

Конспект по дискретной математике
II семестр

Коченюк Анатолий

28 марта 2021 г.

Глава 1

Дискретная теория вероятностей

1.1 Введение

Определение 1 (Вероятностное пространство).

Ω – элементарные исходы, неделимые дальше.

p – дискретная плотность вероятности.

$$p : \Omega \rightarrow [0, 1] \quad \sum_{\omega \in \Omega} p(\omega) = 1$$

Замечание. В случае дискретного вероятностного пространства $|\Omega|$ – не более, чем счётное.

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

Пример (Нечестная монета). $\Omega = \{0, 1\}$ $p(1) = p, p(0) = q$ – различные числа. $p + q = 1$

Ещё одно название – распределение Бернулли

Пример (Честная игральная кость). $\Omega = \{1, 2, 3, 4, 5, 6\}$ $p(\omega) = \frac{1}{6}$

Определение 2. Событие, случайное событие – $A \subseteq \Omega$

Замечание. Неправильное определение – то, что может произойти, а может не произойти.

$\emptyset \subseteq \Omega \quad \Omega \subseteq \Omega$ – примеры, когда никогда не происходит и всегда происходит

Замечание. Для не дискретного случая неверно, что любое подмножество Ω это событие

Определение 3. Вероятность события $P(A) = \sum_{\omega \in A} p(\omega)$
 p берёт элементарные исходы. P, \mathbb{P} – вероятность события

Пример. Событие $E = \{2, 4, 6\} \quad P(E) = p(2) + p(4) + p(6) = \frac{3}{6} = \frac{1}{2}$
 $O = \{1, 3, 5\}$

Замечание. Не существует вероятностного пространства с бесконечным числом равновероятных исходов

$$p(\omega) = 0 \quad \sum = 0$$

$$p(\omega) = a > 0 \quad \sum = a \cdot (+\infty) = +\infty$$

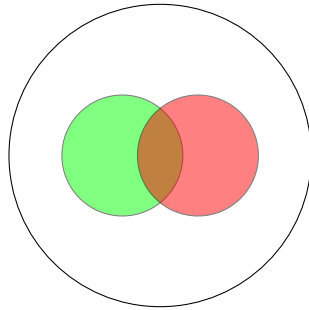
Пример. Событие $B(IG) = \{4, 5, 6\} \quad P(B) = \frac{1}{2}$

Определение 4 (Независимое событие). События A, B независимы, если $P(A \cap B) = P(A) \cdot P(B)$

Пример. $E \cap O = \emptyset \quad B \cap E = \{4, 6\}$

$$P(E \cap O) = 0 \quad P(O) \cdot P(B) = \frac{1}{4} \neq 0$$

$$P(B) \cdot P(E) = \frac{1}{4} \neq \frac{1}{3} = P(B \cap E)$$



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

$$\frac{P(A \cap B)}{P(B)} = \frac{P(A)}{P(\Omega)}$$

Определение 5 (Условная вероятность). $P(A|B) = \frac{P(A \cap B)}{P(B)}$

Замечание. Альтернативное определение независимости, не поддерживающее 0: $P(A|B) = P(A)$

$$V = \{5, 6\}$$

$$P(V \cap E) = \frac{1}{6}$$

$$P(V) = \frac{1}{3} \quad P(E) = \frac{1}{2} \quad P(V) \cdot P(E) = \frac{1}{3} \cdot \frac{1}{2} = \frac{1}{6} = P(V \cap E)$$

Определение 6 (Произведение вероятностных пространств).

$$\Omega_1, p_1 \quad \Omega_2, p_2$$

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

$$p(\langle \omega_1, \omega_2 \rangle) = p_1(\omega_1) \cdot p_2(\omega_2)$$

Теорема 1. $\forall A_1 \subseteq \Omega_1$ и $\forall A_2 \subseteq \Omega_2$

$A_1 \times \Omega_2$ и $\Omega_1 \times A_2$ – независимы

$$\begin{aligned} \text{Доказательство. } P(A_1 \times \Omega_2 \cap \Omega_1 \times A_2) &= P(A_1 \times A_2) = \sum_{\substack{a \in A_1 \\ b \in A_2}} p(\langle a, b \rangle) = \\ &= \sum_{a \in A_1} \sum_{b \in A_2} p_1(a) \cdot p_2(b) = \sum_{a \in A_1} p_1(a) \left(\sum_{b \in A_2} p_2(b) \right) = P_1(A_1) \cdot P_2(A_2) \quad \blacksquare \end{aligned}$$

Определение 7. A_1, A_2, \dots, A_n

1. Попарно независимые A_i и A_j независимы

2. Независимы в совокупности $\forall I \subseteq \{1, 2, \dots, n\} \quad P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

Пример. Кидаем две монеты $\Omega = \{00, 01, 10, 11\}$

$A_1 = \{10, 11\}$ $A_2 = \{01, 11\}$ $A_3 = \{01, 10\}$ – независимы попарно, но не в совокупности

Определение 8 (Формула полной вероятности).

$$\Omega = A_1 \cup A_2 \cup \dots \cup A_n \quad i \neq j \implies A_i \cap A_j = \emptyset$$

Совокупность таких A -шек называется полной системой событий.

Дано: вероятности $P(A_i)$ $P(B|A_i)$ Найти: $P(B)$

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

– формула полной вероятности

Найти: $P(A_j|B)$

A_1 – болен, A_2 – здоров, B – положительный результат теста

$P(A_2|B)$

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

– формула Байеса

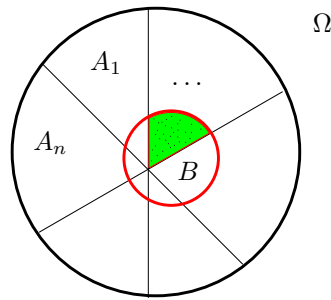


Рис. 1.1: В

1.2 Случайные величины

Замечание. Неправильное (наивное) определение – величина, принимающая случайное значение.

Она может быть константой. Что такое величина?

Определение 9 (Случайная величина). $\xi : \Omega \rightarrow \mathbb{R} - \mathbb{R}$ -значная функция

Ω, p – вероятностное пространство.

Пример. Если взять случайные текст длиной 1Кб. Вариантов текста очень много и бессмысленно их рассматривать отдельно, интересует какое-то свойство, величина.

Графы, $2^{\binom{n}{2}}$ штук. Но нас интересует какая-то (численная) характеристика элементарного исхода.

Пример. $D(ice) = \{1, 2, 3, 4, 5, 6\}$

$$\Omega = D^2 \quad p(\langle i, j \rangle) = \frac{1}{36}$$

$$\xi : \Omega \rightarrow \mathbb{R} \quad \xi(\langle i, j \rangle) = i + j$$

Пример (Случайные графы). $G(4, \frac{1}{2})$ – случайный граф, 4 вершины, каждое ребро существует с вероятностью $\frac{1}{2}$

$$\Omega = \mathbb{B}^6 \quad p(G) = \frac{1}{64}$$

$\xi(G)$ = количество компонент связности

Пример. $\Omega = \{1, 2, 3, 4, 5, 6\}$ $\xi(w) = w$

Пример. $\Omega = \{1, 2, 3, 4, 5, 6\}$ $E = \{2, 4, 6\}$

$$\chi_E(\omega) = \begin{cases} 1, & \omega \in E \\ 0, & \omega \notin E \end{cases} - \text{индикаторная случайная величина}$$

Определение 10. Ω, p, ξ

$$[\xi = i] = \{\omega | \xi(\omega) = i\} \subseteq \Omega$$

$$P([\xi = i]) = P(\xi = i) = f_\xi(i) \quad f : \mathbb{R} \rightarrow \mathbb{R}$$

$f_\xi(i) = P(\xi = i)$ – дискретная плотность вероятности случайной величины ξ

$$F_\xi(i) : \mathbb{R} \rightarrow \mathbb{R} = P(\xi \leq i) - \text{функция распределения}$$

Замечание. Непрерывная vs Дискретная вероятность

Пример.

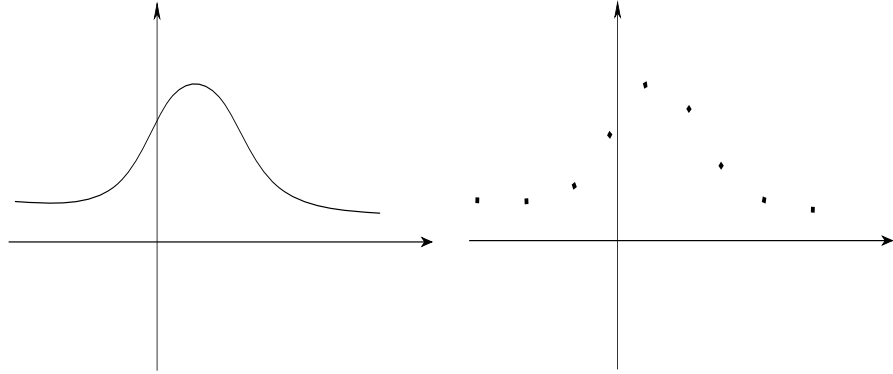


Рис. 1.2: непрерывная и дискретная вероятность

Замечание. $\delta(x) = \begin{cases} 0, & x \neq 0 \\ +\infty, & x = 0 \end{cases}$

$$\int_{-\infty}^{+\infty} \delta(x) dx = 1$$

$$f_{\xi}(i) = P(\xi = i) \quad f_{\xi} = \int_{-\infty}^{+\infty} f_{xi}(x) = F_{\xi}(i)$$

$$f_{\xi}(x) = \sum_i P(\xi = i) f(x - i)$$

Пример. $\Omega = \mathbb{B}^{1000} \quad p(\omega) = \frac{1}{2^{1000}}$

$\xi(w) = \text{число } 1 \text{ в } \omega$

$$|\text{множество значений } \xi| = 1001 \quad p(\xi = i) = \frac{\binom{1000}{i}}{2^{1000}}$$

Замечание. Случайные числа обозначаются строчными греческими или заглавными латинскими из конца алфавита (X, Z)

Замечание (Что можно делать со случайными величинами). ξ, η – функции

$\xi^2 \quad 2\xi \quad \xi + \eta \quad \xi \cdot \eta \quad \xi^\eta \quad \sin \xi \quad e^\eta \quad \frac{1+\xi}{\eta}$ (всё то же, что мы можем делать с функциями).

Пример. $\Omega = D^2 \quad \xi_1(\langle i, j \rangle) = i \quad \xi_2(\langle i, j \rangle) = j$ – одинаково распределённые случайные величины

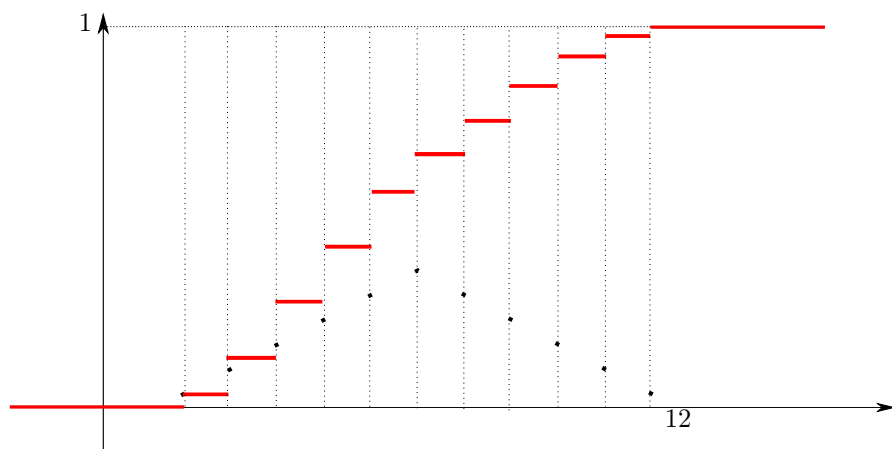


Рис. 1.3: Два-кубика (функция распределения)

Пример. $\Omega = F$ $id(\omega) = \omega$

$1, 2, \dots, 6$ – каждый с вероятностью $\frac{1}{6}$. Другое вероятностное пространство относительно предыдущего примера, но всё равно одинаковое распределение.

$\xi = (i + j) = \xi_1 + \xi_2$ $\xi = (i + j) \% 6 + 1$ – у второй то же распределение, что у верхних, но она уже совсем другая.

Определение 11. Математическое ожидание

$$E_{\xi} = \sum_{\omega} p(\omega) \xi(\omega).$$

Утверждение 1. $E_{\xi} = \sum_i ip(\xi = i)$

Доказательство.

$$\begin{aligned}
 E_{\xi} &= \sum_{\omega} p(\omega) \xi(\omega) \\
 &= \sum_i \sum_{\omega: \xi(\omega)=i} p(\omega) \cdot \xi(\omega) \\
 &= \sum_i \sum_{\omega: \xi(\omega)=i} p(\omega) \cdot i \\
 &= \sum_i i \sum_{\omega: \xi(\omega)=i} p(\omega) \\
 &= \sum_i iP(\xi = i)
 \end{aligned}$$

■

Пример. $\Omega = D \quad \xi = id$

$$E_{\xi} = \frac{1}{6} \cdot 1 + \frac{1}{6} \cdot 2 + \frac{1}{6} \cdot 3 + \frac{1}{6} \cdot 4 + \frac{1}{6} \cdot 5 + \frac{1}{6} \cdot 6 = \frac{21}{6} = \frac{7}{2} = 3,5$$

Пример. $\Omega = D^2 \quad \xi(\langle i, j \rangle) = i + j$

$$E_{\xi} = \frac{1}{36} \cdot (2 \cdot 1 + 3 \cdot 2 + 4 \cdot 3 + 5 \cdot 4 + 6 \cdot 5 + 7 \cdot 6 + 8 \cdot 5 + 9 \cdot 4 + 10 \cdot 3 + 11 \cdot 2 + 12 \cdot 1)$$

– здесь среднее значение оказалось наиболее частым, но так оно не всегда (пример с 3,5)

Теорема 2. $E_{\lambda\xi} = \lambda E_{\xi}$

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

Доказательство. $E_{\lambda\xi} = \sum_{\omega} p(\omega) \lambda\xi(\omega) = \lambda E_{\xi}$

$$E(\xi + \eta) = \sum_{\omega} p(\omega) (\xi(\omega) + \eta(\omega)) = E_{\xi} + E_{\eta}$$

■

Утверждение 2. Если ξ и η одинаково распределены, то $E_{\xi} = E_{\eta}$

Пример. Бросим кубик один раз, ξ_1 – что выпало сверху, ξ_2 – что выпало снизу

$E(\xi_1 + \xi_2) = 7$. – не играет роли как числа друг относительно друга расположены.

МАТОЖИДАНИЕ ЛИНЕЙНО ВСЕГДА

Пример. $\Omega = S_n$ $p(\omega) = \frac{1}{n!}$

$\xi(\pi) = |\{i | \pi[i] = i\}|$ $0 \dots n$, кроме $n-1$

$$E_\xi = \sum_{j=1}^n \xi_j = 1$$

$$\xi_i(\pi) = \begin{cases} 1, & \pi[i] = i \\ 0, & \text{иначе} \end{cases}$$

$$E_{\xi_i} = \frac{1}{n}$$

$$\xi = \sum_{j=1}^n \xi_j$$

1.3 Независимые случайные величины

Определение 12 (удобное). Случайные величины ξ и η независимы, если события $[\xi = \alpha]$ и $[\eta = \beta]$ – независимы $\forall \alpha, \beta$

Определение 13 (нормальное). $[\xi \leq \alpha]$ и $[\eta \leq \beta]$ – независимы для $\forall \alpha, \beta$

Пример. $\Omega = \Omega_1 \times \Omega_2$

$$\xi_1(\langle \omega_1, \omega_2 \rangle) = f(\omega_1)$$

$$\xi_2(\langle \omega_1, \omega_2 \rangle) = g(\omega_2)$$

A и B независимы, χ_A, χ_B – независимы

Теорема 3. ξ, η – независимы $\implies E(\xi \cdot \eta) = E_\xi \cdot E_\eta$

Доказательство. $E\xi \cdot \eta = \sum_{\alpha} \alpha \cdot P(\xi \cdot \eta = \alpha) = \sum_{i,j} \alpha P([\xi = i] \cap [\eta = j]) =$
 $\sum_i \sum_j i j P(\xi = i) P(\eta = j) = E_\xi E_\eta$

$$i \cdot j = \alpha \quad i \in R_\xi \quad j \in R_\eta$$

■

Пример. $\Omega = \{0, 1\}$ $p = \frac{1}{2}$ $\xi(i) = 2i$ $E_\xi = 1$

$\Omega = S_n$ $p = \frac{1}{n!}$ ξ = число неподвижных точек $E_\xi = 1$

Матожидание одно, но ведут себя совершенно по разному.

Определение 14 (Дисперсия). $D_\xi = Var(\xi)$

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

Теорема 4. $D_{c\xi} = c^2 D_\xi$

Если ξ и η независимы, то $D_{\xi+\eta} = D_\xi + D_\eta$

Доказательство. Упражнение ■

Вспомним. $\xi, \eta : \Omega \rightarrow \mathbb{R}$

$$F_\xi(a) = P(\xi \leq a)$$

$$f_\xi(a) = P(\xi = a)$$

$$F_\xi(a) = \sum_{b \leq a} f_\xi(b)$$

$$E_\xi = \sum_{\omega \in \Omega} p(\omega) \xi(\omega) = \sum_a a \cdot P(\xi = a)$$

$$E(\xi + \eta) = E_\xi + E_\eta \quad \blacksquare$$

$E(\xi - E_\xi) = E_\xi - EE_\xi = 0$ матожидание отклонения от матожидания равно нулю..

Хочется смотреть насколько величина отклоняется от своего матожидания. Для этого используется понятие дисперсии:

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

$$D(\xi + \eta) = E\xi^2 + E\eta^2 + 2E\xi\eta - (E\xi)^2 - (E\eta)^2 - 2E\xi E\eta = D_\xi + D_\eta + 2(E_{\xi\eta} - E_\xi E_\eta).$$

В случае независимых случайных величин дисперсия линейна. Иначе она отличается на ковариацию:

Определение 15 (Ковариация).

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

$$D_\xi = Cov(\xi, \xi)$$

Определение 16 (Корреляция).

$$Corr(\xi, \eta) = \frac{Cov(\xi, \eta)}{\sqrt{D_\xi D_\eta}}.$$

Теорема 5. Корреляция двух случайных величин лежит между -1 и 1.

$$-1 \leq Corr(\xi, \eta) \leq 1.$$

Доказательство. $\alpha = \xi - \lambda\eta$

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

$$D_\xi + 2\lambda Cov(\eta, \eta) + \lambda^2 D_\eta$$

$$4Cov(\xi, \eta)^2 - 4D_\xi D_\eta \leq 0 \quad \blacksquare$$

1.4 Хвостовые неравенства

$$\xi \quad E\xi = 10 \quad \xi \geq 0$$

$$P(\xi \geq 100) < \frac{1}{10}$$

Теорема 6 (Неравенство Маркова). $\xi \neq 0 \quad \xi \geq 0 \quad P(\xi \geq a \cdot E_\xi) \leq \frac{1}{a}$

Доказательство. $E_\xi = \sum_v v \cdot P(\xi = v) = \sum_{v < a \cdot E_\xi} v \cdot P(\xi = v) + \sum_{v \geq a \cdot E_\xi} v \cdot P(\xi = v) \geq 0 + a \cdot E_\xi \cdot \sum_{v \geq a \cdot E_\xi} P(\xi = v) = a \cdot E_\xi \cdot P(\xi \geq a \cdot E_\xi) \quad \blacksquare$

Пример. $a = \frac{c}{E_\xi} \quad P(\xi \geq c) \leq \frac{E_\xi}{c}$

$$D_\xi = E(\xi - E\xi)^2$$

$$\eta = (\xi - E_\xi)^2$$

$$P((\xi - E_{xi})^2 \geq a^2 \cdot D_\xi) \leq \frac{1}{a^2}$$

$\sigma = \sqrt{D_\xi}$ – среднее квадратичное отклонение

Теорема 7 (Неравенство Чебышева). $P(|\xi - E_\xi| \geq a\sigma) \leq \frac{1}{a^2}$

$$P(|\xi - E_\xi| \geq c) \leq \frac{D_\xi}{c^2}$$

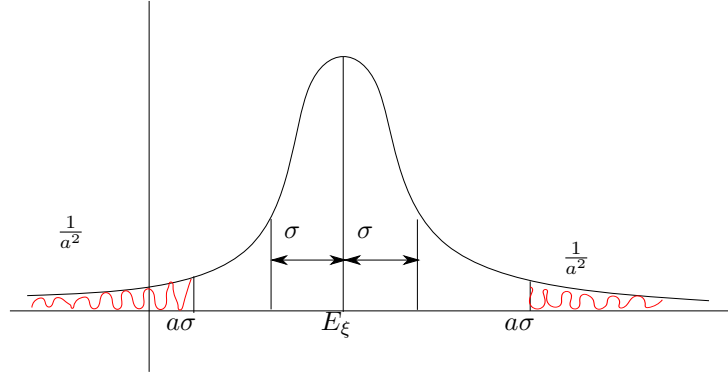


Рис. 1.4: drawing

Задача 1 (10 монет, найти количество “1”).

$$E\xi = 5 \quad D\xi = 2,5$$

$$P(\xi \leq 0) \leq P(|\xi - E\xi| \geq 5) \leq \frac{2,5}{25} = \frac{1}{10}$$

Замечание. С одной стороны, у неравенства есть плюс: оно **всегда** работает; всегда (!)

С другой — иногда оценки получаются, очень грубыми. В нашем примере ответ $\leq \frac{1}{10}$, а в жизни — $\frac{1}{1024}$.

Пример. Нечестная монета $p \neq \frac{1}{2}$. Хотим выяснить чем чаще выпадает.

Бросили: c единиц, $n - c$ нулей. Предположим, что $c < \frac{n}{2}$ $p > \frac{1}{2}$ $pn > \frac{n}{2}$

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

Теорема 8 (Граница Чернова (без доказательства)). ξ_i $P(\xi_i = 1)$ $P(\xi_i = 0) = 1$

$$\xi = \sum_{i=1}^n \xi_i \quad E\xi = np = \mu$$

$$P(|\xi - \mu| \geq \delta\mu) < e^{-\mu \frac{\delta^2}{3}}$$

Пример. Случайная величина ξ . Хотим узнать матожидание. Проведём эксперимент n раз: $\xi_1, \xi_2, \dots, \xi_n$

$$P\left(\left|\frac{\sum \xi_i}{n} - E_\xi\right| > c\right) \leq \frac{D_\xi}{n\varepsilon^2}.$$

$$\xi : \Omega \rightarrow \mathbb{Z}^+$$

$$E_\xi = \sum_{i=0}^n i \cdot P(\xi = i) = \sum_{i=0}^n (P(\xi \geq i) - P(\xi \geq i+1)) = \sum_{i=1}^n P(\xi \geq i)$$

1.5 Теория информации

Определение 17 (Что такое информации).
Информация = – неопределённость

неопределённость Н1. Что-то узнали, стала неопределённость Н2. полученная информация $I = \text{Н1} - \text{Н2} = -\Delta \text{Н}$

Хочется убрать наблюдателя, нас, из определения, чтобы не было кого-то, кто узнаёт и меняет неопределённость. Надо ввести объективную модель:

Определение 18 (Случайный источник). Ω – вероятностное пространство.
Есть исходы p_1, p_2, \dots, p_n

Чёрный ящик с красной кнопкой и дисплеем. Основан на вероятностном пространстве

$$\xi_1, \xi_2, \dots, \xi_m \dots$$

$$P(\xi_i = a) = p_a \quad a = 1 \dots n$$

Случайный источник p_1, p_1, \dots, p_n . Хотим померять сколько информации содержится в одном результате эксперимента.

$$H(p_1, p_2, \dots, p_n) : RS(randomsources) \rightarrow \mathbb{R}^+$$

Частный случай $p_i = \frac{1}{n}$

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

$$1. \quad h(n+1) > h(n)$$

2.

Пример. $\Omega = \{(1, 1), (1, 2), \dots, (1, m_1)(2, 1), \dots, (2, m_2), \dots, (k, 1), \dots, (k, m_k)\}$

$$n = m_1 + m_2 + \dots m_n$$

$$p(i, j) = 1_{ij} \quad p_i = \sum_{j=1}^{m_k} q_{ij}$$

Первый ряд $(1, *) - p_1$. Второй $p_2 \dots$ Последний p_k

Если случайный источник показывает только первое число это эквивалентно $H(p_1, p_2, \dots, p_n)$

Теперь представим, что мы сначала узнаём первую компоненту, а потом открываем вторую

$$\sum_{i=1}^k p_i H\left(\frac{q_{i1}}{p_i}, \dots, \frac{q_{im_i}}{p_i}\right)$$

Если провести эксперимент сразу, получим q_{11}, \dots, q_{im_i}

$$q_{ij} = p_i q_{ij}$$

$$H(p_1 r_{11}, p_1 r_{12}, \dots, p_1 r_{1k_1}, p_2 r_{21}, \dots, p_k r_{km_k}) = H(p_1, p_2, \dots, p_k) + \sum_{i=1}^k p_i H(r_{i1}, \dots, r_{im_i}).$$

3. Для фиксированного n H непрерывная как функция $\mathbb{R}^n \rightarrow \mathbb{R}$

$$\textbf{Теорема 9. } H(p_1, p_2, \dots, p_n) = - \sum_{i=1}^n p_i \log p_i$$

Лемма 1. $h(nm) = H(n) + h(m)$ Следует из второго свойства

$$\textit{Доказательство. } k = n \quad m_i = m \quad p_i = \frac{1}{n} \quad q_{ij} = \frac{1}{nm} \quad r_{ij} = \frac{1}{m}$$

$$h(nm) = H(q_{11}, q_{12}, \dots, q_{nm}) = H\left(\frac{1}{n}, \dots, \frac{1}{n}\right) + \sum_{i=1}^n \frac{1}{n} H\left(\frac{1}{m}, \dots, \frac{1}{m}\right) = h(n) + h(m) \quad \blacksquare$$

Определение 19. $h(2) = \alpha$ (может с точностью до мультипликативной константы задать)

$$\textbf{Лемма 2. } h(2^k) = k\alpha$$

$$\textbf{Лемма 3. } h(n) = \alpha \log_2 n$$

$$\textit{Доказательство. } 2^i \leq n^r < 2^{i+1} \quad r \in \mathbb{N}$$

$$h(i) \leq h(n^r) < \alpha(i+1)$$

$$\alpha \cdot i \leq r \cdot h(n) < \alpha(i+1)$$

$$\alpha \cdot \frac{i}{r} \leq h(n), \alpha \frac{i+1}{r}$$

$$i \leq r \log_2 n < i+1$$

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

$$\forall r \quad |h(n) - \alpha \log_2 n| \leq \frac{\alpha}{r}$$

$$\implies h(n) = \alpha \log_2 n$$

■

Доказательство теоремы. Рациональный $p_i = \frac{a_i}{b}$ $m_i = a_i$ $r_{ij} = \frac{1}{a_i}$ $q_{ij} = \frac{1}{b}$ $q_{ij} = p_i r_{ij}$

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

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

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

Функция непрерывна и она верна для рациональных, следовательно она верна для всех

■

Замечание. $h(2) = \alpha$ – бит

А теперь мы хотим перевести определение информации на неслучайный источник

Ответ 1. Это тогда будет не совсем корректно с математической точки зрения. Когда смотришь на конкретные детерминированные данные.

■

Ответ 2. Изучение среднего не совсем антинаучное занятие. Внешне оно ведёт себя как случайные величины.

■

Пример. Есть строка s , в которой мы хотим померить информацию.

$$s \in \Sigma^* \quad n = |\Sigma|$$

$$|s| = l \quad f_i - \text{количество символов } c_i \text{ в строке } s$$

$$p_i = \frac{f_i}{L}$$

Допустим, что символы выдаёт случайный источник, который выдал символы s_1, s_2, \dots, s_L . Статистически эта строка похожа на s . *натягивание на глобус* Допустим, что количество информации в строке s равно количеству в строке \tilde{s}

$$I(\tilde{s}) = J \cdot H(p_1, p_2, \dots, p_n) = -L \sum_{i=1}^n p_i \log_2 p_i$$

Вспомним арифметическое кодирование

$q = A(s)$ – длина арифметического кодирования

$$\begin{aligned} A(s) &\leq -\log_2(b_L - a_L) = -\log_2(p_{s_1} \cdot \dots \cdot p_{s_L}) = -\log_2\left(\prod_{i=1}^n p_i^{f_i}\right) = -\sum_{i=1}^n \underbrace{\frac{L}{p_i}}_{p_i} \log_2 p_i = \\ &= I(\tilde{s}) = L \cdot H(p_1, p_2, \dots, p_n) \end{aligned}$$

Теорема 10. Длина кода, после арифметического кодирования не превышает энтропию Шеннона

Замечание. Арифметическое кодирование асимптотически оптимально среди тех, которые не учитывают взаимное расположение символов.

Пример (Нижняя оценка для сортировки). Пусть a_1, \dots, a_n – перестановка и мы хотим её отсортировать

Утверждение 3. От одного сравнения мы получаем не больше 1 бита информации

Рассмотрим все перестановки. В каждой содержится

$$\log_2 n! = \sum_{i=1}^n \log_2 i \geq \sum_{i=\frac{n}{2}}^n \log_2 \frac{n}{2} = \Omega(n \log n)$$

1.6 Цепи Маркова

$b = (b_1, b_2, \dots, b_n)$ – b_i вероятность находиться в состоянии i

$C \quad c = (c_1, c_2, \dots, c_n)$ – случайная величина после одного перехода

Матрица перехода p_{ij} – вероятность перейти из i в j

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

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

$b^0 = (1, 0, 0, 0)$ – нулевой шаг

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

Рассмотрим судьбу м.ц. после поглощения. Жизнь происходит внутри одной сильно связанной компоненты – эргадического класса.

1. $d > 1$ длина любого цикла кратна d . Циклический класс
2. $\text{НОД}(\text{длин всех циклов}) = 1$.

Теорема 11 (Эргадическая для регулярных цепей). М.ц. такова, что $p_{ij} > 0 \forall i, j$

Тогда $\exists b \quad \forall b^0 \quad b^0 P^n \rightarrow b$

(b удовлетворяет равенству $b = bP$)

Доказательство. $(b^0 A)_i = \sum_{j=1}^n b_j^0 \cdot A_{ji} = \left(\sum_{j=1}^n b_j^0 \right) \tilde{a}_i = \tilde{a}_i$

$$\square \forall j \quad a_{ji} = \tilde{a}_i$$

$P^n \rightarrow A$, которая удовлетворяет условию выше.

$$m_i^t = \min_j (P^t)_{ji} \quad M_i^t = \min_j (P^t)_{ji}$$

$$M_i^t - m_i^t \rightarrow 0$$

$$\delta = \min_{i,j} \delta > 0$$

$$\begin{aligned} P_{ji}^{t+1} &= \sum_{k=1}^n P_{jk}^t P_{ki} \\ &\leq \overbrace{\sum_{k=1}^n P_{jk}^t M_i^t}^1 + P_{j \text{ posMin}} (m_i^t - M_i^t) \\ &\leq M_i^t + \delta (m_i^t - M_i^t). \end{aligned}$$

Аналогично с максимумом, оцениваем всё снизу минимумов, кроме максимума

$$M_i^{t+1} \leq M_i^t + \delta (m_i^t - M_i^t).$$

$$-m_i^{t+1} \leq -m_i^t + \delta (m_i^t - M_i^t).$$

$$M_i^{t+1} - m_i^{t+1} \leq (M_i^t - m_i^t) (1 - 2\delta) \leq (1 - 2\delta)^{t+1} \rightarrow 0.$$

Теперь у $b = bP$

$$(I - P)b = 0$$

$$\text{Rg}(I - P) = n - 1$$

$$\sum b_i = 1$$

$$P^{2^c}$$

$$bP^n 0 > b \quad bP^{n+1} \rightarrow bP$$

$$b = bP$$



Вернёмся к вопросу что происходит после поглощения.

$$\triangleleft \text{эргадический класс } A \quad \tilde{p} = \sum_{a \in A} (b^o N R)_a$$

$$\tilde{b}^0 = (b^0 N R)_{A - \frac{1}{p}}$$

$$\exists \text{ предельное } b : \quad \tilde{b}^0 A^n \rightarrow b$$

Конечное распределение $b\tilde{p}$

Скрытые Марковские модели. Мы решали до этого прямую задачу – брали м.ц. с известными матрицами перехода и смотрели на их характеристики.

Есть обратная: Есть состояние и мы хотим узнать матрицу перехода.

Ещё задача: Есть немарковский процесс и мы хотим аппроксимировать его марковским.