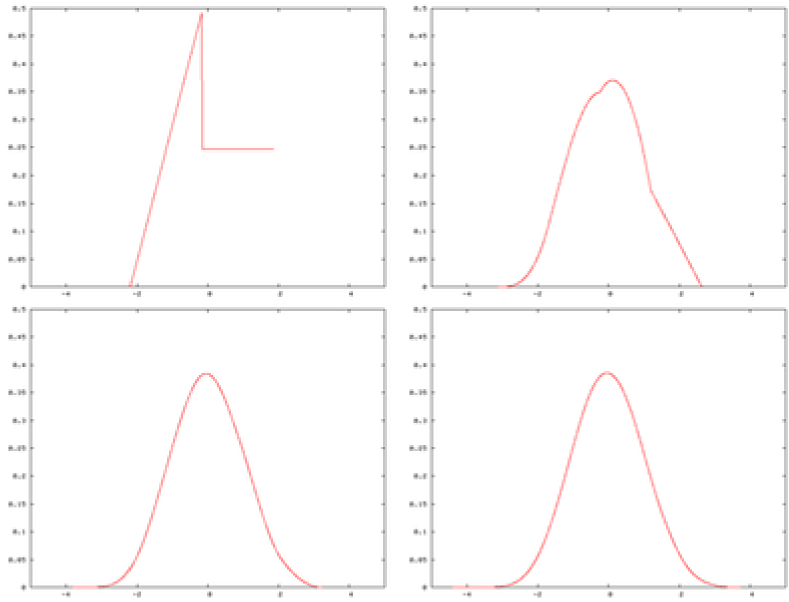
ВикипедиЯ

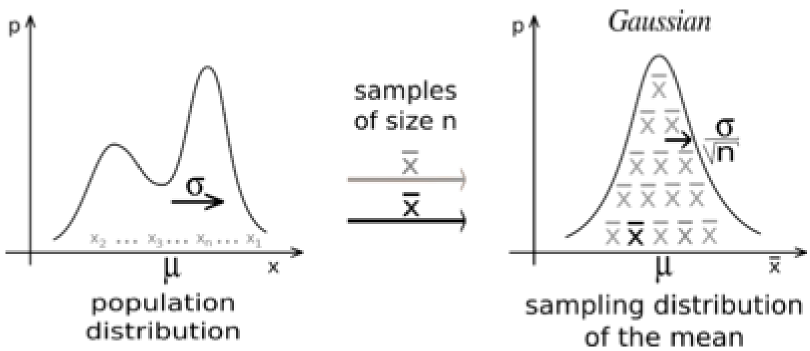
Центральная предельная теорема

Материал из Википедии — свободной энциклопедии

Центра́льные преде́льные теоре́мы (ЦПТ) — класс теорем в теории вероятностей, утверждающих, что сумма достаточно большого количества слабо зависимых случайных величин, имеющих примерно одинаковые масштабы (ни одно из слагаемых не доминирует, не вносит в сумму определяющего вклада), имеет распределение, близкое к нормальному.



«Сглаживание» распределения суммированием. Показана функция плотности вероятности одной случайной величины, а также распределения суммы двух, трёх и четырёх случайных величин с такой же функцией распределения.



Какова бы ни была форма распределения генеральной совокупности, выборочное распределение стремится к нормальному, а его дисперсия задается центральной предельной теоремой.[1]

Так как многие случайные величины в приложениях формируются под влиянием нескольких слабо зависимых случайных факторов, их распределение считают нормальным. При этом должно соблюдаться условие, что ни один из факторов не является доминирующим. Центральные предельные теоремы в этих случаях обосновывают применением нормального распределения.

Содержание

Классическая ЦПТ

Пусть $X_1, \ldots, X_n, \ldots$ есть бесконечная последовательность независимых одинаково распределённых случайных величин, имеющих конечное математическое ожидание μ и дисперсию σ^2. Пусть также

$$S_n = \sum_{i=1}^{n} X_i.$$

Тогда

$$\frac{S_n - \mu n}{\sigma\sqrt{n}} \to N(0,1) \text{ по распределению при } n \to \infty,$$

где $N(0,1)$ — нормальное распределение с нулевым математическим ожиданием и стандартным отклонением, равным единице. Определяя выборочное среднее первых n величин как

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

мы можем переписать результат центральной предельной теоремы в следующем виде:

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \to N(0,1) \text{ по распределению при } n \to \infty.$$

Скорость сходимости можно оценить с помощью неравенства Берри — Эссеена.

Замечания

- Неформально говоря, классическая центральная предельная теорема утверждает, что сумма n независимых одинаково распределённых случайных величин имеет распределение, близкое к $N(n\mu, n\sigma^2)$. Эквивалентно, $\bar{X}_n$ имеет распределение близкое к $N(\mu, \sigma^2/n)$.
- Так как функция распределения стандартного нормального распределения непрерывна, сходимость к этому распределению эквивалентна поточечной сходимости функций распределения к функции распределения стандартного нормального распределения. Положив $Z_n = \frac{S_n - \mu n}{\sigma\sqrt{n}}$, получаем $F_{Z_n}(x) \to \Phi(x),\ \forall x \in \mathbb{R}$, где $\Phi(x)$ — функция распределения стандартного нормального распределения.
- Центральная предельная теорема в классической формулировке доказывается методом характеристических функций (теорема Леви о непрерывности).

- Вообще говоря, из сходимости функций распределения не вытекает сходимость плотностей. Тем не менее в данном классическом случае это имеет место.

Локальная ЦПТ

В предположениях классической формулировки, допустим в дополнение, что распределение случайных величин $\{X_i\}_{i=1}^{\infty}$ абсолютно непрерывно, то есть оно имеет плотность. Тогда распределение Z_n также абсолютно непрерывно, и более того,

$$f_{Z_n}(x) \to \frac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}} \text{ при } n \to \infty,$$

где $f_{Z_n}(x)$ — плотность случайной величины Z_n, а в правой части стоит плотность стандартного нормального распределения.

Обобщения

Результат классической центральной предельной теоремы справедлив для ситуаций гораздо более общих, чем полная независимость и одинаковая распределённость.

ЦПТ Линдеберга

[2]Пусть независимые случайные величины $X_1, \ldots, X_n, \ldots$ определены на одном и том же вероятностном пространстве и имеют конечные математические ожидания и дисперсии: $\mathbb{E}[X_i] = \mu_i,\ \mathrm{D}[X_i] = \sigma_i^2$.

Пусть $S_n = \sum\limits_{i=1}^{n} X_i$.

Тогда $\mathbb{E}[S_n] = m_n = \sum\limits_{i=1}^{n} \mu_i,\ \mathrm{D}[S_n] = s_n^2 = \sum\limits_{i=1}^{n} \sigma_i^2$.

И пусть выполняется **условие Линдеберга**:

$$\forall \varepsilon > 0,\ \lim_{n\to\infty} \sum_{i=1}^{n} \mathbb{E}\left[\frac{(X_i - \mu_i)^2}{s_n^2}\, \mathbf{1}_{\{|X_i - \mu_i| > \varepsilon s_n\}}\right] = 0,$$

где $\mathbf{1}_{\{|X_i - \mu_i| > \varepsilon s_n\}}$ функция — индикатор.

Тогда

$$\frac{S_n - m_n}{s_n} \to N(0,1) \text{ по распределению при } n \to \infty.$$

ЦПТ Ляпунова

Пусть выполнены базовые предположения ЦПТ Линдеберга. Пусть случайные величины $\{X_i\}$ имеют конечный третий момент. Тогда определена последовательность

$$r_n^3 = \sum_{i=1}^{n} \mathbb{E}\left[|X_i - \mu_i|^3\right].$$

Если предел

$$\lim_{n\to\infty} \frac{r_n}{s_n} = 0 \text{ (**условие Ляпунова**)},$$

то

$$\frac{S_n - m_n}{s_n} \to N(0,1) \text{ по распределению при } n \to \infty.$$

ЦПТ для мартингалов

Пусть процесс $(X_n)_{n\in\mathbb{N}}$ является мартингалом с ограниченными приращениями. В частности, допустим, что

$$\mathbb{E}\left[X_{n+1} - X_n \mid X_1, \ldots, X_n\right] = 0,\ n \in \mathbb{N},\ X_0 \equiv 0,$$

и приращения равномерно ограничены, то есть

$$\exists C > 0\, \forall n \in \mathbb{N}\ |X_{n+1} - X_n| \le C \text{ п.н.}$$

Введём случайные процессы σ_n^2 и τ_n следующим образом:

$$\sigma_n^2 = \mathbb{E}\left[(X_{n+1} - X_n)^2 \mid X_1, \ldots, X_n\right]$$

и

$$\tau_n = \min\left\{k \,\middle|\, \sum_{i=1}^{k} \sigma_i^2 \ge n\right\}.$$

Тогда

$$\frac{X_{\tau_n}}{\sqrt{n}} \to N(0,1) \text{ по распределению при } n \to \infty.$$

ЦПТ для случайных векторов

Пусть $\overrightarrow{X_1}, \ldots, \overrightarrow{X_n}, \ldots$ последовательность независимых и одинаково распределённых случайных векторов, каждый из которых имеет среднее $E\overrightarrow{X_1} = \vec{a}$ и невырожденную матрицу ковариаций Σ. Обозначим через $S_n = \overrightarrow{X_1} + \ldots + \overrightarrow{X_n}$ вектор частичных сумм. Тогда при $n \to \infty$ имеет место слабая сходимость распределений векторов

$\overrightarrow{\eta_n} = \frac{S_n - na}{\sqrt{n}} \xrightarrow{weak} \vec{\eta}$, где $\vec{\eta}$ имеет распределение $N(\vec{0}, \Sigma)$.

См. также

- Закон больших чисел
- Закон повторного логарифма

Примечания

1. *Rouaud, Mathieu.* Probability, Statistics and Estimation (http://www.incertitudes.fr/book.pdf) (неопр.). — 2013. — С. 10.
2. *Шуленин В. П.* Математическая статистика. Ч. 3. Робастная статистика: учебник (http://www.fpmk.tsu.ru/sites/default/files/%D0%A5%D1%80%D0%B0%D0%BD%D0%B8%D0%BB%D0%B8%D1%89%D0%B5/%D1%84%D0%BF%D0%BC%D0%BA/%D1%83%D1%87%D0%B5%D0%B1%D0%BD%D0%B8%D0%BA%D0%B8/Shulenin-MS/MS-part3.pdf). — Томск: Издательство НТЛ, 2012. — С. 474. — 520 с. — ISBN 978-5-89503-508-5.

Ссылки

- Предельные теоремы теории вероятностей. Примеры использования (http://mathhelpplanet.com/static.php?p=predelnye-tyeoremy-tyeorii-veroyatnostyei)