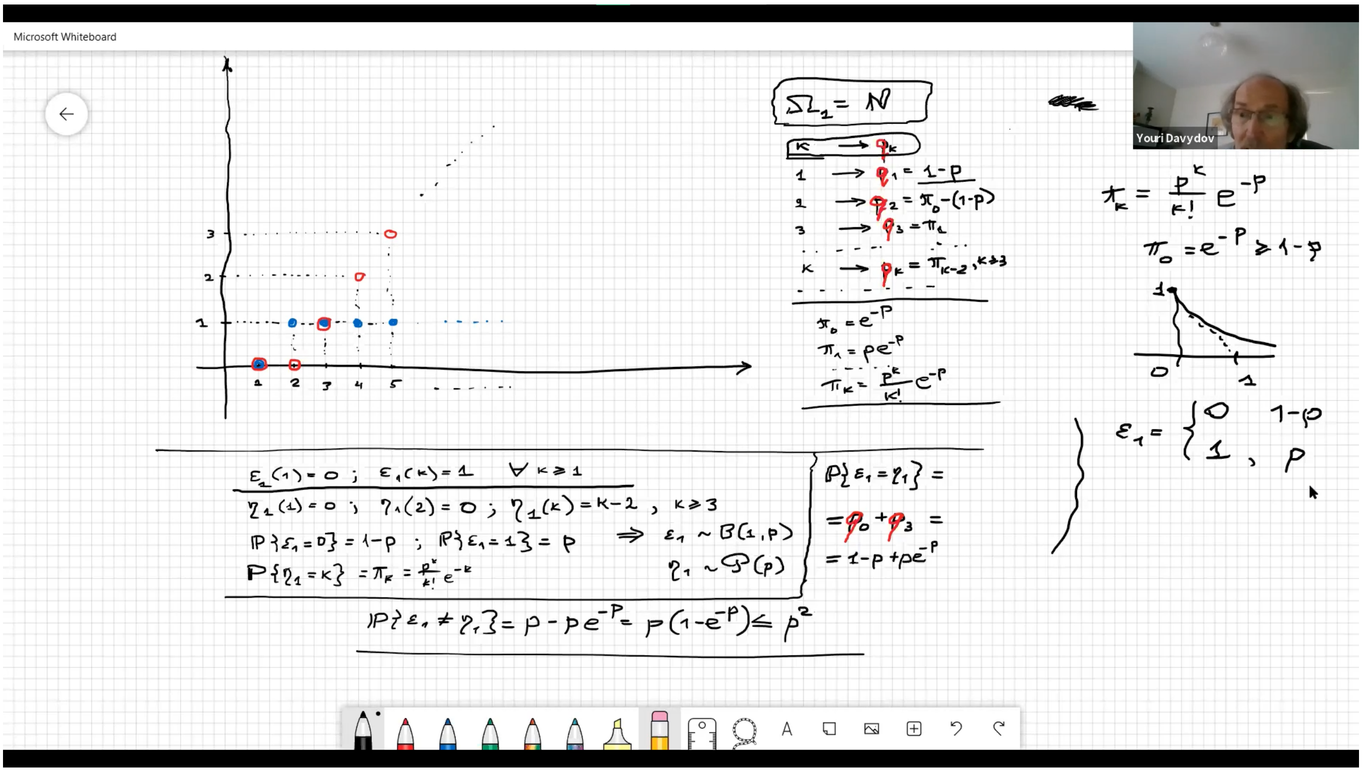
$$\varepsilon_i \stackrel{\mathcal{D}}{=} B(1,n), \qquad \text{а} \qquad \eta_i \stackrel{\mathcal{D}}{=} \mathcal{P}(p).$$

При этом семейства $\{\varepsilon_i\}_{i=1}^n$ и $\{\eta_i\}_{i=1}^n$ независимы. Значит

$$S_n \stackrel{\mathcal{D}}{=} B(n,p) \stackrel{\mathcal{D}}{=} X := \sum_{i=1}^n \varepsilon_i \qquad \text{и} \qquad S \stackrel{\mathcal{D}}{=} \mathcal{P}(np) \stackrel{\mathcal{D}}{=} Y := \sum_{i=1}^n \eta_i.$$

Следовательно

$$\{X \neq Y\} \subseteq \bigcup_{i=1}^n \{\varepsilon_i \neq \eta_i\} \quad \Longrightarrow \quad \mathbb{P}\{X \neq Y\} \leqslant \sum_{i=1}^n \mathbb{P}\{\varepsilon_i \neq \eta_i\} \leqslant np^2.$$



Перерисовать.

Отсюда имеем, что

$$
\begin{aligned}
|\mathbb{P}\{S_n \in A\} - \mathbb{P}\{S \in A\}| &= |\mathbb{P}\{X \in A\} - \mathbb{P}\{Y \in A\}| \\
&= |\mathbb{P}\{X \in A \wedge Y \notin A\} - \mathbb{P}\{Y \in A \wedge X \notin A\}| \\
&\leqslant \mathbb{P}\{X \in A \wedge Y \notin A\} + \mathbb{P}\{Y \in A \wedge X \notin A\} \\
&\leqslant \mathbb{P}\{X \neq Y\} \\
&\leqslant np^2.
\end{aligned}
$$

□

Следствие 15.1. *Пусть дано* $n \in \mathbb{N} \cup \{0\}$ *и случайные события* $S_n \overset{\mathcal{D}}{=} B(n,p)$ *и* $S \overset{\mathcal{D}}{=} \mathcal{P}(np)$. *Тогда*

$$
\sum_{k=0}^{\infty} |\mathbb{P}\{S_n = k\} - \mathbb{P}\{S = k\}| \leqslant 2np^2.
$$

Доказательство. Обозначим

$$
B_+ := \{k \mid \mathbb{P}\{S_n = k\} \geqslant \mathbb{P}\{S = k\}\} \qquad \text{и} \qquad B_- := (\mathbb{N} \cup \{0\}) \setminus B_+.
$$

Тогда понятно, что

$$
\begin{aligned}
\sum_{k=0}^{\infty} |\mathbb{P}\{S_n = k\} - \mathbb{P}\{S = k\}| &= \sum_{k \in B_+} (\mathbb{P}\{S_n = k\} - \mathbb{P}\{S = k\}) - \sum_{k \in B_-} (\mathbb{P}\{S_n = k\} - \mathbb{P}\{S = k\}) \\
&= (\mathbb{P}\{S_n \in B_+\} - \mathbb{P}\{S \in B_+\}) - (\mathbb{P}\{S_n \in B_-\} - \mathbb{P}\{S \in B_-\}) \\
&\leqslant |\mathbb{P}\{S_n \in B_+\} - \mathbb{P}\{S \in B_+\}| + |\mathbb{P}\{S_n \in B_-\} - \mathbb{P}\{S \in B_-\}| \\
&\leqslant np^2 + np^2 = 2np^2
\end{aligned}
$$

□

2 Математическое ожидание

Определение 12 (для дискретной случайной величины)**.** Пусть даны вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$ и дискретная случайная величина X в нём на $\mathbb{R}$ (т.е. $X : \Omega \to \mathbb{R}$ — измеримая функция, что $X(\Omega)$ не более чем счётно). Тогда математическим ожиданием величины X называется сумма

$$\sum_{a \in X(\Omega)} a\mathbb{P}\{X = a\},$$

если она абсолютно сходится. Обозначение: $M(X)$ или $\mathbb{E}(X)$.

Определение 13 (через интеграл Лебега)**.** Пусть даны вероятностное пространство $(\Omega, \mathcal{F}, \mathbb{P})$ и случайная величина X в нём на $\mathbb{R}$ (т.е. $X : \Omega \to \mathbb{R}$ — измеримая функция). Тогда математическим ожиданием величины X называется интеграл Лебега

$$\int_\Omega X(\omega)\mathbb{P}(d\omega),$$

если он сходится. Обозначение: $M(X)$ или $\mathbb{E}(X)$.

Утверждение 16. *Определения 12 и 13 равносильны в случаях, когда оба определены.*

Замечание 5. Несмотря на равносильность этих определений, мы пока не знаем интеграл Лебега. Поэтому будем определять математическое ожидание для дискретного случая. Но для непрерывного случая будем представлять, что математическое ожидание означает именно интеграл случайной величины по пространству.

Также в случае дискретного пространства интеграл Лебега вырождается в сумму

$$\sum_{\omega \in \Omega} X(\omega)\mathbb{P}(\omega),$$

что, понятно, равно определению для дискретной величины.

Лемма 17.

1. $\mathbb{E}(X)$ *определено тогда и только тогда, когда* $\mathbb{E}(|X|)$ *определено.*
2. ***Линейность.*** *Для всяких случайных величин X и Y и констант a и b верно, что*
$$\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y).$$
3. ***Положительность.*** *Если $X \geqslant 0$, то $\mathbb{E}(X) \geqslant 0$.*
4. *Если $X \leqslant Y$, то $\mathbb{E}(X) \leqslant \mathbb{E}(Y)$.*
5. $|\mathbb{E}(X)| \leqslant \mathbb{E}(|X|)$.
6. ***Неравенство Йенсена.*** *Пусть $\varphi : \mathbb{R} \to \mathbb{R}$ — выпуклая функция. Тогда $\mathbb{E}(\varphi(X)) \geqslant \varphi(\mathbb{E}(X))$.*
7. *Если X и Y имеют одинаковые распределения, то $\mathbb{E}(X) = \mathbb{E}(Y)$.*

Теорема 18. *Пусть даны независимые случайные величины X и Y со сходящимися мат. ожиданиями. Тогда*

$$\mathbb{E}(X \cdot Y) = \mathbb{E}(X) \cdot \mathbb{E}(Y).$$

Доказательство.

$$\begin{aligned}\mathbb{E}(X\cdot Y) &= \sum_{\substack{x\in X(\Omega)\\ y\in Y(\Omega)}} (x\cdot y)\mathbb{P}\{X=x\wedge Y=y\}\\ &= \sum_{\substack{x\in X(\Omega)\\ y\in Y(\Omega)}} x\cdot y\cdot \mathbb{P}\{X=x\}\cdot\mathbb{P}\{Y=y\}\\ &= \left(\sum_{x\in X(\Omega)} x\cdot\mathbb{P}\{X=x\}\right)\left(\sum_{y\in Y(\Omega)} y\cdot\mathbb{P}\{Y=y\}\right)\\ &= \mathbb{E}(X)\cdot\mathbb{E}(Y)\end{aligned}$$

Поскольку последние суммы — мат. ожидания X и Y — сходятся, то сходятся абсолютно, а значит абсолютно сходится и ряд-произведение. Значит мат. ожидание $X\cdot Y$ определено. □

Пример 6.

1. Пусть случайная величина $X \overset{\mathcal{D}}{=} B(n,p)$. Тогда по определению

   $$\mathbb{E}(X) = \sum_{k=0}^{n} k\binom{n}{k}p^k(1-p)^{n-k},$$

   что считать непросто. С другой стороны пусть $\{\varepsilon_i\}_{i=1}^n$ — независимые случайные величины с распределением $B(1,p)$. Тогда

   $$\mathbb{E}(X) = \mathbb{E}\left(\sum_{i=1}^{n}\varepsilon_i\right) = \sum_{i=1}^{n}\mathbb{E}(\varepsilon_i) = np.$$

2. Пусть случайная величина $X \overset{\mathcal{D}}{=} \mathcal{P}(\alpha)$. Тогда

   $$\mathbb{E}(X) = \sum_{n=0}^{\infty} n\frac{\alpha^n}{n!}e^{-\alpha} = \alpha\sum_{n=0}^{\infty}\frac{\alpha^n}{n!}e^{-\alpha} = \alpha.$$

Теорема 19 (неравенство Маркова)**.** *Пусть даны случайное событие $X \geqslant 0$ и некоторое значение $t > 0$. Тогда*

$$\mathbb{P}\{X\geqslant t\} \leqslant \frac{\mathbb{E}(X)}{t}.$$

Доказательство. Рассмотрим множество (событие)

$$A := \{\omega\in\Omega \mid X(\omega)\geqslant t\}$$

Заметим, что

$$t\mathbb{1}_A \leqslant X,$$

а значит

$$\mathbb{E}(X) \geqslant \mathbb{E}(t\mathbb{1}_A) = t\mathbb{E}(\mathbb{1}_A) = t\mathbb{P}\{X\geqslant t\}.$$

□

Следствие 19.1. *Пусть даны случайное событие X, произвольное подмножество $S \subseteq \mathbb{R}$, что $X(\Omega) \subseteq S$, неотрицательно определённая неубывающая функция $f : S \to \mathbb{R}$ и некоторая величина $t \in S$, что $f(t) > 0$. Тогда*

$$\mathbb{P}\{X \geqslant t\} \leqslant \frac{\mathbb{E}(f(X))}{f(t)}$$

Доказательство. Заметим, что из неубываемости f следует, что

$$X \geqslant t \quad \Longrightarrow \quad f(X) \geqslant f(t),$$

следовательно

$$\mathbb{P}\{X \geqslant t\} \leqslant \mathbb{P}\{f(X) \geqslant f(t)\}.$$

Применяя неравенство Маркова к $f(X)$ и $f(t)$ получаем

$$\mathbb{P}\{X \geqslant t\} \leqslant \mathbb{P}\{f(X) \geqslant f(t)\} \leqslant \frac{\mathbb{E}(f(X))}{f(t)}.$$

□

Пример 7.

1. Пусть $f(x) = x^2$. Тогда применяя теорему к $|X|$ и $t > 0$, получаем, что

$$\mathbb{P}\{|x| \geqslant t\} \leqslant \frac{\mathbb{E}(X^2)}{t^2}.$$

2. Пусть $f(x) = e^{ax}$. Тогда применяя теорему к X и t, получаем, что

$$\mathbb{P}\{X \geqslant t\} \leqslant \frac{\mathbb{E}(e^{aX})}{e^{at}}.$$

2.1 Дисперсия

Определение 14. *Дисперсия* случайной величины X — значение

$$\mathbb{E}\Big((X - \mathbb{E}(X))^2\Big).$$

Обозначение: $D(X)$ (в русской литературе), $\mathrm{Var}(X)$ (в английской литературе).

Лемма 20.

1. *$D(X)$ сходится тогда и только тогда, когда сходятся $E(X)$ и $E(X^2)$.*
2. *$D(X) \geqslant 0$.*
3. *$D(X) = 0 \Leftrightarrow \mathbb{P}\{X \neq \mathbb{E}(X)\} = 0$.*
4. *$D(X + a) = D(X)$.*
5. *$D(\lambda X) = \lambda^2 D(X)$.*
6. *$D(X) = E(X^2) - E(X)^2$.*
7. *Если X и Y независимы, то $D(X + Y) = D(X) + D(Y)$.*

Пример 8.

1. Пусть случайная величина $X \overset{\mathcal{D}}{=} B(n, p)$. Тогда по определению

$$D(X) = \sum_{k=0}^{n} (k - np)^2 \binom{n}{k} p^k (1-p)^{n-k},$$

что считать непросто. С другой стороны пусть $\{\varepsilon_i\}_{i=1}^{n}$ — независимые случайные величины с распределением $B(1, p)$. Тогда

$$D(\varepsilon_i) = \mathbb{E}(\varepsilon_i^2) - \mathbb{E}(\varepsilon_i)^2 = \mathbb{E}(\varepsilon_i) - \mathbb{E}(\varepsilon_i)^2 = p - p^2 = p(1-p)$$

$$D(X) = D\left(\sum_{i=1}^{n} \varepsilon_i\right) = \sum_{i=1}^{n} D(\varepsilon_i) = np(1-p).$$

2. Пусть случайная величина $X \overset{\mathcal{D}}{=} \mathcal{P}(\alpha)$. Тогда

$$\mathbb{E}(X(X-1)) = \sum_{n=0}^{\infty} n(n-1)\frac{\alpha^n}{n!}e^{-\alpha} = \alpha^2 \sum_{n=0}^{\infty} \frac{\alpha^n}{n!}e^{-\alpha} = \alpha^2,$$

и следовательно

$$D(X) = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2 = \alpha^2 + \alpha - \alpha^2 = \alpha.$$

Теорема 21. *Пусть даны случайная величина X и некоторая константа $t > 0$. Тогда*

$$\mathbb{P}\{|X - \mathbb{E}(X)| \geqslant t\} \leqslant \frac{D(X)}{t^2}.$$

Доказательство. Применяя следствие 19.1 к $f(x) = x^2$, $Y := |X - \mathbb{E}(X)|$ и $t > 0$, получаем

$$\mathbb{P}\{|X - \mathbb{E}(X)| \geqslant t\} = \mathbb{P}\{Y \geqslant t\} \leqslant \frac{\mathbb{E}(Y^2)}{t^2} = \frac{D(X)}{t^2}.$$

□

2.2 Момент и ковариация

Определение 15. Пусть X — случайная величина.

- $\mathbb{E}(X^n)$ $(n \in \mathbb{N})$ — *момент порядка n.*
- $\mathbb{E}(|X|^\alpha)$ $(\alpha \in \mathbb{R})$ — *абсолютный момент порядка α.*
- $\mathbb{E}((X - \mathbb{E}(X))^n)$ $(n \in \mathbb{N})$ — *центральный момент порядка n.*
- $\mathbb{E}(|X - \mathbb{E}(X)|^\alpha)$ $(\alpha \in \mathbb{R})$ — *абсолютный центральный момент порядка α.*

Пример 9. Дисперсия — (абсолютный) центральный момент второго порядка.

Определение 16. Пусть X, Y — случайные величины.

- $\mathbb{E}(X^n Y^m)$ — *смешанный момент порядка (n, m).*
- $\operatorname{cov}(X, Y) := \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y)))$ — *ковариация.*
- $\rho(X, y) := \frac{\operatorname{cov}(X,Y)}{\sqrt{D(X) \cdot D(Y)}}$ — *коэффициент корреляции*, мера линейной зависимости.

2.3 Закон больших чисел (ЗБЧ)

Определение 17. Пусть даны последовательность случайных величин $(X_n)_{n=0}^{\infty}$ и случайная величина X. Тогда говорится, что $(X_n)_{n=0}^{\infty}$ *сходится по вероятности* к X и обозначается

$$X_n \xrightarrow{\mathbb{P}} a \quad \text{при} \quad n \to \infty,$$

Теорема 22 (ЗБЧ)**.** *Пусть дана последовательность независимых одинаково распределённых величин* $(X_n)_{n=0}^{\infty}$*, что* $\mathbb{E}(X_k) = a$*,* $D(X_k) = \sigma^2$*. Пусть также* $S_n := \sum_{k=0}^{n} X_k$*. Тогда для всякого* $\varepsilon > 0$

$$\lim_{n\to\infty} \mathbb{P}\left\{ \left| \frac{S_n}{n} - a \right| \geqslant \varepsilon \right\} = 0,$$

т.е.

$$\frac{S_n}{n} \xrightarrow{\mathbb{P}} a \quad \text{при} \quad n \to \infty.$$

Доказательство. По неравенству Чебышёва

$$\mathbb{P}\left\{ \left| \frac{S_n}{n} - a \right| \geqslant \varepsilon \right\} \leqslant \frac{D(S_n/n)}{\varepsilon^2} = \frac{1}{\varepsilon^2} \cdot \frac{1}{n^2} \cdot D(S_n) = \frac{1}{\varepsilon^2} \cdot \frac{1}{n^2} \cdot \sigma^2 = \frac{\sigma^2}{\varepsilon^2} \cdot \frac{1}{n^2} \to 0$$

□

Теорема 23 (Хинчина)**.** *Пусть дана последовательность независимых одинаково распределённых величин* $(X_n)_{n=0}^{\infty}$*, что* $\mathbb{E}(X_k) = a$*. Пусть также* $S_n := \sum_{k=0}^{n} X_k$*. Тогда*

$$\frac{S_n}{n} \xrightarrow{\mathbb{P}} a.$$

Теорема 24 (Колмогорова)**.** *Пусть дана последовательность независимых одинаково распределённых величин* $(X_n)_{n=0}^{\infty}$*. Пусть также* $S_n := \sum_{k=0}^{n} X_k$*.*

1. *Пусть известно, что* $\mathbb{E}(X_k) = a$*. Тогда* $\frac{S_n}{n} \xrightarrow{\mathbb{P}} a$*.*

2. *Пусть известно, что* $\frac{S_n}{n} \xrightarrow{\mathbb{P}} a$*. Тогда* $\mathbb{E}(X_k) = a$*.*

2.4 Практика

Теорема 25 (аппроксимационная теорема Вейерштрасса)**.** *Пусть дана непрерывная функция* $f : [0;1] \to \mathbb{R}$*. Тогда для всякого* $\varepsilon > 0$ *есть многочлен* $Q \in \mathbb{R}[x]$*, что* $|f - Q| < \varepsilon$ *на отрезке* $[0;1]$*.*

Говоря иначе, есть последовательность многочленов $(Q_n)_{n=0}^{\infty}$ *из* $\mathbb{R}[x]$*, что*

$$\lim_{n\to\infty} \sup_{[0;1]} |f - Q_n| = 0.$$

Доказательство Бернштейна. Рассмотрим последовательность независимых случайных величин $(\xi_k)_{k=1}^{\infty}$ с распределением $B(1,p)$. Также обозначим $S_n := \sum_{k=1}^{n} \xi_k$. Поскольку $S_n \overset{\mathcal{D}}{=} B(n,p)$, то рассмотрим

$$Q_n(p) := \mathbb{E} f\left(\frac{S_n}{n}\right) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) P_{n,k}(p) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \binom{n}{k} p^k (1-p)^{n-k}$$

— многочлен от p, где

$$P_{n,k}(p) := \binom{n}{k} p^k (1-p)^{n-k}.$$

Тогда

$$\Delta_n(p) = |Q_n(p) - f(p)| = \left|\mathbb{E}f\left(\frac{S_n}{n}\right) - \mathbb{E}f(p)\right| = \left|\mathbb{E}\left(f\left(\frac{S_n}{n}\right) - f(p)\right)\right| \leqslant \mathbb{E}\left|f\left(\frac{S_n}{n}\right) - f(p)\right|.$$

Обозначим за

$$\omega_f(h) := \sup_{\substack{t,s\in[0;1]\\|t-s|\leqslant h}} |f(t) - f(s)|.$$

Зафиксируем какое-нибудь $\delta > 0$. Заметим, что

$$\mathbb{E}\left|f\left(\frac{S_n}{n}\right) - f(p)\right| = \underbrace{\sum_{k\colon \left|\frac{k}{n}-p\right|\leqslant\delta}\left|f\left(\frac{S_n}{n}\right) - f(p)\right| P_{n,k}(p)}_{\Sigma_1} + \underbrace{\sum_{k\colon \left|\frac{k}{n}-p\right|>\delta}\left|f\left(\frac{S_n}{n}\right) - f(p)\right| P_{n,k}(p)}_{\Sigma_2}.$$

Тогда

$$\Sigma_1 \leqslant \sum_{k\colon \left|\frac{k}{n}-p\right|\leqslant\delta} \omega_f(\delta)\cdot P_{n,k}(p) \leqslant \omega_f(\delta).$$

Пусть также $M := \operatorname{osc}_{[0;1]} f$, тогда

$$\Sigma_2 \leqslant M \sum_{k\colon \left|\frac{k}{n}-p\right|>\delta} P_{n,k}(p) = M\mathbb{P}\left\{\left|\frac{S_n}{n} - p\right| > \delta\right\} \leqslant M\frac{D\left(\frac{S_n}{n}\right)}{\delta^2} = \frac{Mp(1-p)}{\delta^2 n} \leqslant \frac{M}{4\delta^2}\cdot\frac{1}{n}.$$

Таким образом

$$\Delta_n(p) \leqslant \omega_f(\delta) + \frac{M}{4\delta^2}\cdot\frac{1}{n}.$$

По равномерной непрерывности f для всякого $\varepsilon > 0$ есть $\delta > 0$, что $\omega_f(\delta) < \varepsilon/2$. А тогда есть и $N \in \mathbb{N}$, что для всякого $n \geqslant N$

$$\frac{M}{4\delta^2}\cdot\frac{1}{n} < \frac{\varepsilon}{2},$$

т.е. для всех $n \in \mathbb{N}$

$$|Q_n(p) - f(p)| < \varepsilon.$$

А значит

$$\lim_{n\to\infty}\sup_{[0;1]}|f - Q_n| = 0.$$

□

Написать про нормальные числа (или нет смысла?). См. лекцию №5 от 16.03, примерно с 40:05.

3 Случайные блуждания (?)

3.1 Простое одномерное блуждание

Пусть мы находимся на $\mathbb{Z}$ и каждый момент двигаемся либо с вероятностью $1/2$ на $+1$, либо с вероятностью $1/2$ на -1. В таком случае, можно считать, что наше вероятностное пространство задаётся как

- $\Omega := \{1; -1\}^{\mathbb{N}}$ — множество счётных последовательностей с элементами из $\{1; -1\}$,

- $\mathcal{F}$ — минимальная сигма-алгебра, содержащая все $A_{n,s}$,
- $\mathbb{P}$ задаётся так, что $\mathbb{P}(A_{n,s}) = 1/2$ для всех $n \in \mathbb{N}$, $s \in \{1; -1\}$,

где $A_{n,s}$ подмножество всех последовательностей, у которых на месте n стоит s ($n \in \mathbb{N}$, $s \in \{1; -1\}$). (Почему $\mathbb{P}$ существует и единственно, опустим.)

Также рассмотрим случайные величины

$$\varepsilon_k := \begin{cases} 1 & \text{если } \omega \in A_{k,1}, \\ -1 & \text{если } \omega \in A_{k,-1}. \end{cases}$$

(изменения координаты на шаге k) и

$$S_n := \sum_{k=1}^{n} \varepsilon_k$$

координата после шага n.

Лемма 26.

$$\mathbb{P}\{S_n = k\} = \begin{cases} 0 & \textit{если } n \not\equiv k \mod 2, \\ \binom{n}{\frac{n+k}{2}} & \textit{иначе.} \end{cases}$$

Теорема 27. *Пусть играют два игрока. В начале у первого имеется капитал в размере a, а у второго — b ($a, b \in \mathbb{N} \cup \{0\}$). Каждый ход подбрасывается монетка; при выпадении орла первый даёт единицу капитала второму, а при выпадении решки — наоборот.*

1. *Вероятность разорения первого равна* $\frac{b}{a+b}$.
2. *Вероятность разорения кого-нибудь равна* 1.

Доказательство. Будем в каждый момент времени откладывать на целочисленный оси капитал первого. Тогда начинаем мы в a, каждый ход случайно сдвигаемся либо с вероятностью $1/2$ на $+1$, либо с вероятностью $1/2$ на -1, а заканчиваем игру, если попали в 0 (первый обанкротился), либо в $a + b$ (второй обанкротился).

Заметим, что вероятность разорения первого при фиксированном $a + b$ есть функция от a — $f(a)$. Понятно, что

$$f(0) = 1, \qquad f(a + b) = 0.$$

При этом если $a \in (0; a + b)$, то с вероятностью $1/2$ он спустится в $a - 1$, откуда вероятность обанкротится станет равна $f(a - 1)$, а с вероятностью $1/2$ — поднимется и вероятность станет равна $f(a + 1)$. Следовательно

$$f(a) = \frac{1}{2} f(a - 1) + \frac{1}{2} f(a).$$

Таким образом

$$f(a) = Ka + L;$$

подставляя точки 0 и $a + b$, получаем

$$f(a) = \frac{b}{a + b}.$$

По аналогии вероятность того, что второй обанкротится равна $\frac{a}{a+b}$. Поэтому вероятноость того, что уто-то обанкротится равна 1. □

Теорема 28 (следствие теоремы Пойа)**.** *Вероятность вернуться в* 0 *равна* 1.

Доказательство. Рассмотрим вероятности

$$p_n := \mathbb{P}\{S_n = 0\} = \mathbb{P}\{\text{“вернуться в 0 за } n \text{ шагов”}\}, \qquad p_0 = 1$$

и

$$0f_n := \mathbb{P}\left\{S_n = 0 \wedge \bigwedge_{k=1}^{n-1}\right\} = \mathbb{P}\{\text{“вернуться в 0 ровно за } n \text{ шагов”}\}, \qquad f_0 = 1.$$

Понятно, что вероятность вернуться в 0 когда-нибудь равна

$$\sum_{n=1}^{\infty} f_n.$$

Заметим, что для всякого $n \geqslant 1$ вероятность вернуться в 0 за n шагов можно разбить по случаям первого возвращения в 0 (от 1 до n), а значит

$$p_n = \sum_{k=1}^{n} f_k p_{n-k}.$$

При этом поскольку мы взяли $f_0 = 0$, то

$$p_n = \sum_{k=0}^{n} f_k p_{n-k}.$$

Пусть F и P — производящие функции $\{f_n\}_{n=0}^{\infty}$ и $\{p_n\}_{n=0}^{\infty}$. Тогда из предыдущего равенства мы получаем, что

$$P(t) - 1 = F(t)P(t),$$

т.е.

$$P(t) = \frac{1}{1 - F(t)}.$$

Заметим, что F и G сходятся в 1-окрестности нуля, F сходится в 1, и мы ищем значение $F(1)$. Поэтому несложно понять, что $F(1) = 1$ тогда и только тогда, когда $G(1) = \infty$.

При этом несложно понять, что

$$p_{2n+1} = 0, \qquad p_{2n} = \frac{1}{2^{2n}}\binom{2n}{n} \approx \frac{1}{\sqrt{\pi n}},$$

значит $G(1) = \infty$. $\square$

3.2 Непростое одномерное блуждание

Теперь зафиксируем $p \in [0; 1]$ и рассмотрим похожее вероятностное пространство: Ω и $\mathcal{F}$ такие же, а $\mathbb{P}$ определено так, что

$$\mathbb{P}(A_{n,s}) = \begin{cases} p & \text{если } s = 1, \\ 1 - p & \text{если } s = -1. \end{cases}$$

Несложно видеть, что по ЗБЧ

$$\frac{S_n}{n} \xrightarrow{\mathbb{P}} \mathbb{E}(\varepsilon_1) = p - (1 - p) = 2p - 1,$$

т.е. почти везде $S_n = (2p - 1)n(1 + o(1))$. Таким образом “понятно”, что возвратность будет только при $p = 1/2$.

Теорема 29. *В таком случае случайное блуждание невозвратно.*

Доказательство. Рассуждение с формальными рядами F и P не меняется. В таком случае

$$p_{2n} = \binom{2n}{n} p^n (1-p)^n \approx \frac{(4p(1-p))^n}{\sqrt{\pi n}}.$$

Но $4p(1-p) \leqslant 1$, а равенство достигается только для $p = 1/2$. Следовательно при $p \neq 1/2$ ряд $P(1)$ сходится, а тогда $F(1) < 1$, т.е. блуждание невозвратно. □

Теорема 30. *Если блуждание возвратно, то вероятность вернуться в* 0 *бесконечно много раз равна* 1. *Иначе эта же вероятность равна* 0.

Доказательство. Пусть R_n — событие возвращения n раз, а $r_n := \mathbb{P}(R_n)$. Понятно, что $R_{n+1} \subseteq R_n$.

$$r_{n+1} = \mathbb{P}(R_{n+1}) = \mathbb{P}(R_{n+1} \mid R_n)\mathbb{P}(R_n) = \mathbb{P}(R_1)\mathbb{P}(R_n) = r_1 r_n = r_1^{n+1}.$$

Значит вероятность возвращения бесконечное число раз равна

$$\lim_{n\to\infty} \mathbb{P}(R_n) = \lim_{n\to\infty} r_1^n = \begin{cases} 1 & \text{если } r_1 = 1, \\ 0 & \text{если } r_1 < 1. \end{cases}$$

□

3.3 Простой многомерный случай

Теперь мы рассмотрим ещё одно вероятностное пространство. Зафиксируем размерность d, пространство $\mathbb{R}^d$ и множество шагов $S := \{\pm e_k\}_{k=0}^d$. Тогда по аналогии определим $\Omega := S^{\mathbb{N}}$, $A_{n,s}$ (для всех $n \in \mathbb{N}$, $s \in S$), $\mathcal{F}$, $\mathbb{P}$, ε_n и S_n.

Теорема 31 (Пойа)**.** *Блуждание возвратно тогда и только тогда, когда* $d \in \{1; 2\}$.

Основные наброски доказательства. Рассуждение с рядами P и F остаются неизменными. Поэтому нужно понять, когда $P(1) = \infty$.

В случае $d = 2$ имеем, что

$$p_n = \sum_{k=0}^{n} \frac{\binom{2n}{k,k,n-k,n-k}}{4^{2n}} = \frac{\binom{2n}{n}}{4^{2n}} \sum_{k=0}^{n} \binom{n}{k}^2 = \frac{\binom{2n}{n}^2}{4^{2n}} \approx \frac{1}{\sqrt{\pi n}^2},$$

откуда сразу получается, что $P(1)$ рассходится.

В общем случае можно показать, что

$$p_n = \Theta(n^{-\frac{d}{2}}),$$

откуда так же сразу получается, что $P(1)$ сходится тогда и только тогда, когда $d > 2$. □

Также точно так же как в случае $d = 1$ можно показать, что если $f : \mathbb{Z}^d \to [0; 1]$ — вероятность вернуться из x в 0, то для всякого $x \in \mathbb{Z}^d$

$$f(x) = \frac{1}{2d} \sum_{s \in S} f(x + s).$$

В таком случае можно показать, что $f = 1$. Отсюда сразу можно получить, что случайное блуждание из 0 в каждую точку попадёт с вероятностью 1 (вероятность для каждой точки по отдельности).

3.4 Закон арксинуса

Будем рассматривать простое одномерное блуждание. Обозначим

$$T_+(n) := \sum_{k=1}^{n} \mathbb{1}_{\{S_{k-1}\geqslant 0 \wedge S_k \geqslant 0\}} \qquad \text{и} \qquad T_-(n) := \sum_{k=1}^{n} \mathbb{1}_{\{S_{k-1}\leqslant 0 \wedge S_k \leqslant 0\}}$$

— "время пребывания на $\mathbb{R}_+$" и "время пребывания на $\mathbb{R}_-$". Понятно, что T_+ и T_- одинаково распределены и зависимы (так как $T_+(n) + T_-(n) = n$).

Упражнение 1. Для всякого $n \in \mathbb{N}$

$$T_+(2n)$$

принимает только чётные значения.

Лемма 32 (см. Ширяев, "Вероятность", гл. I, §10 (стр. 129, лемма 2 по изданию 2007 года))**.**

$$\mathbb{P}\{T_+(2n) = 2k\} = \mathbb{P}\{S_{2k} = 0\} \cdot \mathbb{P}\{S_{2(n-k)} = 0\}$$

Теорема 33 (закон арксинуса)**.** *Для всяких* $0 < a < b < 1$

$$\lim_{n\to\infty} \mathbb{P}\left\{\frac{T_+(2n)}{2n} \in [a;b]\right\} = \frac{1}{\pi}\int_a^b \frac{dt}{\sqrt{t(1-t)}} = \frac{2}{\pi}(\sin^{-1}(\sqrt{b}) - \sin^{-1}(\sqrt{a}))$$

Доказательство. Из утверждения 32 понятно, что

$$\mathbb{P}\left\{\frac{T_+(2n)}{2n} \in [a;b]\right\} = \sum_{2na\leqslant 2k\leqslant 2nb} \mathbb{P}\{T_+(2n) = 2k\}$$

Следовательно все рассматриваемые $k \sim n$, а поэтому

$$\mathbb{P}\{T_+(2n) = 2k\} = \mathbb{P}\{S_{2k} = 0\}\cdot\mathbb{P}\{S_{2(n-k)} = 0\} = \frac{\binom{2k}{k}}{2^{2k}}\cdot\frac{\binom{2(n-k)}{n-k}}{2^{2(n-k)}} \approx \frac{1}{\pi\sqrt{k(n-k)}} = \frac{1}{\pi}\cdot\frac{1}{n}\cdot\frac{1}{\sqrt{\frac{k}{n}(1-\frac{k}{n})}}.$$

Следовательно при $n \to \infty$

$$\mathbb{P}\left\{\frac{T_+(2n)}{2n} \in [a;b]\right\} \approx \frac{1}{\pi}\sum_{\frac{k}{n}\in[a;b]} \frac{1}{n}\cdot\frac{1}{\sqrt{\frac{k}{n}(1-\frac{k}{n})}},$$

т.е. вырождается в интегральную сумму функции $\frac{1}{\pi}\cdot\frac{1}{t(1-t)}$ на отрезке $[a;b]$. Таким образом

$$\lim_{n\to\infty} \mathbb{P}\left\{\frac{T_+(2n)}{2n} \in [a;b]\right\} = \frac{1}{\pi}\int_a^b \frac{dt}{t(1-t)}.$$

□

3.5 Распределение максимума

Рассмотрим теперь случайную величину

$$M_n := \max_{k=0}^{n} S_k.$$

Понятно, что $M_n \geqslant 0$. Также несложно заметить, что

$$\begin{aligned}\mathbb{P}\{M_n \geqslant j\} &= \mathbb{P}\{M_n \geqslant j \wedge S_n \geqslant j\} + \mathbb{P}\{M_n \geqslant j \wedge S_n < j\}\\ &= \mathbb{P}\{S_n \geqslant j\} + \mathbb{P}\{M_n \geqslant j \wedge S_n < j\}\\ &= \mathbb{P}\{S_n \geqslant j\} + \mathbb{P}\{S_n > j\}\\ &= 2\mathbb{P}\{S_n \geqslant j\} - \mathbb{P}\{S_n > j\}.\end{aligned}$$

(Так как если $S_n < j$, то путь после первого достигания j можно отразить и получить биекцию со случаями $S_n > j$.)

3.6 Центральная предельная теорема (ЦПТ)

Теперь будем рассматривать испытания Бернулли с распределением $B(1,p)$, где $p \in (0;1)$, и определённые в таком случае $\{\xi_n\}_{n=1}^{\infty}$ и $\{S_n\}_{n=0}^{\infty}$. Вспомним, что тогда

$$\mathbb{P}\{S_n = k\} = \binom{n}{k} p^k (1-p)^{n-k}.$$

Теорема 34 (локальная теорема Муавра-Лапласа)**.** *Для всякой последовательности* $\{\varepsilon_n\}_{n=1} \to 0$

$$\mathbb{P}\{S_n = k\} = \frac{1}{\sqrt{2\pi np(1-p)}} \cdot e^{-\frac{x_k^2}{2}}(1 + o(1))$$

равномерно по всем $k \in [np - \varepsilon_n n^{2/3}; np + \varepsilon_n n^{2/3}]$*, где*

$$x_k = \frac{k - np}{\sqrt{np(1-p)}}.$$

Доказательство. Понятно, что $k = np(1 + o(1))$ и $n - k = n(1-p)(1+o(1))$. Следовательно, применяя формулу Стирлинга,

$$\begin{aligned}\mathbb{P}\{S_n = k\} &\approx \frac{1}{\sqrt{2\pi n}} \cdot \frac{1}{\sqrt{\frac{k}{n}(1 - \frac{k}{n})}} \cdot \frac{n^n}{k^k (n-k)^{n-k}} \cdot p^k (1-p)^k \\ &\approx \frac{1}{\sqrt{2\pi np(1-p)}} \cdot \left(\frac{np}{k}\right)^k \cdot \left(\frac{n(1-p)}{n-k}\right)^{n-k}.\end{aligned}$$

Пусть $v := k - np = o(n^{2/3})$. Заметим, что

$$\ln\left(\frac{np}{k}\right)^k = -k \ln\left(1 + \frac{v}{np}\right) = -(np + v)\left(\frac{v}{np} - \frac{v^2}{2(np)^2} + o\left(\frac{v^3}{(np)^3}\right)\right) = v + \frac{v^2}{2np} + o(1)$$

Аналогично

$$\ln\left(\frac{n(1-p)}{n-k}\right)^{n-k} = -v + \frac{v^2}{2n(1-p)} + o(1),$$

откуда

$$\ln\left(\left(\frac{np}{k}\right)^k \cdot \left(\frac{n(1-p)}{n-k}\right)^{n-k}\right) = -\frac{v^2}{2np(1-p)} + o(1) = -\frac{x_k^2}{2} + o(1).$$

Следовательно

$$\mathbb{P}\{S_n = k\} \approx \frac{1}{\sqrt{2\pi np(1-p)}} \cdot e^{-\frac{x_k^2}{2}}.$$

Равномерность получается внимательным всматриванием в доказательство. □

Теорема 35 (интегральная теорема Муавра-Лапласа)**.** *Пусть*

$$Z_n := \frac{S_n - np}{\sqrt{np(1-p)}}.$$

Тогда

$$\lim_{n\to\infty} \mathbb{P}\{Z_n \in [a;b]\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

и сходимость равномерна по a *и* b*.*

Доказательство. По локальной теореме Муавра-Лапласа имеем, что

$$\mathbb{P}\{Z_n \in [a;b]\} = \sum_{k:\frac{k-np}{\sqrt{np(1-p)}}} \mathbb{P}\{S_n = k\} = \frac{1}{2\pi} \cdot \frac{1}{\sqrt{np(1-p)}} \sum_{k:\frac{k-np}{\sqrt{np(1-p)}}} e^{-\frac{x_k^2}{2}}(1 + r_{n,k}),$$

где $r_{n,k} = o(1)$ равномерно на всём рассматриваемом отрезке. Следовательно

$$\lim_{n\to\infty} \mathbb{P}\{Z_n \in [a;b]\} = \frac{1}{2\pi} \cdot \lim_{n\to\infty} \frac{1}{\sqrt{np(1-p)}} \sum_{k:\frac{k-np}{\sqrt{np(1-p)}}} e^{-\frac{x_k^2}{2}} = \frac{1}{2\pi} \int_a^b e^{-\frac{x^2}{2}} dx.$$

Равномерность получается из того, что

$$\varphi_n(b) := \mathbb{P}\{Z_n \in [-\infty; b]\} \qquad \text{и} \qquad \varphi_n(b) := \frac{1}{2\pi} \int_a^b e^{-\frac{x^2}{2}} dx$$

— монотонно неубывающие ограниченные (так как лежат в $[0;1]$) функции на $\mathbb{R}$, и имеется поточечная сходимость $\{\varphi_n\}_{n=0}^{\infty} \to \varphi$; поэтому должна быть и равномерная сходимость, а следовательно и равномерная сходимость $\varphi_n(b) - \varphi_n(a)$. □

Теорема 36 (Paul Lévy, анонс)**.** *Пусть* $\{X_n\}_{n=0}^{\infty}$ *— независимые одинаково распределённые случайные величины, что* $\mathbb{E}(X_n) = A$ *и* $D(X_n) = \sigma^2$ *(*$\sigma^2 < \infty$*). Пусть также*

$$S_n = \sum_{k=1}^{n} X_k.$$

Тогда

$$\lim_{n\to\infty} \mathbb{P}\left\{\frac{S_n - \mathbb{E}(S_n)}{\sqrt{D(S_n)}} \in [a;b]\right\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{x^2}{2}} dx$$

и сходимость равномерно по a *и* b*.*

Замечание 6. Понятно, что эта теорема является значительным расширением интегральной теоремы Муавра-Лапласа.

3.7 Производящие функции

Определение 18. Пусть дана случайная величина X с множеством значений $X(\Omega) = \mathbb{N} \cup \{0\}$. Обозначим $p_k := \mathbb{P}\{X = k\}$. Тогда *производящей функцией* X называется функция (или ряд)

$$\varphi_X(z) = \mathbb{E}(z^X) = \sum_{k=0}^{\infty} p_k z^k.$$

Лемма 37.

1. φ *определена в* $U_1(0)$*.*
2. $\varphi_X(t)$ *выпукла на* $[0;1]$*.*
3. $\frac{\varphi_X^{(n)}(0)}{n!} = \mathbb{P}\{X = n\}$*.*
4. $\varphi_X(1) = 1$*.*
5. *Для независимых* X *и* Y *верно, что* $\varphi_{X+Y} = \varphi_X \varphi_Y$*.*
6. ***Связь с моментами.*** *Если* $\mathbb{E}(X^n) < \infty$*, то*

$$\varphi_X^{(n)}(1) = \mathbb{E}\left(\prod_{k=0}^{n-1} X - k\right).$$

3.7.1 Сумма случайного числа случайных величин

Теорема 38. *Пусть даны набор независимых случайных величин $\{X_n\}_{n=1}^{\infty}$ с множеством значений $\mathbb{N} \cup \{0\}$ и распределением вероятностей $\mathbb{P}\{X_i = k\} = p_k$, их производящая функция φ, независящая от них случайная величина τ с множеством значений $\mathbb{N} \cup \{0\}$. Обозначим*

$$S_n := \sum_{k=1}^{n} X_k \qquad \text{и} \qquad U = S_\tau = \sum_{k=1}^{\tau} X_k.$$

Тогда

$$\varphi_U(z) = \varphi_\tau(\varphi(z)).$$

Доказательство.

$$\begin{aligned}
\varphi_U(z) &= \sum_{k=0}^{\infty} \mathbb{P}\{U = k\} z^k & &= \sum_{k=0}^{\infty} \sum_{n=0}^{\infty} \mathbb{P}\{U = k \wedge \tau = n\} z^k \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \mathbb{P}\{\tau = n \wedge S_n = k\} z^k & &= \sum_{n=0}^{\infty} \mathbb{P}\{\tau = n\} \sum_{k=0}^{\infty} \mathbb{P}\{S_n = k\} z^k \\
&= \sum_{n=0}^{\infty} \mathbb{P}\{\tau = n\} \varphi_{S_n}(z) & &= \sum_{n=0}^{\infty} \mathbb{P}\{\tau = n\} \varphi(z)^n \\
&= \varphi_\tau(\varphi(z)).
\end{aligned}$$

□

4 Цепи Маркова

4.1 Ветвящиеся процессы

Определение 19 (ветвящийся процесс Гальтона-Ватсона)**.** Пусть фиксирован набор вероятностей $\{p_n\}_{n=0}^{\infty}$, что $\sum_{n=0}^{\infty} p_n = 1$. Рассмотрим вероятностное пространство, в котором

- $\Omega := \mathbb{N}^{\mathbb{N} \times \mathbb{N}}$ — множество двумерных последовательностей с натуральными индексами с элементами из $\mathbb{N}$,
- $\mathcal{F}$ — минимальная сигма-алгебра, содержащая все $A_{n,m,c}$,
- $\mathbb{P}$ задаётся так, что $\mathbb{P}(A_{n,m,c}) = p_c$ для всех $n \in \mathbb{N}$, $s \in \{1; -1\}$,

где $A_{n,m,c}$ — множество последовательностей, у которых на месте (n, m) стоит c. Рассмотрим случайные величины

$$X_{n,m} = \begin{cases} c & \text{если } \omega \in A_{n,m,c} \\ \ldots, \end{cases}$$

и

$$M_0 = 1, \qquad M_{n+1} = \sum_{k=1}^{M_n} X_{n,k}.$$

Также обозначим φ — характеристическую функцию X.

Это можно интерпретировать следующим образом. Пусть у нас есть частица. В первый момент она делится на некоторое количество таких же частиц, равное случайной величине $X_{0,1}$. Во второй момент каждая k-ая частица делится на $X_{1,k}$ частиц; и так происходит далее. При этом все $X_{n,m}$ независимы и одинаково распределены. И таким образом M_n будет количеством частиц после момента n.

Лемма 39.

$$\varphi_n(z) := \varphi_{M_n}(z) = \varphi^n(z).$$

Доказательство. $\varphi_0(z) = \varphi^0(z)$ и

$$\varphi_{n+1}(z) = \varphi_{M_{n+1}}(z) = \varphi_{M_n}(\varphi(z)) = \varphi^n(z).$$

□

Определение 20.

$$R_n := \{M_n = 0\} \qquad \text{и} \qquad R := \bigcup_{n=0}^{\infty} R_n$$

— события вырождения процесса в момент n и вырождения (когда-нибудь), а

$$q_n := \mathbb{P}(R_n) \qquad \text{и} \qquad q := \mathbb{P}(R)$$

— вероятности вырождения процесса в момент n и вырождения (когда-нибудь).

Лемма 40.

1. $R_n \subseteq R_{n+1}$.
2. $q_n \leqslant q_{n+1}$.
3. $q = \lim_{n\to\infty} q_n$.

Теорема 41. *q является наименьшим корнем уравнения $\varphi(t) = t$ на $[0;1]$.*

Доказательство. Для начала заметим, что φ как функция $[0;1] \to [0;1]$ обладает следующими свойствами.

1. $\varphi(0) = p_0$ и $\varphi(1) = 1$.
2. φ монотонна
3. φ выпукла.
4. φ либо прямая, либо пересекает всякую прямую не более чем в двух точках.
5.
   (a) Если $\varphi \geqslant t$ и равенство достигается только в 1, то корнем будет только 1.
   (b) Если φ — прямая, не подходящая под первый пункт, то уравнения $\varphi(t) = t$ множеством корней будет весь $[0;1]$.
   (c) Если φ — не прямая, не подходящая под первый пункт, то у уравнения $\varphi(t) = t$ есть ровно один корень отличный от 1.
6. $\varphi(t) - t$ до первого корня положительна, а между корнями отрицательна.

Теперь заметим, что

$$q_0 = p_0 \qquad \text{и} \qquad q_{n+1} = \varphi^{n+1}(0) = \varphi(q_n).$$

Следовательно, поскольку φ непрерывна на $[0;1]$, переходя к пределу, получаем, что $\varphi(q) = q$. При этом так как $q_n \leqslant q_{n+1}$, то $\varphi(t) - t$ в точках q_n положительно, откуда следует, что q_n меньше наименьшего корня. Значит q является наименьшим корнем. □

Теорема 42. *TFAE*

1. $q < 1$.
2. *Либо* $\mathbb{E}(X) > 1$, *либо* $\mathbb{P}\{X = 1\} = 1$.

Доказательство.

$2 \Rightarrow 1$) Если $\mathbb{P}\{X = 1\} = 1$, то $\varphi(t) = t$. В таком случае $M_n = 1$ с вероятностью 1, а тогда $q = 0 < 1$.

Если же $\mathbb{E}(X) > 1$, то $\varphi'(1) = \mathbb{E}(X) > 1$, значит левее точки 1 локально функция $\varphi(t) - t$ отрицательна, откуда 1 — не минимальный корень. Т.е. $q < 1$.

$1 \Rightarrow 2$) Если $q < 1$, то либо $\varphi(t) = t$, т.е. $\mathbb{P}X = 1 = 1$, либо $\varphi(t) - t$ на промежутке $(q; 1)$ отрицательна и выпукла, т.е. $\varphi'(t)$ возрастает и следовательно $\varphi'(1) > 1$, откуда $\mathbb{E}(X) = \varphi'(1) = 1$.

□

Определение 21. Ветвящийся процесс называется

- *надкритическим*, если $\mathbb{E}(X) > 1$,
- *критическим*, если $\mathbb{E}(X) = 1$,
- *подкритическим*, если $\mathbb{E}(X) < 1$.

4.2 Цепи Маркова

Определение 22. Последовательность случайных величин $\{X_n\}_{n=0}^{\infty}$ называется *марковской цепью* (или *цепью Маркова*), если

1. $\{X_n\}_{n=0}^{\infty}$ независимы,
2. для всяких $\{v_k\}_{k=0}^{n+1}$, что $v_k \in X_k(\Omega)$,

$$\mathbb{P}\{X_{n+1} = v_{n+1} \mid X_n = v_n\} = \mathbb{P}\left\{X_{n+1} = v_{n+1} \mid \bigwedge_{k=0}^{n} X_k = v_k\right\}.$$

Если в цепи Маркова $X_n(\Omega)$ не зависит от n и $\mathbb{P}\{X_{n+1} = v_{n+1} \mid X_n = v_n\}$ не зависит от n, то такая цепь называется однородной.

Замечание. Мы будем рассматривать только однородные дискретные цепи Маркова (т.е. случайные величины дискретны). В таком случае будем считать, что $X_0(\Omega) = \mathbb{N}$, обозначать

$$P_{i,j} := \mathbb{P}\{X_{n+1} = j \mid X_n = i\}$$

и называть все $P_{i,j}$ вероятностями перехода. Также скажем, что $P_{i,j}$ образуют матрицу P. Также обозначим

$$p_i := \mathbb{P}\{X_0 = i\}$$

и порождённый ими вектор-строка $\bar{p}$ — вектор начального распределения.

Лемма 43. *Матрица P стохастическая справа, т.е.*

1. $P_{i,j} \geqslant 0$,

2. $\sum_j P_{i,j} = 1$.

Пример 10.

1. Если $\{X_n\}_{n=0}^{\infty}$ — независимые одинаково распределённые случайные величины, то $\{X_n\}_{n=0}^{\infty}$ — цепь Маркова, а
$$P_{i,j} = \mathbb{P}\{X_1 = j\} = p_j.$$
В таком случае в P все строки совпадают.

2. Если $\{S_n\}_{n=0}^{\infty}$ — случайное блуждание на $\mathbb{Z}$ с вероятностью перехода p, то оно — цепь Маркова, а
$$P_{i,j} = \mathbb{P}\{X_{n+1} = j \mid X_n = i\} = \begin{cases} p & \text{если } j = i+1, \\ 1-p & \text{если } j = i-1, \\ 0 & \text{иначе.} \end{cases}$$

Лемма 44.
$$\mathbb{P}\left\{\bigwedge_{k=0}^{n} X_n = i_n\right\} = p_i \cdot \prod_{k=1}^{n} p_{i_{k-1}, i_k}.$$

Теорема 45 (частный случай теоремы Колмогорова о существовании случайного процесса с заданными конечномерными распределениями)**.** *Пусть заданы вектор $\bar{p} = (p_n)$ и стохастическая справа матрица P. Тогда существует вероятностное пространство и цепь Маркова в нём с данными начальным распределением и матрицей переходов.*

4.3 Вероятность перехода за n шагов

Лемма 46.

1. *Матрица P^m определена корректно и является матрицей перехода за m шагов, т.е. для всех n.*
$$\mathbb{P}\{X_{n+m} = j \mid X_n = i\} = (P^m)_{i,j}.$$

2. *Вектора-строка $\bar{p} \cdot P^n$ определён корректно и является вектором распределения вероятностей в момент n, т.е.*
$$\mathbb{P}\{X_n = j\} = (\bar{p}P^n)_j.$$

Определение 23. Распределение вероятностей $\bar{\pi}$ называется *стационарным* (*инвариантным*), если
$$\bar{\pi} \cdot P = \bar{\pi}.$$

Теорема 47 (Маркова о финальных вероятностях)**.** *Рассмотрим цепь $\{X_n\}_{n=0}^{\infty}$ Маркова с конечным числом значений m. Пусть известно, что существует $n_0 \in \mathbb{N}$ и $\delta \in (0;1)$, что для всех i, j*
$$P_{i,j}(n_0) \geqslant \delta.$$
Тогда существует $C > 0$, $\rho \in (0;1)$ и стационарный вектор $\bar{\pi} = (\pi_j)$, что
$$\max_{i,j} |p_{i,j}(n) - \pi_j| \leqslant C\rho^n.$$

Доказательство.

Написать, как применяется теорема о сжимающем отображении.

Получили $\bar{\pi}$. Подставляя в качестве $\bar{x}_i = (P_{i,j})_{j=1}^{m}$, получаем требуемую сходимость.

Написать переход от $n_0 = 1$ к общему случаю.

$\square$

4.4 Классификация состояний

Определение 24. Состояние j называется *достижимым* из i, если есть $n \in \mathbb{N}$, что $p_{i,j}(n) > 0$. Обозначение: $i \to j$.

Состояния i и j *сообщаются*, если $i \to j$ и $j \to i$. Обозначение: $i \leftrightarrow j$.

Состояние i *существенно*, если из $i \to j$ следует, что $j \to i$.

Определение 25. Будем обозначать:

1. E — множество всех состояний.

2. S — множество существенных состояний.

3. N — множество несущественных состояний.

4. E_i — эргодические классы, т.е. классы эквивалентности по отношению эквивалентности $\leftrightarrow$ на S.

Определение 26. Для всякого состояния i обозначим

$$L_i := \{n \geqslant 0 \mid p_{i,i}(n) > 0\} \qquad \text{и} \qquad d(i) := \text{НОД}(L_i).$$

$d(i)$ называется периодом состояния i.

Лемма 48. *Если $i \leftrightarrow j$, то $d(i) = d(j)$.*

Доказательство. Понятно, что есть такие k и l, что

$$p_{i,j}(k) > 0 \qquad \text{и} \qquad p_{j,i}(l) > 0.$$

Тогда

$$p_{i,i}(n) \geqslant p_{i,j}(k) \cdot p_{j,j}(n) \cdot p_{j,i}(l).$$

Пусть $\{n_t\}_{t=0}^m$ — конечное подмножество $L(j)$, что НОД$(\{n_k\}_{k=0}^m) = d(j)$. Значит

$$\{k+l\} \cup \{k+l+n_k\}_{k=0}^m \subseteq L(i),$$

откуда сразу следует, что $d(i) \mid d(j)$. Аналогично $d(j) \mid d(i)$. Значит $d(i) = d(j)$. $\square$

Следствие 48.1. *В одном эргодическом классе все периоды одинаковы. Этот период называется* периодом класса*.*

Теорема 49. *Пусть E — эргодический класс с периодом d. Тогда есть его разбиение $E = \bigsqcup_{j=0}^{d-1} C_j$, что для всякого $i \in C_k$*

$$\sum_{j \in C_{k+1}} p_{i,j} = 1.$$

Теорема 50 (Колмогоров, анонс)**.** *Пусть цепь маркова имеет единственный эргодический класс, и порядок этого класса — 1. Тогда существуют пределы*

$$\pi_j = \lim_{n \to \infty} p_{i,j}(n)$$

(т.е. для всяких фиксированных i и j предел существует, а его значение не зависит от i). Тогда возможны 2 случая:

1. *$\pi_j = 0$ для всех j (в таком случае цепь называется* нулевой*).*

2. *$\pi_j > 0$ для всех j и $\sum_j \pi_j = 1$, $\bar{\pi}$ — единственный стационарный вектор (в таком случае цепь называется* положительной*).*

Следствие 50.1. *Если в терминах теоремы цепь конечна, то она положительна.*

Доказательство.
$$\sum_j \pi_j = \lim_{n\to\infty} \sum_j p_{i,j}(n) = \lim_{n\to\infty} 1 = 1.$$
Следовательно нулевой цепь быть не может. □

Теорема 51 (анонс)**.** *Пусть дана цепь Маркова с единственным эргодическим классом, его период — 1, и цепь положительна. Обозначим за η_j случайную величину времени первого возвращения в j. Тогда*
$$\pi_j = \frac{1}{\mathbb{E}(\eta_j)}.$$

4.4.1 Возвратность

Обозначим
$$A_k := \left\{ A_k = j \wedge \bigwedge_{t=1}^{k-1} A_t \neq j \right\} \qquad \text{и} \qquad f_{j,j}(k) = \mathbb{P}(A_k \mid X_0 = j).$$
Тогда
$$f_j := \sum_{k=1}^{\infty} f_{j,j}(k) = \mathbb{P}\{\text{“хотя бы один раз вернуться в } j\text{”} \mid X_0 = j\} \leqslant 1.$$

Определение 27. Состояние j возвратно, если $f_j = 1$.

Лемма 52. *Для всякого $n \geqslant 1$*
$$p_{j,j}(n) = \sum_{k=1}^{n} f_{j,j}(k) p_{j,j}(n-k)$$

Следствие 52.1. *Пусть F и P — производящие функции $\{f_{j,j}(n)\}_{n=0}^{\infty}$ и $\{p_{j,j}(n)\}_{n=0}^{\infty}$. Тогда*
$$G(t) = \frac{1}{1 - F(t)}.$$

Следствие 52.2.
$$\sum_{n=0}^{\infty} = \frac{1}{1 - f_j}.$$

Следствие 52.3.
$$f_j = 1 \qquad \iff \sum_{n=0}^{\infty} p_{j,j}(n) = \infty.$$

Следствие 52.4. *Пусть i возвратно и $i \leftrightarrow j$. Тогда j возвратно.*

Следствие 52.5. *Возвратность — свойство эргодического класса.*