Как “бороться” с вероятностями
(слегка экстравагантный взгляд)

Вероятности по определению – это предельные отношения количеств появлений тех или иных событий к общим количествам наблюдений.

Условные вероятности – это предельные отношения количеств появлений тех или иных событий в условиях ограничений.

Т.е. в первом случае рассматриваются отношения числа элементов некоторых подмножеств универсального множества (универсальное множество – это объединение всех подмножеств) к общему числу элементов самого универсального множества. Во втором случае – это отношения числа элементов некоторых подмножеств универсального множества к числу элементов другого подмножества того же самого универсального множества. Вот и все. Это киты (или слоны), на которых держится элементарная теория вероятностей.

Отсюда разворачиваются, подобно рекурсии, следущие понятия:

1. Отношения между подмножествами универсального множества I, включающего пустой элемент $\emptyset$, “регулируются” математическим объектом, который называется **алгеброй множеств**.
2. Алгебра множеств изоморфна алгебре логики высказываний и алгебре событий и исходит из более общего математического объекта под названием **Булева решетка** (или **структура**, см. Биркгоф “Теория решеток”). Эту алгебру еще называют **“Булева алгебра”**.
3. Отношение мощности (числа элементов) некоторого подмножества Aуниверсума Iк мощности этого универсума, т.е. $p = m\,(A)/m\,(I)$– есть некоторое число, нормированное в интервале [0, 1]. Причем, множества могут быть конечными и бесконечными: счетными и несчетными. Подобные отношения тождественны понятию вероятностей на множествах бесконечного числа испытаний. Таким образом, для пустого множества событий мы имеем вероятность $p = m\,(\emptyset)/m\,(I) = 0$, а для универсума (множества всех возможных событий) $p = m\,(I)/m\,(I) = 1$. Для некоторого события Aс вероятностью наступления $p = p(A)$вероятность его **не** наступления, как с очевидностью вытекает из прямого посчета элементов соответствующего множества $\bar{A}$, равна $p(\bar{A}) = 1 - p$.

Из п.3 и свойств Булевой алгебры следуют все правила элементарной теории вероятностей как просто дублирование алгебраических и мощностных свойств алгебры множеств. Их лучше всего выписать в некую таблицу или список и шпаргалить каждый раз, как только захочется что-либо посчитать в плане вероятностей.

Свод правил

Алгебраические правила (этими правилами обладают все дистрибутивные решетки, у других алгебр могут быть другие правила, а с ними и нормы):

1. Объединение событий (сложение, дизъюнкция): $\emptyset \vee \emptyset = \emptyset$; $I \vee \emptyset = I$; $\emptyset \vee I = I$; $I \vee I = I$; $\emptyset \vee A = A$; $A \vee \emptyset = A$; $A \vee A = A$; $A \vee \bar{A} = I$; $I \vee A = I$; $A \vee I = I$, гдеAнекоторое событие.

2. Пересечение событий (совместимость, умножение, конъюнкция): $\emptyset \cdot \emptyset = \emptyset$; $I \cdot \emptyset = \emptyset$; $\emptyset \cdot I = \emptyset$; $I \cdot I = I$; $\emptyset \cdot A = \emptyset$; $A \cdot \emptyset = \emptyset$; $A \cdot A = A$;$A \cdot \bar{A} = \emptyset$; $I \cdot A = A$; $A \cdot I = A$, гдеAнекоторое событие.
3. Свойства коммутативности: $A \vee B = B \vee A$; $A \cdot B = B \cdot A$для любых событий Aи B.
4. Свойства ассоциативности: $(A \vee B) \vee C = A \vee (B \vee C)$; $(A \cdot B) \cdot C = A \cdot (B \cdot C)$для любых событий A, Bи C.
5. Свойства дистрибутивности: $(A \vee B) \cdot C = B \cdot C \vee A \cdot C$; $(A \cdot B) \vee C = (B \vee C) \cdot (A \vee C)$для любых событий A, Bи C.

Правила рассчета мощностей подмножеств (вероятностей событий):

1. Мощность универсума $m(I)$это некоторое действительное число, которым оценивается полное множество всех элементов. Может быть взято довольно произвольно, но может и с чем то ассоциироваться, например, это может быть число элементов универсума, если наш универсум – конечное множество. Например, если некоторая комбинационная логическая схема имеет 5 двоичных входов, то все комбинации их состояний будут составлять универсум с числом элементов равным 32 (эту кухню применяют в теории надежности цифровых схем, например, или применить при создании вероятностных искуственных нейронных сетей). Если же мы скажем, что каждый вход нашей схемы устанавливается в процессе работы в состояние 1 как бы размыто, с такой то вероятностью и эти вероятности могут меняться, то мы уже не сможем сказать, что имеем только 32 элемента в нашем универсуме.
2. Мощность $m(A)$подмножества Aэто некоторое действительное число, которым оценивается часть универсума, охватываемая множеством A.
3. Отношение $p = m\,(A)/m\,(I)$- это некоторая относительная количественная оценка элементов множества пропорциональная числу элементов Aпо отношению к элементам универсума. Эта оценка очень удобна тем, что является некоей универсальной безразмерной величиной в интервале [0, 1], что позволяет без труда применять ее ко всем множествам не зависимо от числа их элементов. Например, к элементам какого то подмножества относительно числа элементов (или их оценкам) относительно другого подмножества того же универсума. Если универсум событий является конечным или счетным множеством событий, то говорят, что p- это частота события Aна множестве I. Если же Iне счетное множество (континуум), то p- это вероятность Aна множестве I. Если же два подмножества AиB- частично пересекающиеся подмножества универсального множества I, то число $m(AB)$, т.е. число общих элементов (или совместная оценка) , отнесенное к числу $m(B)$называют условной вероятностью (условной частотой) и обозначают как $p(A/B)$. Т.е. $p(A/B) = m\,(AB)/m\,(B) = (m\,(AB)/m\,(I))/(m\,(B)/m\,(I)) = p\,(AB)/p\,(B)$- формула Байеса по сути.
4. Если мы возьмем таблицу истинности произвольной булевой функции $F(x_1, x_2, \dots)$(или, что тоже самое, таблицу вхождения **независимых** подмножеств $x_1, x_2, \dots$в некоторое составное множество F) и подставим в строках, где $F = I$(в области вхождения в случае подмножеств) вместо $\emptyset$и Iдля каждой переменной из $x_1, x_2, \dots$соответствующие им вероятности $p(x_i)$при $x_i = I$и $1 - p(x_i)$при $x_i = \emptyset$, то получим вероятность события $F = I$т.е. $p(F)$.
5. Существует важное правило для подсчета вероятности объединения двух событий $A \vee B$. Вероятность $p(A \vee B) = p(A) + p(B)$тогда и только тогда, если $A \cdot B = \emptyset$.Поэтому, чтобы было легко расчитать вероятность произвольного булевого выражения $F(x_1, x_2, \dots)$, его необходимо разложить на ортогональные члены из произведения **независимых** переменных (**независимых** подмножеств). Одной из таких форм является разложение Шеннона – это некий аналог разложения в ряд Тейлора, с

которым мы знакомы из области “непрерывной” математики (из матана). Разложение Шеннона имеет следующий вид: $F(\ldots,x_i,\ldots) = \overline{x_i}\cdot F(\ldots,\emptyset,\ldots) \vee x_i \cdot F(\ldots,I,\ldots)$.

Тогда искомая вероятность $p(F(\ldots,x_i,\ldots)) = p(\overline{x_i})\cdot p(F(\ldots,\emptyset,\ldots)) + p(x_i)\cdot p(F(\ldots,I,\ldots))$.Продолжая раскладывать это выражение по всем оставшимся переменным и, подставляя везде известные значения вероятностей, получим, в итоге, то, что искали.

Примеры (почти из жизни)

Пример 1. Имеется три системы навороченного оборудования: A, Bи C, в которых каким то образом используют одно и тоже поле датчиков. Не уточняя, каким бразом эти системы используют датчики (независимые источники данных) мы понимаем, что сами системы как то зависимы между собой. Их результаты, каким то сложным образом коррелируют друг с другом. Пусть две лаборатории проводят экспериментальные измерения, используя все три системы, по своему секретному алгоритму. Каждая система оборудования не тривиальна и срабатывает с определенной вероятностью отличной от нуля и единицы. И, вдруг, через какое то время, достаточно долгое, при сравнении результатов, обнаружилось, что всегда, при условии, когда регистрируется сложное событие $\bar{C} \vee AB$первой лабораторией, то параллельно обязательно регистрируется сложное событие $\bar{A} \vee B \vee C$другой лабораторией. Т.е. мы имеем условную вероятность

$$p((\bar{A} \vee B \vee C)/(\bar{C} \vee AB))(*),$$

которая равна 1. Это значит, что по формуле условной вероятности получается:

$$p((\bar{A} \vee B \vee C)/(\bar{C} \vee AB)) = \frac{p((\bar{A} \vee B \vee C)\cdot(\bar{C} \vee AB))}{p(\bar{C} \vee AB)} = 1.$$

И означает, что $p((\bar{A} \vee B \vee C)\cdot(\bar{C} \vee AB)) = p(\bar{C} \vee AB)$.Что, в свою очередь, возможно лишь при логическом тождестве:

$$(\bar{A} \vee B \vee C)\cdot(\bar{C} \vee AB) \equiv \bar{C} \vee AB.$$

Или

$$((\bar{A} \vee B \vee C)\cdot(\bar{C} \vee AB))(+mod2)(\bar{C} \vee AB) \equiv \emptyset.$$

Упрощая путем эквивалентных преобразований левую часть последнего уравнения получим простое условие, по которому системы A, Bи Cсвязаны между собой в рамках двух экспериментов. Это условие выражается логическим уравнением

$$A\bar{B}\bar{C} \equiv \emptyset.(**)$$

Можно, конечно, пойти еще дальше и попытаться выразить систему оборудования Aчерез остальные Bи C, т.е. систему Aможно сделать из Bи C. Действительно, логическое уравнение (**) решается относительно системы оборудования A. Теория этого факта выходит за рамки статьи, но это так. Общее решение имеет строго следующий вид:

$$A = (B \vee C)\cdot ƛ,$$

где $ƛ$,- любая, какая угодно, система из тех же датчиков, которые применяли в лаборатории. Это все возможные варианты систем A. Среди всех этих систем имеются, как тривиальные, для которых $A \equiv \emptyset$, так и не тривиальные. Как отделить котлеты от котят это уже другой вопрос.

Пример 2. Просто посчитаем условную вероятность (*), если известно, что системы A, Bи Cреализованы из датчиков a, b, c, d, eвероятности срабатывания, которых соответственно равны 0.1, 0.2, 0.3, 0.4, и 0.5 . Системы организованы из датчиков, каждая согласно своей логике:

$$A = a\bar{b} \vee cd;\ B = \bar{a}\bar{e} \vee d;\ C = ad.\quad (***)$$

Подставим выражения (***) в формулу (*) и после упрощений получим:

$$p((\bar{A} \vee B \vee C)/(\bar{C} \vee AB)) = \frac{p(\bar{a} \vee cd)}{p(\bar{a} \vee \bar{b} \vee c \vee \bar{d})}.$$

Разложим логические выражения в числителе и знаменателе в ряд Шеннона, распишем вероятности по независимым событиям и все посчитаем:

$$p((\bar{A} \vee B \vee C)/(\bar{C} \vee AB)) = \frac{p(\bar{a} \vee cd)}{p(\bar{a} \vee \bar{b} \vee c \vee \bar{d})} = \frac{p(\bar{a} \vee acd)}{p(\bar{a} \vee a(\bar{b} \vee b(\bar{c}\bar{d} \vee c)))} =$$

$$\frac{p(\bar{a}) + p(a)p(c)p(d)}{p(\bar{a}) + p(a)(p(\bar{b}) + p(b)(p(\bar{c})p(\bar{d}) + p(c)))} = \frac{1 - p(a) + p(a)p(c)p(d)}{1 - p(a) + p(a)(1 - p(b) + p(b)((1 - p(c))(1 - p(d)) + p(c)))}$$

$= \frac{1-0.1+0.1\cdot 0.3\cdot 0.4}{1-0.1+0.1\cdot(1-0.2+0.2\cdot((1-0.3)(1-0.4)+0.3))} = 0.912/0.9944 =$ **0.917135961** .

Вообще то по полученной вероятности, вероятностям срабатывания датчиков, указанным выше и выражениям внутри скобок в $p((\bar{A} \vee B \vee C)/(\bar{C} \vee AB))$ можно рашать и обратные задачи, например, задачу поиска выражений (***), решая логические и “множественно-мощностные” уравнения. Иногда решения получаются практически единственными, а иногда получаются множества решений. Все зависит от постановки задач.

Кандидат технических наук Олег Васильевич Маевский