

Случайная величина

Программное обеспечение

- Julia ≥ 1.6
- Jupiter notebook
- Anaconda (Python)
- VS Code

Julia

Ссылка: <https://julialang.org/>

Julia — высокоуровневый высокопроизводительный свободный язык программирования с динамической типизацией, созданный для математических вычислений. Эффективен также и для написания программ общего назначения. Синтаксис языка схож с синтаксисом других математических языков (например, MATLAB и Octave), однако имеет некоторые существенные отличия. Julia написан на Си, C++ и Scheme. Имеет встроенную поддержку многопоточности и распределённых вычислений, реализованные в том числе в стандартных конструкциях.

Jupyter

Ссылка: <https://jupyter.org/>

Jupyter — интерактивный блокнот, первоначально являвшийся веб-реализацией и развитием IPython, ставший самостоятельным проектом, ориентированным на работу со множеством сред выполнения — не только Python, но и R, Julia, Scala и ряда других.

Anaconda

Ссылка: <https://www.anaconda.com/>

Anaconda — дистрибутив языков программирования Python и R, включающий набор популярных свободных библиотек, объединённых проблематиками науки о данных и машинного обучения.

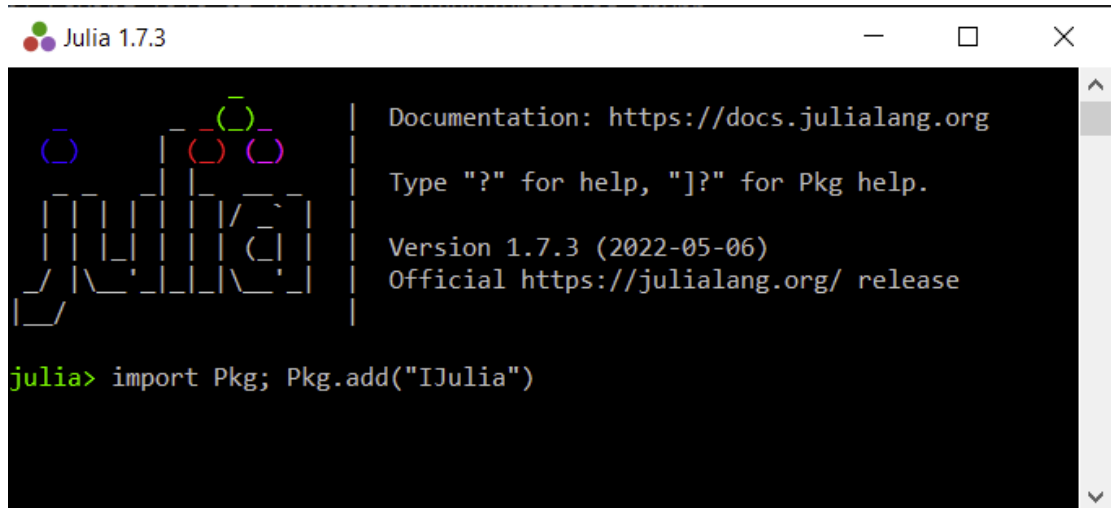
VS Code (Visual Studio Code)

Ссылка: <https://code.visualstudio.com/>

Visual Studio Code — редактор исходного кода, разработанный Microsoft для Windows, Linux и macOS.

Установка Julia

- Загрузить и установить подходящую версию Julia (<https://julialang.org/downloads/>)
- Если не установлен VS Code (или другой редактор с поддержкой Julia), то загрузить с официального сайта и установить
- Для использования Julia совместно с Jupyter необходимо установить IJulia, для этого необходимо запустить REPL и выполнить `import Pkg; Pkg.add("IJulia")`



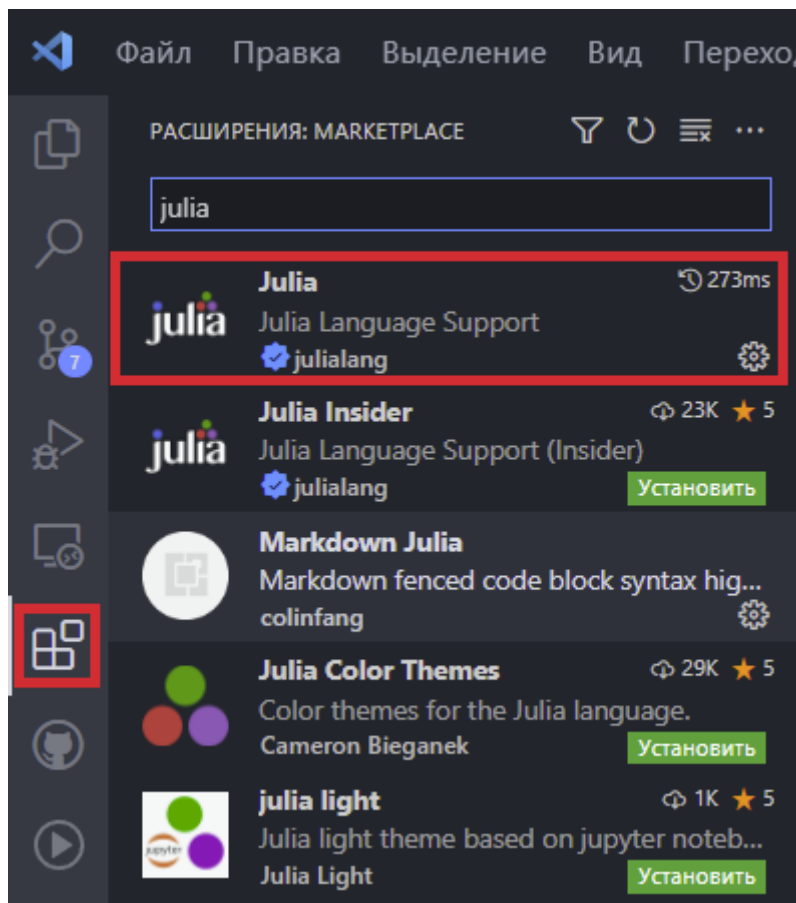
The screenshot shows a terminal window titled "Julia 1.7.3". On the left, there is a colorful ASCII art logo of the letter 'J'. On the right, the following text is displayed:

```
Documentation: https://docs.julialang.org
Type "?" for help, "]"? for Pkg help.
Version 1.7.3 (2022-05-06)
Official https://julialang.org/ release
```

At the bottom of the terminal, the command `julia> import Pkg; Pkg.add("IJulia")` has been entered.

In [1]: `#import Pkg; Pkg.add("IJulia");`

- Для использования в VS Code необходимо в разделе "Расширения" в поисковой строке набрать "julia" и установить расширение. После чего можно будет запускать файлы Julia - *.jl или запускать участки выделенного кода с помощью быстрого сочетания "Ctrl+Enter"



Основные пакеты Julia, которые могут потребоваться во время курса:

- LinearAlgebra
- CSV
- DataFrames
- ForwardDiff
- QuadGK
- HCubature
- DifferentialEquations
- Optim
- Plots

Устанавливаются пакеты по команде: `import Pkg; Pkg.add("CSV")`, установить все: `import Pkg; Pkg.add(["CSV", "DataFrames", "ForwardDiff", "QuadGK", "HCubature", "DifferentialEquations", "Optim", "Plots"])`

Вопросы

Некоторая часть курса будет посвящена практическому применению статистического анализа. Практическое применение в современности невозможно без использования ЭВМ и минимальных навыков программирования. Таким образом, в программу включены базовые основы использования языка программирования Julia.

Почему Julia?

- Достаточное количество пакетов для работы с данными (CSV, DataFrames)
- Простой синтаксис (Fortran-like) и хорошая документация (<https://docs.julialang.org/en/v1/>)
- Множественная диспетчеризация
- Динамическая типизация
- JIT-компиляция (динамическая компиляция)
- Поддержка Юникода, включающая UTF-8

Литература

- **Julia**

- Ссылка: <https://julialang.org/>
- Шеррингтон Малькольм - Осваиваем язык Julia.
- Вадим Никитин - Julia. Язык программирования. Быстрый старт.

- **Документация к основным пакетам Julia**

- *Математика и анализ*
 - Roots.jl (нахождение корней) <https://juliamath.github.io/Roots.jl/stable/>
 - Optim.jl (Поиск минимума/максимума) <https://julianlsolvers.github.io/Optim.jl/stable/#user/minimization/>
 - ForwardDiff.jl (Дифференцирование) <https://juliadiff.org/ForwardDiff.jl/stable/>
 - QuadGK.jl (Численное интегрирование) <https://juliamath.github.io/QuadGK.jl/stable/>
 - DifferentialEquations.jl (Численное решение дифференциальных уравнений) <https://diffeq.sciml.ai/stable/>
 - ExperimentalDesign.jl (Оптимизация дизайна) <https://github.com/phrb/ExperimentalDesign.jl>
- *Статистика*
 - Distributions.jl (Основной пакет для работы с распределениями) <https://juliastats.org/Distributions.jl/stable/>
 - Statistics - Базовый пакет с основными статистическими функциями (не требует установки, но надо подключать `using Statistics`)

- HypothesisTests.jl (основные статистические критерии) <https://juliastats.org/HypothesisTests.jl/stable/>
- GLM.jl (Общая/обобщенная линейная модель) <https://juliastats.org/GLM.jl/stable/>

- Другое

- Plots.jl (Графики) <https://docs.juliaplots.org>

- **Математика**

- Математика. Базовый курс / Б.Ш. Гулиян, Р.Я. Хамидуллин. Москва : Синергия, 2013. - 712 с. - ISBN 978-5-4257-0109-1.
- Д.Т. Письменный - Конспект лекций по высшей математике
- [The Matrix Cookbook](#)
- Линейная алгебра и ее применения - Стренг Г.
- Матричный анализ - Хорн Р., Джонсон Ч.

- **Планирование эксперимента**

- Ч. Хикс - Основные принципы планирования эксперимента.
- Н.И. Сидняев - Теория планирования эксперимента и анализ статистических данных.

- **Теория вероятностей**

- Фадеева Л. Н., Лебедев А. В., - Теория вероятностей и математическая статистика: учебное пособие. - 2-е изд., перераб. и доп. - М.: Эксмо, 2010. - 496 с.
- Гмурман В. Е. - Теория вероятностей и математическая статистика, учебное пособие.
- Гмурман В. Е. - Руководство к решению задач по теории вероятностей и математической статистике – решебник и задачник.

- **Дополнительная литература:**

- Майкл Кохендерфер, Тим Уилер - Алгоритмы оптимизации.
- Кохендерфер Майкл, Рэй Кайл, Уилер Тим - Алгоритмы принятия решений.
- Математическая статистика в медицине, В. А. Медик, М. С. Токмачев, 978-5-279-03195-5.
- Медико-биологическая статистика. Гланц. Пер. с англ. — М., Практика, 1998. — 459 с.

- Уилл Курт - Байесовская статистика: Star Wars, LEGO, резиновые уточки и многое другое.
- Занимательная статистика. Манга. Син Такахаси, 2009, 224с, ISBN: 978-5-97060-179-2.
- Ю.Д. Григорьев - Методы оптимального планирования эксперимента.

Определения

Случайная величина

Под **случайной величиной** понимается переменная, которая в результате испытания в зависимости от случая принимает одно из возможного множества своих значений.

Случайная величина называется **дискретной (прерывной)**, если множество значений конечное, или бесконечное, но счетное.

Под **непрерывной случайной величиной** будем понимать величину, бесконечное несчетное множество значений которой есть некоторый интервал (конечный или бесконечный) числовой оси.

Пространство элементарных исходов

Пространством элементарных исходов («омега», Ω) называется множество, содержащее все возможные результаты данного случайного эксперимента, из которых в эксперименте происходит ровно один. Элементы этого множества называют элементарными исходами (ω), т.е. $\omega \in \Omega$.

Событие

Событиями мы будем называть подмножества множества Ω . Говорят, что в результате эксперимента произошло событие $A \subseteq \Omega$, если в эксперименте произошел один из элементарных исходов, входящих в множество A .

Т.е. A не обязательно конкретный элементарный исход. К примеру, возьмем событие - "Выпадение четного значения кубика", тогда $A = \{2, 4, 6\}$, которое состоит из интересующих нас элементарных исходов $\omega_1 = 2; \omega_2 = 4; \omega_3 = 6$.

Сигма-алгебра событий

Множество событий, которое мы рассматриваем в эксперименте, называется сигма-алгеброй событий и обозначается \mathcal{F} .

Т.е.:

- $\Omega \in \mathcal{F}$
- Если $A \in \mathcal{F}$, то $\overline{A} \in \mathcal{F}$ (т.е. σ -алгебра событий включает событие и его

дополнение)

- Если $A_1, A_2, \dots \in \mathcal{F}$, то $A_1 \cup A_2 \cup \dots \in \mathcal{F}$

Вероятностное пространство

Вероятностью или вероятностной мерой называется функция \mathbb{P} , действующая из \mathcal{F} в множество \mathbb{R} и обладающая следующими свойствами:

- Вероятность всего пространства элементарных исходов Ω равна единице, т.е.
$$P(\Omega) = 1$$
- Для любого события $A \in \mathcal{F}$ вероятность неотрицательна и не больше 1, т.е.
$$0 \leq \mathbb{P}(A) \leq 1$$
- Для любого конечного или счетного набора попарно **непересекающихся** событий A_1, A_2, \dots имеет место равенство: $\mathbb{P}(\bigcup_i A_i) = \sum_i \mathbb{P}(A_i)$

Тройка объектов $(\Omega, \mathcal{F}, \mathbb{P})$ - называется **вероятностное пространство**.

Эксперимент удовлетворяет **классическому определению вероятности**, если пространство элементарных исходов состоит из конечного числа равновозможных исходов.

В таком случае вероятность любого события вычисляется по формуле: $P(A) = \frac{|A|}{|\Omega|}$

Это классическое определение вероятности.

- где $|A|$ - означает мощность множества A , т.е. количество элементов.

Дискретное пространство элементарных исходов (дискретные случайные величины)

Случайной величиной X называется функция, которая каждому элементарному исходу $\omega \in \Omega$ ставит в соответствие действительное число, то есть $X : \Omega \rightarrow \mathbb{R}$

X	1	2	3	4	5	6
\mathbb{P}	1/6	1/6	1/6	1/6	1/6	1/6

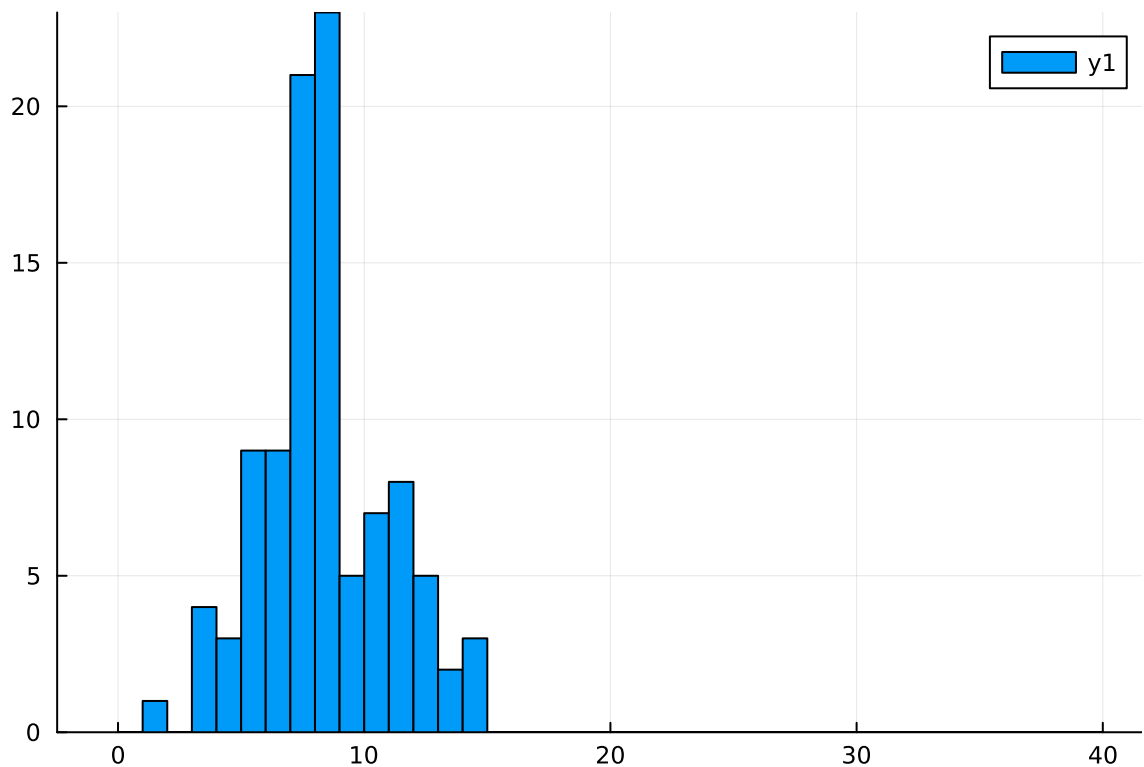
Таблица такого вида называется распределением дискретной случайной величины или рядом распределения дискретной случайной величины. **Функция распределения** случайной величины в этом случае носит ступенчатый характер и может быть записана в виде $F(x) = \mathbb{P}(X \leq x)$.

Случайная величина X имеет дискретное распределение, если множество ее значений не более чем счетно (конечно или счётно). При этом сумма вероятностей элементарных исходов $p_1 + p_2 + \dots + p_n = 1$

In [21]: `using Distributions, Plots`

```
# Пример
# Зададим случайную величину из биномиального распределения
X = Binomial(40, 0.2)
# Генерируем 100 случайных реализация
Xvec = rand(X, 100)
# Построим гистограмму
histogram(Xvec, bins = 0:1:40)
#1*1/6 + 2*1/6 + 3*1/6+ 4*1/6+ 5*1/6+ 6*1/6
```

Out[21]:



In [28]: `#Xvec`

Математическое ожидание дискретной случайной величины X называется число $\mathbb{E}X$:

$$\mathbb{E}X = \sum_i a_i p_i \text{ (в случае если } \sum_i |a_i| p_i < \infty \text{)}$$

Дисперсия случайной величины: $Var(X) = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}[X^2] - \mathbb{E}(X)^2$

In [23]: `1/6 + 2/6 + 3/6 + 4/6 + 5/6 + 6/6`
`((1 - 3.5)^2 + (2 - 3.5)^2 + (3 - 3.5)^2 + (4 - 3.5)^2 + (5 - 3.5)^2 + (6 - 3.5)^2)`

Out[23]: 2.9166666666666665

Случайные величины имеющие плотность (непрерывные случайные величины)

Непрерывной случайной величиной называют случайную величину, которая в результате испытания принимает все значения из некоторого числового подмножества $\subseteq \mathbb{R}$.

К примеру, для нормального распределения $\Omega = \{-\infty; \infty\}$; элементарные события в этом случае - действительные числа, при этом вероятность каждого элементарного события равна нулю (отметим это).

Определение:

Случайная величина X имеет абсолютно непрерывное распределение, если существует неотрицательная функция $f: \mathbb{R} \rightarrow \mathbb{R}$ (функция плотности вероятности) такая, что имеет место равенство:

$$P(X \in \mathbb{A}) = \int_{\mathbb{A}} f(x) dx$$

где \mathbb{A} - не более чем счетное объединение интервалов или отрезков, не обязательно пересекающихся, при чем некоторые могут быть бесконечными (Борелевские множества).

При этом:

- $f(x) \geq 0 \forall x \in \mathbb{R}$
- $\int_{-\infty}^{\infty} f(x) dx = 1$

Функцию f называют плотностью распределения случайной величины X (функция плотности вероятности, PDF)- это первая производная от функции распределения. Можно сказать, что эта функция показывает распределение реализации значений случайной величины.

Просто показать, что исходя из определения, вероятность каждого элементарного события равна интегралу с "нулевой шириной" и соответственно равна нулю.

Функцией распределения случайной величины или **интегральной функцией** называется функция, которая определяет вероятность, что значение случайной величины X меньше или равно граничному значению x , то есть $F(x) = \mathbb{P}(X \leq x)$.

Или

$$F(a) = \mathbb{P}(X \leq a) = \int_{-\infty}^a f(x) dx$$

В общем виде:

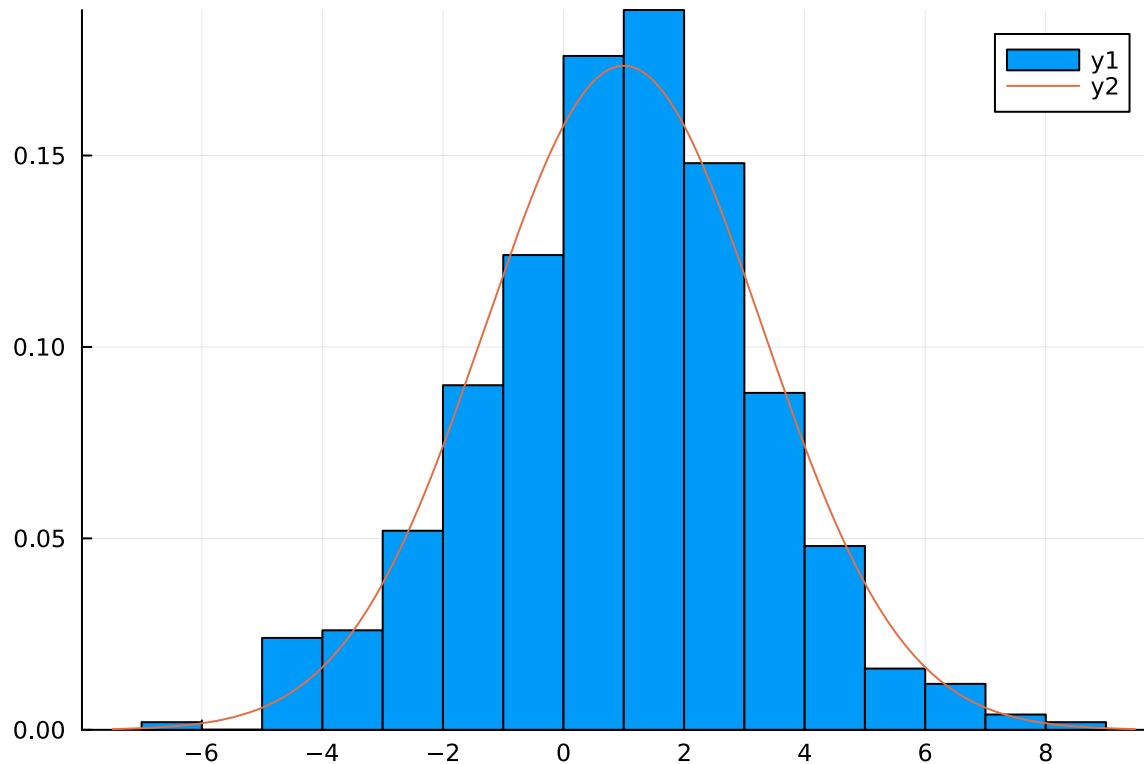
$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x) dx$$

Где f - функция плотности вероятности.

```
In [24]: # Пример
# Зададим случайную величину из нормального распределения
X = Normal(1.0, 2.3)
# Генерируем 500 случайных реализация
Xvec = rand(X, 500)
# Построим гистограмму
histogram(Xvec, normalize=:pdf)
```

```
# Построим график плотности вероятности (PDF) для распределения X
plot!(x->pdf(X, x))
```

Out[24]:



In [29]: `#Xvec`

Математическое ожидание непрерывной случайной величины:

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

где $f(x)$ - функция плотности вероятности случайной величины X .

Дисперсия определяется также, как и в случае дискретной величины:

$$\text{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2 = \mathbb{E}[X^2] - \mathbb{E}(X)^2$$

```
In [7]: using QuadGK

X = Normal(4, 10)

xfx = x -> x*pdf(X, x)

quadgk(xfx, -Inf, Inf)
```

Out[7]: (3.9999999999999987, 3.3535454297896486e-8)

Свойства математического ожидания и дисперсии

Справедливы как для дискретных, так и для непрерывных величин.

Математическое ожидание (МО)

Обозначается как - \mathbb{E} (Expectation), реже \mathbb{M}

- МО константы равно константе
- Постоянный множитель можно вынести за скобки: $\mathbb{E}(cX) = c\mathbb{E}X$
- МО суммы случайных величин равно сумме их МО: $\mathbb{E}(X + Y) = \mathbb{E}X + \mathbb{E}Y$
- МО произведения **независимых** величин равно произведению их МО:
$$\mathbb{E}(X \cdot Y) = \mathbb{E}X \cdot \mathbb{E}Y$$
- Если $X \geq 0$, то $\mathbb{E}X \geq 0$
- Для любых МО произведения величин равно:
$$\mathbb{E}(X \cdot Y) = \mathbb{E}X \cdot \mathbb{E}Y + Cov(X, Y)$$

Дисперсия

Обозначается как - Var , иногда \mathbb{D}

- Умножение на константу: $Var(cX) = c^2Var(X)$
- Дисперсия всегда неотрицательна: $Var(X) \geq 0$
- Дисперсия не зависит от сдвига: $Var(c + X) = Var(X)$
- Если X и Y **независимы**: $Var(X + Y) = Var(X) + Var(Y)$
- Для любых X и Y : $Var(X + Y) = Var(X) + Var(Y) + 2Cov(X, Y)$

Где ковариационный момент: $Cov(X, Y) = \mathbb{E}[(X - \mathbb{E}X)(Y - \mathbb{E}Y)]$

```
In [31]: sqrt(0.2^2 + 0.1^2)
```

```
Out[31]: 0.223606797749979
```

Примеры

Примеры дискретных распределений:

- Вырожденное распределение (константа)
- Р. Бернули Случайная величина X имеет распределение Бернулли, если она принимает всего два значения: 1 и 0 с вероятностями p и $q \equiv 1 - p$ соответственно. Пример - бросок монеты.
- Р. Пуассона Случайная величина отражает число событий, произошедших за фиксированное время, при условии, что данные события происходят с некоторой фиксированной средней интенсивностью и независимо друг от друга. Функция вероятности: $\frac{e^{-\lambda} \lambda^k}{k!}$

- Биномиальное Р. Распределение количества «успехов» в последовательности из n независимых случайных экспериментов, таких, что вероятность «успеха» в каждом из них постоянна и равна p . Пример - серия бросков монеты. Функция вероятности: $\mathbb{P}(X_n = k) = C_n^k \cdot p^k \cdot q^{n-k}$, где $C_n^k = \frac{n!}{k!(n-k)!}$

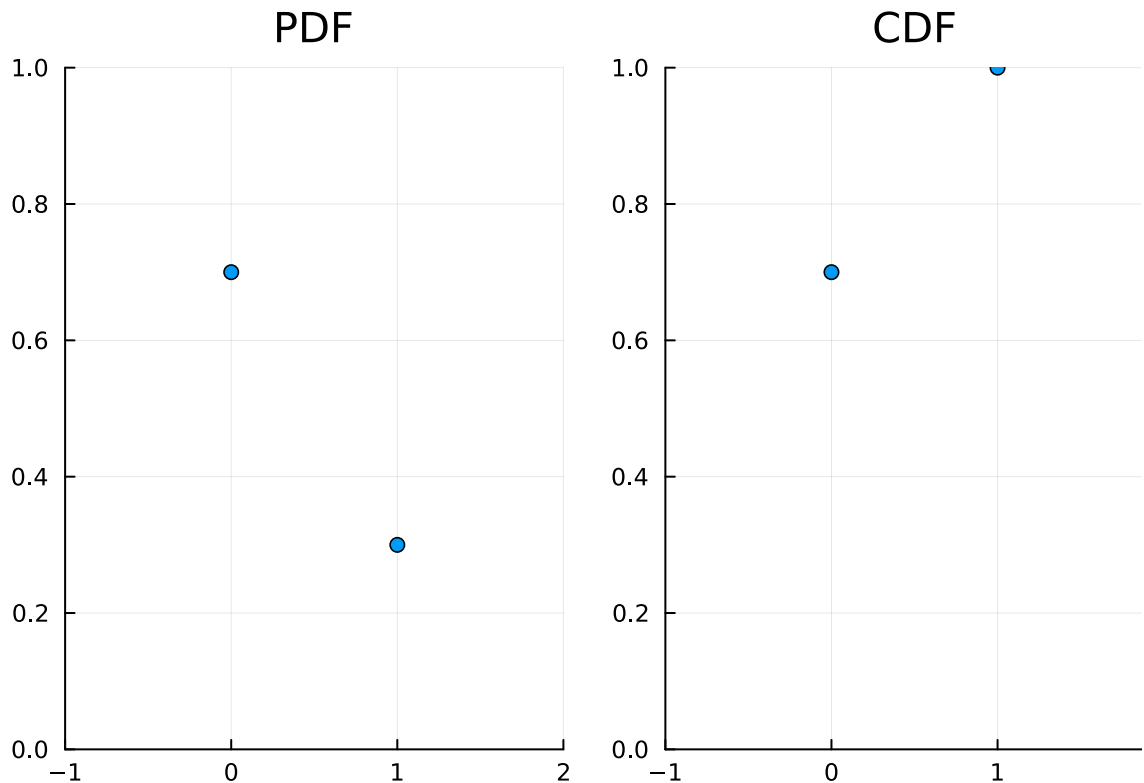
C_n^k - число сочетаний из n по k еще обозначается как $\binom{n}{k}$ - это количество всех (нестрогих) подмножеств (выборок) размера k в n - элементном множестве (биномиальный коэффициент).

\equiv - тождественно равно, т.е. равны при любых значениях входящих параметров.

In [25]: `X = Bernoulli(0.3)`

```
plot(
    scatter([1,0], pdf.(X,[1,0]), xlims = (-1,2), ylims = (0,1), legend = false)
    scatter([1,0], cdf.(X,[1,0]), xlims = (-1,2), ylims = (0,1), legend = false)
    layout = 2, title=["PDF" "CDF"])
```

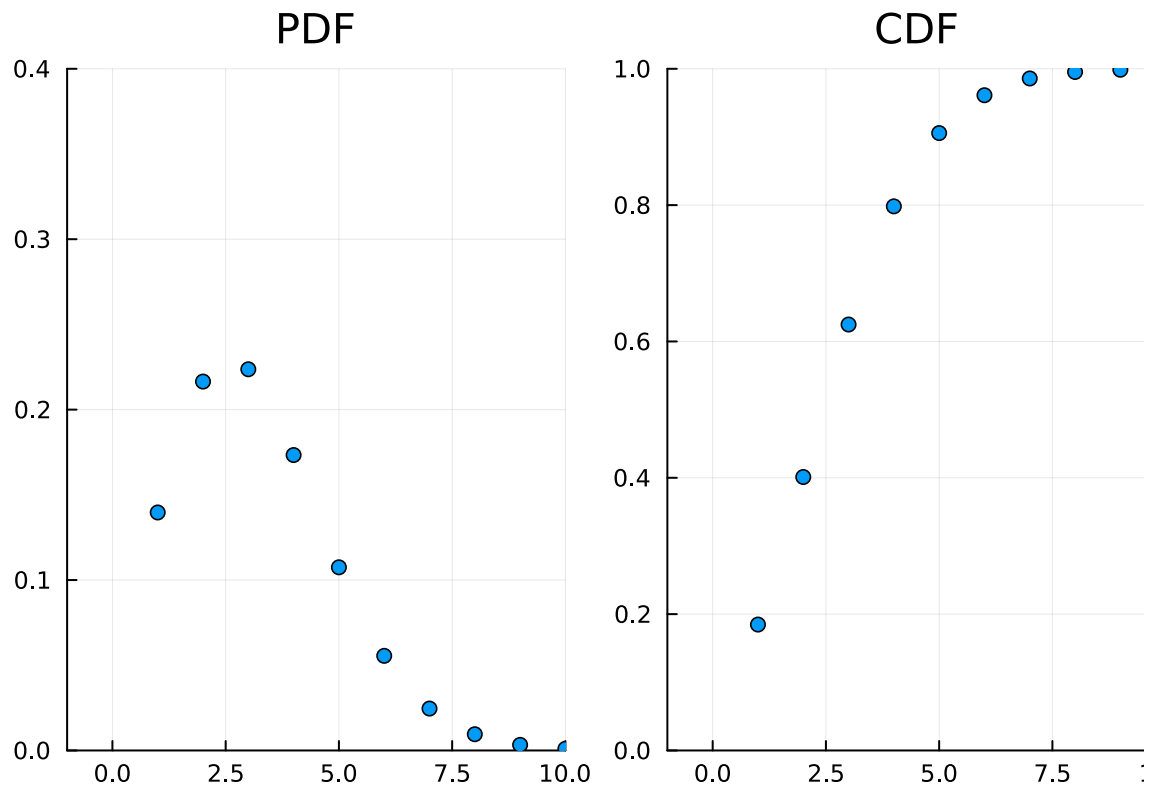
Out[25]:



In [26]: `X = Poisson(3.1)`

```
plot(
    scatter(1:10, pdf.(X, 1:10), xlims = (-1,10), ylims = (0,0.4), legend = false)
    scatter(1:10, cdf.(X, 1:10), xlims = (-1,10), ylims = (0,1), legend = false)
    layout = 2, title=["PDF" "CDF"])
```

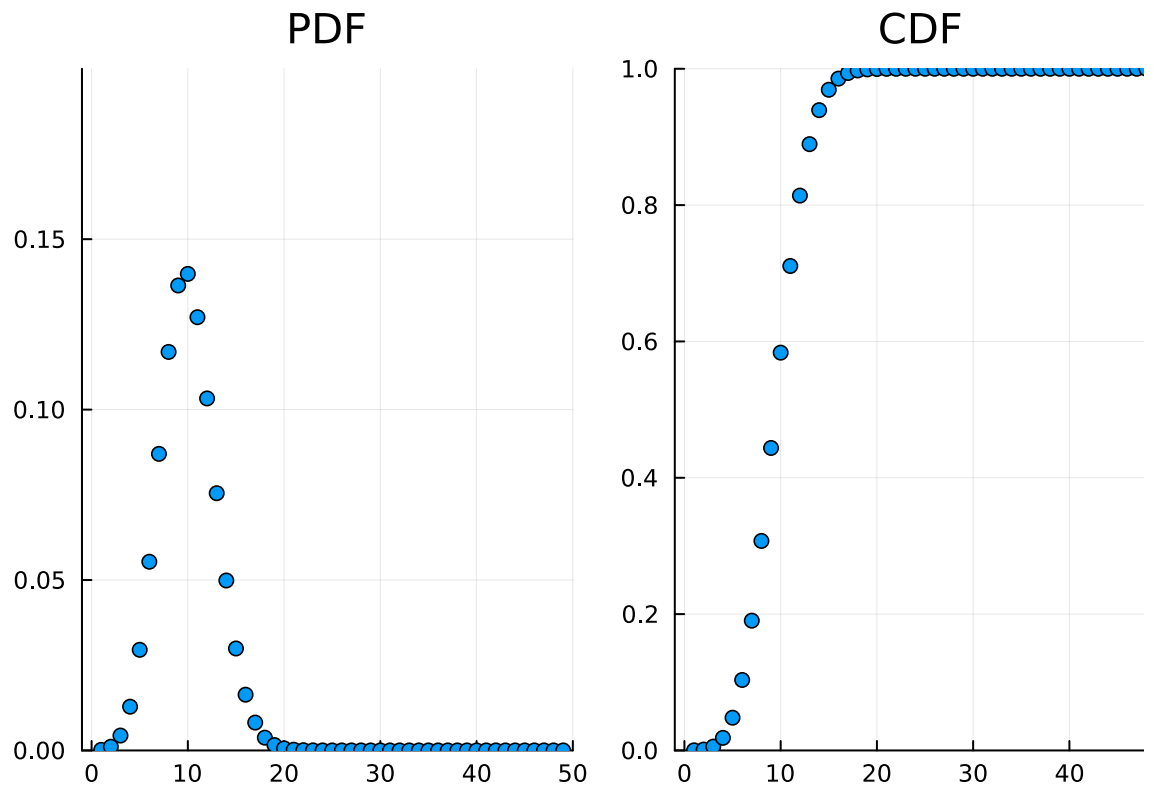
Out[26]:



In [27]: `X = Binomial(50, 0.2)`

```
plot(  
    scatter(1:50, pdf.(X, 1:50), xlims = (-1,50), ylims = (0,0.2), legend = fals  
    scatter(1:50, cdf.(X, 1:50), xlims = (-1,50), ylims = (0,1), legend = false)  
    layout = 2, title=["PDF" "CDF"])
```

Out[27]:

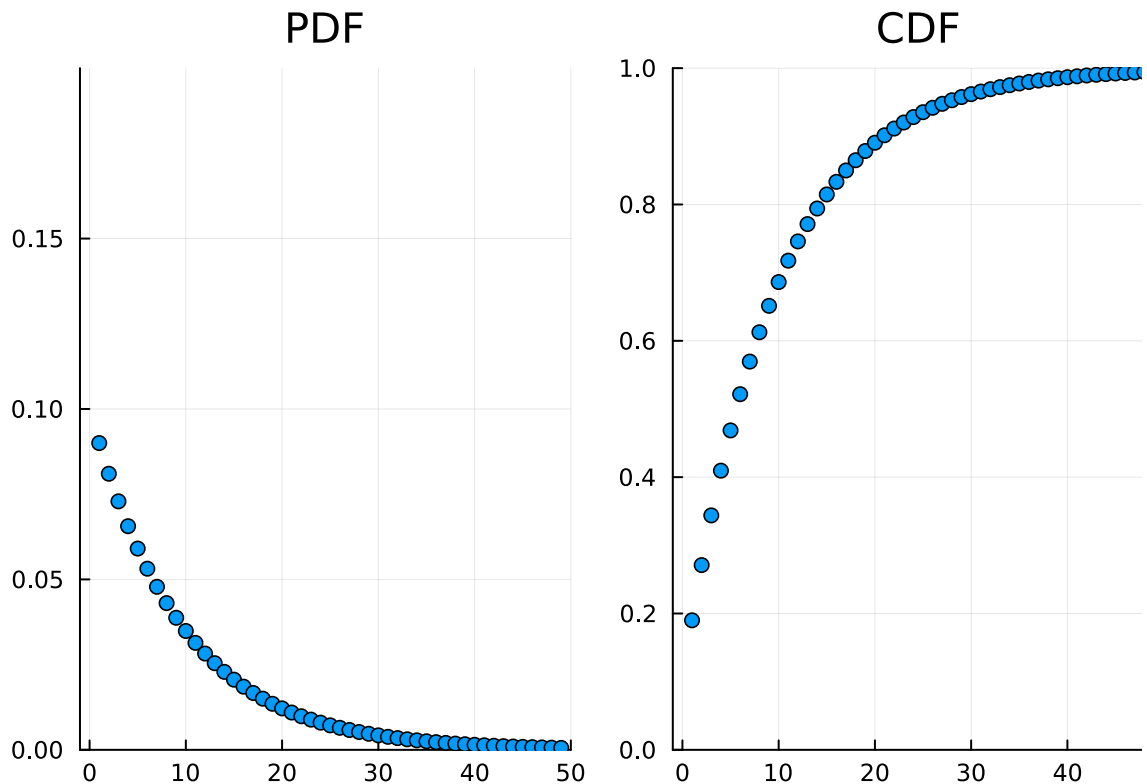


In [28]: `X = Geometric(0.1)`

```
plot(  
    scatter(1:50, pdf.(X, 1:50), xlims = (-1,50), ylims = (0,0.2), legend = fals
```

```
scatter(1:50, cdf.(X, 1:50), xlims = (-1,50), ylims = (0,1), legend = false)
layout = 2, title=["PDF" "CDF"])
```

Out[28]:



Задача Игральная кость брошена 6 раз. Найти вероятность того, что ровно 3 раза выпадет «шестерка».

$$\mathbb{P}(X_n = k) = C_n^k \cdot p^k \cdot q^{n-k} = \frac{6!}{3!(6-3)!} \cdot 1/6^3 \cdot (1 - 1/6)^{6-3} = (5/36)^3 \cdot 4 \cdot 5 = 0.05358$$

Игральная кость брошена 6 раз. Найти вероятность того, что ровно 3 раза выпадет 1 или 2.

```
In [29]: binomial(6, 3) * (1/6)^3 * (1-1/6)^3
#binomial(6, 3) * (2/6)^3 * (1-2/6)^3
```

Out[29]: 0.053583676268861465

Примеры непрерывных распределений:

- Нормальное распределение, функция плотности вероятности

$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \cdot (\frac{x-\mu}{\sigma})^2}$. Стандартным нормальным распределением называется нормальное распределение с параметрами $\mu = 0$ и $\sigma = 1$, функция плотности вероятности такого распределения часто обозначается ϕ , а функция распределения (cdf) - Φ (иногда f и F)

- Хи-квадрат распределение

Распределение χ^2 с ν степенями свободы — это распределение суммы квадратов ν независимых стандартных нормальных случайных величин. Имеет многочисленные приложения при статистических выводах, при использовании

критерия хи-квадрат, при оценке дисперсий, в дисперсионном анализе.

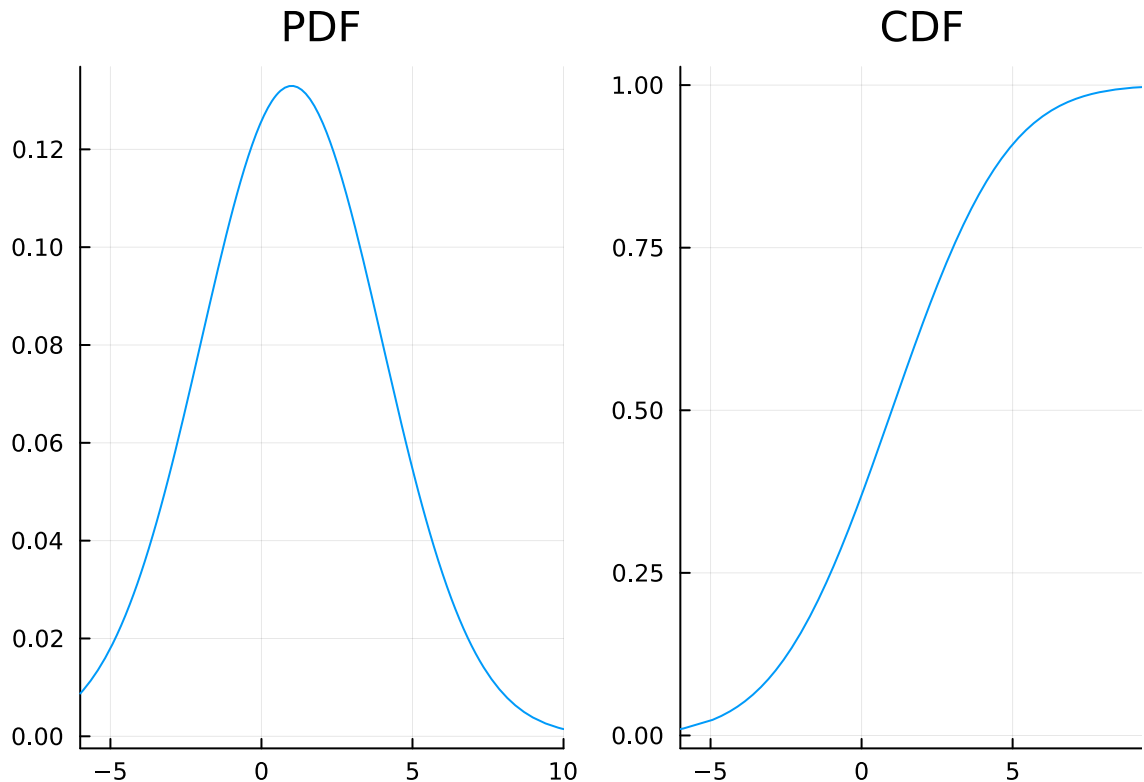
- Экспоненциальное (показательное) распределение

Абсолютно непрерывное распределение, моделирующее время между двумя последовательными свершениями одного и того же события. Функция плотности вероятности: $f(x) = \lambda \cdot e^{-\lambda x}$, функция распределения $F(x) = 1 - e^{-\lambda x}$. Параметр λ может быть интерпретирован как среднее число событий за единицу времени (количество минут между извержениями определенного гейзера, время между землетрясениями, время между обострениями заболевания, длительность безотказной работы).

```
In [30]: X = Normal(1, 3)
```

```
plot(
  plot(x-> pdf(X, x), legend = false, xlims = (-6,10)),
  plot(x-> cdf(X, x), legend = false, xlims = (-6,10)),
  layout = 2, title=["PDF" "CDF"])
```

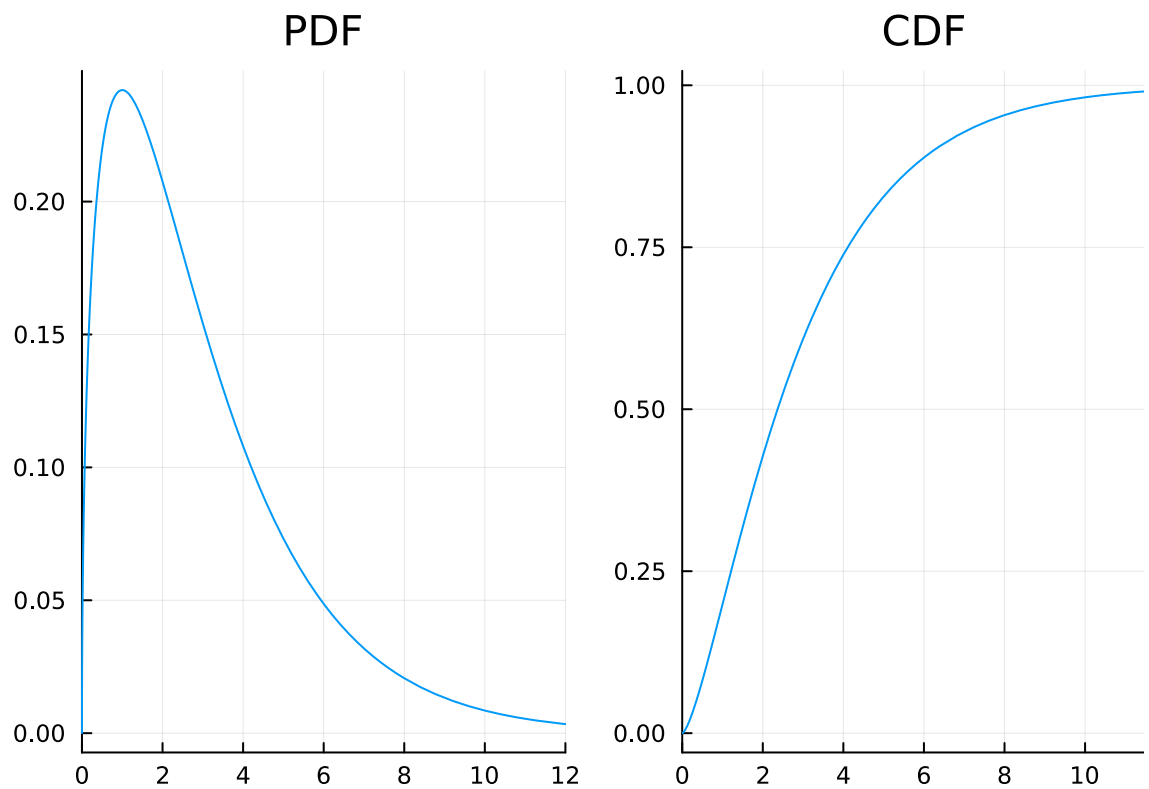
Out[30]:



```
In [31]: X = Chisq(3)
```

```
plot(
  plot(x-> pdf(X, x), legend = false, xlims = (0,12)),
  plot(x-> cdf(X, x), legend = false, xlims = (0,12)),
  layout = 2, title=["PDF" "CDF"])
```

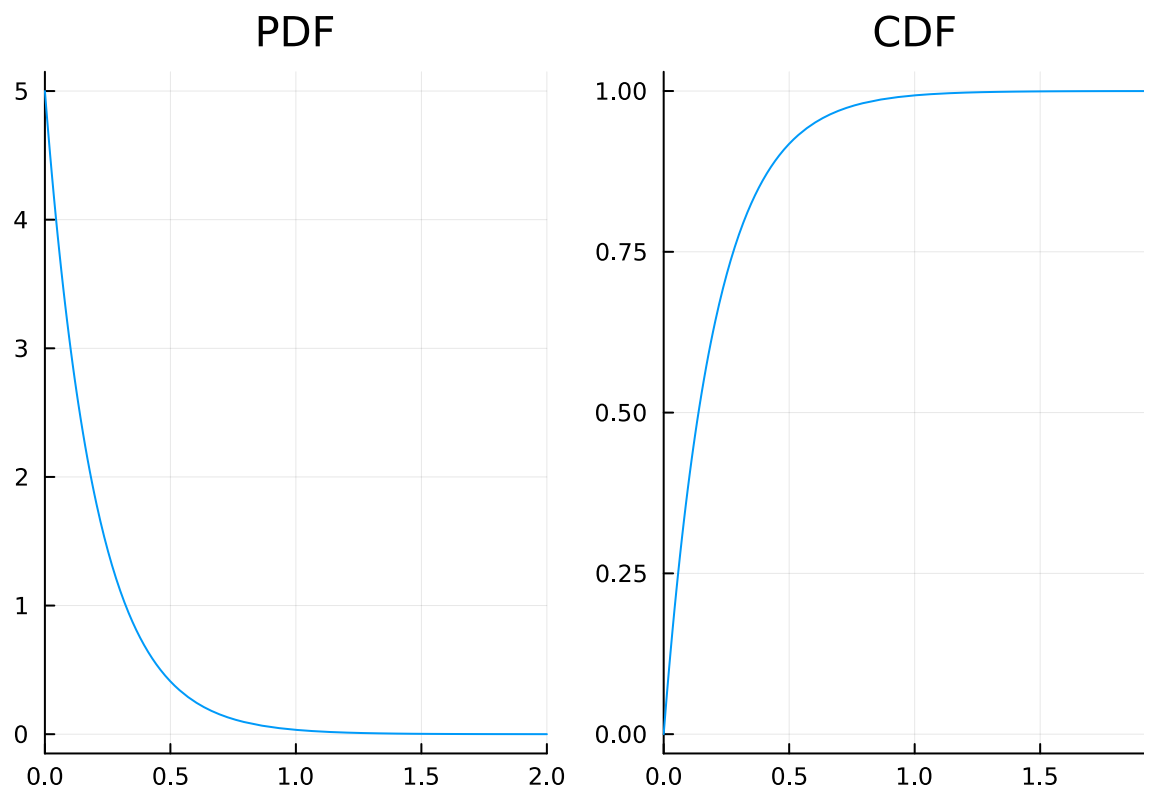
Out[31]:



```
In [32]: X = Exponential(0.2)
```

```
plot(  
    plot(x-> pdf(X, x), legend = false, xlims = (0,2)),  
    plot(x-> cdf(X, x), legend = false, xlims = (0,2)),  
    layout = 2, title=["PDF" "CDF"])
```

Out[32]:



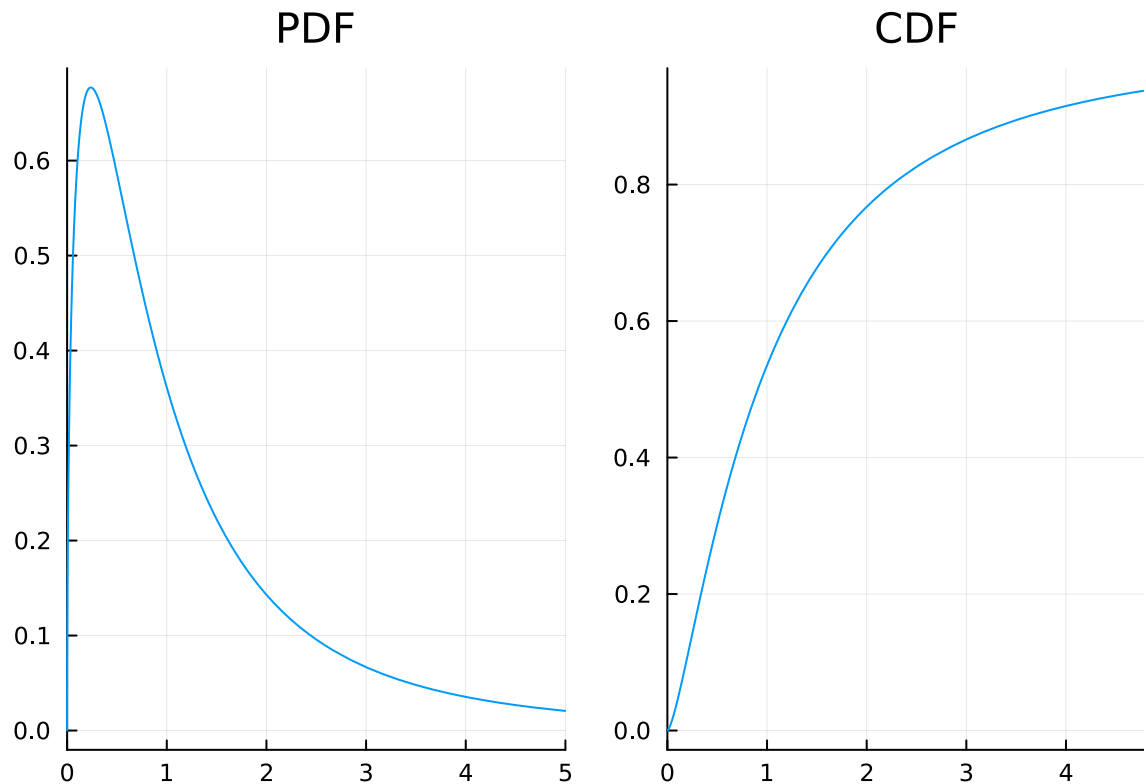
```
In [33]: X = FDist(3, 5)
```

```
plot(  
    plot(x-> pdf(X, x), legend = false, xlims = (0,5)),
```



```
plot(x-> cdf(X, x), legend = false, xlims = (0,5)),
layout = 2, title=["PDF" "CDF"])
```

Out[33]:



Задача

Мы знаем, что срок лужбы электрического чайника распределен согласно показательному распределению с параметром $\lambda = 0.5$. Всего мы продали 100 чайников. Сколько чайников вернут по гарантии через год эксплуатации?

Функция показательного распределения $f(x) = 1 - e^{-\lambda x}$

```
In [34]: λ = 0.5
x = 1
p = 1 - e^(-λ * x)
```

Out[34]: 0.3934693402873666

```
In [35]: round(100 * p)
```

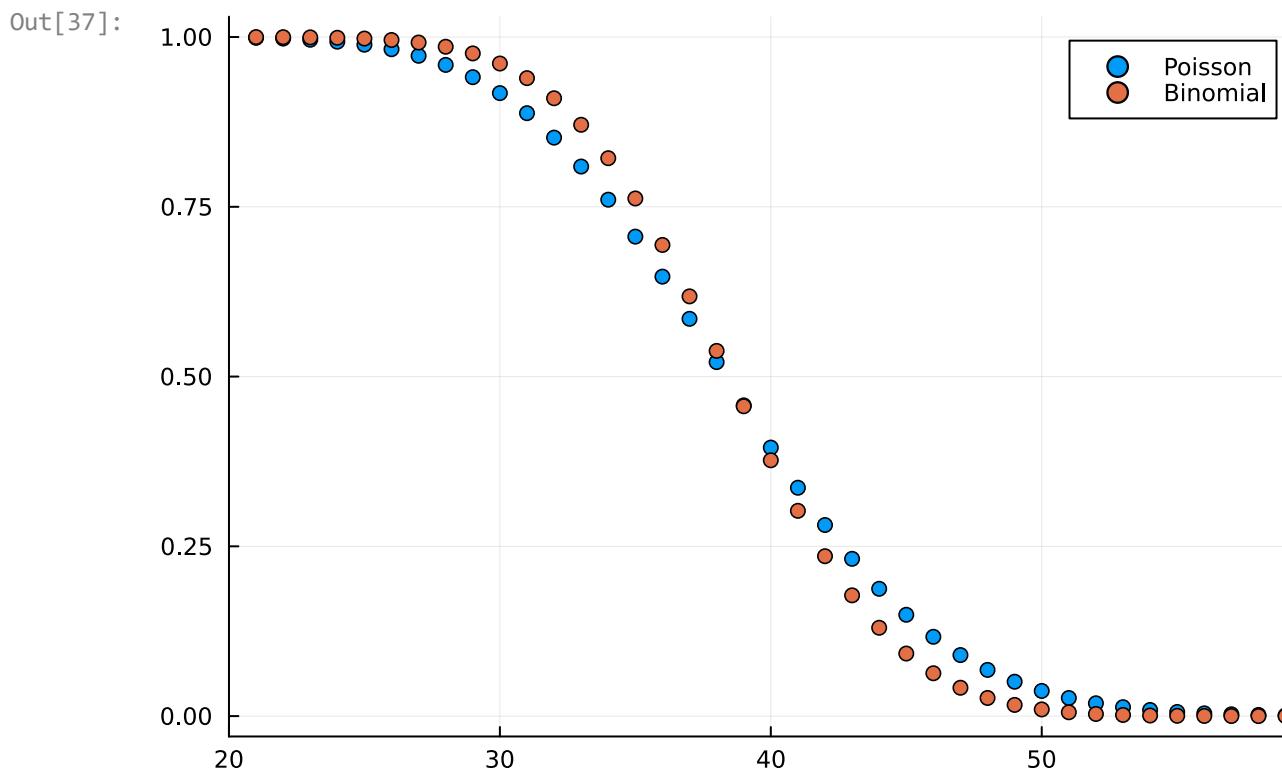
Out[35]: 39.0

```
In [36]: θ = 1/λ

cdf(Exponential(θ), x)
```

Out[36]: 0.3934693402873666

```
In [37]: scatter([1:100], ccdf.(Poisson(39), 1:100), xlims = (20, 60), label = "Poisson")
scatter!([1:100], ccdf.(Binomial(100, 0.39), 1:100), xlims = (20, 60), label = "
```



Аппроксимация биномиального распределения, распределением Пуассона справедлива для случаев с большим количеством наблюдений и низкой частотой возникновения события.

Выборки

Случайная выборка - выборка, обладающая тем свойством, что каждая единица, включенная в нее, могла быть с равными шансами, т.е. с равной вероятностью заменена любой другой.

Характеристики выборки

- Качество выборки (критерии выборки) – описание свойств генеральной совокупности из которой делается выборка.
- Объём выборки - количество случаев.
- Зависимость / независимость - описание контроля отбора зависимых наблюдений (к примеру - близнецы).
- Репрезентативность - насколько итоговая выборка соответствует генеральной совокупности.

Техника выборки

- Выборка с возвращением и без (повторная и бесповторная) - в случае если генеральная совокупность известна полностью, выборка может выполняться с возвращением и без. Даже если выборка достаточно большая, то независимо от этого необходимо определить будет ли контролироваться возврат.

- Рандомизированная - случайная выборка.
- Стратифицированная выборка - выборка выполняется из заранее определенных страт независимо друг от друга.
- Генздная (кластерная) выборка - выбирается не субъект, а группа целиком.
- Механическая выборка - отбор единиц в выборочную совокупность выполняется из генеральной, разбитую по нейтральному признаку на равные интервалы, из каждой такой группы в выборку отбирается лишь одна единица.

Какие еще бывают выборки?

- Неслучайные выборки.
 - Квотная выборка
 - Метод снежного кома
 - Стихийная выборка
 - Выборка типичных случаев

ДЗ

1. Аудитор обнаруживает финансовые нарушения у проверяемой фирмы с вероятностью 0.8. Найти вероятность того, что среди 5 фирм нарушителей будет выявлено больше половины.
 2. Дана случайная величина X , распределенная по нормальному закону с параметрами $\mu=20$, $\sigma=2.5$ см. Найти вероятность того, что реализация x будет в интервале $[21;24]$. В каких интервалах реализуется x с вероятностью 0.92?
- Математическая формулировка
3. Привести пример задачи с гипергеометрическим распределением.

In [1]: `using Distributions`

```
dist = Normal(20, 2.5)
p = cdf(dist, 24) - cdf(dist, 21)
println("Нижн. гр: ", quantile(dist, 0.04))
println("Нижн. гр: ", quantile(dist, 0.96))

#####
#Нижн. зр: 15.623284821869575
#Нижн. зр: 24.376715178130425
ub = 20 + 2.5*3
lb = 20 + 2.5
p = 0.04
while true
    pv = ccdf(dist, lb)
    if abs(p-pv) < 0.0001 break end
    lb += 0.0001
end
println("Bound: ", lb)

#####
```

```

#Нижн. гр: 15.623284821869575
#Нижн. гр: 24.376715178130425
ub = 20 + 2.5*3
lb = 20 + 2.5
p = 0.04
if ccdf(dist, lb) > p && ccdf(dist, ub) < p
    println("Condition OK")
else
    println("Condition FALSE")
end
val = NaN
while true
    val = (lb+ub)/2
    pv = ccdf(dist, val)
    if abs(p-pv) < 0.00000001 break end

    if pv < p
        ub = val
    else
        lb = val
    end
end
println("Bound: ", val)

```

Нижн. гр: 15.623284821869575
 Нижн. гр: 24.376715178130425
 Bound: 24.37381999992906
 Condition OK
 Bound: 24.376715198159218

Обозначения:

- \emptyset - пустое множество
- \in - принадлежит
- \subseteq - подмножество
- \cap - пересечение
- \cup - объединение
- \setminus - разность множеств
- \oplus - симметричная разность
- $|$ - которые
- \wedge - И
- \vee - ИЛИ
- \neg - НЕ
- \oplus - исключающее или
- \exists - существует
- \forall - для всех
- \otimes - Произведение Кронекера