ДТВ.билеты 2 семестр

Кабашный Иван (@keba4ok)
Горбунов Леонид, Савельев Артём
на основе лекций Ю. А. Давыдова
и других материалов

28 мая 2021 г.

Они самые.

Содержание

1 Билеты

1.1 Вероятностное пространство. Аксиомы вероятности. Дискретное вероятностное пространство. Схема равновозможных исходов.

Определение 1. Ω - *пространство элементарных событий* или *множество элементарных исходов*, есть множество, состоящее из ω_i, *элементарных событий*. Нам важно лишь, чтобы это множество было непустым. $\mathcal{F} \subseteq \mathcal{P}(\Omega)$ - некоторая совокупность подмножеств Ω, есть *множество событий*, элементы которого есть A_i - события.

Определение 2. $\mathbb{P}$ - вероятность $A \Rightarrow \mathbb{P}(A)$ - вероятность события A.

Определение 3. Вся же тройка $(\Omega, \mathcal{F}, \mathbb{P})$ называется *вероятностным пространством*.

Для вероятностей существует несколько аксиом:

- $0 \leq \mathbb{P}(a) \leq 1$ для любого события,
- $\mathbb{P}(\Omega) = 1$,
- для любого счётного набора попарно непересекающихся события $\{A_i\}_{i \in N} \subseteq \mathcal{F}$ выполнена *счётная аддитивность*:

$$\mathbb{P}\left(\bigsqcup_{n \in N} A_n\right) = \sum_{n \in N} \mathbb{P}(A_n).$$

Некоторые свойства вероятностей:

- $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$;
- $\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$;
- $\forall A, B \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;
- $\mathbb{P}(\bigcup_n A_n) \leq \sum_n \mathbb{P}(A_n)$.

Теперь перейдём к некоторым примерам вероятностных пространств:

Пример(ы) 1. Пространство $(\Omega, \mathcal{F}, \mathbb{P})$ *дискретно*, если Ω не более, чем счётно. $\mathcal{F} = \mathcal{P}(\Omega)$, элементы $\{\omega\}$ также считаем событиями.

Утверждение 1. Несколько предложений:

- Пусть $\mathbb{P}$ - вероятность в $(\Omega, \mathcal{F}, \mathbb{P})$. Тогда $\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$, где $p_\omega = \mathbb{P}\{\omega\}$. При этом $p_\omega \geq 0$, $\sum_\omega p_\omega = 1$.
- Предположим, что $\{p_\omega\}_{\omega \in \Omega}$ такие, что выполнено последнее предложение предыдущего пункта, тогда $\mathbb{P}(A) = \sum_{\omega \in A} p_\omega$ - вероятность.

Также, можно упомянуть про *равновероятные исходы*, из названия понятно, что это. Если $|\Omega| < \infty$ и $p_\omega = p$ для любого ω, тогда $\mathbb{P}(A) = \frac{|A|}{|\omega|}$.

1.2 Условная вероятность. Формула полной вероятности. Формула Байеса.

Определение 4. *Условная вероятность*: $\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ (при $\mathbb{P}(B) > 0$).

Утверждение 2. Для условной вероятности выполнены аксиомы вероятности.

А теперь - несколько утверждений, которые касаютсся условной вероятности.

Утверждение 3. (B_n) - *разбиение* Ω (дизъюнктный набор, который в объединении даёт всё множество). Тогда для любого A, $\mathbb{P}(A) = \sum_k \mathbb{P}(B_k)\,\mathbb{P}_{B_k}(A)$.

Доказательство.

$$\mathbb{P}(A) = \mathbb{P}(A \cap \Omega) = \sum_n \mathbb{P}(A \cap (\bigcup_n B_n)) = \mathbb{P}(\bigcup_n (A \cap B_n)) = \sum_n \mathbb{P}(A \cap B_n) = \mathbb{P}(B_n) \cdot \mathbb{P}_{B_n}(A).$$

□

Утверждение 4. *Формула Байеса*. Пусть мы знаем событие A, имеется разбиение (B_n), тогда

$$\mathbb{P}_A(B_k) = \frac{\mathbb{P}(A \cap B_k)}{\mathbb{P}(A)} = \frac{\mathbb{P}(B_k)\,\mathbb{P}_{B_k}(A)}{\sum_n \mathbb{P}(B_n)\,\mathbb{P}_{B_n}(A)}$$

Доказательство. Из определения условной вероятности,

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

Тогда просто применим формулу полной вероятности, и заменим $\mathbb{P}(A)$ на сумму. □

1.3 Независимость событий.

Примечание 1. На всякий случай допвопросов, я вставил некоторые рассуждения с лекции.

Перейдём к *независимости событий*. Начнём рассуждения с двух событий: A и B. Если $\mathbb{P}(B) = \mathbb{P}_A(B)$, $\mathbb{P}(A) = \mathbb{P}_B(A)$, или, что равносильно им обоим $\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B)$, то события называются *независимыми*.

Пусть теперь имеется не два, а больше событий $\{A_1, \ldots, A_n\}$. Нельзя сказать, что нам хватает попарной независимости для независимости совокупной.

Пример(ы) 2. (Пирамида Бернштейна). Рассмотрим тетраэдр, у которого стороны покрашены таким образом: белый, синий, красный и флаг России. Рассматриваем события: A_1 - на выпавшем основании есть белый цвет, и так далее A_2 и A_3. Эти события попарно независмы, но не независимы в совокупности.

$$\mathbb{P}(A_i) = \frac{1}{2},\ \mathbb{P}(A_1 \cap A_2) = \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2} = \mathbb{P}(A_1)\,\mathbb{P}(A_2),$$

но тогда

$$\mathbb{P}((A_1 \cap A_2) \cap A_3) = \frac{1}{4} \neq \frac{1}{8}.$$

Таким образом, нужно ввести корректное определение.

Определение 5. События $A_1, \dots, A_n$ *независимы*, если выполнено:

$$\mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\,\mathbb{P}(A_j),\ \forall i \neq j,$$
$$\mathbb{P}(A_i \cap A_j \cap A_k) = \mathbb{P}(A_i)\,\mathbb{P}(A_j)\,\mathbb{P}(A_k),\ \forall i \neq j \neq k,$$
$$\dots$$
$$\mathbb{P}(\bigcap_1^n A_i) = \prod_1^n \mathbb{P}(A_i).$$

Теорема 1. *Пусть имеется* $T_1, \dots, T_m$ *- разбиение* $\{1, \dots, n\}$*, независимые события* $A_1, \dots, A_n$*,* $\{B_j\}_m$ *- комбинация (всякие действия между элементами) событий* $\{A_s, s \in T_j\}$*. Тогда* $\{B_j\}$ *- независимы.*

Доказательство. По индукции. $\square$

1.4 Дискретные случайные величины. Их распределения. Примеры распределений: Бернулли, биномиальное, геометрическое, распределение Пуассона. Независимость случайных величин.

Определение 6. *Случайная величина* - это функция $X : \Omega \to R$.

Пример(ы) 3. Число выпавших решек на n бросках.

Теперь немного о *распределении случайной величины*. Пусть имеется вероятностное пространство и случайная величина X. Нас интересует $\{\omega | X(\omega) \in B \subseteq \mathbb{R}\}$, то есть, мы хотим исследовать попадания случайной величины в те или иные зоны на прямой. Такую вероятность можно рассматривать как вероятность от множества B, но это слишком сложно, поэтому продолжим на таких двух пунктах:

- значения X, $X(\Omega) = \{a_1, \dots\}$, $\{a_k\}$ - значение X,
- $A_k = \{\omega | X(\omega) = a_k\}$; $p_k = \mathbb{P}(A_k)$, причём каждая $p_k \geq 0$, а их сумма равна единице.

Тогда мы можем сделать вывод, что $\mathbb{P}\{X \in B\} = \sum_{k | a_k \in B} p_k$, так как левая часть есть $\mathbb{P}\{\bigcup_{k|a_k \in B}\}$, что равно $\mathbb{P}\{\bigcup_{k|a_k \in B}\{x = a_k\}\} = \sum_{k|a_k \in B} \mathbb{P}\{x = a_k\}$, что уже и равно левой части.

Определение 7. Таким образом, совокупность последовательностей $\frac{\{a_k\}}{\{p_k\}}$ и называется *распередением случайной величины*.

Пример(ы) 4. Приведём примеры распределений:

- *вырожденное*: $X(\omega) = a$ для любого ω.
- *распределение Бернулли*: $B(1, p)$, $p \in [0{,}1]$, причём единица принимается с вероятностью p, 0 - иначе.
- *биномиальное*: $B(n, p)$, $p \in [0{,}1]$, $X \sim B(n, p)$, если принимаются значения от 0 до n, причём $\mathbb{P}\{X = k\} = \mathbb{P}_n(k)$ (просто обозначение), и равно $C_n^k p^k (1-p)^{n-k}$.
- *геометрическое* $X = 1, 2, \dots$, $p \in [0{,}1]$. $\mathbb{P}\{X = k\} = (1-p)^{k-1} p$.
- *пуассоновское* $X \sim \mathcal{P}(\alpha), \alpha > 0$, $X = 0, 1, \dots$, $\mathbb{P}\{X = k\} = \frac{\alpha^k}{k!} e^{-\alpha}$.

Независимость случайных величин.

Определение 8. X и Y - *независимые случайные величины*, если $\forall A, B \subset \mathbb{R}^1$ события $\{X \in A\}$, $\{Y \in B\}$ независимы.

Чуть позже будет критерий независимости, потому что по определению проверить зачастую слишком проблематично.

Определение 9. $X_1, X_2, \ldots, X_n$ - назависимые случайные величины, если $\forall A_1, \ldots, A_n \subset \mathbb{R}^1$, $\{X_i \in A_i\}$ независимы.

Теорема 2. *(Критерий независимости случайных величин). Предположим, что X_k имеет распределение $(a_{kj})_j$, (p_{kj}), $k = 1, \ldots, n$. Тогда X_i назависимы тогда и только тогда, когда*

$$\mathbb{P}\{X_1 = a_{1j_1}, X_2 = a_{2j_2}, \ldots, X_n = a_{nj_n}\} = \prod_{k=1}^{n} p_{kj_k}$$

Доказательство. Из первого во второе - очевидно. В обратную сторону для начала докажем такой факт: если $A_1, \ldots, A_n \subset \mathbb{R}^1$, то тогда

$$\mathbb{P}\{X_1 \in A_1, X_2 \in A_2, \ldots, X_n \in A_n\} = \prod_{k=1}^{n} \mathbb{P}\{X_k \in A_k\}.$$

Докажем все факты для двух, обобщается это всё не сложно.

$$\mathbb{P}\{X_1 \in A_1, X_2 \in A_2\} = \mathbb{P}\Bigg(\bigcup_{j,s|a_{1j}\in A_1, a_{2s}\in A_2} \bigcap \{X_1 = a_{1j}, X_2 = a_{2s}\}\Bigg),$$

Но вероятность объединения можно заменить на сумму вероятностей:

$$\sum_{-//-} \mathbb{P}\{X_1 = a_{1j}, X_2 = a_{2s}\} = \sum_{-//-} p_{1j}p_{2s} = \Bigg(\sum_{j|a_{1j}\in A_1} p_{1j}\Bigg)\Bigg(\sum_{s|a_{2s}\in A_2} p_{2s}\Bigg),$$

что равно

$$\mathbb{P}\{X_1 \in A_1\}\,\mathbb{P}\{X_2 \in A_2\}.$$

Это как раз и даёт нам независимость через немного хитрую вещь. Нам же нужна независимость для любого количества элементов, но казалось бы, у нас есть только для всех X_i. Однако, в лемме можно заменить некоторые A_l на $\mathbb{R}^1$, и тогда выражение как раз показывает нам необходимые соотношения на оставшиеся величины. □

1.5 Испытания Бернулли. Вероятность для числа успехов в схеме Бернулли. Пуассоновское приближение в схеме Бернулли.

Определение 10. Рассмотрим последовательность ε_k - независимых бернуллиевских величин (два исхода), $\varepsilon_k \sim B(1, p)$,

$$\varepsilon_k = \begin{cases} 1, \; p \\ 0, \; q = 1 - p \end{cases}$$

$\{\varepsilon_k = 1\}$ означает успех в каком-то испытании. То есть, *испытание Бернулли* - последовательность из однотипных испытаний.

ν_n - число успехов в n испытаниях., что равно $\sum_{k=1}^n \varepsilon_k$. Покажем, что распределение биномиально. $\nu_n = 0, 1, \ldots, n$.

$$C_n^k p^k q^{n-k} =^? \mathbb{P}\{\nu_n = l\} = \mathbb{P}\{\omega| \text{ в цепочке } \varepsilon_1(\omega), \ldots, \varepsilon_n(\omega) \text{ в точности } k \text{единиц}\} =$$
$$= \mathbb{P}\{\sum_1^n \varepsilon_j = k\} = \sum_{\{i_1,\ldots,i_n\}| \text{ в } \bar{i}\ k \text{ единиц}} \mathbb{P}\{\varepsilon_1 = i_1, \ldots, \varepsilon_n = i_n\} =$$
$$= \sum_{-//-} p^k q^{n-k} = C_n^k p^k q^{n-k}.$$

Теперь встаёт вопрос: существуют ли независимые бернуллиевские случайные величины. Пусть $\Omega = \{(i_1, i_2, \ldots, i_n)|i_k = 0 \text{ или } 1\}$; $|\Omega| = 2^k$. $p_\omega = p^k(1-p)^{n-k}$ для $\omega = (i_1, \ldots, i_n)$, в которой ровно k единиц. $\sum p_n = 1$, $\varepsilon_k(\omega) = \varepsilon_k(i_1, \ldots, i_n) = i_k$.

$P_n(k) = \mathbb{P}\{\nu_n = k\} = C_n^k p^k q^{n-k}$ - можем ли мы как-то упростить это?

Теорема 3 (*Пуассоновская аппроксимация*)**.** *Пусть проводятся n испытаний Бернулли с веростностями успеха $p = p_n$, $n \cdot p_n \to_{n\to\infty} \alpha$, тогда $\forall k \in \mathbb{N}$ $P_n(k) \to \pi_k = \frac{\alpha^k}{k!}e^{-\alpha}$.*

Доказательство.

$$P_n(k) = \frac{n!}{k!(n-k)!} p^k q^{n-k} = \frac{1}{k!} \frac{n(n-1)\ldots(n-k+1)}{n^k} (np)^k q^{n-k} =$$
$$= \frac{1}{k!}(1 - \frac{1}{n}) \ldots (1 - \frac{k-1}{n})(np)^k q^{n-k} \sim \frac{\alpha^k}{k!} q^{n-k}$$

Если что, k - константа.

$$q^{n-k} = (1-p)^{n-k} = (1 - \frac{np_n}{n})^n (1 - \frac{np_n}{n})^{-k}$$
$$\sim (1 - \frac{np_n}{n})^n \sim e^{-np_n} = e^{-\alpha}.$$

В итоге, $P_n(k) \sim \frac{\alpha^k}{k!}e^{-\alpha}$.

□

А теперь - несколько слов о точности аппроксимации. Оценим $|P_n(k) - \pi_k|$. $\pi_k = \frac{(np)^k}{k!}e^{-np}$.

Теорема 4. *Если $X \sim B(n, p)$, $Y \sim \mathcal{P}$, то $\forall A \subset \mathbb{R}$, $|\mathbb{P}\{X \in A\} - \mathbb{P}\{Y \in A\}| \leq 2np^2$.*

Доказательство. (Более наглядно его можно посмотреть с самого начала третьей лекции, от 02.03.2021.)
Построим пространство Ω с такими с.в. $\varepsilon_1, \ldots, \varepsilon_n, \eta_1, \ldots, \eta_n$, чтобы:
они были независимы;
$\varepsilon_k \sim B(1, p)$, $\eta_k \sim \mathcal{P}$;
$\mathbb{P}\{\varepsilon_k \neq \eta_k\} \leq p$.
Это сделаем так:

Шаг 1. При $n = 1$.
$\Omega_1 := \mathbb{N}$, вероятности состояний:
$q_1(1) := 1 - p, q_1(2) := \pi_0 - (1-p), q_1(k) := \pi_{k-2}$ при $k > 2$, где $\pi_k = \frac{p^k}{k!}e^{-p}$

$\varepsilon_1(1) = 0,\ \varepsilon_1(k) = 1$ при $k > 2$,
$\eta_1(1) := 0, \eta_1(2) := 0,\ \eta_1(3) := 1,\ \eta_1(4) := 2, \dots, \eta_1(k) := k - 2$
Проверяем, что всё хорошо: сумма вероятностей равна 1, так как $q_1 + q_2 = \pi_0$, а $\sum_k \pi_k = 1$; $\mathbb{P}\{\varepsilon_1 = 0\} = q_0 = 1 - p$, то есть действительно $\varepsilon_1 \sim B(1,p)$; $\mathbb{P}\{\eta_1 = 0\} = q_0 + q_1 = \pi_0$, $\mathbb{P}\{\eta_1 = k\} = q_{k+2} = \pi_k$, то есть действительно $\eta_1 \sim \mathcal{P}$;
$\mathbb{P}\{\varepsilon_1 \neq \eta_1\} = 1 - q_1 - q_3 = 1 - (1-p) - pe^{-p} = p(1 + e^{-p}) \leq p^2$.

Шаг 2. Получаем много переменных ε_k и η_k.
Сделаем копии нашего Ω_1 вместе с вероятностями и переменными, назовём их $\Omega_1, \Omega_2, \dots, \Omega_n$; ε_k, η_k. И рассмотрим пространство — прямое произведение омег:

$$\Omega = \Omega_1 \times \Omega_2 \times \dots \times \Omega_n$$

Тогда вероятность состояния $\overline{\omega} = (\omega_1, \dots, \omega_n)$ равна $p_{\overline{\omega}} = p_{\omega_1} \cdot \dots \cdot p_{\omega_n}$, случайные величины определяются так: $\tilde{\varepsilon}_k(\overline{\omega}) = \varepsilon_k(\omega_k)$, $\tilde{\eta}_k$ — анологично. Упражнение: доказать, что для этих случайных величин всё хорошо, и они независимы.
Замечание: $\mathbb{P}_k\{S_n \neq S\} = \mathbb{P}\{S_n \neq S\}$, док-во см. на 38:40 в лекции.

Шаг 3. Обозначим $S_n = \tilde{\varepsilon}_1 + \tilde{\varepsilon}_2 + \dots + \tilde{\varepsilon}_n$, $S_n \sim B(n,p)$, $S = \tilde{\eta}_1 + \tilde{\eta}_2 + \dots + \tilde{\eta}_n$, $S \sim (P)(np)$. Тогда легко проверить, что

$$\mathbb{P}\{S_n \neq S\} \leq \mathbb{P}\{\cup_k \{\tilde{\epsilon}_k \neq \tilde{\eta}_k\}\} \leq \sum_{k=1}^{n} \mathbb{P}\{\tilde{\epsilon}_k \neq \tilde{\eta}_k\} \leq np^2$$

И, наконец,

$$\begin{aligned} \mathbb{P}\{S_n \in B\} = \mathbb{P}\{S_n \in B, S_n = S\} + \mathbb{P}\{S_n \in B, S_n \neq S\} \leq \\ \mathbb{P}\{S_n \neq S\} + \mathbb{P}\{S_n \in B, S_n = S\} \\ \leq np^2 + \mathbb{P}\{S \in B\} \end{aligned}$$

Аналогично

$$\mathbb{P}\{S \in B\} \leq np^2 + \mathbb{P}\{S_n \in B\}$$

Откуда $|\mathbb{P}\{X \in B\} - \mathbb{P}\{Y \in B\}| \leq 2np^2$, ч.т.д. □

Следствие 1. $\sum_{k=0}^{\infty} |P_n(k) - \pi_n| \leq 2np^2 = O(\frac{1}{n})$.

Доказательство. Обозначим за $B_+ = \{k | P_n(k) - p_k \geq 0\}$, а за $B_- = B_+^c$. Тогда

$$\begin{aligned} \sum_{k=0}^{\infty} |P_n(k) - \pi_n| = \sum_{k \in B_+} (P_n(k) - \pi_n) + \sum_{k \in B_-} (\pi_n - P_n(k)) = \\ (\mathbb{P}\{S_n \in B_+\} - \mathbb{P}\{S \in B_+\}) + (\mathbb{P}\{S \in B_-\} - \mathbb{P}\{S_n \in B_-\}) \leq 2p^2 n \end{aligned}$$

□

1.6 Математическое ожидание и его свойства. Формула перемножения математических ожиданий для независимых величин. Примеры вычисления математического ожидания для ...

Определение 11. Пусть X - случайная величина, тогда *математическое ожидание* - $EX = \sum_{\omega \in \Omega} X(\omega) p_\omega$, если ряд абсолютно сходится в бесконечном случае.

Свойства:

- Существование математического ожидания эквивалентно существованию математического ожидания $|X|$ (для краткости пишут $E|X| < \infty$).
- Математическое ожидание линейно.
- Если $X \geq 0$, то $EX \geq 0$.
- $X \leq Y$, тогда $EX \leq EY$.
- $|EX| \leq E|X|$ (из $-|X| \leq X \leq |X|$).

А вот это уже можно обобщить до неравенства Йенсена:

Теорема 5. *Если φ - выпуклая функция, тогда $\varphi(EX) \leq E\varphi(X)$.*

Однако доказывать мы это не будем, а просто перейдём к следующей теореме:

Теорема 6. *Пусть X - случайная величина $(a_k);(p_k)$, $f : \mathbb{R}^1 \to \mathbb{R}^1$. Тогда $Ef(x) = \sum_k f(a_k)p_k$ (существуют или не существуют они одновременно).*

Доказательство. Скажем для начала, что $f \geq 0$ (просто возьмём модуль, если что). Тогда $Ef(X) = \sum_{\omega\in\Omega} f(X(\omega))p_\omega = \sum_k \left(\sum_{\omega|X(\omega)=a_k} f(X(\omega))p_\omega \right) = \sum_k f(a_k) \sum_{\omega|X(\omega)=a_k} p_\omega = \sum_k f(a_k)p_k$. □

Отсюда есть несколько следствий:

- $EX = \sum_k a_k p_k$, $f(x) = x$.
- Если X и Y имеют одинаковые распределения, то $EX = EY$.
- EX - центр масс.

Теорема 7. *Пусть $\bar{X}$ - случайный вектор; $(\bar{a}_k);(p_k)$; $f : \mathbb{R}^n \to \mathbb{R}^1$. Тогда $Ef(\bar{X}) = \sum_k f(\bar{a}_k)p_k$.*

Теорема 8. *Если X и Y независимы, и их математические ожидания существуют, то у их произведения также существует матожидание, притом равно произведению матожиданий.*

Доказательство. Применим в предыдущей теореме $f : \mathbb{R}^2 \to \mathbb{R}$ по правилу $f(x,y) = xy$ и распишем $E(XY)$. □

Пример(ы) 5. Пусть $X \sim B(n,p)$ имеет биномиальное распределение. Тогда $EX = \sum_0^n k \cdot C_n^k p^k (1-p)^{n-k} = np$.

К такому выводу можно прийти и другим способом. Мы знаем, что случайная величина X имеет такое же распределение, как сумма бернуллиевских величин, одинаково распределённых: $X =^D \varepsilon_1 + \ldots + \varepsilon_n$, (ε_k) независимы, $\varepsilon_k \sim B(1,p)$. $EX = \sum_1^n E\varepsilon_k = nE\varepsilon_k = np$.

Пример(ы) 6. Пусть $X \sim \mathcal{P}(\alpha)$, тогда

$$EX = \sum_0^\infty k\frac{\alpha^k}{k!}e^{-\alpha} = e^{-\alpha}\sum_0^\infty \frac{\alpha^k}{(k-1)!} = \alpha e^{-\alpha}\sum_1^\infty \frac{\alpha^{k-1}}{(k-1)!} = \alpha.$$

1.7 Дисперсия и ее свойства. Формула сложения дисперсий для независимых величин. Примеры вычисления дисперсии для стандартных распределений . . .

Определение 12. *Дисперсия* - $DX = E[(X - EX)^2]$. Нужно, чтобы $E(X^2)$ было меньше бесконечности, тогда DX определена.

Примечание 2. Дисперсия характеризует меру разброса случайной величины от математического ожидания.

И рассмотрим её свойства:

- Дисперсия неотрицательная.
- $D(X + c) = DX$.
- $D(X \cdot c) = c^2 DX$.
- $DX = E(X^2) - (EX)^2$.
- $DX = \sum_k (a_k - EX)^2 p_k$, где $(a_k), (p_k)$ - распределение X.

Теорема 9. *Если X, Y - независимые случайные величины, то $D(X + Y) = DX + DY$.*

Доказательство. Для начала рассмотрим случай $EX = EY = 0$, он достаточно очевиден, а теперь начнём сводить все остальные случаи к этому. Введём $\tilde{X} = X - EX$; $\tilde{Y} = Y - EY$. Тогда $D(X + Y) = D(\tilde{X} + \tilde{Y}) = DX + DY$, что и требовалось. □

Рассмотрим теперь несколько примеров дисперсий.

Пример(ы) 7. Пусть $X \sim B(n, p)$, тогда $DX = np(1 - p)$.

Пример(ы) 8. Пусть $X \sim \mathcal{P}(\alpha)$, тогда $DX = \alpha$.

Теорема 10. *(Неравенство Маркова). Пусть $X \geq 0$. Тогда $\forall t > 0$, $\mathbb{P}\{X \geq t\} \leq \frac{EX}{t}$.*

Доказательство. Введём понятие *индикатора*, то есть, функции $\mathbb{I}_A(\omega)$, котораяа равна 1, если $\omega \in A$, и 0 иначе (в нашем случае собыние A - $\geq t$). Эта случайная величина имеет Бернуллиевское распределение, 1 с вероятностью $p = \mathbb{P}(A)$, и 0 с вероятностью $1 - p$. Ясно, что левую часть можно записать как $E\mathbb{I}_A$. Посмотрим теперь на $\frac{X}{t}$. Эта вещь не меньше 1 при $\omega \in A$, тогда $\mathbb{I}_A \leq \frac{X}{t}$, а значит, $E\mathbb{I}_A \leq E\left(\frac{X}{t}\right) = \frac{EX}{t}$. □

Это неравенство имеет кучу следствий, рассмотрим некоторые из них.

Следствие 2. Пусть $f \uparrow$, $f(x) \geq 0$. Тогда $\mathbb{P}\{X \geq t\} = \mathbb{P}\{f(X) \geq f(t)\} \leq \frac{Ef(X)}{f(t)}$.

Следствие 3. Если $E(X^2) \leq \infty$, и $f(x) = x^2$, то, как следствие, получим $\mathbb{P}\{|X| \geq t\} \leq \frac{E(X^2)}{t^2}$, тогда при $t \to \infty$ получим более сильную оценку (на месте X^2 можно брать ещё большие степени). Также можно рассмотреть $f(x) = e^{ax}$, $a > 0$, тогда получим $\mathbb{P}\{x \geq t\} \leq \frac{E(e^{aX})}{e^{at}}$.

Теорема 11. *(Неравенство Чебышёва). Пусть X - случайная величина, $EX^2 < \infty$ (DX определена). Тогда $\forall t > 0$,*

$$\mathbb{P}\{|Y| \geq t\} \leq \frac{E(Y^2)}{t^2},$$

при $Y = X - \mathbb{E}X$.

Доказательство. Из следствия 2, при подстановке $(X - EX)$. □

Определение 13. *Моменты* - матожидания следующего вида. Пусть X - случайная величина. Мы можем рассмотреть $E(X^n)$, $E(X-EX)^n$, $n \in \mathbb{N}$, $E|X|^p$, $E|X-EX|^p$, $p \in \mathbb{R}$. Все эти матожидания называются моментами, первые два называются моментами моментами n-го порядка, вторые два - *абсолютными моментами*, моменты с разностями называются *центрированными*.

Определение 14. Если случайных величин несколько, то вводятся *смешанные моменты*, моменты вида $E(X^n \cdot Y^m)$, или $E[(X-EX)(Y-EY)] = \operatorname{cov}(X,Y)$ - частный случай, *ковариация*. Можем также через него определить *коэффициент корреляции* - $\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{DX \cdot DY}}$, который является мерой линейной зависимости между X и Y.

1.8 Случайные векторы. Ковариация. Коэффициент корреляции.

Определение 15. *Случайный вектор* - вектор, который состоит из случайных величин?

Определение 16. Если случайных величин несколько, то вводятся *смешанные моменты*, моменты вида $E(X^n \cdot Y^m)$, или $E[(X-EX)(Y-EY)] = \operatorname{cov}(X,Y)$ - частный случай, *ковариация*. Можем также через него определить *коэффициент корреляции* - $\rho(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{DX \cdot DY}}$, который является мерой линейной зависимости между X и Y.

1.9 Закон больших чисел (ЗБЧ). Пример применения ЗБЧ: теорема Вейерштрасса.

Теорема 12. *(ЗБЧ Чебышева). Пусть $\{X_j\}_{j\geq 1}$ - последовательность некоррелированных с. в. с равномерно ограниченными дисперсиями:*

$$\sigma^2 := \sup_j \mathbb{D}\, X_j < \infty.$$

Тогда случайные величины

$$Z_n := \frac{\sum_{j=1}^n X_j}{n} - \frac{\sum_{j=1}^n \mathbb{E}\, X_j}{n}$$

удовлетворяют соотношению

$$\lim_{n\to\infty} \mathbb{P}\{|Z_n| \geq r\} = 0, \ \ \forall r > 0.$$

Доказательство. Имеем $\mathbb{E}\, Z_n = 0$ и

$$\mathbb{D}\, Z_n = \mathbb{D}\, \frac{\sum_{j=1}^n X_j}{n} = \frac{\sum_{j=1}^n \mathbb{D}\, X_j}{n^2} \leq \frac{\sigma^2}{n}.$$

В соответствии с неравенством Чебышева,

$$\mathbb{P}\{|Z_n| \geq r\} \leq \frac{\mathbb{D}\, Z_n}{r^2} \leq \frac{\sigma^2}{nr^2} \to 0,$$

при $n \to \infty$. □

Следствие 4. $\{X_j\}_{j\geq 1}$ - последовательность одинаково распределённых независимых с.в. с конечной дисперсией. Пусть a - их общее математическое ожидание. Тогда

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{\sum_{j=1}^n X_j}{n} - a\right| \geq 0\right) = 0, \ \ \forall r > 0.$$

Теорема 13. *(ЗБЧ Бернулли). Пусть S_n - число успехова в схеме Бернулли из n испытаний с вероятностью успеха p. Тогда*

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq r\right) = 0, \ \ \forall r > 0.$$

Доказательство. Рассмотрим случайные величины X_j, такие, что они принимают значение 1, если был успех на j испытании, и 0 иначе. Тогда X_j независимы, имеют конечную дисперсию $p(1-p)$ и математическое ожидание p. Поэтому ЗБЧ Бернулли вытекает из следствия к теореме Чебышева. □

Теорема 14. *(Теорема Вейерштрасса). Любая непрерывная функция не интервале может быть равномерно приближена полиномами.*

Доказательство. Итак, пусть $f : [0,1] \to \mathbb{R}^1$ - непрерывная функция. Обозначим

$$w_f(r) := \max_{|s-t|<r} |f(s) - f(t)|$$

её модуль равномерной непрерывности. Как известно, $\lim_{r\to 0} w_f(r) = 0$. Пусть S_n - число успехов в схеме Бернулли с вероятностью успеха p. Тогда выражение

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

является полиномом степени n по переменной p. Тогда полагая $Z_n = \frac{S_n}{n}$, получим из [строчка из доказательства какой-то хуйни, следствия ЗБЧ, можно написать] для любых r, n

$$\begin{aligned} |B_{n,f}(p) - f(p)| &\leq \max_{|z-a|<r} |f(z) - f(a)| + 2 \max_{[0,1]} |f(\cdot)\, \mathbb{P}(|\frac{S_n}{n} - p| \geq r)| \leq \\ &\leq w_f(r) + 2\max |f(\cdot)| \frac{p(1-p)}{nr^2} \leq \\ &\leq w_f(r) + \max |f(\cdot)| \frac{1}{2nr^2}. \end{aligned}$$

Наконец, выбирая $r = n^{-1/3}$, найдём

$$|B_{n,f}(p) - f(p)| \leq w_f(n^{-1/3}) + \max_{[0,1]} |f(\cdot)| \frac{1}{2n^{1/3}},$$

причём эта оценка стремится к нулю с ростом n и не зависит от p.

□

1.10 Локальная и интегральная предельная теорема Муавра в схеме Бернулли.

Теорема 15. *(Локальная теорема Муавра-Лапласа). Пусть S_n - число успехов в схеме Бернулли из n испытаний с вероятностью успеха $p \in (0,1)$. Тогда для любой последовательности $\varepsilon_n \to 0$ верно, что*

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

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

Доказательство. По формуле Стирлинга

$$\begin{aligned}\mathbb{P}(S_n = k) = C_n^k p^k (1-p)^{n-k} = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \sim \\ \sim \frac{\sqrt{2\pi n}}{\sqrt{2\pi k}\sqrt{2\pi(n-k)}} \frac{n^n e^k e^{n-k}}{e^n k^k (n-k)^{n-k}} p^k (1-p)^{n-k} \sim \\ \sim \frac{1}{\sqrt{2\pi p(1-p)n}} \frac{n^n}{k^k (n-k)^{n-k}} p^k (1-p)^{n-k} = \\ = \frac{1}{\sqrt{1\pi p(1-p)n}} \frac{(np)^k}{k^k} \frac{[n(1-p)]^{n-k}}{(n-k)^{n-k}}.\end{aligned}$$

Определим параметр $v = v(n,k)$ соотношением $k = np+v$. Тогда имеем $n-k = n(1-p)-v$, причём $v = o(n^{2/3})$. Анализируя логарифмы, получим, что

$$\begin{aligned}\ln\left(\frac{k^k}{(np)^k}\right) = k\ln\left(1 + \frac{v}{np}\right) = (np+v)\left(\frac{v}{np} - \frac{v^2}{2(np)^2} + o\left(\frac{v^3}{n^3}\right)\right) = \\ = v + \frac{v^2}{2np} + 0(1),\end{aligned}$$

и аналогично,

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

Складывая оценки логарифмов, получим (старшие члены сокращаются!)

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

по определению, $\frac{v^2}{2np(1-p)} = \frac{(k-np)^2}{2np(1-p)}$, так что теорема доказана. □

Теорема 16. *(Интегральная теорема Муавра-Лапласа). Пусть S_n - число успехов в схеме Бернулли из n испытаний с вероятностью успеха $p \in (0,1)$. Положим $Z_n = \frac{S_n - np}{\sqrt{p(1-p)n}}$. Тогда для любых вещественных $a < b$ и $n \to \infty$ верно*

$$\lim_{n\to\infty} \mathbb{P}\{a \le Z_n \le b\} = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-u^2/2} du.$$

Доказательство. Утверждение теоремы получается простым суммированием асимптотик из локальной теоремы. □

1.11 Производящие функции. Связь с математическим ожиданием и дисперсией. Производящая функция суммы независимых величин. Производящая функция суммы случайного числа независимых одинаково распределенных величин.

Определение 17. *Производящая функция*: $X \in Z_+; p_k = \mathbb{P}\{X = k\}; \varphi(z) = E(z^x) = \Sigma_{k=0}^{\infty} p_k z^k, |z| < 1$

Свойства:

- Если X,Y - сл.нез.величины, то $\varphi_{X+Y}(z) = \varphi_X(z)\varphi_Y(z)$, так как $\varphi_{X+Y}(z) = E(z^{X+Y}) = E(z^X z^Y) = E(z^X)E(z^Y) = \varphi_X(z)\varphi_Y(z)$.
- (Следствие Если X_1,X_2,...X_n - н.о.р. сл.величины, то $\varphi_{S_n}(z) = [\varphi(z)]^n$, где $S_n = X_1 + X_2 + ... + X_n, \varphi$ - производящая функция X_1.
- $\varphi^{'}(1) = EX, \varphi^{''}(1) = E[X(X-1)]$, откуда $DX = \varphi^{''}(1) + \varphi^{'}(1) - [\varphi^{'}(1)]^2$.

  Доказательство. $\lim \frac{\varphi(1)-\varphi(t)}{1-t} = \Sigma_{k=0}^{\infty} \frac{1-t^k}{1-t} p_k = \Sigma_{k=0}^{\infty} p_k k t_k^{k-1}, t \le t_k \le 1, \Sigma_{k=1}^{\infty} p_k k t^{k-1} \le \Sigma_{k=1}^{\infty} p_k k t_k^{k-1} \le \Sigma_{k=1}^{\infty} p_k k = EX$, значит $\varphi^{'}(1) = EX$, повторив аналогичное рассуждение еще раз, получим соотношение на $\varphi^{''}(1)$. □

- Пусть τ и (X_k) - н.о.р.сл.величины, $S_n = X_1 + X_2 + ... + X_n, U = S_\tau, \varphi$ - общая производящая функция, тогда $\varphi_U(z) = \varphi_\tau(\varphi(z))$

  Доказательство.

$$\begin{aligned} \varphi_U(z) &= \Sigma_{k=0}^{\infty}\mathbb{P}\{U = k\}z^k = \Sigma_{k=0}^{\infty}[\Sigma_{n=0}^{\infty}\mathbb{P}\{U = k|\tau = n\}]z^k && = \\ &= \Sigma_{k=0}^{\infty}[\Sigma_{n=0}^{\infty}\mathbb{P}\{\tau = n|S_n = k\}]z^k = \Sigma_{k=0}^{\infty}[\Sigma_{n=0}^{\infty}\mathbb{P}\{\tau = n\}\mathbb{P}\{S_n = k\}]z^k && = \\ &= \Sigma_{n=0}^{\infty}[\Sigma_{k=0}^{\infty}\mathbb{P}\{\tau = n\}\mathbb{P}\{S_n = k\}z^k] = \Sigma_{n=0}^{\infty}\mathbb{P}\{\tau = n\}[\Sigma_{k=0}^{\infty}\mathbb{P}\{S_n = k\}z^k] && = \\ &= \Sigma_{n=0}^{\infty}\mathbb{P}\{\tau = n\}\varphi_{S_n}(z) = \Sigma_{n=0}^{\infty}\,\mathbb{P}\{\tau = n\}[\varphi(z)]^n = \varphi_\tau(\varphi(z)). \end{aligned}$$

□

1.12 Ветвящийся процесс. Определение. Вероятность вырождения. Производящая функция ветвящегося процесса как итерация производящей функции числа непосредственных . . .

(*Описание ветявщегося процесса*). Мы последовательно строим дерево. В 0 уровне 1 вершина, из которой появляются потомки, которые составляют 1 уровень, из каждого появляются потомки, составляющие 2 уровень итд. M_n - число вершин на n-ом уровне, X - общая случайная величина количества потомков вершины, φ - производящая функция X, φ_n - производящая функция M_n. Заметим, что $M_{n+1} = \Sigma_{i=1}^{M_n} X_{n,i}$, где $X_{n,i}$ - количество потомков i-ой вершины на n-ом уровне. По последнему пункту 11 билета из этого следует $\varphi_{n+1} = \varphi_n(\varphi(z))$, таким образом $\varphi_n = \varphi^{n\circ}$, где $\varphi^{n\circ}$ - n-ая композиция φ.

$q_n = \mathbb{P}\{M_n = 0\}, q = \mathbb{P}\{\bigcup\{M_n = 0\}\}$. $\{M_n = 0\} \subset \{M_{n+1} = 0\} \Rightarrow q_n \uparrow q$.

Теорема 17. *Вероятность вырождения процесса q совпадает с наименьшим корнем уравнения $\varphi(t) = t$.*

Доказательство. $q_n = \mathbb{P}\{M_n = 0\} = \varphi_n(0) = \varphi^{n\circ}(0)$. $\quad q_{n+1} = \varphi(q_n) \Rightarrow q = \varphi(q)$, так как φ - непрерывная. То есть q - корень $\varphi(t) = t$. Пусть t^* - наименьший корень $\varphi(t) = t, q_0 = 0 \le t^*$, если $q_n \le t^*$, то $q_{n+1} = \varphi(q_n) \le \varphi(t^*) = t^*$, так как φ - монотонна, а значит $q \le t^* \Rightarrow q = t^*$. □

В последующих рассуждениях мы будем использовать, что $\varphi(1) = 1$,$\varphi(0) = p_0$ и φ - неубывающая выпуклая функция

Теорема 18. *(Критерий невырожденности).*
Следующие условия равносильны:

- $q < 1$
- $EX > 1$ *или* $\mathbb{P}\{X = 1\} = 1$

Доказательство.
(2 $\Rightarrow$ 1) 2 случай очевиден; если $EX > 1$, то $\varphi^{'}(1) > 1$, а значит $\exists\delta$, такое что $\varphi(t) < t, \forall 1-\delta < t < 1 \Rightarrow q = t^* < 1$.
(2 $\Leftarrow$ 1) Пусть ни то ни другое неверно, тогда $EX \le 1 \Rightarrow \varphi^{'}(1) \le 1 \Rightarrow$ либо $\varphi(t) > t\ \forall t < 1$, тогда $q = t^* = 1$, либо $\varphi(t) = t \forall t \in [a;1]$ и $\varphi(t) > t \forall t \in [0;a)$, тогда $\varphi^{'}(1) = 1, \varphi^{''}(1) = 0$, а значит $EX = 1, DX = 0 \Rightarrow X = 1$, то есть $\mathbb{P}\{X = 1\} = 1$. □

Определение 18. Если $EX < 1$, то процесс подкритический,
Если $EX = 1$, то процесс критический,
Если $EX > 1$, то процесс надкритический.

1.13 Цепи Маркова со счетным множеством состояний. Основные определения. Вероятность конкретной траектории цепи. Вероятности перехода за n шагов. Инвариантные (стационарные) и предельные распределения цепи. Предельное распределение как инвариантное. Примеры несуществования (простое симметричное случайное блуждание).

Определение 19. Пусть $(\Omega, F, \mathbb{P})$ - вероятностное пространство. Последовательность случайных величин $(X_n)_{n=0,1...}$ (с не более чем счётным множеством состояний L) называется *цепью Маркова*, если $\mathbb{P}(X_{n+1} = l_{i_{n+1}}|X_n = l_{i_n}, X_{n-1} = l_{i_{n-1}}, ...X_0 = l_{i_0}) = \mathbb{P}(X_{n+1} = l_{i_{n+1}}|X_n = l_{i_n})$ для любых $l_{i_0}, ...l_{i_{n+1}} \in L$, т.е., неформально говоря, если каждая случайная величина зависит только от предыдущей.

Далее везде мы состояние цепи Маркова будем просто обозначать числами, писать i вместо l_i.

Определение 20. Если $\mathbb{P}(X_{n+1} = j|X_n = i)$ не зависит от n, то цепь Маркова называется *однородной*. Отныне считаем, что все рассматриваемые цепи однородны.

Определение 21. *Начальное распределение* $\pi_0(x) = \mathbb{P}(X_0 = x)$. Можно записать в виде вектора: $\overline{\pi_0} = (\pi_0(x_i))_{i=0,1...}$

Определение 22. *Вероятность перехода* из i в j - $p_{i,j} = \mathbb{P}(X_{n+1} = j|X_n = i)$. *матрица перехода* $P = (p_{i,j})_{i,j=0,1...}$.

Теорема 19. $\mathbb{P}(X_0 = x_0, ...X_n = x_n) = \pi_0(x_0)p_{x_0,x_1}...p_{x_{n-1},x_n}$

Доказательство. Индукция по n. База $n = 0$ очевидна. Переход: $p_{x_{n-1},x_n} = \mathbb{P}(X_n = X_n|X_{n-1} = x_{n-1}) = \mathbb{P}(X_n = X_n|X_{n-1} = x_{n-1}, ...X_0 = x_0) = \frac{\mathbb{P}(X_n=x_n, X_{n-1}=x_{n-1},...X_0=x_0)}{\mathbb{P}(X_{n-1}=x_{n-1},...X_0=x_0)}$. □

Определение 23. Определим *вероятности перехода за n шагов:* $p_{i,j}(n) = \mathbb{P}(X_{t+n} = j|X_t = i)$ Им соответствует матрица $P^{(n)}$.

Корректность данного определения интуитивно очевидна из однородности цепи Маркова, но докажем её строго:

Утверждение 5. Это определение корректно:

Доказательство.

$$\mathbb{P}(X_{t+n} = j|X_t = i) = \frac{\mathbb{P}(X_{t+n} = j, X_t = i)}{\mathbb{P}(X_t = i)} =$$
$$= \frac{\sum_{(i_1,...i_{n-1})} \mathbb{P}(X_t = i, X_{t+1} = i_1, ...X_{t+n-1} = i_{n-1}, X_{t+n} = j)}{\mathbb{P}(X_t = i)} =$$
$$= \frac{\sum_{...} \mathbb{P}(X_t = i)p_{i,i_1}p_{i_1,i_2}...p_{i_{n-1},j}}{\mathbb{P}(X_t = i)} = \mathrm{Const},$$

т.е. не зависит от t □

Теорема 20.

1. $p_{i,j}(n + m) = \sum_k p_{i,k}(n)p_{k,j}(m)$
2. $P^{(n+m)} = P^{(n)}P^{(m)}$
3. $P^{(n)} = P^n$

Доказательство. Докажем первый пункт, из которого сразу будут следовать остальные. $p_{i,j}(n + m) = \mathbb{P}(X_{n+m} = j|X_0 = i) = \sum_k \mathbb{P}(X_n = k|X_0 = i)\mathbb{P}(X_{n+m} = i|X_n = k, X_0 = i) = \sum_k p_{i,k}(n)\mathbb{P}(X_{n+m} = i|X_n = k) = \sum_k p_{i,k}(n)p_{k,j}(m)$ □

Определение 24. Распределение π называется *инвариантным (стационарным)* для переходной матрицы P, если $\pi \cdot P = \pi$

Следствие 5. $\pi^n \cdot P = \pi$

Утверждение 6. У простого симметричного случайного блуждания нет стационарного распределения.

Доказательство. Если бы стационарное распределение $\overline{\pi}$ существовало, то оно удовлетворяло бы функциональному уравнению $\pi(n) = \frac{\pi(n-1)+\pi(n+1)}{2}$, откуда $\pi(n + 1) - \pi(n) = \pi(n) - \pi(n - 1)$, и $\pi(n) = cn + \pi(0)$. Но такое решение не удовлетворяет соотношению $\sum_{n\in\mathbb{Z}} \pi(n) = 1$, противоречие. □

Определение 25. Предположим, что последовательность $(\overline{p(n)}) = (\overline{\pi}P^n)_{n=0,1...}$ сходится к некоторому распределению τ по метрике из теоремы Маркова (сумма модулей разностей координат) (для случая конечного пространства значений это равносильно покоординатной сходимости). Тогда τ называется *предельным распределением* для начального состояния. π.

Теорема 21.

1. *τ - действительно распределение.*

2. *τ стационарно.*

Доказательство. 1. $1-\epsilon < \sum_i |\tau_i - p_i(n)| - \sum_i p_i(n) \le \sum_i \tau_i \le \sum_i |\tau_i - p_i(n)| + \sum_i p_i(n) < \epsilon + 1$ для любого $\epsilon > 0$ и $n > N(\epsilon)$. Неотрицательность же τ_i очевидна (предел неотрицательных величин неотрицателен).

2. $d(\overline{\tau}P, \overline{\tau}) \le d(\overline{\tau}P, \overline{p(n+1)}) + d(\overline{p(n+1)}, \overline{p(n)}) + d(\overline{p(n)}, \overline{\tau}) < 3\epsilon$ при $n > N(\epsilon)$. Здесь неочевидным является факт, что $d(\overline{x}P, \overline{y}P) \le d(\overline{x}, \overline{y})$. Докажем его по аналогии с рассуждениями из теоремы Маркова: $d(\overline{x}P, \overline{y}P) = \sum_j |\underbrace{\sum_i x_i p_{i,j} - \sum_i y_i p_{i,j}}_{A_j}| = \sum_{j\in J_+} \sum_i (x_i - y_i)p_{i,j} - \sum_{j \in J_-} \sum_i (x_i - y_i)p_{i,j} \le \sum_j \sum_i |x_i - y_i| p_{i,j} = \sum_i |x_i - y_i| \sum_j p_{i,j} = d(\overline{x}, \overline{y})$, что и требовалось (мы поменяли в выкладках знаки суммирования, потому что ряд сходится абсолютно).

$\square$

1.14 Асимптотическое поведение вероятностей перехода – предельная теорема Маркова для конечной цепи с положительными вероятностями перехода и ее непосредственное обобщение.

Теорема 22. *(Теорема Маркова о финальных вероятностях). Пусть (X_n) - цепь Маркова, у которой $m < \infty$ состояний. Предположим, что $\forall i, j$ верно неравенство $p_{i,j} \ge \delta > 0$. Тогда существет единственное стационарное распределение $\overline{\pi}$, причём $\max_{i,j} |p_{i,j}(n) - \pi_j| \le C(\max\{0, 1-2\delta\})^n$, где $C > 0$ - константа.*

Доказательство. Мы будем применять теорему о сжимающем отображении. Для этого определим на множестве $H = \{\overline{x} \in \mathbb{R}^m \,|x_i \ge 0, \sum x_i = 1\}$ метрику $d(\overline{x}, \overline{y}) = \sum_{i=1}^m |x_i - y_i|$. Полученное метрическое пространство будет полным, так как метрика d эквивалентна стандартной евклидовой метрике пространства $\mathbb{R}^m$, которое полно. Также определим отображение $T\overline{x} := \overline{x}P$. Проверим, что оно сжимающее. $d(T\overline{x}, T\overline{y}) = \sum_{j=1}^m |\underbrace{\sum_i x_i p_{i,j} - \sum_i y_i p_{i,j}}_{A_j}| = \sum_{j\in J_+} \sum_i (x_i - y_i)p_{i,j} - \sum_{j\in J_-} \sum_i (x_i - y_i)p_{i,j} = \sum_i (x_i - y_i)(\sum_{j\in J_+} p_{i,j} - \sum_{j \in J_-} p_{i,j})$ (Здесь $J_+ = \{s|A_s \ge 0\}$, $J_- = \{s|A_s < 0\}$). Возможны три случая:

- $J_- = \emptyset$. Тогда $d(T\bar(x), T\overline{y}) = \sum_i (x_i - y_i) \sum_j p_{i,j} = 0 \le C(1-\delta)^n$
- $J_+ = \emptyset$ - аналогичен предыдущему.
- В J_- и J_+ есть хотя бы одно по одному состоянию. Тогда легко понять, что $\delta \le \sum_{j\in J_+} p_{i,j} \le 1-\delta$ и аналогично с J_-, и $d(T\overline{x}, T\overline{y}) = \sum_i (x_i - y_i)(\sum_{j\in J_+} p_{i,j} - \sum_{j\in J_-} p_{i,j}) \le \sum_i |(x_i - y_i)| \cdot |(\sum_{j\in J_+} p_{i,j} - \sum_{j \in J_-} p_{i,j})| \le (1-2\delta) \sum_i |x_i - y_i| = (1-2\delta) d(\overline{x}, \overline{y})$. (заметим, что если этот случай реализовался, то $\delta \le \frac{1}{2}$)

Мы убедились, что отображение сжимающее. Тогда выполнены все условия теоремы о сжимающем отображении, и в нашем метрическом пространстве существует единственная неподвижная точка $\overline{\pi}$. По индукции получаем, что $d(T^n\overline{x}, \overline{\pi}) \le (1-2\delta)^n d(\overline{x}, \overline{\pi}) \le 2(1-2\delta)^n$, так как $d(\overline{x}, \overline{y}) = \sum_{i=0}^m |x_i - y_i| \le \sum_{i=0}^m |x_i| + |y_i| = 2$. Зафиксируем теперь произвольное i и рассмотрим вектор $\overline{x} = (p_{i,j})$. Получим, что $\max_j |p_{i,j}(n) - \pi_j| \le \sum_j |p_{i,j}(n) - \pi_j| \le 2(1-2\delta)^n$ - равномерная оценка по i. $\square$

Следствие 6. Пусть (X_n) - цепь Маркова, у которой $m < \infty$ состояний. Предположим, что для некоторого $N > 0$ и $\forall i, j$ верно неравенство $p_{i,j}(N) \geq \delta > 0$. Тогда существет единственное стационарное распределение $\overline{\pi}$, причём $\max_{i,j} |p_{i,j}(n) - \pi_j| \leq C\rho^n$, где $C > 0$ и $0 \leq \rho < 1$ - некоторые константы.

Доказательство. Будем использовать обозначения из теоремы Маркова и рассмотрим отображение $U = T^N$. По теореме Маркова оно сжимающее, существует едintвенное стационарное распределение π и $d(U^{kN}\overline{x}, \overline{\pi}) \leq 2(\max\{0, 1-2\delta\})^k$. Так как всегда верно неравенство $d(T\overline{x}, T\overline{y}) \leq d(\overline{x}, \overline{y})$, то $\sup_{\overline{x},\overline{y}} d(T\overline{x}, T\overline{y}) \leq \sup_{\overline{x},\overline{y}} d(\overline{x}, \overline{y})$. Обозначим через $\Delta_n = \sup_{\overline{x},\overline{y}} d(T^n\overline{x}, T^n\overline{y})$. Из вышесказанного следуют следующие свойства:

1. (Δ_n) нестрого убывает.
2. $\Delta_{kN} \leq 2(\max\{0, 1-2\delta\})^k$.

Тогда легко показать, что эта последовательность убывает к нулю экспоненциально быстро всегда, для этого достаточно взять $\rho = \max\{0, 1-2\delta\}^{\frac{1}{N}} C = \frac{2}{\rho^{N-1}}$, тогда $\Delta_{kN+r} \leq \Delta_{kN} \leq 2(\max\{0, 1-2\delta\})^k = 2\rho^{kN} \leq \frac{2}{\rho^{N-1}}\rho^{kN+r} = C\rho^{kN+r}$. В итоге получили оценку $d(T^n\overline{x}, T^n\overline{y}) \leq C\rho^n$, откуда требуемое неравенство выводится так же, как в оригинальной теореме. Для полного доказательства осталось лишь проверить, что $\overline{\pi}$ - стационарное распределение для T. Но $d(T\overline{\pi}, \overline{\pi}) = d(T(T^{Nk}\overline{\pi}), T^{Nk}\overline{\pi}) = d(T^{Nk}(T\overline{\pi}), T^{Nk}\overline{\pi}) \leq \Delta_{Nk} \to 0$, и значит, $d(T\overline{\pi}, \overline{\pi}) = 0$. □

1.15 Классификация состояний: существенные, несущественные состояния. Эргодические классы существенных состояний. Период состояния ...

Определение 26. Состояние j называется *достижимым* из i, если для некоторого натурального n $p_{i,j}(n) > 0$ (в таком случае мы буде писать $i \to j$)

Определение 27. Состояния i и j *сообщаются*, если они достижимы друг из друга (обозначим это через $i \leftrightarrow j$)

Определение 28. Состояние i *существенно*, если оно сообщается с любым достижимым из него состоянием (т.е. $\forall j$ из $i \to j$ следует, что $j \to i$), и *несущественно* в противном случае.

Утверждение 7. $\leftrightarrow$ - отношение эквивалентности на множестве существенных состояний S.

Определение 29. Классы эквивалентности на S по $\leftrightarrow$ называются *эргодическими классами*.

Определение 30. Для каждого состояния i рассмотрим множество $L_i = \{n > 0 | p_{i,i}(n) > 0\}$. Если $L_i \neq \emptyset$, то можно определить *период состояния* i как $d(i) = \gcd(L_i)$.

Утверждение 8. Если $i \leftrightarrow j$, то $d(i) = d(j)$.

Доказательство. Из существенности этих состояний следует, что $\exists k, l$ такие, что $p_{i,j}(k) > 0$ и $p_{j,i}(l) > 0$. Тогда для $n = d(i) \in L_i$ $p_{j,j}(l+n+k) \geq p_{j,i}(l)p_{i,i}(n)p_{i,j}(k) > 0$ и $p_{j,j}(l+k) \geq p_{j,i}(l)p_{i,j}(k) > 0$, откуда получаем, что числа $n+k+l$ и $k+l$ лежат в L_j. Тогда $d(j)|n+k+l$, $d(j)|k+l \implies d(j)|n = d(i)$. Аналогично, $d(i)|d(j)$. □

Определение 31. Одинаковый период состояний в одном эргодическом классе называется *периодом класса*.

Теорема 23. *Пусть E - эргодический класс с периодом $d > 1$. Тогда $E = \bigsqcup_{j=0}^{d-1} C_j$, при этом если $x \in C_k$, и $p_{x,y} > 0$, то $y \in C_{k+1 \mod d}$. Более того, это разбиение единственно с точностью до циклического сдвига индексов.*

Доказательство. Выберем произвольное состояние i_0 и определим $C_k = \{j | \exists n p_{i,j}(k+nd) > 0\}$. Понятно, что $E = \bigcup_{k=0}^{d-1} C_k$.

Сначала докажем, что эти классы попарно дизъюнктны. Предположим противное: существуют состояние x и различные числа $0 \le k_1, k_2 < d$ такие, что $p_{x_0,x}(n_1 d + k_1) > 0$, $p_{x_0,x}(n_2 d + k_2) > 0$. Так как x, x_0 лежат в одном эргодическом классе, то для некоторого r верно $p_{x,x_0}(r) > 0$. Тогда из x_0 в себя же можно добраться как за $n_1 d + k_1 + r$ шаг, так за $n_2 d + k_2 + r$ шаг, а тогда $d | n_1 d + k_1 + r$, $d | n_2 d + k_2 + r \implies d | k_1 - k_2$, что невозможно.

Осталось доказать, что другого такого разбиения нет. Действительно, пусть $E = \bigsqcup_{k=0}^{d-1} D_k$ (переходы здесь так же из i-го множества в $i+1$-ое), и $D_0 \cap C_0 \supset i$ (так как мы можем циклически сдвигать индексы). Если для некоторого m теперь найдётся элемент j такой, что $j \in C_m, j \notin D_m$ (другой случай аналогичен), то рассмотрим путь из i в j. По индукции легко показывается, что его t-ое звено (нумеруем с нуля) лежит и в C_t, и в D_t, противоречие. $\square$

Определение 32. Эти C_i называются *циклическими подклассами* данного эргодического класса.

1.16 Возвратные и невозвратные состояния, критерий возвратности. Возвратность простого случайного блуждания в $\mathbb{Z}^d$ (теорема Пойа). Простое случайное блуждание на $\mathbb{Z}$ ...

Для данного состояния j определим $f_{j,j}(k) = \mathbb{P}(\underbrace{X_k = j, X_{k-1} \neq j, ...X_1 \neq j}_{A_k} | X_0 = j)$ - вероятность впервые вернуться обратно на k-ом шаге. Так как события (A_k) попарно дизъюнктны, то $\mathbb{P}(\bigcup_k A_k | X_0 = j) = \sum_{k=1}^{\infty} f_{j,j}(k) = f_j \le 1$ - вероятность хотя бы раз вернуться обратно. Также пусть $p_{j,j}(n)$ - вероятность перехода из j в j за n шагов.

Определение 33. Если $f_j = 1$, то состояние j называется *возвратным*.

Лемма 1. $p_{j,j}(n) = \sum_{k=1}^{n} f_{j,j}(k) p_{j,j}(n-k)$.

Доказательство.

$$p_{j,j}(n) = \mathbb{P}(X_n = j | X_0 = j) = \sum_{k-1}^{n} \mathbb{P}((X_n = j) \cap A_k | X_0 = j) =$$
$$= \sum_{k=1}^{n} \mathbb{P}(A_k | X_0 = j) \mathbb{P}(X_n = j | X_k = j, X_{k-1} \neq j, ...X_1 \neq j, X_0 = j) =$$
$$= \sum_{k=1}^{n} f_{j,j}(k) \mathbb{P}(X_n = j | X_k = j) = \sum_{k=1}^{n} f_{j,j}(k) p_{j,j}(n-k)$$

$\square$

Доопределим $f_{j,j}(0) = 0$, $p_{j,j}(0) = 1$, и рассмотрим производящие функции F и G последовательностей $(f_{j,j}(n))$ и $(p_{j,j}(n))$ соответственно. Они определены при $|t| < 1$. Из утверждения выше следует, что $G(t) - 1 = F(t)G(t) \iff G(t) = \frac{1}{1-F(t)}$.

Лемма 2. *Последовательность $(a_k)_{k=0}^{\infty}$ неотрицательна, ограничена сверху константой C и имеет производящую функцию $A(t)$. Тогда $\lim_{t\nearrow 1} A(t) = A(1) = \sum_{k=0}^{\infty} a_k$.*

Доказательство. Во-первых, $A(t) \le A(1) \implies \limsup_{t\nearrow 1} A(t) \le A(1)$. Во-вторых, для любого натурального N верно, что $A(t) \ge \sum_{k=0}^{N} a_k t^k \implies \liminf_{t\nearrow 1} A(t) \ge \liminf_{t\nearrow 1}(\sum_{k=0}^{N} a_k t^k) = \sum_{k=0}^{N} a_k \implies \liminf_{t\nearrow 1} A(t) \ge A(1)$. Значит, $\lim_{t\nearrow 1} A(1) = A(1)$. □

Из леммы следует, что $\sum_{k=0}^{\infty} p_{j,j}(k) = G(1) = \frac{1}{1-F(1)} = \frac{1}{1-f_j}$. В итоге получаем следующий критерий возвратности состояния:

Теорема 24. *Состояние j возвратно $\iff$ ряд $\sum_{k=0}^{\infty} p_{j,j}(k)$ расходится.*

Следствие 7. Если i возвратно, и $i \leftrightarrow j$, то j тоже возвратно.

Доказательство. $\exists k, l \in \mathbb{N}$ такие, что $p_{i,j}(k) > 0$, $p_{j,i}(l) > 0$. Тогда $p_{j,j}(n+k+l) \ge p_{j,i}(l)p_{i,i}(n)p_{i,j}(k) = Cp_{i,i}(n)$, где $C > 0$, Значит, ряд $\sum_{n=0}^{\infty} p_{j,j}(n+k+l)$ тоже расходится, а вместе с ним и ряд $\sum_{n=0}^{\infty} p_{j,j}(n)$ □

Теорема 25. *(Теорема Пойа).:*

1. *Простое случайное блуждание на $\mathbb{Z}$ возвратно только при $p = q = \frac{1}{2}$*
2. *Простое случайное симметричное блуждание на $\mathbb{Z}^2$ возвратно*
3. *Простое случайное блуждание на $\mathbb{Z}^d$, $d > 2$, невозвратно*

Во всех случаях мы выходим из начала координат.

Доказательство.

1. $p(2k+1) = 0$, $p(2k) = \binom{2k}{k}(pq)^k = \left(\frac{\binom{2k}{k}}{2^{2k}}\right)(4pq)^k \sim \frac{1}{\sqrt{\pi k}}(4pq)^k$. если $p = q = \frac{1}{2}$, то ряд из p_k расходится, и блуждание возвратно. Иначе $0 \le 4pq < 1$, и ряд сходится даже быстрее, чем геометрическая прогрессия.
2. $p(2k+1) = 0$, $p(2k) = \sum_{m=0}^{k} \mathbb{P}(X_{2k} = 0 \text{ и мы сделали ровно m шагов вправо} | X_0 = 0) = \sum_{m=0}^{k} \frac{1}{4^{2k}}\binom{2k}{m}\binom{2k-m}{m}\binom{2k-2m}{k-m} = \frac{1}{4^{2k}}\sum_{m=0}^{k} \frac{(2k)!}{m!(2k-m)!}\frac{(2k-m)!}{m!(2k-2m)!}\frac{(2k-2m)!}{((k-m)!)(k-m)!} = \frac{1}{4^{2k}}\sum_{m=0}^{k}\frac{(2k)!}{(m!)^2((k-m)!)^2} = \frac{1}{4^{2k}}\binom{2k}{k}\sum_{m=0}^{k}\binom{k}{m}^2 = \left(\frac{\binom{2k}{k}}{4^k}\right)^2 \sim \frac{1}{\pi k}$, и ряд из p_k расходится
3. Не доказывали.

□

Следствие 8. Для любой точки $x \in \mathbb{Z}$ $\mathbb{P}(\text{хотя бы один раз прийти в } 0 | X_0 = x) = 1$.

Доказательство. Обозначим $f(x) = \mathbb{P}(\text{хотя бы один раз прийти в } 0 | X_0 = x)$. По формуле полной вероятности, получаем $f(x) = \mathbb{P}(\text{хотя бы один раз прийти в } 0 | X_0 = x) = \mathbb{P}(\text{хотя бы один раз прийти в } 0 | X_0 = x, X_1 = x-1)\mathbb{P}(X_1 = x-1) + \mathbb{P}(\text{хотя бы один раз прийти в } 0 | X_0 = x, X_1 = x+1)\mathbb{P}(X_1 = x+1) = \frac{1}{2}(f(x-1) + f(x+1))$. Легко понять, что решение данного функционального уравнения над $\mathbb{Z}$ - $f(x) = ax + b$, причём $0 \le f(x) \le 1$ для любого x, и $f(0) = 1$ по теореме Пойа. из первого свойства получаем, что $a = 0$, а из второго, что $b = 1$. Значит, $f(x) \equiv 1$ □

Следствие 9. Написав аналогичное функциональное уравнение и рассмотрев $(0,0)$, его соседей, соседей соседей и т.д., можно доказать данное утверждение для блуждания на $\mathbb{Z}^2$.

Определим $M_n = \max_{0\le k\le n}(X_n)$.

Теорема 26. *При $n\to\infty$ и $j\ge 0$ $\mathbb{P}(M_n\ge j)\approx 2\mathbb{P}(X_n\ge j)$.*

Доказательство. $\mathbb{P}(M_n\ge j) = \mathbb{P}(M_n\ge j, X_n\ge j) + \mathbb{P}(M_n\ge j, X_n<j) = \mathbb{P}(X_n\ge j) + \mathbb{P}(X_n>j) = 2\,\mathbb{P}(X_n\ge j) - \mathbb{P}(X_n=j)\approx 2\,\mathbb{P}(X_n\ge j)$. В обосновании нуждается равенство $\mathbb{P}(M_n\ge j, X_n<j) = \mathbb{P}(S_n>j)$. Применим *принцип отражения*: рассмотрим траекторию на $\mathbb{Z}^2$, попадающую под первое условие, отметим момент, когда она впервые пересекла прямую $y=j$, и отразим часть справа от неё. Получим траекторию, попадающую под второе условие. И наоборот, любая траектория, заканчивающаяся в точке с ординатой $>j$, где-то пересекает прямую $y=j$, и мы моожем по ней построить тракеторию такую, что $M_n\ge j, X_n<j$. Получили взаимоoднозначное соответствие между данными двумя множествами, а значит, и вероятности тоже равны. □

Теорема 27. *(Задача о разорении игрока). Есть два игрока, у одного из них a монет, у другого b $(a,b\in\mathbb{Z}_{\ge 0})$. За каждый ход либо первый игрок отдаёт одну моенту второму, либо наоборот, причём оба события имеют вероятность $\frac{1}{2}$. Проигрывает игрок, у которого не остаётся монет. Тогда вероятность выигрыша первого игрока равна $\frac{a}{a+b}$.*

Доказательство. Пусть в игре находится фиксированное число монет $(a+b)$. Обозначим через ϵ_i последовательность ходов ($\epsilon_i=1$, если на i-ом ходе монету получил первый игрок и -1 иначе). Рассмотрим событие $A=\{$первый игрок проиграл$\}$. $\mathbb{P}(A)=f(x)$. Тогда $f(0)=1$, $f(a+b)=0$, а в остальных случаях применим формулу полной вероятности и получим соотношение $f(a)=\mathbb{P}(A)=\frac{1}{2}\mathbb{P}(A|\epsilon_1=+1)+\frac{1}{2}\mathbb{P}(A|\epsilon_1=-1)=\frac{1}{2}(f(a-1)+f(a+1))$. Решение данного функционального уравнения - $f(x)=px+q$, причём мы знаем, что $q=1$, $0=p(a+b)+1$. В итоге получаем $f(a)=\frac{b}{a+b}$, а требуемая вероятность выигрыша равна $\frac{a}{a+b}$. □

Пусть (X_n) - простое симметричное случайное блуждание на $\mathbb{Z}$. $T_+(n)=\sum_{j=1}^{n}\mathbb{I}_{\{X_j\ge 0, X_{j-1}\ge 0\}}$ - время пребывания на положительной полуоси. Аналогично определяется $T_-(n)$. Понятно, что $T_+(n)+T_-(n)=n$

Лемма 3. $\mathbb{P}(T_+(2n)=2k)=p_k p_{n-k}$*, где* $p_k=\mathbb{P}(X_{2k}=0)=\frac{\binom{2k}{k}}{4^k}$

Доказательство. Давыдов ссылается на доказательство в книге: Ширяев “Вероятность”, глава 1, параграф 10 (страница 125?)
или, кажется, Феллер “Введение в теор.вер. и её применения”, том 1, второй или третий параграф, посвящённый случайным блужданиям. □

Теорема 28 (*Закон арксинуса*)**.** *Для любых $1\ge b>a\ge 0$ при $n\to\infty$*

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

Доказательство. $\mathbb{P}(\frac{T_+(2n)}{2n}\in[a,b])=\sum_{2na\le 2k\le 2nb}\mathbb{P}(T_+(2n)=2k)\approx\sum_{2na\le 2k\le 2nb}\frac{1}{\sqrt{\pi k}}\frac{1}{\sqrt{\pi(n-k)}}=\sum_{na\le k\le nb}\frac{1}{n\pi}\frac{1}{\sqrt{\frac{k}{n}(1-\frac{k}{n})}}=\sum_{a\le\frac{k}{n}\le b}\frac{1}{n}h(\frac{k}{n})\approx\int_a^b h(x)dx=\frac{2}{\pi}\Big(\arcsin\sqrt{b}-\arcsin\sqrt{a}\Big)$, где $h(t)=\frac{1}{\pi}\frac{1}{\sqrt{t(1-t)}}$ □

2 Указатель

Начали спрашивать быстрые вопросы, и ты пришёл сюда? Ну ладно ебать, удачи