Урок 5. Распределения случайных величин (часть 1)

Хакимов Р.И. + ChatGPT

Введение в распределения случайных величин

Распределения случайных величин описывают, как вероятности распределяются среди возможных значений случайной величины. Введение в распределения случайных величин включает понимание их основных типов, функций и свойств. Далее преведен обзор ключевых понятий и распределений.

Основные понятия

1. Случайная величина: Функция, которая каждому элементарному исходу случайного эксперимента сопоставляет числовое значение.

2. Функция распределения вероятностей:
- **Для дискретных случайных величин:** Функция вероятностей $P(X = x)$, где X — случайная величина, а x — конкретное значение.
- **Для непрерывных случайных величин:** Функция плотности вероятности $f(x)$, такая что вероятность того, что случайная величина X лежит в интервале $[a, b]$ вычисляется как интеграл от функции плотности:

$$P(a \leq X \leq b) = \int_a^b f(x)\, dx$$

Основные понятия

3. Функция распределения (CDF, Cumulative Distribution Function): Функция $F(x)$ определяет вероятность того, что случайная величина X примет значение меньше или равно x.

- Для дискретных случайных величин:

$$F(x) = P(X \leq x)$$

- Для непрерывных случайных величин:

$$F(x) = \int_{-\infty}^{x} f(t)\, dt$$

где $f(x)$ - функция плотности распределения случайной величины (probability density function).

Свойства распределений

Математическое ожидание (среднее)

Для дискретных величин:

$$E(X) = \sum_x x \cdot P(X = x)$$

Для непрерывных величин:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\, dx$$

Свойства распределений

Дисперсия

Для дискретных величин:

$$\mathrm{Var}(X) = E(X^2) - (E(X))^2$$

Для непрерывных величин:

$$\mathrm{Var}(X) = \int_{-\infty}^{\infty} (x - E(X))^2 \cdot f(x)\, dx$$

Свойства распределений

Функция распределения (CDF) указывает вероятность того, что случайная величина примет значение меньше или равно заданному значению.

Дискретные распределения

Дискретные распределения описывают вероятности, связанные с дискретными случайными величинами, которые могут принимать конечное или счётное количество значений. Рассмотрим следующие основные дискретные распределения: равномерное, биномиальное, геометрическое, гипергеометрическое и распределение Пуассона.

Дискретные распределения

Равномерное распределение описывает случайную величину, которая может принимать n различных значений, каждое из которых имеет одинаковую вероятность. Это распределение используется, когда все исходы равновероятны.

Функция вероятностей: Если случайная величина X принимает значения $x_1, x_2, \ldots, x_n$, то вероятность каждого значения:

$$P(X = x_i) = \frac{1}{n}$$

где $i = 1, 2, \ldots, n$.

Математическое ожидание (среднее):

$$E(X) = \frac{1}{n} \sum_{i=1}^{n} x_i$$

Дисперсия:

$$\mathrm{Var}(X) = \frac{1}{n} \sum_{i=1}^{n} (x_i - E(X))^2$$

Дискретные распределения

Равномерное распределение

Пример: Если мы бросаем симметричную игральную кость, то вероятности выпадения каждого числа от 1 до 6 равны $\frac{1}{6}$, математическое ожидание $E(X)$ равно 3.5, а дисперсия $\mathrm{Var}(X)$ равна 2.92.

Дискретные распределения

Биномиальное распределение моделирует количество успехов в n независимых испытаниях, каждое из которых имеет вероятность успеха p.
Функция вероятностей: Для случайной величины X, которая обозначает количество успехов:

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

где:
- $\binom{n}{k}$ — биномиальный коэффициент, равный $\frac{n!}{k!(n-k)!}$,
- k — число успехов,
- n — число испытаний,
- p — вероятность успеха в каждом испытании.

Дискретные распределения

Биномиальное распределение
Математическое ожидание (среднее):

$$E(X) = n \cdot p$$

Дисперсия:

$$\mathrm{Var}(X) = n \cdot p \cdot (1 - p)$$

Пример: Если вы бросаете монету 10 раз, и вероятность выпадения орла в каждом броске равна 0.5, вероятность получить ровно 6 орлов:

$$P(X = 6) = \binom{10}{6}(0.5)^6(0.5)^4 = \frac{10!}{6!4!} \cdot (0.5)^{10} \approx 0.205$$

Дискретные распределения

Геометрическое распределение описывает количество испытаний до первого успеха, где вероятность успеха в каждом испытании постоянна и равна p.
Функция вероятностей: Для случайной величины X, которая обозначает количество испытаний до первого успеха:

$$P(X = k) = (1 - p)^{k-1} p$$

где k — номер испытания, на котором произошел первый успех.
Математическое ожидание (среднее):

$$E(X) = \frac{1}{p}$$

Дисперсия:

$$\mathrm{Var}(X) = \frac{1 - p}{p^2}$$

Дискретные распределения

Геометрическое распределение

Пример: Если вероятность успеха в каждом испытании равна 0.3, вероятность того, что первый успех произойдет на 4-м испытании:

$$P(X = 4) = (1 - 0.3)^{4-1} \cdot 0.3 = 0.7^3 \cdot 0.3 \approx 0.1029$$

Дискретные распределения

Гипергеометрическое распределение моделирует количество успехов в выборке без замены из конечного популяционного набора, где в популяции N элементов, из которых K являются успехами.

Функция вероятностей: Для случайной величины X, которая обозначает количество успехов k в выборке из n элементов:

$$P(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

где:

- $\binom{K}{k}$ — число способов выбрать k успехов из K,
- $\binom{N-K}{n-k}$ — число способов выбрать $n-k$ неуспехов из оставшихся $N-K$,
- $\binom{N}{n}$ — число способов выбрать n элементов из N.

Дискретные распределения

Гипергеометрическое распределение
Математическое ожидание (среднее):

$$E(X) = n \cdot \frac{K}{N}$$

Дисперсия:

$$\mathrm{Var}(X) = n \cdot \frac{K}{N} \cdot \frac{N-K}{N} \cdot \frac{N-n}{N-1}$$

Пример: В коробке 20 шаров, из которых 7 красные и 13 синие. Если мы выбираем 5 шаров без замены, вероятность того, что из выбранных 5 шаров ровно 3 будут красными:

$$P(X=3) = \frac{\binom{7}{3}\binom{13}{2}}{\binom{20}{5}} \approx 0.176$$

Дискретные распределения

Распределение Пуассона моделирует количество событий, происходящих в фиксированном интервале времени или пространства, если события происходят с постоянной средней частотой и независимо друг от друга.
Функция вероятностей: Для случайной величины X, которая обозначает количество событий:

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

где λ — среднее число событий в интервале, k — число событий.
Математическое ожидание (среднее):

$$E(X) = \lambda$$

Дисперсия:

$$\mathrm{Var}(X) = \lambda$$

Дискретные распределения

Распределение Пуассона

Пример: Если в среднем в ресторане за час приходит 3 клиента, вероятность того, что за следующий час придет ровно 5 клиентов:

$$P(X = 5) = \frac{3^5 e^{-3}}{5!} \approx 0.1008$$