Теория вероятностей и математическая статистика

Вебинар 2

Дискретные случайные величины. Закон распределения вероятностей.
Биномиальный закон распределения. Распределение Пуассона

Случайные величины

Случайная величина — величина, которая в результате опыта принимает некоторое значение, неизвестное заранее.

Дискретные случайные величины, как правило, принимают целые или рациональные значения. Эти значения отделены друг от друга, т.е. если случайная величина принимает значения 1 и 2, то она не обязана принимать промежуточные значения.

Непрерывные случайные величины принимают вещественные значения. Здесь значения уже не отделены друг от друга, т.е. если непрерывная случайная величина принимает значения 1 и 2, то она также может принять и *любое* значение между ними.

Случайная величина — величина, которая в результате опыта принимает некоторое значение, неизвестное заранее.

Дискретные случайные величины, как правило, принимают целые или рациональные значения. Эти значения отделены друг от друга, т.е. если случайная величина принимает значения 1 и 2, то она не обязана принимать промежуточные значения.

Непрерывные случайные величины принимают вещественные значения. Здесь значения уже не отделены друг от друга, т.е. если непрерывная случайная величина принимает значения 1 и 2, то она также может принять и *любое* значение между ними.

Примеры дискретных случайных величин:

1. Сумма очков при 100-кратном подбрасывании игрального кубика.
2. Число метеоритов, упавших на Землю за год.
3. Количество машин, которые успевают проехать через данный светофор за один цикл.

GeekBrains

Пусть X — дискретная случайная величина. *Закон распределения* этой случайной величины — это соответствие между значениями, которые принимает эта величина, и вероятностями, с которыми она их принимает.

Закон распределения

Пусть X — дискретная случайная величина. *Закон распределения* этой случайной величины — это соответствие между значениями, которые принимает эта величина, и вероятностями, с которыми она их принимает.

Например, пусть X — сумма значений двух подбрасываемых игральных кубиков. Вот её закон распределения:

x	2	3	4	5	6	7	8	9	10	11	12
$P(X=x)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

Пусть X — дискретная случайная величина. *Закон распределения* этой случайной величины — это соответствие между значениями, которые принимает эта величина, и вероятностями, с которыми она их принимает.

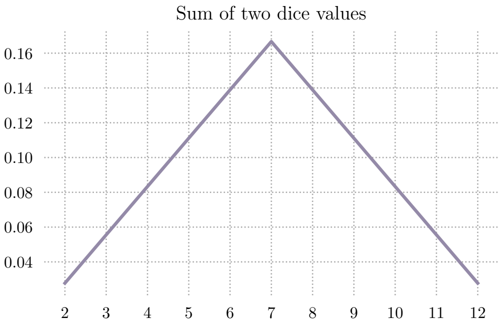
Например, пусть X — сумма значений двух подбрасываемых игральных кубиков. Вот её закон распределения:

x	2	3	4	5	6	7	8	9	10	11	12
$P(X = x)$	$1/36$	$2/36$	$3/36$	$4/36$	$5/36$	$6/36$	$5/36$	$4/36$	$3/36$	$2/36$	$1/36$

Отметим, что сумма вероятностей дискретной случайной величины всегда равна 1.

Закон распределения также удобно изобразить графически: откладываем на оси x значения случайной величины, а на оси y — соответствующие им вероятности.

Например, вот график распределения случайной величины, заданной ранее.



Пусть X, Y — дискретные случайные величины, причём X принимает значения x_i с вероятностями $P(X = x_i)$, $i = 1, 2, \dots$, а Y принимает значения y_j с вероятностями $P(Y = y_j)$, $j = 1, 2, \dots$.

- Их *сумма* $Z = X + Y$ — случайная величина, которая принимает значения $z_{ij} = x_i + y_j$ с вероятностями $P(X = x_i, Y = y_j)$.
- Аналогично считаются *разность* и *произведение* случайных величин, надо лишь заменить соответствующие символы операций.
- *Квадрат* $Z = X^2$ — случайная величина, которая принимает значения $z_i = x_i^2$ по тому же закону распределения, что и X.

Пусть X, Y — дискретные случайные величины, причём X принимает значения x_i с вероятностями $P(X = x_i)$, $i = 1, 2, \ldots$, а Y принимает значения y_j с вероятностями $P(Y = y_j)$, $j = 1, 2, \ldots$.

- Их *сумма* $Z = X + Y$ — случайная величина, которая принимает значения $z_{ij} = x_i + y_j$ с вероятностями $P(X = x_i, Y = y_j)$.
- Аналогично считаются *разность* и *произведение* случайных величин, надо лишь заменить соответствующие символы операций.
- *Квадрат* $Z = X^2$ — случайная величина, которая принимает значения $z_i = x_i^2$ по тому же закону распределения, что и X.

Замечание: не стоит путать сумму случайных *событий* и сумму случайных *величин*.

Пусть X — случайная величина. *Математическим ожиданием* называется среднее значение величины X при стремлении количества испытаний к бесконечности. Обозначается $M(X)$.

Если X — дискретная случайная величина, принимающая значения x_i с вероятностями $p_i = P(X = x_i)$, $i = 1, 2, \ldots$, то

$$M(X) = \sum_i p_i x_i = p_1 x_1 + p_2 x_2 + \ldots$$

Дисперсия

Дисперсией случайной величины X называется математическое ожидание квадрата отклонения случайной величины от её математического ожидания:

$$D(X) = M\left((X - M(X))^2\right)$$

Дисперсия является мерой разброса случайной величины относительно её среднего значения.

Законы распределения случайных величин

Биномиальное распределение

Пусть имеется некоторое событие A, которое наступает с вероятностью p.
Биномиальный закон описывает распределение случайной величины X, задающей число наступлений события A в ходе проведения n независимых опытов.

Биномиальный закон распределения описывается *формулой Бернулли*:

$$P(X = k) = C_n^k \cdot p^k \cdot (1 - p)^{n-k}$$

Пусть имеется некоторое событие A, которое наступает с вероятностью p. *Биномиальный закон* описывает распределение случайной величины X, задающей число наступлений события A в ходе проведения n независимых опытов.

Биномиальный закон распределения описывается *формулой Бернулли*:

$$P(X = k) = C_n^k \cdot p^k \cdot (1 - p)^{n-k}$$

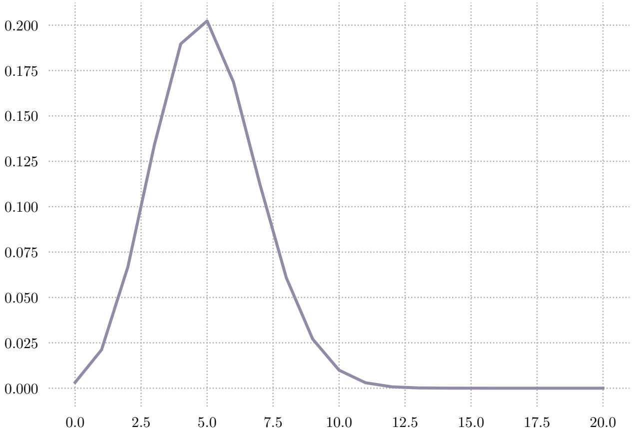
Математическое ожидание биномиального распределения:

$$M(X) = np$$

Дисперсия:

$$D(X) = np(1 - p)$$

Закон биномиального распределения с параметрами $n = 20$, $p = 0.25$:



Допустим теперь, что имеется некоторый поток событий, такой, что в среднем за единицу времени событие наступает λ раз (т.е. с *интенсивностью* λ). Тогда случайная величина X, равная количеству наступлений события за единицу времени, имеет *распределение Пуассона* с параметром λ.

Случайная величина X принимает значения $0, 1, 2, \dots$ (счётное множество значений), а соответствующие вероятности выражаются *формулой Пуассона*:

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Здесь λ — положительное вещественное число.

Распределение Пуассона

Как мы уже отметили, распределение Пуассона описывает счётчики событий, наступивших за единицу времени. Например, распределение Пуассона описывает:

1. число бракованных деталей в партии фиксированного размера,
2. число опечаток в тексте фиксированного размера,
3. число автобусов, проехавших за фиксированное время мимо автобусной остановки.

Распределение Пуассона

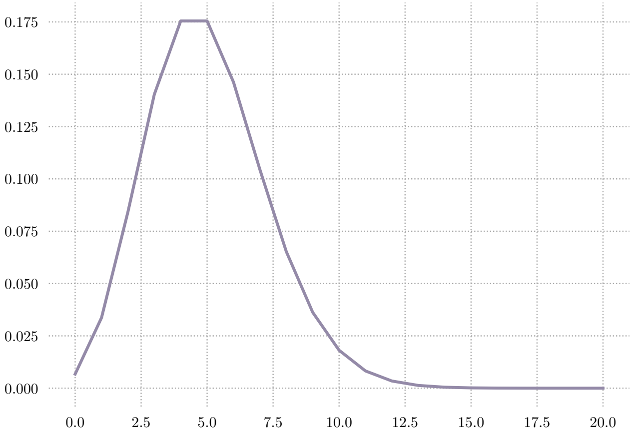
Как мы уже отметили, распределение Пуассона описывает счётчики событий, наступивших за единицу времени. Например, распределение Пуассона описывает:

1. число бракованных деталей в партии фиксированного размера,
2. число опечаток в тексте фиксированного размера,
3. число автобусов, проехавших за фиксированное время мимо автобусной остановки.

Математическое ожидание и дисперсия распределения Пуассона равны:

$$M(X) = D(X) = \lambda$$

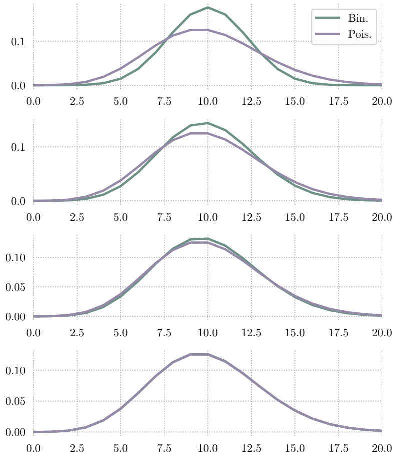
Закон распределения Пуассона с параметром $\lambda = 5$.



Распределение Пуассона

Распределение Пуассона является предельным случаем биномиального.
Если в последнем имеется очень большое число экспериментов ($n \to \infty$), а вероятность наступления события A достаточно мала (можно считать, что $p \approx \lambda/n$), то такое распределение становится очень похоже на распределение Пуассона с параметром $\lambda = np$.

Например, справа изображены графики биномиального распределения (мятный) и распределения Пуассона (фиолетовый). Во всех четырёх случаях $\lambda = 10$ и $p = 10/n$. Параметр n сверху вниз: 20, 40, 100, 1000.



- *Распределение Бернулли*. Событие A наступает с вероятностью p. Индикатор наступления этого события, т.е.

$$X = \begin{cases} 1, & \text{событие } A \text{ произошло}, \\ 0 & \text{иначе}, \end{cases}$$

имеет распределение Бернулли. Вероятности:

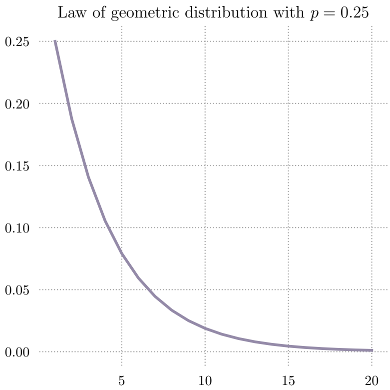
$$P(X = 1) = p, \ \ P(X = 0) = 1 - p$$

Замечание. Биномиальное распределение с параметрами n, p является суммой n распределений Бернулли с параметром p.

- *Дискретное равномерное распределение*. Случайная величина X принимает n различных значений с одинаковой вероятностью $1/n$. Не путать с *непрерывным равномерным*.

- *Геометрическое распределение*. Событие A наступает с вероятностью p. Случайная величина X, равная числу независимых испытаний до первого наступления события A, имеет геометрическое распределение. Вероятности:

$$P(X = k) = (1 - p)^k \cdot p$$



На следующем занятии

Описательная статистика. Качественные и количественные характеристики популяции. Графическое представление данных