

ИССЛЕДОВАНИЕ УТВЕРЖДЕНИЯ ЦЕНТРАЛЬНОЙ ПРЕДЕЛЬНОЙ ТЕОРЕМЫ С ИСПОЛЬЗОВАНИЕМ СРЕДСТВ ЯЗЫКА PYTHON

Исследуется утверждение центральной предельной теоремы на примере треугольного и бета-распределений с использованием языка Python и развитых возможностей Jupyter Notebook.

ВВЕДЕНИЕ

Центральная предельная теорема является важной составляющей теории вероятностей и математической статистики. Но порой не всегда можно найти ее графическое представление. В связи с этим проведем исследование с целью проверки следующего утверждения: если имеется случайная величина X из практически любого распределения, и из этого распределения случайным образом сформирована выборка объемом N , то выборочное среднее, определенное на основании выборки, можно приблизить нормальным распределением со средним значением, которое совпадает с математическим ожиданием исходной совокупности.

I. ПОДХОД К ИССЛЕДОВАНИЮ

Задачей исследования является моделирование распределения выборочного среднего СВ X при разных объемах выборок и оценка его аппроксимации с нормальной кривой. Для проведения эксперимента требуется выбрать распределение, из которого случайным образом будет формироваться выборка. Воспользуемся треугольным и бета-распределениями. Формирование выборок, подсчет их средних, построение графиков и гистограмм осуществляется с помощью инструментария библиотек языка Python: *scipy*, *numpy*, *matplotlib*.

II. ХОД ИССЛЕДОВАНИЯ

Рассмотрим сперва треугольное распределение непрерывной случайной величины X , математическое и дисперсия которого вычисляются следующим образом:

$$E(X) = \frac{a + b + c}{3}. \quad (1)$$

$$D(X) = \frac{a^2 + b^2 + c^2 - ab - ac - bc}{18}. \quad (2)$$

Где $[a, c]$ – область определения X , $b \in [a, c]$ – абсцисса перегиба прямой плотности вероятности. В библиотеке *scipy* распределение задается параметром d . В нашем случае $d = 0.2$ и X определена на отрезке $[0, 1]$, $b = 0.2$. Из данного распределения выберем 100 псевдослучайных значений. Сравним полученные результаты выборки с теоретической плотностью вероятности, гра-

фик которой соответствует голубой линии (*pdf* – *probability distribution function*) на рисунке 1.

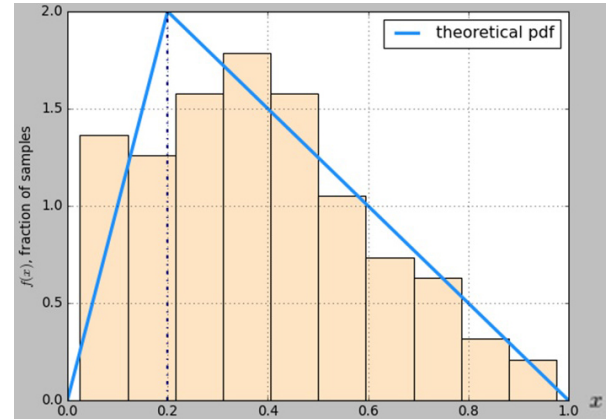


Рис. 1 – Треугольное распределение с объемом выборки 100

Далее, и это самое главное. При трёх и более значениях n генерируется 1000 выборок объёма n , вычисляется для каждой выборки среднее арифметическое. Строится гистограмма полученных выборочных средних и поверх нее накладывается график плотности соответствующего нормального распределения с параметрами:

$$\mu = E(X). \quad (3)$$

$$\sigma = \sqrt{\frac{D(X)}{n}}. \quad (4)$$

Реализуем функцию *build_hist_norm(n)* генерации нормального распределения и визуализации гистограмм по параметру объёма выборки n . Осуществим ее вызов 6 раз со следующими значениями n : 3, 5, 10, 50, 150, 300. Получаем результаты, представленные на рисунках 2-7.

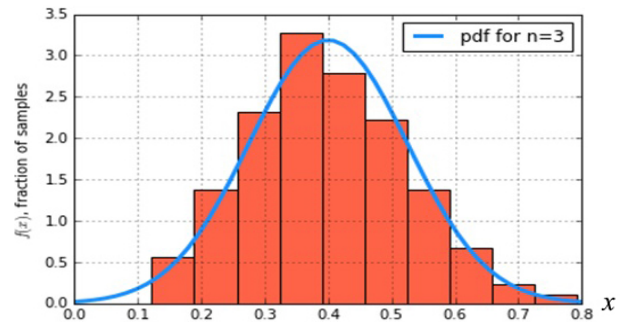


Рис. 2 – *build_hist_norm(3)*

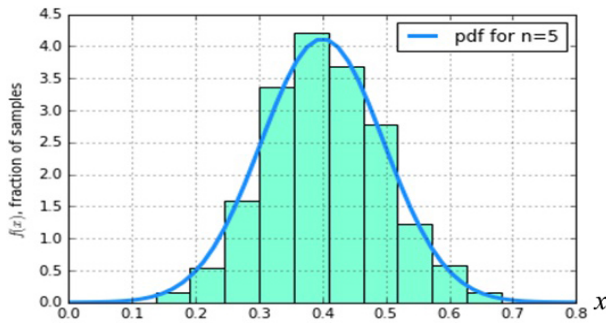


Рис. 3 – *build_hist_norm(5)*

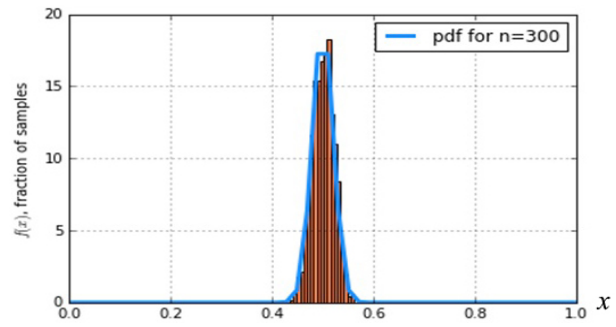


Рис. 7 – *build_hist_norm(300)*

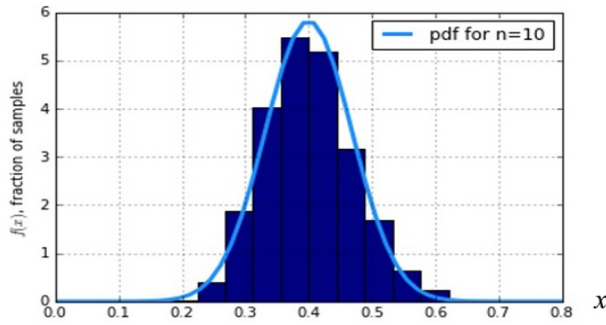


Рис. 4 – *build_hist_norm(10)*

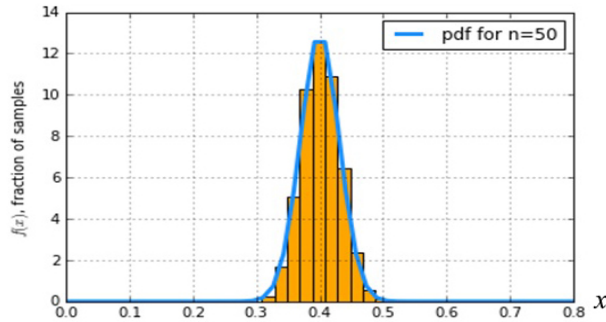


Рис. 5 – *build_hist_norm(50)*

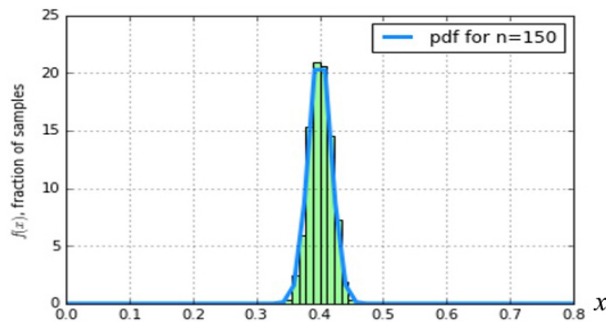


Рис. 6 – *build_hist_norm(150)*

Также рассмотрим бета-распределение со следующими числовыми характеристиками:

$$E(X) = \frac{\alpha}{\alpha + \beta}. \quad (5)$$

$$D(X) = \frac{\alpha\beta}{(\alpha + \beta)^2 (\alpha + \beta + 1)}. \quad (6)$$

Где $\alpha > 0$, $\beta > 0$. Зададим распределение с $\alpha = \beta = 0.5$. Тогда функция плотности распределения и гистограмма выборки объема 100 примут вид как на рисунке 8. Распределение выборочных средних аналогично треугольному.

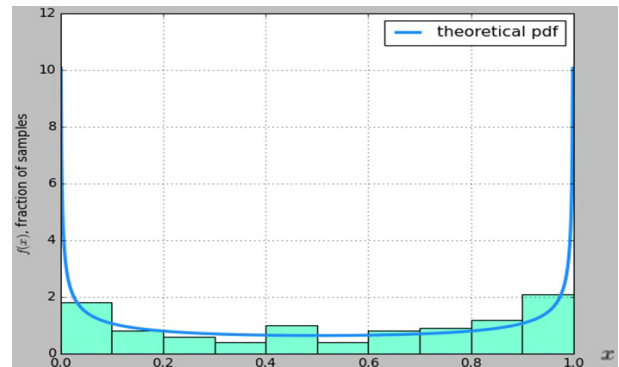


Рис. 8 – Бета-распределение с объемом выборки 100

III. Выводы

В соответствии с графическим представлением результатов хорошо прослеживается следующая закономерность: с ростом объема выборки степень аппроксимации распределения выборочных средних с нормальным распределением также растет и происходит концентрация псевдослучайных величин вокруг математического ожидания исходного распределения, что обосновывает утверждение центральной предельной теоремы.

1. Вентцель Е. В. Теория вероятностей / Е. В. Вентцель // Москва: «Высшая школа». – 2006. – 578 с.

Гусев Станислав Александрович, Кучко Никита Сергеевич, Гудков Алексей Сергеевич, студенты 2 курса факультета информационных технологий и управления БГУИР.

Научный руководитель: Гуринович Алевтина Борисовна, заместитель декана ФИТиУ, кандидат физико-математических наук, доцент, gurinovich@bsuir.by.