ЦПТ и ЗБЧ

Неравенства Чебышёва и Маркова. Некоторые виды сходимости случайных величин. Производящие функции распределений, производящие функции моментов, характеристические функции. Центральная предельная теорема. Закон больших чисел.

Даниил Корбут
Специалист по Анализу Данных

Даниил Корбут
DL Researcher
Insilico Medicine, Inc

Окончил бакалавриат ФИВТ МФТИ (Анализ данных) в 2018г
Учусь на 2-м курсе магистратуры ФИВТ МФТИ
Работал в Statsbot и Яндекс. Алиса.
Сейчас в Insilico Medicine, Inc, занимаюсь генерацией активных молекул и исследованиями старения с помощью DL.

Неравенство Маркова

Мы уже умеем считать математическое ожидание случайной величины, её разброс (дисперсию) вокруг него. Хочется понимать вероятность того, 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**:

!

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

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

Неравенство Чебышёва

Пусть мы знаем про случайную величину не только её математическое ожидание (первый момент), но и дисперсию (второй центральный момент), и эти величины конечны. 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 Чёбышева показывает, что случайная величина принимает значения близкие к среднему (математическому ожиданию) и дает оценку вероятности больших отклонений.

Неравенство Чебышёва (пример)

Положим a=kσ, где σ - стандартное отклонение, тогда получим оценку вероятности того, что случайная величина отклонится по модулю от среднего больше чем на kσ:

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

Для значения k=2 вероятность отклонения меньше 25%, для k=3 - уже 11.12%.

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

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

https://www.matburo.ru/

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, определённых 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 $\xi(\omega)$.

Сходимости случайных величин

!

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

!

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

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

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

https://mipt.ru/diht/students/courses/probability_zhukovsky.pdf

Сходимости случайных величин

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: $\xi_n \xrightarrow{L^p} \xi$.

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: $\xi_n \xrightarrow{d} \xi$

https://mipt.ru/diht/students/courses/probability_zhukovsky.pdf

Сходимости случайных величин

!

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: $\xi_n \xrightarrow{L^p} \xi$.

!

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: $\xi_n \xrightarrow{d} \xi$

НЕТОЛОГИЯ

https://mipt.ru/diht/students/courses/probability_zhukovsky.pdf

Сходимости случайных величин

Хотя и кажется, что определения сходимостей (быть может, за исключением сходимости по распределению) эквивалентны, это далеко не так. Есть теорема о взаимосвязей 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. $\xi_n \xrightarrow{\mathsf{P}} \xi \Rightarrow \xi_n \xrightarrow{d} \xi$

!

НЕТОЛОГИЯ

https://mipt.ru/diht/students/courses/probability_zhukovsky.pdf

Характеристические 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.

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

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

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

НЕТОЛОГИЯ

https://mipt.ru/diht/students/courses/probability_zhukovsky.pdf
https://www.matburo.ru/

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.

Центральная предельная теорема (ЦПТ)

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

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

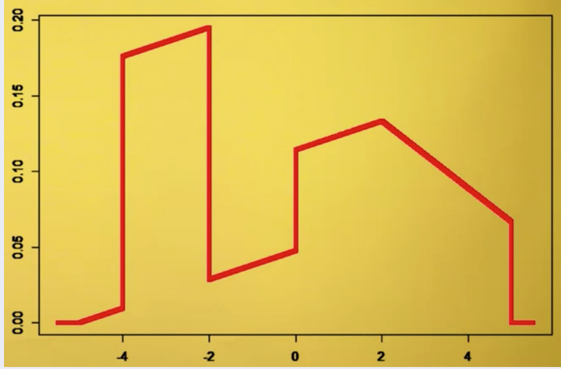
Выборочное среднее: $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

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

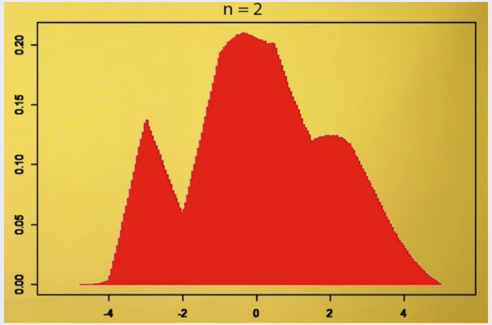
$$\bar{X}_n \sim ?$$

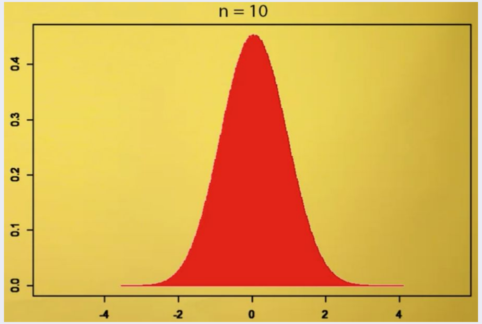
https://www.coursera.org/learn/mathematics-and-python

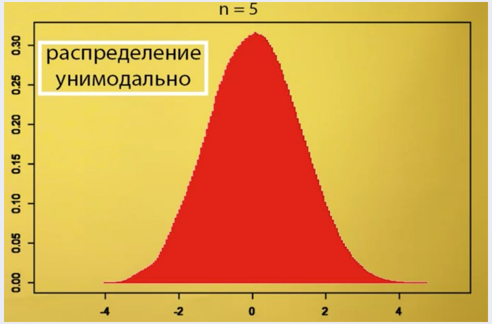
Центральная предельная теорема (ЦПТ)

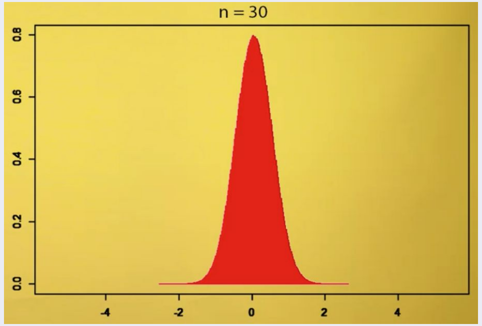


Будем работать с таким “странным” распределением.
Давайте будем семплировать выборки объёма n, считать по ним выборочные средние и повторять так много-много раз. И давайте построим гистограмму этих выборочных средних.









На плотность какого распределения похожи полученные графики?

НЕТОЛОГИЯ

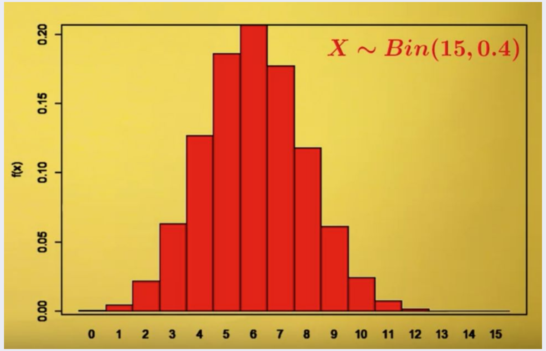
https://www.coursera.org/learn/mathematics-and-python

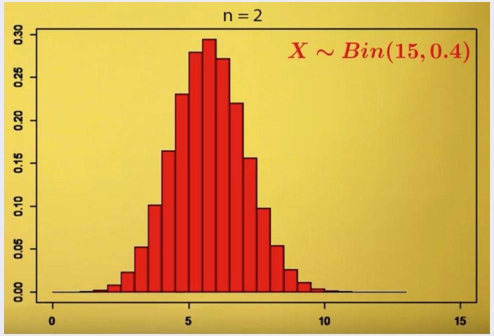
Центральная предельная теорема

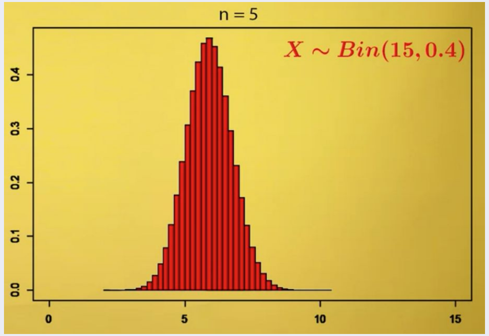
$$X \sim F(x),$$
$$X^n = (X_1, X_2, \dots, X_n) \Rightarrow$$
$$\bar{X}_n \approx\sim N\left(\mathbb{E}X, \frac{\mathbb{D}X}{n}\right)$$

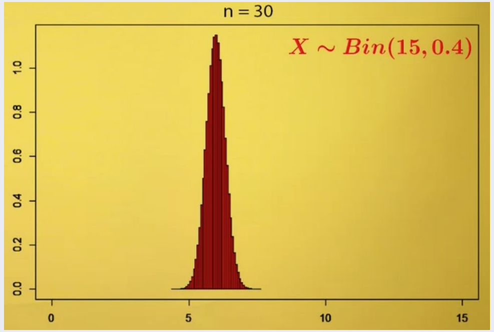
С ростом n точность аппроксимации увеличивается

Интересно, что это справедливо не только для абсолютно непрерывных распределений, но и для дискретных.



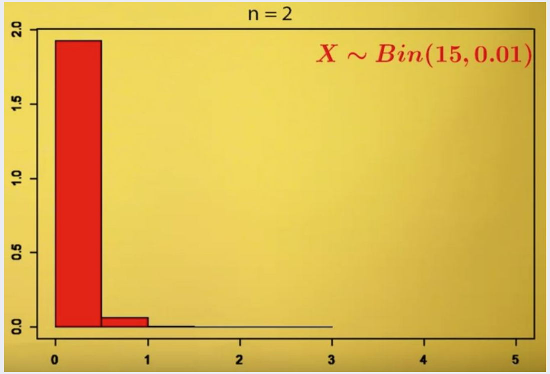


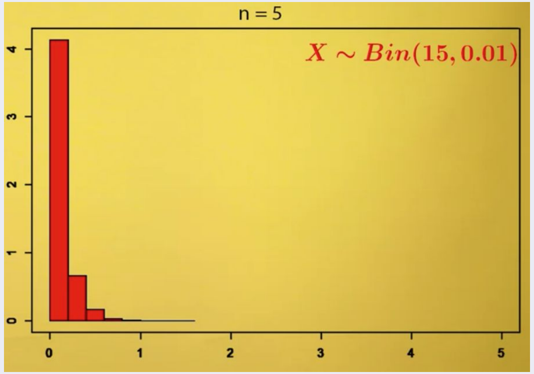


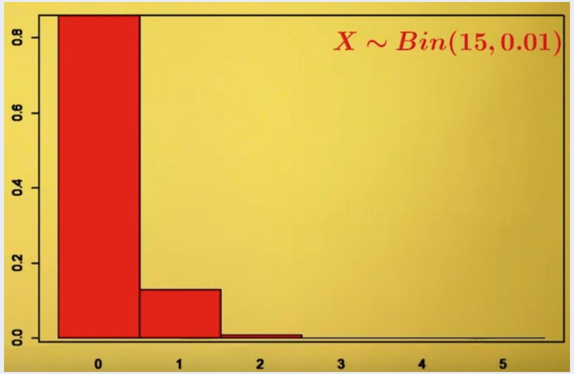


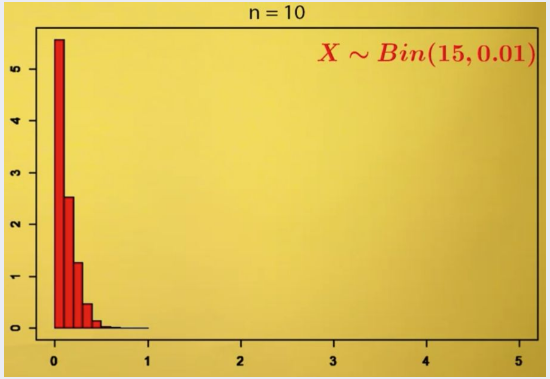
https://www.coursera.org/learn/mathematics-and-python

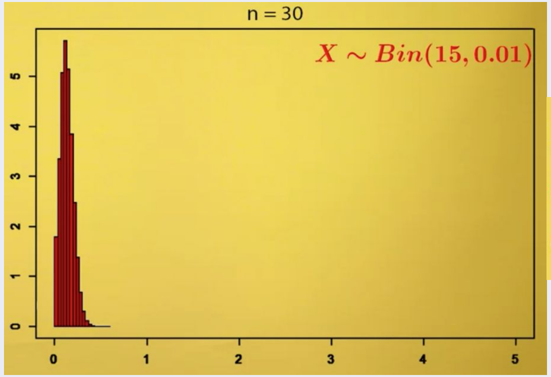
Центральная предельная теорема











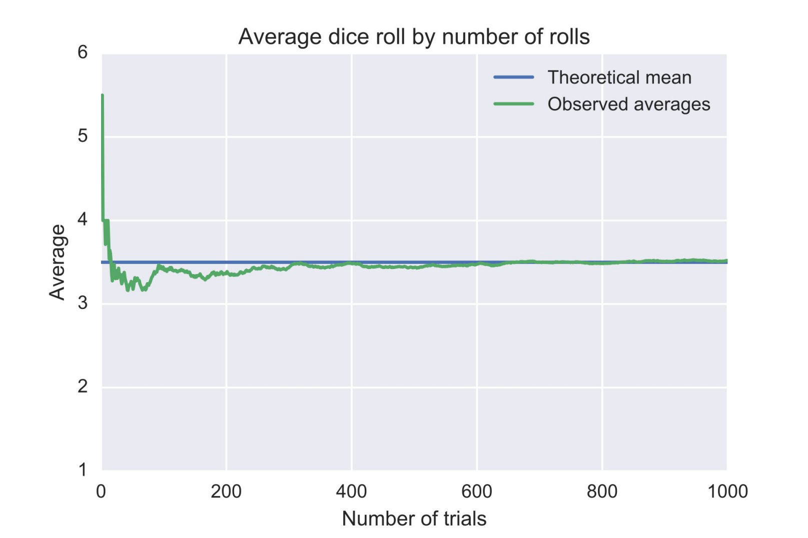
Когда распределение X не слишком скошено, распределение $\bar{X}_n$ хорошо описывается нормальным при $n \geq 30$.

https://www.coursera.org/learn/mathematics-and-python

Закон больших чисел (ЗБЧ)

ЗБЧ - 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 распределения.

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



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

https://ru.wikipedia.org/

Закон больших чисел (ЗБЧ)

Давайте сформулируем Закон больших чисел более формально.

Рассмотрим последовательность независисимых в совокупности случайных величин $X_1, X_2, \ldots,$ 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.

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

То есть $\forall \varepsilon > 0$:

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

НЕТОЛОГИЯ

https://ru.wikipedia.org/

Спасибо за внимание!