1. Формула полной вероятности и формула Байеса

Если событие *A* может произойти только при выполнении одного из событий $B_1, B_2, \ldots, B_n$, которые образуют *полную группу несовместных событий*, то вероятность события *A* вычисляется по формуле

$$P(A) = P(B_1)P(A \mid B_1) + P(B_2)P(A \mid B_2) + \ldots + P(B_n)P(A \mid B_n).$$

Эта формула называется *формулой полной вероятности*.

Вновь рассмотрим полную группу несовместных событий $B_1, B_2, \ldots, B_n$, вероятности появления которых $P(B_1), P(B_2), \ldots, P(B_n)$. Событие *A* может произойти только вместе с каким-либо из событий $B_1, B_2, \ldots, B_n$, которые будем называть *гипотезами*. Тогда по формуле полной вероятности

$$P(A) = P(B_1)P(A \mid B_1) + P(B_2)P(A \mid B_2) + \ldots + P(B_n)P(A \mid B_n)$$

Если событие *A* произошло, то это может изменить вероятности гипотез $P(B_1), P(B_2), \ldots, P(B_n)$.

По теореме умножения вероятностей

$$P(AB_1) = P(B_1)P(A \mid B_1) = P(A)P(B \mid A),$$

откуда

$$P(B_1 \mid A) = \frac{P(B_1)P(A \mid B_1)}{P(A)}.$$

Аналогично, для остальных гипотез

$$P(B_i \mid A) = \frac{P(B_i)P(A \mid B_i)}{P(A)}, \; i = 1, \ldots, n.$$

Полученная формула называется *формулой Байеса* (*формулой Бейеса*). Вероятности гипотез $P(B_i \mid A)$ называются *апостериорными вероятностями*, тогда как $P(B_i)$ - *априорными вероятностями*.

Пример. В магазин поступила новая продукция с трех предприятий. Процентный состав этой продукции следующий: 20% - продукция первого предприятия, 30% - продукция второго предприятия, 50% - продукция третьего предприятия; далее, 10% продукции первого предприятия высшего сорта, на втором предприятии - 5% и на третьем - 20% продукции высшего сорта. Найти вероятность того, что случайно купленная новая продукция окажется высшего сорта.

Решение. Обозначим через *B* событие, заключающееся в том, что будет куплена продукция высшего сорта, через A_1, A_2, A_3 обозначим события, заключающиеся в покупке продукции, принадлежащей соответственно первому, второму и третьему предприятиям.

Можно применить формулу полной вероятности, причем в наших обозначениях:

$$P(A_1) = 0,2 \qquad P(B|A_1) = 0,1$$
$$P(A_2) = 0,3 \qquad P(B|A_2) = 0,05$$
$$P(A_3) = 0,5 \qquad P(B|A_3) = 0,2$$

Подставляя эти значения в формулу полной вероятности, получим искомую вероятность:

$$P(B) = 0,2 \cdot 0,1 + 0,3 \cdot 0,05 + 0,5 \cdot 0,2 = 0,135.$$

Пример. Один из трех стрелков вызывается на линию огня и производит два выстрела. Вероятность попадания в мишень при одном выстреле для первого стрелка равна 0,3, для второго - 0,5; для третьего - 0,8. Мишень не поражена. Найти вероятность того, что выстрелы произведены первым стрелком.

Решение. Возможны три гипотезы:

A_1 - на линию огня вызван первый стрелок,

A_2 - на линию огня вызван второй стрелок,

A_3 - на линию огня вызван третий стрелок.

Так как вызов на линию огня любого стрелка равновозможен, то
$$P(A_1) = P(A_2) = P(A_3) = \frac{1}{3}.$$

В результате опыта наблюдалось событие B - после произведенных выстрелов мишень не поражена. Условные вероятности этого события при сделанных гипотезах равны:

$$P(B \mid A_1) = 0,7 \cdot 0,7 = 0,49;$$
$$P(B \mid A_2) = 0,5 \cdot 0,5 = 0,25;$$
$$P(B \mid A_3) = 0,2 \cdot 0,2 = 0,04.$$

по формуле Байеса находим вероятность гипотезы A_1 после опыта:

$$P(A_1|B) = \frac{0,49 \cdot 1/3}{1/3 \cdot 0,49 + 1/3 \cdot 0,25 + 1/3 \cdot 0,04} = \frac{0,49}{0,78} = 0,628.$$

Пример. На трех станках-автоматах обрабатываются однотипные детали, поступающие после обработки на общий конвейер. Первый станок дает 2% брака, второй – 7%, третий –

10%. Производительность первого станка в 3 раза больше производительности второго, а третьего – в 2 раза меньше, чем второго.

а) Каков процент брака на конвейере?

б) Каковы доли деталей каждого станка среди бракованных деталей на конвейере?

Решение. Возьмем с конвейера наудачу одну деталь и рассмотрим событие А – деталь бракованная. Оно связано с гипотезами относительно того, где была обработана эта деталь: H_k – взятая наудачу деталь обработана на k -ом станке, $k = 1, 2, 3$.

Условные вероятности (в условии задачи они даны в форме процентов):

$$P(A|H_1) = 0,02,\ P(A|H_2) = 0,07,\ P(A|H_3) = 0,1.$$

Зависимости между производительностями станков означают следующее:

$$P(H_1) = 3P(H_2),\ P(H_3) = 0,5P(H_2).$$

А так как гипотезы образуют полную группу, то $P(H_1) + P(H_2) + P(H_3) = 1$.

Решив полученную систему уравнений, найдем: $P(H_1) = 6/9,\ P(H_2) = 2/9,\ P(H_3) = 1/9$.

а) Полная вероятность того, что взятая наудачу с конвейера деталь – бракованная:

$$P(A) = \sum P(A|H_i)P(H_i) = \frac{6}{9}0,02 + \frac{2}{9}0,07 + \frac{1}{9}0,1 = 0,04.$$

Другими словами, в массе деталей, сходящих с конвейера, брак составляет 4%.

б) Пусть известно, что взятая наудачу деталь – бракованная. Пользуясь формулой Байеса, найдем условные вероятности гипотез:

$$P(H_1|A) = \frac{P(A|H_1)P(H_1)}{P(A)} = \frac{6/9 \cdot 0,02}{0,04} = 0,33,$$

$$P(H_2|A) = \frac{P(A|H_2)P(H_2)}{P(A)} = \frac{2/9 \cdot 0,07}{0,04} = 0,39,$$

$$P(H_3|A) = \frac{P(A|H_3)P(H_3)}{P(A)} = \frac{1/9 \cdot 0,1}{0,04} = 0,28.$$

Таким образом, в общей массе бракованных деталей на конвейере доля первого станка составляет 33%, второго – 39%, третьего – 28%.

Формула Бернулли

Воспользуемся понятием **сложного события**, под которым подразумевается совмещение нескольких элементарных событий, состоящих в появлении или непоявлении события A в i–м испытании. Пусть проводится n независимых испытаний, в каждом из которых событие A может либо появиться с вероятностью p, либо не появиться с вероятностью $q = 1 - p$. Рассмотрим событие B_m, состоящее в том, что событие A в этих n испытаниях наступит ровно m раз и, следовательно, не наступит ровно $(n - m)$ раз. Обозначим A_i $(i = 1, 2, \ldots, n)$ появление события A, а $\overline{A}_i$ — непоявление события A в i–м испытании. В силу постоянства условий испытания имеем

$$P\{A_1\} = P\{A_2\} = \cdots = P\{A_n\} = p,$$
$$P\{\overline{A}_1\} = P\{\overline{A}_2\} = \cdots = P\{\overline{A}_n\} = 1 - p = q$$

Событие A может появиться m раз в разных последовательностях или комбинациях, чередуясь с противоположным событием $\overline{A}$. Число возможных комбинаций такого рода равно числу сочетаний из n элементов по m, т. е. C_n^m. Следовательно, событие B_m можно представить в виде суммы сложных несовместных между собой событий, причем число слагаемых равно C_n^m:

$$B_m = A_1 A_2 \cdots A_m \overline{A}_{m+1} \cdots \overline{A}_n + \cdots + \overline{A}_1 \overline{A}_2 \cdots \overline{A}_{n-m} A_{n-m+1} \cdots A_n, \tag{3.1}$$

где в каждое произведение событие A входит m раз, а $\overline{A}$ — $(n - m)$ раз.

Вероятность каждого сложного события, входящего в формулу (3.1), по теореме умножения вероятностей для независимых событий равна $p^m q^{n-m}$. Так как общее количество таких событий равно C_n^m, то, используя теорему сложения вероятностей для несовместных событий, получаем вероятность события B_m (обозначим ее $P_{m,n}$)

$$P_{m,n} = C_n^m p^m q^{n-m} \quad \text{or} \quad P_{m,n} = \frac{n!}{m!(n-m)!} p^m q^{n-m}. \qquad (3.2)$$

Формулу (3.2) называют **формулой Бернулли**, а повторяющиеся испытания, удовлетворяющие условию независимости и постоянства вероятностей появления в каждом из них события A, называют **испытаниями Бернулли**, или **схемой Бернулли**.

Пример 1. Вероятность выхода за границы поля допуска при обработке деталей на токарном станке равна 0,07. Определить вероятность того, что из пяти наудачу отобранных в течение смены деталей у одной размеры диаметра не соответствуют заданному допуску.

Решение. Условие задачи удовлетворяет требования схемы Бернулли. Поэтому, полагая $n = 5,\ m = 1,\ p = 0{,}07$, по формуле (3.2) получаем

$$P_{1,5} = C_5^1 (0{,}07)^1 (0{,}93)^{5-1} \approx 0{,}262.$$

Пример 2. Наблюдениями установлено, что в некоторой местности в сентябре бывает 12 дождливых дней. Какова вероятность того, что из случайно взятых в этом месяце 8 дней 3 дня окажутся дождливыми?

Решение.

$$P_{3;8} = C_8^3 \left(\frac{12}{30}\right)^3 \left(1 - \frac{12}{30}\right)^{8-3} = \frac{8!}{3!(8-3)!} \left(\frac{2}{5}\right)^3 \left(\frac{3}{5}\right)^5 = 56 \cdot \frac{8}{125} \cdot \frac{243}{3125} = \frac{108\,864}{390\,625} \approx 0{,}2787.$$