Математические методы в биологии

Блок 1. Основы теории вероятностей, или случайные события

Лекция 2

Козлова Ольга Сергеевна
89276755130, olga-sphinx@yandex.ru

Следствия теорем сложения и умножения

- **Теорема сложения вероятностей для несовместных событий.** Вероятность появления одного из двух несовместных событий (безразлично, какого), равна сумме вероятности этих событий: $P(A+B) = P(A) + P(B)$

Теорема сложения вероятностей для совместных событий (одно не исключает другое). Вероятность появления хотя бы одного из двух совместных событий А и В равна сумме вероятностей этих событий без вероятности их совместного появления: $P(A+B) = P(A) + P(B) - P(AB)$

Доказательство. Т.к. события совместны, то событие А+В эквивалентно наступлению одного из событий (причём все они уже несовместны):

- А наступило, В – нет
- В наступило, А – нет
- Наступило и А, и В

Пример совместных событий: А – появление 4х очков;
В – появление чётного числа очков.

По формуле сложения событий $P(A+B) = P(A\bar{B}) + P(\bar{A}B) + P(AB)$.

По той же формуле $P(A) = P(A\bar{B}) + P(AB) \Rightarrow P(A\bar{B}) = P(A) - P(AB)$

По той же формуле $P(B) = P(\bar{A}B) + P(AB) \Rightarrow P(\bar{A}B) = P(B) - P(AB)$

Т.о. $P(A+B) = P(A) - P(AB) + P(B) - P(AB) + P(AB) = P(A) + P(B) - P(AB)$

Теорема сложения вероятностей **трёх** совместных событий

Т.к. события совместны, то событие A+B+C эквивалентно наступлению <u>одного из</u> событий (причём все они уже несовместны):

- A и B наступили, C – нет
- A и C наступили, B – нет
- B и C наступили, A – нет
- A наступило, B и C – нет
- B наступило, A и C – нет
- C наступило, A и B – нет
- Наступили все три события

По формуле сложения событий $P(A+B+C) = P(AB\bar{C}) + P(A\bar{B}C) + P(\bar{A}BC) + P(A\overline{BC}) + P(\bar{A}B\bar{C}) + P(\overline{AB}C) + P(ABC)$

- $P(A) = P(AB\bar{C}) + P(A\bar{B}C) + P(A\overline{BC}) + P(ABC)$
  $\Rightarrow P(A\overline{BC}) = P(A) - P(AB\bar{C}) - P(A\bar{B}C) - P(ABC)$
- $P(B) = P(AB\bar{C}) + P(\bar{A}BC) + P(\bar{A}B\bar{C}) + P(ABC)$
  $\Rightarrow P(\bar{A}B\bar{C}) = P(B) - P(AB\bar{C}) - P(\bar{A}BC) - P(ABC)$
- $P(C) = P(A\bar{B}C) + P(\bar{A}BC) + P(\overline{AB}C) + P(ABC)$
  $\Rightarrow P(\overline{AB}C) = P(C) - P(A\bar{B}C) - P(\bar{A}BC) - P(ABC)$

После подстановки:

$$P(A+B+C) = \cancel{P(AB\bar{C})} + \cancel{P(A\bar{B}C)} + \cancel{P(\bar{A}BC)} + P(A) - \cancel{P(AB\bar{C})} - \cancel{P(A\bar{B}C)} - P(ABC)$$
$$+P(B) - P(AB\bar{C}) - \cancel{P(\bar{A}BC)} - P(ABC) + P(C) - P(A\bar{B}C) - P(\bar{A}BC) - P(ABC)$$
$$+P(ABC) = P(A) + P(B) + P(C) - P(AB\bar{C}) - P(A\bar{B}C) - P(\bar{A}BC) - 2 * P(ABC)$$

- $P(AC) = P(A\bar{B}C) + P(ABC) \Rightarrow P(A\bar{B}C) = P(AC) - P(ABC)$
- $P(BC) = P(\bar{A}BC) + P(ABC) \Rightarrow P(\bar{A}BC) = P(BC) - P(ABC)$
- $P(AB) = P(AB\bar{C}) + P(ABC) \Rightarrow P(AB\bar{C}) = P(AB) - P(ABC)$

После подстановки: $P(A+B+C) = P(A) + P(B) + P(C) - P(AB) + P(ABC) - P(AC) + P(ABC) - P(BC) + P(ABC) - 2 * P(ABC) = \underset{0,1}{P(A)} + \underset{0,2}{P(B)} + \underset{0,3}{P(C)} - \underset{0,02}{P(AB)} - \underset{0,03}{P(AC)} - \underset{0,06}{P(BC)} + \underset{0,006}{P(ABC)} = 0,496$

По формуле для вероятности противоположного события:

Следствия теорем сложения и умножения

Формула полной вероятности. Пусть событие A наступает только при условии наступления одного из несовместных событий B_1, B_2, B_3, …, B_n, образующих вместе полную группу событий. Пусть также известны вероятности этих событий B и условные вероятности наступления события A для каждого из событий B ($P_{B_n}(A)$). Тогда вероятность события A равна сумме произведений вероятностей B на соответствующую условную вероятность A:

$$P(A) = P(B_1) * P_{B_1}(A) + P(B_2) * P_{B_2}(A) + P(B_3) * P_{B_3}(A) + \cdots + P(B_n) * P_{B_n}(A)$$

Доказательство. Так как A наступает только в случае наступления одного из B (всё равно, какого), событие «появление A» эквивалентно одному из несовместных событий $\{B_1A, B_2A, B_3A, \ldots, B_nA\}$. По формуле сложения вероятностей для несовместных событий $P(A) = P(B_1A) + P(B_2A) + P(B_3A) + \cdots + P(B_nA)$

Распишем каждое из слагаемых

по формуле условной вероятности: $P(B_1A) = P(B_1) * P_{B_1}(A)$

$$P(B_2A) = P(B_2) * P_{B_2}(A)$$

…

$$P(B_nA) = P(B_n) * P_{B_n}(A)$$

Решение задач через формулу полной вероятности

Задача 1. Имеется два набора деталей. Вероятность того, что деталь из первого набора без брака – 0,8; вероятность того, что деталь из второго набора без брака – 0,9. Найти вероятность того, что случайно взятая деталь из случайно взятого набора – без брака.

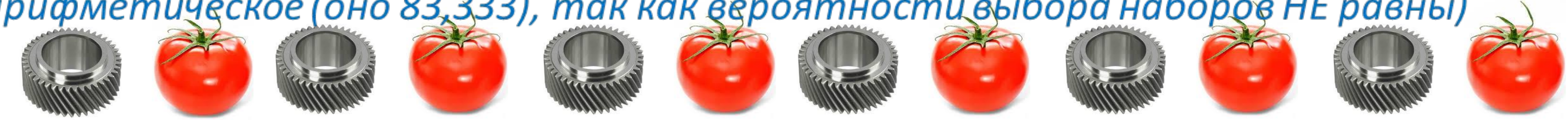
Решение. Событие А – «Извлечённая деталь без брака». Естественно, событие А наступает только после того, как произойдёт одно из событий группы В – «выбор одного или другого набора».

Применим формулу: $P(A) = P(B_1) * P_{B_1}(A) + P(B_2) * P_{B_2}(A) = 0{,}5 * 0{,}8 + 0{,}5 * 0{,}9 = 0{,}85$ *(среднее арифметическое, так как вероятности выбора наборов равны)*

Задача2. Три поставщика поставляют в магазин помидоры в соотношении 5:8:7 (на каждые 5 помидоров 1го поставщика второй поставляет 8, а третий – 7). Среди помидоров 1го поставщика – 90% свежие, у второго – 85% свежие, у третьего – 75% свежие. Найти вероятность того, что случайным образом выбранный помидор будет свежим.

Решение. Событие А («помидор свеж») наступает только после того, как произойдёт одно из событий группы В – «выбор одного из поставщиков». Вероятность выбора поставщика 1 равна 5/20=1/4=0,25; вероятность выбора поставщика 2 равна 8/20=0,4; вероятность выбора поставщика 3 равна 7/20=0,35.

Применим формулу: $P(A) = P(B_1) * P_{B_1}(A) + P(B_2) * P_{B_2}(A) + P(B_3) * P_{B_3}(A) = 0{,}25 * 0{,}9 + 0{,}4 * 0{,}85 + 0{,}35 * 0{,}75 = 0{,}225 + 0{,}34 + 0{,}2625 = 0{,}8275$ *(НЕ среднее арифметическое (оно 83,333), так как вероятности выбора наборов НЕ равны)*



Формула Байеса

Пусть событие А наступает только при условии наступления одного из несовместных событий $B_1, B_2, B_3, \ldots, B_n$, образующих вместе полную группу событий.

ВОПРОС: Можем ли мы узнать, как изменились вероятности событий группы B при условии, что событие А наступило?

ОТВЕТ: Да, для этого надо найти условные вероятности $P_A(B_1), P_A(B_2), P_A(B_3), \ldots, P_A(B_n)$

По формуле произведения вероятностей в общем виде: $P(AB_1) = P(A) * P_A(B_1) = P(B_1) * P_{B_1}(A)$, значит, $P_A(B_1) = \frac{P(B_1) * P_{B_1}(A)}{P(A)}$

В то же время, по формуле полной вероятности $P(A) = P(B_1) * P_{B_1}(A) + P(B_2) * P_{B_2}(A) + P(B_3) * P_{B_3}(A) + \cdots + P(B_n) * P_{B_n}(A)$

Отсюда $P_A(B_1) = \frac{P(B_1) * P_{B_1}(A)}{P(B_1) * P_{B_1}(A) + P(B_2) * P_{B_2}(A) + P(B_3) * P_{B_3}(A) + \cdots + P(B_n) * P_{B_n}(A)}$

Аналогично для любого $i, i = \overline{1, n}$:

$$P_A(B_i) = \frac{P(B_i) * P_{B_i}(A)}{P(B_1) * P_{B_1}(A) + P(B_2) * P_{B_2}(A) + P(B_3) * P_{B_3}(A) + \cdots + P(B_n) * P_{B_n}(A)}$$

Компактный вид формулы Байеса:

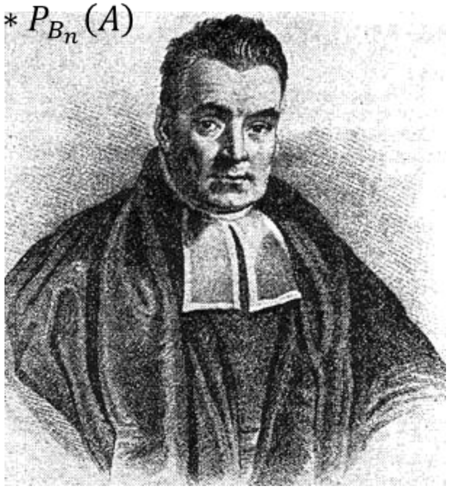
$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Априорная, «до-опытная» вероятность B)

Вероятность B при наступлении A (апостериорная, «после-опытная» вероятность)

Событие B также часто называют гипотезой

Полная вероятность A



Решение задач по формуле Байеса

Пусть в задаче про помидоры купленный помидор оказался свежим. Какова вероятность того, что он попал к нам в руки от третьего поставщика?

<u>Решение.</u> Событие А – «помидор свеж». Всего можно сделать три предположения (гипотезы):

- Это помидор от первого поставщика
- Это помидор от второго поставщика
- Это помидор от третьего поставщика

Априорная вероятность B_3

По формуле Байеса вероятность того, что этот помидор – от третьего поставщика, равна

$$P_A(B_3) = \frac{P(B_3) * P_{B_3}(A)}{P(B_1) * P_{B_1}(A) + P(B_2) * P_{B_2}(A) + P(B_3) * P_{B_3}}$$

Усл.вер-ть А

Апостериорная вероятность события B_3

Полная вероятность события А

Вероятность выбора 3го поставщика

Вероятность того, что у 3го поставщика помидор свежий

Подставляем значения:

$$P_A(B_3) = \frac{0{,}35 * 0{,}75}{0{,}25 * 0{,}9 + 0{,}4 * 0{,}85 + 0{,}35 * 0{,}75} = \frac{0{,}2625}{0{,}225 + 0{,}34 + 0{,}2625} = \frac{0{,}2625}{0{,}8275} \approx 0{,}317$$

Вывод. До испытания (покупки помидора) вероятность того, что случайный помидор – от третьего поставщика, составляла 0,35. Сейчас, для уже свершившегося события, она составляет 0,317. Таким образом, использование формулы Байеса позволило нам **<u>переоценить вероятность гипотезы</u>**.

Формула Бернулли

Пусть производится n независимых испытаний, в каждом из которых событие А может либо появиться, либо нет (причём вероятность его появления p одинакова!).
ВОПРОС: Можем ли мы вычислить вероятность того, что из этих n испытаний событие наступит <u>ровно k раз</u>?
ОТВЕТ: Да, и эта величина будет записана как $P_n(k)$
По теореме умножения вероятностей вероятность того, что событие А наступит k раз и не наступит (n-k) раз, равна $p^k (1-p)^{(n-k)}$.
Вопрос: сколькими способами можно выбрать k испытаний, в которых наступило интересующее нас событие А, среди общего количества испытаний n?
Формула Бернулли: $P_n(k) = C_n^k * p^k (1-p)^{n-k}$

<u>Пример.</u>
Вероятность того, что расход электроэнергии в продолжение одних суток не превысит установленной нормы, равна 0,75.
Найти вероятность того, что в ближайшие 6 суток расход электроэнергии в течение 4 суток не превысит нормы.
Решение. Каждый день – это испытание (n=6), а k=4.
Распишем формулу Бернулли:

$$P_n(k) = C_n^k * p^k (1-p)^{n-k} = C_6^4 * 0{,}75^4 * (1-0{,}75)^{4-2}$$
$$= \frac{6!}{4!\,(6-4)!} * 0{,}75^4 * 0{,}25^2 \approx 15 * 0{,}316 * 0{,}0625 \approx 0{,}3$$

Весь классический тер-вер на одном слайде

Что?	Как считать?	Условия применения	Пример
Классическое определение вероятности		Равновероятность элементарных исходов	Выпадение чётного числа очков
Число перестановок	$P_n = n!$	-	Размещение 3х книг на 3х местной полке
Число размещений из n по m			Составление 2х-букв.слов из 4х букв
Число размещений из *n* по *m* с повторением		Неограниченное количество элементов типа *n*	Составление 3х-значн.чисел из 1,2,3,4,5
Число сочетаний из n по m			Выбор 2х деталей из 4х
Формула сложения вероятностей		События несовместны	Устроит 1 событие из 2х (всё равно, какое)
Формула условной вероятности			Вер-ть B при усл-ии, что A уже случилось
Общий вид формулы умножения вер-тей		-	Устроит только пара совместно случившихся событий
Ф-ла противоположной вероятности		События независимы в совокупности	Вероятность появления хотя бы одного события
Формула полной вероятности		Событие A наступает только после одного из B; B образуют полную группу событий	Когда вероятности наступления A после B разные
Формула Байеса		Событие A наступает только после одного из B; B образуют полную группу событий	Переоценить вероятность B после того, как наступило A
Формула Бернулли			Вероятность того, что в 8 семьях родится 5 девочек

Парадокс дней рождения

Ну надо же, мы родились в один день! Какое невероятное совпадение!

Вовсе нет! Анна пригласила столько народу, что вероятность такого совпадения больше 50%!

ВОПРОС: Сколько же людей минимум должна была пригласить на вечеринку Анна?

Парадокс дней рождения

- Пусть дни рождения распределены равномерно по всему году
- Пусть в году ровно 365 дней, и нет високосных лет
- На вечеринке нет близнецов и людей, заведомо родившихся в один и тот же день

Рассчитаем вероятность противоположного события, то есть того, что дни рождения всех участников вечеринки различны.

Для второго человека эта вероятность равна $1-\frac{1}{365}$=$\frac{364}{365}$, для третьего $1-\frac{2}{365}$=$\frac{363}{365}$, для n-го $1-\frac{n-1}{365}$=$\frac{365-n+1}{365}$.

По формуле умножения вероятностей вероятность того, что дни рождения n участников вечеринки различны, равна

$$\frac{364}{365}*\frac{363}{365}*\ldots*\frac{365-n+1}{365}=\frac{364*363*\cdots*(365-n+1)}{365^{n-1}}=\frac{365*364*363*\cdots*(365-n+1)}{365^{n}}=\frac{365!}{365^{n}*(365-n)!}$$

Значит, вероятность того, что хотя бы у двух человек дни рождения в один день, равна $1-\frac{365!}{365^{n}*(365-n)!}$, и она должна быть больше 0,5.

Такое произойдёт при $n \geq 23$, значит, на вечеринке у Анны присутствует не менее 23 человек, и этого числа достаточно, чтобы вероятность совпадения дней рождения была больше вероятности несовпадения!