

## Централна Гранична Теорема (ЦГТ)

**ЗГЧ.**  $(X_i)_{i=1}^{\infty}$  е редица от независими и еднакво разпределени (i.i.d.) случайни величини с  $\mathbb{E}|X_1| < \infty$ ,  $\mathbb{E}X_1 = \mu$  и  $\sigma = \sqrt{\mathbb{D}X_1}$ .

$$S_n = \sum_{j=1}^n X_j \Rightarrow \frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{\text{п.с.}(\mathbb{P})} \mu; \quad \frac{S_n}{n} - \mu \approx \frac{\sigma}{\sqrt{n}} H_n$$

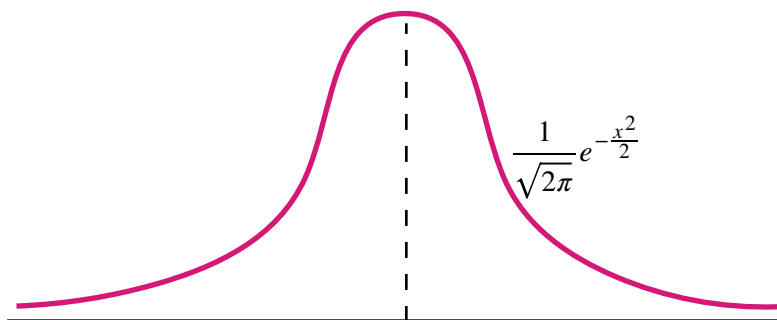
**Теорема (ЦГТ).** Нека  $X = (X_i)_{i=1}^{\infty}$  е редица от независими, еднакво разпределени случайни величини със  $\sigma^2 = \mathbb{D}X_1 < \infty$  и  $\mu = \mathbb{E}X_1$ . Тогава

$$Z_n \stackrel{\text{def.}}{=} \frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{\frac{S_n}{n} - \mu}{\frac{\sigma}{\sqrt{n}}} \xrightarrow[n \rightarrow \infty]{d} Z \in \mathcal{N}(0,1).$$

$$f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad \mathbb{P}(Z < x) = \Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt.$$

$$\mathbb{P}(Z \geq x) = 1 - \Phi(x) = \bar{\Phi}(x).$$

ЦГТ измерва каква е грешката в ЗГЧ. Т.е. ние знаем, че  $\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{} \mu$ , 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.

### Следствия.

⊕  $Z_n := \frac{S_n - n\mu}{\sigma\sqrt{n}}$ , където  $\mu = \mathbb{E}X_1$ ,  $\sigma = \sqrt{\mathbb{D}X_1}$ . ЦГТ гласи следното:

Ако се интересуваме от

$$\begin{aligned}\mathbb{P}(Z_n \in (a, b)) &= \mathbb{P}(S_n \in (n\mu + a\sigma\sqrt{n}, n\mu + b\sigma\sqrt{n})) \sim \mathbb{P}(Z \in (a, b)) = \\ &= \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \times \underbrace{\int_a^b e^{-\frac{y^2}{2}} dy}_{\text{този интеграл не се интересува от това с какви случайни величини сме стартирали. Той зависи само от } a \text{ и } b.}\end{aligned}$$

$$\mathbb{P}(Z_n < a) \xrightarrow{n \rightarrow \infty} \Phi(a) = \mathbb{P}(Z < a); \quad \mathbb{P}(Z_n \geq a) \xrightarrow{n \rightarrow \infty} \bar{\Phi}(a) = 1 - \Phi(a)$$

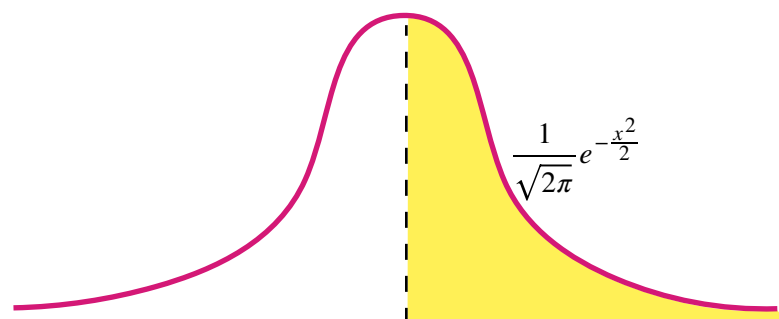
$$\mathbb{P}(Z_n < \mu) \xrightarrow{n \rightarrow \infty} \Phi(\mu).$$

Т.е. ЦГТ дава инструмент, с който може да приближаваме вероятности.

⊕  $\mu = 0, \sigma^2 = 1$ , тогава  $\frac{S_n}{\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} Z \in \mathcal{N}(0,1)$ , където  $S_n = \sum_{j=1}^n X_j$ .

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

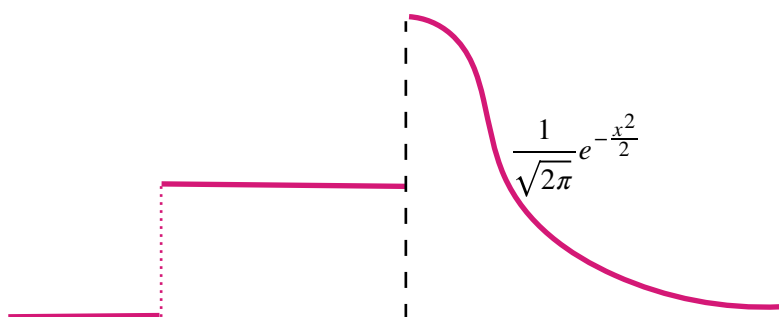
$$\mathbb{P}(S_n > 0) = \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > 0\right) \sim \mathbb{P}(Z > 0) = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{y^2}{2}} dy = \frac{1}{2}.$$



⊕ Хвърляме зарче 6 000 000 пъти. Каква е вероятността измежду тези 6 000 000 пъти да сме хвърлили повече от 1 000 000 пъти 6-ца? Решение: (виж последния пример от зад. 10 от файла [SEM HW](#)).

$$\mu = 0, \sigma^2 = 1, S_n = \sum_{j=1}^n X_j$$

$$\mathbb{P}(S_n > 0) \sim \frac{1}{2}, f_{X_1}(x) = \begin{cases} \frac{1}{2}e^{-x}, & x > 0 \\ \frac{1}{4}, & x \in (-2, 0) \\ 0, & \text{иначе} \end{cases}$$



$$\begin{aligned} \mathbb{E}X_1 &= \int_{-2}^0 x \frac{1}{4} dx + \int_0^{\infty} x \frac{1}{2} e^{-x} dx = \frac{1}{4} \frac{x^2}{2} \Big|_{-2}^0 - \frac{1}{2} \int_0^{\infty} x de^{-x} = \\ &= -\frac{1}{4} \times \frac{4}{2} - \underbrace{\frac{1}{2} x e^{-x} \Big|_0^{\infty}}_{=0} + \frac{1}{2} \int_0^{\infty} e^{-x} dx = -\frac{1}{2} + \frac{1}{2} (-e^{-x}) \Big|_0^{\infty} = \\ &= -\frac{1}{2} + \frac{1}{2} = 0. \end{aligned}$$

Т.е. дори и за случайни величини, които са много далеч от симетрия, ако ги сумираме всички от тях, то вероятността да видим нещо положително е  $\sim \frac{1}{2}$ .

$$X_i = \begin{cases} 1, & \frac{1}{2} \\ -1, & \frac{1}{2} \end{cases} \quad S_n = \mathcal{D}_n - \Lambda_n = \sum_{j=1}^n X_j; \quad \mathbb{P}\left(\frac{S_n}{\sqrt{n}} > a\right) \sim \overline{\Phi}(a),$$

$$\mathbb{P}\left(\mathcal{D}_n > \Lambda_n + a\sqrt{n}\right) = \mathbb{P}\left(\mathcal{D}_n > \frac{n}{2} + \frac{a}{2}\sqrt{n}\right) \sim \overline{\Phi}(a).$$

## Функция на моментите (ФМ)

**Дефиниция.** Нека  $X$  е непрекъснатата случайна величина. Ако  $\mathbb{E}e^{tX}$  съществува за  $t \in (-\varepsilon, \varepsilon)$  и някое  $\varepsilon > 0$ , то  $M_X(t) = \mathbb{E}e^{tX}$  за  $t \in (-\varepsilon, \varepsilon)$  се нарича функция на моментите.

$\oplus$   $\mathbb{E}e^{tX} = \sum_i e^{tx_i} \mathbb{P}(X = x_i)$  – винаги съществува за  $\forall t$ , ако стойностите са краен брой. Но ако не са, тази сума може да не е крайна и да отива към  $\infty$ .

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

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

$$X \sim Unif(0, 1).$$

$$M_X(t) = \mathbb{E}e^{tX} = \int_0^1 e^{tx} \times f_X(x) dx = \int_0^1 e^{tx} \frac{1}{1-0} dx = \frac{1}{t} \int_0^1 e^{tx} dt = \frac{1}{t} e^{tx} \Big|_0^1 = \frac{e^t - 1}{t}$$

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 компактен или някакво крайно множество (в случая са от 0 до 1), то винаги функцията на моментите е добре дефинирана за всяко  $t$ .

**Дефиниция.**  $X$  е случайна величина. Тогава:

- a)  $\mathbb{E}X^k$  се нарича момент от ред  $k \geq 1$ ;
- b)  $\mathbb{E}|X|^k$  се нарича абсолютен момент от ред  $k$ ;
- c)  $\mathbb{E}(X - \mathbb{E}X)^k$  се нарича централен момент от ред  $k$ ;  $DX = \mathbb{E}(X - \mathbb{E}X)^2$
- d)  $\mathbb{E}|X - \mathbb{E}X|^k$  се нарича абсолютен централен момент от ред  $k$ .

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

- a)  $M_X(0) = \mathbb{E}e^{0 \cdot X} = \mathbb{E}1 = 1$ ;
- b)  $M_X^{(k)}(0) = \mathbb{E}X^k$ , за  $\forall k \geq 1$ ;

ред на  
Маклорен

$$c) M_X(0) \stackrel{\text{Маклорен}}{=} \mathbb{E} \sum_{k=0}^{\infty} \frac{t^k}{k!} X^k = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathbb{E} X^k;$$

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

e)  $X \stackrel{d}{=} Y \Leftrightarrow M_X = M_Y$ ;

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

Нека например  $X, Y$  са непрекъснати с плътности  $f_X, f_Y$ . Как да докажем, че  $M_X(t) = M_Y(t)$ ?

$$1. f_{X+Y}(x) = \int_{-\infty}^{\infty} f_X(x-y) f_Y(y) dy;$$

$$2. \int_{-\infty}^{\infty} e^{tX} f_{X+Y}(x) dx = \int_{-\infty}^{\infty} e^{tX} f_X(x) dx \times \int_{-\infty}^{\infty} e^{tY} f_Y(y) dy;$$

g) Ако  $Y = aX + b$ , то  $M_Y(t) = e^{bt} M_X(at)$ , за всяко  $t$ , такава 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  $-\frac{\varepsilon}{|a|} < t < \frac{\varepsilon}{|a|}$ .

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

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

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

$$M_Z(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tX} e^{-\frac{x^2}{2}} dx = \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{1}{2}(x-t)^2} dx \stackrel{x-t=y}{=} e^{\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = e^{\frac{t^2}{2}}$$

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

**Доказателство (ЦГТ).**  $(X_i)_{i=1}^{\infty}$  е редица от независими, еднакво разпределени случайни величини с  $\mathbb{E} X_1 = \mu$ ,  $\mathbb{D} X_1 = \sigma^2$ .

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} Z.$$

$X_1$  има функция на моментите.  $M_{X_1}(t)$  е добре дефинирана за  $t \in (-\varepsilon, \varepsilon)$

$$\frac{S_n - n\mu}{\sigma\sqrt{n}} = \frac{1}{\sqrt{n}} \sum_{j=1}^n \frac{X_j - \mu}{\sigma} = \frac{1}{\sqrt{n}} \sum_{j=1}^n Y_j = \frac{V_n}{\sqrt{n}} =: W_n, \text{ където сме положили}$$

$Y_j = \frac{X_j - \mu}{\sigma}, \forall j \geq 1$  и  $(Y_j)_{j=1}^\infty$  са независими с еднакво разпределение случайни величини.

$$\mathbb{E}Y_1 = \frac{\mathbb{E}X_1 - \mu}{\sigma} = 0; \quad \mathbb{D}Y_1 = \frac{\mathbb{D}X_1}{\sigma^2} = 1.$$

$$M_{Y_1}(t) = e^{-\frac{\mu}{\sigma}t} M_{X_1}\left(\frac{t}{\sigma}\right), \quad \varepsilon < \frac{t}{\sigma} < t\varepsilon \Rightarrow -\sigma\varepsilon < t < \sigma\varepsilon$$

$M_{Y_1}$  е добре дефинирана за  $|t| < \sigma\varepsilon$ .

Нека фиксираме  $t$ . Ще докажем, че  $M_{W_t} \xrightarrow[n \rightarrow \infty]{} M_Z(t) = e^{\frac{t^2}{2}}$ .

$$M_{W_n}(t) = \mathbb{E}e^{\frac{t}{\sqrt{n}}V_n} \stackrel{\text{деф.}}{=} \mathbb{E}e^{\frac{t}{\sqrt{n}}\sum_{j=1}^n Y_j} \stackrel{\text{незав.}}{=} \prod_{j=1}^n M_{Y_j}\left(\frac{t}{\sqrt{n}}\right) \stackrel{Y_j \stackrel{d}{=} Y_1, \forall j}{=} \left[M_{Y_1}\left(\frac{t}{\sqrt{n}}\right)\right]^n.$$

Ако  $\left|\frac{t}{\sqrt{n}}\right| < \sigma\varepsilon$ , то  $M_{W_n}(t)$  е добре дефинирано.

$$M_{W_n}(t) = \left[M_{Y_1}\left(\frac{t}{\sqrt{n}}\right)\right]^n, \quad M_{Y_1}\left(\frac{t}{\sqrt{n}}\right) = \mathbb{E}e^{\frac{t}{\sqrt{n}}Y_1}.$$

$$e^{\frac{t}{\sqrt{n}}Y_1} = 1 + \frac{t}{\sqrt{n}}Y_1 + \frac{t^2}{2n}Y_1^2 + \frac{\theta(Y_1)t^3Y_1^3}{3!n^{\frac{3}{2}}}; \quad |\theta(Y_1)| \leq 1;$$

$$M_{Y_1}(t) = \mathbb{E}e^{\frac{t}{\sqrt{n}}Y_1} = 1 + \frac{t^2}{2n} + \frac{\mathbb{E}\theta(Y_1)t^3Y_1^3}{3!n^{\frac{3}{2}}};$$

$$M_{Y_1} \left( \frac{t}{\sqrt{n}} \right) = 1 + \frac{t^2}{2n} + \frac{t^3}{3!n^{\frac{3}{2}}} \mathbb{E} \theta(Y_1) Y_1^3;$$

$$|\theta(Y_1) Y_1^3| \leq |Y_1|^3 \Rightarrow |\mathbb{E} (\theta(Y_1) Y_1^3)| \leq \mathbb{E} |Y_1|^3 = \rho_3.$$

$$\begin{aligned} M_{W_a}(t) &= \left[ M_{Y_1} \left( \frac{t}{\sqrt{n}} \right) \right]^n = \left[ 1 + \frac{t^2}{2n} + \frac{t^2}{2n} \left( \frac{\rho_3}{\sqrt{n}} \right) \right]^n = \\ &= \left[ 1 + \frac{t^2}{2n} \left( 1 + \underbrace{\frac{\rho_3}{\sqrt{n}}}_{\rightarrow 0} \right) \right]^n \sim \left( 1 + \frac{t^2}{2n} \right)^n \xrightarrow{n \rightarrow \infty} e^{\frac{t^2}{2}} \end{aligned}$$

$$\Rightarrow \lim_{n \rightarrow \infty} M_{W_n}(t) = e^{\frac{t^2}{2}} = M_Z(t)$$

$$\text{СВОЙСТВО} \Rightarrow W_n \xrightarrow[n \rightarrow \infty]{d} Z; \quad Z_n = \frac{S_n - n\mu}{\sigma\sqrt{n}} \xrightarrow[n \rightarrow \infty]{d} Z.$$

$$\mathbb{P} (Z_n \in (a, b)) \xrightarrow[n \rightarrow \infty]{} \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{y^2}{2}} dy = \Phi(b) - \Phi(a).$$

$$\oplus \quad X_i = \begin{cases} 1, & p \\ 0, & 1-p=q \end{cases}, \quad X_i \in \text{Ber}(p)$$

$$\frac{S_n}{n} \xrightarrow[n \rightarrow \infty]{\text{п.с.}} p, \quad \mathbb{E} n = \left| \frac{S_n}{n} - p \right|$$

$$\begin{aligned} \mathbb{P} (|\mathbb{E} n| > \varepsilon) &= \mathbb{P} \left( \underbrace{\left| \frac{S_n - np}{\sqrt{n}\sqrt{pq}} \right|}_{n \rightarrow \infty} > \frac{\sqrt{n}\varepsilon}{\sqrt{pq}} \right) = \mathbb{P} \left( |Z| > \frac{\sqrt{n}\varepsilon}{\sqrt{pq}} \right) = 2\overline{\Phi} \left( \frac{\sqrt{n}\varepsilon}{\sqrt{pq}} \right) \leq \\ &\leq 2\overline{\Phi} \left( \frac{\sqrt{n}\varepsilon}{\frac{1}{2}} \right) = 2\overline{\Phi}(2\sqrt{n}\varepsilon) = 2 \int_{2\sqrt{n}\varepsilon}^{+\infty} e^{-\frac{y^2}{2}} dy. \end{aligned}$$

**Тоерема (Бери-Есен).** Нека  $(X_i)_{i=1}^{\infty}$  са независими и еднакво разпределени случайни величини с  $\mathbb{E}X_1 = \mu$ ,  $DX_1 = \sigma^2$  и  $\mathbb{E}|X_1 - \mathbb{E}X_1|^3 = \rho_3$ .

$$\text{Тогава } \sup_{x \in \mathbb{R}} \left| \mathbb{P} \left( \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x \right) - \underbrace{\Phi(x)}_{=\mathbb{P}(Z < x)} \right| \leq 0,4748 \frac{\rho^3}{\sigma^3\sqrt{n}}.$$

**Следствие.**  $X \in \text{Bin}(n, p)$ , то  $\lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{X - np}{\sqrt{npq}} \leq x \right) = \Phi(x)$

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