ПРАКТИЧЕСКОЕ ЗАНЯТИЕ 2

Подходы к измерению информации

Основное содержание:
Вероятностный подход

Ход занятия

Теоретический материал

Вероятностный подход к измерению информации предполагает, что количество информации — это мера уменьшения неопределённости знаний при получении информационных сообщений. Этот подход учитывает вероятность появления различных событий или символов.

Первая формула вероятностного подхода была предложена в 1928 году Ральфом Хартли, вторая — в 1948 году Клодом Шенноном.

Принцип

Сообщение содержит информацию, если оно приводит к уменьшению неопределённости знаний. Например: infourok.ru

- **Бросание монеты**. Возможен один результат из двух: монета окажется в положении «орёл» или «решка». Каждое из этих событий равновероятно, ни одно из них не имеет преимущества перед другим. Сообщение о положении монеты уменьшает неопределённость знания в два раза, так как из двух равновероятных событий произошло одно.
- **Кидание шестигранного кубика**. Перед броском не известно, какой стороной кубик упадёт на поверхность. В этом случае возможно получить один результат из шести равновероятных. Неопределённость знаний о результате бросания кубика равна шести, так как именно шесть равновероятных событий может произойти.

Формула

Связь между количеством возможных событий (N) и количеством информации (I) определяется **формулой Хартли**: **N = 2I**.
Если возможные события имеют различные вероятности реализации, количество информации определяют по **формуле Шеннона**: **I = −∑i=1Npilog2pi**, где:

- I — количество информации;
- N — количество возможных событий;
- pi — вероятность i-го события.

Важно: чем менее вероятно событие, тем более «ценно» оно с точки зрения информации. Редкое событие несёт больше информации, и наоборот, частое, вероятное событие — меньше.

Примеры

- **Шахматная доска состоит из 64 полей**: 8 столбцов на 8 строк. Какое количество бит несёт сообщение о выборе одного шахматного поля? **Решение**: поскольку выбор любой из 64 клеток равновероятен, то количество бит находится из формулы: **2i = 64**, **i = log2 64 = 6**, так как 26 = 64. Следовательно, **i = 6 бит**.

Ограничения

Вероятностный подход субъективен — разные люди, получившие одно и то же сообщение, по-разному оценивают количество информации, содержащееся в нём. Это происходит из-за того, что знания людей о событиях, о которых идёт речь в сообщении, различны.

Вероятностный подход к измерению информации

Существует три подхода к измерению информации: алфавитный, содержательный и вероятностный. Для того, чтобы изучить последний, необходимо ознакомиться с такой наукой, как теория вероятностей.

Теория вероятности – это раздел математики, изучающий случайные события.

Основным понятием в этой науке является вероятность события. Когда мы говорим о вероятности события, в обычной жизни мы подразумеваем некоторую меру возможности его возникновения. В теории вероятности есть четкое определение вероятности, которого нужно придерживаться. Допустим, у вас есть несколько исходов какого-то события, например, вы подбрасываете монету. Может выпасть орел, а может – решка, при этом монета идеальна и симметрична, и вероятность выпадения орла и решки одинакова (в теории вероятности говорят, что эти события «равновероятны»). В описанном нами случае вероятность каждого исхода равна 1/2. Интуитивно мы это понимаем, а с точки зрения теории вероятности это число получилось из определения вероятности события. По определению

Вероятность события равна количеству «нужных» исходов, деленному на количество всех возможных:

Вероятность=Количество подходящих вариантов/Количество всех возможных вариант ов

Вероятность изменяется в пределах от 0 до 1.

Практический материал

Пример. Кидают игральный кубик. Какова вероятность, что выпадет четное число?
Решение:

Всего исходов 6 (6 сторон кубика), а нужных нам – 3 (2, 4, 6).

Вероятность = 3/6 = 1/2.

Пример. В вагоне поезда 20 мест. Вам выдают билет случайным образом. На кассе сказали, что вероятность того, что вы получите место у окна, равна 2/5. Сколько мест у окна в вагоне?
Решение:

Количество мест у окна – x. Вероятность равна x/20 = 2/5. Отсюда получаем x = 8.

Вероятность нескольких событий

Пускай есть не одна монета, а три. И мы подбрасываем их одновременно. Мы уже знаем, что для каждой монеты вероятность выпадения орла или решки равна 1/2. Найдем, чему равна вероятность того, что на всех трех монетах сразу выпадет орел.

Можем просто посчитать по известной нам формуле. При подбрасывании трех монет возможно всего 8 исходов: ооо, оор, оро, орр, роо, рор, рро, ррр.

Выпадение на всех трех монетах орла — это 1 исход. Получается, вероятность равна 1/8.

А сейчас посчитаем, чему равна вероятность выпадения орла хотя бы на одной монете?

Выпадение хотя бы на одной орла — это 7 исходов (все ситуации, кроме ррр). Как видите, такой способ считать вероятность составных событий не очень удачный. В теории вероятности есть правила, которые позволяют избежать таких сложностей.

События происходят одновременно

Если необходимо найти вероятность И первого, И второго, И третьего … событий (т.е. вероятность одновременного выполнения исходов), то вероятности перемножаются.

Вернемся к примеру с подбрасыванием трех монет.

Вероятность выпадения орла на каждой монете отдельно равна 1/2, значит, вероятность одновременного выполнения этих трех событий равна 1/2 × 1/2 × 1/2 = 1/8.

Должно произойти хотя бы одно из нескольких событий

Если необходимо найти вероятность ИЛИ первого, ИЛИ второго, ИЛИ третьего … событий (т.е. вероятность выполнения хотя бы исходов), то вероятности складываются.

Это правило хорошо иллюстрируется тем, как мы считали вероятность того, что орел выпадет хотя бы на одной монете. Вероятность каждого из исходов (ооо, оор, оро, орр, роо, рор, рро, ррр) была равна 1/8, а всего их 7. Складываем 1/8 7 раз и получаем 7/8.

Вероятность того, что произойдёт хотя бы одно из независимых событий, можно вычислить путём вычитания из 1 произведения вероятностей противоположных событий, то есть для примера нахождения вероятности выпадения орла хотя бы на одной монете достаточно решить выражение: 1 – 1/8 = 7/8. От единицы мы отняли вероятность неподходящего исхода (ррр), и получили тот же ответ.

Пример. Бросают два игральных кубика. Какова вероятность, что на обоих кубиках выпадет четное число?
Решение:

Вероятность того, что на каждом кубике в отдельности выпадет четное число, равна 1/2 (см. пример выше). По условию надо, чтобы события произошли одновременно, поэтому перемножаем вероятности и получаем 1/4.

Пример. Какова вероятность, что хотя бы на одном из двух кубиков выпадет четное число?
Решение:
Теперь нам подходят варианты, когда четное число: на первом ИЛИ на втором ИЛИ на обоих кубиках сразу. Вероятность каждого из этих событий в отдельности равна 1/4. Чтобы найти итоговую вероятность в этом случае, необходимо сложить вероятности трех перечисленных случаев: 1/4 + 1/4 + 1/4 = 3/4.

Сущность вероятностного подхода к измерению информации

Теория вероятности и теория информации тесно связаны. Когда мы узнаем о выполнении какого-то события, мы получаем определенную информацию, которую можно измерить.

Чем менее вероятно событие, тем более «ценно» оно с точки зрения информации.

Редкое событие несет больше информации, и наоборот, частое, вероятное событие – меньше. Это выражает следующая формула, которая связывает вероятность события с количеством информации, которое оно несет:

2i=1p,$2i=p1$,

где p — вероятность события, а *i* — количество информации (в бит).
Пример. В корзине есть 4 шара, из них 1 синий, остальные — красные. Наугад достают два шара. Какое количество информации несет в себе сообщение о том, что один из шаров — красный, а второй — синий?
Решение:
Сначала посчитаем вероятность. Вероятность достать красный шар: 3/4. Вероятность достать после этого синий шар: 1/3 (потому что шаров стало на 1 меньше).

Надо, чтобы эти события произошли вместе поэтому перемножаем вероятности и получаем: 3/4 × 1/3 = 1/4.

Теперь посчитаем количество информации:

2^i = 1/(1/4).
Значит, 2^i = 4, откуда i = 2 бита.

Домашнее задание:
Решите задачу:
№1. В корзине лежат 8 мячей. Все разного цвета, в т.ч. красный. Какое количество информации несёт в себе сообщение, что из корзины будет вынут красный мяч? Решение. Т.к. вероятность вынуть мяч любого цвета одинакова, то для решения задачи воспользуемся формулой ... $I=\log_2 N$ ... $I=\log_2 8=3$ (бита)
Ответ: 3 бита

№2. В корзине лежат 16 мячей разного цвета: 4 красных, 8 синих и 4 жёлтых. Какое кол-во информации несёт в себе сообщение, что из корзины извлечён 1 мяч?
Решение: Т.к. количество мячей разного цвета в корзине не одинаково, подсчитаем вероятности их извлечения:
красного шара $p_к$=4/16=0,25
синего шара $p_с$=8/16=0,5
жёлтого шара $p_ж$=4/16=0,25
Т.к. события не являются равновероятными, воспользуемся формулой ... $I=-(p_к*\log_2 p_к + p_с*\log_2 p_с + p_ж*\log_2 p_ж)$= $=-(0,25 *\log_2 0,25+0,5*\log_2 0,5+0,25 *\log2\ 0,25)$= $=-(0,25*(-2)+0,5*(-1)+ 0,25*(-2))$=1,5 бита
Ответ:1,5 бита