

**ПОЛНАЯ  
ВЕРОЯТНОСТЬ.  
ВЕРОЯТНОСТНЫЕ  
АЛГОРИТМЫ В  
ДИАГНОСТИКЕ.  
ФОРМУЛА БАЙЕСА.**

# ТЕОРЕМА ПОЛНОЙ ВЕРОЯТНОСТИ

Пусть событие  $A$  может произойти в результате появления одного и только одного события  $H_i$  ( $H_1, H_2, \dots, H_n$ ).

События этой группы  $H_i$  ( $H_1, H_2, \dots, H_n$ ) обычно называются *гипотезами*.

$P(H_i)$  – вероятность гипотезы;

$P(A/H_i)$  – условная вероятность события (т.е. вероятность события при данной гипотезе.)

**Теорема:** *Вероятность события  $A$ , которое может наступить лишь при появлении одного из несовместных событий (гипотез)  $H_1, H_2, H_3, \dots, H_n$ , образующих полную группу, равна сумме произведений вероятностей каждой из гипотез на соответствующую условную вероятность события  $A$ .*

$$P(A) = \sum_{i=1}^n P(H_i) \cdot P(A/H_i)$$

Или

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + \dots + P(H_n) \cdot P(A/H_n)$$

## ЗАДАЧИ

1. В отделении осуществляется лечение больных с тремя видами заболеваний: ревматизмом, бронхиальной астмой и сердечной недостаточностью. Из 120 больных отделения страдают ревматизмом 52 пациента, бронхиальной астмой 18 человек, остальные страдают сердечной недостаточностью. Вероятность излечения в первой группе больных равна 0,4; в другой – 0,3; в третьей – 0,1. Определить вероятность того, что наугад выбранный больной излечится полностью.

Дано:

$n=120$  – общее количество больных в отделении;

$m_1=52$  – ревматизм;

$m_2=18$  – астма;

$P(A/H_1)=0,4$  -вероятность излечения при

ревматизме;

$P(A/H_2)=0,3$  вероятность излечения при

бронхиальной астме;

$P(A/H_3)=0,1$  - вероятность излечения при сердечной недостаточности;

---

$P(A)$ -?

Решение:

$P H_1 = \frac{52}{120} = 0,43$  – вероятность (гипотеза) того, что пациент болен ревматизмом;

$P H_2 = \frac{18}{120} = 0,15$  - вероятность (гипотеза) того, что пациент болен бронхиальной астмой;

$P H_3 = \frac{120 - (52+18)}{120} = \frac{50}{120} = 0,42$                       вероятность (гипотеза) того, что пациент болен сердечной недостаточностью;

Искомую вероятность найдем по формуле полной вероятности:

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + P(H_3) \cdot P(A/H_3);$$

$$P(A) = 0,43 \cdot 0,4 + 0,15 \cdot 0,3 +$$

$$+ 0,42 \cdot 0,1 = 0,173 + 0,045 + 0,042 = 0,26$$

Ответ: вероятность того, что наугад выбранный больной излечится полностью равна 0,26

1. Вероятность того, что при работе ЭВМ произойдет сбой в АЛУ – 0,2, в ОЗУ – 0,8. Вероятность обнаружения сбоя в АЛУ составляет 0,9, в ОЗУ – 0,95. Определить вероятность того, что возникший в машине сбой будет обнаружен.

**Дано:**

$P(A/H_1)=0,9$  - вероятность обнаружения сбоя в АЛУ;

$P(A/H_2)=0,95$  - вероятность обнаружения сбоя в ОЗУ;

$P(H_1)=0,2$  - вероятность сбоя в АЛУ (гипотеза) ;

$P(H_2)=0,82$  - вероятность сбоя в ОЗУ (гипотеза) ;

---

$P(A)$ -?

**Ответ:**  $P(A)=0,94$

**Решение:**

Искомую вероятность найдем по формуле полной вероятности:

$$P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2);$$

$$P(A) = 0,2 \cdot 0,9 + 0,8 \cdot 0,95 = 0,18 + 0,76 = 0,94;$$

## ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. На участке у врача 120 человек, у которых с вероятностью 0,3 встречается инфекционно-аллергический полиартрит. На другом участке – 160 человек, и это заболевание встречается с вероятностью 0,2. Определить вероятность того, что наугад выбранный пациент болен инфекционно-аллергическим полиартритом (0,243).

2. В поликлинике принимают два врача-стоматолога. Вероятность попасть на прием к первому врачу – 0,7, ко второму – 0,3. Вероятность повторного обращения к стоматологу для первого врача равна 0,02, для второго – 0,15. Определить вероятность того, что наугад выбранный пациент повторно обратится к врачу (0,059).

# ДИАГНОСТИЧЕСКИЕ ТАБЛИЦЫ

Таблица 1

Признаки  (Симптомы) $S_i$	Вероятности наличия симптомов при болезнях			
	Инфаркт миокарда  $D_1$	Перитонит  $D_2$	Крупозная пневмония  $D_3$	Тромбоземболия легочной артерии $D_4$
Число больных	100	100	600	30
$S_1$ - Боли в грудной клетке	0.9	0.05	0.9	0.5
$S_2$ - Боли в животе	0.3	0.95	0.4	0.03
$S_3$ - Повышение температуры	0.95	0.8	0.9	0.03
$S_4$ - Лейкоцитоз	0.95	0.83	0.92	0.03
$S_5$ - Нарушение сердечного ритма	0.92	0.01	0.05	0.1
$S_6$ - Повышение артериального давления	0.5	0.03	0.05	0.03
$S_7$ - Снижение артериального давления	0.1	0.95	0.78	0.87
$S_8$ - Шум трения перикарда	0.86	0.07	0.05	0.03
$S_9$ - Изменение ЭКГ	0.98	0.17	0.15	0.6
$S_{10}$ - Бледность кожи	0.98	0.83	0.78	0.9
$S_{11}$ - Учащение пульса	0.5	0.96	0.99	0.93
$S_{12}$ - Учащение дыхания	0.03	0.07	0.91	0.97
$S_{13}$ - Угнетение рефлексов	0.01	0.02	0.1	0.93
$S_{14}$ - Напряжение брюшной стенки	0.02	0.8	0.1	0.03
$S_{15}$ - Вздутие живота	0.2	0.95	0.13	0.13
$S_{16}$ - Шум трения плевры	0.1	0.01	0.95	0.1



В диагностическую таблицу входит набор симптомов  $S_i$ , относящихся к определенному классу заболеваний (по вертикали), болезни данного класса (по горизонтали), и набор  $P(S_i/D_j)$  для различных заболеваний. Диагноз ставится не по одному, а по нескольким симптомам, обнаруженным у больного (например, у больного обнаружены симптомы  $S_2, S_3, S_4, S_7, S_9, S_{10}, S_{11}, S_{14}, S_{15}$ ) - этот набор симптомов называется симптомокомплексом. Будем обозначать его  $S_{ci}$ .  $S_{ci} = (S_2, S_3, S_4, S_7, S_9, S_{10}, S_{11}, S_{14}, S_{15})$ .

Первое, что делается при рассматриваемом диагностическом методе, это выборка вероятностей всех симптомов для предполагаемых заболеваний. Если заболеваний четыре ( $D_1, D_2, D_3, D_4$ ), то должно появиться четыре группы чисел:

$$P(S_2/D_1) \quad P(S_2/D_2) \quad P(S_2/D_3) \quad P(S_2/D_4)$$

$$P(S_3/D_1) \quad P(S_3/D_2) \quad P(S_3/D_3) \quad P(S_3/D_4)$$

$$P(S_4/D_1) \quad P(S_4/D_2) \quad P(S_4/D_3) \quad P(S_4/D_4)$$

$$P(S_7/D_1) \quad P(S_7/D_2) \quad P(S_7/D_3) \quad P(S_7/D_4)$$

...

$$P(S_{15}/D_1) \quad P(S_{15}/D_2) \quad P(S_{15}/D_3) \quad P(S_{15}/D_4)$$

Если симптомов много и много возможных диагнозов, что и бывает на практике, то один этот этап выборки осуществить без привлечения ЭВМ трудно.

Примечание: Если появление одного события влияет на вероятность появления второго, то такие события называются **зависимыми**.

Курение и заболевание раком легкого - события зависимые. Наличие первого события увеличивает вероятность появления второго.

В случае зависимых событий ( $B$  зависит от  $A$ ) для события  $B$  вводят понятие **условной вероятности**, под которой понимается вероятность события  $B$  при условии, что событие  $A$  произошло.

Обозначение:  $P(B/A)$  или  $P_A(B)$ .

**Правило умножения вероятностей для зависимых событий:** Вероятность совместного появления зависимых событий равна произведению вероятности первого события на условную вероятность второго события

Для двух зависимых событий  $A$  и  $B$  (причем  $B$  зависит от  $A$ ) это правило записывается так:

$$P(A \text{ и } B) = P(A) \times P(B/A) \text{ ИЛИ } P(A \text{ и } B) = P(A) \times P_A(B).$$

Условную вероятность симптомокомплекса  $P(S_{ci}/D_j)$  вычисляют по формуле умножения вероятностей:

$$P(S_{ci}/D_j) = P(S_1/D_j) \cdot P(S_2/D_j) \cdot \dots \cdot P(S_n/D_j) \quad (1)$$

Например, для описанного выше симптомокомплекса  $S_{ci}$  имеем:

$$P(S_{ci}/D_j) = P(S_2/D_j) \cdot P(S_3/D_j) \cdot \dots \cdot P(S_{15}/D_j)$$

Причем условную вероятность симптомокомплекса вычисляют для каждой патологии отдельно. Например, для перитонита формула будет следующей:

$$P(S_{ci}/D_2) = P(S_2/D_2) \cdot P(S_3/D_2) \cdot \dots \cdot P(S_{15}/D_2)$$

задача диагностики заключается в том, чтобы на основании симптомокомплекса, установленного у больного, и данных диагностической таблицы определить вероятности  $P(D_j/S_{ci})$  каждой из имеющихся в таблице болезней  $D_1, D_2, D_3, \dots, D_n$  т.е. по сути дела нужно перейти от  $P(S_{ci}/D_j)$  к  $P(D_j/S_{ci})$ . Болезнь, имеющая наибольшую вероятность, и будет рассматриваться как искомый диагноз. Этот переход осуществляется по известной в теории вероятностей формуле Байеса.

\*

## ФОРМУЛА БАЙЕСА

**Теорема:** Пусть событие  $A$ , может наступить лишь при условии появления одного из несовместных событий (гипотез)  $H_1, H_2, H_3, \dots, H_n$ , которые образуют полную группу событий. Если событие  $A$  уже произошло, то вероятности гипотез могут быть переоценены по формуле Байеса:

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{\sum_{i=1}^n P(H_i) \cdot P(A/H_i)}$$

ИЛИ

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + \dots + P(H_n) \cdot P(A/H_n)}$$

## ЗАДАЧИ

1. В отделении осуществляется лечение двух нозологических единиц: пневмонии и бронхита. Из 40 больных отделения больны пневмонией 15 человек. Вероятность повторного поступления в отделение при пневмонии – 0,03; при бронхите – 0,25. Определить вероятность бронхита у пациента, повторно попавшего в больницу.

Дано:

$m_1=15$ -пневм.

$m_2=25$ -бронхит

$n=40$

$P(A/H_1)=0,03$ ;

$P(A/H_2)=0,25$ ;

$P(H_2/A)$ -?

Решение:

Найдем вероятности гипотез:

$$P(H_1) = \frac{15}{40} = 0,375$$

$$P(H_2) = \frac{25}{40} = 0,625$$

По теореме Байеса:

$$P(H_2/A) = \frac{P(H_2) \cdot P(A/H_2)}{P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2)}$$

$$P(H_1/A) = \frac{0,03 \cdot 0,375 + 0,25 \cdot 0,625}{0,03 \cdot 0,375 + 0,25 \cdot 0,625} = \frac{0,15625}{0,1675} = 0,933$$

Ответ: вероятность бронхита у пациента, повторно попавшего в больницу 0,933.

2. В диагностический центр в равных количествах поступают пациенты из трех консультативных пунктов. Диагноз был подтвержден для пациентов первого диагностического пункта в 90% случаев, для второго составляет в 85% случаев, а для третьего – в 75% случаев. Найдите вероятность того, что наугад выбранный пациент будет с подтвержденным диагнозом.

<p>Дано:</p> <p><math>P(A/H_1)=0,9</math>;</p> <p><math>P(A/H_2)=0,85</math>;</p> <p><math>P(A/H_3)=0,75</math>;</p>	<p>Решение:</p> <p>Событие А – выбранный пациент будет с подтвержденным диагнозом. Всего 3 гипотезы: <math>H_1</math> – пациент прибыл из 1-го пункта; <math>H_2</math> – пациент прибыл из 2-го пункта; <math>H_3</math> – пациент прибыл из 3-го пункта. По условию они равновероятны, следовательно сумма вероятностей гипотез равна 1, таким образом вероятность каждой гипотезы равна: <math>P(H_1) = P(H_2) = P(H_3) = \frac{1}{3}</math>.</p> <p><math>P(A/H_1)=0,9</math>, <math>P(A/H_2)=0,85</math>, <math>P(A/H_3)=0,75</math> – это условные вероятности того, что диагноз будет подтвержден для 1-го, 2-го и 3-го пункта, соответственно.</p> <p>Искомую вероятность найдем по формуле полной вероятности:</p> $P(A) = P(H_1) \cdot P(A/H_1) + P(H_2) \cdot P(A/H_2) + P(H_3) \cdot P(A/H_3)$ $P(A) = \frac{1}{3} (0,9 + 0,85 + 0,75) = 0,833$
<p><math>P(A)</math>-?</p>	<p>Ответ: вероятность того, что наугад выбранный пациент будет с подтвержденным диагнозом 0,833</p>

## ЗАДАЧИ ДЛЯ САМОСТОЯТЕЛЬНОГО РЕШЕНИЯ

1. Из 200 больных отделения 61 пациент болен диабетом. Вероятность повторного обращения в больницу в группе больных диабетом составляет 0,8; в другой группе – 0,01. Больной повторно обратился к врачу. Определит вероятность того, что у него диабет.

2. Предприятие выпускает изделия определенного вида на трех поточных линиях. Первая изготавливает 60% всех изделий, вторая – 30%, третья – 10%. Вероятность изготовления бракованного изделия этими линиями равна 0,01; 0,02; 0,1 соответственно. Определить вероятность того, что наугад выбранное бракованное изделие было изготовлено третьей линией.

3. Две аптечные работницы развесили по одинаковому комплекту порошков. Вероятность того, что первая работница допустит неточность, равна 0,05; для второй работницы эта вероятность равна 0,01. При контроле правильности расфасовки была выявлена ошибка. Найти вероятность того, что ошиблась первая работница.