

Математические методы в биологии

Блок 2. Случайные величины

Лекция 3

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

Дискретная случайная величина

- Случайная величина – величина, которая в результате испытания принимает одно и только одно возможное значение, наперёд не известное и зависящее от случайных обстоятельств, которые заранее не могут быть учтены.

Пример. Выпадение определённого числа очков на игральной кости (от 1 до 6). Число очков – случайная величина.

- Дискретная случайная величина – случайная величина, которая может принимать конечные, изолированные значения из некоторого числового промежутка
- Непрерывная случайная величина – случайная величина, которая может принимать все значения из некоторого числового промежутка

Пример. Содержание какого-либо фермента в крови

- Закон распределения вероятностей дискретной случайной величины – это сопоставление всех возможных значений случайной величины и их вероятностей.

Тривиальный пример. Случайная величина – сторона монетки. Она принимает два изолированных значения – либо «орёл», либо «решка» и

подчинена следующей

X	«Орёл» (0)	«Решка» (1)	Контроль
P	0,5	0,5	$0,5+0,5=1$

Название
случайной
Векторы
вероятностей

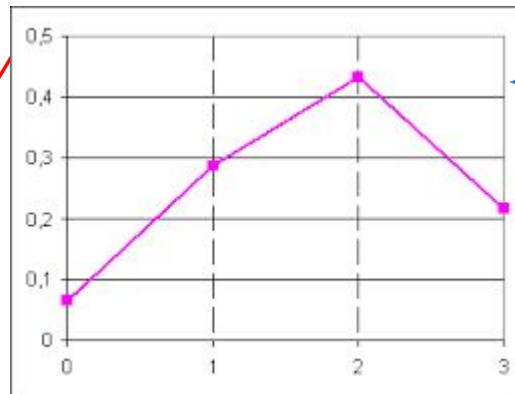
Способы задания распределения вероятностей дискретной случайной величины

- Таблично

Из ящика, в котором лежат 2 белых и 8 чёрных шаров, последовательно вынимают шары до тех пор, пока не появится чёрный шар. Число вынутых шаров – есть дискретная случайная величина X , которая может принимать изолированные значения на промежутке от 1 до 3. Зададим закон её распределения таблично.

X	1	2	3	Контроль
P	$8/10$ (уже первый шар – чёрный)	$2/10 * 8/9 = 8/45$ (первый – белый, второй – чёрный)	$2/10 * 1/9 * 8/8 = 1/45$ (два первых – белые, третий – чёрный)	$8/10 + 8/45 + 1/45 = 1$

- Графически
- Аналитически



← Многоугольник распределения дискретной случайной величины, принимающей изолированные значения на промежутке от 0 до 3.

Биномиальное распределение

- n независимых испытаний
- Событие A появляется в каждом из них с вероятностью p
- Дискретная случайная величина X – число испытаний, в которых появилось событие A

ВОПРОС: Можем ли мы задать закон распределения дискретной случайной величины X ?

ОТВЕТ: Да, для этого нужно знать:

- Возможные значения случайной величины X
- Вероятности того, что X примет каждое из этих возможных значений

Найдём а): $X=0,1,2,\dots,n$ (всего $n+1$ значений)

Найдём б): $P(X = k) = C_n^k * p^k (1 - p)^{n-k}$, где $k=X=0,1,2,\dots,n$ (формула Бернулли)

Бином Ньютона для $(p + (1 - p))^n$

Аналитическое выражение
закона распределения д.с.в.

Определение. Биномиальное распределение – это то распределение, которое определяется формулой Бернулли с заданными n и p .

Биномиальное распределение. Пример.

Монета брошена 5 раз. Задать распределение случайной величины X – числа выпадения «гербов» аналитически, таблично и графически.

Решение. $n=5, X=0,1,2,\dots,5$ (всего 6 возможных значений).

Аналитически: $P(X = X_i) = C_5^{X_i} * 0,5^{X_i}(1 - 0,5)^{5-X_i}, i = \overline{1,6}$

Таблично: $i=1, X_i=0, P_5(0) = C_5^0 * 0,5^0(1 - 0,5)^{5-0} = \frac{5!}{5!} * 1 * 0,5^5 = 0,03125$

$i=2, X_i=1. P_5(1) = C_5^1 * 0,5^1(1 - 0,5)^{5-1} = \frac{5!}{5! 1! 4!} * 0,5 * 0,5^4 = 0,15625$

$i=3, X_i=2. P_5(2) = C_5^2 * 0,5^2(1 - 0,5)^{5-2} = \frac{5!}{2! 3!} * 0,5^2 * 0,5^3 = 0,3125$

$i=4, X_i=3. P_5(3) = C_5^3 * 0,5^3(1 - 0,5)^{5-3} = \frac{5!}{3! 2!} * 0,5^3 * 0,5^2 = 0,3125$

$i=5, X_i=4. P_5(4) = C_5^4 * 0,5^4(1 - 0,5)^{5-4} = \frac{5!}{4! 1!} * 0,5^4 * 0,5^1 = 0,15625$

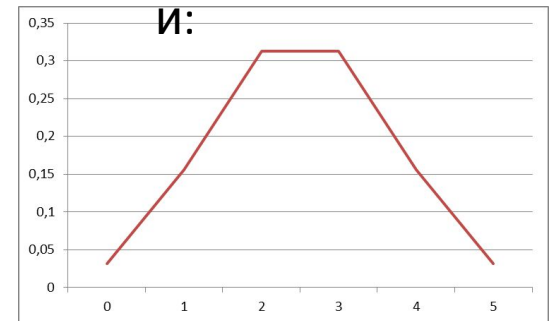
$i=6, X_i=5. P_5(5) = C_5^5 * 0,5^5(1 - 0,5)^{5-5} = \frac{5!}{5! 0!} * 0,5^5 * 0,5^0 = 0,03125$

$$\sum P_5(i) = 1$$

X	0	1	2	3	4	5
P	0,03125	0,15625	0,3125	0,3125	0,15625	0,03125

Графическ

и:



Геометрическое распределение

- Событие А появляется в каждом из независимых испытаний с вероятностью p
- Пусть испытания заканчиваются, как только появляется событие А
- X – дискретная случайная величина – число событий до первого появления А (включая само событие, в котором А появилось)

ВОПРОС: Можем ли мы задать закон распределения дискретной случайной величины X ?

ОТВЕТ: Да, для этого для каждого номера испытания k надо задать вероятность того, что в предыдущих $(k-1)$ испытаниях событие не наступило, а наступило оно только в k -ом испытании!

Так как все испытания независимы в совокупности, по формуле умножения вероятностей $P(X = k) = (1 - p)^{k-1}p$

Геометрическая прогрессия – степени числа $(1-p)$

Определение. Геометрическое распределение – это то распределение, которое определяется формулой геометрической прогрессии со знаменателем $(1-p)$ и первым членом p

Геометрическое распределение. Пример.

Стрелок стреляет по цели до первого попадания. Вероятность попадания в цель = 0,6. Задать распределение дискретной случайной величины X – числа выстрелов до первого попадания.

Решение. $P(X = 1) = (1 - 0,6)^{1-1}0,6 = 0,6$

$$P(X = 2) = (1 - 0,6)^{2-1}0,6 = 0,24$$

$$P(X = 3) = (1 - 0,6)^{3-1}0,6 = 0,096$$

$$P(X = 4) = (1 - 0,6)^{4-1}0,6 = 0,0384$$

$$P(X = 5) = (1 - 0,6)^{5-1}0,6 = 0,01536$$

...

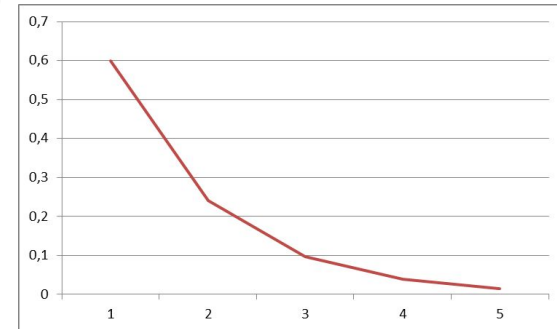
X	P
1	0,6
2	0,24
3	0,096
4	0,0384
5	0,01536

Сходимость ряда вероятностей. По признаку Даламбера, если существует такое $0 < c < 1$, что, начиная с некоторого n $\frac{a_{n+1}}{a_n} \leq c$, то ряд сходится.

Действительно, $\frac{(1-p)^{(n+1)-1} * p}{(1-p)^{n-1} * p} = 1 - p \leq c$

Чему равна сумма ряда? Для геометрической прогрессии со знаменателем $|a| < 1$

$$\sum_{n=0}^{\infty} a^n = \frac{1}{1-a} \Rightarrow \sum_{k=1}^{\infty} (1-p)^{k-1} * p = p * \frac{1}{1-(1-p)} = 1$$



Гипергеометрическое распределение

- Есть N деталей, из них M – без брака ($M < N$)
- Случайным образом отбираем n деталей
- X – дискретная случайная величина – число деталей без брака среди n

ВОПРОС: Можем ли мы задать распределение дискретной случайной величины X ?

ОТВЕТ: Да, для этого нужно знать:

- а) Возможные значения случайной величины
- б) Вероятности того, что X примет каждое из этих возможных значений

Найдём а): $X=0,1,2,\dots,\min(n,M)$

Вероятность того, что среди n деталей будет ровно m без брака

Найдём б): по классическому определению вероятности $P(X = m) = \frac{C_M^m * C_{N-M}^{n-m}}{C_N^n}$

(число эл. исходов, благоприятствующих событию / общее число эл.исх.)

C_M^m - сколькими способами можно выбрать m деталей из M без брака

C_{N-M}^{n-m} - сколькими способами можно выбрать $n-m$ деталей из $N-M$ с браком

Гипергеометрическое распределение определяется тремя параметрами –

N, M и n .

Гипергеометрическое распределение. Пример.

Среди 50 деталей 20 окрашенных. Задать распределение дискретной случайной величины X – числа окрашенных деталей среди наудачу извлечённых пяти?

Решение. $N=50$, $M=20$, $n=5$. Д.с.в. X принимает значения $0, 1, 2, \dots, \min(5, 20)$, т.е. $X = 0, 5$ (всего 6 значений).

Аналитически:
$$P(X = m) = \frac{C_{20}^m * C_{30}^{5-m}}{C_{50}^5}$$

Таблично:
$$P(X = 0) = \frac{C_{20}^0 * C_{30}^{5-0}}{C_{50}^5} = \frac{1 * \frac{30!}{5! * 25!}}{\frac{50!}{5! * 45!}} = \frac{142506}{2118760} \approx 0,067$$

$$P(X = 1) = \frac{C_{20}^1 * C_{30}^{5-1}}{C_{50}^5} = \frac{20 * \frac{30!}{4! * 26!}}{\frac{50!}{5! * 45!}} = \frac{548100}{2118760} \approx 0,259$$

$$P(X = 2) = \frac{C_{20}^2 * C_{30}^{5-2}}{C_{50}^5} = \frac{\frac{20!}{2! * 18!} * \frac{30!}{3! * 27!}}{\frac{50!}{5! * 45!}} = \frac{190 * 4060}{2118760} \approx 0,36$$

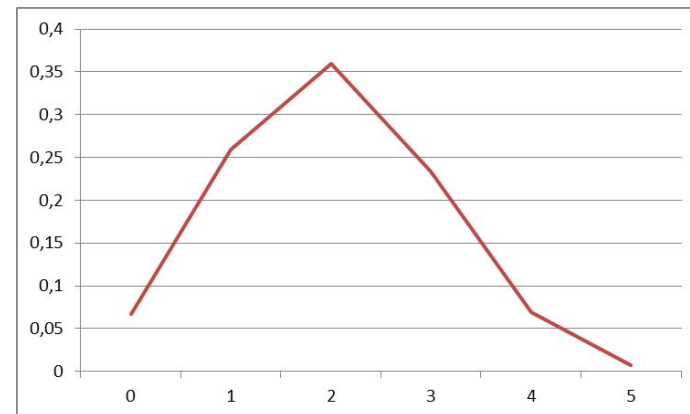
$$P(X = 3) = \frac{C_{20}^3 * C_{30}^{5-3}}{C_{50}^5} = \frac{\frac{20!}{3! * 17!} * \frac{30!}{2! * 28!}}{\frac{50!}{5! * 45!}} = \frac{1140 * 435}{2118760} \approx 0,234$$

$$P(X = 4) = \frac{C_{20}^4 * C_{30}^{5-4}}{C_{50}^5} = \frac{\frac{20!}{4! * 16!} * \frac{30!}{1! * 29!}}{\frac{50!}{5! * 45!}} = \frac{4845 * 30}{2118760} \approx 0,069$$

$$P(X = 5) = \frac{C_{20}^5 * C_{30}^{5-5}}{C_{50}^5} = \frac{\frac{20!}{5! * 15!} * 1}{\frac{50!}{5! * 45!}} = \frac{15504}{2118760} \approx 0,0073$$

Контроль: $0,067 + 0,259 + 0,36 + 0,234 + 0,069 + 0,0073 \approx 1$

X	P
0	0,067
1	0,013
2	0,36
3	0,234
4	0,069
5	0,0073



Числовые характеристики дискретной случайной величины

● Математическое ожидание д.с.в. - сумма произведений всех её возможных значений на их вероятности. $M(X) = \sum_{i=1}^n X_i * P(X = X_i)$

Пример. Биномиальное распределение числа выпадения гербов в 5 подбросах.

X	0	1	2	3	4	5
P	0,03125	0,15625	0,3125	0,3125	0,15625	0,03125

$$M(X) = 0 * 0,03125 + 1 * 0,15625 + 2 * 0,3125 + 3 * 0,3125 + 4 * 0,15625 + 5 * 0,03125 = 2,5$$

ВАЖНО:

- С увеличением числа испытаний среднее арифметическое значений д.с.в. стремится к математическому ожиданию.
- Мат.ожидание числа появлений события в одном испытании равно его вероятности.

СВОЙСТВА МАТ.ОЖИДАНИЯ:

- $M(C) = C$
- $M(C * X) = C * M(X)$
- Для независимых случайных величин X и Y $M(XY) = M(X) * M(Y)$
- Для n независимых случайных величин $M(\prod_{i=1}^n X_i) = \prod_{i=1}^n M(X_i)$
- $M(X + Y) = M(X) + M(Y)$
- $M(\sum_{i=1}^n X_i) = \sum_{i=1}^n M(X_i)$

Независ.д.с.в. – те д.с.в., закон распределения которых не зависит от значений др. д.с.в.

Доказательства свойств M.c

X	C
---	---

Распределение такой «случайной» величины в табл.

1. $M(C) = C$

P	1
---	---

 $\Rightarrow M(X) = C * 1 = C$

2. $M(C * X) = C * M(X)$

X	X ₁	X ₂	...	X _n
P	p ₁	p ₂	...	p _n

X'	C*X ₁	C*X ₂	...	C*X _n
P	p ₁	p ₂	...	p _n

$\Rightarrow M(X') = C * X_1 * p_1 + C * X_2 * p_2 + \dots + C * X_n * p_n = C(X_1 p_1 + \dots + X_n p_n) = C * M(X)$

3. $M(XY) = M(X) * M(Y)$

X	X ₁	X ₂
P	p ₁	p ₂

*

Y	Y ₁	Y ₂
P	q ₁	q ₂

=

XY	X ₁ Y ₁	X ₁ Y ₂	X ₂ Y ₁	X ₂ Y ₂
P	p ₁ q ₁	p ₁ q ₂	p ₂ q ₁	p ₂ q ₂

$\Rightarrow M(XY) = X_1 Y_1 * p_1 q_1 + X_1 Y_2 * p_1 q_2 + X_2 Y_1 * p_2 q_1 + X_2 Y_2 * p_2 q_2$
 $= X_1 p_1 (Y_1 q_1 + Y_2 q_2) + X_2 p_2 (Y_1 q_1 + Y_2 q_2) = (Y_1 q_1 + Y_2 q_2) * (X_1 p_1 + X_2 p_2)$
 $= MX * MY$

4. $M(X + Y) = M(X) + M(Y)$

X	X ₁	X ₂
P	p ₁	p ₂

+

Y	Y ₁	Y ₂
P	q ₁	q ₂

=

X+Y	X ₁ +Y ₁	X ₁ +Y ₂	X ₂ +Y ₁	X ₂ +Y ₂
P	p ₁₁	p ₁₂	p ₂₁	p ₂₂

$\Rightarrow M(X + Y) = (X_1 + Y_1) * p_{11} + (X_1 + Y_2) * p_{12} + (X_2 + Y_1) * p_{21} + (X_2 + Y_2) * p_{22}$
 $= X_1(p_{11} + p_{12}) + X_2(p_{21} + p_{22}) + Y_1(p_{11} + p_{21}) + Y_2(p_{12} + p_{22})$

$X = X_1 \Rightarrow X + Y = X_1 + Y_1$ или $X + Y = X_1 + Y_2$. $P(X + Y = X_1 + Y_1) = p_{11}, P(X + Y = X_1 + Y_2) = p_{12} \Rightarrow p_1 = p_{11} + p_{12}$

Математическое ожидание числа появления события в независимых испытаниях

- n независимых испытаний
- Событие A появляется в каждом из них с вероятностью p
- Дискретная случайная величина X – число появления события A в этих испытаниях

ВОПРОС: Чему равно среднее число (математическое ожидание случайной величины X) появлений события A в испытаниях?

ОТВЕТ: Математическое ожидание числа появлений события A в n испытаниях равно произведению n на p : $M(X) = n * p$

Доказательство. Пусть X_1 – число появления события A в первом испытании, X_2 – во втором и т.д., X_n – в n -ом. Всего событие A появилось $X_1 + X_2 + \dots + X_n$ раз. По свойству математического ожидания суммы, $M(X) = M(X_1) + M(X_2) + \dots + M(X_n)$. Так как математическое ожидание числа появлений события в одном испытании равно его вероятности, $M(X_1) = M(X_2) = \dots = M(X_n) = p$. Отсюда $M(X) = n * p$.

*Данная случайная величина X распределена по биномиальному закону, поэтому можно сказать, что **МАТЕМАТИЧЕСКОЕ ОЖИДАНИЕ БИНОМИАЛЬНОГО РАСПРЕДЕЛЕНИЯ С ПАРАМЕТРАМИ n И p РАВНО ПРОИЗВЕДЕНИЮ $n * p$.***

Резюме

- $P(X = k) = C_n^k * p^k (1 - p)^{n-k}$ - формула биномиального распределения с параметрами n и p
- $P(X = k) = (1 - p)^{k-1} p$ - формула геометрического распределения с параметром p
- $P(X = m) = \frac{C_M^m * C_{N-M}^{n-m}}{C_N^n}$ - формула гипергеометрического распределения с параметрами N , M и n
- $M(X) = \sum_{i=1}^n X_i * P(X = X_i)$ - формула математического ожидания дискретной случайной величины