

Centralne Twierdzenie Graniczne

Gabriel Tyszką

25 czerwca 2025

Spis treści

1	Czym jest Centralne Twierdzenie Graniczne?	1
2	Kluczowe fakty dotyczące sum zmiennych	1
3	Centralne Twierdzenie Graniczne : Ścisłe sformułowanie	1
4	Przykład Zastosowania: Suma Rzutów Kostką	2
5	Założenia Centralnego Twierdzenia Granicznego	2

1 Czym jest Centralne Twierdzenie Graniczne?

Ogólna idea CTG jest niezwykle prosta, a jednocześnie potężna: jeśli weźmiemy sumę wielu niezależnych zmiennych losowych, to rozkład prawdopodobieństwa tej sumy będzie coraz bardziej przypominał krzywą dzwonową, czyli rozkład normalny, w miarę wzrostu liczby tych zmiennych. Jest to prawdą, niezależnie od początkowego rozkładu poszczególnych zmiennych.

2 Kluczowe fakty dotyczące sum zmiennych

Jeśli mamy zmienną losową z początkową średnią μ_X i odchyleniem standardowym σ_X , to dla sumy n niezależnych realizacji tej zmiennej:

- **Średnia sumy:** $\mu_{sum} = n \times \mu_X$. Oznacza to, że średnia sumy przesuwą się liniowo w prawo w miarę dodawania kolejnych zmiennych.
- **Wariancja sumy:** $\sigma_{sum}^2 = n \times \sigma_X^2$. To wariancje dodają się do siebie, a nie odchylenia standardowe.
- **Odchylenie standardowe sumy:** $\sigma_{sum} = \sqrt{n} \times \sigma_X$. Rozkład sumy rozszerza się, ale wolniej, proporcjonalnie do pierwiastka kwadratowego z liczby zmiennych.

3 Centralne Twierdzenie Graniczne : Ścisłe sformułowanie

Niech X_1, X_2, \dots, X_n będzie próbą losową o liczebności n pobraną z populacji o wartości oczekiwanej μ i skończonej, dodatniej wariancji σ^2 . Niech \bar{X}_n oznacza średnią arytmetyczną próby, która sama jest zmienną losową.

Wtedy, gdy liczebność próby n dąży do nieskończoności ($n \rightarrow \infty$), rozkład zmiennej losowej:

$$\sqrt{n}(\bar{X}_n - \mu)$$

zbliża się do rozkładu normalnego o wartości oczekiwanej 0 i wariancji σ^2 , tzn.:

$$\sqrt{n}(\bar{X}_n - \mu) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$$

Oznacza to, że niezależnie od rozkładu populacji, z której pobierana jest próba, rozkład średniej z próby będzie dla dużych n przybliżał rozkład normalny. Dzięki temu rozkład normalny ma fundamentalne znaczenie w statystyce matematycznej, szczególnie w konstrukcji przedziałów ufności i testach statystycznych.

4 Przykład Zastosowania: Suma Rzutów Kostką

Rozważmy przykład: rzucamy uczciwą kostką 100 razy i dodajemy wyniki. Jaki jest zakres wartości, w którym suma wyników znajdzie się z 95% pewnością?

Krok 1: Właściwości pojedynczego rzutu kostką:

- Średnia (μ_X): $\frac{1+2+3+4+5+6}{6} = \mathbf{3.5}$.
- Wariancja (σ_X^2): $\approx \mathbf{2.92}$.
- Odchylenie standardowe (σ_X): $\sqrt{2.92} \approx \mathbf{1.71}$.

Krok 2: Właściwości sumy 100 rzutów:

- Średnia sumy (μ_{sum}): $100 \times \mu_X = 100 \times 3.5 = \mathbf{350}$.
- Odchylenie standardowe sumy (σ_{sum}): $\sqrt{100} \times \sigma_X = 10 \times 1.71 = \mathbf{17.1}$.

Krok 3: Wykorzystanie reguły 3 sigm dla rozkładu normalnego:

- Około 68% wartości mieści się w zakresie ± 1 odchylenia standardowego od średniej.
- Około 95% wartości mieści się w zakresie ± 2 odchylen standardowych od średniej.
- Około 99.7% wartości mieści się w zakresie ± 3 odchylen standardowych od średniej.

Szukany zakres dla 95% pewności to $\mu_{sum} \pm 2 \times \sigma_{sum}$:

- Dolna granica: $350 - (2 \times 17.1) = 350 - 34.2 = \mathbf{315.8}$.
- Górna granica: $350 + (2 \times 17.1) = 350 + 34.2 = \mathbf{384.2}$.

Zaokrąglając, z 95% pewnością suma wyników ze 100 rzutów kostką znajdzie się w zakresie od około 316 do 384.

Warto zauważyć, że CTG pozwala również przewidzieć, jak blisko średnia empiryczna (suma podzielona przez liczbę rzutów) będzie wartości oczekiwanej (w tym przypadku 3.5).

5 Założenia Centralnego Twierdzenia Granicznego

CTG opiera się na trzech kluczowych założeniach, które muszą być spełnione, aby twierdzenie było prawdziwe:

1. **Niezależność zmiennych:** Wszystkie dodawane zmienne losowe muszą być niezależne od siebie. Wynik jednego procesu nie może wpływać na wynik żadnego innego procesu. W przykładzie z kostkami, każdy rzut jest niezależny od poprzednich.
2. **Identyczny rozkład:** Wszystkie zmienne muszą pochodzić z tego samego rozkładu prawdopodobieństwa*. W przykładzie z kostkami, każda kostka ma ten sam rozkład prawdopodobieństwa wyników. Te dwa pierwsze założenia są często określane akronimem IID (Independent and Identically Distributed).
3. **Skończona wariancja:** Wariancja rozkładu, z którego pochodzą zmienne, musi być skończona. W przypadku kostki, gdzie jest tylko sześć możliwych wyników, wariancja jest zawsze skończona. Istnieją jednak rozkłady z nieskończoną wariancją, a w takich sytuacjach CTG może nie obowiązywać, a rozkład sumy może nie dążyć do rozkładu normalnego.