Теория информации

0.1 Теория вероятности

0.1.1 Основные сведения из теории вероятности

◄ 05.09.13 ►

Введение

Термин **информация** в курсе будет пониматься в узком научном смысле.

Теория информации — специальная математическая дисциплина. Её содержанием является абстрактно формулируемые теоремы и модели. ТИ имеет обширное применение к теории передачи сообщений, записывающих устройств, матлингвистике, компьютерной технике.

В самом общем виде теория информации понимается как теория передачи сигналов по линиям связи. Наиболее важное понятие ТИ — сама информация. В нашей жизни большую роль играет информация и связанные с ней операции: передача, получение, обработка, хранение.

Информация имеет две стороны: количественную и качественную. Иногда важно получение общего количества информации (количественная сторона), иногда важно конкретное содержание самой ИИ. Отметим, что переработка ИИ является технически сложной процедурой, которая усложняет разработку общей теории информации.

Важнейшим этапом в открытии основных закономерностей ТИ были работы американского инженера-связиста, математика Клода Шеннона (1947-49гг).

Для вычисления количества информации была предложена т.н. **логарифмическая мера**. Понятие **количества информации** тесно связано с понятием энтропии как меры степени неопределённости. Приобретение информации сопровождается уменьшением неопределённости, следовательно, количество информации можно измерять количеством "исчезнувшей неопределённости" (энтропии).

Теория информации является математической теорией, использующей понятия и методы теории вероятности.

Вероятность. Случайные события и величины

Пусть производится серия из N опытов, причём некоторое событие A происходит в $N_a < N + 1$. Тогда $h_n(A) = N_a/N$ называется частотой появления события A в серии из N опытов. Известный факт: с ростом $N \quad h_n(A) \to p$ (постоянная p - вероятность появления случайного события A).

Наука, изучающая свойства вероятности и применение этого понятия называется **теория вероятности**.

Событие, которое при выполнении некоторого комплекса условий обязательно выполняется называется **достоверным событием**.

Событие, которое при выполнении некоторого комплекса условий не выполняется называется **невозможным событием**.

Пример: Выпадение определённого числа очков на грани игральной кости – достоверное событие.

Выпадение семи очков на грани игральной кости – невозможное событие.

Случайное событие – событие, которое может произойти, а может и не произойти.

Задача: В урне 10 шаров : 5 белых, 3 чёрных и 2 красных. Найти вероятность выпадения шара определённого цвета (шары одинаковы).

Решение: Выписать случайные события:
A — {вынутый шар белый} $P(A) = 5/10 = 1/2$;
B — {вынутый шар чёрный} $P(B) = 3/10$;
C — {вынутый шар красный} $P(C) = 2/10 = 1/5$.

Задача: Какова вероятность, что при бросании кости выпадет число очков, кратное 3?

Решение:
Кратны 3 $\{3, 6\}$.
N исходов $= 6$.
$P(A) = 2/6 = 1/3$.

Общий принцип решения задач сводится к понятию равновероятности или равновозможности. (например, все грани кости одинаковы, и вероятность выпадения той или иной грани равна 1/6).

Классическое определение вероятности Пусть из N возможных исходов опыта случайное событие A появляется M раз. Тогда вероятность случайного события A в модели с равновероятными исходами вычисляется по формуле $P(A) = M/N$

Каждому опыту отвечает своя таблица вероятности. К примеру, в задаче с урнами и шарами таблица вероятности имеет вид:

События	A	B	C
Вероятности	$P(A) = 1/2$	$P(B) = 3/10$	$P(C) = 1/5$

Можно сказать, что в опыте с бросанием кости число очков, выпадающих на грани является случайной величиной, которая может принимать одно из возможных 6 числовых значений в зависимости от случая.

Итак, случайная величина — числовая функция, принимающая то или иное числовое значение в зависимости от случая.

Например, количество рождений в городе за год – случайная величина.

Свойства вероятности. Сложение и умножение случайных событий.

Несовместные и независимые случайные события. Из определения вероятности $\sim$ основные свойства вероятности случайного события A:

1 $0 \leqslant P(A) \leqslant 1$.

$P(A) = 1$ – достоверное событие;

$P(A) = 0$ – невозможное событие.

2 Пусть опыт приводит к двум взаимоисключающим событиям или исходам A или B. В этом случае B называют противоположным A событием ($B = \bar{A}$)

Пусть $P(A) = \frac{m}{n}$;

Тогда $P(\bar{A}) = \frac{(n-m)}{n} = 1 - \frac{m}{n} = 1 - P(A) \Rightarrow P(A) = 1 - P(\bar{A})$.

3 Пусть случайное событие $A_1 \subset A$ влечёт появление события $A \Rightarrow P(A1) < P(A)$ $\sim\sim\sim$. ???Тогда эти события **совместны**.

4 Правило сложения вероятностей для двух событий:

4.1. Пусть A и B – несовместны.

Тогда $A \cap B = \varnothing$;

$P(A) = \frac{m_1}{n}$;

$P(B) = \frac{m_2}{n}$;

$P(A+B) = \frac{(m_1+m_2)}{n} = \frac{m_1}{n} + \frac{m_2}{n} = P(A) + P(B)$.

Таким образом, $P(A+B) = P(A) + P(B)$.

В примере с урной вероятность извлечь чёрный или белый шар равна $P(A+B) = P(A) + P(B) = \frac{1}{2} + \frac{3}{10} = \frac{4}{5}$;

Замечание: Пусть некоторый опыт проиводит к появлению K различных (взаимоисключающих) исходов:

Исходы	A1	A2	...	An
Вероятности	P1	P2	...	Pn

Заметим, что бывают случаи, когда

$$\sum_{i=1}^{k} P(A_i) = P(A_1 + A_2 + \ldots + A_n) = 1$$

В этом случае говорят, что события $A_1, A_2, \ldots, A_n$ составляют **полную группу** случайных событий, то есть $A_1, A_2, \ldots, A_n$ попарно несовместны.

$A_1, A_2, \ldots, A_n : A_i \cap A_j = \varnothing \, \forall \, i, j : i \neq j$; если $A_1 + A_2 + \ldots + A_n$ – достоверное событие.

4.2. Пусть A и B совместны. $P(A+B) = P(A) + P(B) - P(AB)$, где $P(AB)$ – вероятность одновременного происхождения двух случайных событий A и B.

Теорема сложения вероятности для совместных случайных событий. (диаграмма Венна: $A = m_1; B = m_2, A \cap B = l$).

$P(AB) = \frac{l}{n}$

$P(A+B) = \frac{(m_1+m_2-l)}{n} =$ (в m_1 и m_2 входит l) $= \frac{m_1}{n} + \frac{m_2}{n} - \frac{l}{n}$

События A и B называются **независимыми**, если результат выполнения события A не связан с результатом события B. (извлечение двух чёрных шаров из разных урн – независимые события)

5 Теорема умножения вероятности для двух независимых событий:

Если A и B независимы, то $P(AB) = P(A) * P(B)$

Пример 1: Какова вероятность при двух бросках монеты оба раза выпадет орёл?
$P(AB) = ?$
A\{орёл\} $P(A) = \frac{1}{2}$;
B\{решка\} $P(B) = \frac{1}{2}$;
$P(AB) = \frac{1}{4}$.

Пример 2: В колоде 52 карты, 4 масти, 2 козыря. Какова вероятность того, что взятая наугад карта 2 является тузом или козырем?
A\{туз\} $P(A) = 1/13$;
B\{козырь\} $P(B) = 1/4$;
$P(AB) = 1/52$;
A и B совместны, независимы.
$P(A+B) = P(A) + P(B) - P(AB) = 1/13 + 1/4 - 1/52 = 4/13$.

◄ 19.09.13 ►

Условная вероятность.

Рассмотрим пример: В урне M чёрных шаров и $N-M$ белых. Случайное событие A \{извлечение чёрного шара\} и
B \{извлечение чёрного шара из той–же урны после того, как из неё уже вынут один шар\}

$$P(B|A) = \frac{(m-1)}{(n-1)}$$

Поскольку, если событие A имело место, то в урне осталось $M-1$ чёрных шаров.

$$P(B|\bar{A} = \frac{m}{(n-1)}\,\bar{A}$$

Вероятность события B здесь разная. Вероятность, которую имеет событие B в том, случае, когда известно, что событие A имело место называется **условной вероятностью** события B при условии выполнения события A.

$$P(B|A)P(B|A) \sim\sim\sim Pa(B)$$

Условные вероятности можно вычислять аналогично вычислению безусловных вероятностей.

В случае если A и B независимы, $P(A|) = P(A) * P(B)$.

В случае зависимости $P(AB) = P(A) * P(B|A) = P(B) * P(A|B)$.

В обоих случаях мы имеем правило умножения вероятностей. В одном случае для независимых событий, в другом для зависимых. Последнее соотношение часто кладут в определение условной вероятности.

$$P(B|A) = \frac{P(AB)}{P(A)} \qquad P(A|B) = \frac{P(AB)}{P(B)}$$

Из предыдущей формулы можем составить пропорцию:

$$\frac{P(B|A)}{P(B)} = \frac{P(A|B)}{P(A)}$$

Из определения условной вероятности вытекают ее основные свойства:

1 $0 \leqslant P(B|A) \leqslant 1$, причём $P(B|A) = 1$ когда $A \subset B$; B — достоверное случайное событие.

$P(B|A) = 0 \Longleftrightarrow A, B$ несовместны, или известно, что B – невозможное событие.

2 Пусть $B_1 \subset B$ (появление B_1 вызывает событие B) $P(B1|A) \leqslant P(B|A)$

3 Если B и C несовместны $P(B + C|A) = P(B|A) + P(C|A)$ (теорема сложения вероятностей для несовместных событий)

4 $P(\bar{B}|A) = 1 - P(B|A)$

Замечание: Пусть имеется K (и только K) попарно несовместных исходов некоторого опыта $A_1, A_2, \ldots, A_k$, называемых гипотезами. Пусть некоторое случайное событие B может произойти при выполнении одной из гипотез. Тогда очевидно, что $B = A_1B + A_2B + \ldots + A_kB$ (все события A_iB несовместны, поэтому можно воспользоваться теоремой сложения вероятностей)

$$P(B) = P\left(\sum_{i=1}^{k} A_iB\right) = \sum_{i=1}^{k} P(A_iB) = \sum_{i=1}^{k} \big(P(A_i) * P(B|A_i)\big)$$

Формула носит название формулы полной вероятности

$$P(B) = \sum_{i=1}^{k} P(A_i) * P(B|A_i)$$

Задача: Имеется 5 урн : в двух по одному белому и пять чёрных шаров; в одной – 2 белых, 5 чёрных; в двух – 3 белых, 5 чёрных шаров. Наудачу выбирается одна урна. Из неё извлекается один шар. Какова вероятность того, что шар белый?

Решение: Выберем в качестве гипотез 3 способа

$A_1 : \{$Выбрана урна с 1 б.ш$\}$ $\quad P(A_1) = 2/5 \quad P(B|A_1) = 1/6$

$A_2 : \{$Выбрана урна с 2 б.ш$\}$ $\quad P(A_2) = 1/5 \quad P(B|A_2) = 2/7$

$A_3 : \{$Выбрана урна с 3 б.ш$\}$ $\quad P(A_3) = 2/5 \quad P(B|A_3) = 3/8$

$B =$ извлечён белый шар

$P(B) = \frac{1}{6} * \frac{2}{5} + \frac{2}{7} * \frac{1}{5} + \frac{3}{8} * \frac{2}{5} = \frac{23}{84}$

Математическое ожидание случайной величины. Основные свойства математического ожидания

Введение. Важнейшей числовой характеристикой ξ является её математическое ожидание или среднее значение, вычисляемое по правилу $M\xi = \sum_{i=1}^{n} x_ip_i)$, где x_i – принимаемые ξ значения, p_i – вероятности их выпадения.

С помощью математического ожидания мы можем сравнивать между собой две случайные величины (например, из двух стрелков лучший тот, кто выбивает в среднем наибольшее число очков), однако встречаются задачи, в которых знание одного лишь $M\xi$ недостаточно.

Пример: Пушка ведёт прицельный огонь по мишени, удалённой от пушки на расстояние a. Обозначим дальность полёта снаряда через ξ километров; $M\xi = a$

Отклонение $M\xi$ от a свидетельствует о наличии систематической ошибки (производственный дефект, неправильный угол наклона). Ликвидация систематической ошибки достигается изменением угла наклона орудия.

Вместе с тем, отсутствие систематической ошибки ещё не гарантирует высокую точность стрельбы. Чтобы оценить точность надо знать, насколько близко ложатся снаряды к цели.

Как определить точность стрельбы и сравнить между собой качество стрельбы двух орудий?

Отклонение снаряда от цели — $\xi - a$

$M(\xi - a) = M\xi - a = a - a = 0$

В среднем, положительные и отрицательные значения $M\xi$ сокращаются. Поэтому принято характеризовать разброс значений случайной величины математическим ожиданием квадрата её отклонения от своего математического ожидания. Полученное таким образом число называется дисперсией случайной величины ξ.

$D\xi = M(\xi - a)^2 = M[\xi - M\xi]^2$

Ясно, что в случае орудий, ведущих стрельбу, лучшим следует считать орудие, у которого $D\xi$ будет наименьшей.

Пусть ξ характеризуется таблицей вероятностей

$x_i:$	x_1	x_2	$\dots$	x_n
$p_i:$	p_1	p_2	$\dots$	p_n

$$M\xi = \sum_{i=1}^{n} x_i p_i; \qquad D\xi = M(\xi - M\xi)^2 = \sum_{i=1}^{n}(x_i - M\xi)^2 * p_i$$

Определение математического ожидания Пусть есть некоторое пространство, в котором имеется некоторое $\xi = \xi(\omega_i)$.

$\omega_i, (i = 1, \bar{n})$. ω_i – неразделимое событие (пример: исходы броска монеты)

Совокупность ω_i образует пространство элементарных событий $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$

Математическим ожиданием случайной величины ξ называется число, обозначаемое $M\xi$ и равное

$$M\xi = \sum_{\omega_i \in \Omega} \{(\omega_i) * P(\omega_i)\} = \sum_{i=1}^{n} \xi(\omega_i) * p(\omega_i), \text{ где } p_i \text{ - элементарные вероятности. } \sim\sim\sim$$

Из определения математического ожидания вытекают следующие свойства:

1 Аддитивность. $M(\xi + \eta) = M\xi + M\eta$.

Следствие $M\left(\sum_{k=1}^{n} \xi_k\right) = \sum_{k=1}^{n}(M\xi_k)$.

2 $\forall C = const : M(C * \xi) = C * M\xi$. Совокупность свойств 1 и 2 даёт нам свойство линейности математического ожидания:

$$M(C_1\xi_1 + C_2\xi_2 + \dots + C_n\xi_n) = C_1M(\xi_1) + C_2M(\xi_2) + \dots + C_nM(\xi_n)$$

3 Математическое ожидание индикатора случайного события равно вероятности этого случайного события.

Индикатор $[\chi]$: $M\chi_A(\omega) = P(A)$.

$$\chi_A(\omega) = \{1, \omega \in A \,|\, 0, \omega \notin A\}$$

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

$$M\chi_A(\omega) = \sum_{\omega\in A} 1 * p(\omega) \sim\sim\sim$$

4 Свойство монотонности $\xi \geqslant \eta \Rightarrow M\xi \geqslant M\eta$.

Докажем свойство $\xi \geqslant 0 \Rightarrow M\xi \geqslant 0$ (при разложении по определению неотрицательны).

$\xi - \eta \geqslant 0 \Rightarrow M(\xi - \eta) \geqslant 0 :: M\xi - M\eta \geqslant 0 \Rightarrow M\xi \geqslant M\eta$.

Формулы вычисления математического ожидания Пусть $x_1, x_2, \ldots, x_n$ — значения случайной величины ξ, принимаемые с вероятностями $p_1, \ldots, p_i$. Тогда имеет место следующая формула для вычисления математического ожидания :

$$M\xi = \sum_{i=1}^{n} x_i * P(\xi = x_i)$$

Чтобы доказать формулу будем исходить из того, что ξ может быть представлена в виде линейной комбинации индикаторов сл собй

$$\xi = \sum_{i=1}^{n} x_i * \chi_{A_i}(\omega)) :: A_i\omega_i : \xi = x_i$$

Левые и правые части соотношения совпадают. Применим к написанному равенству операцию математического ожидания:

$$M\left(\sum_{i=1}^{n} x_i\chi_{A_i}(\omega)\right) = \sum_{i=1}^{n} M\left(x_i\chi_{A_i}(\omega)\right) = \sum_{i=1}^{n} x_i M\left(\chi_{A_i}(\omega)\right) = \sum_{i=1}^{n} x_i P(\xi = x_i)$$

Рассуждая аналогично, нетрудно получить формулы вычисления математического ожидания от величин, представляющих собой функции случайных величин.

Пусть заданы $f(\xi), g(\xi, \eta)$.

В этом случае

$$M(f(\xi)) = \sum_{i=1}^{n} (f(x_i) * P(\xi = x_i))$$

$$M(g(\xi,\eta)) = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} g(x_i, y_j) * P\left(\xi = x_i,\, \eta = y_i\right)\right)$$

,где $P(\xi, \eta)$ – совместная вероятность.

5 Мультипликативное свойство математического ожидания

Пусть ξ, η – независимые случайные величины, то $M(\xi, \eta) = M\xi * M\eta$

Доказательство:

$$M(\xi,\eta) = \sum_{i=1}^{n}\left(\sum_{j=1}^{m} x_i * y_j * P\left(\xi = x_i, \eta = y_j\right)\right)$$

Если ξ, η независимы, то для них применима теорема умножения вероятности.

$$P(\xi = x_i, \eta = y_j) = (\xi, \eta \text{ независимы}) = P(\xi = x_i) * P(\eta = y_j)$$

$$\sum_{i=1}^{n} \left(\sum_{j=1}^{m} x_i * y_j * P(\xi = x_i) * P(\eta = y_j) \right) =$$

$$= \sum_{i=1}^{n} x_i * P(\xi = x_i) \; * \; \sum_{j=1}^{m} y_i * P(\eta = y_j) = M\xi * M\eta$$

Замечание: Все написанные формулы имеют место, если вероятностное пространство конечно, т.е. число элементарных событий конечно $\omega_i = (1, \bar{n})$.

В случае, если вероятностное пространство счётно, количество элементарных сообщений бесконечно, тогда для случайной величины

$$\xi(\omega), \omega \in (\text{счетное вероятностное пространство})$$

имеют место следующие формулы:

$$\omega_i, i = [\overline{1, \infty}]$$

$$M\xi = \sum_{i=1}^{\infty} (x_i * P(\xi = x_i))$$

$$Mf(\xi) = \sum_{i=1}^{\infty} (f(xi) * P(\xi = x_i))$$

В формулах справа стоят ряды. Чтобы математические ожидания существовали надо, чтобы эти ряды сходились. Ряд сходится, если он имеет конечную сумму.

Задача: Вычислить $M\xi$, распределённой по закону Пуассона. $P(\xi = k) = (a^k/k!)e^{-a}$, где $k = \{0, 1, 2, 3, 4, \ldots, \infty\}$; $\quad a > 0$ – заданный заранее характер распределения.

Решение:

$$M\xi = \sum_{k=0}^{\infty} k * \frac{(a^k * e^{-a})}{k!} = e^{-a} \sum_{k=0}^{\infty} (k * \frac{(ka^k)}{k!} =$$

$$= e^{-a} \sum_{k=0}^{\infty} \frac{(k * a^{k-1} a)}{(k-1)!} = e^{-a} a \sum_{s=0}^{\infty} \frac{a^s}{s!} = e^{-a} a e^{a} = a$$

(формула Маклорена).

Математическое ожидание случайной величины, распределённой по закону Пуассона с параметром распределения a равно этому параметру распределения.