Лекция 9.

Центральная предельная теорема и закон больших чисел.

Ранее мы рассмотрели дискретные случайные величины и непрерывные случайные величины. Здесь мы поговорим о таких величинах, распределение которых нам вообще неизвестно.

Слайд 3.

Неравенство Маркова

Мы уже умеем считать математическое ожидание случайной величины, ее разброс (дисперсию) вокруг него. Хочется понимать вероятность того, что насколько далеко отклонится случайная величина и с какой вероятностью.

Пусть X – случайная величина, принимающая неотрицательные значения, M(X) – математическое ожидание этой случайной величины. Тогда для любого a›0 справедливо **неравенство Маркова**

$$P(X \geq a) \leq \frac{M(X)}{a}$$

или

$$P(X < a) > 1 - \frac{M(X)}{a}$$

То есть для случайной величины неизвестного распределения, можно оценить вероятность того, что она принимает определенные значения.

Слайд 4.

Неравенство Чебышёва

Пусть мы знаем про случайную величину не только ее математическое ожидание (первый момент), но и дисперсию (второй центральный момент), и они конечны (!)это важно). Тогда можно воспользоваться следствием неравенства Маркова - неравенством Чебышёва

$$P(|X - M(X)| \geq a) \leq \frac{D(X)}{a^2}$$

$$P(|X - M(X)| < a) > 1 - \frac{D(X)}{a^2}, \quad a > 0.$$

Неравенство Чебышёва показывает, что случайная величина принимает значения близкие к среднему (математическому ожиданию) и дает оценку вероятности больших отклонений от него.

Видим, что чем больше величина отклонения а, тем меньше вероятность этого отклонения (a^2 в знаменателе).

Слайд 5.

Неравенство Чебышева пример.

Положим а=кб, где б – стандартное отклонение, тогда получим оценку вероятности того, что случайная величина отклонится по модулю от среднего больше чем на кб:

$$P(|X - M(X)| \geq k\sigma) \leq \frac{1}{k^2}.$$

Для значения к=2 вероятность отклонения меньше 25%, а для к=3 уже 11.12%.

Для случайной величины, распределенной по биномиальному закону с параметрами n, p неравенство Чебышева принимает вид:

$$P(|X - np| < a) > 1 - \frac{npq}{a^2}.$$

Слайд 6.

Сходимости случайных величин.

Как мы помним, случайная величина – (измеримая) функция из некоторого абстрактного множества Ω в множество действительных чисел.

Последовательность случайных величин $\{\xi_n\}_{n=1}^{\infty}$ есть тем самым последовательность функций, определенных на одном и том же пространстве элементарных исходов.

Будем изучать разные виды сходимостей!

Последовательность $\{\xi_n\}$ сходится *поточечно* к ξ, если для любого $\omega \in \Omega$ числовая

последовательность $\{\xi_n(\omega)\}$ сходится к $\xi(\omega)$.

Слайд 7.

Событие $A \in \mathscr{F}$ выполнено *почти наверное*, если $\mathsf{P}(A) = 1$

Последовательность $\{\xi_n\}$ сходится к ξ *почти наверное*, если событие $\{\xi_n \to \xi\}$

выполнено почти наверное. (т.е. $\mathsf{P}(\{\xi_n \to \xi\}) = 1$).

Обозначение: $\xi_n \stackrel{\text{п.н.}}{\to} \xi$.

Последовательность $\{\xi_n\}$ сходится к ξ *по вероятнсоти*, если выполнено:

$$\forall \varepsilon > 0: \ \mathsf{P}(|\xi_n - \xi| > \varepsilon) \to 0$$

Обозначение: $\xi_n \stackrel{\mathsf{P}}{\to} \xi$

Слайд 8.

Последовательность $\{\xi_n\}$ сходится к ξ в L^p, $(p > 0)$, если $\mathsf{E}|\xi_n - \xi|^p \to 0$ при $n \to \infty$.

Обозначение: $\xi_n \stackrel{L^p}{\to} \xi$.

Последовательность $\{\xi_n\}$ *слабо сходится* к ξ (или сходится *по распределению*),

если для любой ограниченной непрерывной функции $f : \mathbb{R} \to \mathbb{R}$ верно:

$$\mathsf{E}f(\xi_n) \to \mathsf{E}f(\xi)$$

Обозначение: $\xi_n \stackrel{d}{\to} \xi$

Слайд 9.

Последовательность $\{\xi_n\}$ сходится к ξ в L^p, $(p > 0)$, если $\mathsf{E}|\xi_n - \xi|^p \to 0$ при $n \to \infty$.

Обозначение: $\xi_n \stackrel{L^p}{\to} \xi$.

Последовательность $\{\xi_n\}$ *слабо сходится* к ξ (или сходится *по распределению*),

если для любой ограниченной непрерывной функции $f : \mathbb{R} \to \mathbb{R}$ верно:

$$\mathsf{E}f(\xi_n) \to \mathsf{E}f(\xi)$$

Обозначение: $\xi_n \stackrel{d}{\to} \xi$

Слайд 10.

Хотя и кажется, что определения сходимости эквивалентны, но это не так. Есть теоремы о взаимосвязи различных видов сходимости случайных величин.

Теорема 8.1. *(взаимосвязь между различными видами сходимости случайных величин)*

Пусть $\xi, \xi_1, \xi_2, \ldots$ — *случайные величины на* $(\Omega, \mathscr{F}, \mathsf{P})$. *Тогда верны следующие импликации:*

1. $\xi_n \xrightarrow{п.н.} \xi \Rightarrow \xi_n \xrightarrow{\mathsf{P}} \xi$

2. $\xi_n \xrightarrow{L^p} \xi \Rightarrow \xi_n \xrightarrow{\mathsf{P}} \xi$

3. $\xi_n \xrightarrow{\mathsf{P}} \xi \Rightarrow \xi_n \xrightarrow{d} \xi$

Слайд 11.

Введем понятие характеристической функции.

Определение 11.1. Пусть ξ — случайная величина на $(\Omega, \mathscr{F}, \mathsf{P})$. Тогда ее характеристической функцией называется функция

$$\varphi_\xi(t) = \mathsf{E}e^{it\xi} \text{ (прямое преобразование Фурье)}$$

Если ξ — случайный вектор, то

$$\varphi_\xi(t) = \mathsf{E}e^{i\langle t,\, \xi\rangle},\ t \in \mathbb{R}^n$$

где $\langle t, \xi\rangle$ — скалярное произведение.

$$\phi_\xi(t) = M[e^{it\xi}] = \sum_k e^{itx_k} \cdot p_k.$$

Характеристическая функция однозначно определяет распределение случайной величины. Характеристическая функция суммы независимых случайных величин равна произведению их характеристических функций.

Слайд 12.

Производящие функции

Для случайной величины ξ **производящая функция моментов** (сокращенно ПФМ) определяется следующим образом:

$$M_\xi(t) = M[e^{t\xi}].$$

$$M_\xi(t) = M[e^{t\xi}] = \sum_i e^{tx_i} \cdot p_i.$$

$$M_\xi(t) = M[e^{t\xi}] = \int_{-\infty}^{\infty} e^{tx} \cdot f(x)\, dx.$$

По известной ПФМ можно вычислять моменты случайной величины по формуле

$$M[\xi^n] = \frac{d^n}{dx^n} M_\xi(t)|_{t=0}$$

ПФМ однозначно определяет распределение случайной величины. ПФМ суммы независимых случайных величин равна произведению их производящих функций моментов. Производящая функция существует только в случае существования всех моментов, а характеристическая функция – всегда.

Слайд 13.

Центральная предельная теорема (ЦПТ)

Вернемся к случайным величинам, но будем рассматривать уже не отдельные значения, а их выборки.

Выборка из $X \sim F(x)$:
$X^n = (X_1, X_2, ..., X_n)$

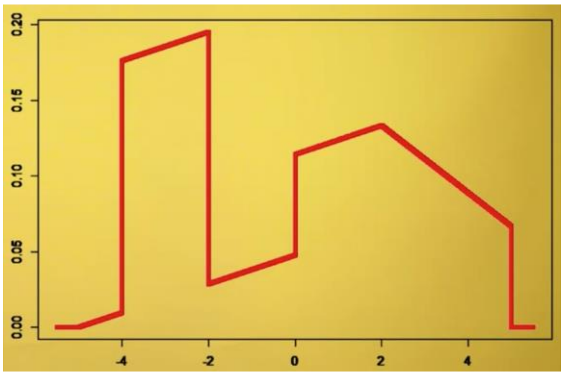
Выборочное среднее: $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$

У выборочного среднего пишем индекс n, просто чтобы понимать с выборкой какого размера мы работаем. Давайте подумаем, как связано выборочное среднее с исходным распределением?

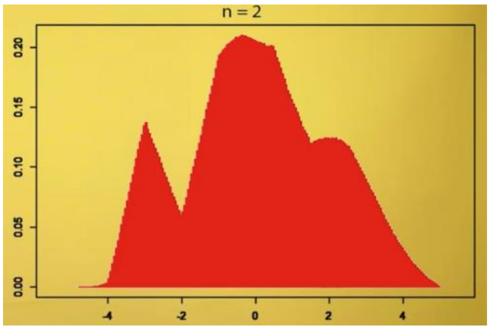
$\bar{X}_n \sim ?$

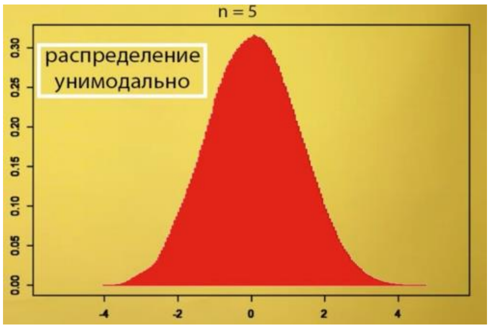
Слайд 14.

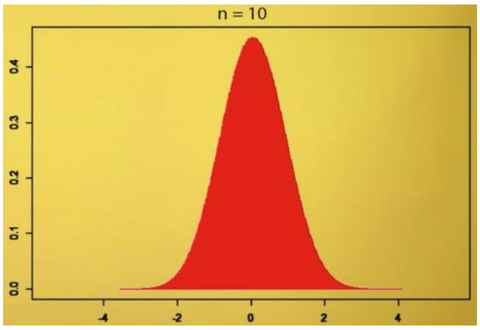
Будем работать с таким “странным” распределением.

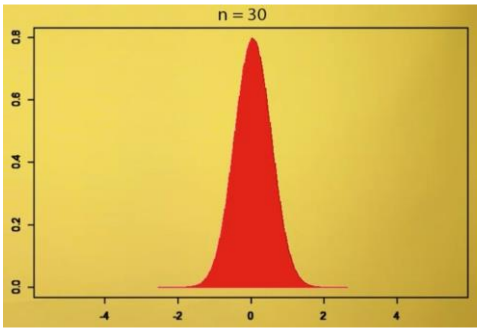


Давайте будем семплировать выборки объёма n, считать по ним выборочные средние и повторять так много-много раз. И давайте построим гистограмму этих выборочных средних.



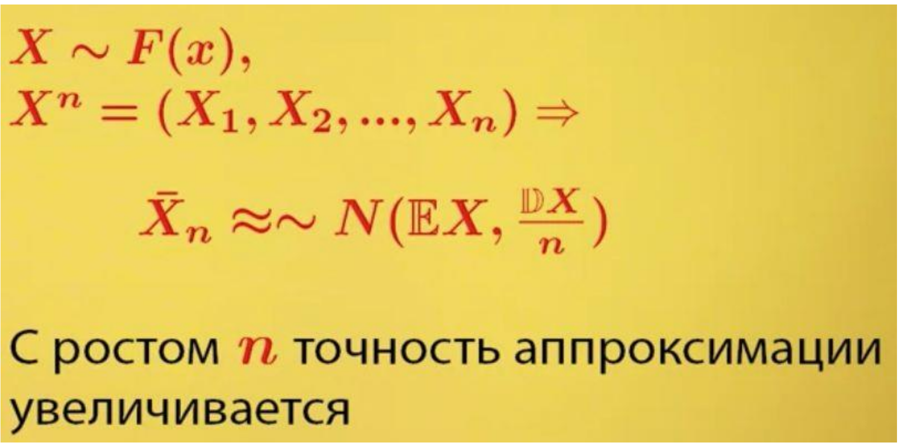




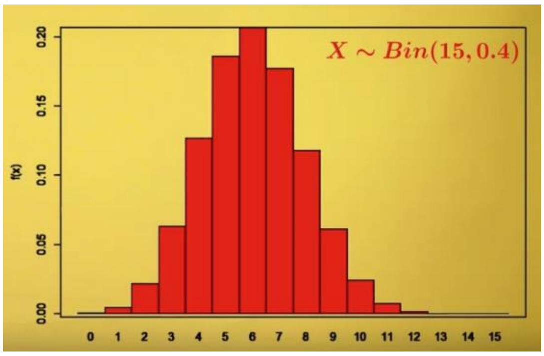


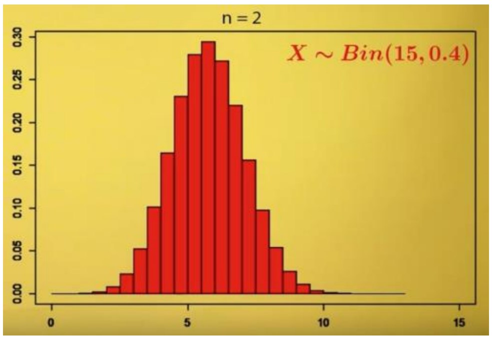
На плотность какого распределения похожи эти графики?

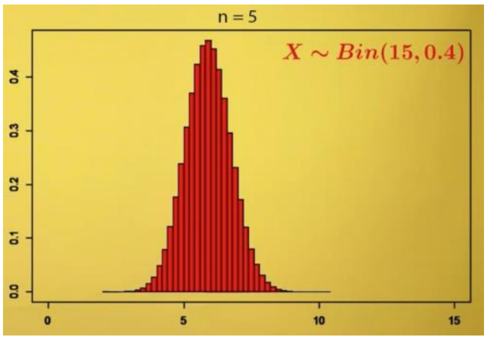
Слайд 15.

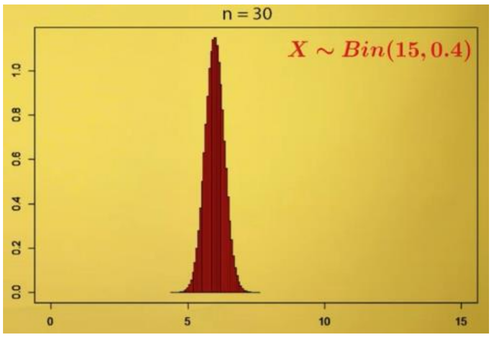


Интересно. Что это справедливо не только для абсолютно непрерывных величин, но и для дискретных.

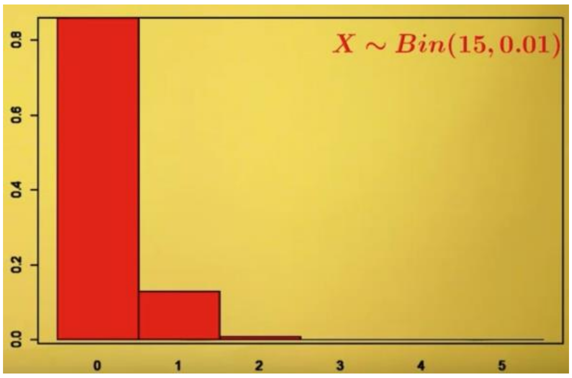


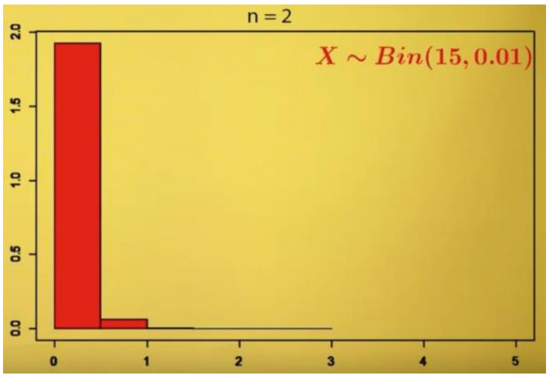


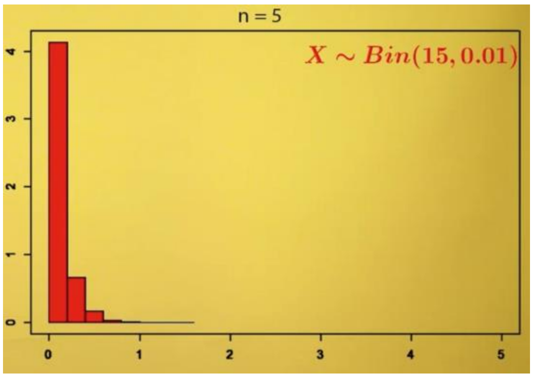

n = 5
$X \sim Bin(15, 0.4)$

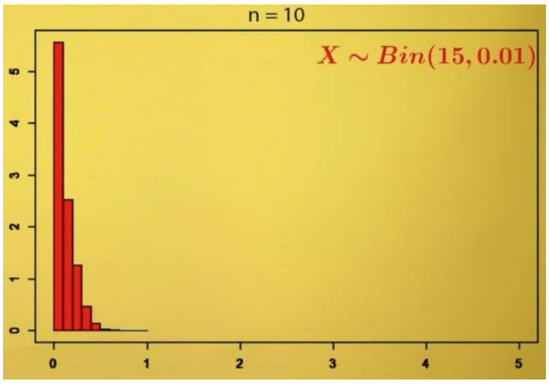

n = 30
$X \sim Bin(15, 0.4)$

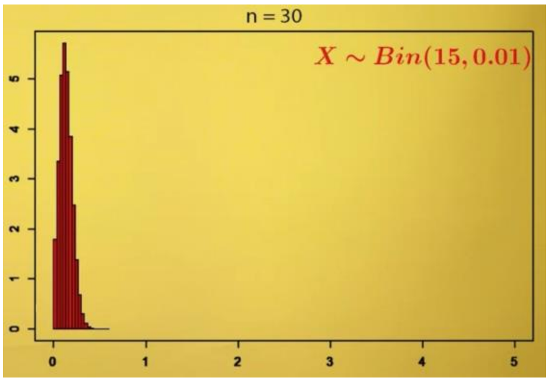
Слайд 16.


$X \sim Bin(15, 0.01)$


n = 2
$X \sim Bin(15, 0.01)$



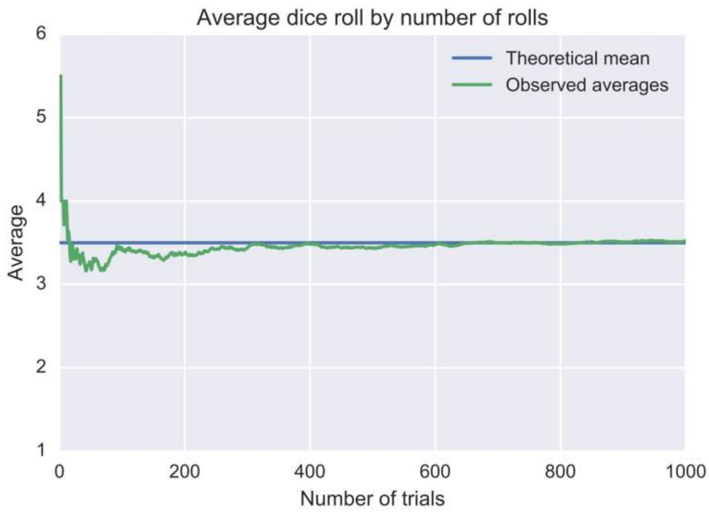




Слайд 17.

Закон больших чисел – принцип, описывающий результат выполнения одного и того же эксперимента много раз. Согласно закону. Среднее значение конечной выборки из фиксированного распределения близко к математическому ожиданию этого эксперимента

$$\frac{1+2+3+4+5+6}{6} = 3{,}5$$



Среднее значение очков при подбрасывании игральной кости - 3,5. Согласну закону больших чисел при большом количестве бросков их среднее значение, вероятно, будет близким к 3.5, при этом точность буде возрастать по мере увеличения числа бросков.

Слайд 18.

Закон больших чисел (ЗБЧ)

Давайте сформулируем Закон больших чисел более формально.

Рассмотрим последовательность независимых в совокупности случайных величин X1, X2, …, , которые имеют одинаковые распределения, следовательно и одинаковые математические ожидания $E(X_1) = E(X_2) = \ldots = \mu.$

$$\overline{X}_n = \frac{1}{n}(X_1 + \ldots + X_n)$$

- их среднее арифметическое.

Слабый закон больших чисел гласит, что среднее значение выборки сходится по вероятности к математическому ожиданию.

$\overline{X}_n \xrightarrow{P} \mu$ при $n \to \infty$

То есть $\forall \varepsilon > 0$:

$$\lim_{n\to\infty} P\left(\left|\overline{X}_n - \mu\right| > \varepsilon \right) = 0.$$

.

Центральная предельная теорема

В этом задании нужно убедиться, что ЦПТ действительно работает

In [ ]:

```
import pandas as pd
```

```python
import numpy as np
import matplotlib.pyplot as plt
import scipy.stats as sts
%matplotlib inline
```

Создайте случайную величину из любого выбранного вами распределения, для разных значений n сгенерируйте 1000 выборок размера n и постройте гистрограммы средних этих выборок.

In [ ]:

```python
rv = ### YOUR CODE HERE
sample = rv.rvs(size=1000)
```

In [ ]:

```python
x = np.linspace(0,2,100)
pdf = rv.pdf(x)
plt.hist(sample, density=True, bins=30, label='sample histogram')
plt.plot(x, pdf, label='theoretical pdf', alpha=0.5)
plt.legend()
plt.ylabel('fraction of samples, $f(x)$')
plt.xlabel('$x$')
plt.show()
```

In [ ]:

```python
#параметры запуска: размеры выборок и количество повторов генерации
n1 = 2
# YOUR CODE HERE

samples_count = 1000
#массивы для сохранения средних по выборкам
a1 = np.array([])

for number in range(samples_count):
    # на каждом цикле генерируем выборки разных размеров
    sample1 = rv.rvs(size=n1)
    # YOUR CODE HERE
    # считаем среднее по каждой выборке и добавляем в массив средних
    a1 = np.append(a1, sample1.mean())
    # YOUR CODE HERE
```

In [ ]:

```python
#строим графики
#подписи для графиков
label_1 = 'sample of ' + str(n1) + ' histogram'
plt.hist(a1, density=True, label=label_1)

plt.legend()
plt.ylabel('samples means')
plt.xlabel('$x$')
```