Содержание

Инструкция

Программа моделирует создание структур разной периодичности в стиле бус Бахо. На замкнутой красной нити последовательно появляются бусы, расстояние между которыми определяется значением случайной величины. Бусы можно создавать в ручном режиме (**"Получить новое значение"**) либо запустить (**"Запустить"**) процесс с определенным периодом обновления (**"Замедление процесса"**) и позже остановить (**"Остановить"**) либо сбросить (**"Сбросить"**).

Введем некоторые обозначения:
ξ - случайная величина равная расстоянию между последовательно сгенерированным бусами
$p(x) := p(\xi = x)$ - плотность распределения случайной величины
Δx_i - расстояние между i и $i+1$ бусами
$\Rightarrow x_k = \sum_{i=1}^{k-1} \Delta x_i$ - координата k бусины
Исходя из наших определений Δx_i - реализация случайной величины ξ

Пользователь может отслеживать следующие выборочные характеристики, где выборкой является $\{\Delta x_i\}_{i=1}^{N}$:

1. Среднее
2. Дисперсия
3. Энтропия
4. Периодичность
5. Эмпирическая плотность распределения

В соответствующем списке можно выбрать вид распределения ξ, были реализованы следующие:

1. Нормальное
2. Экспоненциальное
3. Гамма
4. Парето
5. Вейбулла
6. Бернулли
7. Пуассона

Формулы

Для вычисления выборочных значений были использованы следующие формулы:
$\mu = \frac{1}{N}\sum_{i=1}^{N}\Delta x_i$ - выборочное среднее
$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(\Delta x_i - \mu)^2$ - выборочная дисперсия
$T = 1 - \frac{\sigma^2}{\mu^2}$ - периодичность

Определим эмпирическую плотность распределения следующим образом:
$p(\Delta x, x, h) = \frac{1}{N}\sum_{i=1}^{N} I(x < \Delta x_i < x + h)$, где $I(x)$ - индикаторная функция, которая равняется 1 $\iff$ выполнено x
$S = -\int_R p\,(t, x, h)\, log\,(p\,(t, x, h))\, dt$ - выборочная энтропия

Распределения

Главная цель программы - показать как от распределения ξ зависит вид структуры, поэтому самой важным является выбор распределения и его параметра. Ниже описаны основные распределения и их плотности распределения.

Нормальное распределение

Нормальное распределение задается двумя параметрами
$m \in R$ - коэффициент сдвига
$\sigma > 0$ - коэффициент масштаба
$x \in R$
$\xi \sim N(m, \sigma) \iff p(\xi = x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}}$
$E_p(\xi) = m$
$D_p(\xi) = \sigma^2$

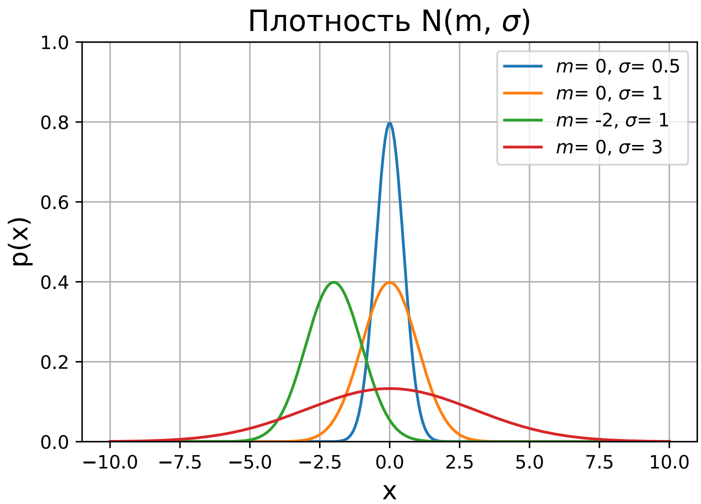

Рис. 1: Плотность нормального распределения

Гамма распределение

Гамма распределение задается параметрами
$\alpha > 0$ - параметр формы
$\beta > 0$ - коэффициент масштаба
$x \geq 0$
$\xi \sim \Gamma(\alpha, \beta) \equiv \Gamma(k, \theta) \iff p(\xi = x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}$
$E_p(\xi) = \beta\alpha$
$D_p(\xi) = \beta\alpha^2$

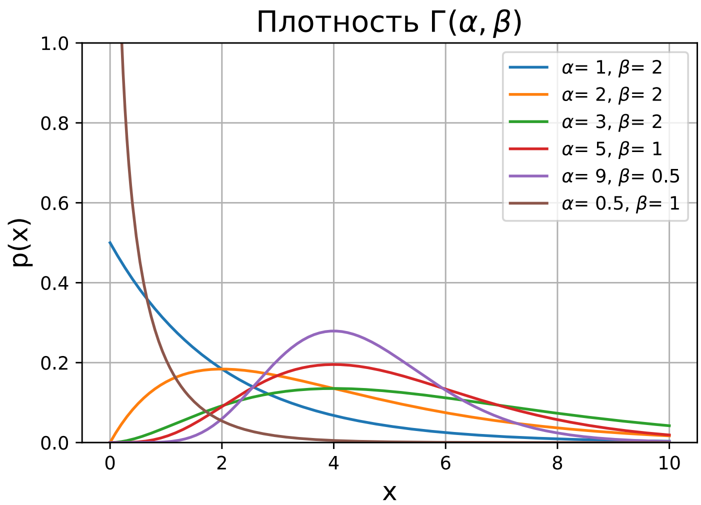

Рис. 2: Плотность гамма распределения

Распределение Парето

Распределение Парето задается параметрами
$x_m > 0$ - коэффициент масштаба
$\alpha > 0$ - параметр формы
$x \geq x_m$
$\xi \sim Pareto(x_m, \alpha) \iff p(\xi = x) = \frac{\alpha x_m^\alpha}{x^{\alpha+1}}$
$E_p(\xi) = \frac{\alpha x_m}{\alpha - 1},\ \alpha > 1$
$D_p(\xi) = \left(\frac{x_m}{\alpha-1}\right)^2 \frac{\alpha}{\alpha-2},\ \alpha > 2$

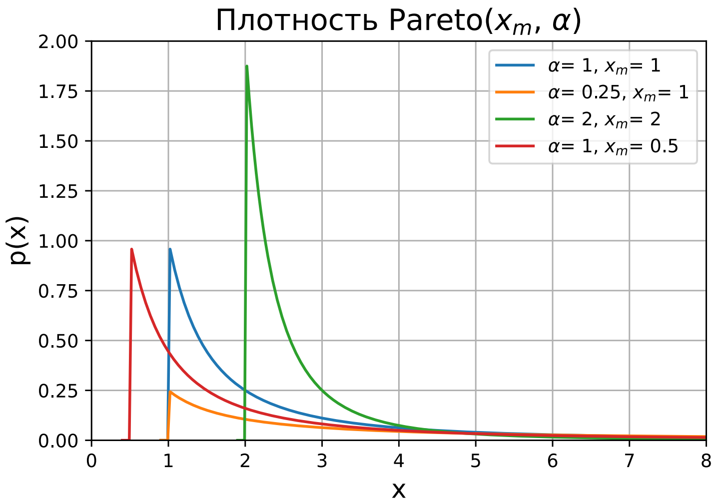

Рис. 3: Плотность распределения Парето

Распределение Вейбулла

Распределение Парето задается параметрами
$\lambda > 0$ - коэффициент масштаба
$k > 0$ - параметр формы
$x \geq 0$
$\xi \sim W(\lambda, k) \iff p(\xi = x) = \frac{k}{\lambda}(\frac{x}{\lambda})^{k-1}e^{-(x/\lambda)^k}$
$E_p(\xi) = \lambda\Gamma(1 + 1/k)$
$D_p(\xi) = \lambda^2\Gamma(1 + 2/k) - (E_p\,(\xi))^2$

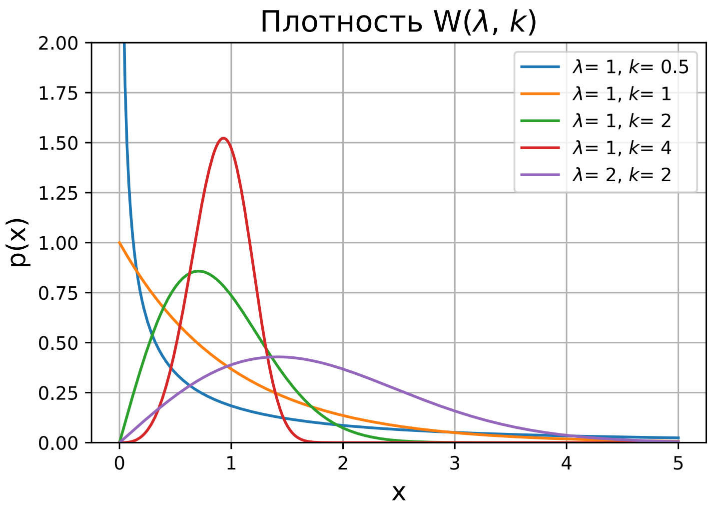

Рис. 4: Плотность распределения Вейбулла

Распределение Бернулли

Распределение Бернулли задается параметром p
$p \in [0, 1]$
$x \in \{0, 1\}$
$\xi \sim Br(p) \iff p(\xi = 1) = p; \quad p(\xi = 0) = 1 - p$
$E_p(\xi) = p$
$D_p(\xi) = p(1 - p)$

Распределение Пуассона

Распределение Пуассона задается параметром
$\lambda > 0$
$k \in \{0, 1, 2, \dots\}$
$\xi \sim Poiss(\lambda) \iff p(\xi = k) = \frac{e^{-\lambda}\lambda^k}{k!}$
$E_p(\xi) = \lambda$
$D_p(\xi) = \lambda$

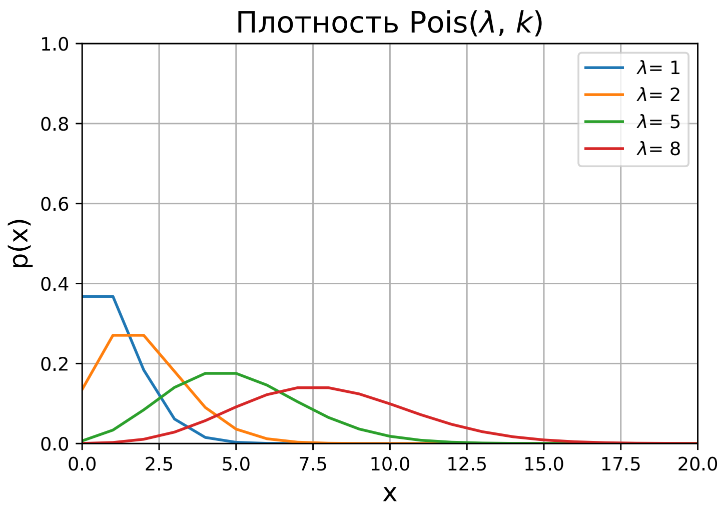

Рис. 5: Плотность распределения Пуассона

Замечание

Для некоторых распределений в программе введен коэффициент масштаба $k > 0$ идейно он отображает $\xi \to k\xi$
$\Rightarrow$
$E_p(k\xi) = kE_p(\xi)$
$D_p(k\xi) = k^2 D_p(\xi)$

Субпуассоновские и суперпуассоновские процессы

Выделим 2 вида процесса, на основе их периодичности:

1. Процесс, для которого $T \in (0, 1]$ будем называть субпуассоновским
2. Процесс, для которого $T < 0$ будем называть суперпуассоновский

В первом случае мы видим отрицательную корреляцию между импульсами. Появление одного уменьшает вероятность появления другого в ближайшем будущем, а при суперпуассоновской статистике корреляции положительные. В программе соответствующие процессы получаются, например, с помощью распределений $Pareto(30, 100)$ и $\Gamma(0.01, 0.01)$.
Наглядный пример двух статистик дают лисицы и лисички. Первые тщательно охраняют свою территорию определенного размера. Эта территория обозначена метками, все другие лисицы с нее прогоняются. Поэтому, встретив одну лису маловероятно встретить другую в ближайшее время. А лисички растут колониями, принадлежащими к одной микоризе, поэтому, найдя одну, есть большая вероятность найти следующую. Это приводит к суперпуассоновской статистике.