

# Конспект по Дискретной математике.

Чепелин В.А.

## Содержание

### 1 Лекция 1.

1.1 Аксиоматическое вероятное пространство. . . . .

### 2 Лекция 2.

2.1 Случайная величина. . . . .

2.2 Мат. ожидание. . . . .

2.3 Незав. случайные величины . . . . .

2.4 Дисперсия случайной величины. . . . .

### 3 Лекция 3.

3.1 Ковариация . . . . .

3.2 Корреляция . . . . .

3.3 Хвостовые неравенства . . . . .

### 4 Лекция 4.

4.1 Введение в теорию информации. . . . .

4.2 Энтропия . . . . .

### 5 Лекция 5.

5.1 Определение цепи Маркова . . . . .

5.2 Эргодический класс . . . . .

5.3 Уходим в математику. . . . .

5.4 Мат. ожидание времени до поглощения . . . . .

### 6 Информация о курсе

# 1 Лекция 1.

## 1.1 Аксиоматическое вероятное пространство.

Пусть у нас есть  $\Omega$  - элементарные исходы и связанная с ним функция  $p : \Omega \rightarrow [0, 1]$  - дискретная вероятностная мера (плотность вероятности) - функция, которая по элементарному исходу возвращает вероятность.

А также  $\sum_{w \in \Omega} p(w) = 1$ , а также  $0 \leq p_i \leq 1$  А также мы считаем, что  $|\Omega|$  не более чем счетно. Для множеств мощности континуума нам нужна более сложная теория.

Рассмотрим **примеры**:

1. Честная монета:

$$\Omega = \{0, 1\}. p(0) = p(1) = \frac{1}{2}.$$

2. Нечестная монета или распределение Бернулли:

$$\Omega = \{0, 1\}. p(0) = 1 - p(1) = q.$$

3. Честная игральная кость:

$$\Omega = \{1, 2, 3, 4, 5, 6\}. p(w) = \frac{1}{6}. p(w) = \frac{1}{52}$$

4. Колода карт:

$$\Omega = \{\langle c, r \rangle \mid 1 \leq c \leq 4, 1 \leq r \leq 15\}$$

5. Геометрическое распределение:

$$\Omega = \mathbb{N}, p(i) = \frac{1}{2^i}$$

**Замечание.** Не существует равномерного распределения на счетном множестве.

**Событие** — множество  $A \subset \Omega$ .  $P(A) = \sum_{w \in A} p(w)$ . (Иногда используют  $\Pr$ ).

$P(A) = 1$  — достоверное событие.

$P(A) = 0$  — невозможное событие.

Рассмотрим примеры на честной игральной кости:

1. Только четные:  $P(A) = \frac{3}{6} = \frac{1}{2}$ .

2. Больше 4-ех:  $P(A) = \frac{2}{6} = \frac{1}{3}$ .

**Замечание:** нельзя с равной вероятностью выбрать случайное целое число.

**Независимые события** — A, B независимы, если  $P(A \cap B) = P(A) \cdot P(B)$ .

$\frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(\Omega)}$  — независимы (если произошло B, то вероятность не поменялась)

$P(A|B) = \frac{P(A \cap B)}{P(B)}$  — вероятность A при условии B — условная вероятность.

Произведение вероятностных пространств.

Пусть у нас есть  $\Omega_1, p_1$ , а также  $\Omega_2, p_2$ , тогда произведение вероятностных пространств:

$$\Omega = \Omega_1 \times \Omega_2$$

$$p(\langle w_1, w_2 \rangle) = p_1(w_1) \cdot p_2(w_2)$$

**Утв.**  $\forall A \subset \Omega_1, B \subset \Omega_2$ .

$A \times \Omega_2$  и  $\Omega_1 \times B$  независимы.

Пусть у нас есть  $n$  событий:  $A_1, A_2, \dots, A_n$ .

Тогда обычно **независимость  $n$  событий** подразумевает:

1.  $A_i, A_j$  - независимы  $\forall i, j, \quad i \neq j$
2.  $\forall I \subset \{1, 2, 3, \dots, n\}: P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$

### Формула полной вероятности

$\Omega = A_1 \cup A_2 \cup \dots \cup A_n, \forall i \neq j: A_i \cap A_j = \emptyset$  — **полная система событий**.

Возьму  $B$  - какое-то событие.

$$P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i) \cdot P(A_i)$$

Пример: урна с шариками. Сначала выбираете урну, потом достаете шарик.

### Формула Байеса.

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(B|A_i) \cdot P(A_i)}{\sum_{j=1}^n P(B|A_j) \cdot P(A_j)}$$

## 2 Лекция 2.

### 2.1 Случайная величина.

**Случайная величина** или численная характеристика каждого элементарного исхода — это отображение  $\xi : \Omega \rightarrow \mathbb{R}$ , которое сопоставляет каждому элементарному исходу какое-то число.

**Пример:**

1.  $D = \{1, 2, \dots, 6\}$ . Возьмем  $\Omega = D^2$ . Например, человек бросает два игральных кубика. Тогда, очевидно,  $p(\langle i, j \rangle) = \frac{1}{36}$ . И тогда он задает функцию случайной величины, например, как  $\xi(\langle i, j \rangle) = i + j$ .
2. Возьмем случайный граф  $G$  на  $n$  вершинах.  $\xi(G) =$  количеству компонент связности. Или  $\xi(G) =$  количеству ребер в этом графе.
3. Давайте кидать игральный кубик и сопоставим каждой выпадающей грани число, равное количеству точек на этой грани. То есть  $\Omega = \{1, 2, \dots, 6\}$ ,  $\xi(i) = i$ .
4.  $\Omega = \{1, 2, \dots, 6\}$ ;  $E = \{2, 4, 6\}$ .  $x_E(w) = \begin{cases} 1, w \in E \\ 0, w \notin E \end{cases}$

Возьмем какие-то  $\Omega, p, \xi$ :

$[\xi = i] = \{w | \xi(w) = i\} \subset \Omega$  — множество элементарных исходов, случайная величина которых равна  $i$ .

**def:**  $f_\xi : \mathbb{R} \rightarrow \mathbb{R}$  — дискретная плотность вероятности  $\xi$ .

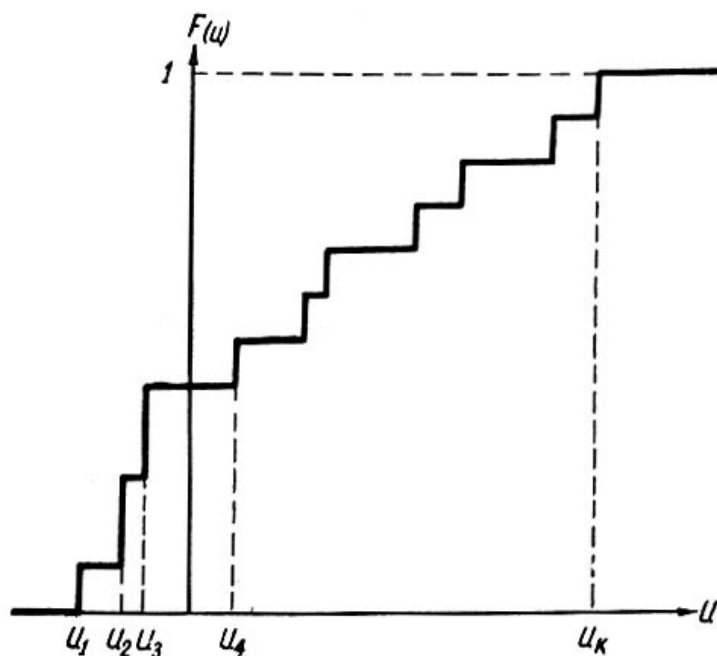
$$P([\xi = i]) = P(\xi = i) = f_\xi(i) = \sum_{w \in [\xi=i]} p(w)$$

**Дискретная плотность вероятности** — это функция, которая говорит нам, насколько вероятно каждое из этих отдельных значений, которые может принимать случайная величина. Другими словами, она присваивает вероятность каждому возможному исходу.

Немного поменяем и получим  $[\xi \leq i] = \{w | \xi(w) \leq i\} \subset \Omega$ .

$$P([\xi \leq i]) = P(\xi \leq i) = F_\xi(i)$$

**def:**  $F_\xi : \mathbb{R} \rightarrow \mathbb{R}$  — функция распределения. У дискретной случайной величины функция распределения ступенчатая. Например:



## 2.2 Мат. ожидание.

Математическое ожидание — среднее значение случайной величины.

$$E_{\xi} = \sum_w p(w)\xi(w) = \sum_i i \cdot P(\xi = i).$$

Дальше А.С. использует 3 вида обозначений:

1.  $E_{\xi}$  2.  $E(\xi)$  3.  $E\xi$  — не боимся, это одно и то же.

Теорема (линейность мат ожидания)

$$E\lambda\xi = \lambda E_{\xi} \quad E_{(\xi+\eta)} = E_{\xi} + E_{\eta}$$

**Доказательство:**

$$E\lambda\xi = \sum_w p(w) \cdot \lambda\xi(w) = \lambda \sum_w p(w)\xi(w) = \lambda E_{\xi}$$

$$E(\xi + \eta) = \sum_w p(w)(\xi(w) + \eta(w)) = \sum_w p(w)\xi(w) + \sum_w p(w)\eta(w) = E(\xi) + E(\eta)$$

Q.E.D.

МАТ. ОЖИДАНИЕ ВСЕГДА ЛИНЕЙНО!!!

## 2.3 Незав. случайные величины

$\xi, \eta$  - независимы, если  $[\xi = a], [\eta = b]$  — независимы  $\forall a, b$ .

Эквивалентное утверждение —  $[\xi \leq a], [\eta \leq b]$  — независимы  $\forall a, b$ .

Иначе говоря, две случайные величины называются *независимыми*, если по значению одной нельзя сделать выводы о значении другой.

**Теорема (о мультипликативности мат. ожидания)**

$\xi, \eta$  — независимы  $\Rightarrow E(\xi \cdot \eta) = E_\xi \cdot E_\eta$ .

**Доказательство:**

$$\begin{aligned} E_{(\xi \cdot \eta)} &= \sum_a a P(\xi, \eta = a) = \sum_a a \sum_{\forall i, j: i \cdot j = a} P(\xi = i, \eta = j) = \\ &= \sum_a \sum_i \sum_j a P(\xi = i) P(\eta = j) = \sum_i i P(\xi = i) \cdot \sum_j j P(\eta = j) = E_\eta \cdot E_\xi \end{aligned}$$

Q.E.D.

**2.4 Дисперсия случайной величины.**

$D_\xi = Var(\xi)$  — дисперсия случайной величины.

$$D_\xi = E((\xi - E_\xi)^2) = E_{\xi^2} - (E_\xi)^2$$

**Дисперсия случайной величины** — это мера того, насколько сильно разбросаны значения этой случайной величины вокруг её математического ожидания (среднего значения). Другими словами, она показывает, насколько "широко" распределение вероятностей случайной величины.

**Теорема (свойства дисперсии).** Если  $\xi, \eta$  - независимы:

$$D_{c\xi} = c^2 D_\xi \quad D_{\xi+\eta} = D_\xi + D_\eta$$

Доказательство тривиально из линейности мат. ожидания.

## 3 Лекция 3.

### 3.1 Ковариация

$$Cov(\xi, \eta) = E_{\xi\eta} - E_{\xi}E_{\eta}$$

**Ковариация** или **корреляционный момент** показывает на сколько зависимы случайные величины это мера зависимости двух случайных величин.

Если  $\xi, \eta$  - независимые случайные величины

$$Cov(\xi, \eta) = 0$$

:

$$Cov(\xi, \xi) = D_{\xi} = Var_{\xi} - \text{вариация}$$

### 3.2 Корреляция

$$Corr(\xi, \eta) = \frac{E_{\xi\eta} - E_{\xi}E_{\eta}}{\sqrt{D_{\xi} \cdot D_{\eta}}} = \frac{Cov(\xi, \eta)}{\sqrt{D_{\xi} \cdot D_{\eta}}}$$

**Корреляция** - статистическая взаимосвязь двух случайных величин. Корреляция является **нормированной** версией ковариации, что позволяет сравнивать силу линейной зависимости между различными парами переменных, независимо от их масштаба.

Теорема (об ограниченности корреляции)

$$-1 \leq Cor(\xi, \eta) \leq 1$$

**Доказательство:**

Возьму  $\alpha = \xi - \lambda\eta$ :

$$D\alpha = D(\alpha) = E\xi^2 - 2\lambda E_{\xi\eta} + \lambda^2 E\eta^2 - (E\xi)^2 + 2\lambda E_{\xi}E_{\eta} - \lambda^2 (E\eta)^2 \geq 0$$

$$D\xi - 2\lambda Cov(\xi, \eta) + \lambda^2 D\eta \geq 0$$

Откуда, если рассматривать это, как уравнение относительно  $\lambda$ , то  $D \leq 0$ , то есть:

$$4Cov(\xi, \eta) - 4D_{\eta}D_{\xi} \leq 0$$

А если присмотреться, то это и есть то, что нам надо.

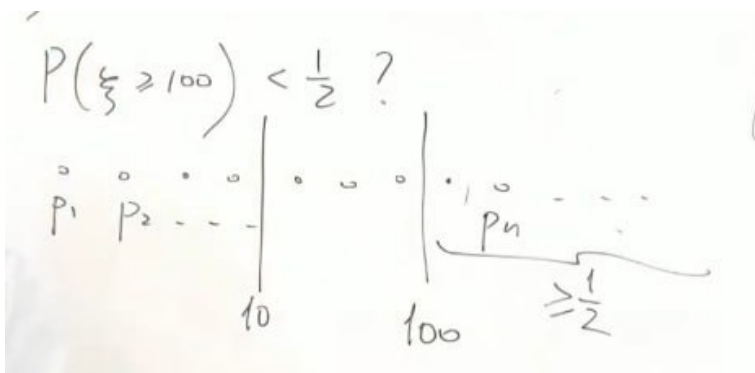
Q.E.D.

### 3.3 Хвостовые неравенства

Рассмотрим азартную игру. ~~не одобряем, не играем.~~

Проводится случайный эксперимент, смотрится значение  $\xi$ . Если оно получилось 100 или больше, то мы платим 100 рублей, а иначе наш друг платит нам 100 рублей. Мы знаем  $E\xi = 10, \xi \geq 0$ .

Хотим оценить  $P(\xi \leq 100)$ :



Давайте посмотрим, является ли наша вероятность меньше  $\frac{1}{2}$ . Тогда всё, что правее 100 имеет вероятность выпадения  $\geq \frac{1}{2}$ . Все левое оценивается нулем, откуда мат ожидание хотя бы 50. Такого быть не может. В общем случае:

### Теорема (Неравенство Маркова)

$$\xi \neq 0, \xi \geq 0 : \forall a \geq 1 : P(\xi \geq a \cdot E\xi) \leq \frac{1}{a}$$

**Доказательство:**

$$E\xi = \sum_v v \cdot P(\xi = v) = \sum_{v < a \cdot E\xi} v P(\xi = v) + \sum_{v \geq a \cdot E\xi} v P(\xi = v) \geq \sum_{v \geq a \cdot E\xi} a E\xi P(\xi = v) = a E\xi \cdot P(\xi \geq a \cdot E\xi)$$

Q.E.D.

### Теорема (Неравенство Чебышева)

Абсолютная версия и относительная версия ( $\alpha = \lambda\sigma$ ):

$$P(|\xi - E\xi| \geq \alpha) \leq \frac{D\xi}{\alpha^2} \quad P(|\xi - E\xi| \geq \lambda\sigma) \leq \frac{1}{\lambda^2}$$

**Доказательство:**

Возьму вот такие величины:

$$D\xi = E(\xi - E\xi)^2 \quad \eta = (\xi - E\xi)^2$$

Заметим, что  $E\eta = D\xi$ . Используем неравенство Маркова для оценки дисперсии:

$$P(\eta \geq c \cdot E\eta) \leq \frac{1}{c}$$

Возьму  $c = \frac{D\xi}{\alpha^2}$  и получу искомое.

Q.E.D.

**Нечестная монета.** Вот вам дали домашку, вместе с вопросом  $p > \frac{1}{2}$  или  $p < \frac{1}{2}$ . Что вы можете делать? Только кидать ее, но при этом бесконечное количество раз вы не кинете, у вас дедлайн домашки через час.



Пусть мы бросили  $n$  раз. Выпало  $c$  единиц и  $n - c$  нулей. Пусть  $c \leq \frac{n}{2}$ :

$$P(\xi = c) \leq P(\xi \leq c) \leq P(|\xi - pn| \geq pn - c) \leq P(|\xi - pn| \geq \frac{n}{2} - c) \leq \frac{n}{4} \cdot \frac{1}{\left(\frac{n}{2} - c\right)^2}$$

Что это концептуально значит? На самом деле, это дает нам оценку на распределение. Зачем? Чтобы **СДАТЬ** домашку.

### Теорема (Граница Чернова)

$$P(\xi \geq (1 + \varepsilon)p) \leq e^{-\frac{\varepsilon^2}{2+\varepsilon}np} \quad \Leftrightarrow \quad P(\xi \geq (1 + \varepsilon)p) \leq e^{-\frac{\varepsilon^2}{2}np}$$

$$e^{-\frac{\varepsilon^2}{3}np} \leq \delta \quad \Leftrightarrow \quad -\frac{\varepsilon^2}{3}np \leq \ln \delta$$

$$n \geq \frac{3}{p\varepsilon^2} \ln \frac{1}{\delta}$$

Не знаю, что это концептуально, напишите пж

## 4 Лекция 4.

### 4.1 Введение в теорию информации.

информация = - неопределенность - сказал дяденька Шеннон

Для осознания нам поможет рисунок АС:



Есть что-то - неизвестное - облачко. Затем, вы с помощью глаза заглядываете туда, и ваша неопределенность уменьшается. Соответственно вы получили информацию. То есть сначала была неопределенность  $H_1$ , потом  $H_2$ .  $I = H_1 - H_2$ , откуда и получается наша формула. У него есть глубокий смысл, но создается вопрос: «И че? И что это за неопределенность?»

Ну наличие глаза мешает, непонятно, фу фу фу. Поэтому хотим ввести что-то более формальное и менее абстрактное.

Пусть у нас есть какой-то случайный эксперимент  $\Omega$ , с вероятностями  $p_1, \dots, p_n$ . И вот мы получили информацию что выпало (например орел на монетке).

**Случайный источник** — черный ящик с красной кнопкой, который показывает номер эл. исхода, когда вы нажимаете на красную кнопку.

Возьмем монетку. Кинули, получили 0 или 1. Теперь возьмем кубик, получим число от 1 до 6. Когда мы кидаем кубик, мы получаем больше информации. И вот Шеннон решил систематизировать все это...

### 4.2 Энтропия

Пусть у нас есть случайный источник и вероятности  $p_1, p_2, \dots, p_n$ . Мы хотим померить численно сколько информации содержится в одном эксперименте:

$$H(p_1, \dots, p_n) : RS \rightarrow R^+$$

**Энтропия Шеннона** ( $H$ ) - это мера неопределенности или случайности, связанная с случайной переменной. Она измеряет среднее количество информации, необходимое для описания результата случайной переменной. Иными словами, энтропия показывает, насколько непредсказуемым является источник информации.

Возьму пример  $p_i = \frac{1}{n}$ . Введем новое обозначение:

$$h(n) = H\left(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n}\right)$$

Очевидно, что  $h(n+1) > h(n)$ .

Теперь рассмотрим вероятностное пространство и источник на нем:

$$\Omega = \{(1, 1), (1, 1), \dots, (1, m_1), (2, 1), \dots, (k, 1), \dots, (k, m_k)\}$$

И давайте теперь каждому причислим какую-то  $q_{ij}$ , так, что в сумме 1.  $p_i = \sum_{j=1}^{m_i} q_{ij}$ . Пусть наш случайный источник сломан и показывает только одно число. Если я возьму сломанный случайный источник от  $\Omega$ , то мы получим столько же информации сколько и у случайного источника сделанного из  $p$ .

Теперь давайте делить это на 2 части. Что вот мы сначала видим первую часть информации, а потом хоба и видим вторую часть информации. И того мы получаем, что когда мы открываем вторую часть мы получим  $p_i H(\frac{q_{i1}}{p_i}, \dots, \frac{q_{im_i}}{p_i})$  информации. Откуда благодаря таким рассуждение получаем свойство, которое называется аддитивностью энтропии:

$$H(p_1, \dots, p_k) + \sum_{i=1}^k p_i H(\frac{q_{i1}}{p_i}, \dots, \frac{q_{im_i}}{p_i}) = H(q_{11}, \dots, q_{mk})$$

Также для фиксированного  $n$ ,  $H$  непр из  $\mathbb{R}^n \rightarrow \mathbb{R}$ .

Теорема.(Формула энтропии Шеннона)

$$H(p_1, \dots, p_n) = -\alpha \sum_{i=1}^n p_i \log_2 p_i$$

$\alpha$  отвечает за выбор единицы измерений.

**Доказательство:**

Лемма 1.  $h(n \cdot m) = h(n) + h(m)$ .

**Доказательство:**

Возьмем  $k = n, m_i = m, p = \frac{1}{n}, q_{ij} = \frac{1}{nm}$ . Из утверждения сверху это верно!

Q.E.D.

Фиксируем  $h(2) = \alpha$ . Тогда:

Лемма 2.  $h(2^k) = k\alpha$ . тривиально из Леммы 1.

Лемма 2,5.  $h(n^r) = rh(n)$ . тривиально из Леммы 1.

Лемма 3.  $h(n) = \alpha \log_2 n$

**Доказательство:**

Найду  $i$  такое, что  $2^i \leq n^r < 2^{i+1}$ , где  $r \in \mathbb{N}$ .

Из монотонности  $h$  следует:  $\alpha i \leq h(n^r) < \alpha(i+1)$ . Поэтому:

$$\alpha i \leq rh(n) < \alpha \Leftrightarrow a \frac{i}{r} \leq h(n) \leq a \frac{i+1}{r}$$

Также мы знаем, что  $i \leq r \log_2 n < i + 1$ . Получим, что:

$$\alpha \frac{i}{r} \leq \alpha \log_2 n < \alpha \frac{i+1}{r}$$

То есть  $\forall r : |h(n) - \alpha \log_2 n| \leq \frac{\alpha}{r}$ . Откуда, получаем требуемое равенство.

Q.E.D.

Возвращаемся к доказательству теоремы. Пусть  $p_i$  рациональные. Приведем все  $p$  к общему знаменателю и пусть теперь  $p_i = \frac{a_i}{b_i}$ . Возьму  $m_i = a_i$ ,  $r_{ij} = \frac{1}{a_i}$ ,  $q_{ij} = \frac{1}{n}$ . Подставим во второе неравенство получим:

$$H\left(\frac{1}{b}, \frac{1}{b}, \dots, \frac{1}{b}\right) = H(p_1, p_2, \dots, p_k) + \sum_{i=1}^k p_i H\left(\frac{1}{a_i}, \dots, \frac{1}{a_i}\right)$$

Что тут происходит? Я разбиваю каждый исход изначальный, на  $a_i$  исходов по  $\frac{1}{b_i}$ . С одной стороны я получаю  $b$  исходов по  $\frac{1}{b}$ . С другой стороны я могу выбрать исход, а потом его разбить. Откуда по аддитивности и получается такая формула. А она в свою очередь уже удобная, так как в ней повторяются значения внутри  $H$ , так что можем заменить на  $h$ :

$$h(b) = H(p_1, \dots, p_k) + \sum_{i=1}^k p_i h(a_i)$$

Заметим, что  $\sum_{i=1}^n p_i = 1$ , так что левую часть на эту сумму:

$$\sum_{i=1}^n p_i h(b) = H(p_1, \dots, p_k) + \sum_{i=1}^k p_i h(a_i)$$

$$H(p_1, \dots, p_k) = \sum_{i=1}^n p_i (h(b) - h(a_i))$$

$$H(p_1, \dots, p_k) = \sum_{i=1}^n p_i (\alpha \log_2 b - \alpha \log_2 a_i) = -\alpha \sum_{i=1}^n p_i \log_2 p_i$$

Эта формула верна и не для рац. исходя непрерывности (любое не рац. можно зажать с двух сторон сходящимися последовательностями и мы победили)

Q.E.D.

$\alpha$  — бит, единица информации.

Обычно используется логарифм по основанию 2, тогда энтропия измеряется в битах (или "Шеннонах"). Если используется натуральный логарифм (основание  $e$ ), то энтропия измеряется в натах. Использование логарифма по основанию 10 даёт единицы измерения в децитах (Hartleys). Выбор основания влияет только на масштаб энтропии, а не на её относительные значения.

Энтропия Шеннона имеет широкое применение в различных областях, включая:

- Теория информации: Является фундаментальным понятием для измерения количества информации.
- Сжатие данных: Используется для оценки теоретического предела сжатия данных.
- Криптография: Оценка случайности ключей и стойкости шифров.
- Машинное обучение: В деревьях решений используется для выбора признаков, которые лучше всего разделяют данные.
- Обработка естественного языка (NLP): Оценка неопределенности в языковых моделях.
- Термодинамика: Аналогична термодинамической энтропии, отражает меру беспорядка в системе.

Также есть такие понятия, как взаимная энтропия и условная энтропия. Их определения появятся в конспекте после того, как пройдет неделя со сдачей домашки.

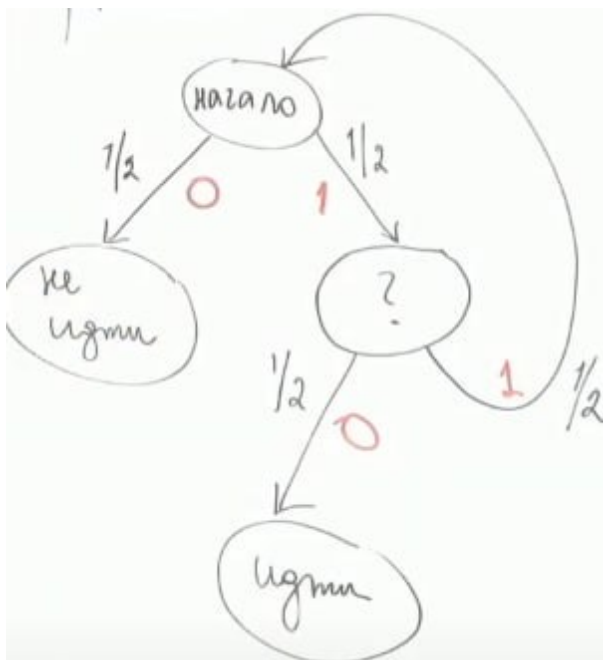
## 5 Лекция 5.

### 5.1 Определение цепи Маркова

Вспомним задачу с домашних заданий:

Петя хочет пойти в кино с вероятностью ровно  $\frac{1}{3}$ , а у него есть только честная монета. Может ли он осуществить свой замысел?

Как вы знаете она обладает вот таким решением:



И действия Пети задаются этим ориентированным графом с весами на ребрах. Добавим петли в ребра идти и не идти с вероятностью 1.

Граф с конечным числом вершин, где на каждом ребре написаны вероятности, и где сумма чисел на ребрах исходящих из каждой вершины равна 1, называется цепь Маркова или Марковская цепь.

Посмотрим на распределение вероятностей. Посмотрим на вектор:

$b = (b_1, b_2, \dots, b_n)$ , где  $b_i$  - вероятность оказаться в  $i$ -ой позиции.

Возьмем матрицу  $P$ , где  $p_{ij}$  - вероятность перейти из  $i$  в  $j$ . Тогда на нашем примере была вот такая матрица:

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Поглощающее состояние - то, которое переходит само в себя с вероятностью 1.

Это матрица перехода для цепи Маркова.

Пусть мы перешли в новое состояние:  $c = (c_1, \dots, c_n)$ .

$$c_i = P(c = i) = \sum_{j=1}^n P(c = i | B = j) P(b = j) = \sum_{j=1}^n p_{ji} b_j$$

Откуда  $c^T = b^T P$ . Есть еще другое обозначение  $\vec{c} = c^T$ .

Посмотрим на наши шаги:

$b^0 = (1, 0, 0, 0)$  - наш вектор  $b$  изначально

$b^1 = (0, \frac{1}{2}, \frac{1}{2}, 0)$

$b^2 = (\frac{1}{4}, \frac{1}{2}, 0, \frac{1}{4})$

$b^3 = (0, \frac{5}{8}, \frac{1}{8}, \frac{1}{4})$

и так далее. Логично, что  $\vec{b}^n = \vec{b} \cdot P^n$ .

Мы смотрели с точки зрения линейной алгебры. Давайте посмотрим граф цепи Маркова (смотрим только ненулевые ребра). Тогда в нем есть такие виды вершин:

1. поглощающие (существенное)
2. непоглощающие (не существенное) - несколько выходов

Цепь Маркова - поглощающая, если из любого состояние можно дойти до поглощающего состояния.

## 5.2 Эргодический класс

Эргодический класс - компонента сильной связности графа марковской цепи.

todo: вставить определение сильной связности.

Эргодический класс называется поглощающим, если из него не исходит ребер в другие эргодические классы.

Конденсацией марковской цепи называется конденсация графа.

Эргодическая марковская цепь - состоит из одного эргодического класса.

Эргодический класс периодическим с периодом  $d \neq 1$ , если длина любого цикла в этом эргодическом классе делится на  $d$  ( $d$  - максимально).

### Теорема (о классификации марковских цепей)

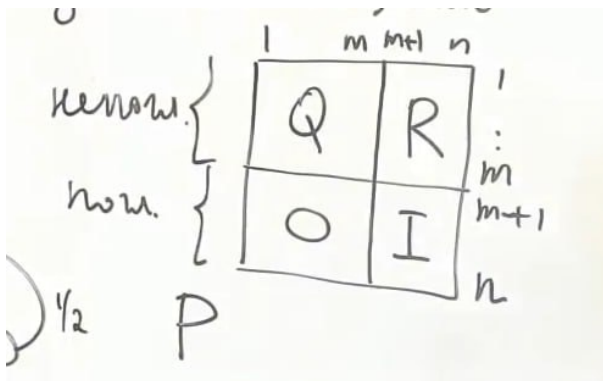
1.  $\forall$  марковская цепь содержит поглощающий эргодический класс.
2. Марковская цепь с вероятностью 1 рано или поздно оказывается в состоянии из поглощающего эргодического класса.
3. для непериодического поглощающего эргодического класса в случае попадания в него существует стационарное предельное распределение вероятностей  $b : b = bP$ .  $\forall$  начального распределение  $b^0, b^0 : b^0 P^n \rightarrow b$ .

Жизнь Марковской цепи крайне скучна - А.С. Станкевич

### 5.3 Уходим в математику.

$$b^1 = b \cdot P, b^n = b \cdot P^n$$

Дана поглощающая марковская цепь. Занумеруем, чтобы изначально шли непогл., а потом погл:



Где  $I$  = единичная матрица.

Возьмем  $a$  = началу вектора  $b$ , где любое состояние не поглощающее.

Заметим, что:  $a^n = a \cdot Q^n$  из-за нулей. Докажем, что  $Q^n \rightarrow 0$ .

#### Теорема о поглощении.

Поглощающая М.Ц переходит в погл. состояние с вероятностью 1.

todo: написать доказательство, оно несложное, я просто с температурой и не склеил

Хотим теперь понять, а где же мы поглотимся?

### 5.4 Мат. ожидание времени до поглощения

Давайте рассмотрим мат. ожидание времени до поглощения.

Есть  $b_0$  - начальное распределение.  $T$  - сл. величина: число шагов до погл:

$$T = \sum_{i=1}^m T_i, \text{ где } T_i - \text{число посещений } i\text{-ого состояния.}$$

$$T_i = \sum_{j=0}^{\infty} T_{ij}, \text{ такая что } T_{ij} = \begin{cases} 1, & \text{если на } j\text{-ом ходу в сост } i \\ 0, & \text{иначе} \end{cases}$$

$$ET = \sum_{i=1}^m ET_i = \sum_{i=1}^m \sum_{j=0}^{\infty} ET_{ij} = \sum_{i=1}^m \left( \sum_{j=0}^{\infty} a^0 Q^j \right)_i = \sum_{i=1}^m \left( a_0 \sum_{j=0}^{\infty} Q^j \right)_i = \sum_{i=1}^m (a_0 N)_i;$$

$$\text{Как мы знаем из линейной алгебры: } \sum_{j=0}^{\infty} Q^j = (I - Q)^{-1}$$

$N = (I - Q)^{-1}$  - фундаментальная матрица поглощающей марковской цепи.

$A = a^0 N R$  - распределение вероятностей поглощения в состояниях



## 6 Информация о курсе

Поток — у2024.

Группы М3138-М3142.

Преподаватель — Станкевич Андрей Сергеевич.

В данном семестре фокусируются 2 темы: Дискретная теория вероятности и представление слов (токенов) в компьютере.

