Теория вероятностей и математическая статистика

Кракашова Ольга Анатольевна

доцент, канд. экон. наук,

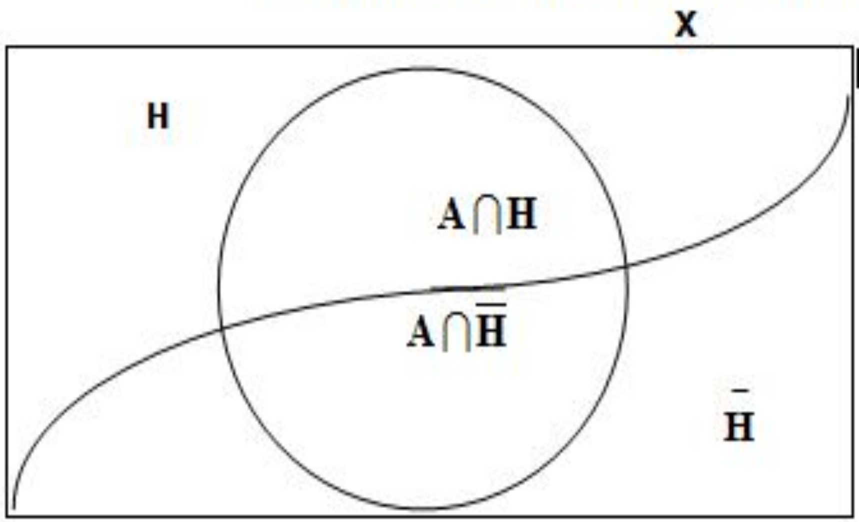
доцент кафедры Статистики, эконометрики и оценки рисков РГЭУ (РИНХ)

Лекция № 4

Формулы полной вероятности и Байеса

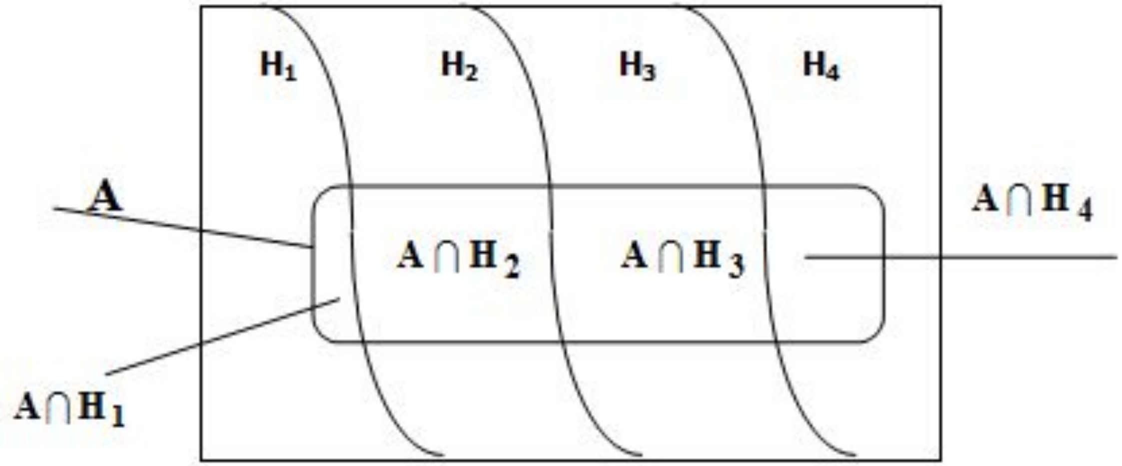
Рассмотрим два события А и Н, изображенные на диаграмме Венна. Наборы **Н** и $\bar{\mathbf{H}}$ - форма разбиения набора **А** на два подмножества взаимно несовместных событий. События **Н** и $\bar{\mathbf{H}}$ взаимно противоположны. Событие **А** может произойти либо с **Н**, либо с $\bar{\mathbf{H}}$, но не с двумя вместе. Каковы бы ни были взаимосвязи между событиями А и Н, всегда можно сказать, что вероятность события А равна вероятности пересечения событий А и Н плюс вероятность пересечения А и дополнения Н (событие $\bar{\mathbf{H}}$).

Разложение **А** *на части зависит от* **Н** *и* $\bar{\mathbf{H}}$.



$$P(A) = P(A \cap H) + P(A \cap \bar{H})$$

Пусть событие A может осуществляться лишь вместе с одним из событий H_1, H_2, H_3, ..., H_n, образующих полную группу (то есть они единственно возможные и несовместные).



Так как заранее неизвестно, какое из событий H_1, H_2, ..., H_n наступит, то их называют **гипотезами**.

Пусть также известны вероятности гипотез $P(H_1)$, $P(H_2)$, ..., $P(H_n)$ и условные вероятности события A: $P(A/H_1)$, $P(A/H_2)$, ..., $P(A/H_n)$.

Так как гипотезы образуют полную группу, то $\sum_{i=1}^{n} P(H_i) = 1$. Рассмотрим событие A - это (или H_1A, ..., или H_nA). События H_1A, H_2A, ..., H_nA - несовместные попарно, так как события H_1, H_2, ..., H_n попарно несовместны. К этим событиям применяем теорему сложения вероятностей для несовместных событий:

$$P(A) = P(H_1A) + P(H_2A) + \dots + P(H_nA), \text{ или}$$

$$\mathbf{P(A) = \sum_{i=1}^{n} P(A \cap H_i)}$$

События $H_1A, H_2A, \ldots, H_nA$ - зависимые. Применяя теорему умножения вероятностей для зависимых событий, получим:

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

или

$$P(A) = \sum_{i=1}^{n} P(H_i)\cdot P(A/H_i)$$ – **формула полной вероятности.**

Если событие A может наступить только вместе с одним из событий $H_1, H_2, H_3, \ldots, H_n$, образующих полную группу несовместных событий и называемых гипотезами, то вероятность события A равна сумме произведений вероятностей каждого из событий $H_1, H_2, \ldots, H_n$ на соответствующую условную вероятность события A.

Случай двух событий:

$$P(A) = P(H)\cdot P(A/H) + P(\bar{H})\cdot P(A/\bar{H})$$

Случай более двух событий:

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

$i=1,2,\ldots,n$

Пример 1. Экономист полагает, что вероятность роста стоимости акций некоторой компании в следующем году равна 0.75, если экономика страны будет на подъёме; и эта же вероятность равна 0.30, если экономика страны не будет успешно развиваться. По его мнению, вероятность экономического подъёма в будущем году равна 0.80. Используя предположения экономиста, оцените вероятность того, что акции компании поднимутся в цене в следующем году?

Решение. Событие A – "акции компании поднимутся в цене в будущем году".

Составим рабочую таблицу:

Таблица 1

№ H_i	Гипотезы H_i	$P(H_i)$	$P(A/H_i)$	$P(H_i) \bullet P(A/H_i)$
1	H1 - "подъем экономики"	0.80	0.75	0.60
2	H2 - "спад экономики"	0.20	0.30	0.06
Σ		**1.00**	–	**P(A) = 0.66**

Пример 2. В каждой из двух урн содержится 6 черных и 4 белых шара. Из первой урны во вторую наудачу переложен один шар. Найти вероятность того, что шар, извлеченный из второй урны после перекладывания, окажется черным?

Решение. Событие А -"шар, извлеченный из второй урны - черный".

Составим рабочую таблицу:

Таблица 2

№ H_i	Гипотезы H_i	$P(H_i)$	$P(A/H_i)$	$P(H_i) \bullet P(A/H_i)$
1	H1 - "из 1 урны во 2 урну переложили черный шар"	6/10	7/11	42/110
2	H2 - "из 1 урны во 2 урну переложили белый шар"	4/10	6/11	24/110
Σ		**1.00**	--	**P(A) = 0.60**

Вычисление вероятностей гипотез (формулы Байеса)

В тех случаях, когда стало известно, что событие А произошло, возникает потребность в определении (уточнении) условной вероятности $P(H_i/A)$. Пусть событие А может осуществляться лишь вместе с одной из гипотез H_i ($i = 1, 2, ..., n$). Известны вероятности гипотез $P(H_1)$, ..., $P(H_n)$ и условные вероятности А: $P(A/H_1)$, $P(A/H_2)$, ..., $P(A/H_n)$. Так как $AH_i = H_iA$, то $P(AH_i) = P(H_iA)$ или $P(A) \cdot P(A/H_i) = P(H_i) \cdot P(H_i/A)$, а отсюда по правилу пропорций:

$$P(H_i/A) = \frac{P(H_i) \cdot P(A/H_i)}{P(A)}$$

Итак:

Случай двух событий:

$$P(H/A) = \frac{P(H) \cdot P(A/H)}{P(H) \cdot P(A/H) + P(\overline{H}) \cdot P(A/\overline{H})}$$

Случай более двух событий:

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

Формулы Бейеса позволяют переоценить вероятности гипотез после того как становится известным результат испытания, в итоге которого появилось событие А.

Вероятность события H, задаваемая при условии появления события A, получается из вероятностей событий H и $\overline{H}$ и из условной вероятности события A при заданном H. Вероятности событий H и $\overline{H}$ называют априорными (предшествующими), вероятность P(H/A) называют апостериорной (последующей). Формулы Бейеса позволяют переоценить наши априорные вероятности гипотез в апостериорные вероятности, вытекающие из знания появления события A.

Пример 3. Экономист полагает, что в течение периода активного экономического роста американский доллар будет расти в цене с вероятностью 0.7, в период умеренного экономического роста доллар подорожает с вероятностью 0.4 и при низких темпах экономического роста доллар подорожает с вероятностью 0.2. В течение любого периода времени вероятность активного экономического роста 0.3, умеренного экономического роста 0.5 и низкого роста - 0.2. Предположим, что доллар дорожает в течение текущего периода, чему равна вероятность того, что анализируемый период совпал с периодом активного экономического роста.

Решение. Определим гипотезы:

H_1 - "активный экономический рост"

H_2 - "умеренный экономический рост"

H_3 - "низкий экономический рост"

Определим событие А - "доллар дорожает"

Имеем: $P(H_1) = 0.3$, $P(H_2) = 0.5$, $P(H_3) = 0.2$,

$P(A/H_1) = 0.7$,

$P(A/H_2) = 0.4$ и $P(A/H_3) = 0.2$

Найти: $P(H_1/A)$

Используя формулу Байеса и подставляя заданные значения вероятностей, имеем:

$$P(H_1/A) = \frac{P(H_1)\cdot P((A/H_1)}{P(H_1)\cdot P(A/H_1)+P(H_2)\cdot P(A/H_2)+P(H_3)\cdot P(A/H_3)} =$$

$$= \frac{0.30\cdot 0.70}{0.30\cdot 0.70+0.50\cdot 0.40+0.20\cdot 0.20} = 0.467$$

Мы можем получить тот же результат при помощи таблицы:

Таблица 3

Гипотезы H_i	**Априорные вероятности $P(H_i)$**	**Условные вероятност и $P(A/H_i)$**	**Совместные вероятности $P(A\cap H_i)$**	**Апостериорные вероятности $P(H_i/A)$**
H_1	0.30	0.70	0.21	0.21 / 0.45 = 0.467
H_2	0.50	0.40	0.20	0.20 / 0.45 = 0.444
H_3	0.20	0.20	0.04	0.04 / 0.45 = 0.089
Сумма	1.00	--	0.45	1

Пример 3. Экономист полагает, что в течение периода активного экономического роста американский доллар будет расти в цене с вероятностью 0.7, в период умеренного экономического роста доллар подорожает с вероятностью 0.4 и при низких темпах экономического роста доллар подорожает с вероятностью 0.2. В течение любого периода времени вероятность активного экономического роста 0.3, умеренного экономического роста 0.5 и низкого роста - 0.2. Предположим, что доллар дорожает в течение текущего периода, чему равна вероятность того, что анализируемый период совпал с периодом активного экономического роста.

Решение. Определим гипотезы:

H_1 - "активный экономический рост"

H_2 - "умеренный экономический рост"

H_3 - "низкий экономический рост"

Определим событие А - "доллар дорожает"

Имеем: $P(H_1) = 0.3$, $P(H_2) = 0.5$, $P(H_3) = 0.2$,

$P(A/H_1) = 0.7$,

$P(A/H_2) = 0.4$ и $P(A/H_3) = 0.2$

Найти: $P(H_1/A)$

Используя формулу Байеса и подставляя заданные значения вероятностей, имеем:

$$P(H_1/A) = \frac{P(H_1)\cdot P((A/H_1)}{P(H_1)\cdot P(A/H_1)+P(H_2)\cdot P(A/H_2)+P(H_3)\cdot P(A/H_3)} =$$

$$= \frac{0.30\cdot 0.70}{0.30\cdot 0.70+0.50\cdot 0.40+0.20\cdot 0.20} = 0.467$$

Мы можем получить тот же результат при помощи таблицы:

Таблица 3

Гипотезы H_i	Априорные вероятности $P(H_i)$	Условные вероятности $P(A/H_i)$	Совместные вероятности $P(A\cap H_i)$	Апостериорные вероятности $P(H_i/A)$
H_1	0.30	0.70	0.21	0.21 / 0.45 = 0.467
H_2	0.50	0.40	0.20	0.20 / 0.45 = 0.444
H_3	0.20	0.20	0.04	0.04 / 0.45 = 0.089
Сумма	1.00	--	0.45	1

Мы можем также построить дерево решений для нашей задачи: