СЕМ, Л-12
(2020-12-17)

Централна гранична теорема (ЦГТ)

Централната гранична теорема (ЦГТ) [9] е един от най-фундаменталните резултати в теорията на вероятностите и статистиката. В своята най-класическа форма тя гласи, че при определени условия разпределението на сумата (или средната стойност) на голям брой независими случайни величини се доближава до нормално разпределение (разпределение на Гаус или „камбановидна крива“), дори ако самите величини не са нормално разпределени.

Тази теорема обяснява защо нормалното разпределение възниква толкова често в природата, науката и индустрията – много сложни явления са резултат от сумирането на огромен брой малки и независими влияния.

Дефиниция и формулировка (ЦГТ)

Нека $X_1, X_2, \dots, X_n$ е редица от независими и еднакво разпределени (i.i.d.) случайни величини със крайно очакване μ и крайна дисперсия $\sigma^2 > 0$.

Да разгледаме тяхната сума $S_n = X_1 + X_2 + \dots + X_n$ и средната стойност $\overline{X}_n = \dfrac{S_n}{n}$.

Тогава:

- Очакване: $\mathbb{E}[\overline{X}_n] = \mu$
- Дисперсия: $\mathrm{Var}[\overline{X}_n] = \dfrac{\sigma^2}{n}$

Централна гранична теорема гласи, че стандартизираната случайна величина

$$Z_n = \frac{\overline{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

клони по разпределение към стандартно нормално разпределение $\mathcal{N}(0, 1)$ при $n \to \infty$.

Математически това се записва като:

$$\lim_{n\to\infty} \mathbb{P}(Z_n \le z) = \Phi(z),\ \forall z \in \mathbb{R}$$

Където $\Phi(z)$ е функцията на разпределение (кумулативната функция) на стандартното нормално разпределение.

С други думи, при достатъчно голям размер на извадката n разпределението на извадковата средна стойност $\overline{X}_n$ е приблизително нормално с параметри $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$.

ЦГТ измерва каква е грешката в ЗГЧ. Тоест ние знаем, че $\frac{S_n}{n} \xrightarrow[n\to\infty]{\text{п.с.}} \mu$, почти сигурно (траекторно) и скалираме грешката по $\frac{\sigma}{\sqrt{n}}$ и виждаме, че се държи като нормално разпределение.

Често се дава за пример пресмятането на средно IQ за група хора или тяхна височина и т.н., но за да бъде IQ-то нормална разпределена случайна величина е необходимо да бъде средно аритметичен ефект от множество независими характеристики. А дали това е така?

Във физиката ЦГТ може да ни даде каква е грешката между акумулиран ефект и среден ефект, който сме измерили в някаква система.

$\oplus$ $Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}$, където $\mu = \mathbb{E}[X_1]$, $\sigma^2 = \mathrm{Var}[X_1]$. ЦГТ гласи следното:

Ако се интересуваме от:

$$\begin{aligned}\mathbb{P}\left(Z_n \in (a,b)\right) &= \mathbb{P}\left(S_n \in (n\mu + a\sigma\sqrt{n}, n\mu + b\sigma\sqrt{n})\right) \\ &\sim \mathbb{P}\left(Z \in (a,b)\right) = \Phi(b) - \Phi(a) = \\ &= \frac{1}{\sqrt{2\pi}} \cdot \int_a^b e^{-\frac{y^2}{2}}\,\mathrm{d}y.\end{aligned}$$

$$\mathbb{P}(Z_n < a) \xrightarrow[n\to\infty]{} \Phi(a) = \mathbb{P}(Z < a)$$

$$\mathbb{P}(Z_n \geq a) \xrightarrow[n\to\infty]{} \overline{\Phi}(a) = 1 - \Phi(a)$$

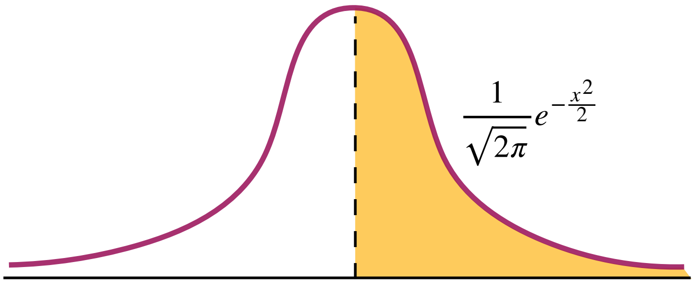
$$\mathbb{P}(Z_n < \mu) \xrightarrow[n\to\infty]{} \Phi(\mu).$$

Тоест ЦГТ дава инструмент, с който може да приближаваме вероятности.

$\oplus$ $\mu = 0$, $\sigma^2 = 1$, тогава $\frac{S_n}{\sqrt{n}} \xrightarrow[n\to\infty]{d} Z \in \mathcal{N}(0,\, 1)$, където $S_n = \Sigma_{i=1}^n X_i$.

$$\mathbb{P}\left(S_n \in (a\sqrt{n}, b\sqrt{n})\right) \sim \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}} \mathrm{d}\, y.$$

$$\mathbb{P}(S_n > 0) = \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > 0\right) \sim \mathbb{P}(Z > 0) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{y^2}{2}} \mathrm{d}\, y = \frac{1}{2}.$$



⊕ Хвърляме зарче 6 млн. пъти. Каква е вероятността измежду тези 6 млн. пъти да сме получили повече от 1 млн. пъти 6-ца?

Решение: виж последния пример от задача 10 от файла SEM HW.

Функция на моментите

Дефиниция. Нека X е случайна величина. Функцията на моментите на X, означавана като $M_X(t)$, се дефинира като:

$$M_X(t) = \mathbb{E}[e^{tX}]$$

където $\mathbb{E}[e^{tX}]$ съществува за реално $t \in (-\varepsilon, \varepsilon)$ и някое $\varepsilon > 0$.

Алтернативна форма за дискретни и непрекъснати случайни величини

1. За дискретна случайна величина:

$$M_X(t) = \sum_x e^{tx} \cdot p_X(x)$$

   където $p_X(x)$ е вероятностната масова функция.

2. За непрекъсната случайна величина:

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \cdot f_X(x) \mathrm{d}\, x$$

   където $f_X(x)$ е плътността на вероятността.

$\oplus$ $\mathbb{E}[e^{tX}] = \Sigma_i e^{tx_i} \cdot \mathbb{P}(X = x_i)$: винаги съществува за всяко t, ако стойностите са краен брой. Но ако не са, тази сума може да не е крайна и да клони към безкрайност.

Но ако вземем $x_i = j$ и $\mathbb{P}(X = j) = \dfrac{1}{j^2}$, то няма да може да направим сумирането за $t > 0$.

Ако имаме непрекъсната случайна величина, функцията на моментите се изчислява чрез интеграла $\mathbb{E}[e^{tX}] = \displaystyle\int_{-\infty}^{\infty} e^{tx} f_X(x) \mathrm{d}\,x$, който може да съществува само за някаква част от t, но е достатъчно да съществува за някаква малко околност около t например: $t \in (-\varepsilon, \varepsilon)$, за да може да го наречем функция на моментите.

Нека $X \sim \mathrm{Uni}(0, 1)$.

Тогава:

$$\begin{aligned} M_X(t) = \mathbb{E}[e^{tX}] &= \int_0^1 e^{tx} \cdot f_X(x) \mathrm{d}\,x \\ &= \int_0^1 e^{tx} \cdot \frac{1}{1-0} \mathrm{d}\,x = \frac{1}{t} \int_0^1 e^{tx} \mathrm{d}\,tx \\ &= \frac{1}{t} e^{tx} \Big|_0^1 = \frac{e^t - 1}{t}. \end{aligned}$$

е добре дефинирано за всяко $t \in \mathbb{R}$.

Винаги когато нашата случайна величина може да приема краен брой стойности или възможните стойности са в някакъв компакт или някакво крайно множество (в случая са от 0 до 1), то винаги функцията на моментите е добре дефинирана за всяко t.

Свойства. Нека X е случайна величина. Тогава:

- $\mathbb{E}[X^k]$ се нарича момент от ред $k \geq 1$
- $\mathbb{E}[\,|X|^k\,]$ се нарича абсолютен момент от ред k
- $\mathbb{E}[(X - \mathbb{E}[X])^k]$ се нарича централен момент от ред k
- $\mathbb{E}[\,|X - \mathbb{E}[X]|^k\,]$ се нарича абсолютен централен момент от ред k

Ще допускаме, че $M_X(t)$ е добре дефинирана за $t \in (-\varepsilon, \varepsilon)$

- $M_X(0) = \mathbb{E}[e^{0 \cdot X}] = \mathbb{E}[1] = 1$

- $M_X(0) \overset{\text{ред на Маклорен}}{=} \mathbb{E}\sum_{k=0}^{\infty}\frac{t^k}{k!}X^k = \sum_{k=0}^{\infty}\frac{t^k}{k!}\mathbb{E}[X^k]$

- $M_X^{(k)}(0) = \mathbb{E}[X^k],\ \forall k \geq 1$

- Ако $M_{X_n}(t) \underset{n\to\infty}{\longrightarrow} M_X(t)$, за $t \in (-\varepsilon, \varepsilon)$, където $\{X_n\}$ е редица от случайни величини, то $X_n \xrightarrow[n\to\infty]{d} X$

- $X \overset{d}{=} Y \Leftrightarrow M_X = M_Y$;

- Ако $X \perp\!\!\!\perp Y$ и M_X, M_Y са добре дефинирани за $t \in (-\varepsilon, \varepsilon)$, то $M_{X+Y}(t) = \mathbb{E}[e^{t(X+Y)}] = \mathbb{E}[e^{tX}e^{tY}] \overset{X\perp\!\!\!\perp Y}{=} \mathbb{E}[e^{tX}]\mathbb{E}[e^{tY}] = M_X(t)\cdot M_Y(t)$.

- Ако $Y = aX + b$, то $M_Y(t) = e^{bt}M_X(at)$, за всяко t, такова че $M_X(at)$ е добре дефинирано.

  Ако M_X е добре дефинирано за $t \in (-\varepsilon, \varepsilon)$, то $M_X(at)$ е добре дефинирано за $-\varepsilon < at < \varepsilon$ и следователно $M_Y(t)$ е добре дефинирано за $-\dfrac{\varepsilon}{|a|} < t < \dfrac{\varepsilon}{|a|}$.

  $M_Y(t) = \mathbb{E}[e^{t(aX+b)}] = \mathbb{E}[e^{bt}e^{taX}] = e^{bt}\mathbb{E}[e^{atX}] = e^{bt}M_X(at)$.

Твърдение. Ако $X \sim \mathcal{N}(\mu, \sigma^2)$, то за $\forall t \in \mathbb{R}$, то $M_X(t) = e^{\mu t}e^{\frac{\sigma^2 t^2}{2}}$.

Доказателство:

$$X = \mu + \sigma Z, \text{ където } Z \in \mathcal{N}(0, 1).$$

$$M_Z(t) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{tX}e^{-\frac{x^2}{2}}\,\mathrm{d}\,x = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}}\int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-t)^2}\,\mathrm{d}\,x$$
$$\overset{x-t=y}{=} e^{\frac{t^2}{2}}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}}\,\mathrm{d}\,y = e^{\frac{t^2}{2}}.$$

Следователно: $M_X(t) = e^{\mu t}M_Z(\sigma t) = e^{\mu t}e^{\frac{\sigma^2 t^2}{2}}, \forall t \in \mathbb{R}$.

$\square$

Характеристична функция

Дефиниция. Нека X е случайна величина с функция на разпределение $F_X(x)$. Характеристичната функция $\varphi_X(t)$ се дефинира като:

$$\varphi_X(t) = \mathbb{E}[e^{itX}] = \int_{-\infty}^{\infty} e^{itx}\,\mathrm{d}\,F_X(x),\ t \in \mathbb{R}$$

където i е имагинерната единица ($i^2 = -1$).

Ако X има плътност $f_X(x)$, тогава $\varphi_X(t) = \int_{-\infty}^{\infty} e^{itx} f_X(x)\mathrm{d}\,x$

Това е преобразуване на Фурие на функцията на разпределение (или плътността).

Прилики между функция на моментите и характеристична функция

1. Структура

   И двете са очаквани стойности на експоненциални функции на случайната величина:
   - $\varphi_X(t) = \mathbb{E}[e^{itX}]$
   - $M_X(t) = \mathbb{E}[e^{tX}]$

2. Свойства, свързани с моментите

   Ако $\mathbb{E}[\,|X|^n\,] < \infty$, то и двете функции позволяват получаване на моментите:

   $$\mathbb{E}[X^n] = i^{-n}\frac{\mathrm{d}^n}{\mathrm{t}^\mathrm{n}}\varphi_X(t)\Big|_{t=0} = \frac{\mathrm{d}^n}{\mathrm{d}\,t^n}M_X(t)\Big|_{t=0}$$

   при съответните условия за диференцируемост.

3. Сума на независими случайни величини

   $$\varphi_{X_1+X_2}(t) = \varphi_{X_1}(t)\cdot\varphi_{X_2}(t)$$

   $$M_{X_1+X_2}(t) = M_{X_1}(t)\cdot M_{X_2}(t)$$

Разлики между функция на моментите и характеристична функция		
Аспект	Характеристична функция	Функция на моментите
Аргумент	Реално число t, умножено по i	Реално число t
Област на съществуване	Винаги съществува за всяко $t \in \mathbb{R}$, тъй като $\|\|e^{itX}\|\| = 1$	Може да не съществува за $t \neq 0$ (особено при тежки опашки)
Връзка с преобразувания	Преобразуване на Фурие на разпределението	Преобразуване на Лаплас (двустранно) на разпределението
Комплексни стойности	Комплекснозначна функция с $\|\|\varphi_X(t)\|\| \leq 1$	Реалнозначна (при реални t)
Еднозначност	Еднозначно определя разпределението (теорема на Бохнер)	Не винаги еднозначно определя разпределението, дори ако съществува
Приложение в гранични теореми	Основен инструмент в доказателствата на централна гранична теорема	По-рядко се използва за гранични теореми поради ограниченията на съществуване

Доказателство (ЦГТ):

Теорема (Класическа ЦГТ)

Нека $X_1, X_2, \dots$ е редица от независими, еднакво разпределени случайни величини с крайно очакване $\mathbb{E}[X_k] = \mu$ и крайна положителна дисперсия $\mathrm{Var}[X_k] = \sigma^2 > 0$.

Нека $S_n = \sum_{k=1}^{n} X_k$ е частична сума. Тогава нормираната сума клони по разпределение към стандартно нормално разпределение:

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{d} Z \sim \mathcal{N}(0,1)$$

Където $\xrightarrow{d}$ означава сходимост по разпределение.

Доказателство (с характеристични функции):

Стъпка 1: Подготовка и центриране

Първо центрираме еличините. Нека $Y_k = X_k - \mu$. Тогава:

- $\mathbb{E}[Y_k] = 0$
- $\mathrm{Var}[Y_k] = \sigma^2$

Нормираната сума, която разглеждаме, става:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n}(X_n - \mu) = \frac{1}{\sigma\sqrt{n}}\sum_{k=1}^{n} Y_k$$

Нека $\varphi(t) = \mathbb{E}[e^{itY_k}]$ е характеристичната функция на всяко от центирраните събираеми Y_k (те са еднакво разпределени, така че дункцията е една и съща за всички).

Стъпка 2: Характеристична функция на Z_n

Поради независимостта и еднаквото разпределение, характеристичната функция на Z_n е :

$$\begin{aligned}\varphi_{Z_n}(t) = \mathbb{E}[e^{itZ_n}] &= \mathbb{E}\left[\exp\left(\frac{it}{\sigma\sqrt{n}}\sum_{k=1}^{n} Y_k\right)\right] \\ &= \left[\mathbb{E}\left(\exp\left(i\frac{t}{\sigma\sqrt{n}}Y_1\right)\right)\right]^n = \left[\varphi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n\end{aligned}$$

Стъпка 3: Развитие на $\varphi(s)$ в ред на Тейлор

Тъй като $\mathbb{E}[X_1] = 0$ и $\mathbb{E}[Y_1^2] = \sigma^2 < \infty$, характеристичната функция $\varphi(s)$ има непрекъсната втора производна. Можем да я развием в ред на Тейлор около нулта:

$$\begin{aligned}\varphi(s) = \mathbb{E}[e^{isY_1}] &= 1 + is\mathbb{E}[Y_1] + \frac{(is)^2}{2!}\mathbb{E}[Y_1^2] + o(s^2) \\ &= 1 - \frac{\sigma^2 s^2}{2} + o(s^2),\ s \to 0\end{aligned}$$

Тук използваме факта, че $\mathbb{E}[Y_1] = 0$, $\mathbb{E}[Y_1^2] = \sigma^2$, а $o(s^2)$ означава член, който клони към 0 по-бързо от s^2.

Стъпка 4: Прилагане към $\varphi_{Z_n}(t)$

За да намерим $\varphi_{Z_n}(t)$, заместваме $s = \dfrac{t}{\sigma\sqrt{n}}$:

$$\begin{aligned}\varphi\left(\frac{t}{\sigma\sqrt{n}}\right) &= 1 - \frac{\sigma^2}{2}\cdot\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + o\left(\frac{t^2}{n}\right) \\ &= 1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right),\ n \to \infty\end{aligned}$$

Стъпка 5: Граничен преход

Характеристичната функция на Z_n е:

$$\varphi_{Z_n}(t) = \left[1 - \frac{t^2}{2n} + o\left(\frac{t^2}{n}\right)\right]^n$$

Използваме добре известното гранично равенство:

$$\lim_{n\to\infty}\left(1 + \frac{a_n}{n}\right)^n = e^{\lim_{n\to\infty} a_n}, \text{ ако } a_n \to a$$

В нашия случай:

$$a_n = -\frac{t^2}{2} + n \cdot o\left(\frac{t^2}{n}\right) \longrightarrow -\frac{t^2}{2}$$

Следователно:

$$\lim_{n\to\infty} \varphi_{Z_n}(t) = e^{-\frac{t^2}{2}}$$

Стъпка 6: Прилагане на теоремата за непрекъснатост

Позната е Теоремата за непрекъснатост (на Леви): Ако характеристичните функции са поредица от случайни величини сходящи по точки към функция $\psi(t)$, която е непрекъсната в нулата, то $\psi(t)$ е характеристична функция на някакво разпределение, и величините клонят по разпределение към това разпределение.

В нашия случай:

- $\varphi_{Z_n}(t) \to e^{-t^2/2}$ за всяко t.
- Функцията $e^{-t^2/2}$ е непрекъсната в нулата и точно тя е характеристичната функция на стандартното разпределение $\mathcal{N}(0,1)$.

От теоремата на Леви следва, че:

$$Z_n \xrightarrow[n\to\infty]{d} Z \sim \mathcal{N}(0,1)$$

С това доказателството е завършено.

$\square$

Условие на Линдеберг – Обобщение

Класическата ЦГТ има строги ограничения: независимост, еднакво разпределение, крайна дисперсия. Тези ограничения биват отстранени от по-обща теорема на Линдеберг-Фелер.

- Ситуация: $X_1, X_2, \ldots$ са независими (но не задължително еднакво разпределени) с $\mathbb{E}[X_k] = \mu_k$ и $\mathrm{Var}[X_k] = \sigma_k^2 > 0$. Нека $s_n^2 = \Sigma_{k=1}^n \sigma_k^2$.
- Условие за пренебрежимост: Никое от събираемите не трябва да доминира сумата. Това се изразява като[10]:

$$\max_{1 \le k \le n} \frac{\sigma_k^2}{s_n^2} \xrightarrow[n \to \infty]{} 0$$

- Условие на Линдеберг: За всяко $\varepsilon > 0$ трябва да е изпълнено:

$$\lim_{n \to \infty} \frac{1}{s_n^2} \sum_{k=1}^{n} \mathbb{E}\left[(X_k - \mu_k)^2 \cdot \mathbf{1}_{\{|X_k - \mu_k| > \varepsilon s_n\}}\right] = 0$$

  Това условие означава, че „отделните големи отклонения“ са статистически незначителни в сумата.

Теорема на Линдеберг-Фелер: Ако е изпълнено условието за пренебрежимост, то условието на Линдеберг е необходимо и достатъчно за валидността на ЦГТ, тоест за сходимостта на $(S_n - \Sigma\mu_k)/s_n$ към $\mathcal{N}(0,1)$. Доказателстовот следва същата схема с харакеристични функции, но техниките са по-сложни поради разнородността.

По подразбиране под ЦГТ се има предвид класическата теорема за еднакво разпределени величини, доказана по-горе. Същинската мощ на теоремата идва от факта, че тя успява да опише граничното поведение на суми от случайни величини без значение какво е тяхното конкретно разпределение. Методът с характеристични функции е решаващ за доказателството. Обобщената теорема на Линдеберг показва, че това явление е по-универсално и се проявява при много по-общи условия.

Ярл Валдемар Линдеберг (швед. Jarl Waldemar Lindeberg, 4 август 1876 – 12 декември 1932)[11] е финландски математик от шведски произход, известен с основополагащия си принос в теорията на вероятностите и математическата статистика. Той е личността, на която се дължи Условието на Линдеберг, което беше споменато в горното доказателство на Централната гранична теорема.

Уилям Фелер (1906–1970)[12] е хърватско-американски математик, считан за един от най-влиятелните учени в областта на теорията на вероятностите през 20 век. Той дава ключови допълнения към работата на Ярл Линдеберг, поради което централната гранична теорема за разнородни случайни величини носи името „теорема на Линдеберг–Фелер“.

Тоерема (Бери-Есен). Теоремата на Бери-Есен (Berry-Esseen Theorem) е теорема за скоростта на сходимостта в Централната гранична теорема (ЦГТ). Тя дава ясна количествена граница за това колко бързо и колко точно разпределението на стандартизираната сума от случайни величини се доближава до нормалното разпределение.

Нека $\{X_i\}_{i=1}^{n}$ са независими и еднакво разпределени случайни величини с :

- Крайно очакване: $\mathbb{E}[X_k] = \mu$
- Крайна дисперсия: $\mathrm{Var}[X_k] = \sigma^2 > 0$
- Краен абсолютен трети момент: $\rho_3 = \mathbb{E}[\,|X_k - \mu|^3\,] < \infty$

Нека $F_n(x)$ е функцията на разпределение на стандартизираната сума:

$$Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}}, \text{ където } S_n = \sum_{k=1}^{n} X_k$$

А $\Phi(x)$ е функцията на разпределение на стандартното нормално разпределение $\mathcal{N}(0,1)$.

Твърдение на теоремата

Съществува универсално положителна константа C, такава че за всички $x \in \mathbb{R}$ и за всяко n е в сила неравенството:

$$\sup_{x\in\mathbb{R}} |F_n(x) - \Phi(x)| \leq C \cdot \frac{\rho_3}{\sigma^3\sqrt{n}}$$

Тук $\sup\limits_{x\in\mathbb{R}}$ означава супремум или най-голямото отклонение (разстоянието по равномерната норма).

Обяснение на параметрите

Величина	**Значение на сходимостта**
$\frac{\rho}{\sigma^3}$	Коефициент на асиметрия (стандартизиран). Характеризира скостта и асиметрията на първоначалното разпределение на X_k. Колкото по-симетрично и по-близко до нормалното е разпределението, толкова по-малък е този коефициент.
$\frac{1}{\sqrt{n}}$	Скорост на сходимост. Това е най-важният извод от теоремата: грешката в ЦГТ намалява като $\mathcal{O}(1/\sqrt{n})$, тоест много бавно. За да намалим грешката наполовина, трябва да учетворим броя на наблюденията n.
C	Абсолютна константа. От най-голям теоретичен интерес е да се намери най-добрата (най-малката) възможна такава константа.

История и оптимална константа

- История: Резултатът е доказан независимо от:
  - Андрю Бери (1941) – за случая на Бернулиеви величини.
  - Карл-Густав Есен (1942) – за общия случай на еднакво разпределени величини.
- Оптимална константа C:
  - Първоначалните доказателства даваха $C \approx 7.59$.

- С години тази оценка се подобрява. Днес е известно, че константата лежи в интервала [13] :

$$0.4097 \approx \frac{\sqrt{10}+3}{6\sqrt{2\pi}} \leq C_{opt} \leq 0.4748$$

- Широко приета стойност за практически оценки е $C = 0.5$.

Следствие. Нека $X \sim \mathrm{Bin}(n, p)$. Тогава $\lim\limits_{n\to\infty} \mathbb{P}\left(\frac{X - np}{\sqrt{npq}} \leq x\right) = \Phi(x)$

Доказателство:

$X = \sum\limits_{j=1}^{n} X_j$, $X_j \in \mathrm{Ber}(p)$. Тогава прилагаме ЦГТ с $\mu = p$ и $\sigma^2 = pq = p(1-p)$.