Modelowanie Stochastyczne

Krótki opis nierówności Gaussa wraz z dowodem i zastosowaniami

Nazwisko i imię prowadzącego kurs: Dr Marek Skarupski

Wykonawca:	
Imię i Nazwisko	**Andrzej Puć**
nr Indeksu, wydział	249748, Wydział matematyki
Termin zajęć: dzień tygodnia, godzina	Poniedziałek, 9:15-11:00
Numer grupy ćwiczeniowej	T00-66b
Data oddania pracy	20 kwietnia 2020

Spis treści

1 Wstęp z krótkim rysem historycznym **3**

2 Treść i dowód nierówności **3**

3 Symulacja **6**

4 Podsumowanie **10**

1 Wstęp z krótkim rysem historycznym

Carl Friedrich Gauss (1777-1855) jest powszechnie uznawany za jednego z najwybitniejszych matematyków dziewiętnastego wieku. Jego zainteresowania naukowe sięgały także takich dziedzin jak elektromagnetyzm, optyka (gdzie wprowadził kilka nowych pojęć) oraz astronomia. Urodzony w Brunszwiku, w niezamożnej rodzinie, przejawiał duże zdolności już od dziecka. Znana jest anegdota o tym, jak w wieku dziesięciu lat w szybki sposób obliczył sumę ciągu arytmetycznego. Należy też podkreślić, że prawdopodobnie duża część jego prac nie została opublikowana, gdyż nie chciał on publikować teorii niepełnych.[4]

Nierówność, którą przedstawimy, Gauss udowodnił w roku 1823 [3]. Prześledzimy więc jej dowód oraz porównamy z szerzej znaną nierównością Czebyszewa.

2 Treść i dowód nierówności

Poniżej zapiszemy dowód oparty na oryginalnym pomyśle Gaussa [3].
Dowiedziemy więc, że

Twierdzenie 1. *dla unimodalnej zmiennej losowej X o istniejącej wartości oczekiwanej v i wariancji σ mamy [3]*

$$P(|X - v| \geqslant r) \leqslant \frac{4\sigma^2}{9r^2} \ dla\ r \geqslant \sqrt{\frac{4}{3}}\sigma,$$

$$P(|X - v| \geqslant r) \leqslant 1 - \frac{r}{\sqrt{3}\sigma} \ dla\ r \leqslant \sqrt{\frac{4}{3}}\sigma.$$

Dowód. Jeśli lewa strona nierówności jest równa 0 całość staje się trywialna. Taki sam wynik daje nam podstawienie $\sigma^2 = \infty$.
Rozpatrzmy więc przypadek, kiedy $P(|X - v| \geqslant r) > 0$ i $\sigma < \infty$. Zdefiniujmy najpierw nową zmienną losową $Z = |X - EX| = |X - v|$ oraz jej gęstość $g(z) = f(v + z) + f(v - z)$ na $(0, \infty)$. Z unimodalności X wiemy, że f jest niemalejąca poniżej v i nierosnąca powyżej. Stąd i z parzystości zmiennej losowej Z, g jest nierosnąca na $(0, \infty)$ z modą zero. Mamy zatem $g(r) > 0$ i $\sigma^2 = \int_{\mathbb{R}^+} z^2 g(z)dz$ (z tego, że zajmujemy się przypadkiem nietrywialnym).

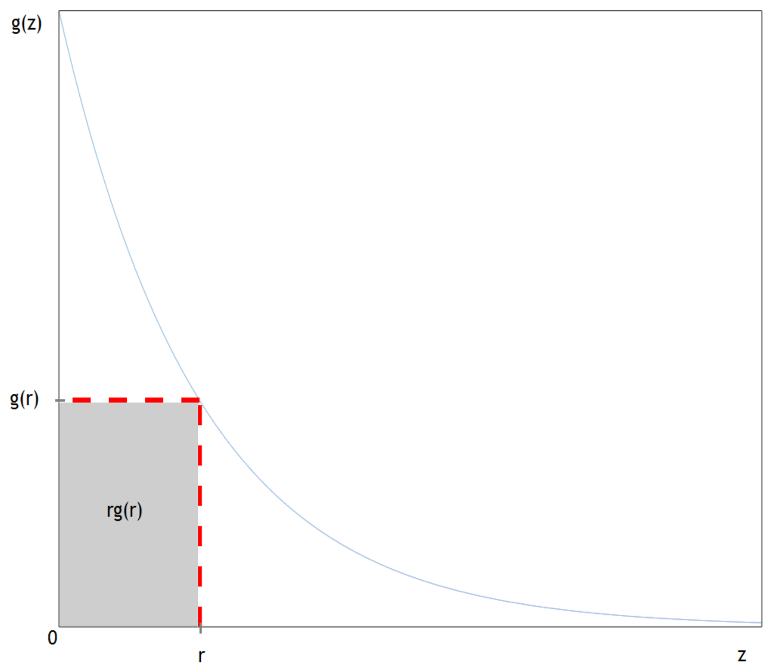
Zdefiniujmy teraz dystrybuantę oraz dystrybuantę odwrotną zmiennej losowej Z.
Dla $a = \sup\{z \geqslant 0 : g(z) > 0\} \in (0, \infty)$ mamy odpowiednio funkcję

$$F(x) = \int_0^x g(z)dz \ : (0, a) \to (0, 1)$$

różniczkowalną i rosnącą ($x_1 > x_2 \Rightarrow F(x_1) > F(x_2)$) oraz

$$\Psi(y) = F^{-1}(y) \ : (0, 1) \to (0, a)$$

spełniającą te same warunki. Pochodna $\Psi'(y) = \frac{1}{F'(\Psi(y))} = \frac{1}{g(\Psi(y))}$ jest niemalejąca (co wynika z monotoniczności gęstości g). Stąd Ψ jest wypukła.
Z tego, że $F(r) < 1$ wynika, że $r \in (0, a)$. Narysujmy teraz poglądowy rysunek gęstości g zgodnej z poprzednimi założeniami.



Rysunek 1: Wykres przedstawia szkic gęstości g oraz prostokąta o bokach r, g(r). Widzimy, że $g(z) \geqslant g(r)$ dla $z \in (0, r)$.

Mamy zatem

$$F(r) = \int_0^r g(z)dz \geqslant r \cdot g(r) > 0,$$

$$\epsilon = 1 - \frac{rg(r)}{F(r)}$$

gdzie ϵ określa różnicę między polem prostokąta a polem pod wykresem g(z) na (0, r). Obliczmy teraz parametry stycznej L do wykresu Ψ w punkcie F(r).

$$\Psi'(F(r)) = \frac{1}{g(\Psi(F(r)))} = \frac{1}{g(r)},$$

$$L(y) - \Psi(F(r)) = \frac{1}{g(r)}(y - F(r)),$$

$$L(y) = \frac{1}{g(r)}(y - \epsilon F(r)).$$

Wypukłość Ψ gwarantuje, że $L(y) \leqslant \Psi(y)$ dla każdego $y \in (0,1)$. Możemy zatem wziąć $y_0 = \epsilon F(r)$ i dla $y \in (y_0, 1)$ zapisać

$$[L(y)]^2 \leqslant [\Psi(y)]^2,$$

$$\int_{y_0}^{1} [L(y)]^2 \, dy \leqslant \int_{Y_0}^{1} [\Psi(y)]^2 \, dy.$$

Prawa strona powyższej nierówności jest oczywiście ograniczona przez

$$\int_0^1 [\Psi(y)]^2 \, dy = E[(\Psi(Y))^2] = E(Z^2) = \sigma^2$$

gdzie zmienna losowa $Y \sim U(0,1)$. Zgodnie z powyższym σ^2 jest ograniczona z dołu przez całkę

$$\int_{y_0}^{1} [L(y)]^2 \, dy \leqslant \sigma^2.$$

Obliczając ją otrzymamy więc

$$\sigma^2 \geqslant \frac{1}{[g(r)]^2} \int_{\epsilon F(r)}^{1} (y - \epsilon F(r))^2 dy = \frac{(1 - \epsilon F(r))^3}{3(g(r))^2}.$$

Po podstawieniu $g(r) = \frac{1}{r}(F(r) - \epsilon F(r))$

$$\sigma^2 \geqslant (1 - \epsilon F(r))^3 \cdot \frac{r^2}{3\,[F(r)]^2} \cdot \frac{1}{(1-\epsilon)^2} = B(\epsilon).$$

Pochodna B to funkcja określona na $\epsilon \in (-\infty, 1)$

$$B'(\epsilon) = \frac{r^2\,(1 - \epsilon F(r))^2}{3F(r)\,(1-\epsilon)^3} \left(\epsilon - 3 + \frac{2}{F(r)}\right),$$

która zmienia znak w punkcie $\epsilon_0 = 3 - \frac{2}{F(r)} < 1$. Stąd minimum globalne B to

$$B(\epsilon_0) = \frac{9r^2}{4}\,(1 - F(r)).$$

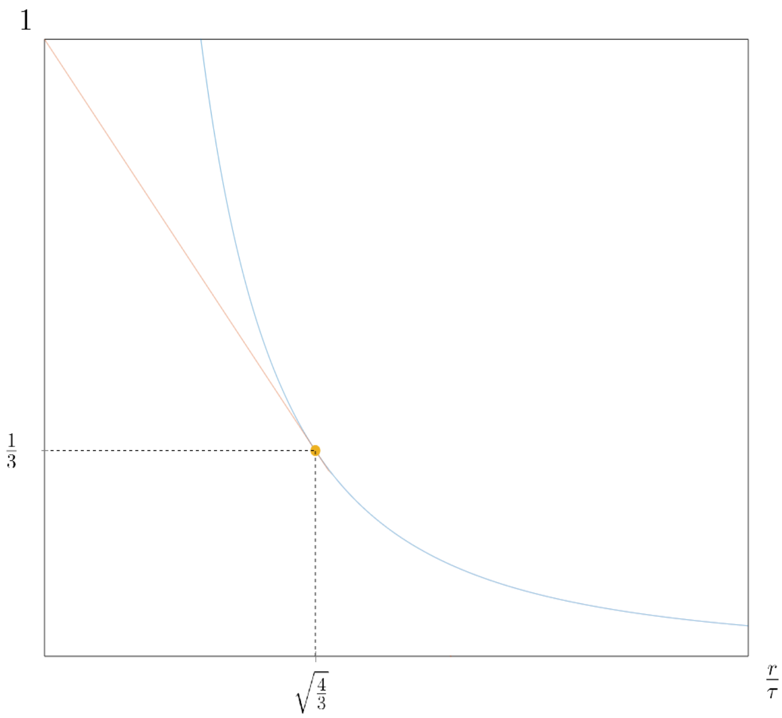
Warunek $\sigma^2 \geqslant B(\epsilon_0)$ można więc przedstawić jako

$$1 - F(r) \leqslant \frac{4\sigma^2}{9r^2}$$

co z definicji dystrybuanty pozwala zapisać ostatecznie

$$P(|X - v| \geqslant r) = 1 - F(r) \leqslant \frac{4\sigma^2}{9r^2}, \ r > 0.$$

Aby przejść z powyższej nierówności do pełnej postaci twierdzenia (1) zauważmy, że funkcja wypukła $1 - F$ nie przecina żadnego odcinka łączącego punkt $(0,1)$ z punktem $(r, \frac{4\sigma^2}{9r^2})$. Stąd najlepszym ograniczeniem dla $1 - F$ będzie najbardziej "stroma" prosta, $1 - \frac{r}{\sqrt{3}\sigma}$ aż do punktu przecięcia w $r = \sqrt{\frac{4}{3}}\sigma$ co zostało zobrazowane na wykresie 2.



Rysunek 2: Wykres przedstawia funkcję z prawej strony nierówności 2 (niebieski) oraz oraz fragment funkcji $1 - \frac{r}{\sqrt{3}\sigma}$ (czerwony).

$\square$

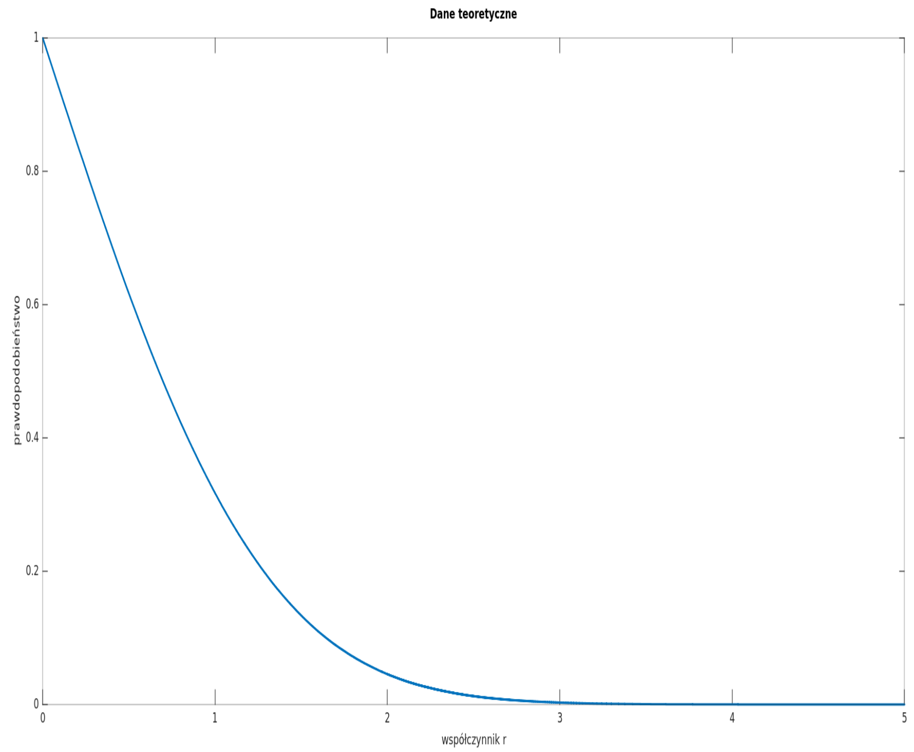
3 Symulacja

Ostatnia część dowodu dobitnie wskazuje na dużą przewagę nierówności Gaussa w stosunku do innej, powszechnie przez nas stosowanej nierówności.[1][2]

Twierdzenie 2. ***Nierówność Czebyszewa.*** *Jeżeli istnieje* $m = E(X)$, $\sigma^2 = Var(X)$, *to*

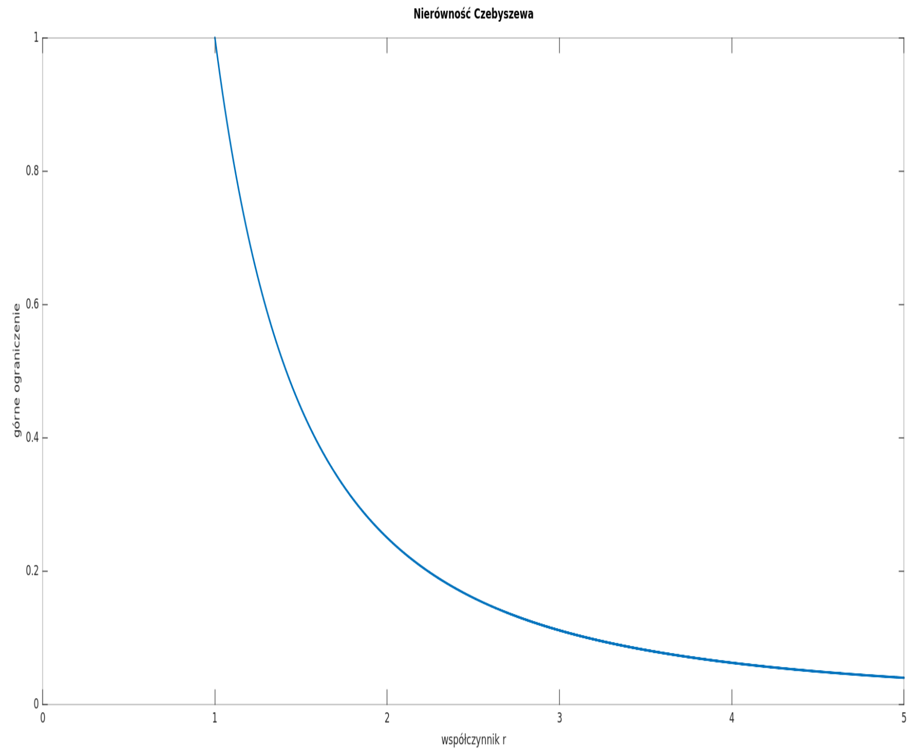
$$P(|X - m| \geqslant r) \leqslant \frac{1}{t^2}\sigma^2.$$

Aby wyraźniej zobrazować różnice między tymi ograniczeniami przeprowadzimy prostą symulację na podstawie dużej próby z rozkładu normalnego.

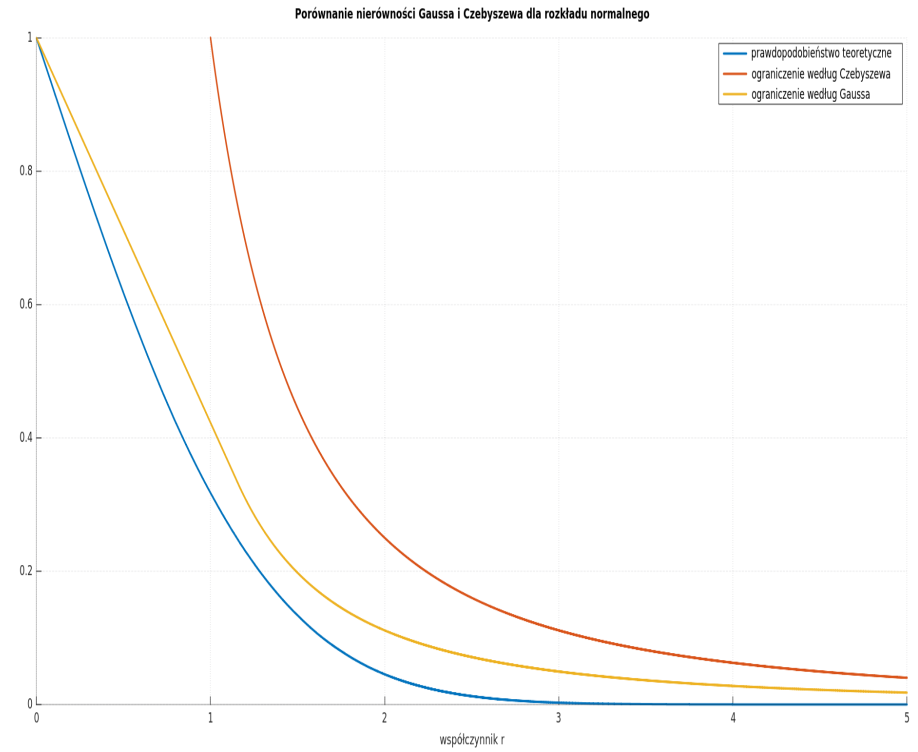


Rysunek 3: Graficzne przedstawienie $P(|X - v| \geqslant r)$ dla r z przedziału $(0.001, 5)$.

Uznajmy na potrzeby eksperymentu, że obliczone powyżej prawdopodobieństwa przedstawiają teoretyczną postać $1 - F(r)$ (gdzie F to dystrybuanta rozkładu normalnego). Spróbujmy teraz ograniczyć je za pomocą nierówności Czebyszewa.



Prawdopodobieństwo, zdefiniowane jak na poprzednim wykresie, ograniczone za pomocą twierdzenia 2. Wykres nie był rysowany dla ograniczeń większych od 1. Z definicji funkcji prawdopodobieństwa wiemy, że jest to niepotrzebne. Warto zauważyć, że nie jest możliwe narysowanie takiego wykresu dla podobnej dziedziny $[0, 5)$. Ograniczenie to jest bardzo oddalone od teoretycznych wartości dla małych r. Z uwagi na fakt, że wykres dla nierówności Gaussa został już narysowany (2) przejdziemy do porównania graficznego trzech funkcji.



Symulacja ta dobitnie pokazuje jak dokładne potrafi być przybliżenie za pomocą nierówności Gaussa. Maksymalna odległość między linią żółtą a niebieską wynosi zaledwie ~ 0.11509. Dla porównania minimalna odległość od danych teoretycznych dla nierówności Czebyszewa to w tym przypadku 0.01.

Powyższe rozważania sugerują dużą przydatność tak dobrego ograniczenia. Rzeczywiście, za jego pomocą można uzasadnić nawet tak popularne tożsamości jak reguła trzech sigm.

Twierdzenie 3. ***Reguła trzech sigm.*** *Jeśli istnieją $E(X)$, $Var(X) = \sigma^2 < \infty$, to*

$$P\left(|X - E(X)| > 3\sigma\right) \leqslant \frac{1}{9}.$$

Dowód. Zauważmy, że aby udowodnić powyższe twierdzenie na podstawie nierówności Gaussa wystarczy podstawić 3σ w miejsce r w poprzednio udowodnionych wzorach

$$P\left(|X - E(X)| > 3\sigma\right) \leqslant \frac{4\sigma^2}{9 \cdot 3^3\sigma^2}$$

ponieważ $3\sigma \geqslant \sqrt{\frac{4}{3}}\sigma$. Skracając σ^2 w ułamku po prawej stronie nierówności otrzymamy

$$P\left(|X - E(X)| > 3\sigma\right) \leqslant \frac{4}{81} \leqslant \frac{1}{9}.$$

$\square$

Warto podkreślić, że za pomocą tej nierówności można także w łatwy sposób dowieść prawdziwości słabego prawa wielkich liczb [2].

4 Podsumowanie

Po zapoznaniu z wieloma niekwestionowanymi zaletami nierówności Gaussa oraz przejrzystym dowodem, opartym na elementarnych pojęciach, pozostaje pytanie o przyczynę małej popularności tej tożsamości. Nasuwającą się odpowiedzią jest spora liczba warunków, które zmienna losowa musiałaby spełnić. W przypadku nierówności Czebyszewa mamy natomiast tylko warunek dotyczący istnienia drugiego momentu danej zmiennej losowej.

Literatura

[1] William Feller. *Wstęp do rachunku prawdopodobieństwa*, volume 2. Wydawnictwo Nukowe PWN SA, The address, 4 edition, 2009.

[2] Gerard Hooghiemstra and Piet Mieghem. An inequality of Gauss. *NAW*, 5/16:123–126, 06 2015.

[3] Friedrich Pukelsheim. The three sigma rule. *The American Statistician*, 48(2):88–91, 1994. Accessed: 30-03-2020 20:58 UTC. URL: `https://www.jstor.org/stable/2684253`.

[4] Karolee Weller. Carl Friedrich Gauss. `http://www.math.wichita.edu/history/men/gauss.html`. Accessed: 2020-04-30.