МДК.01.02 Математический аппарат для построения компьютерных сетей, занятие 24

Полная вероятность. Формула Байеса и применение неравенств Маркова и Чебышева для решения комбинаторных задач

Практическая работа

Формула полной вероятности

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**.

Формула Байеса (Бейеса)

Вновь рассмотрим полную группу несовместных событий

$B_1, B_2, \dots, B_n$, вероятности появления которых

$P(B_1), P(B_2), \dots, P(B_n)$.

Событие ***A*** может произойти только вместе с каким-либо из событий $B_1, B_2, \dots, B_n$, которые будем называть **гипотезами**.

Тогда по формуле полной вероятности

$$P(A) = P(B_1)P(A|B_1) + P(B_2)P(A|B_2) + \dots + P(B_n)P(A|B_n).$$

Как определить вероятность того, что имела место **та или иная гипотеза**?

Формула Байеса (Бейеса)

При условии, что событие уже произошло, вероятности гипотез переоцениваются по формулам, которые получили фамилию английского священника Томаса Байеса:

$P_A(B_1) = \frac{P(B_1) \cdot P_{B_1}(A)}{P(A)}$ – вероятность того, что имела место гипотеза B_1

$P_A(B_2) = \frac{P(B_2) \cdot P_{B_2}(A)}{P(A)}$ – вероятность того, что имела место гипотеза B_2

$P_A(B_3) = \frac{P(B_3) \cdot P_{B_3}(A)}{P(A)}$ – вероятность того, что имела место гипотеза B_3

…

$P_A(B_n) = \frac{P(B_n) \cdot P_{B_n}(A)}{P(A)}$ – вероятность того, что имела место гипотеза B_n

Формула Байеса (Бейеса)

На первый взгляд кажется полной нелепицей – зачем пересчитывать вероятности гипотез, если они и так известны? Но на самом деле разница есть:

$P(B_1), P(B_2), P(B_3), .., P(B_n)$ - это ***априорные*** (оцененные **до** испытания) вероятности.

$P_A(B_1), P_A(B_2), P_A(B_3), .., P_A(B_n)$ - это ***апостериорные***

(оцененные **после** испытания) вероятности тех же гипотез, пересчитанные в связи «со вновь открывшимися обстоятельствами» – с учётом того факта, что событие ***А*** достоверно произошло.

Формула Байеса (Бейеса)

Рассмотрим пример.

На склад поступило 2 партии изделий: первая – 4000 штук, вторая – 6000 штук.
Средний процент нестандартных изделий в первой партии составляет 20%, а во второй – 10%.
Наудачу взятое со склада изделие оказалось стандартным.

Найти вероятность того, что оно:
а) из первой партии,
б) из второй партии.

Формула Байеса (Бейеса)

Решение.

Первая часть решения состоит в использовании формулы полной вероятности. Иными словами, вычисления проводятся в предположении, что испытание ещё не произведено и событие «изделие оказалось стандартным» пока не наступило.

Рассмотрим две гипотезы:

B_1– наудачу взятое изделие будет из 1-й партии;

B_2– наудачу взятое изделие будет из 2-й партии.

Всего: 4000 + 6000 = 10000 изделий на складе.

Решение.

По классическому определению:

$$P(B_1)=\frac{4000}{10000}=0,4; \quad P(B_2)=\frac{6000}{10000}=0,6.$$

Рассмотрим зависимое событие: ***A*** – наудачу взятое со склада изделие будет стандартным.
В первой партии 100% – 20% = 80% стандартных изделий,

поэтому: $P_{B_1}(A)=\frac{80}{100}=0,8$ - вероятность того, что наудачу взятое на складе изделие будет стандартным **при условии**, что оно принадлежит 1-й партии. Аналогично, во второй партии 100% – 10% = 90% стандартных изделий и

$P_{B_2}(A)=\frac{90}{100}=0,9$ - вероятность того, что наудачу взятое на складе изделие будет стандартным при условии, что оно принадлежит 2-й партии.

Решение.

По формуле полной вероятности:

$P(A)=P(B_1)\cdot P_{B_1}(A)+P(B_2)\cdot P_{B_2}(A)=0{,}4\cdot 0{,}8+0{,}6\cdot 0{,}9=0{,}32+0{,}54=0{,}86$ – вероятность того, что наудачу взятое на складе изделие будет стандартным.

Часть вторая. Пусть наудачу взятое со склада изделие оказалось стандартным. Эта фраза прямо прописана в условии, и она констатирует тот факт, что событие произошло. По формулам Байеса вероятность того, что выбранное стандартное изделие принадлежит 1-й партии:

а) $P_A(B_1)=\dfrac{P(B_1)\cdot P_{B_1}(A)}{P(A)}=\dfrac{0{,}4\cdot 0{,}8}{0{,}86}=\dfrac{0{,}32}{0{,}86}=\dfrac{32}{86}=\dfrac{16}{43}\approx 0{,}37$

вероятность того, что выбранное стандартное изделие принадлежит 2-й партии:

б) $P_A(B_2)=\dfrac{P(B_2)\cdot P_{B_2}(A)}{P(A)}=\dfrac{0{,}6\cdot 0{,}9}{0{,}86}=\dfrac{0{,}54}{0{,}86}=\dfrac{54}{86}=\dfrac{27}{43}\approx 0{,}63$

Задание 1.

На склад поступило 3 партии изделий: первая – 2000 штук, вторая – 3000 штук, третья – 5000 штук.
Средний процент нестандартных изделий в первой партии составляет 5%, во второй – 10%, а в третьей – 15%.
Наудачу взятое со склада изделие оказалось стандартным.

Найти вероятность того, что оно:
а) из первой партии,
б) из второй партии,
в) из третьей партии.

Неравенство Маркова

Неравенство Маркова дает вероятностную оценку того, что значение неотрицательной случайной величины превзойдет некоторую константу через известное математическое ожидание.

Когда никаких других данных о распределении нет, неравенство дает некоторую информацию, хотя зачастую оценка груба или тривиальна.

Неравенство Маркова

Пусть X - случайная величина, принимающая неотрицательные значения, M(X) - ее конечное математическое ожидание, то для любых a>0 выполняется условие

$$P(X \geq a) \leq \frac{M(X)}{a}.$$

Альтернативная форма записи (когда нужно оценить вероятность того, что случайная величина меньше некоторой константы):

$$P(X < a) > 1 - \frac{M(X)}{a}.$$

Неравенство Маркова

Пример.

Среднее количество вызовов, поступающих на коммутатор завода в течение часа, равно 300.

Оценить вероятность того, что в течение следующего часа число вызовов на коммутатор:

а) превысит 400;
б) будет не более 500.

Неравенство Маркова

Решение.

По условию M(X)=300.

а) Воспользуемся формулой (неравенством Маркова)

$$P(x > A) \le \frac{M(X)}{A}.$$

Тогда $P(x > 400) \le \frac{300}{400} = 0,75$, т.е. вероятность того,

что число вызовов превысит 400, будет не более 0,75.

Неравенство Маркова

Решение.
б) воспользуемся неравенством Маркова в альтернативном виде:

$$P(x \le A) \ge 1 - \frac{M(X)}{A}.$$

Тогда $P(x \le 500) \ge 1 - \frac{300}{500} = 0,4$, т.е. вероятность того,

что число вызовов не более 500, будет не менее 0,4.

Задание 2.

Количество потребляемой за сутки электроэнергии предприятием является случайной величиной с математическим ожиданием 8 мегаватт.

Оценить вероятность того, что в ближайшие сутки потребление электроэнергии окажется:

а) более 10 мегаватт,
б) не более 12 мегаватт.

Неравенство Чебышева

Когда известны не только математическое ожидание, но и дисперсия для случайной величины, можно применять следствие неравенства Маркова — **неравенство Чебышева**, которое дает оценку вида:

$$P(|X - M(X)| \geq a) \leq \frac{D(X)}{a^2}, \quad a > 0.$$

Также его можно записать в другой форме:

$$P(|X - M(X)| < a) > 1 - \frac{D(X)}{a^2}, \quad a > 0.$$

Неравенство Чебышева показывает, что случайная величина принимает значения близкие к среднему (математическому ожиданию) и дает оценку вероятности больших отклонений.

Неравенство Чебышева

Пример.

Генератор обеспечивает выходное напряжение, которое может отклоняться от номинального назначение, не превышающее 1 В, с вероятностью 0,95. Какие значения дисперсии выходного напряжения можно ожидать?

Решение.

Пусть X - величина выходного напряжения. Применим неравенство Чебышева

$$P\left(|X - MX|\right) < \varepsilon) \geq 1 - \frac{DX}{\varepsilon^2}.$$

Неравенство Чебышева

Подставляя наши данные ε=1, P=0,95P, имеем:

$$0,95 = P\left(\left|X - \mathrm{M}X\right|\right) < 1) \geq 1 - \frac{D(X)}{1^2},$$

$$1 - D(X) \leq 0,95,$$

$$D(X) \geq 0,05.$$

Ответ: дисперсия не менее 0,05 B^2.

Задание 3.

Завод каждый день выпускает определённое количество изделий.
Количество выпускаемых изделий может отклоняться от номинального на 250 штук с вероятностью 0,93.

Определите величину дисперсии количества выпускаемых изделий в день.